2

Einstein Geometrodynamics

If Einstein gave us a geometric account of motion and gravity, if according to his 1915 and still-standard geometrodynamics spacetime tells mass how to move and mass tells spacetime how to curve, then his message requires mathematical tools to describe position and motion, curvature and the action of mass on curvature. The tools (see the mathematical appendix) will open the doorways to the basic ideas—equivalence principle, geometric structure, field equation, equation of motion, equation of geodesic deviation—and these ideas will open the doorways to more mathematical tools—exact solutions of Einstein's geometrodynamics field equation, equations of conservation of source, and the principle that the boundary of a boundary is zero. The final topics in this chapter—black holes, singularities, and gravitational waves—round out the interplay of mathematics and physics that is such a central feature of Einstein's geometrodynamics.

2.1 THE EQUIVALENCE PRINCIPLE

At the foundations of Einstein[1–10] geometrodynamics[11–21] and of its geometrical structure is one of the best-tested principles in the whole field of physics (see chap. 3): the equivalence principle.

Among the various formulations of the **equivalence principle**[16,21] (see § 3.2), we give here three most important versions: the **weak form**, also known as the *uniqueness of free fall* or the *Galilei equivalence principle* at the base of most known viable theories of gravity; the **medium strong form**, at the base of metric theories of gravity; and the **very strong form**, a cornerstone of Einstein geometrodynamics.

Galilei in his *Dialogues Concerning Two New Sciences*[22] writes: "The variation of speed in air between balls of gold, lead, copper, porphyry, and other heavy materials is so slight that in a fall of 100 cubits a ball of gold would surely not outstrip one of copper by as much as four fingers. Having observed this, I came to the conclusion that in a medium totally void of resistance all bodies would fall with the same speed."

We therefore formulate the **weak equivalence principle**, or *Galilei equivalence principle*[22,23] in the following way: *the motion of any freely falling test* *particle is independent of its composition and structure*. A test particle is defined to be electrically neutral, to have negligible gravitational binding energy compared to its rest mass, to have negligible angular momentum, and to be small enough that inhomogeneities of the gravitational field within its volume have negligible effect on its motion.

The weak equivalence principle—that all test particles fall with the same acceleration—is based on the principle[24] that the ratio of the inertial mass to the gravitational—passive—mass is the same for all bodies (see chap. 3). The principle can be reformulated by saying that in every local, nonrotating, freely falling frame the line followed by a freely falling test particle is a straight line, in agreement with special relativity.

Einstein generalized[10] the weak equivalence principle to all the laws of special relativity. He hypothesized that in no local freely falling frame can we detect the existence of a gravitational field, either from the motion of test particles, as in the weak equivalence principle, or from any other special relativistic physical phenomenon. We therefore state the **medium strong form of the equivalence principle**, also called the *Einstein equivalence principle*, in the following way: *for every pointlike event of spacetime, there exists a sufficiently small neighborhood such that in every local, freely falling frame in that neighborhood, all the nongravitational laws of physics obey the laws of special relativity*. As already remarked, the medium strong form of the equivalence principle is satisfied by Einstein geometrodynamics and by the so-called metric theories of gravity, for example, Jordan-Brans-Dicke theory, etc. (see chap. 3).

If we replace[18] *all the nongravitational laws of physics* with *all the laws of physics* we get the **very strong equivalence principle**, which is at the base of Einstein geometrodynamics.

The medium strong and the very strong form of the equivalence principle differ: the former applies to all phenomena except gravitation itself whereas the latter applies to all phenomena of nature. This means that according to the medium strong form, the existence of a gravitational field might be detected in a freely falling frame by the influence of the gravitational field on local gravitational phenomena. For example, the gravitational binding energy of a body might be imagined to contribute differently to the inertial mass and to the passive gravitational mass, and therefore we might have, for different objects, different ratios of inertial mass to gravitational mass, as in the Jordan-Brans-Dicke theory. This phenomenon is called the Nordtvedt effect[26,27] (see chap. 3). If the very strong equivalence principle were violated, then Earth and Moon, with different gravitational binding energies, would have different ratios of inertial mass to passive gravitational mass and therefore would have different accelerations toward the Sun; this would lead to some polarization of the Moon orbit around Earth. However, the Lunar Laser Ranging[28] experiment has put strong limits on the existence of any such violation of the very strong equivalence principle.

The equivalence principle, in the medium strong form, is at the foundations of Einstein geometrodynamics and of the other metric theories of gravity, with a "locally Minkowskian" spacetime. Nevertheless, it has been the subject of many discussions and also criticisms over the years.[13,25,29,30]

First, the equivalence between a gravitational field and an accelerated frame in the absence of gravity, and the equivalence between a flat region of spacetime and a freely falling frame in a gravity field, has to be considered valid only locally and not globally.[29] However, the content of the strong equivalence principle has been criticized even "locally." It has been argued that if one puts a spherical drop of liquid in a gravity field, after some time one would observe a tidal deformation from sphericity of the drop. Of course, this deformation does not arise in a flat region of spacetime. Furthermore, let us consider a freely falling frame in a small neighborhood of a point in a gravity field, such as the cabin of a spacecraft freely falling in the field of Earth. Inside the cabin, according to the equivalence principle, we are in a local inertial frame, without any observable effect of gravity. However, let us take a gradiometer, that is, an instrument which measures the gradient of the gravity field between two nearby points with great accuracy (present room temperature gradiometers may reach a sensitivity of about 10^{-11} (cm/s^2)/cm per Hz$^{-1/2}$ $\equiv$ 10^{-2} Eötvös per Hz$^{-1/2}$ between two points separated by a few tens of cm; future superconducting gradiometers may reach about 10^{-5} Eötvös Hz$^{-1/2}$ at certain frequencies, see §§ 3.2 and 6.9). No matter if we are freely falling or not, the gradiometer will eventually detect the gravity field and thus will allow us to distinguish between the freely falling cabin of a spacecraft in the gravity field of a central mass and the cabin of a spacecraft away from any mass, in a region of spacetime essentially flat. Then, may we still consider the strong equivalence principle to be valid?

From a mathematical point of view, at any point P of a pseudo-Riemannian, Lorentzian, manifold (see § 2.2 and mathematical appendix), one can find coordinate systems such that, at P, the metric tensor $g_{\alpha\beta}$ (§ 2.2) is the Minkowski metric $\eta_{\alpha\beta} = \mathrm{diag}(-1, +1, +1, +1)$ and the first derivatives of $g_{\alpha\beta}$, with respect to the chosen coordinates, are zero. However, one cannot in general eliminate certain combinations of second derivatives of $g_{\alpha\beta}$ which form a tensor called the Riemann curvature tensor: $R^{\alpha}{}_{\beta\gamma\delta}$ (see § 2.2 and mathematical appendix). The Riemann curvature tensor represents, at each point, the intrinsic curvature of the manifold, and, since it is a tensor, one cannot transform it to zero in one coordinate system if it is nonzero in some other coordinate system. For example, at any point P on the surface of a sphere one can find coordinate systems such that, at P, the metric is $g_{11}(P) = g_{22}(P) = 1$. However, the Gaussian curvature of the sphere (see mathematical appendix), that is, the R_{1212} component of the Riemann tensor, is, at each point, an intrinsic (independent of coordinates) property of the surface and therefore cannot be eliminated with a coordinate transformation. The metric tensor can indeed be written using the Riemann tensor, in a neighborhood of a spacetime event, in a freely falling, nonrotating, local inertial frame, to second order in the separation, δx^α, from the origin:

$$
\begin{aligned}
g_{00} &= -1 - R_{0i0j}\delta x^i \delta x^j \\
g_{0k} &= -\frac{2}{3} R_{0ikj}\delta x^i \delta x^j \\
g_{kl} &= \delta_{kl} - \frac{1}{3} R_{kilj}\delta x^i \delta x^j .
\end{aligned}
\tag{2.1.1}
$$

These coordinates are called *Fermi Normal Coordinates*.

In section 2.5 we shall see that in general relativity, and other metric theories of gravity, there is an important equation, the *geodesic deviation equation*, which connects the physical effects of gravity gradients just described with the mathematical structure of a manifold, that is, which connects the physical quantities measurable, for example with a gradiometer, with the mathematical object representing the curvature: the Riemann curvature tensor. We shall see via the geodesic deviation equation that the relative, covariant, acceleration between two freely falling test particles is proportional to the Riemann curvature tensor, that is, $\ddot{\delta x}^\alpha \sim R^\alpha{}_{\beta\mu\nu}\delta x^\mu$, where δx^α is the "small" spacetime separation between the two test particles. On a two-surface, this equation is known as the Jacobi equation for the second derivative of the distance between two geodesics on the surface as a function of the Gaussian curvature.

The Riemann curvature tensor, however, cannot be eliminated with a coordinate transformation. Therefore, the relative, covariant, acceleration cannot be eliminated with a change of frame of reference. In other words, by the measurement of the second rate of change of the relative distance between two test particles, we can detect, in every frame, the gravitational field, and indeed, at least in principle, we can measure all the components of the Riemann curvature tensor and therefore completely determine the gravitational field. Furthermore, the motion of one test particle in a local freely falling frame can be described by considering the origin of the local frame to be comoving with another nearby freely falling test particle. The motion of the test particle in the local frame, described by the separation between the origin and the test particle, is then given by the geodesic deviation equation of section 2.5. This equation gives also a rigorous description of a falling drop of water and of a freely falling gradiometer, simply by considering two test particles connected by a spring, that is, by including a force term in the geodesic deviation equation (see § 3.6.1).

From these examples and arguments, one might think that the strong equivalence principle does not have the content and meaning of a fundamental principle of nature. Therefore, one might think to restrict to interpreting the equivalence principle simply as the equivalence between inertial mass M_i and gravitational mass M_g. However, $M_i = M_g$ is only a part of the medium (and strong) equivalence principle whose complete formulation is at the basis of the locally Minkowskian spacetime structure.

In general relativity, the content and meaning of the strong equivalence principle is that *in a sufficiently small neighborhood of any spacetime event, in a local freely falling frame, no gravitational effects are observable*. Here, neighborhood means neighborhood in *space* and *time*. Therefore, one might formulate the medium strong equivalence principle, or Einstein equivalence principle, in the following form: for every spacetime event (then excluding singularities), for any experimental apparatus, with some limiting accuracy, there exists a neighborhood, in space and time, of the event, and infinitely many local freely falling frames, such that for every nongravitational phenomenon the difference between the measurements performed (assuming that the smallness of the spacetime neighborhood does not affect the experimental accuracy) and the theoretical results predicted by special relativity (including the Minkowskian character of the geometry) is less than the limiting accuracy and therefore undetectable in the neighborhood. In other words, in the spacetime neighborhood considered, in a freely falling frame all the nongravitational laws of physics agree with the laws of special relativity (including the Minkowskian character of spacetime), apart from a small difference due to the gravitational field that is; however, unmeasurable with the given experimental accuracy. We might formulate the very strong equivalence principle in a similar way.

For a test particle in orbit around a mass M, the geodesic deviation equation gives

$$\delta\ddot{x}^{\alpha} \sim R^{\alpha}{}_{0\beta 0}\delta x^{\beta} \sim \omega_0^2 \delta x^{\alpha} \tag{2.1.2}$$

where ω_0 is the orbital frequency. Thus, one would sample large regions of the spacetime if one waited for even one period of this "oscillator." We must limit the dimensions in space and time of the domain of observation to values small compared to one period if we are to uphold the equivalence principle.

A liquid drop which has a surface tension, and which resists distortions from sphericity, supplies an additional example of how to interpret the equivalence principle. In order to detect a gravitational field, the *measurable* quantity—the *observable*—is the tidal deformation $\boldsymbol{\delta x}$ of the drop. If a gravity field acts on the droplet and if we choose a small enough drop, we will not detect any deformation because the tidal deformations from sphericity are proportional to the size D of the small drop, and even for a self-gravitating drop of liquid in some external gravitational field, the tidal deformations $\boldsymbol{\delta x}$ are proportional to its size D. This can be easily seen from the geodesic deviation equation with a springlike force term (§ 3.6.1), in equilibrium: $\frac{k}{m}\delta x \sim R^{i}{}_{0j0}D \sim \frac{M}{R^3}D$, where M is the mass of an external body and $R^{i}{}_{0j0} \sim \frac{M}{R^3}$ are the leading components of the Riemann tensor generated by the external mass M at a distance R. Thus, in a spacetime neighborhood, with a given experimental accuracy, the deformation $\boldsymbol{\delta x}$, is unmeasurable for sufficiently small drops.

We overthrow yet a third attempt to challenge the equivalence principle—this time by use of a modern gravity gradiometer—by suitably limiting the scale or time of action of the gradiometer. Thus either one needs large distances over which to measure the gradient of the gravity field, or one needs to wait a period of time long enough to increase, up to a detectable value, the amplitude of the oscillations measured by the gradiometer. Similarly, with gravitational-wave detectors (resonant detectors, laser interferometers, etc.; see § 3.6), measuring the time variations of the gravity field between two points, one may be able to detect very small changes of the gravity field (present relative sensitivity to a metric perturbation or fractional change in physical dimensions $\sim 10^{-18}$ to 10^{-19}, "near" future sensitivity $\sim 10^{-21}$ to 10^{-22}; see § 3.6) during a small interval of time (for example a burst of gravitational radiation of duration $\sim 10^{-3}$ s). However, all these detectors basically obey the geodesic deviation equation, with or without a force term, and in fact their sensitivity to a metric perturbation decreases with their dimensions or time of action (see § 3.6).

In a final attempt to challenge the equivalence principle one may try to measure the *local* deviations from geodesic motion of a spinning particle, given by the Papapetrou equation described in section 6.10. In agreement with the geodesic deviation equation, these deviations are of type $\delta\ddot{x}^i \sim R^i{}_{0\mu\nu} J^{\mu\nu}$, where $J^{\mu\nu}$ is the spin tensor of the particle and $u^0 \cong 1$, defined in section 6.10. However, general relativity is a classical—nonquantized—theory. Therefore, in the formulation of the strong equivalence principle one has to consider only *classical* angular momentum of finite size particles. However, the classical angular momentum of a particle goes to zero as its size goes to zero, and we thus have a case analogous to the previous ones: sufficiently limited in space and time, no observations of motion will reveal any violation of the equivalence principle.

Of course, the local "eliminability" of gravitational effects is valid for gravity only. Two particles with arbitrary electric charge to mass ratios, $\frac{q_1}{m_1} \neq \frac{q_2}{m_2}$, for example $q_1 = 0$ and $\frac{q_2}{m_2} \gg 1$ (in geometrized units), placed in an external electric field, will undergo a relative acceleration that can be very large independently from their separation going to zero.

In summary, since the gravitational field is represented by the Riemann curvature tensor it cannot be transformed to zero in some frame if it is different from zero in some other frame; however, the measurable effects of the gravitational field, that is, of the spacetime curvature, between two nearby events, go to zero as the separation in space and time between the two events, or equivalently as the separation between the space and time origin of a freely falling frame and another local event. Thus, *any effect of the gravitational field is unmeasurable, in a sufficiently small spacetime neighborhood in a local freely falling frame of reference*.

2.2 THE GEOMETRICAL STRUCTURE

In 1827 Carl Friedrich Gauss (1777–1855) published what is thought to be the single most important work in the history of differential geometry: *Disquisitiones generales circa superficies curvas* (General Investigations of Curved Surfaces).[31] In this work he defines the curvature of two-dimensional surfaces, the Gaussian curvature, from the intrinsic properties of a surface. He concludes that all the properties that can be studied within a surface, without reference to the enveloping space, are independent from deformations, without stretching, of the surface—*theorema egregium*—and constitute the intrinsic geometry of the surface. The distance between two points, measured along the shortest line between the points within the surface, is unchanged for deformations, without stretching, of the surface.

The study of non-Euclidean geometries really began with the ideas and works of Gauss, Nikolai Ivanovich Lobačevskij (1792–1856),[32] and János Bolyai (1802–1860). In non-Euclidean geometries, Euclid's 5th postulate on straight lines is not satified (that through any point not lying on a given straight line, there is one, and only one, straight line parallel to the given line; see § 1.1).

In 1854 Georg Friedrich Bernhard Riemann (1826–1866) delivered his qualifying doctoral lecture (published in 1866): *Über die Hypothesen, welche der Geometrie zu Grunde liegen* (On the Hypotheses Which Lie at the Foundations of Geometry).[33] This work is the other cornerstone of differential geometry; it extends the ideas of Gauss from two-dimensional surfaces to higher dimensions, introducing the notions of what we call today Riemannian manifolds, Riemannian metrics, and the Riemannian curvature of manifolds, a curvature that reduces to the Gaussian curvature for ordinary two-surfaces. He also discusses the possibilities of a curvature of the universe and suggests that space geometry may be related to physical forces (see § 1.1).

The absolute differential calculus is also known as tensor calculus or Ricci calculus. Its development was mainly due to Gregorio Ricci Curbastro (1853–1925) who elaborated the theory during the ten years 1887–1896.[34,35] Riemann's ideas and a formula (1869) of Christoffel[36] were at the basis of the tensor calculus. In 1901 Ricci and his student Tullio Levi-Civita (1873–1941) published the fundamental memoir: *Méthods de calcul différential absolu et leurs applications* (Methods of Absolute Differential Calculus and their Applications),[35] a detailed description of the tensor calculus; that is, the generalization, on a Riemannian manifold, of the ordinary differential calculus. At the center of attention are geometrical objects whose existence is independent of any particular coordinate system.

From the medium strong equivalence principle, it follows that spacetime must be at an event, in suitable coordinates, Minkowskian; furthermore, it may be possible to show some theoretical evidence for the existence of a curvature of the spacetime.[37] The **Lorentzian**, **pseudo-Riemannian**[38–43] character of spacetime is the basic ingredient of general relativity and other metric theories of gravity; we therefore assume the **spacetime** to be a **Lorentzian manifold**: that is, a four-dimensional pseudo-Riemannian manifold, with signature $+2$ (or -2, depending on convention); that is, a smooth manifold M^4 with a continuous two-index tensor field $\boldsymbol{g}$, the **metric tensor**, such that $\boldsymbol{g}$ is covariant (see the mathematical appendix), symmetric, and nondegenerate or, simply, at each point of M, in components:

$$\begin{aligned} &g_{\beta\alpha} = g_{\alpha\beta} \\ &\det(g_{\alpha\beta}) \neq 0; \qquad \text{and signature}(g_{\alpha\beta}) = +2 \text{ (or } -2). \end{aligned} \tag{2.2.1}$$

The metric $g_{\alpha\beta}(\boldsymbol{x})$ determines the spacetime squared "distance" ds^2 between two nearby events with coordinates x^α and $x^\alpha + dx^\alpha$: $ds^2 \equiv g_{\alpha\beta}dx^\alpha dx^\beta$. On a pseudo-Riemannian manifold (the spacetime), for a given vector $\boldsymbol{v}_P$ in P, the squared norm $g_{\alpha\beta}v_P^\alpha v_P^\beta$ can be positive, negative, or null, the vector is then respectively called spacelike, timelike, or null. The metric tensor with both indices up, that is, *contravariant*, $g^{\alpha\beta}$, is obtained from the *covariant* components, $g_{\alpha\beta}$, by $g^{\alpha\beta}g_{\beta\gamma} \equiv \delta^\alpha{}_\gamma$, where $\delta^\alpha{}_\gamma$ is the Kronecker tensor, 0 for $\alpha \neq \gamma$ and 1 for $\alpha = \gamma$.

Let us briefly recall the definition of a few basic quantities of tensor calculus on a Riemannian manifold;[38–43] for a more extensive description see the mathematical appendix. We shall mainly use quantities written in components and referred to a coordinate basis on an n-dimensional Riemannian manifold.

A ***p*-covariant tensor** $T_{\alpha_1\cdots\alpha_p}$, or $\boldsymbol{T}$, is a mathematical object made of n^p quantities that under a coordinate transformation, $x'^\alpha = x'^\alpha(x^\alpha)$, change according to the transformation law $T'_{\alpha_1\cdots\alpha_p} = \partial^{\beta_1\cdots\beta_p}_{\alpha'_1\cdots\alpha'_p}T_{\beta_1\cdots\beta_p}$, where $\partial^{\beta_1\cdots\beta_p}_{\alpha'_1\cdots\alpha'_p} \equiv \frac{\partial x^{\beta_1}}{\partial x'^{\alpha_1}} \cdots \frac{\partial x^{\beta_p}}{\partial x'^{\alpha_p}}$ denotes the partial derivatives of the old coordinates x^α with respect to the new coordinates x'^α : $\partial^\beta_{\alpha'} \equiv \frac{\partial x^\beta}{\partial x'^\alpha}$.

A ***q*-contravariant tensor** $T^{\alpha_1\cdots\alpha_q}$ is a mathematical object made of n^q quantities that transform according to the rule $T'^{\alpha_1\cdots\alpha_q} = \partial^{\alpha'_1\cdots\alpha'_q}_{\beta_1\cdots\beta_q}T^{\beta_1\cdots\beta_q}$ where $\partial^{\alpha'_1\cdots\alpha'_q}_{\beta_1\cdots\beta_q} \equiv \frac{\partial x'^{\alpha_1}}{\partial x^{\beta_1}} \cdots \frac{\partial x'^{\alpha_q}}{\partial x^{\beta_q}}$. The covariant and contravariant components of a tensor are obtained from each other by lowering and raising the indices with $g_{\alpha\beta}$ and $g^{\alpha\beta}$.

The **covariant derivative** ∇_γ of a tensor $T^{\alpha\cdots}{}_{\beta\cdots}$, written here with a semicolon "; γ" is a tensorial generalization to curved manifolds of the standard partial derivative of Euclidean geometry. Applied to an n-covariant, m-contravariant tensor $T^{\alpha\cdots}{}_{\beta\cdots}$ it forms a $(n+1)$-covariant, m-contravariant tensor $T^{\alpha\cdots}{}_{\beta\cdots;\gamma}$ defined as

$$T^{\alpha\cdots}{}_{\beta\cdots;\gamma} \equiv T^{\alpha\cdots}{}_{\beta\cdots,\gamma} + \Gamma^\alpha_{\sigma\gamma}T^{\sigma\cdots}{}_{\beta\cdots} - \Gamma^\sigma_{\beta\gamma}T^{\alpha\cdots}{}_{\sigma\cdots} \tag{2.2.2}$$

where the $\Gamma^{\alpha}_{\beta\gamma}$ are the **connection coefficients**. They can be constructed, on a Riemannian manifold, from the first derivatives of the metric tensor:

$$\Gamma^{\alpha}_{\gamma\beta} = \Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2} g^{\alpha\sigma} (g_{\sigma\beta,\gamma} + g_{\sigma\gamma,\beta} - g_{\beta\gamma,\sigma}) \equiv \left\{ \begin{matrix} \alpha \\ \beta\gamma \end{matrix} \right\}. \tag{2.2.3}$$

On a Riemannian manifold, in a **coordinate basis** (**holonomic basis**), the connection coefficients have the above form, $\left\{ \begin{smallmatrix} \alpha \\ \beta\gamma \end{smallmatrix} \right\}$, as a function of the metric and of its first derivatives, and are usually called Christoffel symbols (see § 2.8 and mathematical appendix). The **Christoffel symbols** $\Gamma^{\alpha}_{\beta\gamma}$ are not tensors, but transform according to the rule $\Gamma'^{\alpha}_{\beta\gamma} = \partial^{\alpha'}_{\sigma} \partial^{\mu}_{\beta'} \partial^{\nu}_{\gamma'} \Gamma^{\sigma}_{\mu\nu} + \partial^{\alpha'}_{\delta} \partial^{\delta}_{\beta'\gamma'}$ where $\partial^{\delta}_{\beta'\gamma'} \equiv \frac{\partial^2 x^{\delta}}{\partial x'^{\beta} \partial x'^{\gamma}}$.

The **Riemann curvature tensor** $R^{\alpha}{}_{\beta\gamma\delta}$ is the generalization to n-dimensional manifolds of the Gaussian curvature K of a two-dimensional surface; it is defined as the commutator of the covariant derivatives of a vector field $\boldsymbol{A}$,

$$A^{\alpha}{}_{;\beta\gamma} - A^{\alpha}{}_{;\gamma\beta} = R^{\alpha}{}_{\sigma\gamma\beta} A^{\sigma}. \tag{2.2.4}$$

In terms of the Christoffel symbols (2.2.3) the curvature is given by

$$R^{\alpha}{}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\beta\delta,\gamma} - \Gamma^{\alpha}_{\beta\gamma,\delta} + \Gamma^{\alpha}_{\sigma\gamma} \Gamma^{\sigma}_{\beta\delta} - \Gamma^{\alpha}_{\sigma\delta} \Gamma^{\sigma}_{\beta\gamma}. \tag{2.2.5}$$

The various symmetry properties of the Riemann curvature tensor are given in the mathematical appendix.

2.3 THE FIELD EQUATION

In electromagnetism[44] the four components of the electromagnetic vector potential A^{α} are connected with the density of charge ρ and with the three components of the density of current, $j^i = \rho v^i$, by the Maxwell equation

$$F^{\alpha\beta}{}_{,\beta} \equiv (A^{\beta,\alpha} - A^{\alpha,\beta})_{,\beta} = 4\pi j^{\alpha} \equiv 4\pi\rho u^{\alpha} \tag{2.3.1}$$

in flat spacetime. Here $F^{\alpha\beta} \equiv A^{\beta,\alpha} - A^{\alpha,\beta}$ is the electromagnetic field tensor, $j^{\alpha} \equiv \rho u^{\alpha}$ is the charge current density four-vector, and $u^{\alpha} \equiv \frac{dx^{\alpha}}{ds}$ is the four-velocity of the charge distribution. The comma ", β" means partial derivative with respect to x^{β}: $\frac{\partial A^{\alpha}}{\partial x^{\beta}} \equiv A^{\alpha}{}_{,\beta}$.

We search now for a field equation that will connect the gravitational tensor potential $g_{\alpha\beta}$ with the density of mass-energy and its currents. Let us follow David Hilbert[45] (1915) to derive this Einstein field equation[6] from a variational principle, or principle of least action. We are motivated by Richard Feynman's later insight that classical action for a system reveals and follows the phase of the quantum mechanical wave function of that system (see below, refs. 128 and 129). We write the total action over an arbitrary spacetime region Ω as

$$I = \int_{\Omega} (\mathcal{L}_G + \mathcal{L}_M) d^4x \tag{2.3.2}$$

where $d^4x = dx^1 \cdot dx^2 \cdot dx^3 \cdot dx^0$ and $\mathcal{L}_G$ and $\mathcal{L}_M$ are the Lagrangian densities for the geometry and for matter and fields, respectively, $\mathcal{L}_G \equiv L_G\sqrt{-g}$ and $\mathcal{L}_M \equiv L_M\sqrt{-g}$, and g is the determinant of the metric $g_{\alpha\beta}$: $g = \det(g_{\alpha\beta})$. The field variables describing the geometry, that is, the gravitational field, are the ten components of the metric tensor $g_{\alpha\beta}$. In order to have a tensorial field equation for $g_{\alpha\beta}$, we search for a $(\mathcal{L}_G + \mathcal{L}_M)$ that is a scalar density, that is, we search for an action I that is a scalar quantity. By analogy with electromagnetism we then search for a field equation of the second order in the field variables $g_{\alpha\beta}$, which, to be consistent with the observations, in the weak field and slow motion limit, must reduce to the classical Poisson equation. Therefore, the **Lagrangian density for the geometry** should contain the field variables $g_{\alpha\beta}$ and their first derivatives $g_{\alpha\beta,\gamma}$ only. In agreement with these requirements we assume

$$\mathcal{L}_G = \frac{1}{2\chi}\sqrt{-g} \cdot R. \tag{2.3.3}$$

Here $\frac{1}{2\chi}$ is a constant to be determined by requiring that we recover classical gravity theory in the weak field and slow motion limit, $R \equiv R^\alpha{}_\alpha \equiv g^{\alpha\beta}R_{\alpha\beta}$ is the **Ricci** or **curvature scalar**, and $R_{\alpha\beta}$ is the **Ricci tensor** constructed by contraction from the Riemann curvature tensor, $R_{\alpha\beta} = R^\sigma{}_{\alpha\sigma\beta}$. The curvature scalar R has a part linear in the second derivatives of the metric; however, it turns out that the variation of this part does not contribute to the field equation (see below).

Before evaluating the variation of the action I, we need to introduce a few identities and theorems, valid on a Riemannian manifold, that we shall prove at the end of this section.

1. The covariant derivative (defined by the Riemannian connection, see § 2.8) of the metric tensor $g^{\alpha\beta}$ is zero (Ricci theorem):
$$g^{\alpha\beta}{}_{;\gamma} = 0. \tag{2.3.4}$$
2. The variation, δg, with respect to $g_{\alpha\beta}$, of the determinant of the metric g is given by
$$\delta g = g \cdot g^{\alpha\beta} \cdot \delta g_{\alpha\beta} = -g \cdot g_{\alpha\beta} \cdot \delta g^{\alpha\beta}. \tag{2.3.5}$$
3. For a vector field v^α, we have the useful formula
$$v^\alpha{}_{;\alpha} = \left(\sqrt{-g}v^\alpha\right)_{,\alpha}\frac{1}{\sqrt{-g}}, \tag{2.3.6}$$
and similarly for a tensor field $T^{\alpha\beta}$
$$T^{\alpha\beta}{}_{;\beta} = \left(\sqrt{-g}T^{\alpha\beta}\right)_{,\beta}\frac{1}{\sqrt{-g}} + \Gamma^\alpha_{\sigma\beta}T^{\sigma\beta}. \tag{2.3.7}$$
4. Even though the Christoffel symbols $\Gamma^\alpha_{\beta\gamma}$ are not tensors and transform according to the rule that follows expression (2.2.3), $\Gamma'^\alpha_{\beta\gamma} = \partial^{\alpha'}_\sigma\partial^\mu_{\beta'}\partial^\nu_{\gamma'}\Gamma^\sigma_{\mu\nu} +$

$\partial^{\alpha'}_{\delta}\partial^{\delta}_{\beta'\gamma'}$, the difference between two sets of Christoffel symbols on the manifold M, $\delta\Gamma^{\alpha}_{\beta\gamma}(x) \equiv \Gamma^{*\alpha}_{\beta\gamma}(x) - \Gamma^{\alpha}_{\beta\gamma}(x)$, is a tensor. This immediately follows from the transformation rule for the $\Gamma^{\alpha}_{\beta\gamma}(x)$. The two sets of Christoffel symbols on M, $\Gamma^{*\alpha}_{\beta\gamma}(x)$ and $\Gamma^{\alpha}_{\beta\gamma}(x)$, may, for example, be related to two tensor fields, $g^{*}_{\alpha\beta}(x)$ and $g_{\alpha\beta}(x)$, on M.

5. The variation $\delta R_{\alpha\beta}$ of the Ricci tensor $R_{\alpha\beta}$ is given by

$$\delta R_{\alpha\beta} = \left(\delta\Gamma^{\sigma}_{\alpha\beta}\right)_{;\sigma} - \left(\delta\Gamma^{\sigma}_{\alpha\sigma}\right)_{;\beta}. \tag{2.3.8}$$

6. The generalization of the **Stokes divergence theorem**, to a **four-dimensional** manifold M, is

$$\int_{\Omega} v^{\sigma}{}_{;\sigma}\sqrt{-g}\,d^4x = \int_{\Omega}\left(v^{\sigma}\sqrt{-g}\right)_{,\sigma}d^4x = \int_{\partial\Omega}\sqrt{-g}\,v^{\sigma}d^3\Sigma_{\sigma}. \tag{2.3.9}$$

Here v^{σ} is a vector field, Ω is a four-dimensional spacetime region, $d^4x = dx^0\,dx^1\,dx^2\,dx^3$ its four-dimensional integration element, $\partial\Omega$ is the three-dimensional boundary (with the induced orientation; see § 2.8 and mathematical appendix) of the four-dimensional region Ω, and $d\Sigma_{\sigma}$ the three-dimensional integration element of $\partial\Omega$ (see § 2.8).

We now require the action I to be stationary for arbitrary variations $\delta g^{\alpha\beta}$ of the field variables $g^{\alpha\beta}$, with certain derivatives of $g^{\alpha\beta}$ fixed on the boundary of Ω: $\delta I = 0$. By using expression (2.3.5) we then find that

$$\delta I = \frac{1}{2\chi}\int_{\Omega}\left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\right)\sqrt{-g}\,\delta g^{\alpha\beta}d^4x + \frac{1}{2\chi}\int_{\Omega}g^{\alpha\beta}\sqrt{-g}\,\delta R_{\alpha\beta}d^4x + \int_{\Omega}\frac{\delta\mathcal{L}_M}{\delta g^{\alpha\beta}}\,\delta g^{\alpha\beta}d^4x = 0. \tag{2.3.10}$$

The second term of this equation can be written

$$\begin{aligned}\frac{1}{2\chi}\int_{\Omega}g^{\alpha\beta}\sqrt{-g}\,\delta R_{\alpha\beta}d^4x &= \frac{1}{2\chi}\int_{\Omega}g^{\alpha\beta}\sqrt{-g}\left[\left(\delta\Gamma^{\sigma}_{\alpha\beta}\right)_{;\sigma} - \left(\delta\Gamma^{\sigma}_{\alpha\sigma}\right)_{;\beta}\right]d^4x\\ &= \frac{1}{2\chi}\int_{\Omega}\sqrt{-g}\left[\left(g^{\alpha\beta}\delta\Gamma^{\sigma}_{\alpha\beta}\right)_{;\sigma} - \left(g^{\alpha\beta}\delta\Gamma^{\sigma}_{\alpha\sigma}\right)_{;\beta}\right]d^4x\\ &= \frac{1}{2\chi}\int_{\Omega}\left[\left(\sqrt{-g}\,g^{\alpha\beta}\delta\Gamma^{\sigma}_{\alpha\beta}\right) - \left(\sqrt{-g}\,g^{\alpha\sigma}\delta\Gamma^{\rho}_{\alpha\rho}\right)\right]_{,\sigma}d^4x.\end{aligned} \tag{2.3.11}$$

where $\delta\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}g^{\alpha\sigma}[(\delta g_{\beta\sigma})_{;\gamma} + (\delta g_{\sigma\gamma})_{;\beta} - (\delta g_{\gamma\beta})_{;\sigma}]$. This is an integral of a divergence and by the four-dimensional Gauss theorem can be transformed into an integral over the boundary $\partial\Omega$ of Ω, where it vanishes if certain derivatives of $g_{\alpha\beta}$ are fixed on the boundary $\partial\Omega$ of Ω. Then, this term gives no contribution to the field equation. Indeed, the integral over the boundary $\partial\Omega = \sum_I S_I$ of Ω can be rewritten (see York 1986)[46] as

$$\sum_I \frac{\varepsilon_I}{2\chi} \int_{S_I} \gamma_{\alpha\beta}\delta N^{\alpha\beta} d^3x \tag{2.3.12}$$

where $\varepsilon_I \equiv \boldsymbol{n}_I \cdot \boldsymbol{n}_I = \pm 1$ and $\boldsymbol{n}_I$ is the unit vector field normal to the hypersurface S_I, $\gamma_{\alpha\beta} = g_{\alpha\beta} - \varepsilon_I n_\alpha n_\beta$ is the three-metric on each hypersurface S_I of the boundary $\partial\Omega$ of Ω, and

$$N^{\alpha\beta} \equiv \sqrt{|\gamma|}(K\gamma^{\alpha\beta} - K^{\alpha\beta}) = -\frac{1}{2} g\gamma^{\alpha\mu}\gamma^{\beta\nu} \pounds_{\boldsymbol{n}}(g^{-1}\gamma_{\mu\nu}) \tag{2.3.13}$$

where γ is the three-dimensional determinant of $\gamma_{\alpha\beta}$, $K_{\alpha\beta} = -\frac{1}{2}\pounds_{\boldsymbol{n}}\gamma_{\alpha\beta}$ is the so-called second fundamental form or "extrinsic curvature" of each S_I (see § 5.2.2 and mathematical appendix), $K \equiv \gamma^{\alpha\beta}K_{\alpha\beta}$, and $\pounds_{\boldsymbol{n}}$ is the Lie derivative (see § 4.2 and mathematical appendix) along the normal $\boldsymbol{n}$ to the boundary $\partial\Omega$ of Ω. Therefore, if the quantities $N^{\alpha\beta}$ are fixed on the boundary $\partial\Omega$, for an arbitrary variation $\delta g^{\alpha\beta}$, from the first and last integrals of (2.3.10), we have the **field equation**

$$\boxed{G_{\alpha\beta} = \chi T_{\alpha\beta}} \tag{2.3.14}$$

where $G_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta}$ is the **Einstein tensor**, and—following the last integral of 2.3.10—we have defined the **energy-momentum tensor** of matter and fields $T_{\alpha\beta}$ (see below) as:

$$T_{\alpha\beta} \equiv -2\frac{\delta L_M}{\delta g^{\alpha\beta}} + L_M g_{\alpha\beta}. \tag{2.3.15}$$

Let us now determine the constant χ by comparison with the classical, weak field, Poisson equation, $\Delta U = -4\pi\rho$, where U is the standard Newtonian gravitational potential. We first observe that in any metric theory of gravity (see chap. 3), without any assumption on the field equations, in the weak field and slow motion limit (see § 3.7), the metric $\boldsymbol{g}$ can be written at the lowest order in U, $g_{00} \cong -1 + 2U$, $g_{ik} \cong \delta_{ik}$, and $g_{i0} \cong 0$ and the energy-momentum tensor, at the lowest order, $T_{00} \cong -T \cong \rho$. From the definition of Ricci tensor $R_{\alpha\beta}$, it then follows that $R_{00} \cong -\Delta U$. From the field equation (2.3.14) we also have

$$R^\alpha{}_\alpha - \frac{1}{2}R\delta^\alpha{}_\alpha = -R = \chi T^\alpha{}_\alpha \equiv \chi T \tag{2.3.16}$$

where $T \equiv T^\alpha{}_\alpha$ is the trace of $T^{\alpha\beta}$. Therefore, the field equation can be rewritten in the alternative form

$$R_{\alpha\beta} = \chi\left(T_{\alpha\beta} - \frac{1}{2}Tg_{\alpha\beta}\right). \tag{2.3.17}$$

From the 00 component of this equation, in the weak field and slow motion limit, we have

$$R_{00} \cong \chi\Big(T_{00} + \frac{1}{2}T\Big), \tag{2.3.18}$$

and therefore

$$\Delta U \cong -\frac{\chi}{2}\rho. \tag{2.3.19}$$

Requiring the agreement of the very weak field limit of general relativity with the classical Newtonian theory and comparing this equation (2.3.19) with the classical Poisson equation, we finally get $\chi = 8\pi$.

An alternative method of variation—the **Palatini method**[47]—is to take as independent field variables not only the ten components $g^{\alpha\beta}$ but also the forty components of the affine connection $\Gamma^{\alpha}_{\beta\gamma}$, assuming, a priori, no dependence of the $\Gamma^{\alpha}_{\beta\gamma}$ from the $g^{\alpha\beta}$ and their derivatives. Taking the variation with respect to the $\Gamma^{\alpha}_{\beta\gamma}$ and the $g^{\alpha\beta}$, and assuming L_M to be independent from any derivative of $g^{\alpha\beta}$, we thus have

$$\begin{aligned}
&\frac{1}{2\chi}\int_\Omega \Big(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\Big)\delta g^{\alpha\beta}\sqrt{-g}\,d^4x \\
&\quad + \frac{1}{2\chi}\int_\Omega g^{\alpha\beta}\Big(\delta\Gamma^{\sigma}_{\alpha\beta;\sigma} - \delta\Gamma^{\sigma}_{\alpha\sigma;\beta}\Big)\sqrt{-g}\,d^4x \\
&\quad + \int_\Omega \Big(\frac{\delta L_M}{\delta g^{\alpha\beta}} - \frac{1}{2}g_{\alpha\beta}L_M\Big)\delta g^{\alpha\beta}\sqrt{-g}\,d^4x = 0.
\end{aligned} \tag{2.3.20}$$

From the second integral, after some calculations,[11] one then gets

$$g_{\alpha\beta;\gamma} = g_{\alpha\beta,\gamma} - g_{\alpha\sigma}\Gamma^{\sigma}_{\beta\gamma} - g_{\sigma\beta}\Gamma^{\sigma}_{\alpha\gamma} = 0, \tag{2.3.21}$$

and therefore, by calculating from expression (2.3.21): $g^{\alpha\sigma}(g_{\beta\sigma,\gamma} + g_{\sigma\gamma,\beta} - g_{\beta\gamma,\sigma})$, on a Riemannian manifold, one gets the expression of the affine connection as a function of the $g_{\alpha\beta}$, that is, the Christoffel symbols $\left\{ {\alpha \atop \beta\gamma} \right\}$

$$\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}g^{\alpha\sigma}\big(g_{\beta\sigma,\gamma} + g_{\sigma\gamma,\beta} - g_{\beta\gamma,\sigma}\big) \equiv \left\{ {\alpha \atop \beta\gamma} \right\}. \tag{2.3.22}$$

From the first and third integral in expression (2.3.20), we finally have the field equation (2.3.14).

Let us give the expression of the energy-momentum tensor in two cases: an electromagnetic field and a matter fluid.

In special relativity the energy-momentum tensor for an electromagnetic field[44] is $T^{\alpha\beta} = \frac{1}{4\pi}(F^{\alpha}{}_{\sigma}F^{\beta\sigma} - \frac{1}{4}\eta^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta})$, where $F^{\alpha\beta}$ is the electromagnetic field tensor. Moreover the *energy-momentum tensor*[131,132] of a matter fluid can be written $T^{\alpha\beta} = (\varepsilon + p)u^{\alpha}u^{\beta} + (q^{\alpha}u^{\beta} + u^{\alpha}q^{\beta}) + p\eta^{\alpha\beta} + \pi^{\alpha\beta}$, where ε is the *total energy density* of the fluid, u^α its *four-velocity*, q^α the *energy flux* relative to u^α (*heat flow*), p the *isotropic pressure*, and $\pi^{\alpha\beta}$ the tensor representing *viscous stresses* in the fluid. Therefore, by replacing $\eta_{\alpha\beta}$ with $g_{\alpha\beta}$ (in agreement with the equivalence principle), we define in Einstein geometrodynamics:

$$T^{\alpha\beta} = \frac{1}{4\pi}\left(F^{\alpha}{}_{\sigma}F^{\beta\sigma} - \frac{1}{4}g^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right) \tag{2.3.23}$$

for an electromagnetic field, and

$$T^{\alpha\beta} = (\varepsilon + p)u^\alpha u^\beta + (q^\alpha u^\beta + u^\alpha q^\beta) + pg^{\alpha\beta} + \pi^{\alpha\beta} \tag{2.3.24}$$

for a matter fluid, where $\pi^{\alpha\beta}$ may be written:[11] $\pi^{\alpha\beta} = -2\eta\sigma^{\alpha\beta} - \zeta\Theta(g^{\alpha\beta} + u^\alpha u^\beta)$, where η is the *coefficient of shear viscosity*, ζ the *coefficient of bulk viscosity*, and $\sigma^{\alpha\beta}$ and Θ are the *shear tensor* and the *expansion scalar* of the fluid (see § 4.5).

In the case of a perfect fluid, defined by $\pi_{\alpha\beta} = q_\alpha = 0$, we then have

$$T^{\alpha\beta} = (\varepsilon + p)u^\alpha u^\beta + pg^{\alpha\beta}. \tag{2.3.25}$$

The general relativity expressions (2.3.23) and (2.3.24), for the energy-momentum tensor of an electromagnetic field and for a matter fluid, agree with the previous definition (2.3.15) of energy-momentum tensor, with a proper choice of the matter and fields Lagrangian density $\mathcal{L}_M$.

Let us finally prove the identities used in this section.

1. From the definition of covariant derivative and Christoffel symbols, we have

$$\begin{aligned} g^{\alpha\beta}{}_{;\gamma} &= g^{\alpha\beta}{}_{,\gamma} + \frac{1}{2}g^{\alpha\mu}g^{\beta\nu}(g_{\gamma\nu,\mu} + g_{\nu\mu,\gamma} - g_{\mu\gamma,\nu}) \\ &\quad + \frac{1}{2}g^{\mu\beta}g^{\alpha\nu}(g_{\mu\nu,\gamma} + g_{\nu\gamma,\mu} - g_{\gamma\mu,\nu}) \\ &= g^{\alpha\beta}{}_{,\gamma} + g^{\alpha\mu}g^{\beta\nu}g_{\nu\mu,\gamma} \\ &= g^{\alpha\beta}{}_{,\gamma} + g^{\alpha\beta}{}_{,\gamma} - g^{\beta\nu}{}_{,\gamma}g^{\alpha\mu}g_{\nu\mu} - g^{\alpha\mu}{}_{,\gamma}g^{\beta\nu}g_{\nu\mu} = 0. \end{aligned} \tag{2.3.26}$$

2. By using the symbol $\delta^{\alpha\beta\gamma\lambda}_{\mu\nu\rho\sigma}$, defined to be equal to $+1$ if $\alpha\beta\gamma\lambda$ is an even permutation of $\mu\nu\rho\sigma$, equal to -1 if $\alpha\beta\gamma\lambda$ is an odd permutation of $\mu\nu\rho\sigma$, and 0 otherwise (see § 2.8), we can write the determinant of a 4×4 tensor, $g_{\alpha\beta}$, in the form

$$g \equiv \det g_{\alpha\beta} = \delta^{\alpha\beta\gamma\lambda}_{0\,1\,2\,3}g_{\alpha 0}\,g_{\beta 1}\,g_{\gamma 2}\,g_{\lambda 3}. \tag{2.3.27}$$

By taking the variation of g we then have

$$\delta g = \delta g_{\alpha\beta} \cdot (g^{\alpha\beta} \cdot g) \tag{2.3.28}$$

and therefore

$$\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}\, g^{\alpha\beta}\delta g_{\alpha\beta}. \tag{2.3.29}$$

Moreover, from $\delta(g^{\alpha\beta}g_{\alpha\beta}) = 0$, we also have

$$\delta g_{\alpha\beta}\cdot g^{\alpha\beta} = -\delta g^{\alpha\beta}\cdot g_{\alpha\beta}. \tag{2.3.30}$$

3. From the definition (2.3.22) of Christoffel symbols, we have

$$\Gamma^{\sigma}_{\sigma\alpha} = \frac{1}{2} g_{\mu\nu,\alpha}g^{\mu\nu}, \tag{2.3.31}$$

and therefore, from the rule for differentiation of a determinant, $g_{,\alpha} = g g^{\mu\nu} g_{\mu\nu,\alpha}$, we get

$$\left(\ln\sqrt{-g}\right)_{,\alpha} = \Gamma^{\sigma}_{\sigma\alpha} \tag{2.3.32}$$

and finally

$$v^{\alpha}{}_{;\alpha} = v^{\alpha}{}_{,\alpha} + v^{\sigma}\Gamma^{\alpha}_{\alpha\sigma} = \left(\sqrt{-g}\, v^{\alpha}\right)_{,\alpha}\frac{1}{\sqrt{-g}}. \tag{2.3.33}$$

4. From the rule for transformation of the connection coefficients, it immediately follows that the difference between two sets of connection coefficients is a tensor.
5. At any event of the spacetime Lorentzian manifold, we can find infinitely many local inertial frames of reference where $\overset{(i)}{g}_{\alpha\beta} = \eta_{\alpha\beta}$, $\overset{(i)}{g}_{\alpha\beta,\gamma} = 0$ and therefore $\overset{(i)}{\Gamma}{}^{\alpha}_{\mu\nu} = 0$. From the definition of Ricci tensor (contraction of the Riemann tensor $R^{\alpha}{}_{\beta\gamma\delta}$ of expression (2.2.5) on the two indices α and γ) we then have at the event in any such local inertial frame

$$\delta\overset{(i)}{R}_{\alpha\beta} = \left(\delta\overset{(i)}{\Gamma}{}^{\sigma}_{\alpha\beta}\right)_{,\sigma} - \left(\delta\overset{(i)}{\Gamma}{}^{\sigma}_{\alpha\sigma}\right)_{,\beta}, \tag{2.3.34}$$

or equivalently

$$\delta\overset{(i)}{R}_{\alpha\beta} = \left(\delta\overset{(i)}{\Gamma}{}^{\sigma}_{\alpha\beta}\right)_{;\sigma} - \left(\delta\overset{(i)}{\Gamma}{}^{\sigma}_{\alpha\sigma}\right)_{;\beta}, \tag{2.3.35}$$

and since this is a tensorial equation, it is valid in any coordinate system

$$\delta R_{\alpha\beta} = \left(\delta\Gamma^{\sigma}_{\alpha\beta}\right)_{;\sigma} - \left(\delta\Gamma^{\sigma}_{\alpha\sigma}\right)_{;\beta}. \tag{2.3.36}$$

2.4 EQUATIONS OF MOTION

According to the *field equation*, $G^{\alpha\beta} = \chi T^{\alpha\beta}$, mass-energy $T^{\alpha\beta}$ "tells" geometry $g^{\alpha\beta}$ how to "curve"; furthermore, from the field equation itself, geometry "tells" mass-energy how to move. The key to the proof is **Bianchi's second**

identity[48,49] (for the "boundary of a boundary interpretation" of which see § 2.8):

$$R^{\alpha}{}_{\beta\gamma\nu;\mu} + R^{\alpha}{}_{\beta\nu\mu;\gamma} + R^{\alpha}{}_{\beta\mu\gamma;\nu} = 0.$$

Raising the indices β and ν and summing over α and γ, and over β and μ, we get the **contracted Bianchi's identity**:

$$G^{\nu\beta}{}_{;\beta} = \left(R^{\nu\beta} - \frac{1}{2}Rg^{\nu\beta}\right)_{;\beta} = 0. \tag{2.4.1}$$

By taking the covariant divergence of both sides of the field equation (2.3.14), we get

$$\boxed{T^{\nu\beta}{}_{;\beta} = 0.} \tag{2.4.2}$$

This statement summarizes the dynamical equations for matter and fields described by the energy-momentum tensor $T^{\alpha\beta}$. Therefore, as a consequence of the field equation, we have obtained the **dynamical equations** for matter and fields.

There exists an alternative approach to get the contracted Bianchi's identity. Consider an infinitesimal coordinate transformation:

$$x'^{\alpha} = x^{\alpha} - \xi^{\alpha}. \tag{2.4.3}$$

Under this transformation the metric tensor changes to (see § 4.2)

$$g'_{\alpha\beta} = g_{\alpha\beta} + \delta g_{\alpha\beta} = g_{\alpha\beta} + \xi_{\alpha;\beta} + \xi_{\beta;\alpha}. \tag{2.4.4}$$

This coordinate change bringing with it no real change in the geometry or the physics, we know that the action cannot change with this alteration. In other words, from the variational principle, $\delta \int \mathcal{L}_G \, d^4x = 0$, corresponding to the variation $\delta g^{\alpha\beta} = \xi^{\alpha;\beta} + \xi^{\beta;\alpha}$, we have

$$\delta I_G = \frac{1}{2\chi} \int G_{\alpha\beta}\left(\xi^{\alpha;\beta} + \xi^{\beta;\alpha}\right)\sqrt{-g}\,d^4x = 0. \tag{2.4.5}$$

We translate

$$G_{\alpha\beta}\xi^{\alpha;\beta} = -G_{\alpha\beta}{}^{;\beta}\xi^{\alpha} + (G_{\alpha\beta}\xi^{\alpha})^{;\beta} = -G_{\alpha\beta}{}^{;\beta}\xi^{\alpha} + \frac{1}{\sqrt{-g}}\left(\sqrt{-g}G_{\alpha}{}^{\beta}\xi^{\alpha}\right)_{,\beta}$$

and use the four-dimensional divergence theorem (2.3.9), to get

$$\delta I_G = -\frac{1}{\chi} \int G_{\alpha\beta}{}^{;\beta}\xi^{\alpha}\sqrt{-g}\,d^4x = 0. \tag{2.4.6}$$

Since I_G is a scalar its value is independent of coordinate transformations; therefore this expression must be zero for every infinitesimal ξ^{α}, whence the contracted Bianchi identities (2.4.1).

For a pressureless perfect fluid, $p = 0$, that is, for dust particles, from expression (2.3.25) we have

$$T^{\alpha\beta} = \varepsilon u^\alpha u^\beta, \tag{2.4.7}$$

and from the equation of motion $T^{\alpha\beta}{}_{;\beta} = 0$,

$$T^{\alpha\beta}{}_{;\beta} = (\varepsilon u^\alpha u^\beta)_{;\beta} = u^\alpha{}_{;\beta}\varepsilon u^\beta + \left(\varepsilon u^\beta\right)_{;\beta} u^\alpha = 0. \tag{2.4.8}$$

Multiplying this equation by u_α (and summing over α), recognizing $u^\alpha u_\alpha = -1$, and $(u^\alpha u_\alpha)_{;\beta} = 0$ or $u^\alpha{}_{;\beta} u_\alpha = 0$, we get $(\varepsilon u^\beta)_{;\beta} = 0$. Then, on substituting this result back into equation (2.4.8) we obtain the **geodesic equation**

$$u^\alpha{}_{;\beta} u^\beta = 0. \tag{2.4.9}$$

Therefore, each particle of dust follows a geodesic,[50,51] in agreement with the equivalence principle and with the equation of motion of special relativity, $\frac{du^\alpha}{ds} = u^\alpha{}_{,\beta} u^\beta = 0$. In a local inertial frame, from expression (2.4.8), we get to lowest order the classical equation of continuity, $\rho_{,0} + (\rho v^i)_{,i} = 0$, and also the Euler equations for fluid motion, $\rho (v^i)_{,0} + \rho (v^i)_{,k} v^k = 0$, where ρ is the fluid mass density.

In general, we assume that the equation of motion of any test particle is a geodesic, where we define[39] a **geodesic** as the **extremal curve**, or history, $x^\alpha(t)$ that extremizes the integral of half of the squared interval E^b_a between two events $a = x(t_a)$ and $b = x(t_b)$:

$$E^b_a(x(t)) \equiv \frac{1}{2} \int_{t_a}^{t_b} g_{\alpha\beta}\Big(x(t)\Big) \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}\, dt. \tag{2.4.10}$$

In this sense a geodesic counts as a critical point in the space of all histories. We demand that any first-order small change $\delta x^\alpha(t)$ of the history, that keeps the end point fixed $\delta x^\alpha(t_a) = \delta x^\alpha(t_b) = 0$, shall cause a change in the integral $E^b_a(x(t))$ that is of higher order. The first-order change is required to vanish: $\delta E^b_a(x^\alpha(t)) = 0$. It is the integral of the product of $\delta x^\alpha(t)$ with the Lagrange expression:

$$\frac{\partial L}{\partial x^\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \left(\frac{dx^\alpha}{dt}\right)} = 0, \tag{2.4.11}$$

where $L = \frac{1}{2} g_{\alpha\beta}(x(t)) \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}$, and we have

$$g_{\alpha\beta} \frac{d^2 x^\beta}{dt^2} + g_{\alpha\beta,\gamma} \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt} - \frac{1}{2} g_{\beta\gamma,\alpha} \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt} = 0. \tag{2.4.12}$$

This **equation for a geodesic** translates into the language (2.3.22) of the Christoffel symbols:

$$\boxed{\frac{d^2x^\alpha}{dt^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{dt}\frac{dx^\gamma}{dt} = 0.} \tag{2.4.13}$$

The geodesic equation keeps the standard form (2.4.13) for every transformation of the parameter t of the type $s = ct + d$, where $c \neq 0$ and d are two constants; when the geodesic equation has the standard form (2.4.13), t is called **affine parameter**. A special choice of parameter p is the **arc-length** itself $s(p)$ along the curve $s(p) = L^p_a(x) = \int_a^p \sqrt{\pm g_{\alpha\beta}(x(p'))\frac{dx^\alpha}{dp'}\frac{dx^\beta}{dp'}}\,dp'$ (+ sign for spacelike geodesics and $-$ sign for timelike geodesics), where p is a parameter along the curve. When $p = s$, the geodesic is said to be parametrized by arc-length. For a timelike geodesic, $s \equiv \tau$ is the **proper time** measured by a clock comoving with the test particle ("wrist-watch time").

On a proper Riemannian manifold there is a variational principle that gives the geodesic equation parametrized with any parameter. This principle defines a geodesic[39] as the **extremal curve for the length** $L^a_b(x(p))$:

$$L^b_a\big(x(p)\big) = \int_{p_a}^{p_b}\sqrt{g_{\alpha\beta}\big(x(p)\big)\frac{dx^\alpha}{dp}\frac{dx^\beta}{dp}}\,dp. \tag{2.4.14}$$

From

$$\delta L^b_a\big(x(p)\big) = 0 \tag{2.4.15}$$

for any variation $\delta x^\alpha(p)$ of the curve $x^\alpha(p)$, such that $\delta x^\alpha(p_a) = \delta x^\alpha(p_b) = 0$, taking the variation of $L^b_a(x(p))$, from the Lagrange equations, we thus find

$$\frac{d^2x^\alpha}{dp^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{dp}\frac{dx^\gamma}{dp} - \frac{dx^\alpha}{dp}\left(\frac{d^2s/dp^2}{ds/dp}\right) = 0 \tag{2.4.16}$$

where $s(p)$ is the arc-length.

Extremal curve for the quantity E^b_a and extremal curve for the length L^b_a? When are the two the same on a proper Riemannian manifold? When and only when the two equations (2.4.13) and (2.4.16) are both satisfied: that is, when the quantity $\frac{d^2s}{dp^2}$ vanishes—that is, when the parameter p grows linearly with arc-length. Therefore, an extremal curve for the quantity E^b_a is also an extremal curve for the length, L^b_a, and vice versa; it is always possible[39] to reparametrize a curve that on a proper Riemannian manifold is an extremal curve for the length and with $\frac{dx^\alpha}{dp} \neq 0$ everywhere, to give an extremal curve for the quantity E^b_a.

For a test particle with proper mass different from zero, the geodesic equation of motion is the curve that extremizes the proper time $\tau = \int d\tau = \int\sqrt{-g_{\alpha\beta}dx^\alpha dx^\beta}$ along the world line of the particle. For a photon, the equation of motion follows from the variational principle for E^b_a, (2.4.10), and is a null geodesic (with $ds^2 = 0$), in agreement with special relativity and with the equivalence principle. On a timelike geodesic, we can write

$$\frac{D}{d\tau} u^\alpha = 0 \tag{2.4.17}$$

where τ is the proper time measured by a clock moving on the geodesic, $u^\alpha \equiv \frac{dx^\alpha}{d\tau}$ its four-velocity, and $u^\alpha u_\alpha = -1$.

Parallel transport of a vector v^α along a curve $x^\alpha(t)$, with tangent vector $n^\alpha(t) \equiv \frac{dx^\alpha}{dt}(t)$, is defined by requiring $\boldsymbol{n} \cdot \boldsymbol{v}$ to be covariantly constant along the curve:

$$\frac{D}{dt}(n^\alpha v_\alpha) = (n^\alpha v_\alpha)_{;\beta} n^\beta = 0. \tag{2.4.18}$$

Therefore, for a geodesic, from equation (2.4.13), we have that $v^\alpha{}_{;\beta} n^\beta = 0$.

In particular, *a geodesic is a curve with tangent vector*, n^α, *transported parallel to itself all along the curve*: $n^\alpha{}_{;\beta} n^\beta = 0$.

Finally, from the definition (2.2.5) of Riemann tensor, one can derive[39] the formula for the change of a vector v^α parallel transported around an infinitesimal closed curve determined by the infinitesimal displacements δx^α and $\widetilde{\delta x}^\alpha$ (infinitesimal "quadrilateral" which is closed apart from higher order infinitesimals in $\delta x \cdot \widetilde{\delta x}$):

$$\delta v^\alpha = -R^\alpha{}_{\beta\mu\nu} v^\beta \delta x^\mu \widetilde{\delta x}^\nu. \tag{2.4.19}$$

This equation shows that, on a curved manifold, the vector obtained by parallel transport along a curve depends on the path chosen and on the curvature (and on the initial vector; see fig. 2.1).

2.5 THE GEODESIC DEVIATION EQUATION

A fundamental equation of Einstein geometrodynamics and other metric theories of gravity is the **equation of geodesic deviation**.[38,52] It connects the spacetime curvature described by the Riemann tensor with a measurable physical quantity: the relative "acceleration" between two nearby test particles.

The equation of geodesic deviation, published in 1925 by Levi-Civita,[38,52] gives the second covariant derivative of the distance between two infinitesimally close geodesics, on an arbitrary n-dimensional Riemannian manifold:

$$\frac{D^2(\delta x^\alpha)}{ds^2} = -R^\alpha{}_{\beta\mu\nu} u^\beta \delta x^\mu u^\nu. \tag{2.5.1}$$

Here, δx^α is the infinitesimal vector that connects the geodesics, $u^\mu = \frac{dx^\mu[s]}{ds}$ is the tangent vector to the base geodesic, and $R^\alpha{}_{\mu\nu\delta}$ is the Riemann curvature tensor. This equation generalizes the classical **Jacobi equation** for the distance

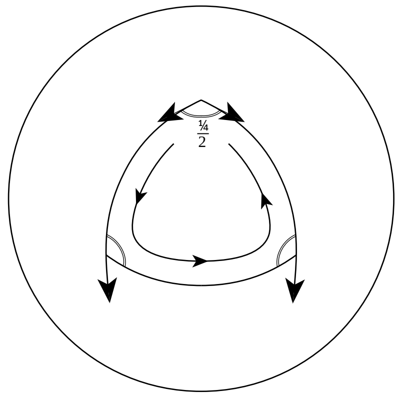


FIGURE 2.1. A vector transported parallel to itself around the indicated circuit, on the surface of a sphere of radius R, comes back to its starting point rotated through an angle of $\frac{\pi}{2}$. The curvature of the surface is given by

$$\text{(curvature)} = \frac{\text{(angle of rotation)}}{\text{(area circumnavigated)}} = \frac{\frac{\pi}{2}}{\frac{1}{8}(4\pi R^2)} = \frac{1}{R^2}.$$

y between two geodesics on a two-dimensional surface:

$$\frac{d^2 y}{d\sigma^2} + K y = 0 \tag{2.5.2}$$

where σ is the arc of the base geodesic and $K[\sigma]$ is the *Gaussian curvature* of the surface.[31,39]

The equation of geodesic deviation can be derived from the second variation of the quantity $E^b_a(x(t))$, defined by expression (2.4.10), set equal to zero. However, we follow here a more intuitive approach.

In order to derive the geodesic deviation equation (2.5.1) let us assume the following:

1. The two curves are geodesics:

$$\frac{Du_1^\alpha}{d\tau} = 0 \qquad \text{and} \qquad \frac{Du_2^\alpha}{d\sigma} = 0 \tag{2.5.3}$$

where τ, σ are affine parameters.

2. The law of correspondence between the points of the two geodesics—that is, the definition of the connecting vector $\delta x^\alpha[\tau]$—is such that, if $d\tau$ is an infinitesimal arc on geodesic 1 and $d\sigma$ the arc on geodesic 2 corresponding to the connecting vectors $\delta x^\alpha[\tau]$ and $\delta x^\alpha[\tau + d\tau]$, we have[38]

$$\frac{d\sigma}{d\tau} = 1 + \lambda, \qquad \text{where} \quad \frac{d\lambda}{d\tau} = 0 \tag{2.5.4}$$

3. The geodesics are infinitesimally close in a neighborhood U:

$$x_2^{\alpha}[\sigma] = x_1^{\alpha}[\tau] + \delta x^{\alpha}[\tau] \tag{2.5.5}$$

where $x_2^{\alpha} \epsilon U$ and $x_1^{\alpha} \epsilon U$, and where the relative change in the curvature is small:

$$\left| \frac{\mathcal{R}_{,\alpha} \delta x^{\alpha}}{\mathcal{R}} \right| \ll 1, \tag{2.5.6}$$

and $\mathcal{R}^{-2}$ is approximately the typical magnitude of the components of the Riemann tensor.

4. The difference between the tangent vectors to the two geodesics is infinitesimally small in the neighborhood U:

$$\left| \frac{\|\delta u^{\alpha}\|}{\|u^{\alpha}\|} \right| \ll 1 \tag{2.5.7}$$

where

$$\delta u^{\alpha} \equiv u_2^{\alpha}[\sigma] - u_1^{\alpha}[\tau]. \tag{2.5.8}$$

5. Equation (2.5.1) is derived neglecting terms higher than the first-order, ϵ^1, in δx^{α} and in δu^{α}. Furthermore, for simplicity, we define the connecting vector δx^{α} as connecting points of equal arc-lengths s on the two geodesics,* then, $\delta\tau = \delta\sigma = ds$ and s satisfies

$$u_1^{\alpha}[s] u_{1\alpha}[s] = -1, \qquad \text{where} \quad u_1^{\alpha}[s] \equiv \frac{dx_1^{\alpha}[s]}{ds} \tag{2.5.9}$$

and

$$u_2^{\alpha}[s] u_{2\alpha}[s] = -1, \qquad \text{where} \quad u_2^{\alpha}[s] \equiv \frac{dx_2^{\alpha}[s]}{ds}. \tag{2.5.10}$$

Physically s is the proper time measured by two observers comoving with two test particles following the two geodesics.

The equation of geodesic 1 is

$$\frac{Du_1^{\alpha}}{ds} = \frac{du_1^{\alpha}}{ds} + \Gamma^{\alpha}_{\mu\nu}[x_1] u_1^{\mu} u_1^{\nu} = 0, \tag{2.5.11}$$

and the equation of geodesic (2) is

$$\begin{aligned} \frac{Du_2^{\alpha}}{ds} &= \frac{du_2^{\alpha}}{ds} + \Gamma^{\alpha}_{\mu\nu}[x_1 + \delta x] u_2^{\mu} u_2^{\nu} = \frac{d^2}{ds^2}\left(x_1^{\alpha} + \delta x^{\alpha}\right) \\ &+ \Gamma^{\alpha}_{\mu\nu}[x_1 + \delta x] \frac{d}{ds}\left(x_1^{\mu} + \delta x^{\mu}\right) \frac{d}{ds}\left(x_1^{\nu} + \delta x^{\nu}\right) = 0. \end{aligned} \tag{2.5.12}$$

*For simplicity, in this derivation we do not consider null geodesics.

We also have

$$\frac{d}{ds}\left(\delta x^{\mu}[s]\right) \equiv \frac{d}{ds}\left(x_2^{\mu}[s] - x_1^{\mu}[s]\right) = u_2^{\mu}[s] - u_1^{\mu}[s] \equiv \delta u^{\mu}[s] \quad (2.5.13)$$

with this notation, and writing $u_1^{\mu} \equiv u^{\mu}$, we can rewrite equation (2.5.12), with a Taylor expansion to first order in δx^{α} and δu^{α}, as

$$\frac{d^2}{ds^2}\left(x_1^{\alpha}\right) + \frac{d^2}{ds^2}\left(\delta x^{\alpha}\right) + \left(\Gamma^{\alpha}_{\mu\nu} + \Gamma^{\alpha}_{\mu\nu,\rho}\delta x^{\rho}\right)\left(u^{\mu}u^{\nu} + 2u^{\mu}\delta u^{\nu}\right) = 0. \quad (2.5.14)$$

Taking the difference between equations (2.5.14) and (2.5.11) we find, to first order,

$$\frac{d^2(\delta x^{\alpha})}{ds^2} + \Gamma^{\alpha}_{\mu\nu,\rho}\delta x^{\rho}u^{\mu}u^{\nu} + 2\Gamma^{\alpha}_{\mu\nu}u^{\mu}\delta u^{\nu} = 0, \quad (2.5.15)$$

and using the definition $\frac{Dv^{\alpha}}{ds} = \frac{dv^{\alpha}}{ds} + \Gamma^{\alpha}_{\mu\nu}u^{\mu}v^{\nu}$ and the expression (2.2.5) of the Riemann tensor in terms of the Christoffel symbols and their derivatives, we have, to first order, the law of change of the geodesic separation,

$$\frac{D^2(\delta x^{\alpha})}{ds^2} = -R^{\alpha}{}_{\beta\mu\nu}u^{\beta}\delta x^{\mu}u^{\nu}. \quad (2.5.16)$$

In electromagnetism,[44] one can determine all the six independent components of the antisymmetric electromagnetic field tensor $F^{\alpha\beta}$, by measuring the accelerations of test charges in the field, and by using the Lorentz force equation

$$\frac{d^2x^{\alpha}}{ds^2} = \frac{e}{m}F^{\alpha}{}_{\beta}u^{\beta} \quad (2.5.17)$$

where e, m, and u^{β} are charge, mass, and four-velocity of the test particles. In electromagnetism, it turns out that the minimum number of test particles, with proper initial conditions, necessary to fully measure $F^{\alpha\beta}$ is two.[11]

Similarly, on a Lorentzian n-dimensional manifold, in any metric theory of gravity (thus with geodesic motion for test particles), one can determine all the $\frac{n^2(n^2-1)}{12}$ independent components of the Riemann tensor, by measuring the relative accelerations of a sufficiently large number of test particles and by using the equation of geodesic deviation (2.5.1).

However, which is the *minimum number of test particles necessary to determine the spacetime curvature fully*? As we observed, in a four-dimensional spacetime the Riemann tensor has twenty independent components. However, when the metric of the spacetime is subject to the Einstein equation in vacuum, $R_{\alpha\beta} = R^{\sigma}{}_{\alpha\sigma\beta} = 0$, the number of independent components of the Riemann tensor is reduced to ten, and they form the **Weyl tensor**[11] which is in general defined by

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + g_{\alpha[\delta}R_{\gamma]\beta} + g_{\beta[\gamma}R_{\delta]\alpha} + \frac{1}{3}Rg_{\alpha[\gamma}g_{\delta]\beta} \quad (2.5.18)$$

where $R = R_{\alpha\beta} g^{\alpha\beta}$.

Synge in his classic book on the general theory of relativity[13] describes a method of measuring the independent components of the Riemann tensor. Synge calls his device a five-point curvature detector. The five-point curvature detector consists of a light source and four mirrors. By performing measurements of the distance between the source and the mirrors and between the mirrors, one can determine the curvature of the spacetime. However, in order to measure all the independent components of the Riemann tensor with Synge's method, the experiment must be repeated several times with different orientations of the detector; equivalently—and when the spacetime is not stationary—it is necessary to use several curvature detectors at the same time.

Instead, one can measure the relative accelerations of test particles moving on infinitesimally close geodesics and use equation (2.5.1). However, in order to minimize the number of test particles necessary to determine all the independent components of the Riemann tensor at one event, it turns out that one has to use nearby test particles, moving with arbitrarily different four-velocities.

It is then possible to derive a generalized geodesic deviation equation,[53] valid for any two geodesics, with arbitrary tangent vectors, not necessarily parallel, and describing the relative acceleration of two test particles moving with any four-velocity on neighboring geodesics. This generalized equation can be derived by dropping the previous condition (4): $\left| \frac{\|\delta u^{\alpha}\|}{\|u^{\alpha}\|} \right| \ll 1$, and by retaining the conditions (1), (2), (3), and (5) only,[53] and it is valid in any neighborhood in which the change of the curvature is small (condition 3). Of course, when the two geodesics are locally parallel one recovers the classical geodesic deviation equation. Physically, one would measure the relative acceleration of two test particles moving with arbitrary four-velocities (their difference $(u_2^{\alpha} - u_1^{\alpha})$ need not necessarily be small) in an arbitrary gravitational field (in an arbitrary Riemannian manifold), in a region where the relative change of the gravitational field is small. The spacetime need not necessarily satisfy the Einstein field equation so long as the test particles follow geodesic motion (metric theories). It turns out[54] that the minimum number of test particles can be drastically reduced by using the generalized geodesic deviation equation instead of the standard geodesic deviation equation (2.5.1). This number is reduced either (1) under the hypothesis of an arbitrary four-dimensional Lorentzian manifold or (2) when we have an empty region of the spacetime satisfying the Einstein equations, $R_{\alpha\beta} = 0$ (the measurement of the Riemann tensor reduces then to the measurement of the Weyl tensor $C^{\alpha}{}_{\beta\mu\delta}$).

It turns out[54] that to fully determine the curvature of the spacetime in vacuum, in general relativity, it is *sufficient* to use four test particles, and in general spacetimes (twenty independent components of the Riemann tensor) it is sufficient to use six test particles. It is easy to show that in a vacuum, to fully determine the curvature, it is also *necessary* to use at least four test particles. With four test particles we have three independent geodesic deviation equations leading to twelve relations between the ten independent components of the Riemann tensor and the relative accelerations of the test particles. In general spacetimes it is necessary to use at least six test particles. Of course, it is possible to determine the curvature of the spacetime using test particles having approximately equal four-velocities and using the standard geodesic deviation equation. However, it turns out then that the minimum number of test particles which is required in general relativity increases to thirteen in general spacetimes and to six in vacuum.

2.6 SOME EXACT SOLUTIONS OF THE FIELD EQUATION

A Rigorous Derivation of a Spherically Symmetric Metric

Given a **three-dimensional Riemannian manifold** M^3, one may define M^3 to be **spherically symmetric**[20,38,41] about one point O (for the definition based on the isometry group see § 4.2), if, in some coordinate system, x^i, every rotation about O, of the type $x'^i = O^{i'}_k x^k$ where $\delta_{ij} = O^{m'}_i O^{n'}_j \delta_{m'n'}$, and $\det O^{i'}_k = +1$, is an **isometry** for the metric $\boldsymbol{g}$ of M^3. In other words, the metric $\boldsymbol{g}$ in M^3 is defined spherically symmetric if it is **formally invariant** *for rotations*; that is, the new components of $\boldsymbol{g}$ are the same functions of the new coordinates x'^α as the old components of $\boldsymbol{g}$ were of the old coordinates x^α for rotations

$$g_{\alpha\beta}(y^\alpha \equiv x^\alpha) = g'_{\alpha\beta}(y^\alpha \equiv x'^\alpha). \tag{2.6.1}$$

A **Lorentzian manifold** M^4 may then be defined **spherically symmetric** about one point O, if, in some coordinate system, the metric $\boldsymbol{g}$ is formally invariant for three-dimensional spatial rotations about $O:\ x'^i = O^{i'}_k x^k$ (as defined above), that is, three-dimensional spatial rotations are isometries for $\boldsymbol{g}:\ g_{\alpha\beta}(x^0, x^i) = g'_{\alpha\beta}(x^0, x^i)$. (In general, on a Lorentzian manifold a geometrical quantity $G(x^0, x^i)$ may be defined to be spherically symmetric if G is formally invariant for three-dimensional spatial rotations: $G(x^0, x^i) = G'(x^0, x^i)$.)

Formal invariance of the metric $\boldsymbol{g}$ under the infinitesimal coordinate transformation $x'^\alpha = x^\alpha + \varepsilon\xi^\alpha$, where $|\varepsilon| \ll 1$, is equivalent to the requirement that the **Lie derivative**[55,56] (see § 4.2 and mathematical appendix) of the metric tensor $\boldsymbol{g}$, with respect to $\boldsymbol{\xi}$, be zero:

$$\mathcal{L}_{\boldsymbol{\xi}} g_{\alpha\beta} \equiv g_{\alpha\beta,\sigma}\xi^\sigma + g_{\sigma\beta}\xi^\sigma{}_{,\alpha} + g_{\alpha\sigma}\xi^\sigma{}_{,\beta} = 0. \tag{2.6.2}$$

This requirement follows from the definition (2.6.1) of formal invariance under the infinitesimal coordinate transformation $x'^{\alpha} = x^{\alpha} + \varepsilon\xi^{\alpha}$, thus

$$\begin{aligned} 0 &= g_{\alpha\beta}(x'^{\gamma}) - g'_{\alpha\beta}(x'^{\gamma}) \\ &= g_{\alpha\beta}(x^{\gamma}) + g_{\alpha\beta,\sigma}\varepsilon\xi^{\sigma} - \partial^{\sigma}_{\alpha'}{}^{\rho}_{\beta'} g_{\sigma\rho}(x^{\gamma}) \\ &= g_{\alpha\beta,\sigma}\varepsilon\xi^{\sigma} + \varepsilon\xi^{\sigma}{}_{,\alpha}g_{\sigma\beta} + \varepsilon\xi^{\rho}{}_{,\beta}g_{\alpha\rho}. \end{aligned} \tag{2.6.3}$$

As follows from the definition (2.3.22) of the Christoffel symbols that enter into covariant derivatives, this condition on the metric is equivalent (see § 4.2) to the **Killing equation**:

$$\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = 0. \tag{2.6.4}$$

Therefore, the **Killing vector $\boldsymbol{\xi}$** describes the symmetries of the metric tensor $\boldsymbol{g}$ by defining the isometric mappings of the metric onto itself, that is, the isometries.[57] We have just defined a metric $\boldsymbol{g}$ to be spherically symmetric if it is formally invariant under three-dimensional spatial rotations, therefore a metric is spherically symmetric if it satisfies the Killing equation for every Killing vector $\boldsymbol{\xi}_{ss}$ that represents a three-dimensional spatial rotation. The Killing vector representing **spherical symmetry**, in "generalized Cartesian coordinates," is

$$\xi^{0}_{ss} = 0, \qquad \xi^{i}_{ss} = c^{ij}x^{j} \tag{2.6.5}$$

where $c^{ik} = -c^{ki}$ are three constants. In other words, spherical symmetry about the point O is equivalent to axial symmetry around each of the three-axes Ox^{a}, represented by the Killing vector:

$$\xi^{0} = \xi^{a} = 0; \qquad \xi^{b} = x^{c}; \qquad \xi^{c} = -x^{b} \tag{2.6.6}$$

where (a, b, c) is some permutation of $(1, 2, 3)$. In particular, using generalized Cartesian coordinates, we have

$$\begin{aligned} \xi'^{\alpha}_{1} &= (0, 0, z, -y) \\ \xi'^{\alpha}_{2} &= (0, -z, 0, x) \\ \xi'^{\alpha}_{3} &= (0, y, -x, 0) \end{aligned} \tag{2.6.7}$$

or using "generalized polar coordinates," defined by the usual transformation $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ and $z = r\cos\theta$, we have

$$\begin{aligned} \xi^{\alpha}_{1} &= (0, 0, \sin\phi, \cot\theta\cos\phi) \\ \xi^{\alpha}_{2} &= (0, 0, -\cos\phi, \cot\theta\sin\phi) \\ \xi^{\alpha}_{3} &= (0, 0, 0, -1). \end{aligned} \tag{2.6.8}$$

From the Killing equation (2.6.2), using the Killing vector $\boldsymbol{\xi}_{3}$, we get[58]

$$g_{\alpha\beta,\phi} = 0, \tag{2.6.9}$$

and using the Killing vectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ in equation (2.6.2), we then get

$$g_{11,\theta} = 0, \qquad g_{00,\theta} = 0, \qquad g_{10,\theta} = 0, \tag{2.6.10}$$

and by applying equation (2.6.2) to $\boldsymbol{\xi}_1$:

$$\begin{aligned}
&g_{22,\theta} \sin\phi = 2g_{23} \frac{\cos\phi}{\sin^2\theta} \\
&(g_{33,\theta} - 2g_{33}\cot\theta)\sin\phi = -2g_{23}\cos\phi \\
&g_{12,\theta} \sin\phi = g_{13} \frac{\cos\phi}{\sin^2\theta} \\
&(g_{13,\theta} - g_{13}\cot\theta)\sin\phi = -g_{12}\cos\phi \\
&(g_{23,\theta} - g_{23}\cot\theta)\sin\phi = \big(-g_{22} + g_{33}\frac{1}{\sin^2\theta}\big)\cos\phi \\
&g_{20,\theta} \sin\phi = g_{30} \frac{\cos\phi}{\sin^2\theta} \\
&(g_{30,\theta} - g_{30}\cot\theta)\sin\phi = -g_{20}\cos\phi,
\end{aligned} \tag{2.6.11}$$

plus the seven similar equations for $\boldsymbol{\xi}_2$ obtained by replacing both $\sin\phi$ with $-\cos\phi$ and $\cos\phi$ with $\sin\phi$ in the equations (2.6.11). From equations (2.6.9), (2.6.10), and (2.6.11) and the seven similes we get

$$\begin{aligned}
&g_{00} = g_{00}(r,t), \qquad g_{11} = g_{11}(r,t), \qquad g_{22} = g_{22}(r,t), \\
&g_{33} = g_{22}(r,t)\sin^2\theta \qquad \text{and} \qquad g_{01} = g_{01}(r,t),
\end{aligned} \tag{2.6.12}$$

that is, g_{00}, g_{11}, g_{22}, $g_{33}/\sin^2\theta$, and g_{01} are functions of r and t only; all the other components of $\boldsymbol{g}$ are identically equal to zero.

The general form of a **four-dimensional spherically symmetric metric** is then

$$\begin{aligned}
ds^2 = {} & A(r,t)dt^2 + B(r,t)dr^2 + C(r,t)drdt \\
& + D(r,t)(d\theta^2 + \sin^2\theta d\phi^2).
\end{aligned} \tag{2.6.13}$$

This we simplify by performing the coordinate transformation

$$t' = t \qquad \text{and} \qquad r'^2 = D(r,t) \tag{2.6.14}$$

where we assume $D(r,t) \neq$ constant. We then get (dropping the prime in t' and r')

$$ds^2 = E(r,t)dt^2 + F(r,t)dr^2 + G(r,t)drdt + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{2.6.15}$$

With the further coordinate transformation

$$t' = H(r,t) \qquad \text{and} \qquad r' = r \tag{2.6.16}$$

where we assume $H_{,t} \neq 0$, we have

$$g_{01} = \partial_0^{0'} \partial_1^{1'} g'_{01} + \partial_0^{0'} \partial_1^{0'} g'_{00} = H_{,t} g'_{01} + H_{,t} H_{,r} g'_{00} \tag{2.6.17}$$

and

$$g_{00} = \partial_0^{0'} \partial_0^{0'} g'_{00} = (H_{,t})^2 g'_{00}; \tag{2.6.18}$$

to simplify the metric in its new form, we impose the condition

$$g'_{01} = \frac{H_{,t} \cdot G}{2(H_{,t})^2} - \frac{H_{,r} \cdot E}{(H_{,t})^2} \equiv 0. \tag{2.6.19}$$

This condition can always be satisfied, for any function G and $E \neq 0$, by finding a solution to the differential equation:

$$\frac{1}{2} H_{,t} \cdot G - H_{,r} \cdot E = 0. \tag{2.6.20}$$

Therefore, we finally have (dropping the prime in t' and r')

$$ds^2 = -e^{m(r,t)} dt^2 + e^{n(r,t)} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \tag{2.6.21}$$

as **metric of a spherically symmetric spacetime** in a *particular* coordinate system. The signs were determined according to the Lorentzian character of the Riemannian manifold, in agreement with the equivalence principle: $\overset{(i)}{g}_{\alpha\beta} \to \eta_{\alpha\beta}$.

Let us now find the expression of a spherically symmetric metric satisfying the vacuum Einstein field equation (2.3.14), with $T^{\alpha\beta} = 0$:

$$G^{\alpha\beta} = 0 \qquad \text{or, equivalently,} \qquad R^{\alpha\beta} = 0. \tag{2.6.22}$$

From the definition of Ricci tensor, that we symbolically write here

$$R^{\sigma}{}_{\alpha\sigma\beta} = \begin{vmatrix} {}_{,\sigma} & {}_{,\beta} \\ \Gamma^{\sigma}_{\alpha\sigma} & \Gamma^{\sigma}_{\alpha\beta} \end{vmatrix} + \begin{vmatrix} \Gamma^{\sigma}_{\rho\sigma} & \Gamma^{\sigma}_{\rho\beta} \\ \Gamma^{\rho}_{\alpha\sigma} & \Gamma^{\rho}_{\alpha\beta} \end{vmatrix} \equiv \Gamma^{\sigma}_{\alpha\beta,\sigma} - \Gamma^{\sigma}_{\alpha\sigma,\beta} + \cdots, \tag{2.6.23}$$

and from the definition (2.3.22) of Christoffel symbols, we then get

$$\begin{aligned} R_{00} = {} & -e^{m-n}\Big(\frac{1}{2} m_{,rr} - \frac{1}{4} m_{,r} n_{,r} + \frac{1}{4} m_{,r}{}^2 + \frac{m_{,r}}{r}\Big) \\ & + \frac{1}{2} n_{,tt} + \frac{1}{4} n_{,t}{}^2 - \frac{1}{4} m_{,t} n_{,t} = 0 \end{aligned} \tag{2.6.24}$$

$$\begin{aligned} R_{11} = {} & \frac{1}{2} m_{,rr} - \frac{1}{4} m_{,r} n_{,r} + \frac{1}{4} m_{,r}{}^2 - \frac{n_{,r}}{r} \\ & - e^{n-m}\Big(\frac{1}{2} n_{,tt} + \frac{1}{4} n_{,t}{}^2 - \frac{1}{4} m_{,t} n_{,t}\Big) = 0 \end{aligned} \tag{2.6.25}$$

$$R_{22} = -1 + e^{-n} + \frac{1}{2} e^{-n} r (m_{,r} - n_{,r}) = 0 \tag{2.6.26}$$

$$R_{33} = R_{22} \sin^2 \theta = 0 \tag{2.6.27}$$

and

$$R_{01} = -\frac{n_{,t}}{r} = 0 \tag{2.6.28}$$

with all the other nondiagonal components of $R_{\alpha\beta}$ identically zero. From the 00 and 11 components we then have

$$(m + n)_{,r} = 0, \tag{2.6.29}$$

and from the 01 component (2.6.28) $\frac{\partial n}{\partial t} = 0$; therefore,

$$m + n(r) = f(t) \qquad \text{or} \qquad e^m = e^{f(t)} e^{-n(r)}. \tag{2.6.30}$$

The time dependence $f(t)$ can be absorbed in the definition of t with a coordinate transformation of the type $t' = \int e^{\frac{1}{2} f(t)} dt$. Therefore, in the new coordinates (dropping the prime in n' and m'), we have the result

$$\frac{\partial n}{\partial t} = \frac{\partial m}{\partial t} = 0 \qquad \text{and} \qquad e^{m(r)} = e^{-n(r)}. \tag{2.6.31}$$

Therefore, a spherically symmetric spacetime satisfying the vacuum Einstein field equation (2.6.22) is static, that is, there is a coordinate system in which the metric is time independent, $g_{\alpha\beta,0} = 0$, and in which $g_{0i} = 0$.

We recall that a **spacetime** is called **stationary** if it admits a timelike Killing vector field, $\boldsymbol{\xi}_t$. For it, there exists some coordinate system in which $\boldsymbol{\xi}_t$ can be written $\boldsymbol{\xi}_t = (1, 0, 0, 0)$. In this system, from the Killing equation (2.6.2), the metric $\boldsymbol{g}$ is then time independent, $g_{\alpha\beta,0} = 0$. A **spacetime** is called **static** if it is stationary and the timelike Killing vector field $\boldsymbol{\xi}_t$ is orthogonal to a foliation (§ 5.2.2) of spacelike hypersurfaces. Therefore, there exists some coordinate system, called adapted to $\boldsymbol{\xi}_t$, in which the metric $\boldsymbol{g}$ satisfies both $g_{\alpha\beta,0} = 0$ and $g_{0i} = 0$.

From the 22, or the 33, component of the vacuum field equation, plus equation (2.6.29), we then have

$$-1 + e^{-n} - re^{-n} n_{,r} = 0 \tag{2.6.32}$$

and therefore

$$(re^{-n})_{,r} = 1 \tag{2.6.33}$$

with the solution

$$e^{-n} = 1 + \frac{C}{r}. \tag{2.6.34}$$

By writing the constant $C \equiv -2M$, we finally have

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 \left(d\theta^2 + \sin^2\theta \, d\phi^2\right). \tag{2.6.35}$$

This is the **Schwarzschild** (1916) **solution**.[59] In conclusion, any spherically symmetric solution of the vacuum Einstein field equation must be static and in some coordinate system must have the Schwarzschild form (**Birkhoff Theorem**).[60] By assuming that the spacetime geometry generated by a spherically symmetric object is itself spherically symmetric, and by requiring that we recover the classical gravity theory, for large r, in the weak field region, we find that M is the mass of the central body (see § 3.7).

However, inside a hollow, static, spherically symmetric distribution of matter, for $r \to 0$, to avoid $g_{00} \to \infty$ and $g_{11} \to 0$, we get $C \equiv 0$. Therefore, the solution internal to a nonrotating, empty, spherically symmetric shell is the Minkowski metric $\eta_{\alpha\beta}$ (for the weak field, slow motion solution inside a rotating shell, see § 6.1 and expression 6.1.37).

Other One-Body Solutions

A solution of the field equation with no matter but with an electromagnetic field, with three parameters M, Q, and J that in the weak field limit are identified with the mass M, the charge Q, and the angular momentum J of a central body, is the **Kerr-Newman solution**,[61,62] that in the t, r, θ, ϕ Boyer-Lindquist coordinates[63] can be written

$$\begin{aligned} ds^2 = &-\left(1 - \frac{(2Mr - Q^2)}{\rho^2}\right)dt^2 \\ &- \left(\frac{(4Mr - 2Q^2)a \sin^2\theta}{\rho^2}\right) dt d\phi + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 \\ &+ \left(r^2 + a^2 + \frac{(2Mr - Q^2)a^2 \sin^2\theta}{\rho^2}\right)\sin^2\theta d\phi^2 \end{aligned} \tag{2.6.36}$$

where

$$\begin{aligned} \Delta &\equiv r^2 - 2Mr + a^2 + Q^2 \\ \rho^2 &\equiv r^2 + a^2\cos^2\theta \end{aligned} \tag{2.6.36$'$}$$

and $a \equiv \frac{J}{M}$ = angular momentum per unit mass.

In the case $Q = J = 0$ and $M \neq 0$ we have the Schwarzschild metric (2.6.35); when $J = 0$, $M \neq 0$ and $Q \neq 0$, we have the **Reissner-Nordstrøm metric**:[64,65]

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1}dr^2 \\ &+ r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right). \end{aligned} \tag{2.6.37}$$

This solution describes a spherically symmetric spacetime satisfying the Einstein field equation in a region with no matter, but with a radial electric field to be included in the energy-momentum tensor $T_{\alpha\beta}$ (see § 2.3),

$$\boldsymbol{E} = \frac{Q}{r^2}\boldsymbol{e}_r \qquad \boldsymbol{B} = \boldsymbol{0} \tag{2.6.38}$$

where $\boldsymbol{e}_r$ is the radial unit vector of a static orthonormal tetrad. In the weak field region, M and Q are identified with the mass and the charge of the central object.

Finally, when $Q = 0$ and $M \neq 0$, $J \neq 0$ we have the Kerr solution.[61] In the **weak field** and slow motion limit,[66–69] $M/r \ll 1$, $(J/M)/r \ll 1$, in Boyer-Lindquist coordinates, the **Kerr metric** (2.6.36) can be written

$$ds^2 \cong -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 + \frac{2M}{r}\right)dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right) - \frac{4J}{r}\sin^2\theta d\phi dt. \tag{2.6.39}$$

This is the weak field metric generated by a central body with mass M and angular momentum J; we shall return to this important solution in chapter 6.

2.7 CONSERVATION LAWS

In classical electrodynamics[44] one defines the total charge on a three-dimensional spacelike hypersurface Σ, corresponding to t = constant: $Q = \int_\Sigma j^0 d^3\Sigma_0$. From the Maxwell equations with source $F^{\alpha\beta}{}_{,\beta} = 4\pi j^\alpha$ and from the antisymmetry of the electromagnetic tensor $F^{\alpha\beta}$, one has the differential conservation law of charge $j^\alpha{}_{,\alpha} = 0$. Therefore, by using the four-dimensional divergence theorem (2.3.9), we verify that Q is conserved:

$$0 = \int_\Omega j^\alpha{}_{,\alpha}\, d^4\Omega = \int_{\partial\Omega} j^\alpha d^3\Sigma_\alpha \tag{2.7.1}$$

where Ω is a spacetime region and $\partial\Omega$ its three-dimensional boundary, and where $d^4\Omega$ and $d^3\Sigma_\alpha$ are respectively the four-dimensional and the three-dimensional integration elements defined by expressions (2.8.21) and (2.8.20) below. By choosing $\partial\Omega$ composed of two spacelike hypersurfaces Σ and Σ', corresponding to the times t = constant and t' = constant$'$, plus an embracing hypersurface Λ, away from the source, on which j^α vanishes (see fig. 2.2), we then have

$$Q = \int_\Sigma j^0 d^3\Sigma_0 = \int_{\Sigma'} j'^0 d^3\Sigma'_0 = Q', \tag{2.7.2}$$

that is, the total charge Q = constant, or $\frac{dQ}{dt} = 0$.

Similarly, in special relativity, one defines the total four-momentum of a fluid described by energy momentum tensor $T^{\alpha\beta}$ (see § 2.3), on a spacelike

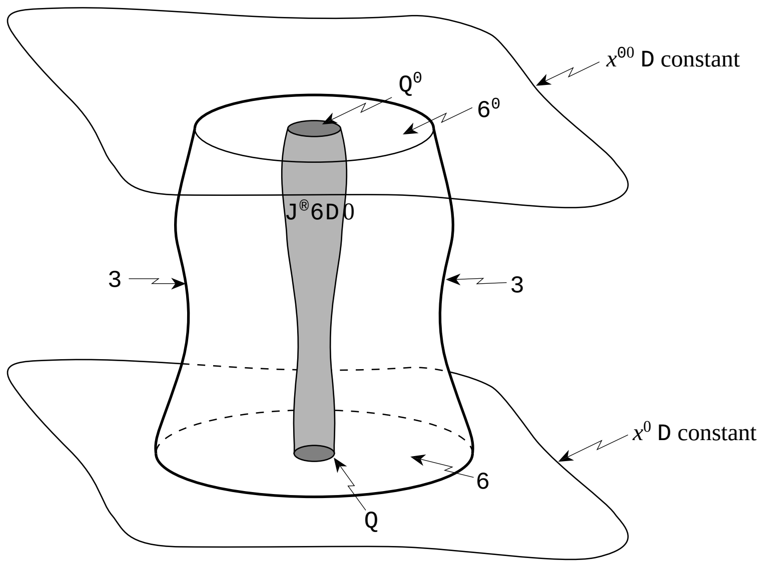


FIGURE 2.2. The hypersurface of integration $\partial\Omega^{(3)}$, boundary of $\Omega^{(4)}$ (see equation (2.7.2)).

hypersurface Σ, as

$$P^\alpha = \int_\Sigma T^{\alpha\beta} d^3\Sigma_\beta \tag{2.7.3}$$

where $E \equiv P^0 = \int T^{0\beta} d^3\Sigma_\beta$ is the energy, and the angular momentum of the fluid is defined (see also § 6.10) on a spacelike hypersurface Σ:

$$J^{\alpha\beta} = \int_\Sigma (x^\alpha T^{\beta\mu} - x^\beta T^{\alpha\mu}) d^3\Sigma_\mu. \tag{2.7.4}$$

From the special relativistic, differential conservation laws $T^{\alpha\beta}{}_{,\beta} = 0$, it then follows that these quantities are conserved:

$$0 = \int_\Omega T^{\alpha\beta}{}_{,\beta} d^4\Omega = \int_{\partial\Omega} T^{\alpha\beta} d^3\Sigma_\beta \tag{2.7.5}$$

and

$$P^\alpha = \int_\Sigma T^{\alpha 0} d^3\Sigma_0 = \int_{\Sigma'} T'^{\alpha 0} d^3\Sigma'_0 = P'^\alpha \tag{2.7.6}$$

(zero total outflow of energy and momentum), or

$$\frac{dP^\alpha}{dt} = 0, \tag{2.7.7}$$

and similarly, for the angular momentum:

$$\frac{dJ^{\alpha\beta}}{dt} = 0 \tag{2.7.8}$$

where, in formula (2.7.6), we have chosen the hypersurface $\partial\Omega$ as shown in figure 2.2, with Λ away from the source where $T^{\alpha\beta}$ vanishes, and Σ and Σ' corresponding to t = constant and t' = constant$'$.

In this section we generalize these Minkowski-space definitions to geometrodynamics, to get conserved quantities in curved spacetime. In geometrodynamics, the special relativistic dynamical equation generalize to the tensorial equation, $T^{\alpha\beta}{}_{;\beta} = 0$, consequence of the field equation and of the Bianchi identities—that is, of the fundamental principle that the boundary of the boundary of a region is zero (§ 2.8). However, the divergence theorem does not apply to the covariant divergence of a tensor, therefore the geometrodynamical conserved quantities cannot involve only the energy-momentum tensor $T^{\alpha\beta}$.

Before describing the mathematical details of the definition of the conserved quantities in general relativity, let us first discuss what one would expect from the fundamental analogies and differences between electrodynamics and geometrodynamics. First, the gravitational field $g_{\alpha\beta}$ has energy and momentum associated with it. We know that, in general relativity, gravitational waves carry energy[133–135] and momentum (see § 2.10); this has been experimentally indirectly confirmed with the observations of the decrease of the orbital period of the binary pulsar PSR 1913+1916, explained by the emission of gravitational waves, in agreement with the general relativistic formulae (§ 3.5.1). Two gravitons may create matter, an electron and a positron, by the standard Ivanenko process;[70] therefore, for the conservation of energy, gravitons and gravitational waves must carry energy. We also know that the gravitational geon,[71] made of gravitational waves (see § 2.10), carries energy and momentum. Therefore, since gravitational waves are curvature perturbations of the spacetime, the spacetime geometry must have energy and momentum associated with it. In general relativity the geometry $g_{\alpha\beta}$, where the various physical phenomena take place, is generated by the energy and the energy-currents in the universe, through the field equation. Since the gravity field $g_{\alpha\beta}$ has energy and momentum, the gravitational energy contributes itself, in a loop, to the spacetime geometry $g_{\alpha\beta}$. However, in special relativistic electrodynamics the spacetime geometry $\eta_{\alpha\beta}$ where the electromagnetic phenomena take place, is completely unaffected by these phenomena. Indeed, the fundamental difference between electrodynamics and geometrodynamics is the equivalence principle: locally, in a suitable spacetime neighborhood, it is possible to eliminate every *observable* effect of the gravitational field (see § 2.1). This is true for gravity only.

Therefore, what should one expect from this picture, before one defines the conserved quantities in geometrodynamics? First, one should not expect the conserved quantities to involve only the energy and momentum of matter and nongravitational fields, described by the energy-momentum tensor $T^{\alpha\beta}$ (see expressions 2.3.24 and 2.3.23 for the energy-momentum tensor of a fluid and of an electromagnetic field). Indeed, since the gravitational field $g_{\alpha\beta}$ itself carries energy and momentum, it must, somehow, be included in the definition of energy, momentum, and angular momentum. However, because of the equivalence principle, we should not expect any definition of the energy of the gravitational field to have any local validity; in general relativity, gravitational energy and momentum should only have nonlocal (or quasi-local)[74] validity. Indeed, the gravity field can be locally eliminated, in every freely falling frame, in the sense of eliminating the first derivatives of the metric $g_{\alpha\beta}$ and have $\overset{(i)}{g}_{\alpha\beta} \longrightarrow \eta_{\alpha\beta}$ at a pointlike event; and in the sense of locally (in a spacetime neighborhood of the event) eliminating any measurable effect of gravity, this should also apply to the gravitational energy.

Let us now define the general relativistic conserved quantities. In special relativity, one defines quantities that can be shown to be conserved by using the four-dimensional divergence theorem applied to the differential conservation laws $j^{\alpha}{}_{,\alpha} = 0$ and $T^{\alpha\beta}{}_{,\beta} = 0$. On a curved manifold, from the covariant divergence of the charge current density we can still define conserved quantities by using formula (2.3.6):

$$\begin{aligned} &\int j^{\alpha}{}_{;\alpha}\sqrt{-g}\,d^4\Omega = \\ &\int \left(j^{\alpha}\sqrt{-g}\right)_{,\alpha} d^4\Omega = \\ &\int j^{\alpha}\sqrt{-g}\,d^3\Sigma_{\alpha}. \end{aligned} \tag{2.7.9}$$

However, the four-dimensional divergence theorem is valid for standard divergences but not for the vanishing covariant divergence of the tensor $T^{\alpha\beta}$ in geometrodynamics, $T^{\alpha\beta}{}_{;\beta} = 0$; for a tensor field $T^{\alpha\beta}$, expression (2.3.7) holds, and we cannot directly apply the divergence theorem.

Therefore, we should define quantities $t^{\alpha\beta}$, representing the energy and momentum of the gravitational field, such that the sum of these quantities and of the energy-momentum tensor $T^{\alpha\beta}$

$$T^{\alpha\beta} + t^{\alpha\beta} \equiv T^{\alpha\beta}_{\text{eff}} \tag{2.7.10}$$

will satisfy an equation of the type $T^{\alpha\beta}_{\text{eff},\beta} = 0$. We could then apply the four-dimensional divergence theorem. Of course, on the basis of what we have just observed, we should not expect these quantities $t^{\alpha\beta}$ to form a tensor, since locally the gravity field and its energy should be eliminable.

There are several possible choices for $t^{\alpha\beta}$. We follow here the useful convention of Landau-Lifshitz.[17] By our making zero the first derivatives of the metric tensor at a pointlike event, the gravity field can be "eliminated" in a local inertial frame. Therefore, the quantities $t^{\alpha\beta}$ representing energy and momentum of the gravity field should go to zero in every local inertial frame, and should then be a function of the first derivatives of $g_{\alpha\beta}$. Indeed, at any event, in a local inertial frame, one can reduce the differential conservation laws to $T^{\alpha\beta}{}_{,\beta} = 0$. Therefore, in order to define the pseudotensor, $t^{\alpha\beta}$, for the gravity field, we first write the field equation at an event, in a local inertial frame, where the first derivatives of the metric are zero. At this event the field equation will involve only the metric and its second derivatives. After some rearrangements, the field equation can then be written

$$\overset{(i)}{\Lambda}{}^{\alpha\beta\mu\nu}{}_{,\nu\mu} = (-\overset{(i)}{g})\overset{(i)}{T}{}^{\alpha\beta} \tag{2.7.11}$$

where

$$\overset{(i)}{g}_{\alpha\beta,\mu} = 0 \tag{2.7.12}$$

and

$$\Lambda^{\alpha\beta\mu\nu} \equiv \frac{1}{16\pi}(-g)\left(g^{\alpha\beta}g^{\mu\nu} - g^{\alpha\mu}g^{\beta\nu}\right). \tag{2.7.13}$$

We may now rewrite the field equation in a general coordinate system, where the first derivatives of $g_{\alpha\beta}$ are in general different from zero, by defining a quantity $(-g)t^{\alpha\beta}$ that represents the difference between the field equation written in the two systems (2.7.11 and 2.3.14), depending on the first derivatives of the metric:

$$(-g)t^{\alpha\beta} \equiv \Lambda^{\alpha\beta\mu\nu}{}_{,\nu\mu} - (-g)T^{\alpha\beta}. \tag{2.7.14}$$

Then this Einstein field equation (2.7.14) lets itself be translated into the language of the effective energy-momentum pseudotensor of expression (2.7.10); that is,

$$(-g)T^{\alpha\beta}_{\text{eff}} \equiv (-g)\left(T^{\alpha\beta} + t^{\alpha\beta}\right) = \Lambda^{\alpha\beta\mu\nu}{}_{,\nu\mu}. \tag{2.7.15}$$

From expression (2.7.13) we know that $\Lambda^{\alpha\beta\mu\nu}$ is antisymmetric with respect to β and μ. Hence the quantity $\Lambda^{\alpha\beta\mu\nu}{}_{,\nu\mu\beta}$ is zero, and therefore from the field equation we have

$$\left((-g)T^{\alpha\beta}_{\text{eff}}\right)_{,\beta} = \Lambda^{\alpha\beta\mu\nu}{}_{,\nu\mu\beta} = 0. \tag{2.7.16}$$

The explicit expression of the pseudotensor $t^{\alpha\beta}$ can be found after some cumbersome calculations. $t^{\alpha\beta}$ can be symbolically written in the form

$$\left(\begin{array}{c}\textbf{energy-momentum}\\ \textbf{pseudotensor for the}\\ \textbf{gravity field}\end{array}\right) = t^{\alpha\beta} \sim \sum\left(g \cdot g \cdot \Gamma \cdot \Gamma\right), \tag{2.7.17}$$

that is, $t^{\alpha\beta}$ is the sum of various terms, each quadratic in both $g^{\alpha\beta}$ and $\Gamma^{\alpha}_{\mu\nu}$. The precise expression of $t^{\alpha\beta}$ is (see Landau-Lifshitz)[17]

$$\begin{aligned} t^{\alpha\beta} = \frac{1}{16\pi} \Big[& \left(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\beta}g^{\mu\nu}\right)\left(2\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho} - \Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\sigma}\Gamma^{\rho}_{\nu\rho}\right) \\ & + g^{\alpha\mu}g^{\nu\sigma}\left(\Gamma^{\beta}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} + \Gamma^{\beta}_{\nu\sigma}\Gamma^{\rho}_{\mu\rho} - \Gamma^{\beta}_{\sigma\rho}\Gamma^{\rho}_{\mu\nu} - \Gamma^{\beta}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho}\right) \\ & + g^{\beta\mu}g^{\nu\sigma}\left(\Gamma^{\alpha}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} + \Gamma^{\alpha}_{\nu\sigma}\Gamma^{\rho}_{\mu\rho} - \Gamma^{\alpha}_{\sigma\rho}\Gamma^{\rho}_{\mu\nu} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\rho}_{\sigma\rho}\right) \\ & + g^{\mu\nu}g^{\sigma\rho}\left(\Gamma^{\alpha}_{\mu\sigma}\Gamma^{\beta}_{\nu\rho} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\sigma\rho}\right)\Big]. \end{aligned} \tag{2.7.18}$$

Using the effective **energy-momentum pseudotensor for** matter, fields and **gravity field**, in analogy with special relativity and electromagnetism, we finally define the conserved quantities on an asymptotically flat spacelike hypersurface Σ (see below):

$$P^{\alpha} \equiv \int_{\Sigma} \left(T^{\alpha\beta} + t^{\alpha\beta}\right)(-g)d^3\Sigma_{\beta}: \quad \textbf{four-momentum} \tag{2.7.19}$$

$$E \equiv P^0: \quad \textbf{energy} \tag{2.7.20}$$

$$J^{\alpha\beta} \equiv \int_{\Sigma} \left(x^{\alpha}T^{\beta\mu}_{\text{eff}} - x^{\beta}T^{\alpha\mu}_{\text{eff}}\right)(-g)d^3\Sigma_{\mu}: \quad \textbf{angular momentum}. \tag{2.7.21}$$

From equations (2.7.16), as in special relativity, we then have that E, P^{α}, and $J^{\alpha\beta}$ are conserved.

Of course $t^{\alpha\beta}$ (and therefore $T^{\alpha\beta}_{\text{eff}}$) is not a tensor; however, it transforms as a tensor for linear coordinate transformations, as is clear from its expression (2.7.18). Even if the spacetime curvature is different from zero, the pseudotensor for the gravity field $t^{\alpha\beta}$ can be set equal to zero at an event. Vice versa, even in a flat spacetime, $t^{\alpha\beta}$ can be made different from zero with some simple nonlinear coordinate transformation, not even a physical change of frame of reference, but just a mathematical transformation of the spatial coordinates, for example, a simple spatial transformation from Cartesian to polar coordinates. However, the fact that $t^{\alpha\beta}$ can be made different from zero in a flat spacetime, and that it can be made zero, at an event, in a spacetime with curvature, is what we expected, even before defining $t^{\alpha\beta}$, on the basis of the equivalence principle, that is, on the basis that, locally, we can eliminate the observable effects of the gravity field, and therefore, locally, we should not be able to define an energy associated with the gravity field.

However, the situation is different nonlocally; for example, one can define the effective energy carried by a gravitational wave by integrating over a region large compared to a wavelength (see next section). In fact, the energy, momentum, and angular momentum, $E \equiv P^0$, P^{α}, and $J^{\alpha\beta}$, as defined by expressions (2.7.20), (2.7.19), and (2.7.21), have the fundamental property that in an asymptotically flat spacetime, if evaluated on a large region extending far from the source, have a value independent from the coordinate system chosen near the source, and behave as special relativistic four-tensors for any transformation that far from the source is a Lorentz transformation. This happy feature appears when the integrals are transformed to two-surface integrals evaluated far from the source. We have, in fact,

$$P^{\alpha} = \int_{\Sigma} T^{\alpha\beta}_{\text{eff}}(-g) d^3\Sigma_{\beta} = \int_{\Sigma} \Lambda^{\alpha\beta\mu\nu}{}_{,\nu\mu} d^3\Sigma_{\beta}. \qquad (2.7.22)$$

By choosing a hypersurface $x^0 = $ constant, with volume element $d^3\Sigma_0$, and by using the divergence theorem, we find

$$P^{\alpha} = \int_{\Sigma} \Lambda^{\alpha 0 i \nu}{}_{,\nu i} d^3\Sigma_0 = \int_{\partial\Sigma \equiv S} \Lambda^{\alpha 0 i \nu}{}_{,\nu} d^2 S_i, \qquad (2.7.23)$$

and similarly for $J^{\alpha\beta}$, where $d^2 S_i \equiv (^*dS)_{0i}$ is defined by expression (2.8.19) below. Therefore, P^{α} is invariant for any coordinate transformation near the source, that far from the source, and thus on $\partial\Sigma$, leaves the metric unchanged. Then, since $t^{\alpha\beta}$ behaves as a tensor for linear coordinate transformations (see expression 2.7.18) and P^{α} and $J^{\alpha\beta}$ have a value independent from the coordinates chosen near the source, P^{α} and $J^{\alpha\beta}$ behave as special relativistic four-tensors for any transformation that far from the source is a Lorentz transformation.

In an asymptotically flat manifold, in the weak field region far from the source, where $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, and $|h_{\alpha\beta}| \ll 1$, from expression (2.7.23), we have the **ADM** formula for the **total energy**:[72]

$$E \equiv P^0 = \frac{1}{16\pi} \int_S \left(g_{ij,j} - g_{jj,i} \right) d^2 S_i. \qquad (2.7.24)$$

In a spacetime that in the weak field region matches the Schwarzschild (or the Kerr) solution, one then gets, from the post-Newtonian expression (3.4.17) of chapter 3, in asymptotically Minkowskian coordinates, $E = M$, where M is the observed (Keplerian) mass of the central object.

If the interior of the hypersurface of integration Σ contains singularities with apparent horizons or wormholes, one can still prove[73] the gauge invariance and the conservation of P^{α}, without the use of the divergence theorem.

Penrose[74] has given an interesting *quasi-local definition* of energy-momentum and angular momentum, using twistors (a type of spinor field), valid, unlike the ADM formula,[72] even if the integration is done over a finite spacelike two-surface on a manifold *not necessarily asymptotically flat*.

One may now ask an important question. In general, when dealing with arbitrarily strong gravitational fields at the source and with arbitrary matter distributions as sources, is the total energy E of an isolated system positive in general relativity? The solution of this problem is given by the so-called Positive Energy Theorem.

The **Positive Energy Theorem** of Schoen and Yau[75–79] (see also Choquet-Bruhat, Deser, Teitelboim, Witten, York, etc.)[73,80–83] states that given a spacelike, asymptotically Euclidean, hypersurface Σ, and assuming the so-called dominance of energy condition, that is, $\varepsilon \geq (j^i j_i)^{\frac{1}{2}}$, where ε is the energy density on Σ and j^i is the momentum-density on Σ (the dominance of energy condition implies also the weak energy condition $\varepsilon \geq 0$; see § 2.9), and the validity of the Einstein field equation (2.3.14), then: (1) $|E| \equiv |P^0| > |\boldsymbol{P}|$, that is, *the ADM four-momentum is timelike*, and (2) *future-pointing*, $E > 0$, unless $P^\alpha = 0$ (occurring only for Minkowskian manifolds).

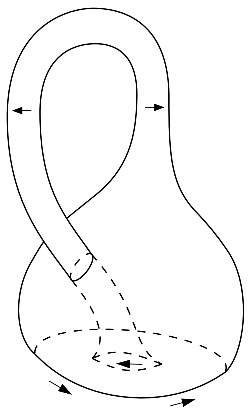
2.8 [THE BOUNDARY OF THE BOUNDARY PRINCIPLE AND GEOMETRODYNAMICS]

Einstein's "general relativity," or geometric theory of gravitation, or "geometrodynamics," has two central ideas: (1) Spacetime geometry "tells" mass-energy how to move; and (2) mass-energy "tells" spacetime geometry how to curve.

We have just seen that the way spacetime tells mass-energy how to move is automatically obtained from the Einstein field equation (2.3.14) by using the identity of Riemannian geometry, known as the Bianchi identity, which tells us that the covariant divergence of the Einstein tensor is zero.

According to an idea of extreme simplicity of the laws at the foundations of physics, what one of us has called "the principle of austerity" or "law without law at the basis of physics,"[84] in geometrodynamics it is possible to derive[85,11] the dynamical equations for matter and fields from an extremely simple but central identity of algebraic topology:[86,39] *the principle that the* **boundary of the boundary of a manifold is zero**. Before exploring the consequences of this principle in physics, we have to introduce some concepts and define some quantities of topology and differential geometry.[39–43,86,87]

An n-dimensional **manifold**, M, **with boundary** is a topological space, each of whose points has a neighborhood homeomorphic (two topological spaces are homeomorphic if there exists a mapping between them that is bijective and bicontinuous, called a homeomorphism; see mathematical appendix), that is, topologically equivalent, to an open set in half $\Re^n$, that is to the subspace H^n of all the points $(x^1, x^2, \cdots, x^n)$ of $\Re^n$ such that $x^n \geq 0$. The boundary ∂M of this manifold M is the $(n-1)$-dimensional manifold of all points of M whose images under one of these homeomorphisms lie on the submanifold of H^n corresponding to the points $x^n = 0$. An **orientable manifold** is a manifold that can be covered by a family of charts or coordinate systems $(x^1, \cdots, x^n)$, $(\bar{x}^1, \cdots, \bar{x}^n)$, ..., such that in the intersections between the charts, the Jacobian, that is, the determinant $\left| \frac{\partial x^i}{\partial \bar{x}^j} \right| \equiv \det\left(\frac{\partial x^i}{\partial \bar{x}^j} \right)$ of the derivatives of the coordinates, is positive. Examples of nonorientable manifolds are the Möbius strip and the



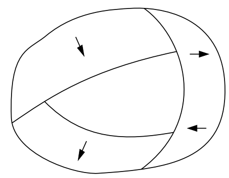


FIGURE 2.3. Two examples of nonorientable manifolds: the Klein bottle or twisted torus and the Möbius strip.

Klein bottle or twisted torus (see fig. 2.3). In the theory of integration[39] on a manifold M which is smooth (that is, differentiable, or which is covered by a family of charts, such that in their intersections the $\frac{\partial x^i}{\partial \bar{x}^j}$ are C^∞ functions) and orientable, one defines a singular n-cube (see fig. 2.4) as a smooth map in the manifold M of an n-cube in the Euclidean $\Re^n$; singular means that the correspondence between a standard n-cube of $\Re^n$ and its image in the manifold M is not necessarily one to one. Then, n-chains c of n-cubes are formally defined as finite sums of n-cubes (multiplied by integers).[39] On these n-chains one defines integration. The boundary ∂c (see figs. 2.4, 2.5, and 2.6) of an n-chain c of n-cubes is the sum of all the properly oriented singular $(n-1)$-cubes which are the boundary of each singular n-cube of the n-chain c. One can then define an operator ∂ that gives the boundary, with a definite orientation, of an n-cube or of an n-chain. It is in general possible to prove[39] that the boundary of the boundary of any n-chain c is zero (see figs. 2.5 and 2.6), that is,

$$\partial(\partial c) = 0 \qquad \text{or formally} \qquad \partial^2 = 0. \tag{2.8.1}$$

Next, let us consider a **differential *n*-form *θ*** that is, a completely antisymmetric covariant n-tensor, in components $\theta_{\alpha_1 \cdots \beta\gamma \cdots \alpha_n} = -\theta_{\alpha_1 \cdots \gamma\beta \cdots \alpha_n}$, against exchange of any pair of nearby indices such as β, γ; n is the degree of the form. Similarly one can consider a completely antisymmetric contravariant n-tensor called ***n*-polyvector**. The operation of **antisymmetrization** of an n-tensor $T_{\alpha_1 \cdots \alpha_n}$, that we shall denote by writing the indices of the tensor within square brackets, is defined as

$$T_{[\alpha_1 \cdots \alpha_n]} = \frac{1}{n!} \sum_{\substack{\text{all} \\ \text{permutations, } p}} \epsilon_p T_{\alpha_1 \cdots \alpha_n} \tag{2.8.2}$$

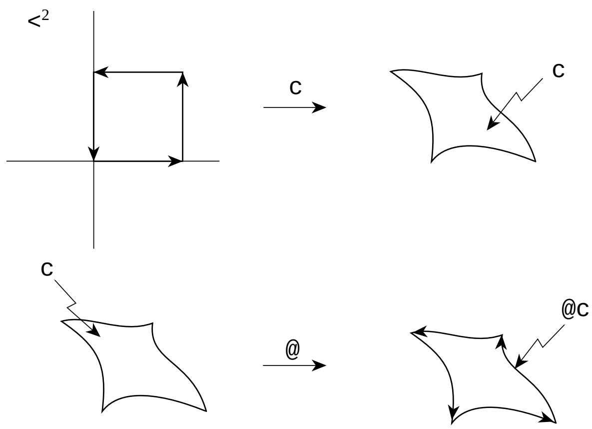


FIGURE 2.4. A standard two-cube c and its $(2-1)$-dimensional boundary ∂c.

where the sum is extended to all the permutations of $\alpha_1 \cdots \alpha_n$, with a plus sign for even permutations, $\epsilon_{p\ \text{even}} \equiv +1$, and minus sign for odd permutations, $\epsilon_{p\ \text{odd}} \equiv -1$. An n-form $\boldsymbol{\theta}$ can then be defined in components as

$$\theta_{\alpha_1 \cdots \alpha_n} = \theta_{[\alpha_1 \cdots \alpha_n]}. \tag{2.8.3}$$

From a p-form $\theta_{\alpha_1 \cdots \alpha_p}$ and from a q-form $\omega_{\alpha_1 \cdots \alpha_q}$, one can construct a $(p+q)$-form, by defining the **wedge product** or **exterior product** $\wedge$ between the two forms, in components

$$(\boldsymbol{\theta} \wedge \boldsymbol{\omega})_{\alpha_1 \cdots \alpha_{p+q}} = \frac{(p+q)!}{p!\,q!}\,\theta_{[\alpha_1 \cdots \alpha_p} \omega_{\alpha_{p+1} \cdots \alpha_{p+q}]} \tag{2.8.4}$$

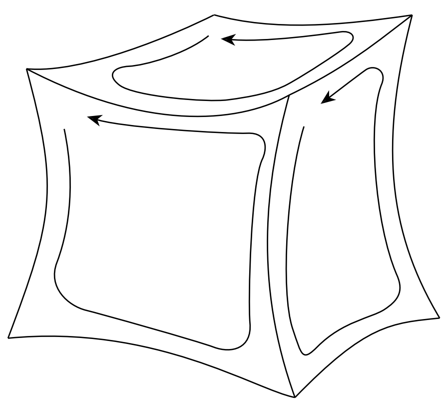


FIGURE 2.5. The oriented one-dimensional boundary of the two-dimensional boundary of a three-cube is zero.

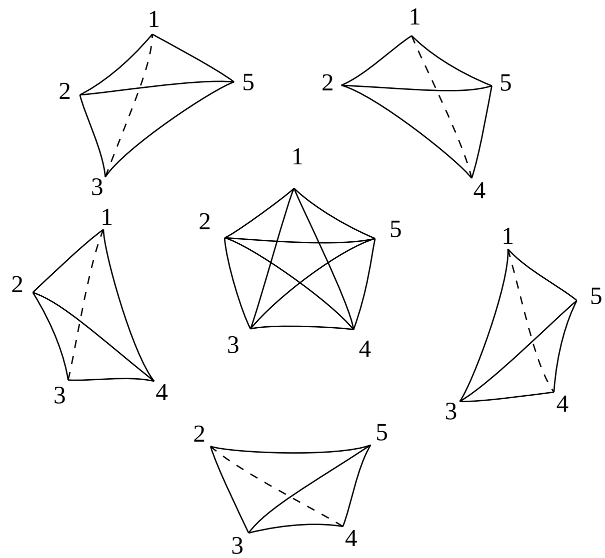


FIGURE 2.6. The two-dimensional boundary of the three-dimensional boundary of a four-dimensional singular four-cube, here a four-simplex, is zero. A two-dimensional projection of the four-simplex is shown in the center. A four-simplex has five vertices, ten edges, ten triangles, and five tetrahedrons. The three-dimensional boundary of the four-simplex is made out of the five tetrahedrons shown in the figure. Each of the ten, two-dimensional, triangles is counted twice with opposite orientations. Therefore, the two-dimensional boundary of the three-dimensional boundary of the four-simplex is zero (adapted from W. Miller 1988).[88]

where $[\alpha_1 \cdots \alpha_{p+q}]$ means antisymmetrization (2.8.2), with respect to the indices within square brackets. The wedge product satisfies the properties

$$\begin{aligned}(\boldsymbol{\theta}_1 \wedge \boldsymbol{\theta}_2) \wedge \boldsymbol{\theta}_3 &= \boldsymbol{\theta}_1 \wedge (\boldsymbol{\theta}_2 \wedge \boldsymbol{\theta}_3)\\ (\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2) \wedge \boldsymbol{\omega} &= \boldsymbol{\theta}_1 \wedge \boldsymbol{\omega} + \boldsymbol{\theta}_2 \wedge \boldsymbol{\omega}\\ \boldsymbol{\theta} \wedge (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) &= \boldsymbol{\theta} \wedge \boldsymbol{\omega}_1 + \boldsymbol{\theta} \wedge \boldsymbol{\omega}_2\\ \boldsymbol{\theta} \wedge \boldsymbol{\omega} &= (-1)^{pq} \boldsymbol{\omega} \wedge \boldsymbol{\theta}.\end{aligned} \tag{2.8.5}$$

Then, from an n-form $\theta_{\alpha_1 \cdots \alpha_n} = \theta_{[\alpha_1 \cdots \alpha_n]}$, one can construct an $(n+1)$-form, by defining the **exterior derivative** $\boldsymbol{d\theta}$ of $\boldsymbol{\theta}$, that is the exterior product of $\frac{\partial}{\partial x^\alpha}$

with $\theta_{\alpha_1\cdots\alpha_n}$, in components

$$\begin{aligned} d\theta_{\alpha_1\cdots\alpha_{n+1}} &= (n+1)\frac{\partial}{\partial x^{[\alpha_1}}\theta_{\alpha_2\cdots\alpha_{n+1}]} \\ &= \frac{1}{n!}\sum_{\substack{\text{all} \\ \text{permutations, } p}} \epsilon_p \frac{\partial}{\partial x^{\alpha_1}}\theta_{\alpha_2\cdots\alpha_{n+1}}. \end{aligned} \tag{2.8.6}$$

The exterior derivative of the exterior product (where $\boldsymbol{\theta}$ is a p-form) satisfies the property

$$\boldsymbol{d}(\boldsymbol{\theta}\wedge\boldsymbol{\omega}) = \boldsymbol{d\theta}\wedge\boldsymbol{\omega} + (-1)^p\boldsymbol{\theta}\wedge\boldsymbol{d\omega}. \tag{2.8.7}$$

We introduce the **Levi-Civita pseudotensor**, $\epsilon_{\alpha\beta\gamma\lambda} \equiv \sqrt{-g}[\alpha\beta\gamma\lambda]$, where $\sqrt{-g}$ is the square root of minus the determinant of the metric (equal to one when $g_{\alpha\beta} = \eta_{\alpha\beta} = \mathrm{diag}(-1, +1, +1, +1) =$ Minkowski tensor), and the symbol $[\alpha\beta\gamma\lambda]$ is equal to $+1$ for even permutations of $(0, 1, 2, 3)$, -1 for odd permutations of $(0, 1, 2, 3)$, and 0 when any indices are repeated. We then have $\epsilon^{\alpha\beta\gamma\lambda} = -\frac{1}{\sqrt{-g}}[\alpha\beta\gamma\lambda]$, and the Levi-Civita pseudotensor satisfies the following relations:

$$\epsilon^{\alpha\beta\gamma\lambda}\epsilon_{\alpha\beta\gamma\lambda} = -4! \tag{2.8.8}$$

$$\epsilon^{\rho\sigma\tau\alpha}\epsilon_{\rho\sigma\tau\beta} = -3!\,\delta^{\alpha}{}_{\beta} \tag{2.8.9}$$

$$\begin{aligned} \epsilon^{\rho\sigma\alpha\beta}\epsilon_{\rho\sigma\gamma\lambda} &= -2!\left(\delta^{\alpha}{}_{\gamma}\delta^{\beta}{}_{\lambda} - \delta^{\alpha}{}_{\lambda}\delta^{\beta}{}_{\gamma}\right) \\ &= -2!\left(2!\delta^{\alpha}{}_{[\gamma}\delta^{\beta}{}_{\lambda]}\right) \\ &\equiv -2!\delta^{\alpha\beta}{}_{\gamma\lambda} \end{aligned} \tag{2.8.10}$$

$$\epsilon^{\alpha\beta\gamma\sigma}\epsilon_{\lambda\mu\nu\sigma} = -3!\delta^{\alpha}{}_{[\lambda}\delta^{\beta}{}_{\mu}\delta^{\gamma}{}_{\nu]} \equiv -\delta^{\alpha\beta\gamma}{}_{\lambda\mu\nu} \tag{2.8.11}$$

and

$$\epsilon^{\alpha\beta\gamma\lambda}\epsilon_{\mu\nu\rho\sigma} = -4!\delta^{\alpha}{}_{[\mu}\delta^{\beta}{}_{\nu}\delta^{\gamma}{}_{\rho}\delta^{\lambda}{}_{\sigma]} \equiv -\delta^{\alpha\beta\gamma\lambda}{}_{\mu\nu\rho\sigma} \tag{2.8.12}$$

where $\delta^{\alpha_1\cdots\alpha_n}{}_{\beta_1\cdots\beta_n}$ is equal to $+1$ if $\alpha_1\cdots\alpha_n$ is an even permutation of $\beta_1\cdots\beta_n$ with no repeated indices ($1 \le n \le 4$), equal to -1 if it an odd permutation, and 0 otherwise. The δ-tensors satisfy

$$\begin{aligned} &\delta^{\alpha\beta\gamma\sigma}{}_{\lambda\mu\nu\sigma} = \delta^{\alpha\beta\gamma}{}_{\lambda\mu\nu}; \quad \delta^{\alpha\beta\sigma}{}_{\mu\nu\sigma} = 2\delta^{\alpha\beta}{}_{\mu\nu}; \\ &\delta^{\alpha\sigma}{}_{\beta\sigma} = 3\delta^{\alpha}{}_{\beta} \quad \text{and} \quad \delta^{\alpha}{}_{\alpha} = 4. \end{aligned} \tag{2.8.13}$$

They can be used to antisymmetrize a tensor

$$T_{[\alpha_1\cdots\alpha_n]} = \frac{1}{n!}T_{\beta_1\cdots\beta_n}\delta^{\beta_1\cdots\beta_n}{}_{\alpha_1\cdots\alpha_n} \tag{2.8.14}$$

(where in a four-manifold: $1 \leq n \leq 4$) and to write the determinant of a tensor $T^{\alpha}{}_{\beta}$

$$\det(T^{\alpha}{}_{\beta}) = \frac{1}{4!}\,\delta^{\alpha\beta\gamma\lambda}{}_{\mu\nu\rho\sigma}T^{\mu}{}_{\alpha}T^{\nu}{}_{\beta}T^{\rho}{}_{\gamma}T^{\sigma}{}_{\lambda} = [\mu\nu\rho\sigma]T^{\mu}{}_{0}T^{\nu}{}_{1}T^{\rho}{}_{2}T^{\sigma}{}_{3}. \tag{2.8.15}$$

Finally, by using the δ-tensors, one can compactly rewrite a **two-dimensional surface element** $dS^{\alpha\beta}$, a **three-dimensional hypersurface element** $d\Sigma^{\alpha\beta\gamma}$, and a **four-dimensional volume element** $d\Omega^{\alpha\beta\gamma\lambda}$, respectively built on two, three, and four infinitesimal displacements $dx^{\alpha}_{(\rho)}$:

$$dS^{\alpha\beta} \equiv \delta^{\alpha\beta}{}_{\mu\nu}dx^{\mu}_{(1)}dx^{\nu}_{(2)} = \begin{vmatrix} dx^{\alpha}_{(1)} & dx^{\alpha}_{(2)} \\ dx^{\beta}_{(1)} & dx^{\beta}_{(2)} \end{vmatrix} \tag{2.8.16}$$

$$d\Sigma^{\alpha\beta\gamma} \equiv \delta^{\alpha\beta\gamma}{}_{\mu\nu\rho}dx^{\mu}_{(1)}dx^{\nu}_{(2)}dx^{\rho}_{(3)} \tag{2.8.17}$$

$$d\Omega^{\alpha\beta\gamma\lambda} \equiv \delta^{\alpha\beta\gamma\lambda}{}_{\mu\nu\rho\sigma}dx^{\mu}_{(0)}dx^{\nu}_{(1)}dx^{\rho}_{(2)}dx^{\sigma}_{(3)}. \tag{2.8.18}$$

The **duals** of these elements, for $\sqrt{-g} = 1$, are defined as

$$({}^{*}dS)_{\alpha\beta} \equiv \frac{1}{2}[\rho\sigma\alpha\beta]dS^{\rho\sigma} \tag{2.8.19}$$

$$d^{3}\Sigma_{\alpha} \equiv \frac{1}{3!}[\alpha\mu\nu\rho]d\Sigma^{\mu\nu\rho} \tag{2.8.20}$$

$$d^{4}\Omega \equiv \frac{1}{4!}[\mu\nu\rho\sigma]d\Omega^{\mu\nu\rho\sigma}. \tag{2.8.21}$$

In particular, for the four infinitesimal coordinate displacements, $dx^{\alpha}_{(\rho)} = \delta^{\alpha}{}_{\rho}dx^{\alpha}$ (no summation over α), with $\rho\epsilon(0, 1, 2, 3)$, we have

$$d^{4}\Omega \equiv d^{4}x = dx^{0}dx^{1}dx^{2}dx^{3}, \tag{2.8.22}$$

and corresponding to a hypersurface $x^{0} =$ constant:

$$d^{3}\Sigma_{0} \equiv d^{3}V = dx^{1}dx^{2}dx^{3}. \tag{2.8.23}$$

On an n-dimensional manifold, we can then define the $(n - p)$-polyvector ${}^{*}\boldsymbol{\theta}$ **dual** to the $\boldsymbol{p}$**-form** $\boldsymbol{\theta}$ in components

$$({}^{*}\boldsymbol{\theta})^{\alpha_1\cdots\alpha_{n-p}} = \frac{1}{p!}\epsilon^{\beta_1\cdots\beta_p\alpha_1\cdots\alpha_{n-p}}\theta_{\beta_1\cdots\beta_p} \tag{2.8.24}$$

with a similar definition for the $(n - p)$-form, ${}^{*}\boldsymbol{v}$ dual of a p-polyvector $\boldsymbol{v}$.

Now, on an n-dimensional manifold M, we have the beautiful and fundamental **Stokes theorem** (for the mathematical details see Spivak 1979, vol. 2)[39]

$$\int_{c} d\boldsymbol{\theta} = \int_{\partial c}\boldsymbol{\theta} \qquad \textbf{Stokes theorem} \tag{2.8.25}$$

where c is an n-chain on the manifold M, ∂c the $(n-1)$-chain oriented boundary of c, $\boldsymbol{\theta}$ a $(n-1)$-form on M, and $\boldsymbol{d\theta}$ the n-form exterior derivative of $\boldsymbol{\theta}$. For an oriented, n-dimensional manifold M with boundary ∂M (with the induced orientation)[39] and for an $(n-1)$-form $\boldsymbol{\theta}$ on M, with compact support (i.e., the smallest closed set containing the region of M where $\boldsymbol{\theta}$ is nonzero is compact), we then have

$$\int_M \boldsymbol{d\theta} = \int_{\partial M} \boldsymbol{\theta} \qquad \textbf{Stokes theorem}. \tag{2.8.26}$$

Furthermore, as a consequence of the boundary of the boundary principle (2.8.1), for every $(n-2)$-form $\boldsymbol{\theta}$ on an n-dimensional, differentiable, oriented manifold M, we have

$$\int_{\partial\partial M} \boldsymbol{\theta} = 0. \tag{2.8.27}$$

Therefore, from the boundary of the boundary principle (2.8.1) and from Stokes theorem:

$$\int_c \boldsymbol{dd\theta} = \int_{\partial c} \boldsymbol{d\theta} = \int_{\partial\partial c} \boldsymbol{\theta} = 0. \tag{2.8.28}$$

By applying this result to an arbitrary neighborhood of an arbitrary point, one has then, automatically,

$$\boldsymbol{dd\theta} = \mathbf{0}, \qquad \text{or formally} \qquad d^2 = 0. \tag{2.8.29}$$

The exterior derivative of the exterior derivative of any form is zero. In other words, the exterior derivative of any exact form is zero, where **exact** is any ***n*-form** that can be written as $\boldsymbol{d\theta}$ and $\boldsymbol{\theta}$ is an $(n-1)$-form. Therefore, any exact form is **closed**, that is, with null exterior derivative (as one can also directly calculate from the definition of $\boldsymbol{d}$). For a vector field $\boldsymbol{W}$ in the three-dimensional Euclidean space $\Re^3$, from Stokes theorem we get two well-known corollaries, the so-called *divergence theorem* (*Ostrogradzky-Green formula* or *Gauss theorem*):

$$\int_V \boldsymbol{\nabla} \cdot \boldsymbol{W} d^3 V = \int_{\partial V = S} \boldsymbol{W} \cdot \boldsymbol{n}\, d^2 S \tag{2.8.30}$$

and the *Riemann-Ampère-Stokes formula*:

$$\int_S (\boldsymbol{\nabla} \times \boldsymbol{W}) \cdot \boldsymbol{n}\, d^2 S = \int_{\partial S = l} \boldsymbol{W} \cdot d^1 \boldsymbol{l} \tag{2.8.31}$$

where d^3V, d^2S and $d^1\boldsymbol{l}$ are the standard Euclidean volume, surface, and line elements, and $\boldsymbol{n}$ is the normal to the surface S.

We are now ready to investigate on some physical consequences[89,90] of the boundary of the boundary principle.

In electrodynamics, one defines (see § 2.3) the electromagnetic field tensor $\boldsymbol{F}$ as the 2-form:

$$\boldsymbol{F} = \boldsymbol{dA} \tag{2.8.32}$$

or in components, $F_{\alpha\beta} = A_{\beta,\alpha} - A_{\alpha,\beta,}$ where $\boldsymbol{A}$ is the four-potential 1-form, with components A_α.

From the boundary of the boundary principle, in the form $d^2 = 0$, we *automatically* get the **sourceless Maxwell equations** for $\boldsymbol{F}$:

$$\boldsymbol{dF} = \boldsymbol{ddA} = \boldsymbol{0} \tag{2.8.33}$$

in components

$$F_{[\alpha\beta,\gamma]} = 0. \tag{2.8.34}$$

The **Maxwell equations with source** are

$$F^{\alpha\beta}{}_{,\beta} = 4\pi j^\alpha \tag{2.8.35}$$

where $j^\alpha = \rho u^\alpha$ is the charge current density four-vector. This equation can be rewritten by defining the dual form, ${}^\star\boldsymbol{F}$, of the form $\boldsymbol{F}$ and the dual form, ${}^\star\boldsymbol{j}$, of the charge current density 1-form $\boldsymbol{j}$ (see expression 2.8.63 for the general definition of ${}^\star(\cdots)$):

$$({}^\star\boldsymbol{F})_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta} \tag{2.8.36}$$

$$({}^\star\boldsymbol{j})_{\beta\mu\nu} \equiv \epsilon_{\alpha\beta\mu\nu}j^\alpha; \tag{2.8.37}$$

therefore

$$\begin{aligned}(\boldsymbol{d}{}^\star\boldsymbol{F})_{\alpha\beta\gamma} &= \frac{3}{2}\epsilon_{\mu\nu[\alpha\beta}F^{\mu\nu}{}_{,\gamma]} \\ &= 4\pi\epsilon_{\sigma\alpha\beta\gamma}j^\sigma\end{aligned} \tag{2.8.38}$$

or

$$\boldsymbol{d}{}^\star\boldsymbol{F} = 4\pi{}^\star\boldsymbol{j}. \tag{2.8.39}$$

From the boundary of the boundary principle, in the form $d^2 = 0$, we then *automatically* get the **dynamical equations for $\boldsymbol{j}$**:

$$4\pi\boldsymbol{d}{}^\star\boldsymbol{j} = \boldsymbol{dd}{}^\star\boldsymbol{F} = \boldsymbol{0} \tag{2.8.40}$$

in components

$$\left(j^\alpha\epsilon_{\alpha[\beta\mu\nu}\right)_{,\gamma]} = 0, \tag{2.8.41}$$

that is, multiplying by $\epsilon^{\beta\mu\nu\gamma}$ and summing over all its indices,

$$j^\alpha{}_{,\alpha} = 0. \tag{2.8.42}$$

Summarizing, in electrodynamics we have

$$\boldsymbol{F} \equiv \boldsymbol{dA} \begin{pmatrix} \text{definition} \\ \text{of } \boldsymbol{F} \end{pmatrix} \overset{d^2=0}{\Longrightarrow} \left\lceil \boldsymbol{dF} = \boldsymbol{0} \begin{pmatrix} \text{sourceless} \\ \text{Maxwell} \\ \text{equations} \end{pmatrix} \right. \tag{2.8.43}$$

and

$$\left\lfloor \boldsymbol{d}^{\star}\boldsymbol{F} = 4\pi^{\star}\boldsymbol{j} \begin{pmatrix} \text{Maxwell} \\ \text{equations} \\ \text{with source} \end{pmatrix} \right. \overset{d^2=0}{\Longrightarrow} \boldsymbol{d}^{\star}\boldsymbol{j} = \boldsymbol{0} \begin{pmatrix} \text{dynamical} \\ \text{equations} \\ \text{for } \boldsymbol{j} \end{pmatrix} \tag{2.8.44}$$

In geometrodynamics, the Riemann curvature tensor satisfies the so-called **first Bianchi identity**:

$$R^{\alpha}{}_{[\beta\gamma\delta]} = 0, \tag{2.8.45}$$

and the **second Bianchi identity** (§ 2.4)

$$R^{\alpha}{}_{\beta[\gamma\delta;\mu]} = 0. \tag{2.8.46}$$

Consequently the Einstein tensor $G_{\alpha\beta}$ satisfies the contracted second Bianchi identities

$$G^{\sigma}{}_{\alpha;\sigma} \equiv \left(R^{\sigma}{}_{\alpha} - \frac{1}{2} R\delta^{\sigma}{}_{\alpha}\right)_{;\sigma} = 0. \tag{2.8.47}$$

As in electrodynamics, these identities can be derived from the boundary of the boundary principle, $\partial^2 = 0$, directly from its consequence that the second exterior derivative of any form is zero, $d^2 = 0$.

Let us first consider,[91,39,43] on an n-dimensional manifold, n linearly independent vector fields $\boldsymbol{X}_{1,\dots,}\boldsymbol{X}_n$, called a *moving frame* (the *Cartan's Repère Mobile*). We can then consider the 1-forms $\boldsymbol{\theta}^{\alpha}$ which define the dual basis (different concept from the dual of a form (2.8.24) or the dual of a polyvector), that is, the forms $\boldsymbol{\theta}^{\alpha}$ such that $\theta^{\alpha}{}_{\sigma} X^{\sigma}{}_{\beta} = \delta^{\alpha}{}_{\beta}$. Furthermore, by using the exterior product (2.8.4), on a Riemannian manifold M with metric $g_{\alpha\beta}$, one can construct the **connection 1-forms** $\boldsymbol{\omega}^{\alpha}{}_{\beta} = \Gamma^{\alpha}_{\beta\gamma}\boldsymbol{\theta}^{\gamma}$, defined by

$$\boldsymbol{dg}_{\alpha\beta} = g_{\alpha\sigma}\boldsymbol{\omega}^{\sigma}{}_{\beta} + g_{\sigma\beta}\boldsymbol{\omega}^{\sigma}{}_{\alpha} \tag{2.8.48}$$

where $\boldsymbol{dg}_{\alpha\beta} = \boldsymbol{X}_{\rho}(g_{\alpha\beta})\boldsymbol{\theta}^{\rho}$, and in a coordinate basis $\boldsymbol{dg}_{\alpha\beta} = g_{\alpha\beta,\rho}\boldsymbol{dx}^{\rho}$, and by

$$\boldsymbol{\Theta}^{\alpha} = \boldsymbol{0} \tag{2.8.49}$$

where $\boldsymbol{\Theta}^{\alpha} = \boldsymbol{d\theta}^{\alpha} + \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\theta}^{\sigma}$ (first Cartan structure equation), and $\boldsymbol{\Theta}^{\alpha}$ are the **torsion 2-forms** (see below).

Using the connection 1-forms $\boldsymbol{\omega}^{\alpha}{}_{\beta}$, the exterior derivative (2.8.6) and the exterior product (2.8.4), one can then construct the **curvature 2-forms** $\boldsymbol{\Omega}^{\alpha}{}_{\beta}$, for the moving frame $\boldsymbol{X}^{\alpha}$:

$$\boldsymbol{\Omega}^{\alpha}{}_{\beta} = \boldsymbol{d\omega}^{\alpha}{}_{\beta} + \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\beta} \qquad \text{(second Cartan structure equation).} \tag{2.8.50}$$

By taking the exterior derivative of expression (2.8.49), from the boundary of the boundary principle, in the form $d^2 = 0$, we get the **first Bianchi identity**:

$$\begin{aligned} \mathbf{0} = \boldsymbol{d}\boldsymbol{d}\boldsymbol{\theta}^{\alpha} &= \boldsymbol{d}\left(-\boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\theta}^{\sigma}\right) = -\left(\boldsymbol{\Omega}^{\alpha}{}_{\sigma} - \boldsymbol{\omega}^{\alpha}{}_{\rho} \wedge \boldsymbol{\omega}^{\rho}{}_{\sigma}\right) \wedge \boldsymbol{\theta}^{\sigma} \\ &- \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\rho} \wedge \boldsymbol{\theta}^{\rho} = -\boldsymbol{\Omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\theta}^{\sigma} \end{aligned} \tag{2.8.51}$$

and by taking the exterior derivative of expression (2.8.50), from $d^2 = 0$, we get

$$\boldsymbol{d}\boldsymbol{\Omega}^{\alpha}{}_{\beta} = \boldsymbol{d}\boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\beta} - \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{d}\boldsymbol{\omega}^{\sigma}{}_{\beta}. \tag{2.8.52}$$

By substituting $\boldsymbol{d}\boldsymbol{\omega}^{\alpha}{}_{\beta} = \boldsymbol{\Omega}^{\alpha}{}_{\beta} - \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\beta}$, we then have

$$\boldsymbol{d}\boldsymbol{\Omega}^{\alpha}{}_{\beta} + \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\Omega}^{\sigma}{}_{\beta} - \boldsymbol{\Omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\beta} = \mathbf{0}. \tag{2.8.53}$$

This is the **second Bianchi identity**. Finally, by defining the **exterior covariant derivative**, D: $\boldsymbol{D}\boldsymbol{\Omega}^{\alpha}{}_{\beta} \equiv \boldsymbol{d}\boldsymbol{\Omega}^{\alpha}{}_{\beta} + \boldsymbol{\omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\Omega}^{\sigma}{}_{\beta} - \boldsymbol{\Omega}^{\alpha}{}_{\sigma} \wedge \boldsymbol{\omega}^{\sigma}{}_{\beta}$, which maps a tensor-valued p-form (a p-form with tensor indices) into a tensor-valued $(p+1)$-form, we can rewrite the second Bianchi identity as:

$$\boldsymbol{D}\boldsymbol{\Omega}^{\alpha}{}_{\beta} = \mathbf{0}. \tag{2.8.54}$$

Equation (2.8.49) expresses that the torsion $\boldsymbol{\Theta}^{\alpha}$ is zero, and equation (2.8.48) that the connection is metric-compatible, that is, the covariant derivative of the metric is zero. It follows that the connection is uniquely[39] determined to be the standard Riemannian connection. Using the **natural coordinate basis**, $\{\overset{(c)}{X}_{\alpha}\} = \{\frac{\partial}{\partial x^{\alpha}}\}$ (a coordinate basis is also called **holonomic,** and a noncoordinate basis **anholonomic**), on a Riemannian manifold, one has then

$$\left(\overset{(c)}{\boldsymbol{\omega}}{}^{\alpha}{}_{\beta}\right)_{\gamma} = \Gamma^{\alpha}_{\beta\gamma} = \text{Christoffel symbols (expression 2.2.3)} \tag{2.8.55}$$

(for the expression of $\Gamma^{\alpha}_{\beta\gamma}$ in a general basis see the mathematical appendix),

$$\left(\overset{(c)}{\boldsymbol{\Theta}}{}^{\alpha}\right)_{\beta\gamma} = \Gamma^{\alpha}_{\gamma\beta} - \Gamma^{\alpha}_{\beta\gamma} \equiv T^{\alpha}_{[\gamma\beta]} = 0, \qquad \text{i.e., no torsion,} \tag{2.8.56}$$

and

$$\left(\overset{(c)}{\boldsymbol{\Omega}}{}^{\alpha}{}_{\beta}\right)_{\gamma\delta} = R^{\alpha}{}_{\beta\gamma\delta} = \text{Riemann curvature tensor (expression 2.2.5)} \tag{2.8.57}$$

and we can rewrite equations (2.8.51) and (2.8.54), in components, as

$$R^{\alpha}{}_{[\beta\gamma\delta]} = 0, \qquad \text{(eq. 2.8.45)} \tag{2.8.58}$$

and

$$R^{\alpha}{}_{\beta[\gamma\delta;\mu]} = 0, \qquad \text{(eq. 2.8.46).} \tag{2.8.59}$$

Equations (2.8.50) and (2.8.57) define the curvature tensor $R^{\alpha}{}_{\beta\gamma\delta}$ without the use of the covariant derivatives as in the standard definition (2.2.4).

In geometrodynamics the contracted second Bianchi identity, consequence of $d^2 = 0$, is especially important. In fact, the dynamical equations for matter and fields automatically follow from this identity plus the Einstein field equation (2.3.14). To derive the dynamical equations from the boundary of the boundary principle we first construct[11] the double dual of the Riemann tensor:

$$\left({}^{*}\boldsymbol{R}^{*}\right)_{\alpha\beta}{}^{\gamma\delta} \equiv \frac{1}{4}\,\epsilon_{\alpha\beta\mu\nu} R^{\mu\nu}{}_{\rho\sigma}\epsilon^{\rho\sigma\gamma\delta}. \tag{2.8.60}$$

We can then rewrite the Einstein tensor, $G^{\alpha}{}_{\beta}$, as

$$G^{\alpha}{}_{\beta} = \left({}^{*}\boldsymbol{R}^{*}\right)^{\alpha\sigma}{}_{\beta\sigma}. \tag{2.8.61}$$

We have, in fact,

$$\begin{aligned} \left({}^{*}\boldsymbol{R}^{*}\right)^{\alpha\sigma}{}_{\beta\sigma} &= \frac{1}{4}\,\epsilon^{\alpha\sigma\gamma\lambda}\epsilon_{\mu\nu\beta\sigma} R^{\mu\nu}{}_{\gamma\lambda} = -\frac{1}{4}\,\delta^{\alpha\gamma\lambda}{}_{\mu\nu\beta} R^{\mu\nu}{}_{\gamma\lambda} \\ &= -\frac{1}{4}\left(2\delta^{\alpha}{}_{\beta} R - 2R^{\alpha\gamma}{}_{\beta\gamma} - 2R^{\alpha\lambda}{}_{\beta\lambda}\right) = G^{\alpha}{}_{\beta} \end{aligned} \tag{2.8.62}$$

where we have used the relation (2.8.11). We now define the *star operator* ${}^{\star}(\cdots)$, with a star on the *left*, a duality operator which acts only on m-forms (with $m \le n$ on an n-dimensional manifold) and gives $(n - m)$-forms, that is, a duality operator which acts only on the m $(0 \le m \le n)$ antisymmetric covariant indices of a tensor and generates $n - m$ antisymmetric covariant indices. In other words, the ${}^{\star}(\cdots)$ operator acts only on the antisymmetric covariant indices of a tensor $T^{\alpha\beta\cdots}{}_{\gamma\delta\cdots}$, by first raising each covariant index with $g^{\mu\nu}$ and then by taking the dual, with $\epsilon_{\alpha\beta\cdots\mu}$, of these raised indices:

$$({}^{\star}\boldsymbol{T})^{\alpha\beta\cdots}{}_{\cdots\mu} = \frac{1}{m!}\,\epsilon_{\sigma\rho\cdots\mu} T^{\alpha\beta\cdots}{}_{\gamma\delta\cdots} g^{\sigma\gamma} g^{\rho\delta\cdots}. \tag{2.8.63}$$

Similarly, we define the *star operator* $(\cdots)^{\star}$, with a star on the *right*, as a duality operator which acts only on m-polyvectors (antisymmetric m-contravariant tensors) and gives $(n - m)$-polyvectors, that is, a duality operator which acts only on the m $(0 \le m \le n)$ antisymmetric contravariant indices of a tensor and generates $n - m$ antisymmetric contravariant indices:

$$(\boldsymbol{T}^{\star})^{\cdots\mu}{}_{\gamma\delta\cdots} = \frac{1}{m!}\,\epsilon^{\sigma\rho\cdots\mu} T^{\alpha\beta\cdots}{}_{\gamma\delta\cdots} g_{\alpha\sigma} g_{\beta\rho\cdots}. \tag{2.8.64}$$

We then introduce the vector-valued (a form with a vector index) 1-form, $(\boldsymbol{dP})^{\alpha}{}_{\beta} = \delta^{\alpha}{}_{\beta}$, sometimes called the Cartan unit tensor. By taking the star dual ${}^{\star}(\cdots)$ of both sides of the Einstein field equation $G^{\alpha}{}_{\beta} = \chi T^{\alpha}{}_{\beta}$,

$${}^{\star}\boldsymbol{G} = \chi\,{}^{\star}\boldsymbol{T}, \tag{2.8.65}$$

we have in components

$$\epsilon_{\sigma\beta\gamma\delta} G^{\alpha\sigma} = \chi \epsilon_{\sigma\beta\gamma\delta} T^{\alpha\sigma}. \quad (2.8.66)$$

By defining

$$[\boldsymbol{dP} \wedge \boldsymbol{R}]^{\alpha\mu\nu}_{\beta\rho\sigma} \equiv \frac{3!}{2!}\frac{3!}{2!} \delta^{[\alpha}_{[\beta} R^{\mu\nu]}{}_{\rho\sigma]} \quad (2.8.67)$$

here by $[\boldsymbol{S} \wedge \boldsymbol{T}]$ we mean exterior product of *both* the covariant and the contravariant parts of the antisymmetric tensors $\boldsymbol{S}$ and $\boldsymbol{T}$, that is we mean both antisymmetrization of the covariant indices of the product of $\boldsymbol{S}$ with $\boldsymbol{T}$ times a factor $\frac{(p+q)!}{p!q!}$ and antisymmetrization of the *contravariant* indices times a factor $\frac{(n+m)!}{n!\,m!}$. We can rewrite the left-hand side of the star dual, ${}^\star\boldsymbol{G}$, of the Einstein tensor, $\boldsymbol{G}$,

$$[\boldsymbol{dP} \wedge \boldsymbol{R}]^\star = {}^\star\boldsymbol{G} = \chi {}^\star\boldsymbol{T} \quad (2.8.68)$$

Indeed, we have, in components, using expressions (2.8.14) and (2.8.62):

$$\begin{aligned} &\frac{3}{2} g^{\gamma\tau} \epsilon_{\alpha\mu\nu\tau} \delta^{[\alpha}_{[\beta} R^{\mu\nu]}{}_{\rho\sigma]} = \frac{1}{4} g^{\gamma\tau} \epsilon_{\alpha\mu\nu\tau} \delta^{\alpha\mu\nu}{}_{\lambda\theta\varphi} \delta^{\lambda}_{[\beta} R^{\theta\varphi}{}_{\rho\sigma]} = \\ &-\frac{3}{2} g^{\gamma\tau} \epsilon_{\mu\nu\tau[\beta} R^{\mu\nu}{}_{\rho\sigma]} = -\frac{1}{4} g^{\gamma\tau} \epsilon_{\mu\nu\tau\lambda} \delta_{\beta\rho\sigma}{}^{\lambda\theta\varphi} R^{\mu\nu}{}_{\theta\varphi} = \\ &-\frac{1}{4} g^{\gamma\tau} \epsilon_{\mu\nu\tau\lambda} \epsilon_{\beta\rho\sigma\alpha} \epsilon^{\alpha\lambda\theta\varphi} R^{\mu\nu}{}_{\theta\varphi} = \epsilon_{\alpha\beta\rho\sigma} G^{\alpha\gamma}. \end{aligned} \quad (2.8.69)$$

By taking the exterior covariant derivative of equation (2.8.68) we then have

$$\boldsymbol{D}[\boldsymbol{dP} \wedge \boldsymbol{R}]^\star = \big(\boldsymbol{D}[\boldsymbol{dP} \wedge \boldsymbol{R}]\big)^\star = \big([\boldsymbol{DdP} \wedge \boldsymbol{R}] - [\boldsymbol{dP} \wedge \boldsymbol{DR}]\big)^\star = \boldsymbol{0} \quad (2.8.70)$$

where $\boldsymbol{DdP} = \boldsymbol{0}$, that is, there is no torsion, and $\boldsymbol{DR} = \boldsymbol{0}$ is the second Bianchi identity (2.8.54) as a consequence of $d^2 = 0$, that is, as a consequence of the boundary of the boundary principle. Finally, from the Einstein field equation (2.8.68), we have

$$\boldsymbol{D}^\star\boldsymbol{G} = \boldsymbol{D}^\star\boldsymbol{T} = \boldsymbol{0}, \quad (2.8.71)$$

that is, in components, using (2.3.7),

$$T^{\alpha\beta}{}_{;\beta} = 0. \quad (2.8.72)$$

The quantity $[\boldsymbol{dP} \wedge \boldsymbol{R}]^\star$ has a geometrical interpretation.[91,84,11] It may be thought of as the star dual of the moment of rotation, of a vector, associated with a three-cube and induced by the Riemann curvature (see fig. 2.7). The Einstein field equation may then be geometrically interpreted as identifying the star dual of the moment of rotation associated with a three-cube with the amount of

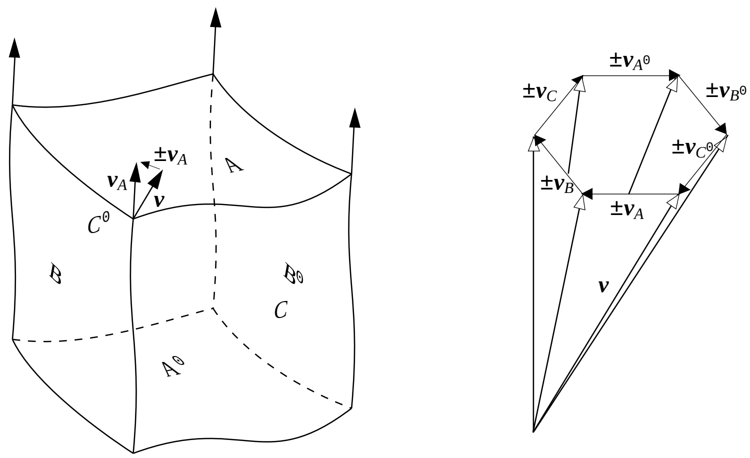


FIGURE 2.7. The rotation of a vector $\boldsymbol{v}$ associated with each face of a three-cube and induced by the Riemann curvature tensor, and the one-boundary of the two-boundary of a three-cube. *Left*: the rotation of a vector transported $\boldsymbol{v}_t$ parallel to itself around the indicated circuit, this rotation measures some components of the spacetime curvature (see eq. 2.4.19). *Right*: the rotations associated with all six faces together add up to zero; the diagram closes. It closes because each edge of the cube is traversed twice, and in opposite directions, in the circumnavigation of the two abutting faces of the cube: $\partial\partial = 0$.

energy-momentum of matter and fields contained in that three-cube:

$$\begin{pmatrix} \text{dual of} \\ \text{moment of} \\ \text{rotation} \\ \text{associated with} \\ \text{a three-cube} \end{pmatrix} = 8\pi \begin{pmatrix} \text{amount of} \\ \text{energy-momentum} \\ \text{in that} \\ \text{three-cube} \end{pmatrix}. \tag{2.8.73}$$

This is the geometrical content of the Einstein equation. Then, by applying to the Einstein field equation the simple but central topological 2-3-4 (in two-three-four dimensions) **boundary of the boundary principle**, $\partial^2 = 0$, one gets the *dynamical equations* for matter and fields.

2.9 BLACK HOLES AND SINGULARITIES

Black Holes and Gravitational Collapse

Collapse of a spherically symmetric star to a dense configuration[92–96] can on occasion put enough mass M inside a spherical surface of circumference $2\pi r$

as to make the terms $-(1-\frac{2M}{r})dt^2$ and $(1-\frac{2M}{r})^{-1}dr^2$ in the metric (2.6.35) reverse sign inside this surface. By analyzing the radial light cones (θ and ϕ constant), as calculated from $ds^2 = 0$ in the Schwarzschild coordinates of expression (2.6.35), we find that $\frac{dr}{dt} = \pm(1-\frac{2M}{r})$ tends to zero as it approaches the region $r = 2M$, and inside this region, where $r < 2M$, the future light cones point inward, toward $r = 0$ (fig. 2.8). Since particles, or photons, propagate within, or on, the light cones, no photon can escape from such a region, nor any particle that follows classical physics. It is no wonder that such a collapsed star received[92] the name "**black hole**"[19,97–101,144] as early as 1967. This strange behavior of the Schwarzschild spacetime geometry extends over the region where r is less than the so-called Schwarzschild radius, $r_s = 2M$. A black hole with Earth mass has a Schwarzschild radius of about 0.88 cm and one of Sun's mass $M_\odot$ of about 3 km.

The X-ray telescope UHURU floating above the atmosphere discovered in 1971 (see ref. 130) the first compelling evidence for a black hole, Cygnus X-1. Its mass is today estimated as of the order of 10 $M_\odot$ (since then, other black hole candidates have been found in X-ray binary systems, for example in nova V404 Cygni[142] and in Nova Muscae[143]). Recently, H. Ford et al., using the Faint Object Spectrograph of the *Hubble Space Telescope*, have observed gas orbiting at high velocity near the nucleus of the elliptical galaxy M87.[147] This observation provides a decisive experimental evidence for a supermassive black hole, source of the strong gravitational field that keeps the gas orbiting (see picture 4.5, p. 203). A star collapses by contraction,[93–96] after the end of the nuclear reactions that kept the star in equilibrium, if the mass of the star is

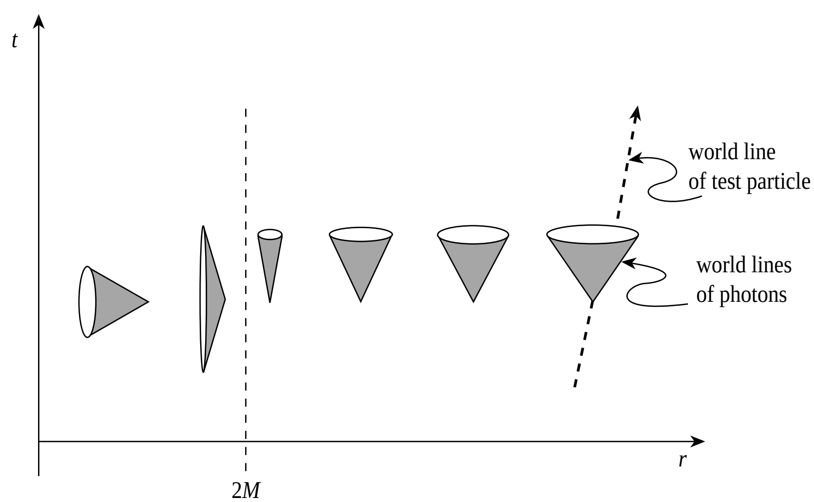


FIGURE 2.8. Future light cones in Schwarzschild coordinates outside, near, and inside the region $r = 2M$.

larger than a critical value, the *critical mass* (in general relativity, for a neutron star and depending from the equation of state used, at most $\sim$ 2–3$M_\odot$; the *Chandrasekhar limit* for the mass of a white dwarf is about 1.2 $M_\odot$).

The first detailed treatment of gravitational collapse within the framework of Einstein geometrodynamics was given in 1939 by Oppenheimer and Snyder.[93] For simplicity they treated the collapsing system as a collection of dust particles ($p = 0$), so that all the problems of pressure and temperature could be overlooked. Each particle would then move freely under the gravitational attraction of the others. More realistic equations of state have been later used,[96,99] without avoiding the collapse.

However, do we know enough about matter to be sure that it cannot successfully oppose collapse? We understand electromagnetic radiation better than we understand the behavior of matter at high density. Then why not consider a star containing no matter at all, an object built exclusively out of light, a "gravitational electromagnetic entity" or "geon," described in section 2.10, deriving its mass solely from photons, and these photons held in orbit solely by the gravitational attraction of that very mass?[71] It turns out that a geon has the stability—and the instability—of a pencil standing on its tip.[95] The geon does not let its individual photons escape any more than the pencil lets its individual atoms escape. But that swarm of photons, collectively, like the assembly of atoms that make up the pencil, collectively, can fall one way or the other. Starting slowly at first, it can expand outward more and more rapidly and explode into its individual photons. Equally easily, it can fall the other way slowly at first, then more and more rapidly to complete gravitational collapse. Thus it does not save one from having to worry about gravitational collapse to turn from matter to "pure" radiation.

A closer look at matter itself shows that "the harder it resists, the harder it falls": pressure itself has weight, and that weight creates more pressure, a "regenerative cycle" out of which again the only escape is collapse (see § 4.5).[96]

Gravitational collapse will have quite a different appearance according as it is studied by a faraway observer or a traveler falling in with, and at the outskirts of, the cloud of dust. The traveler will arrive in a very short time at a condition of infinite gravitational stress. If he sends out a radio "beep" every second of his existence, he will get off only a limited number of messages before the collapse terminates. In contrast, the faraway observer will receive these beeps at greater and greater time intervals; and, wait as long as he will, he will never receive any of the signals given out by the traveler after his crossing of the intangible horizon, $r_s = 2M$. Moreover, the cloud of dust will appear to the faraway observer, not to be falling ever faster, but to slow up and approach asymptotically a limiting sphere with the dimensions of the horizon. As it freezes down to this standard size it grows redder and fainter by the instant, and quickly becomes invisible. In other words, the observer on the surface of the collapsing star will pass through the horizon in a finite amount of his proper time, measured by his clocks. In contrast, an observer far from the collapsing star will see the collapse slow down and only asymptotically reach the horizon. However, since the intensity of the light he receives will exponentially decrease as the surface of the star approaches the horizon, after a short time he essentially will not receive any more light emitted from the collapsing star (however, see the Hawking radiation below). This phenomenon of different speed of the collapse is due to the gravitational time dilation of clocks, explained in section 3.2.2, and experimentally observed in a variety of experiments in weak fields (§ 3.2.2).[21] From the metric (2.6.35), we have

$$\begin{aligned}\Delta\tau|_{r'\cong 2M} &\equiv \begin{pmatrix}\text{interval of proper time}\\ \text{measured by an external}\\ \text{observer, at } r'\text{, near the}\\ \text{horizon } r_s = 2M\end{pmatrix}\\ &= \left(1 - \frac{2M}{r'}\right)^{1/2}\Delta t = \left(1 - \frac{2M}{r'}\right)^{1/2}\Delta\tau_\infty\\ &= \varepsilon \times \begin{pmatrix}\text{interval of proper time}\\ \text{measured by an asymptotic}\\ \text{observer}\end{pmatrix}\end{aligned} \tag{2.9.1}$$

where $\varepsilon \equiv (1 - \frac{2M}{r'})^{1/2} \ll 1$. This is the sense in which time goes slower near a black hole. Put an atomic clock on the surface of a planet. Let it send signals to a higher point. The interval from pulse to pulse of this clock is seen to be greater than the interval between pulse and pulse of an identical clock located at the higher point. In this sense the clock closer to the planet's surface goes slower than the clock further away. Likewise a clock somehow suspended close above a black hole, measuring proper time: $\Delta\tau_{BH} = (1 - \frac{2M}{r'})^{1/2}\Delta t = \varepsilon\Delta t$, will send signals to a faraway observer, equipped with an identical clock, measuring proper time: $\Delta\tau_\infty \cong \Delta t = \Delta\tau_{BH}/\varepsilon$. Therefore, the spacing between ticks of the clock just above the black hole is seen to be much larger than the spacing between ticks of the clock of the faraway observer.

Features of a Black Hole

Not even light signals or radio messages will escape from inside the horizon of the collapsed object. The only feature of the black hole that will be observed is its gravitational attraction[97–101,19] (however, see the Hawking radiation below). What falls into a black hole carries in mass and angular momentum, and it can also carry in electric charge. These are the only attributes that a black hole conserves out of the matter that falls into it. All other particularities, all other details, all other physical properties of "matter" are extinguished. The resulting stationary black hole, according to all available theoretical evidence, is completely characterized by its mass, its charge, and its angular momentum, and by nothing more. Jokingly put, "*a black hole has no hair*."[11]

Of the number of particles that went in not a trace is left, if present physics is safe as our guide. Not the slightest possibility is evident, even in principle, to distinguish between three black holes of the same mass, charge, and angular momentum, the first made from particles, the second made from antiparticles, and the third made primarily from pure radiation. This circumstance deprives us of all possibility to count or even define the number of particles at the end and compare it with the starting count. In this sense the laws of conservation of particle number are not violated; they are transcended.

The typical black hole is spinning and has angular momentum. This is a very strange kind of spin. One cannot "touch one's finger to the flywheel" to find it. The flywheel, the black hole, is so "immaterial," so purely geometrical, so untouchable, that no such direct evidence for its spin is available. Evidence for the spin of the black hole is obtainable by indirect means. For this purpose it is enough to put a gyroscopic compass in polar orbit around the black hole. The gyroscopic compass, pointed originally at a distant star, will slowly sweep about the circuit of the heavens, in sympathy with the rotation of the black hole, but at a far slower rate. At work on the gyro, in addition to the normal direct pull of gravity, is a new feature of geometry predicted by Einstein's theory. This "gravitomagnetic force" is as different from standard gravity as magnetism is different from electricity. An electric charge circling in orbit creates magnetism. A spinning mass creates gravitomagnetism.

We are far from being able today to observe gravitomagnetism of a spinning black hole. However, space experiments are in active development (GP-B and LAGEOS III; chap. 6) to measure the gravitomagnetic effects, on an orbiting gyroscope, due to the slow rotation of Earth.

The Event Horizon

Using the Schwarzschild coordinates of expression (2.6.35), at the Schwarzschild horizon, $r_S = 2M$, we have $g_{11} = -g_{00}^{-1} \xrightarrow{r=2M} \infty$. However, the **Schwarzschild horizon** is not a true singularity but just a **coordinate singularity**.

The quantities that have an intrinsic geometrical meaning, independent from the particular coordinates that are used, are the scalar invariants[15] constructed using the Riemann curvature tensor and the metric tensor. No invariant,[19] built with the curvature and metric tensors, diverges on the horizon, $r_s = 2M$. The Schwarzschild horizon is just a pathology of the coordinates of expression (2.6.35), but not a true geometrical singularity (see below). Indeed, with a coordinate transformation, for example to Eddington-Finkelstein[102,103] coordinates, or to Kruskal-Szekeres[104,105] coordinates, one can extend the solution (2.6.35) to a solution covering the whole Schwarzschild geometry with nonsingular

FIGURE 2.9. Alternative interpretations of the three-dimensional "maximally extended Schwarzschild metric" of Kruskal at time $t' = 0$. (a) A connection in the sense of Einstein and Rosen (Einstein-Rosen bridge)[106] between two otherwise Euclidean spaces. (b) and (c) A **wormhole** connecting two regions in one Euclidean space, in (c) not orientable with the topology of a Möbius strip (in the case where these regions are extremely far apart compared to the dimensions of the throat of the wormhole). Case (a) has the same curvature but different topology from cases (b) and (c). For a discussion on causality in a case of type (b) or (c) see refs. 107–109 and 138–141.

metric components at $r_s = 2M$ (see fig. 2.9). With the transformation to **Kruskal-Szekeres coordinates**:[11,19]

$$\begin{bmatrix} x' = \left(\dfrac{r}{2M} - 1\right)^{\frac{1}{2}} e^{r/4M} \cosh\left(\dfrac{t}{4M}\right) \\ \\ t' = \left(\dfrac{r}{2M} - 1\right)^{\frac{1}{2}} e^{r/4M} \sinh\left(\dfrac{t}{4M}\right) \end{bmatrix} : \quad \text{for} \quad r > 2M$$

$$\begin{bmatrix} x' = \left(1 - \dfrac{r}{2M}\right)^{\frac{1}{2}} e^{r/4M} \sinh\left(\dfrac{t}{4M}\right) \\ \\ t' = \left(1 - \dfrac{r}{2M}\right)^{\frac{1}{2}} e^{r/4M} \cosh\left(\dfrac{t}{4M}\right) \end{bmatrix} : \quad \text{for} \quad r < 2M, \qquad (2.9.2)$$

one thus gets

$$\begin{aligned} ds^2 = &\left(\frac{32M^3}{r}\right) e^{-r/2M} \left(-dt'^2 + dx'^2\right) \\ &+ r^2(t', x') \left(d\theta^2 + \sin^2\theta d\phi^2\right) \end{aligned} \qquad (2.9.3)$$

where r is a function of x' and t' implicitly determined, from expression (2.9.2), by

$$\left(\frac{r}{2M} - 1\right) e^{r/2M} = x'^2 - t'^2. \qquad (2.9.4)$$

The metric (2.9.3), in Kruskal-Szekeres coordinates, explicitly shows that the Schwarzschild geometry is well-behaved at $r_s = 2M$ and that is possible to extend analytically the Schwarzschild solution (2.6.35) to cover the whole Schwarzschild geometry (see fig. 2.9).

Black Hole Evaporation

In 1975 Hawking[110] discovered the so-called process of *black hole evaporation* (fig. 2.10). Quantum theory allows a process to happen at the horizon analogous to the Penrose process.[111] In the Penrose process two already existing particles trade energy in a region outside the horizon of a spinning black hole (see 2.6.36) called the ergosphere, the only domain where macroscopic masses of positive energy and of negative energy can coexist. Because the ergosphere shrinks to extinction when a black hole is deprived of all spin, the Penrose process applies only to a rotating, or "live," black hole. In contrast, the Hawking process takes place at the horizon itself and thus operates as effectively for a nonrotating black

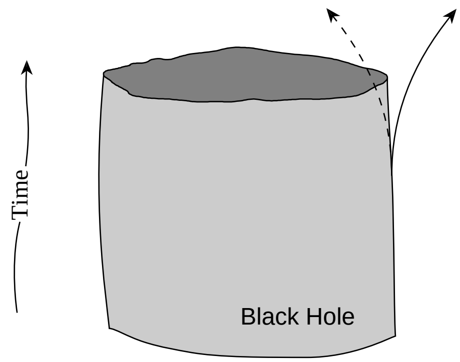


FIGURE 2.10. The Hawking[110] evaporation process capitalizes on the fact that space is nowhere free of so-called quantum vacuum fluctuations, evidence that everywhere latent particles await only opportunity—and energy—to be born. Associated with such fluctuations at the surface of a black hole, a might-have-been pair of particles or photons can be caught by gravity and transformed into a real-life particle or photon (solid arrow) that evaporates out into the surroundings and an antiparticle or counterphoton (dashed arrow) that "goes down" the black hole.

hole as for a rotating one. Furthermore, unlike the Penrose process, it involves a pair of newly created microscopic particles.

According to the uncertainty principle for the energy, $\Delta E \Delta t \gtrsim \hbar$, that is, space—pure, empty, energy-free space—all the time and everywhere experiences so-called quantum vacuum fluctuations at a very small scale of time, of the order of 10^{-44} s and less. During these quantum fluctuations, pairs of particles appear for an instant from the emptiness of space—perhaps an electron and an antielectron pair or a proton and an antiproton pair. Particle-antiparticle pairs are in effect all the time and everywhere being created and destroyed. Their destruction is so rapid that the particles never come into evidence at any everyday scale of observation. For this reason, the pairs of particles everywhere being born and dying are called virtual pairs. Under the conditions at the horizon, a virtual pair may become a real pair.

In the Hawking process, two newly created particles exchange energy, one acquiring negative energy $-E$ and the other positive energy E. Slightly outside the horizon of a black hole, the negative energy photon has enough time Δt to cross the horizon. Therefore, the negative energy particle flies inward from the horizon; the positive energy particle flies off to a distance. The energy it carries with it comes in the last analysis from the black hole itself. The massive object is no longer quite so massive because it has had to pay the debt of energy brought in by the negative energy member of the newly created pair of particles.

Radiation of light or particles from any black hole of a solar mass or more proceeds at an absolutely negligible rate—the bigger the mass the cooler the surface and the slower the rate of radiation. The calculated Bekenstein-Hawking temperature of a black hole of 3 $M_\odot$ is only 2×10^{-8} degrees above the absolute zero of temperature. The total thermal radiation calculated to come from its 986 square kilometers of surface is only about 1.6×10^{-29} watt, therefore this evaporation process would not be able to affect in any important way black holes of about one solar mass or more. A black hole of any substantial mass is thus deader than any planet, deader even than any dead moon—when it stands in isolation.

Singularities

The $r = 2M$ region of the Schwarzschild metric (2.6.35) is a mere coordinate singularity; however, the $r = 0$ region, where $g_{00} = -g_{11}^{-1} \xrightarrow{r=0} \infty$, is a true geometrical singularity,[19] where, as for the big bang and big crunch singularities of some cosmological models (see chap. 4), some curvature invariants diverge; for example the Kretschmann invariant for the Schwarzschild metric is $R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} \sim \frac{m^2}{r^6} \xrightarrow{r=0} \infty$ (see § 6.11).

Indeed, besides coordinate singularities, or pathologies of a coordinate system removable with a coordinate transformation, there are various types of true geometrical singularities.[112–115]

Usually, in a physically realistic solution, a singularity is characterized by diverging curvature.[19] However, on a curved manifold the individual components of the Riemann tensor depend on the coordinates used. Therefore, one defines the true curvature singularities using the invariants built by contracting the Riemann tensor $R^{\alpha}{}_{\beta\mu\nu}$, with $g_{\alpha\beta}$ and with $\epsilon_{\alpha\beta\mu\nu}$. The regions where these invariants diverge are called **scalar polynomial curvature singularities**. One may also measure the components of the Riemann tensor with respect to a local basis parallel transported along a curve. In this case the corresponding curvature singularities are called **parallelly propagated curvature singularities**.

It is usual to assume that spacetime is a differentiable manifold (i.e., a manifold that is covered by a family of charts, such that in the intersections between the charts, the coordinates x^{α} of a chart as a function of the coordinates $\overline{x}^{\alpha}$ of another chart, $x^{\alpha} = x^{\alpha}(\overline{x}^{\alpha})$, are continuous and with continuous derivatives, C^{∞}), where space and time intervals and other physical quantities can be measured, and standard equations of physics hold in a neighborhood of every event. Then a curvature singularity is not part of the differentiable manifold called spacetime. Therefore, in such manifolds with singularities cut out, there will exist curves incomplete in the sense that they cannot be extended.

To distinguish between different types of incompleteness of a manifold, various definitions have been given. First, a manifold is called inextendible if it includes all the nonsingular spacetime points.[19] The definition of geodesic completeness is useful to characterize an incomplete manifold. A manifold is called geodesically complete if every geodesic can be extended to any value of its affine parameter (§ 2.4). In particular a manifold is **not** timelike or null **geodesically complete**, if it has incomplete timelike or null geodesics. In this case the history of a freely moving observer (or a photon), on one of these incomplete geodesics, cannot be extended after (or before) a finite amount of proper time. However, this definition does not include the type of singularity that a nonfreely falling observer, moving with rockets on a nongeodesic curve, may encounter in some manifolds. To describe these types of singularities on nongeodesic curves, one can give the definition of bundle-completeness or b-completeness. One first constructs on any continuous curve, with continuous first derivatives, a generalized affine parameter that in the case of a geodesic reduces to an affine parameter. An inextendible manifold (with all nonsingular points) is called **bundle-complete**, or *b-complete*, if for every curve of finite length, measured by the generalized affine parameter from a point p, there is an endpoint of the curve in the manifold. Bundle-completeness implies geodesic completeness, but not vice versa. Usually, in physically realistic solutions, a spacetime which is bundle-incomplete has curvature singularities on the b-incomplete curves (however, see the Hawking-Ellis discussion[19] of the Taub-NUT space).

In 1965 Roger Penrose proved a theorem about the existence of singularities,[112] of the type corresponding to null geodesic incompleteness, without using any particular assumption of exact symmetry.

Incomplete null geodesics exist on a manifold if:

1. The **null convergence condition** is satisfied: $R_{\alpha\beta}k^{\alpha}k^{\beta} \geq 0$, for every null vector k^{α}.
2. In the manifold there exists a noncompact Cauchy surface, that is, a noncompact spacelike hypersurface such that every causal path without endpoint intersects it once and only once (see chap. 5).
3. In the manifold there exists a **closed trapped surface.** A closed trapped surface is a closed (compact, without boundary) spacelike two-surface such that both the ingoing and the outgoing light rays moving on the null geodesics orthogonal to the surface converge toward each other.

Such a closed trapped surface is due to a very strong gravitational field that attracts back and causes the convergence even of the outgoing light rays. An example of closed trapped surface is a two-dimensional spherical surface inside the Schwarzschild horizon. Even the outgoing photons emitted from this surface are attracted back and converge due to the very strong gravitational field. Since not even the outgoing orthogonal light rays can escape from the closed trapped surface, all the matter, with velocity less than c, is also trapped and cannot escape from this surface. Closed trapped surfaces occur if a star collapses below its Schwarzschild radius. As we have previously observed, this should happen if a cold star, white dwarf, or neutron star, or white dwarf or neutron star core of a larger star, after the end of the nuclear reactions that kept the star in equilibrium, has a mass above a critical value of a few solar masses (in general relativity, for a neutron star, depending from the equation of state used, at most $\sim 3\ M_{\odot}$). Therefore, any such star or star core should collapse within the horizon and generate closed trapped surfaces and singularities, according to various singularity theorems[112–115] and in particular according to the 1965 Penrose theorem[112] and to the 1970 Hawking-Penrose theorem.[115]

Singularities of the type of incomplete timelike and null geodesics occur in a manifold, if:

1. $R_{\alpha\beta}u^{\alpha}u^{\beta} \geq 0$ for every nonspacelike vector u^{α}.
2. Every nonspacelike geodesic has at least a point where:

$$u_{[\alpha}R_{\beta]\gamma\delta[\mu}u_{\nu]}u^{\gamma}u^{\delta} \neq 0,$$

   where u^{α} is the tangent vector to the geodesics (the manifold is not too highly symmetric): this is the so-called **generic condition**.
3. There are no closed timelike curves; this causality condition is called **chronology condition** (see the 1949 Gödel model universe, discussed in § 4.6, as an example of solution violating the chronology condition).
4. There exists a closed trapped surface (or some equivalent mathematical condition is satisfied; see Hawking and Ellis).[19]

We note that the null convergence condition (1) of the Penrose theorem is a consequence of the **weak energy condition**, $T_{\alpha\beta}u^{\alpha}u^{\beta} \geq 0$, for every timelike vector u^{α}, plus the Einstein field equation (2.3.14) (even including a cosmological term), $R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = \chi T_{\alpha\beta}$. The **timelike convergence condition**, $R_{\alpha\beta}u^{\alpha}u^{\beta} \geq 0$, for every timelike vector u^{α}, is a consequence of the Einstein field equation plus the condition $T_{\alpha\beta}u^{\alpha}u^{\beta} \geq u^{\alpha}u_{\alpha}(\frac{1}{2}T - \frac{1}{8\pi}\Lambda)$, for every timelike vector u^{α}; for $\Lambda = 0$ this is called the **strong energy condition** for the energy-momentum tensor.

We conclude this brief introduction to spacetime singularities by observing that, probably, the problem of the occurrence of the singularities in classical geometrodynamics might finally be understood[95] only when a consistent and complete quantum theory of gravity[116,145] is available. Question: Does a proper quantum theory of gravity rule out the formation of such singularities?

2.10 GRAVITATIONAL WAVES

As in electromagnetism in which there are electromagnetic perturbations propagating with speed c in a vacuum—electromagnetic waves—Einstein geometrodynamics predicts curvature perturbations propagating in the spacetime—**gravitational waves**.[117–121]

In this section we derive a simple, weak field, wave solution of the field equation (2.3.14). Let us first consider a perturbation of the flat Minkowski metric $\eta_{\alpha\beta}$:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{2.10.1}$$

where $h_{\alpha\beta}$ is a small perturbation of $\eta_{\alpha\beta}$: $|h_{\alpha\beta}| \ll 1$. We then define

$$\begin{aligned} h^{\alpha}{}_{\beta} &\equiv \eta^{\alpha\sigma} h_{\sigma\beta} \\ h^{\alpha\beta} &\equiv \eta^{\alpha\sigma}\eta^{\beta\rho} h_{\sigma\rho} \\ h &\equiv h^{\alpha}{}_{\alpha} = \eta^{\sigma\rho} h_{\sigma\rho}. \end{aligned} \tag{2.10.2}$$

Therefore, to first order in $|h_{\alpha\beta}|$, we have

$$g^{\alpha\beta} = \eta^{\alpha\beta} - h^{\alpha\beta}. \tag{2.10.3}$$

From the definition of Ricci tensor (§ 2.3), we then have up to first order

$$R^{(1)}_{\alpha\beta} = \Gamma^{\sigma}_{\alpha\beta,\sigma} - \Gamma^{\sigma}_{\alpha\sigma,\beta} = \frac{1}{2}\left(-\Box h_{\alpha\beta} + h^{\sigma}{}_{\beta,\sigma\alpha} + h^{\sigma}{}_{\alpha,\sigma\beta} - h_{,\alpha\beta}\right) \tag{2.10.4}$$

where $\Box = \eta^{\alpha\beta} \frac{\partial^2}{\partial x^\alpha \partial x^\beta}$ is the d'Alambertian operator. Therefore the Einstein field equation, in the alternative form (2.3.17), can be written

$$-\Box h_{\alpha\beta} + h^{\sigma}{}_{\beta,\sigma\alpha} + h^{\sigma}{}_{\alpha,\sigma\beta} - h_{,\alpha\beta} = 16\pi\Big(T_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta} T\Big) \tag{2.10.5}$$

where $T = \eta^{\sigma\rho} T_{\sigma\rho} = -\frac{1}{8\pi} R$. With an infinitesimal coordinate, or gauge, transformation, $x'^{\alpha} = x^{\alpha} + \xi^{\alpha}$ (see § 2.6), we then have

$$h'_{\alpha\beta} = h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} \tag{2.10.6}$$

where, of course, $h'_{\alpha\beta}$ is still a solution of the field equation (gauge invariance of the field equation). Therefore, if for ξ_α we choose a solution of the differential equation

$$\Box \xi_\alpha = h^{\sigma}{}_{\alpha,\sigma} - \frac{1}{2} h^{\sigma}{}_{\sigma,\alpha}\,, \tag{2.10.7}$$

we have

$$h'^{\sigma}{}_{\alpha,\sigma} - \frac{1}{2} h'^{\sigma}{}_{\sigma,\alpha} = h^{\sigma}{}_{\alpha,\sigma} - \frac{1}{2} h^{\sigma}{}_{\sigma,\alpha} - \xi_{\alpha,\sigma}{}^{\sigma} = 0. \tag{2.10.8}$$

In this gauge, $(h'^{\sigma}{}_{\alpha} - \frac{1}{2}\delta^{\sigma}{}_{\alpha} h')_{,\sigma} = 0$, sometimes called the *Lorentz gauge*, the field equation becomes (dropping the prime in $h'_{\alpha\beta}$)

$$\Box h_{\alpha\beta} = -16\pi\Big(T_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta} T\Big). \tag{2.10.9}$$

As in electromagnetism,[44] a solution to this equation is the retarded potential:

$$h_{\alpha\beta}(\boldsymbol{x}, t) = 4 \int \left\{ \frac{\left[T_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta} T\right](\boldsymbol{x}', t - |\boldsymbol{x} - \boldsymbol{x}'|)}{|\boldsymbol{x} - \boldsymbol{x}'|} \right\} d^3x'. \qquad (2.10.10)$$

This solution represents a gravitational perturbation propagating at the speed of light, $c \equiv 1$. When $T_{\alpha\beta} = 0$, we then have, in the Lorentz gauge,

$$\Box h_{\alpha\beta} = 0. \qquad (2.10.11)$$

This is the wave equation for $h_{\alpha\beta}$. We recall that in electromagnetism, in the Lorentz gauge, $A^{\alpha}{}_{,\alpha} = 0$, we have the sourceless wave equation for A^{α}: $\Box A^{\alpha} = 0$. Correspondingly, a simple solution of the wave equation (2.10.11) for $h_{\alpha\beta}$ is a plane wave. By choosing the z-axis as the propagation axis of the plane wave, we then have

$$\left(\frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial t^2}\right) h_{\alpha\beta} = 0 \qquad (2.10.12)$$

where $h_{\alpha\beta} = h_{\alpha\beta}(z \pm t)$, that is, $h_{\alpha\beta}$ is a function of $(z \pm t)$, where $c \equiv 1$. From expression (2.10.6), it follows that the Lorentz condition (2.10.8) and the simple form (2.10.11) of the vacuum field equation for $h_{\alpha\beta}$ are invariant for any infinitesimal transformation $x'^{\alpha} = x^{\alpha} + \xi^{\alpha}$, if ξ^{α} is a solution of $\Box \xi^{\alpha} = 0$. Here gravity is similar to electromagnetism where, with the gauge transformation $A'^{\alpha} = A^{\alpha} + \phi^{,\alpha}$, if $\Box \phi = 0$, the Lorentz condition is preserved, $A^{\alpha}{}_{,\alpha} = A'^{\alpha}{}_{,\alpha} = 0$, and we still have $\Box A'^{\alpha} = 0$. Therefore, by performing an infinitesimal coordinate transformation, with the four components of ξ^{α} solutions of $\Box \xi^{\alpha} = 0$, for a plane gravitational wave, $h_{\alpha\beta} = h_{\alpha\beta}(z \pm t)$, it is possible to satisfy the four conditions: $h_{i0} = 0$ and $h \equiv h^{\sigma}{}_{\sigma} = 0$; that is, the trace of $h_{\alpha\beta}$ equal to zero. Since in this gauge we have $h^{\alpha}{}_{\beta} - \frac{1}{2}\delta^{\alpha}{}_{\beta} h = h^{\alpha}{}_{\beta}$, the Lorentz gauge condition becomes simply $h^{\sigma}{}_{\alpha,\sigma} = 0$. Therefore, for the weak field plane gravitational wave $h_{\alpha\beta}(z \pm t)$, and more generally for any weak field gravitational wave, linear superposition of plane waves, in this gauge, from $h^{\sigma}{}_{\alpha,\sigma} = 0$, we can set $h_{00} = 0$.

Summarizing in this **gauge**, called **transverse-traceless** (transverse because the wave is orthogonal to its direction of propagation), we have

$$h^{\mathrm{TT}}_{\alpha 0} = 0, \text{ i.e., } h^{\mathrm{TT}}_{\alpha\beta} \text{ has spatial components only,} \qquad (2.10.13)$$

and

$$h^{\mathrm{TT}} \equiv h^{\mathrm{TT}\alpha}{}_{\alpha} = 0, \text{ i.e., } h^{\mathrm{TT}}_{\alpha\beta} \text{ is traceless,} \qquad (2.10.14)$$

and

$$h^{\mathrm{TT}k}{}_{i,k} = 0, \text{ i.e., } h^{\mathrm{TT}}_{ij} \text{ is transverse.} \qquad (2.10.15)$$

Finally, from expressions (2.10.13), (2.10.14), and (2.10.15), for the plane wave $h^{\mathrm{TT}}_{\alpha\beta}(z \pm t)$ described by equation (2.10.12), apart from integration constants, in the transverse-traceless gauge we get $h^{\rm TT}_{zz} = h^{\rm TT}_{zx} = h^{\rm TT}_{zy} = 0$, and a solution to $\Box h^{\rm TT}_{\alpha\beta} = 0$ is

$$\begin{aligned} h^{\rm TT}_{xx} &= -h^{\rm TT}_{yy} = A_+ e^{-i\omega(t-z)} \\ h^{\rm TT}_{xy} &= h^{\rm TT}_{yx} = A_\times e^{-i\omega(t-z)} \end{aligned} \tag{2.10.16}$$

where as usual we take the real part of these expressions, with all the other components of $h^{\rm TT}_{\alpha\beta}$ equal to zero to first order. This expression (2.10.16) describes a plane gravitational wave as a perturbation of the spacetime geometry, traveling with speed c. This perturbation of the spacetime geometry corresponds to the curvature perturbation $R_{i0j0} = -R_{i0jz} = R_{izjz} = -\frac{1}{2} h^{\rm TT}_{ij,00}$ traveling with speed c on the flat background, where i and j are 1 or 2.

In this simple case of a weak field, plane gravitational wave, in the transverse-traceless coordinate system (2.10.13)–(2.10.15), one can easily verify that test particles originally at rest in the flat background $\eta_{\alpha\beta}$ before the passage of the gravitational wave will remain at rest *with respect to the coordinate system* during the propagation of the gravitational wave. In fact, from the geodesic equation (2.4.13), to first order in $h^{\rm TT}_{\alpha\beta}$, we have

$$\frac{Du^\alpha}{ds} \cong \frac{du^\alpha}{ds} = 0. \tag{2.10.17}$$

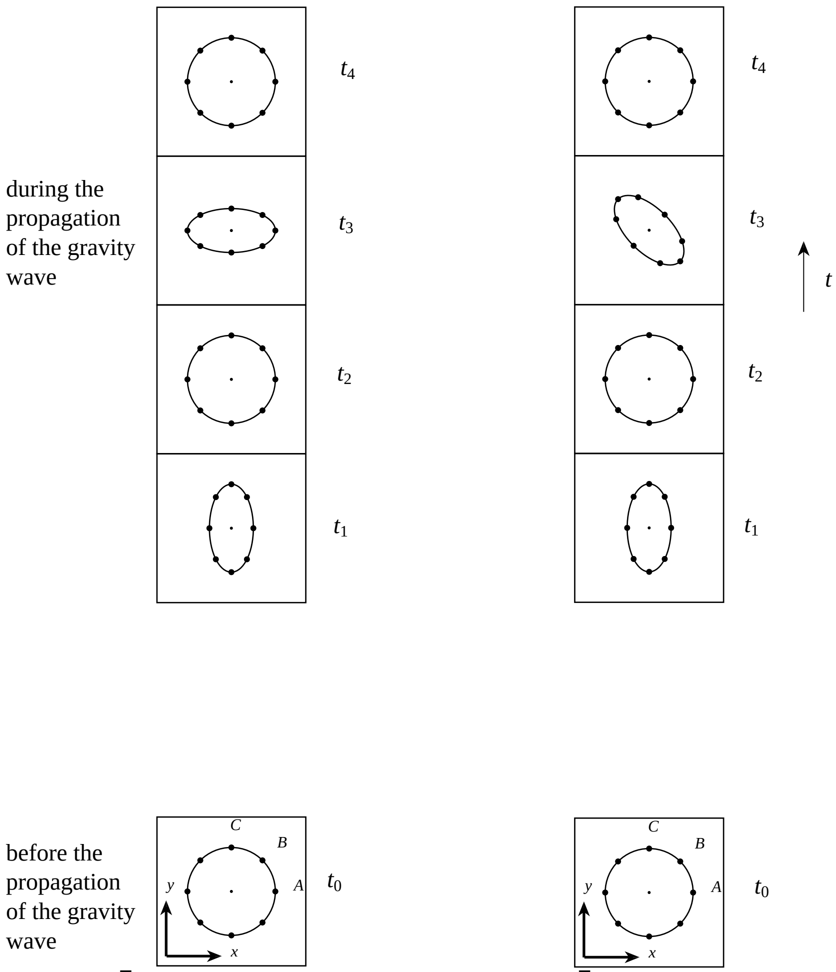
However, the *proper distance* between the two test particles at rest in x^i and $x^i + dx^i$ is given by $dl^2 = g_{ik} dx^i dx^k$. Therefore, since $g_{ik} = \eta_{ik} + h_{ik}$ changes with time, the proper distance between the test particles will *change* with time during the passage of the gravitational wave. For a plane wave propagating along the z-axis in the transverse-traceless gauge, the proper distance between particles in the xy-plane is given by

$$dl = \left[\left(1 + h^{\rm TT}_{xx}\right) dx^2 + \left(1 - h^{\rm TT}_{xx}\right) dy^2 + 2h^{\rm TT}_{xy} dx dy \right]^{\frac{1}{2}}. \tag{2.10.18}$$

For the particles A, B, and C of figure 2.11, on a circumference with center at $x^\alpha = 0$, with coordinates

$$x^i_A \equiv (l, 0, 0); \quad x^i_B \equiv \left(\frac{l}{\sqrt{2}}, \frac{l}{\sqrt{2}}, 0 \right); \quad \text{and} \quad x^i_C \equiv (0, l, 0) \tag{2.10.19}$$

from the expression (2.10.18) for dl and from the expression (2.10.16) for $h^{\rm TT}_{xx}$ and $h^{\rm TT}_{xy}$, we immediately find the behavior of the proper distance between test particles on a circumference due to the passage of a plane gravitational wave perpendicularly to the circumference, behavior that is shown in figure 2.11. Case I, $A_+ \neq 0$ and $A_\times = 0$, and case II, $A_+ = 0$ and $A_\times \neq 0$, describe two waves with polarizations at 45° one from the other. Of course one can get the same result by using the geodesic deviation equation (see § 3.6.1).



$$\begin{bmatrix} A_+ \neq 0 \\ A_\times = 0 \end{bmatrix} \qquad \begin{bmatrix} A_+ = 0 \\ A_\times \neq 0 \end{bmatrix}$$

$$\begin{cases} \delta l_A^+ = \frac{1}{2} l h_{xx}^{\mathrm{TT}} = \frac{1}{2} l A_+ \cos\omega(t-z) \\ \delta l_B^+ = 0 \\ \delta l_C^+ = -\frac{1}{2} l h_{xx}^{\mathrm{TT}} = -\frac{1}{2} l A_+ \cos\omega(t-z) \end{cases} \qquad \begin{cases} \delta l_A^\times = 0 \\ \delta l_B^\times = \frac{1}{2} l h_{xy}^{\mathrm{TT}} = \frac{1}{2} l A_\times \cos\omega(t-z) \\ \delta l_C^\times = 0 \end{cases}$$

FIGURE 2.11. Effect of weak plane gravitational wave, propagating along the z-axis, on the proper distance between a ring of test particles in the xy-plane.

We observe that in general relativity there are gravitational wave pulses that after their passage leave test particles slightly displaced from their original position for a very long time compared to the duration of the pulse (the pulse is characterized by a nonzero curvature tensor); thus, after the propagation of such gravitational-wave pulse, the position of the test particles may represent a record of the passage of the gravitational wave. This phenomenon is sometimes called **position-coded memory** and may be a linear effect[122–125] or an effect due to nonlinear terms[126] in the Einstein field equation. Gravitational-wave pulses with a **velocity-coded memory** have been also inferred in general relativity.[127]

By applying to a plane gravitational wave the definition (2.7.18) for the pseudotensor of the gravitational field,[120] in the TT gauge (2.10.13)–(2.10.15), after some calculations one gets:[11]

$$t^{\mathrm{GW}}_{\alpha\beta} = \frac{1}{32\pi}\left\langle h^{\mathrm{TT}}_{ij,\alpha} h^{\mathrm{TT}ij}{}_{,\beta}\right\rangle \tag{2.10.20}$$

where $\langle \quad \rangle$ means average over a region of several wavelengths. In particular, applying this expression for $t^{\mathrm{GW}}_{\alpha\beta}$ to the case of the plane gravitational wave (2.10.16), traveling along the z-axis with $h_{xx} = -h_{yy} = A_+ \cos\omega(t - z)$ and $h_{xy} = h_{yx} = A_\times \cos\omega(t - z)$, we get

$$t^{\mathrm{GW}}_{zz} = t^{\mathrm{GW}}_{tt} = -t^{GW}_{tz} = \frac{1}{32\pi}\,\omega^2\left(A^2_+ + A^2_\times\right), \tag{2.10.21}$$

that is, the *energy-momentum pseudotensor for a plane gravitational wave* propagating along the z-axis, averaged over several wavelengths. From section 2.7 we find that the expression (2.10.21) represents the flux of energy carried by a plane gravitational wave propagating along the z-axis.

Finally, we give the so-called **quadrupole formula** for the outgoing flux of gravitational wave energy emitted by a system characterized by a weak gravitational field and slow motion, that is, such that its size, R, is small with respect to the reduced wavelength $\frac{\lambda}{2\pi} \equiv \lambda\!\!\!^{-}$ of the gravitational waves emitted: $R \ll \frac{\lambda}{2\pi}$. The transverse and traceless linearized metric perturbation for gravitational waves in the wave zone, $r \gg \lambda\!\!\!^{-}$, and where the background curvature can be ignored,[118,137] has been calculated to be:[5,11,51,118,136,137]

$$h^{\mathrm{TT}}_{ij} = \frac{2}{r}\frac{\partial^2}{\partial t^2}\left[I\!\!\!-^{\mathrm{TT}}_{ij}(t - r)\right] + O\left(\frac{1}{r^2}\frac{\partial}{\partial t}I\!\!\!-^{\mathrm{TT}}_{ij}\right) \tag{2.10.22}$$

where $t - r$ is the retarded time, r the distance to the source center, t the proper time of a clock at rest with respect to the source, and $I\!\!\!-^{\mathrm{TT}}_{ij}$ the transverse (with respect to the radial direction of propagation of the gravitational waves) and traceless part of the mass quadrupole moment of the source. For a source characterized by a weak gravitational field and small stresses, the symmetric **reduced quadrupole moment** (traceless), of the source mass density ρ, is given by[11]:

$$I\!\!\!-_{ij} = \int \rho(x_i x_j - \frac{1}{3}\delta_{ij} r^2) d^3x. \tag{2.10.23}$$

We can expand in powers of $\frac{1}{r}$ the Newtonian gravitational potential U, generated by this source, as a function of the reduced quadrupole moment $I\!\!\!-_{ij}$. By a suitable choice of origin at the source, we have

$$U = \frac{M}{r} + \frac{3}{2}\frac{I\!\!\!-_{ij} n^i n^j}{r^3} + O\left(\frac{1}{r^4}\right) \tag{2.10.24}$$

where $n^i \equiv \frac{x^i}{r}$. By inserting the transverse and traceless metric perturbation (2.10.22) in the expression (2.10.20) for the flux of energy carried by a gravitational wave and by integrating over a sphere of radius r, we then get the rate of gravitational-wave energy from the source crossing, in the wave zone, a sphere of radius r at time t:

$$\begin{aligned} \frac{dE}{dt} &= \int t^{0r} r^2 d\Omega = -\int t^{00} r^2 d\Omega \\ &= \frac{1}{5}\left\langle \sum_{ij}\left[\frac{\partial^3}{\partial t^3} I\!\!\!-_{ij}(t-r)\right]^2 \right\rangle \equiv \frac{1}{5}\left\langle \dddot{I\!\!\!-}_{ij}\dddot{I\!\!\!-}^{ij}\right\rangle \end{aligned} \tag{2.10.25}$$

where $d\Omega = \sin\theta d\theta d\phi$ and $\langle \quad \rangle$ means an average over several wavelengths.

From this formula for the emission of gravitational-wave energy due to the time variations of the quadrupole moment, one can calculate the time decrease of the orbital period of some binary star systems. This general relativistic theoretical calculation agrees with the observed time decrease of the orbital period of the **binary pulsar PSR 1913+1916** (see § 3.5.1).

Geons

In the 1950s one of us[71] found an interesting way to treat the concept of body in general relativity. An object can, in principle, be constructed out of gravitational radiation or electromagnetic radiation, or a mixture of the two, and may hold itself together by its own gravitational attraction. The gravitational acceleration needed to hold the radiation in a circular orbit of radius r is of the order of c^2/r. The acceleration available from the gravitational pull of a concentration of radiant energy of mass M is of the order GM/r^2. The two accelerations agree in order of magnitude when the radius r is of the order

$$r \sim GM/c^2 = (0.742 \times 10^{-28}\ \text{cm/g})M. \tag{2.10.26}$$

A collection of radiation held together in this way is called a **geon** (gravitational electromagnetic entity) and is a purely classical object. It has nothing whatsoever directly to do with the world of elementary particles. Its structure can be treated entirely within the framework of classical geometrodynamics, provided that a size is adopted for it sufficiently great that quantum effects do not come into play. Studied from a distance, such an object presents the same kind of gravitational attraction as any other mass. Moreover, it moves through space as a unit, and undergoes deflection by slowly varying fields of force just as does any other mass. Yet nowhere inside the geon is there a place where one can put a finger and say "here is mass" in the conventional sense of mass. In particular, for a geon made of pure gravitational radiation—**gravitational geon**—there is no local measure of energy, yet there is global energy. The gravitational geon owes its existence to a localized—but everywhere regular—curvature of spacetime, and to nothing more.

In brief, a geon is a collection of electromagnetic or gravitational-wave energy, or a mixture of the two, held together by its own gravitational attraction, that describes *mass without mass*.

REFERENCES CHAPTER 2

1. A. Einstein, Zur allgemeinen Relativitätstheorie, *Preuss. Akad. Wiss. Berlin, Sitzber.* 778–86 (1915).
2. A. Einstein, Zur allgemeinen Relativitätstheorie (Nachtrag), *Preuss. Akad. Wiss. Berlin, Sitzber.* 799–801 (1915).
3. A. Einstein, Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie, *Preuss. Akad. Wiss. Berlin, Sitzber.* 47:831–39 (1915).
4. A. Einstein, Die Feldgleichungen der Gravitation, *Preuss. Akad. Wiss. Berlin, Sitzber.* 844–47 (1915).
5. A. Einstein, Näherungsweise Integration der Feldgleichungen der Gravitation, *Preuss. Akad. Wiss. Berlin, Sitzber.* 688–96 (1916).
6. A. Einstein, Hamiltonsches Prinzip und allgemeine Relativitätstheorie, *Preuss. Akad. Wiss. Berlin, Sitzber.* 1111–16 (1916).
7. A. Einstein, The Foundation of the General Theory of Relativity, *Ann. Physik* 49:769–822 (1916).
8. A. Einstein, Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie, *Preuss. Akad. Wiss. Berlin, Sitzber.* 142–52 (1917).
9. A. Einstein et al., *"The Principle of Relativity": A Collection of Original Papers on the Special and General Theory of Relativity by A. Einstein, H.A. Lorentz, H. Weyl and H. Minkowski*, with notes by A. Sommerfeld, trans. W. Perret and G.B. Jeffery (Dover, New York, 1952).
10. A. Einstein, *The Meaning of Relativity*, 5th ed. (Princeton University Press, Princeton, 1955).

11. C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973); see also refs. 12–21.

12. C. Møller, *The Theory of Relativity* (Oxford University Press, London, 1972).

13. J.L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960).

14. J.A. Wheeler, *Geometrodynamics* (Academic Press, New York, 1962).

15. A.Z. Petrov, *Einstein Spaces*, trans. R.F. Kelleher (Pergamon Press, Oxford, 1969).

16. W. Rindler, *Essential Relativity: Special, General, and Cosmological* (Van Nostrand, New York, 1969).

17. L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields*, 3d rev. ed. (Addison-Wesley, Reading, MA, and Pergamon, London, 1971).

18. S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).

19. S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-time* (Cambridge University Press, Cambridge, 1973).

20. D. Kramer, H. Stephani, M. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1980).

21. C.M. Will, *Theory and Experiment in Gravitational Physics*, rev. ed. (Cambridge University Press, Cambridge, 1993).

22. G. Galilei, *Discorsi e dimostrazioni matematiche intorno a due nuove scienze* (Elzevir, Leiden, 1638); trans. H. Crew and A. de Salvio as *Dialogues Concerning Two New Sciences* (Macmillan, New York, 1914); reprint (Dover, New York, 1954); see also ref. 23.

23. G. Galilei, *Dialogo dei due massimi sistemi del mondo* (Landini, Florence, 1632); trans. S. Drake as *Galileo Galilei: Dialogue Concerning the Two Chief World Systems—Ptolemaic and Copernican* (University of California Press, Berkeley and Los Angeles, 1953).

24. I. Newton, *Philosophiae naturalis principia mathematica* (Streater, London, 1687); trans. A. Motte (1729) and revised by F. Cajori as *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World* (University of California Press, Berkeley and Los Angeles, 1934; paperback, 1962).

25. B. Bertotti and L.P. Grishchuk, The strong equivalence principle, *Class. Quantum Grav.* 7:1733–45 (1990).

26. K. Nordtvedt Jr., Equivalence principle for massive bodies, I: Phenomenology, *Phys. Rev. 169*: 1014–16 (1968).

27. K. Nordtvedt Jr., Equivalence principle for massive bodies, II: Theory, *Phys. Rev. 169*: 1017–25 (1968).

28. P.L. Bender, D.G. Currie, R.H. Dicke, D.H. Eckhardt, J.E. Faller, W.M. Kaula, J.D. Mulholland, H.H. Plotkin, S.K. Poultney, E.C. Silverberg, D.T. Wilkinson, J.G. Williams, and C.O. Alley, The lunar laser ranging experiment, *Science* 182:229–38 (1973).

29. V. Fock, *The Theory of Space, Time and Gravitation* (Moscow, 1961); 2d rev. ed., trans. N. Kemmer (Pergamon Press, Oxford, 1966).

30. J.L. Anderson, *Principles of Relativity Physics* (Academic Press, New York, 1967).

31. K.F. Gauss, Disquisitiones generales circa superficies curvas, in *Karl Friedrich Gauss Werke*, vol. 4, 217–58; trans. J.C. Morehead and A.M. Hiltebeitel as *General Investigations of Curved Surfaces of 1827 and 1825*, reprint (Raven Press, New York, 1965).
32. N.I. Lobačevskij, *Novye Načala Geometrij s Polnoj Teoriej Parallel'nyh*, Učënye Zapiski Kazanskogo Universiteta (1835–1838).
33. G.F.B. Riemann, Über die Hypothesen, welche der Geometrie zu Grunde liegen, in *Gesammelte Mathematische Werke* (1866); reprint of 2d ed., ed. H. Weber (Dover, New York, 1953); see also the translation by W.K. Clifford, *Nature* 8:14–17, 36–37 (1873).
34. G. Ricci Curbastro, Resumé de quelques travaux sur les systèmes variables de fonction, *Bull. Sc. Mathematiques* 16:167–89 (1892).
35. G. Ricci Curbastro and T. Levi-Civita, Méthods du calcul différentiel absolu et leurs applications, *Math. Ann.* 54:125–201 (1901; also published by Blanchard, Paris, 1923); see also ref. 34.
36. E.B. Christoffel, Über die Transformation der homogenen Differentialausdrücke zweiten Grades, *J. Reine Angew. Math.* (Crelle) 70:46–70 (1869).
37. A. Schild, Lectures on general relativity theory, in *Relativity Theory and Astrophysics: I. Relativity and Cosmology; II. Galactic Structure; III. Stellar Structure*, ed. J. Ehlers (American Mathematical Society, Providence, RI, 1967), 1–105.
38. T. Levi-Civita, *Lezioni di calcolo differenziale assoluto*, compiled by E. Persico (Rome, Stock, 1925); trans. M. Long as *The Absolute Differential Calculus* (Dover, New York, 1977).
39. M. Spivak, *A Comprehensive Introduction to Differential Geometry*, 2d ed. (Publish or Perish, Berkeley, 1979).
40. S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Wiley-Interscience, New York, 1963).
41. B. Schutz, *Geometrical Methods of Mathematical Physics* (Cambridge University Press, Cambridge, 1980).
42. Y. Choquet-Bruhat, and C. DeWitt-Morette, with M. Dillard-Bleick, *Analysis, Manifold and Physics* (North-Holland, Amsterdam, 1982).
43. See also C.J.S. Clarke and F. de Felice, *Relativity on Curved Manifolds* (Cambridge University Press, Cambridge, 1990).
44. J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
45. D. Hilbert, Die Grundlagen der Physik, *Königl. Gesell. Wiss. Göttingen, Nachr., Math.-Phys. Kl.* 395–407 (1915); see also *Math. Ann.* 92:1–32 (1924).
46. J.W. York, Jr., Boundary terms in the action principles of general relativity, *Found. Phys.* 16:249–57 (1986).
47. A. Palatini, Deduzione invariantiva delle equazioni gravitazionali dal principio di Hamilton, *Rend. Circ. Mat. Palermo* 43:203–12 (1919).
48. L. Bianchi, Sui simboli a quattro indici e sulla curvatura di Riemann, *Rend. della R. Acc. dei Lincei* 11:3–7 (1902).

49. According to T. Levi-Civita (ref. 38, p. 182 of Dover ed.), G. Ricci first stated these identities without proof. Later, in 1902, L. Bianchi rediscovered and published them with a proof.
50. For a derivation of the geodesic equation of motion of a compact body, see, e.g., the approximation method by A. Einstein, L. Infeld, and B. Hoffmann, The gravitational equations and the problem of motion, *Ann. Math.* 39:65–100 (1938); see also ref. 51.

51. T. Damour, The problem of motion in Newtonian and Einsteinian gravity, in *300 Years of Gravitation*, ed. S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), 128–98.
52. T. Levi-Civita, Sur l'écart géodésique, *Math. Ann.* 97:291–320 (1926).
53. I. Ciufolini, Generalized geodesic deviation equation, *Phys. Rev. D* 34:1014–17 (1986).
54. I. Ciufolini and M. Demianski, How to measure the curvature of space-time, *Phys. Rev. D* 34:1018–20 (1986).
55. K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland, Amsterdam, 1955).
56. S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press, New York, 1962).
57. W. Killing, Über die Grundlagen der Geometrie, *J. Reine Angew. Math.* 109:121–86 (1892).
58. H. Takeno, *The Theory of Spherically Symmetric Space-times*, Scientific Reports of the Research Institute for Theoretical Physics, Hiroshima Univ., No. 3 (1963).
59. K. Schwarzschild, Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie, *Sitzber. Deut. Akad. Wiss. Berlin, Königl. Math.-Phys. Tech.* 189–96 (1916).
60. G.D. Birkhoff, *Relativity and Modern Physics* (Harvard University Press, Cambridge, 1923).
61. R.P. Kerr, Gravitational field of a spinning mass as an example of algebraically special metrics, *Phys. Rev. Lett.* 11:237–38 (1963).
62. E.T. Newman, E. Couch, K. Chinnapared, A. Exton, A. Prakash, and R. Torrence, Metric of a rotating, charged mass, *J. Math. Phys.* 6:918-19 (1965).
63. R.H. Boyer and R.W. Lindquist, Maximal analytic extension of the Kerr metric, *J. Math. Phys.* 8:265–81 (1967).
64. H. Reissner, Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie, *Ann. Physik* 50:106–20 (1916).
65. G. Nordstrøm, On the energy of the gravitational field in Einstein's theory, *Proc. Kon. Ned. Akad. Wet.* 20:1238–45 (1918).
66. J. Lense and H. Thirring, Über den Einfluss der Eigenrotation der Zentralkörper auf die Bewegung der Planeten und Monde nach der Einsteinschen Gravitationstheorie, *Phys. Z.* 19:156–63 (1918).
67. H. Thirring and J. Lense, trans. B. Mashhoon, F.W. Hehl, and D.S. Theiss, as On the gravitational effects of rotating masses: The Thirring-Lense papers, *Gen. Rel. Grav.* 16:711–50 (1984); see also refs. 63 and 64.

68. H. Thirring, Über die Wirkung rotierender ferner Massen in der Einsteinschen Gravitationstheorie, *Phys. Z.* 19:33–39 (1918).

69. H. Thirring, Berichtigung zu meiner Arbeit: Über die Wirkung rotierender ferner Massen in der Einsteinschen Gravitationstheorie, *Phys. Z.* 22:29–30 (1921).

70. D. Ivanenko and A. Sokolov, *Klassische Feldtheorie*, (Akademie Verlag, Berlin, 1953); section on 2 graviton $\rightarrow$ electron pair.

71. J.A. Wheeler, Geons, *Phys. Rev.* 97:511–36 (1955)

72. R. Arnowitt, S. Deser, and C.W. Misner, The dynamics of general relativity, in *Gravitation: An Introduction to Current Research*, ed. L. Witten (Wiley, New York, 1962), 227–65.

73. J.W. York, Jr., Energy and momentum of the gravitational field, in *Essays in General Relativity*, ed. F.J. Tipler (Academic Press, New York, 1980), 39–58.

74. R. Penrose, Quasi-local mass and angular momentum in general relativity, *Proc. Roy. Soc. Lond. A* 381:53–63 (1982).

75. R.M. Schoen and S.-T. Yau, On the proof of the positive mass conjecture in general relativity, *Comm. Math. Phys.* (*Germany*) 65:45–76 (1979).

76. R.M. Schoen and S.-T. Yau, Proof of the positive-action conjecture in quantum relativity, *Phys. Rev. Lett.* 42:547–48 (1979).

77. R.M. Schoen and S.-T. Yau, Positivity of the total mass of a general space-time, *Phys. Rev. Lett.* 43:1457–59 (1979).

78. R.M. Schoen and S.-T. Yau, The energy and the linear momentum of space-times in general relativity, *Comm. Math. Phys.* 79:49–51 (1981).

79. R.M. Schoen and S.-T. Yau, Proof of the positive mass theorem. II, *Comm. Math. Phys.* 79:231–60 (1981).

80. S. Deser and C. Teitelboim, Supergravity has positive energy, *Phys. Rev. Lett.* 39:249–52 (1977).

81. E. Witten, A new proof of the positive energy theorem, *Comm. Math. Phys.* 80:381–402 (1981).

82. E. Witten, Positive energy and Kaluza-Klein theory, in *10th Int. Conf. on General Relativity and Gravitation*, ed. B. Bertotti (Reidel, Dordrecht, 1984), 185–97.

83. Y. Choquet-Bruhat, Positive-energy theorems, in *Relativité, Groupes et Topologie II*, ed. B.S. DeWitt and R. Stora (Elsevier, Amsterdam, 1984), 739–85.

84. J.A. Wheeler, *Physics and Austerity, Law Without Law* (Anhui Science and Technology Publications, Anhui, China, 1982).

85. J.A. Wheeler, Gravitation as geometry. II, in *Gravitation and Relativity*, ed. H.-Y. Chiu and W.F. Hoffmann (Benjamin, New York, 1964), 65–89.

86. E.H. Spanier, *Algebraic Topology* (McGraw-Hill, New York, 1966).

87. M. Göckeler and T. Schücker, *Differential Geometry, Gauge Theories, and Gravity* (Cambridge University Press, Cambridge, 1987).

88. W.A. Miller, The geometrodynamic content of the Regge equations as illuminated by the boundary of a boundary principle, in *Between QUANTUM and COSMOS: Studies and Essays in Honor of John Archibald Wheeler*, ed. W.H. Zurek, A. van

der Merwe, and W. A. Miller (Princeton University Press, Princeton, 1988), 201–27; reprinted from *Foundations of Physics* (1986); see also ref. 89.

89. A. Kheyfets, The boundary of a boundary principle: A unified approach, in *Between QUANTUM and COSMOS: Studies and Essays in Honor of John Archibald Wheeler*, ed. W.H. Zurek, A. van der Merwe, and W.A. Miller (Princeton University Press, Princeton, 1988), 284–98.
90. F.W. Hehl and J.D. McCrea, Bianchi identities and the automatic conservation of energy-momentum and angular momentum in general-relativistic field theories, in *Between QUANTUM and COSMOS: Studies and Essays in Honor of John Archibald Wheeler*, ed. W.H. Zurek, A. van der Merwe, and W.A. Miller (Princeton University Press, Princeton, 1988), 256–82.
91. É. Cartan, *Leçon sur la géométrie des espaces de Riemann*, 2d ed. (Gauthier-Villars, Paris, 1963).
92. J.A. Wheeler, Our universe: The known and the unknown, address before the American Association for the Advancement of Science, New York, 29 December 1967, in *Am. Scholar* 37:248–74 (1968) and *Am. Sci.* 56:1–20 (1968).
93. J.R. Oppenheimer and H. Snyder, On continued gravitational collapse, *Phys. Rev.* 56:455–59 (1939).
94. J.A. Wheeler, The superdense star and the critical nucleon number, in *Gravitation and Relativity*, ed. H.-Y. Chiu and W.F. Hoffman (Benjamin, New York, 1964), 195–230.
95. J.A. Wheeler, Geometrodynamics and the issue of the final state, in *Relativity Groups and Topology*, ed. C. DeWitt and B. DeWitt (Gordon and Breach, New York, 1964), 315–520.
96. B.K. Harrison, K.S. Thorne, M. Wakano, and J.A. Wheeler, *Gravitation Theory and Gravitational Collapse* (University of Chicago Press, Chicago, 1965).
97. R. Ruffini and J.A. Wheeler, Introducing the black hole, *Phys. Today* 24:30–36 (1971).
98. R. Penrose, Black holes, *Sci. Am.* 226:38–46 (1972).
99. C. DeWitt and B. DeWitt, eds., *Black Holes*, Les Houches Summer School, Grenoble 1972 (Gordon and Breach, New York, 1973).
100. J.A. Wheeler, The black hole, in *Astrophysics and Gravitation: Proc. 16th Solvay Conf. on Physics*, Brussels, September 1973, ed. R. Debever (Éditions de l'Université de Bruxelles, Brussels, 1974), 279–316.
101. K.S. Thorne, The search for black holes, *Sci. Am.* 231:32–43 (1974).
102. A.S. Eddington, A comparison of Whitehead's and Einstein's formulas, *Nature* 113:192 (1924).
103. D. Finkelstein, Past-future asymmetry of the gravitational field of a point particle, *Phys. Rev.* 110:965–67 (1958).
104. M.D. Kruskal, Maximal extension of Schwarzschild metric, *Phys. Rev.* 119:1743–45 (1960).
105. G. Szekeres, On the singularities of a Riemannian manifold, *Publ. Mat. Debrecen* 7:285–301 (1960).

106. A. Einstein and N. Rosen, The particle problem in the general theory of relativity, *Phys. Rev.* 48:73–77 (1935).
107. R.W. Fuller and J.A. Wheeler, Causality and multiply-connected space-time, *Phys. Rev.* 128:919–29 (1962).
108. M.S. Morris, K.S. Thorne, and U. Yurtsever, Wormholes, time machines, and the weak energy condition, *Phys. Rev. Lett.* 61:1446–49 (1988).
109. S.W. Hawking, The chronology protection conjecture, in *Proc. 6th Marcel Grossmann Meeting on General Relativity*, Kyoto, June 1991, ed. H. Sato and T. Nakamura (World Scientific, Singapore), 3–13.
110. S.W. Hawking, Particle creation by black holes, *Comm. Math. Phys.* 43:199–220 (1975).
111. R. Penrose, Gravitational collapse: The role of general relativity, *Riv. Nuovo Cimento* 1:252–76 (1969).
112. R. Penrose, Gravitational collapse and space-time singularities, *Phys. Rev. Lett.* 14:57–59 (1965).
113. S.W. Hawking, The occurrence of singularities in cosmology. III. Causality and singularities, *Proc. Roy. Soc. Lond. A* 300:187–201 (1967).
114. R.P. Geroch, Singularities, in *Relativity*, ed. S. Fickler, M. Carmeli, and L. Witten (Plenum Press, New York, 1970), 259–91.
115. S.W. Hawking and R. Penrose, The singularities of gravitational collapse and cosmology, *Proc. Roy. Soc. Lond. A* 314:529–48 (1970).
116. B.S. DeWitt, Quantum gravity: The new synthesis, in *General Relativity: An Einstein Centenary Survey*, ed. S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979), 680–893.
117. J. Weber, *General Relativity and Gravitational Waves* (Wiley-Interscience, New York, 1961).
118. K.S. Thorne, *Gravitational Radiation* (Cambridge University Press, Cambridge), to be published.
119. R.A. Isaacson, Gravitational radiation in the limit of high frequency. I. The linear approximation and geometrical optics, *Phys. Rev.* 166:1263–71 (1968).
120. R.A. Isaacson, Gravitational radiation in the limit of high frequency. II. Nonlinear terms and the effective stress tensor, *Phys. Rev.* 166:1272–80 (1968).
121. L.P. Grishchuk and A.G. Polnarev, Gravitational waves and their interaction with matter and fields, in *General Relativity and Gravitation, One Hundred Years After the Birth of Albert Einstein*, ed. A. Held (Plenum Press, New York and London, 1980), 393–434.
122. Ya.B. Zel'dovich and A.G. Polnarev, Radiation of gravitational waves by a cluster of superdense stars, *Sov. Astron. AJ* (*USA*) 18:17–23 (1974); *Astron. Zh.* (*USSR*) 51:30–40 (1974).
123. S.J. Kovács, Jr., and K.S. Thorne, The generation of gravitational waves. IV. Bremsstrahlung, *Astrophys. J.* 224:62–85 (1978).
124. V.B. Braginsky and L.P. Grishchuk, Kinematic resonance and memory effect in free-mass gravitational antennas, *Sov. Phys.—JETP* 62:427–30 (1985); *Zh. Eksp. Teor. Fiz.* 89:744–50 (1985).

125. V.B. Braginsky and K.S. Thorne, Gravitational-wave bursts with memory and experimental prospects, *Nature* 327:123–25 (1987).

126. D. Christodoulou, Nonlinear nature of gravitation and gravitational-wave experiments, *Phys. Rev. Lett.* 67:1486–89 (1991).

127. L.P. Grishchuk and A.G. Polnarev, Gravitational wave pulses with velocity-coded memory, *Sov. Phys.—JETP* 69:653–57 (1990); *Zh. Eksp. Theor. Fiz.* 96:1153–60 (1989).

128. R.P. Feynman, The principle of least action in quantum mechanics, Ph.D. dissertation, Princeton University (1942).

129. R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

130. M. Oda, P. Gorenstein, H. Gursky, E. Kellogg, E. Schreier, H. Tananbaum, and R. Giacconi, X-ray pulsations from Cygnus X-1 observed from UHURU, *Astrophys. J. Lett.* 166:L1–L7 (1971).

131. J. Ehlers, Contributions to the relativistic mechanics of continuous media, *Akad. Wiss. Lit. Mainz Abh. Math. Nat. Kl.* 793–837 (1961).

132. J. Ehlers, General relativity and kinetic theory, in *General Relativity and Cosmology, Proc. Course 47 of the Int. School of Physics "Enrico Fermi"*, ed. R.K. Sachs (Academic Press, New York, 1971), 1–70.

133. H. Bondi, Plane gravitational waves in general relativity, *Nature* 179:1072–73 (1957).

134. H. Bondi and W.H. McCrea, Energy transfer by gravitation in Newtonian theory, *Proc. Cambridge Phil. Soc.* 56:410–13 (1960).

135. H. Bondi, Some special solutions of the Einstein equations, in *Lectures on General Relativity*, Brandeis 1964 Summer Institute on Theoretical Physics, vol. 1, ed. A. Trautmann, F.A.E. Pirani, and H. Bondi (Prentice-Hall, Englewood Cliffs, NJ, 1965), 375–459.

136. A. Einstein, Über Gravitationswellen, *Preuss. Akad. Wiss. Berlin, Sitzber*. 8, 154–167 (1918).

137. K.S. Thorne, Multipole expansions of gravitational radiation, *Rev. Mod. Phys.* 52:299–339 (1980).

138. S.-W. Kim, Quantum effects of Lorentzian wormhole, *Proc. 6th Marcel Grossmann Meeting on General Relativity*, Kyoto, June 1991, ed. H. Sato and T. Nakamura (World Scientific, Singapore, 1992), 501–3.

139. U.P. Frolov and I.D. Novikov, Physical effects in wormholes and time machines, *Phys. Rev. D* 42:1057–65 (1990).

140. J. Friedman, M.S. Morris, I.D. Novikov, F. Echeverria, G. Klinkhammer, K.S. Thorne, and U. Yurtsever, Cauchy problem in spacetimes with closed timelike curves, *Phys. Rev. D* 42:1915–30 (1990).

141. I.D. Novikov, Time machine and self-consistent evolution in problems with self-interaction, *Phys. Rev. D* 45:1989–94 (1992).

142. J. Casares, P.A. Charles, and T. Naylor, A 6.5-day periodicity in the recurrent nova V404 Cygni implying the presence of a black hole, *Nature* 355:614–17 (1992).

143. R.A. Remillard, J.E. McClintock, and C.D. Bailyn, Evidence for a black hole in the X-ray binary Nova Muscae 1991, *Astrophys. J. Lett*. 399:L145–149 (1992).

144. Ya.B. Zel'dovich and I.D. Novikov, *Relativistic Astrophysics*, vol. 1, *Stars and Relativity* (University of Chicago Press, Chicago, 1971).

145. A. Ashtekar, *Lectures on Non-Perturbative Canonical Gravity* (World Scientific, Singapore, 1991); see also ref. 146.

146. C. Rovelli, Ashtekar formulation of general relativity and loop-space non-perturbative quantum gravity: a report, *Class. Quantum Grav*. 8:1613–75 (1991).

147. H.C. Ford et al., Narrowland *HST* images of M87: Evidence for a disk of ionized gas around a massive black hole, *Astrophys. J. Lett*. 435:L27–30 (1994). See also: R.J. Harms et al., ibid:L35–38.