A curvature-based derivation of the Schwarzschild metric

Mariano Santander, a) Luis M. Nieto, b) and Nicolás A. Cordero c)
Departamento de Física Teórica, Universidad de Valladolid, E-47011 Valladolid, Spain

(Received 19 February 1997; accepted 12 June 1997)

We give a derivation of Schwarzschild and other spherically symmetric solutions in general relativity which starts by computing the curvatures directly from Einstein’s equations and Bianchi identities, and then derives the metric. This transforms the “plausibility” argument for the Schwarzschild metric given by Berry in his book Principles of Cosmology and Gravitation into a complete derivation, which puts the stress from the beginning on the curvatures, the quantities which actually correspond to the gravitational field. © 1997 American Association of Physics Teachers.

I. INTRODUCTION

In the excellent introductory textbook Principles of Gravitation and Cosmology,1 Berry gives an argument to make plausible the Schwarzschild metric, avoiding completely the usual tensor machinery. The reasoning goes as follows: let us consider a static gravitational field in empty space, around a gravitational source of mass M with spherical symmetry. According to Einstein’s theory, one can expect that three space dimensions will be curved, and the sectional curvature along any two-plane direction will be curved, and the sectional curvature along any two-plane. If we do not care about the precise meaning of the coordinate r, simple dimensional analysis determines the r dependence of any spatial sectional curvature as qGM/(c^2r^3), where q is a numerical coefficient, because GM/(c^2r^3) is the unique quantity with dimensions L^-2 built out of G,M,c,r. In particular, along any diametral two-plane, Berry chooses the simplest negative value q = -1.

On the other hand, the most general static metric with spherical spatial symmetry, and radial coordinate chosen so that the area of the sphere at radial coordinate r is exactly 4πr^2, depends on two unknown functions of r:

\[ d\tau^2 = e(r)dt^2 - \frac{1}{e^2}f(r)dr^2 - \frac{r^2}{c^2}(d\theta^2 + \sin^2\theta \, d\phi^2). \]  

(1.1)

The standard argument on the gravitational redshift, based on the equivalence principle, relates (in suitable coordinate systems) the metric element e(r) = g_{tt}(r) to the Newtonian gravitational potential \( \varphi \) as \( g_{tt}(r) = (1 + 2\varphi/c^2) \). Let us take the Newtonian expression \( \varphi = -GM/r \) for \( \varphi \), assuming that the radial coordinate r is the “correct” replacement for the Newtonian radial distance, and let us forget about the ~ sign. In other words, let us make the ansatz \( e(r) = 1 - 2GM/(c^2r^3) \). Now the only unknown is the function f(r).

By spherical symmetry, the diametral surface \( \theta = \pi/2, \phi = \text{const} \) is totally geodesic, so the sectional curvature along any diametral two-plane can be computed as the curvature of the spatial metric restricted to this surface:

\[ dl^2 = -c^2d\tau^2|_{\theta = \pi/2, \phi = \text{const}} = f(r)dr^2 + r^2d\phi^2, \]

which using the standard formula [see (2.3) below] is given by the simple expression

\[
\frac{1}{2rf^2(r)} \frac{df(r)}{dr}.
\]  

(1.2)

Equating this to Berry’s choice, \(-GM/(c^2r^3)\), we get a differential equation whose solution with the good behavior \( f(r) \to 1 \) at spatial infinity is

\[
f(r) = \left(1 - \frac{2GM}{c^2r^3}\right)^{-1}.
\]  

(1.3)

Therefore we obtain the Schwarzschild metric in the so-called curvature coordinates:

\[
d\tau^2 = \left(1 - \frac{2GM}{c^2r} \right)dt^2 - \frac{1}{c^2} \left(1 - \frac{2GM}{c^2r}\right)^{-1}dr^2 - \frac{r^2}{c^2}(d\theta^2 + \sin^2\theta \, d\phi^2).
\]  

(1.4)

The argument, which can also be found in other introductory textbooks,2 is undoubtedly appealing. However, it relies on some unwarranted coincidences and on educated guesswork. Indeed, the great simplicity of the Schwarzschild metric allows simple but incorrect arguments to have the appearance of a correct derivation (see the explicit remarks by Rindler when criticizing the Lenz–Schiff argument3). In this example, the criticisms are:

(i) The relation \( g_{tt} = (1 + 2\varphi/c^2) \) is only approximate although the approximation cancels out with the choice made when the expression for the Newtonian potential is taken as \( \varphi = -GM/r \) in terms of a coordinate r, which is not now the distance to the origin.

(ii) The expression for the diametral sectional curvature determined by dimensional analysis is but the simplest possibility.

(iii) The choice \( q = -1 \) is unfounded.

(iv) The implicit identification of the unspecified radial coordinate giving the \( GM/(c^2r^3) \) dependence with the actual radial coordinate r in the metric (1.1) leads to the right result, but the true reason for this surprising outcome is hidden. In fact, it is not clear beforehand why the coordinate r (which is defined in terms of area of spheres, and is not the radial distance) should play such a distinguished role.

Berry argues for the choice \( q = -1 \) by analogy with a rubber-sheet curving negatively under a weight, but he admits this is rather unconvincing. Should a similar reasoning be applied to a two-plane transversal to the radial direction,
Robertson–Walker metric

spherically symmetric energy–momentum distribution, and

loose end. Section IV is devoted to the case of a source with

following the ideas outlined in Sec. I, but completing every

Section III gives the derivation of the Schwarzschild metric

curvatures, and states Einstein’s equations in terms of them.

introduces some notation, recalls some properties of sectional

tamples the Reissner–Nordstro¨ m metric and the cosmological

Newtonian ansatz.

the one used to calculate

f

r

for example, in the book by Spivak,5 Vol. II, p. 109. For the special case of

an orthogonal coordinate system u, v, where the metric is given by

\( dl^2 = E(u,v) du^2 + G(u,v) dv^2 \), this formula for the curvature is given in some texts (see, e.g., Appendix B in Ref. 1 for a down-to-earth derivation):

\[
K = \frac{1}{2EG} \left[ \frac{\partial^2 E}{\partial u^2} \frac{\partial^2 G}{\partial u \partial v} - \frac{1}{2E} \left( \frac{\partial E}{\partial u} \frac{\partial G}{\partial v} + \left( \frac{\partial E}{\partial v} \right)^2 \right) \right] \\
- \frac{1}{2G} \left( \frac{\partial E}{\partial u} \frac{\partial G}{\partial v} + \left( \frac{\partial G}{\partial u} \right)^2 \right). 
\]

(2.2)

The Riemannian character of the metric, with both functions E and G positive, allows one to write this formula in the still simpler form:

\[
K = -\frac{1}{\sqrt{EG}} \left[ \frac{\partial}{\partial u} \left( \frac{1}{\sqrt{E}} \frac{\partial G}{\partial u} \right) + \frac{\partial}{\partial v} \left( \frac{1}{\sqrt{G}} \frac{\partial E}{\partial v} \right) \right]. 
\]

(2.3)

If we now consider a two-dimensional locally Minkowskian space, where the indefinite pseudo-Riemannian metric is denoted \( d\tau^2 \), the outstanding new fact is that there are timelike and spacelike curves. In the kinematical standard interpretation, the length of a timelike curve is the proper time measured by an ideal clock which follows this world line, while the (space) length of a spacelike curve should be measured by the spatial metric \(-c^{-2}d\tau^2\), which is positive definite when restricted to spacelike curves. There is a formula similar to (2.1) governing the acceleration (relative to the proper time along a temporal geodesic) of the spatial separation \( \delta(\tau) \) between two neighboring timelike geodesics:

\[
\frac{d^2(\delta(\tau))}{d\tau^2} = -K \delta(\tau), 
\]

(2.4)

and the expression for the curvature \( K \) in terms of the metric tensor is the same as already referred to; in particular, in an orthogonal coordinate system \( u,v \) with a metric

\[
d\tau^2 = E(u,v) du^2 + G(u,v) dv^2 \]

(2.5)

the curvature \( K \) is given by the same expression (2.2). However in this pseudo-Riemannian case, either \( E \) or \( G \) will be negative, and the algebraic reduction leading to the formula
determine the second derivative of the separation between two geodesics having direction (1) as a hinge. Their sectional curvatures are related through Eq. (2.7).

(2.3) now requires due care to signs, and leads to the expression for the (Gaussian) curvature $K$:

$$
K = -\frac{1}{\sqrt{|E||G|}} \left\{ \left( \text{sg} E \right) \frac{\partial}{\partial u} \left( \frac{1}{\sqrt{|E|}} \frac{\partial \sqrt{|G|}}{\partial u} \right) + (\text{sg} G) \frac{\partial}{\partial v} \left( \frac{1}{\sqrt{|G|}} \frac{\partial \sqrt{|E|}}{\partial v} \right) \right\},
$$

(2.6)

where $\text{sg} E$ is the sign of $E$. This formula can be found in the textbook by O'Neill; as expected, it includes (2.3) when both $E, G$ are positive.

We note the use of two slightly different symbols, $K$ for the \textquoteleft\textquoteleft space\textquoteright\textquoteleft curvature of a properly Riemannian metric, and $K$ for the curvature of a pseudo-Riemannian metric. This anticipates the different role these curvatures will play in Einstein’s theory.

**B. Sectional curvatures in any Riemannian or pseudo-Riemannian space**

For a space of higher dimension and Riemannian (positive definite) metric, the curvature produces the same kind of behavior described by (2.1), but with the important new property that $K$ depends now on the two-plane spanned by the tangent vectors to the two geodesics; these are called sectional curvatures. For instance, in dimension $n = 3$, if we choose three space directions (1), (2), (3) mutually perpendicular at a point $p$, there are three such sectional curvatures $K_{(12)}$ (respectively $K_{(23)}$, $K_{(31)}$) along the two-planes spanned by the vectors tangent to directions (1), (2) [respectively, (2), (3) or (3), (1)] [see Fig. 1(a)]. These curvatures determine the second derivative of the separation between any two nearby geodesics with tangent vectors at $p$ contained in the two-plane (12) [respectively (23) and (31)]. In any Riemann space the sum of these three sectional curvatures along three mutually orthogonal two-planes does not depend on which three two-planes have been chosen. This can be explicitly seen as a consequence of the equations which describe how sectional curvatures at a given point depend on the two-plane. For instance, if (3') denotes the direction contained in the two-plane (23) and making an angle $\alpha$ with (3), the sectional curvature along the bundle of two-planes (13') which has the direction (1) as a “hinge” [see Fig. 1(b)] depends on the angle $\alpha$ in the following rigid way, which is reminiscent of Euler’s equation for normal curvatures in surfaces:

$$
K_{(3'1)} = K_{(31)} \cos^2 \alpha - 2R^3_{112} \cos \alpha + K_{(12)},
$$

(2.7)

where $K_{(31)}$ and $K_{12}$ are the sectional curvatures along the two-planes (31) and (12), and the coefficient of the crossed term $R^3_{112}$ is an element of the Riemann tensor at the point $p$ in suitable coordinates, with the three different indices corresponding to the three directions involved. There are similar expressions for two-planes in a bundle having (2) [or (3) or any other direction] as a hinge. This formula shows that the sum of sectional curvatures along two orthogonal two-planes in any such bundle does not depend on the particular two-planes chosen; the extension to three mutually orthogonal two-planes goes along the same lines.

*Mutatis mutandis*, this applies in a locally Minkowskian space; in the $(1 + 3)$ case, the spatial separation between two neighboring timelike geodesics which separate in the three space orthogonal to the time direction (0) along the space direction (1) [respectively, (2) and (3)] is governed by an equation like (2.4), with a value for the constant which is called sectional curvature $K_{(01)}$ (respectively, $K_{(02)}$, $K_{(03)}$). Again, the sum of these three temporal sectional curvatures does not depend on the choice of the three orthogonal space directions.

We remark that these curvatures are geometrical quantities, and are independent of the coordinate system; the use of a notation with parentheses tries to underline this fact and to avoid confusion with the standard tensorial language; of course, $K_{(12)} = K_{(21)}$, etc.

**C. Einstein’s equations in terms of sectional curvatures**

In Einstein’s theory, gravitation can be described as curvature: space–time will have “temporal” curvature, as in (2.4), and ordinary three space will also be curved, as in (2.1). By choosing a pseudo-Riemannian space as the mathematical model, all the properties of curvature in such a space are automatically incorporated. This includes the laws (2.7), their analogues for another bundle of two-planes with either spacelike or timelike “hinges,” and Bianchi identities. The use of slightly different symbols for the space curvature $K$ and the temporal curvature $K$ that we have done in the previous expressions, anticipating their interpretation in Einstein’s theory of gravitation, is deliberate and tries to underline they correspond to properties which are perceived rather differently from a 1 + 3 viewpoint: $K$ has dimension $L^{-2}$, while the dimension of $K$ is $T^{-2}$. The naive argument given in Sec. I allows one to produce a quantity $GM/r^3$ with di-
tions. For example, Feynman’s textbooks toward tensor formulation. There are some ex-

or, and provides another reason to distinguish K from K [see (2.10) below; a brief resume of Cartan’s formulation of Newtonian gravitation can be found in Tipler’s paper].

In addition to the relations satisfied by curvature in any pseudo-Riemannian space, Einstein’s equations for the gravitational field are further expressions relating curvature to the physical distribution of energy. These equations are usually written in tensor language as $G_{\mu\nu} = (8\pi G/c^2)T_{\mu\nu}$, where $G_{\mu\nu}$ is the Einstein tensor of space–time, and $T_{\mu\nu}$ is the energy–momentum tensor of the sources creating the gravitational field. The physical meaning of these equations is, however, more easily grasped when they are written in a purely geometrical form, which relates sectional curvatures of space–time to the proper energy and/or momentum densities and/or fluxes. This geometrical form is less familiar than the tensorial form, yet it goes a long way to the physical understanding of Einstein’s theory. In fact, the complete content of Einstein’s theory is embodied in the single equation:

$$K_{(12)} + K_{(23)} + K_{(31)} = \frac{8\pi G}{c^4} W^{(0)}.$$  \hspace{1cm} (2.8)

where $W^{(0)}$ is the proper energy density at the considered space–time point. This relation is assumed to be valid irrespective of the state of motion of the observer (as long as the one keeps measuring energy density in the proper comoving frame, and sectional curvatures in the proper space three).

While it is true that the form (2.8) of Einstein’s equations is unfamiliar, this is only due to the shift of emphasis in most textbooks toward tensor formulation. There are some exceptions. For example, Feynman’s Lectures on Physics’ states Einstein’s equations in a form directly equivalent to (2.8), and more details on this form are given in Lecture 11 of Feynman’s Lectures on Gravitation.

To compare (2.8) with the standard tensor form $G_{\mu\nu} = (8\pi G/c^2)T_{\mu\nu}$ of Einstein’s equations there are two stages. First, the requirement of a single equation for any observer implies a set of equations which should hold true for each observer; this is similar to the situation for electromagnetism, where two basic equations (Gauss’s law and the absence of magnetic monopoles), assumed valid for any observer, contain the complete Maxwell’s theory. When this is done for (2.8), one gets a set of equations which turns out to be equivalent to the complete system of ten Einstein’s equations. We write here only some equations derived from this set, which give simple relations between sectional curvatures, the energy density, and the diagonal components of the stress at the point under consideration. If any three orthogonal space directions are chosen at any point, and the three indices $ijk$ are in cyclic order, these relations are:

$$K_{(12)} + K_{(23)} + K_{(31)} = \frac{8\pi G}{c^4} W^{(0)},$$  \hspace{1cm} (2.9a)

$$K_{(01)} + K_{(02)} + K_{(03)} = \frac{4\pi G}{c^2} \left\{ W^{(0)} + (\sigma^{(11)} + \sigma^{(22)} + \sigma^{(33)}) \right\}.$$  \hspace{1cm} (2.9b)

We recall that each diagonal component of stress $\sigma^{(kk)}$ is the flux of the $k$ component of momentum transferred per unit time across the unit of area along a $(ij)$ two-plane (either a pressure or a tension).

The second translation stage between the usual tensor formulation of Einstein’s equations and the purely geometric form (2.9), is provided by the link between Ricci and Einstein tensors with the sectional curvatures of any pseudo-Riemannian space. In a coordinate system which is orthogonal and reduces the metric at the point $p$ to the Minkowskian form (this will be indicated as usual by indices with a caret), and denoting by $K$ the sectional curvatures of the metric $dT^2$ and by $k$, the ones corresponding to the metric $d\tilde{t}^2 = -c^2 d\tau^2$, which are related by $K = (-1/c^2)K$, then:

(i) $R_{00}^i$ equals the sum of the sectional curvatures $K$ along three mutually orthogonal two-planes all of which contain the $(0)$ direction:

$$R_{00}^i = K_{(01)} + K_{(02)} + K_{(03)}.$$

(ii) $R_{ii}^j$ (without the sum in $i$) equals to $-1/c^2$ times the sum of the sectional curvatures $K$ along three mutually orthogonal two-planes all of which contain the $(i)$ direction:

$$R_{ii}^j = (-1/c^2) (K_{(01)} + K_{(12)} + K_{(31)}) = (-1/c^2) K_{(01)} + K_{(12)} + K_{(31)}; \text{ etc.}$$

(iii) $-G_{00}^i$ equals to the sum of the sectional curvatures $K$ along three mutually perpendicular two-planes orthogonal to the $(0)$ direction:

$$-G_{00}^i = (K_{(12)} + K_{(23)} + K_{(31)})$$

$$= -c^2 (K_{(12)} + K_{(23)} + K_{(31)}).$$

(iv) $-G_{ii}^j$ equals to $-1/c^2$ times the sum of the sectional curvatures $K$ along three mutually orthogonal two-planes perpendicular to the $(i)$ direction:

$$-G_{ii}^j = (-1/c^2) (K_{(01)} + K_{(02)} + K_{(03)} + K_{(23)}) = (-1/c^2) K_{(01)} + (-1/c^2) K_{(02)} + K_{(03)} + K_{(23)}; \text{ etc.}$$

As far as the diagonal components of the energy–momentum tensor are concerned, in a coordinate system which is locally Minkowskian at the point $p$, $T_{00}^i$ equals the proper energy density $W^{(0)}$, and the spatial diagonal components $T_{kk}$ are equal to $1/c^2$ times the diagonal component $\sigma^{(kk)}$ of the stress tensor. It is now a simple exercise to get Eqs. (2.9) by explicitly translating the “diagonal” equations $G_{\mu\nu} = (8\pi G/c^2)T_{\mu\nu}$.
We close this section by remarking that the equations of the Newtonian theory of gravitation, when written in a geometric form, are directly the \( c \to \infty \) limit of the equations (2.9). If \( W^{(0)} \) is due to a mass density, \( W^{(0)} = \rho c^2 \), these equations reduce naturally to
\[
K_{(i)} = 0, \quad K_{(01)} + K_{(02)} + K_{(03)} = 4 \pi G \rho, \tag{2.10}
\]
that is, flat three space and an equation which turns out to be exactly the Poisson equation for the gravitational potential because in this theory the curvatures are related to the Newtonian potential as \( K_{(0)} = (\partial^2 \phi/\partial x^i \partial x^j)^2 \).

### III. DERIVATION OF THE SCHWARZSCHILD SOLUTION

#### A. Sectional curvatures in empty space around a nonrotating spherically symmetric mass

We assume a static gravitational field, and we shall use a \( t \) coordinate in such a way that metric, curvatures, etc., will be \( t \) independent. Spherical spatial symmetry implies the following:

(a) It should be possible to introduce as spatial coordinates two angles \( \theta, \phi \), and a radial coordinate \( r \), still unspecified, so that the spatial surfaces \( \{ t = \text{const}, \ r = \text{const} \} \) will be isometric to ordinary spheres \( S^2 \) parametrized by the two angles \( (\theta, \phi) \) as usual \( dr^2 = f(r)^2 (d\theta^2 + \sin^2 \theta d\phi^2) \). The general expression for a static, spherically symmetric metric in terms of these coordinates is (1.1).

(b) All sectional curvatures should depend only on the coordinate \( r \).

(c) The spatial sectional curvatures at a given point, along any two-plane spatial direction (in the three-space orthogonal to the time direction) can be expressed in terms of the values of two sectional curvatures which are distinguished for symmetry reasons: the diametral sectional curvatures along two-planes containing the radial direction, which should all be equal
\[
K_{(\phi r)} = K_{(\theta r)} = D(r), \tag{3.1a}
\]
and the transversal sectional curvature along the two-plane direction tangent to the spheres \( \{ r = \text{const}, \ t = \text{const} \} \)
\[
K_{(\theta \phi)} : = T(r). \tag{3.1b}
\]

The sectional curvature along any spatial two-plane direction whose normal makes an angle \( \alpha \) with the radial direction, after (2.7), is \( T(r) \cos^2 \alpha + D(r) \sin^2 \alpha \); note that \( D(r) \) is the diametral sectional curvature in Berry’s argument.

(d) The temporal sectional curvatures at a given point, along any two-plane direction containing the time direction can be expressed again in terms of two temporal sectional curvatures which are distinguished by symmetry reasons: the temporal–radial sectional curvature along two-planes containing the time and the radial direction, and the temporal–transversal sectional curvatures along two-planes containing the time and a spatial direction transverse to the radial one:
\[
K_{(\tau r)} = K_{(\tau \phi)} = K_{(\phi r)} = 0, \tag{3.2}
\]

(e) Finally, and as far as the external sources are concerned, the assumption of spherical symmetry means that only the proper energy density and the proper pressures across the spatial two-planes can be different from zero, and again the two diametral pressures (across a diametral two-plane) should be equal. Hence any spherically symmetric source is completely described by three functions of the radial coordinate:
\[
W^{(0)}(r), \quad a^{(r)}(r), \quad a^{(\phi \phi)}(r), \tag{3.3}
\]
which are, respectively, the proper energy density, the proper pressure on a transversal spatial two-plane, and the proper pressure on a diametral two-plane. In empty space, all these quantities are zero.

We remark again that the notation with parentheses underlines that we are dealing with the proper values of the energy density or the pressures, as measured by an observer at rest at the given space–time point. These values should not be understood as the ordinary components of the energy–momentum tensor in the coordinate system \( \{ t, r, \theta, \phi \} \).

Our aim is to derive now the space–time metric in empty space around a nonrotating spherically symmetric mass. In the absence of energy–momentum, for any choice of three mutually orthogonal directions in three–space at any point, Einstein’s equations (2.9) give the following relations between sectional curvatures:
\[
K_{(12)} + K_{(23)} + K_{(31)} = 0, \tag{3.4a}
\]
\[
K_{(23)} = -\frac{1}{c^2} K_{(01)}, \quad K_{(31)} = -\frac{1}{c^2} K_{(02)} , \tag{3.4b}
\]
\[
K_{(12)} = -\frac{1}{c^2} K_{(03)} , \tag{3.4c}
\]

If we take (1) as the radial direction, and (2) and (3) as the directions determined by the coordinates \( \theta, \phi \) (both of which are orthogonal to the radial direction), these equations give simple relations between the principal spatial and temporal sectional curvatures. First, spherical symmetry (3.1)–(3.2) and the Einstein equation (3.4b) imply that all these curvatures can be expressed in terms of the two functions \( D(r) \) and \( T(r) \) as
\[
K_{(\phi r)} = K_{(\theta r)} = D(r), \quad K_{(\theta \phi)} = T(r). \tag{3.5a}
\]
\[
K_{(\phi r)} = K_{(\theta r)} = -c^2 D(r), \quad K_{(\phi \phi)} = -c^2 T(r). \tag{3.5b}
\]

Furthermore, Eq. (3.4a), which gives the sum of the spatial sectional curvatures along any three mutually orthogonal two-planes, leads in empty space to
\[
K_{(\phi r)} + K_{(\theta r)} + K_{(\theta \phi)} = T(r) + 2D(r) = 0. \tag{3.6}
\]
We have to determine the two functions \( T(r) \) and \( D(r) \); therefore, in addition to (3.6), we need another equation. This comes from Bianchi identities. In the Appendix we show that for any spherically symmetric situation (even if there is nonzero energy density or stress), and choosing the
coordinate \( r \) so that the length of a geodesic circle at radial coordinate \( r \) is \( 2\pi r \) and the area of a geodesic sphere at radial coordinate \( r \) is \( 4\pi r^2 \), then the Bianchi equations imply:

\[
\frac{dT(r)}{dr} = \frac{2D(r) - 2T(r)}{r}.
\]

Equations (3.6) and (3.7) together lead to the differential equation

\[
\frac{dT(r)}{dr} = -\frac{3T(r)}{r},
\]

which has the solution

\[
T(r) = \frac{2m}{r^3}, \quad D(r) = -\frac{m}{r^3},
\]

where \( m \) is the integration constant. The four principal sectional curvatures are therefore

\[
\begin{align*}
K_{(\theta\phi)} &= \frac{2m}{r^3}, \quad K_{(\phi r)} = K_{(\theta r)} = -\frac{m}{r^3}, \\
K_{(rr)} &= -c^2\frac{2m}{r^3}, \quad K_{(r\theta)} = K_{(r\phi)} = c^2\frac{m}{r^3}.
\end{align*}
\]

In order to relate the coefficient \( m \) with the mass generating the field, we observe that in the Newtonian theory of gravitation space–time also displays curvature. The pertinent mathematical model is a space with an affine connection and two compatible metrics: the degenerate temporal metric (for which the length of a timelike path is absolute time, and does not depend on the timelike path itself, but only on the endpoints), and the spatial metrics, defined in each of the simultaneity slices, and which is Euclidean everywhere; they are the \( c \rightarrow \infty \) limit of \( d\tau^2 \) and of the restriction of \( -c^2d\tau^2 \) to the spacelike hypersurfaces. This is the so-called Cartan’s formulation of Newtonian gravity.\(^7\) In this theory, space curvatures \( K \) are equal to zero, but temporal curvatures \( K \) are not, and they show up as tidal forces, which can be computed in an elementary way in the framework of Newtonian theory by means of the geodesic deviation (or tidal force) equation (2.4), where in this case time and space length should be measured by the two separate metrics. The result fixes the coefficient \( k = GM/c^2 \), so that we have finally the three principal spatial sectional curvatures:

\[
\begin{align*}
K_{(\theta\phi)} &= \frac{2GM}{c^2r^3}, \quad K_{(\phi r)} = K_{(\theta r)} = -\frac{GM}{c^2r^3}, \\
K_{(rr)} &= -c^2\frac{2GM}{c^2r^3}, \quad K_{(r\theta)} = K_{(r\phi)} = c^2\frac{GM}{r^3}.
\end{align*}
\]

and the three principal temporal sectional curvatures:

\[
\begin{align*}
K_{(r)} &= -2\frac{GM}{r^3}, \quad K_{(\phi)} = K_{(\theta)} = \frac{GM}{r^3}.
\end{align*}
\]

At this point, we have completed the first task proposed in Sec. I. The dependence on \( r^{-3} \) of sectional curvatures foreseen by dimensional analysis is actually the right one only if \( r \) is defined as mentioned before Eq. (3.7), and we get the exact coefficients \( q = 2 \) and \( q = -1 \). Bianchi identity allows us to understand one of the weak points in the plausibility derivation given in Sec. I. The implicit identification of the unspecified radial coordinate giving the \( GML(c^2r^3) \) dependence with the concrete radial coordinate \( r \) in the metric leads to the right result because Bianchi identities imply a radial \( r^{-3} \) dependence of the curvatures only when the radial coordinate has been chosen so that the area of the spheres at radial coordinate \( r \) is \( 4\pi r^2 \). Note that in Newtonian gravitation, an \( r^{-3} \) dependence of the temporal curvatures \( K \) appears in terms of a coordinate \( r \) such that \( 4\pi r^2 \) is also the area of the sphere; in this case \( r \) coincides—accidentally because three space is flat—with the radial distance, and this is why one is used to thinking of Newtonian gravitation as a theory of action at a distance. In Cartan’s formulation this is no longer so, because Newtonian gravitation is described as locally as in Einstein’s theory (see the pertinent comments in Tipler\(^7\)).

### B. The Schwarzschild metric

The two unknown functions in the Schwarzschild metric, \( e(r), f(r) \), must be determined by requiring that the metric (1.1) should produce the sectional curvatures (3.11) and (3.12). Because of spherical symmetry, the two-dimensional submanifolds \( \{ \theta = \pi/2, \ t = \text{const} \} \) and \( \{ \theta = \pi/2, \ \phi = 0 \} \), are both totally geodesic; therefore the sectional curvatures \( K_{(\phi r)} \) and \( K_{(rr)} \) can be computed directly as the Gaussian curvatures of the two-dimensional metrics induced in the corresponding submanifolds by the metric (1.1). Beware: a similar argument cannot be applied to the submanifolds \( \{ t = \text{const}, \ r = \text{const} \} \), which are not totally geodesic. Summarizing, we have:

(a) The curvature of

\[
d^2 = -c^2dr^2 + e(r)(dr^2 + r^2d\phi^2)
\]

should be \( K_{(\phi r)} = -GM/(c^2r^3) \). By using (2.6), this curvature is shown to be equal to (1.2). If this is equated to the value just mentioned, the following differential equation for \( f(r) \) is obtained:

\[
K_{(\phi r)} = -\frac{1}{2rf(r)} \frac{df(r)}{dr} = -\frac{GM}{c^2r^3}.
\]

This equation is easily solved, because the variables are separated. Imposing the condition \( \lim_{r \rightarrow \infty} f(r) = 1 \) (which can be justified as the curvature of space tends to zero for large \( r \) so that the metric should be Euclidean at spatial infinity), we finally have

\[
f(r) = \left(1 - \frac{2GM}{c^2r} \right)^{-1}.
\]

(a) Now the curvature of the metric

\[
d^2 \big|_{\theta = \pi/2, \phi = 0} = e(r)dr^2 - \frac{1}{c^2 f(r)} dr^2
\]

should be \( K_{(rr)} = -2GM/r^3 \). For this metric, (2.6) gives the curvature as

\[
\frac{dT(r)}{dr} = \frac{2D(r) - 2T(r)}{r}.
\]
where \( f(r) \) is given by (3.14). It is possible to find the general solution for this nonlinear differential equation and then apply the boundary condition \( \lim_{r \to \infty} e(r) = 1 \), but it is easier to use this limiting condition on the equation itself in order to get the asymptotic form of the solution. In the \( r \to \infty \) limit, Eq. (3.15) becomes

\[
\frac{c^2}{2} e''(r) = -\frac{2GM}{r^3},
\]  

(3.16)

whose solution, with the proper asymptotics, is

\[
e_u(r) = 1 - \frac{2GM}{c^2r}.
\]  

(3.17)

Now write \( e(r) = e_u(r)u(r) \); Eq. (3.15) gives for \( u(r) \) the equation

\[
2uu'' - (u')^2 + \frac{6GM}{c^2r^2} f(r)uu' = 0,
\]

which has a constant function \( u(r) = \text{const.} \) as a solution. Therefore it turns out that the expression \( e(r) = e_u(r) \) also satisfies Eq. (3.15), and then it is the particular solution with the desired behavior at spatial infinity we looked for, because the temporal curvatures also tend to zero at spatial infinity, and there the metric should be Minkowskian.

This completes the derivation of the Schwarzschild metric using sectional curvatures. Although we have enforced only the correct values for \( K_{(\phi \theta)} \) and \( K_{(rr)} \), it can be checked that the remaining curvatures \( K_{(\phi \theta)} \) and \( K_{(\theta \theta)} \) as computed from this metric equal the values required by Eqs. (3.11) and (3.12).

IV. OTHER SPHERICALLY SYMMETRIC SOLUTIONS

A. The sectional curvatures generated by a source with a spherically symmetric energy–momentum distribution

Let us now investigate the same problem in the presence of a spherically symmetric energy–momentum distribution. This is the case with the electromagnetic energy–momentum produced by a spherically symmetric distribution of charge, which correspond to the Reissner–Nordström metric. Another such situation is the cosmological homogeneous Robertson–Walker models. We first outline these cases by pointing out the changes introduced by the presence of energy density and/or pressure.

All the discussion in Sec. III A remains unchanged, as it is only based on the assumption of spherical symmetry. However, Einstein’s equations (2.9) involve three source functions (energy density, radial pressure, and transversal pressure) which depend only on \( r \). By first using Eq. (2.9c) we get the following relations between spatial and temporal sectional curvatures:

\[
K_{(rr)} = -c^2K_{(\phi \theta)} + \frac{4\pi G}{c^2} (W^{(0)} + \sigma^{(\theta \theta)} + \sigma^{(\phi \phi)} - \sigma^{(rr)}),
\]

\[
K_{(\phi \theta)} = -c^2K_{(\phi \theta)} + \frac{4\pi G}{c^2} (W^{(0)} + \sigma^{(rr)}),
\]

(4.1)

\[
K_{(\phi \phi)} = -c^2K_{(\phi \phi)} + \frac{4\pi G}{c^2} (W^{(0)} + \sigma^{(rr)}).
\]

Therefore, in this case the six principal curvatures can be expressed in terms of the two functions \( D(r) \) and \( T(r) \), and the source functions, as

\[
K_{(\phi \phi)} = K_{(\theta \theta)} = D(r),
\]

(4.2)

\[
K_{(\phi \theta)} = K_{(\theta \phi)} = -c^2D(r) + \frac{4\pi G}{c^2} (W^{(0)} + \sigma^{(rr)}),
\]

(4.3)

\[
K_{(rr)} = -c^2T(r) + \frac{4\pi G}{c^2} (W^{(0)} + \sigma^{(\theta \theta)} + \sigma^{(\phi \phi)} - \sigma^{(rr)}).
\]

Furthermore, Eq. (2.9a) gives the sum of the spatial sectional curvatures along any three mutually orthogonal two-planes:

\[
K_{(\phi \phi)} + K_{(\theta \theta)} + K_{(\theta \phi)} = T(r) + 2D(r) = \frac{8\pi G}{c^4} W^{(0)}.
\]

(4.4)

Equation (3.7), coming from Bianchi identities, remains unchanged:

\[
\frac{dT(r)}{dr} = 2D(r) - 2T(r),
\]

(4.5)

but when considered together with (4.4) it leads to the differential equation

\[
\frac{dT(r)}{dr} = -\frac{3T(r)}{r} + \frac{8\pi G}{c^4} W^{(0)}.
\]

(4.6)

Equations (4.4) and (4.6) allow one to determine the six sectional curvatures, once the source functions are given.

B. The Reissner–Nordström solution

Let us now assume that the source is a spherically symmetric mass which also has an electric charge \( Q \). The energy density is \( W^{(0)} = \frac{|E|^2}{(8 \pi)} \), and transversal and diametral pressures produced by the electric field are given by electromagnetic theory:

\[
W^{(0)}(r) = \frac{1}{8 \pi} \frac{Q^2}{r^2}, \quad \sigma^{(\theta \theta)} = \sigma^{(\phi \phi)} = \frac{1}{8 \pi} \frac{Q^2}{r^2},
\]

\[
\sigma^{(rr)} = -\frac{1}{8 \pi} \frac{Q^2}{r^2},
\]

(4.7)

where the \( r^{-2} \) dependence of the electric field is consequence of Gauss’s law, provided again the radial coordinate \( r \) is defined so that the area of the sphere at radial coordinate \( r \) is \( 4\pi r^2 \). Note that there is a radial tension along the radial flux lines, and a transversal pressure perpendicular to the flux lines.
In this case, the relations (4.2)–(4.3) between sectional curvatures are:

\[ K_{(r)} = -c^2K_{(\phi)} + \frac{2GQ^2}{c^2r^2}, \]

\[ K_{(\theta)} = -c^2K_{(\phi r)}, \quad K_{(\phi r)} = -c^2K_{(\theta r)}, \]

while (4.4) gives

\[ T(r) + 2D(r) = \frac{GQ^2}{c^4r^5}, \]

and Eq. (4.6) takes the form

\[ \frac{dT(r)}{dr} = -\frac{3T(r)}{r} + \frac{GQ^2}{c^4r^5}. \]

This is an inhomogeneous linear differential equation, whose solution, with the suitable zero charge limit, is:

\[ T(r) = \frac{2GM}{c^2r^3} - \frac{GQ^2}{c^4r^5}, \quad D(r) = -\frac{GM}{c^2r^3} + \frac{GQ^2}{c^4r^5}, \]

so that the curvatures are finally given by (4.2) and (4.3) by substituting (4.7) and (4.11). We can now determine the two unknown functions \( e(r) \), \( f(r) \) by requiring that the metric (1.1) should produce these sectional curvatures. This derivation is completely parallel to the Schwarzschild one:

(a) The curvature of

\[ ds^2 = -c^2dt^2 + \frac{1}{c^2}f(r)dr^2 + r^2d\phi^2 \]

should be

\[ K_{\phi r} = D(r) = -\frac{GM}{c^2r^3} + \frac{GQ^2}{c^4r^5}. \]

This leads to the following differential equation for \( f(r) \):

\[ K_{(\phi r)} = \frac{1}{2rf(r)} \frac{df(r)}{dr} = -\frac{GM}{c^2r^3} + \frac{GQ^2}{c^4r^5}. \]

The solution of this equation having the boundary condition \( \lim_{r \to 0} f(r) = 1 \) is easily found to be

\[ f(r) = \left( 1 - \frac{2GM}{c^2r} + \frac{GQ^2}{c^4r^2} \right)^{-1}. \]

(b) Now the curvature of

\[ ds^2 |_{\theta=\pi/2, \phi=0} = e(r)dt^2 - \frac{1}{c^2}f(r)dr^2 \]

should be

\[ K_{(\theta r)} = -c^2T(r) + \frac{2GQ^2}{c^2r^4} = -\frac{2GM}{c^2r^3} + \frac{3GQ^2}{c^4r^4}. \]

The expression for this curvature in terms of \( e(r) \) and \( f(r) \) has already been given in (3.15). Therefore we have to solve the following second-order differential equation:

\[ K_{(\theta r)} = \frac{c^2}{4e(r)f(r)} \left( 2e''(r) - \frac{e'(r)^2}{e(r)} - \frac{e'(r)f'(r)}{f(r)} \right) = -\frac{2GM}{c^2r^3} + \frac{3GQ^2}{c^4r^4}. \]

By using the expression just found for \( f(r) \), the solution with the boundary condition \( \lim_{r \to 0} e(r) = 1 \) can be found in the same way we did for the Schwarzschild case. The result is now:

\[ e(r) = \left( 1 - \frac{2GM}{c^2r} + \frac{GQ^2}{c^4r^2} \right). \]

This is the Reissner–Nordström metric:

\[ ds^2 = \left( 1 - \frac{2GM}{c^2r} + \frac{GQ^2}{c^4r^2} \right)dt^2 - \frac{1}{c^2} \left( 1 - \frac{2GM}{c^2r} + \frac{GQ^2}{c^4r^2} \right)dr^2 + \frac{r^2}{c^2}d\theta^2 + \sin^2\theta \, d\phi^2. \]

C. The Robertson–Walker metrics

The simplest cosmological models are based on the hypothesis that the space–time is spatially homogeneous and isotropic in the large scale. This constraint means that in a synchronous system (in which locally, and on average, the matter producing the gravitational field is at rest) the metric has the form

\[ ds^2 = dt^2 - \frac{a^2(t)}{c^2} \, dl^2, \]

where \( dl^2 \) is the spatial metric in a space of constant curvature \( k = 1, 0, -1 \). The spatial hypersurfaces are characterized by \( t = \text{const} \), and have curvature equal to \( k/a^2(t) \) with respect to the induced spatial metric \( -c^2d\tau^2 \).

Neglecting at this large scale the local inhomogeneities, the source of the gravitational field is taken as a perfect fluid, with energy density \( \rho \), pressure \( p \), and pressure \( P \), both depending only on \( t \). The conventional derivation imposes the metric (14.7) to be a solution of Einstein’s equations; after computing Christoffel coefficients and Riemann, Ricci, and Einstein tensors, one arrives at the equations for \( a(t) \):

\[ \frac{1}{a^2(t)} \left( \frac{da(t)}{dt} \right)^2 = \frac{8\pi G \rho}{3} - \frac{kc^2}{a^2(t)}, \]

\[ \frac{1}{a(t)} \frac{d^2a(t)}{dt^2} = -\frac{4\pi G}{3} \left( \rho + \frac{3p}{c^2} \right). \]

The simplest situation corresponds to a negligible pressure. In this case, Eq. (14.18) leads to Friedmann’s model and all the well-known predictions concerning the cosmological dynamics.

It is possible to arrive at the same results in a straightforward way by observing that spatial isotropy around any point requires that for any choice of three mutually orthogonal space directions (1), (2), (3), the three temporal sectional curvatures \( K_{(01)}, K_{(02)}, K_{(03)} \) must coincide. The same is true for the spatial sectional curvatures \( K_{(12)}, K_{(23)}, K_{(31)} \). On the other side, spatial homogeneity impose all those sec-
The computation of the spatial sectional curvatures is a bit more involved, as the surfaces \( r = \text{const} \) are not totally geodesic, and in addition to the curvature of the metric induced in these surfaces one must consider the extrinsic curvature coming from the second fundamental form of the embedding. The result is:

\[
K_{(xy)} = K_{(yz)} = K_{(zx)} = \frac{\left( \frac{da(t)}{dt} \right)^2}{c^2a^2(t)} + \frac{k}{a^2(t)},
\]

\[K_{(01)} = K_{(02)} = K_{(03)} = \frac{4\pi G}{3} \left( \frac{\rho + 3p}{c^2} \right), \]

\[K_{(12)} = K_{(23)} = K_{(31)} = \frac{8\pi G \rho}{3c^2}.\]

In this example, the situation is similar to the derivation of the Schwarzschild metric: the knowledge of curvatures, taken directly from Einstein’s equations, must provide a differential equation for the unknown function \( a(t) \). The temporal curvatures \( K_{(01)} \) can be calculated more easily by again using spatial spherically adapted coordinates, where the metric of constant curvature can be written as:

\[dt_k^2 = \frac{1}{1-kr^2} dr^2 + r^2(d\theta^2 + \sin^2 \theta \, d\phi^2).\]

In this case the surface \( \{ \theta = \pi/2, \phi = 0 \} \) is totally geodesic and the curvature \( K_{(tr)} \), which must be equal to any of the curvatures \( K_{(01)} = K_{(02)} = K_{(03)} \), can be simply computed as the curvature of the metric

\[d\tau^2|_{\theta=\pi/2,\phi=0} = dt^2 - \frac{a^2(t)}{c^2} \frac{1}{1-kr^2} dr^2.\]

Using (2.6) we get

\[K_{(tr)} = -\frac{d^2a(t)}{dt^2} \quad (4.20)\]

The impossibility of a derivation of the Schwarzschild metric by using arguments taken only from special relativity and Newtonian gravitation has already been pointed out in a paper in this journal (see Ref. 11 and references therein). The trouble with some attempts to get such simple derivations is the presence of space curvature, which is a specifically relativistic effect of gravitation, and which cannot be accounted for in either special relativity (which ignores gravitation) or in Newtonian gravitation (which ignores relativity). This criticism does not apply to the treatment we have given, which starts by recognizing the space curvature; we claim that this gives a correct derivation. As the departure point is Einstein’s equations themselves, we simply provide another derivation of the Schwarzschild metric alternative to the one found in practically all textbooks of gravitation. We feel that this derivation is worthy because:

- It uses the absolute minimum of elaborate mathematics—the basic curvature formula and solving two elementary differential equations is all which is required—and, in particular, can be formulated without the whole machinery of tensor calculus.
- Once Einstein’s equations written in terms of sectional curvatures are taken as the starting point, the derivation can be grasped by any student after an introduction to the geometric description of gravitation, such as the one contained in the beautiful chapter devoted to this topic in Feynman’s lectures.
- The meaning of what is behind the scenes (the roles of the radial curvature coordinate and of Bianchi identities) is more clearly highlighted than in conventional treatments, where admittedly the differential equations to be solved are not more complex than the ones we found, but are obtained after a rather tedious computation (either brute force or elaborate) of Ricci and Einstein tensors for the most general spherically symmetric metric.
- The disappearance of the space curvature and the permanence of the temporal curvatures in the Newtonian limit can be very clearly seen, helping the student to realize that the description of gravitation as curvature is perfectly possible, and even natural, in a nonrelativistic context (see the similar claims in relation to Newtonian cosmology in Ref. 7). Yet the specific properties of curvature are different in the two cases: in the Newtonian theory curvature is only present along temporal space—time two-planes, but in a relativistic theory nothing happens to time which is not accompanied by some similar fact relative to space, so the presence of spatial curvature should be expected beforehand.

ACKNOWLEDGMENTS

This work has been partially supported by a DGICYT project (No. PB94-1115) from the Ministerio de Educació n y Cultura (Spain), and also by the Junta de Castilla y León (Spain).

APPENDIX: RADIAL DEPENDENCE OF SECTIONAL CURVATURES

Feynman\(^4\) gives a clear argument to understand the \( r^{-3} \) dependence of sectional curvatures. In the three-dimensional space (at a fixed instant), spatial curvatures, which are associated to a two-plane, are analogous to stresses, and the Bianchi identity means that in this space these “stresses” lead to a zero net “force” on any closed volume.

Let us consider the small volume obtained by cutting a small sector, with circular cross section, out of a thin spherical shell centered at the symmetry center (refer to Fig. 2 along this discussion). By suitably choosing the orientation of the coordinate system, this region is described by the ranges \((r, r + dr)\) for the radial coordinate, \((0, d\theta)\) for \(\theta\), and \((0, 2\pi)\) for \(\phi\) (Note that \(r\) is still an unspecified radial coordinate). The top and the bottom of this volume are small sectors of a sphere, spanning a solid angle \(4\pi d\theta^2\). If we choose the radial coordinate in such a way that the total area of the sphere at radial coordinate \(r\) is \(4\pi r^2\) and the length of the geodesic circle at radial coordinate \(r\) is \(2\pi r\), then the bottom has an area \(r^2d\theta^2\) and a perimetric length equal to \(2\pi rd\theta\). Likewise, the area of the top is \(\pi (r + d r)^2d\theta^2\).
By symmetry reasons, the "stress" at the bottom $T(r)$ and at the top $T(r + dr)$ are radial. Again by the spherical symmetry, the hoop stress $D(r)$ is always normal to the lateral side. We must compute the net force produced by these stresses and make it equal to zero.

The net force at the bottom is approximately given by the product of the radial stress by the bottom area. The same happens at the top. Both contributions add up to

$$T(r + dr)(r + dr)\pi d\theta^2 - T(r)(r)\pi d\theta^2 \approx \pi \frac{d(T(r)^2)}{dr} dr d\theta^2. \quad (A1)$$

The key to computing the net force due to the lateral side is to observe that the hoop stress on this lateral side is not exactly orthogonal to the radial stress at the center of the top and bottom. If we decompose it into components parallel and orthogonal to the radial stress at the center of the volume, it is clear that the orthogonal components cancel when integrating over the lateral side. So the only contribution comes from the remaining component of the hoop stress. An elementary geometric reasoning shows that this component equals $-D(r)d\theta$, and to the degree of approximation we are working it is constant over the lateral side.

Hence the "force" due to the lateral side of our small volume is

$$-D(r)d\theta \times 2\pi r dr d\theta. \quad (A2)$$

Thus the condition of zero net "force" is

$$\frac{d(r^2T(r))}{dr} - 2rD(r) = 0 \quad (A3)$$

or

$$\frac{dT(r)}{dr} = \frac{2D(r) - 2T(r)}{r}. \quad (A4)$$

It is worthy to remark how the Schwarzschild radial curvature coordinate $r$ enters into this argument. Should we use any other radial coordinate, the areas of the top, bottom, and lateral sides of the small volume would be expressed in terms of the new radial coordinate by more complicated expressions, and the Bianchi identity would still express the balance of forces, tough in no so simple a form.

---

a) Electronic mail: santander@cpd.uva.es
b) Electronic mail: lmnieto@cpd.uva.es
c) Electronic mail: nicolas@fixat8.fam.cie.uva.es
1Michael Berry, Principles of Cosmology and Gravitation (IOP, Bristol, 1989).
5Michael Spivak, A Comprehensive Introduction to Differential Geometry (Publish or Perish, Berkeley, 1979), Vol. II.
10Theodore Frankel, Gravitational Curvature (Freeman, San Francisco, 1979).

---

THESOCRATIC METHOD

I seem to be advocating what is sometimes called the Socratic or do-it-yourself or discovery method, or, especially in Texas, the Moore method. The method is not to tell students but to ask them, and, better yet, to inspire them to ask themselves—make students solve problems, and better yet, train students, by example, encouragement, and generous reinforcement, to construct problems of their own. Problem solving—that is the most highly touted current shibboleth, and that is the flag that I too want to wave. The flag should be kept waving; the important ideas deserve to be emphasized over and over again.

The most effective way to teach mathematics by problem solving is to keep challenging students with problems that are just barely within their reach.