Chapter 3 Field Equations and Curvature

3.0 Introduction

The General Relativity (GR) field equations quantify the relationship between energy and space curvature. In flat Euclidean space, there is a single field equation, Poisson's equation (2.6-11):

$$\nabla^2 V = 4\pi G\rho.$$

This is a field equation because LHS contains the energy potential V, developed in equation (2.6-2), which represents all possible sources of energy, and RHS contains G, which represents gravity, the source of curvature in general relativity.

In Section 3.1 we introduce the stress tensor, $T^{\mu\nu}$, which represents all sources of energy. We show the relationship beween $T^{\mu\nu}$ and the relativistic continuity equation and equation of motion, analogous to Section 2.6 where we deveoped Poisson's equation and the fluid continuity equation from Newton's equation of motion.

In Sections 3.2 – 3.4, we develop curvature, including the various curvature tensors, parallel transport, and geodesic deviation.

In Section 3.5 we present the GR field equations, and in Section 3.6 we show that they meet the constraint that Poisson's equation be satisfied in flat Euclidean space at non-relativistic speeds.

The Schwarzschild exact solution is derived in Section 3.7.

3.1 The stress tensor and fluid motion

We begin in *flat spacetime* using the (ct,x,y,z) inertial coordinate system, then extend the equations to curved spacetime, and then to non-inertial curved spacetime.

Convention Boldface is used for 3-vectors, and non boldface for 4-vectors:

$\lambda^\mu \equiv (\lambda^0, \lambda^1, \lambda^2, \lambda^3) \equiv (\lambda^0, \boldsymbol{\lambda})$.

We start with a particle and use the following notation:

$m \equiv$ **rest or proper mass of particle**

$t \equiv$ **coordinate time**

$\tau \equiv$ **proper time**

$v \equiv$ **speed of particle**

$$\gamma \equiv \frac{\mathrm{d}t}{\mathrm{d}\tau} \overset{(A.42)}{=} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$E \equiv \gamma m c^2 \equiv$ **energy of particle** (3.1-1)

$v^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t} \equiv$ **coordinate velocity of particle**

$u^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \gamma v^\mu \equiv$ **world velocity of particle** (3.1-2)

$p^\mu \equiv m u^\mu \equiv$ **4-momentum of particle** (3.1-3)

Since $\dfrac{\mathrm{d}x^0}{\mathrm{d}t} = \dfrac{\mathrm{d}(ct)}{\mathrm{d}t} = c$, by the convention above we get that

$$v^\mu = (c, \boldsymbol{v}) \tag{3.1-4}$$

where $\boldsymbol{v} = (v^1, v^2, v^3)$ is the velocity 3-vector. Hence, the particle's speed is

$$v = |\boldsymbol{v}|, \text{ and } v^2 = \left(v^1\right)^2 + \left(v^2\right)^2 + \left(v^3\right)^2 \tag{3.1-5}$$

Of the quantities listed above, only $m = m(x^\mu)$ and $\tau = \tau(x^\mu)$ are scalars and only u^μ and p^μ are vectors:

We showed in Section A.6 that, $v^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t}$ is not a vector because $\dfrac{\mathrm{d}\,x^\mu}{\mathrm{d}t'} \neq \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t}$,

whereas $u^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ *is* as vector because $\tau' = \tau$.

E, v, and t are not functions of (x^μ) so, by definition (1.7-4), they are not scalars.

For a moving particle,

$$u^\mu \overset{(3.1-2)}{=} \gamma v^\mu \overset{(3.1-4)}{=} \gamma(c,\boldsymbol{v}) \qquad (3.1\text{-}6)$$

$$p^\mu \overset{(3.1-3)}{=} m u^\mu \overset{(3.1-6)}{=} m\gamma(c,\boldsymbol{v}) = (m\gamma c, m\gamma \boldsymbol{v}) \overset{(3.1-1)}{=} \left(\frac{E}{c}, \boldsymbol{p}\right) \qquad (3.1)$$

where the 3-momentum $\boldsymbol{p} \equiv m\gamma\boldsymbol{v}$. (3.1-7)

For a stationary particle, $\gamma = 1$, $\boldsymbol{v} = 0$, and so

$$u^\mu \overset{(3.1-6)}{=} \gamma(c,\boldsymbol{v}) = (c,\mathbf{0}) \qquad (3.1\text{-}8)$$

$$p^\mu \overset{(3.1-3)}{=} m u^\mu = m(c,\mathbf{0})$$

$p^0 = mc = \dfrac{mc^2}{c} \overset{(3.1-1)}{=} \dfrac{E}{c}$ is proportional to the **rest energy**, $E \equiv mc^2$

Equation (3.1) reminds us that energy and 3-momentum are components of a single momentum 4-vector, just as time and position are components of the position 4-vector.

Another useful formula is

$$u^\mu u_\mu = c^2 : \qquad (3.1\text{-}9)$$

$$u^\mu u_\mu \overset{(3.1-6)}{=} \gamma^2\, v^\mu v_\mu \overset{(\text{Example } A.0.1\,(b))}{=} \gamma^2\,[(v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2]$$

$$\overset{(3.1-4,\,3.1-5)}{=} \gamma^2\,(c^2 - v^2) \overset{(A.42)}{=} \frac{c^2 - v^2}{1 - \frac{v^2}{c^2}} = c^2 \checkmark$$

We next extend these concepts to the continuous case. The simplest such environment is a perfect fluid characterized by two scalar fields, density ρ and pressure P, and a vector field, world velocity u^μ. In order for ρ to be a scalar field, we must define it to be **proper density**, the rest mass per unit rest volume. In calculations we now use the **4-momentum density** ρu^μ in place of the particle 4-momentum $p^\mu = m u^\mu$. We italicize pressure P in order to distinguish it from a point P, although the context should always make the distinction clear. (The book uses *p,* which looks a lot like ρ, and is also the symbol that should represent the magnitude of the 3-momentum, $\boldsymbol{p}$.)

For a particle, p^μ is a tensor that captures both energy and momentum. For a fluid, we seek a tensor that captures the energy and 4-momentum density of the fluid. The derivation is beyond the scope of this book so we just give the tensor and motivate it by describing its properties.

Definition The **energy-momentum-stress tensor** (or, **stress tensor** for short) for a perfect fluid is defined as

$$\boxed{T^{\mu\nu} \equiv \left(\rho + \frac{P}{c^2}\right) u^\mu u^\nu - P\eta^{\mu\nu}} \quad , \tag{3.2}$$

where η is the metric tensor for special relativity defined in equation (A.4).

Observe that $T^{\mu\nu}$ is symmetric and is composed of ρ, P, and u^μ, the scalar and vector fields that characterize the fluid. Next, observe that

$T^{\mu\nu} u_\nu = c^2 \rho u^\mu$: (3.1-10)

$$T^{\mu\nu} u_\nu \overset{(3.2)}{=} (\rho + \frac{P}{c^2}) u^\mu u^\nu u_\nu - P\, \eta^{\mu\nu} u_\nu \overset{(3.1\text{-}9)}{=} (\rho + \frac{P}{c^2}) u^\mu c^2 - P u^\mu = c^2 \rho u^\mu \quad ✓$$

That is, $T^{\mu\nu} u_\nu$ equals the 4-momentum density ρu^μ up to a factor of c^2.

Recalling that $\lambda^\nu{}_{,\mu} = \dfrac{\partial \lambda^\nu}{\partial x^\mu}$ so that, for example, $u^\mu{}_{,\mu} = \dfrac{\partial u^\mu}{\partial x^\mu} = \dfrac{\partial u^0}{\partial x^0} + \dfrac{\partial u^1}{\partial x^1} + \dfrac{\partial u^2}{\partial x^2} + \dfrac{\partial u^3}{\partial x^3}$, we have the following definitions.

Definition The **relativistic continuity equation** is

$$\boxed{(\rho\, u^\mu)_{,\mu} + \frac{P}{c^2}\, u^\mu{}_{,\mu} = 0} \tag{3.5}$$

and the **equation of motion of a perfect fluid** is

$$\boxed{\left(\rho + \frac{P}{c^2}\right) u^\nu{}_{,\mu}\, u^\mu = \left(\eta^{\mu\nu} - \frac{1}{c^2}\, u^\mu u^\nu\right) P_{,\mu}} \,. \tag{3.6}$$

Recall that for 3-vectors, **divergence** was defined as

$$\textbf{div}\ \boldsymbol{v} \overset{(2.61\text{ A})}{\equiv} \nabla \bullet \boldsymbol{v} = \frac{\partial v^i}{\partial x^i} = v^i{}_{,i}\,.$$

Since the stress tensor is symmetric, it also only has one **divergence:**

$$\textbf{div}\ \boldsymbol{T}^{\mu\nu} \equiv \nabla \bullet T^{\mu\nu} \overset{(2.61\text{ D})}{\equiv} T^{\mu\nu}{}_{,\mu}\,. \tag{3.1-11}$$

Theorem 3.1.1 The divergence $T^{\mu\nu}{}_{,\mu} = 0$ iff equations (3.5) and (3.6) hold.

Proof. First, suppose $T^{\mu\nu}{}_{,\mu} = 0$. Since $u^\nu u_\nu \overset{(3.1-9)}{=} c^2$ we get

$$u^\nu{}_{,\mu}\, u_\nu + u^\nu\, u_{\nu,\mu} = 0\,. \tag{3.4}$$

Therefore (Exercise 3.1.4) $u^\nu{}_{,\mu}\, u_\nu = 0$: (3.1-12)

$$\begin{aligned} 0 &\overset{(3.4)}{=} u^\nu{}_{,\mu}\, u_\nu + u^\nu\, u_{\nu,\mu} \\ &\overset{(\text{Ex A0.1 a})}{=} (u^0 u^0{}_{,\mu} - u^1 u^1{}_{,\mu} - u^2 u^2{}_{,\mu} - u^3 u^3{}_{,\mu}) + (u^0 u^0{}_{,\mu} - u^1 u^1{}_{,\mu} - u^2 u^2{}_{,\mu} - u^3 u^3{}_{,\mu}) \\ &= 2\, u^\nu{}_{,\mu}\, u_\mu \quad ✓ \end{aligned}$$

Differentiating equation (3.2) with respect to x^μ yields

$$0 = T^{\mu\nu}{}_{,\mu} = \rho u^\mu{}_{,\mu} u^\nu + \rho u^\mu u^\nu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu} u^\nu + \frac{P}{c^2} u^\mu u^\nu{}_{,\mu} + \frac{1}{c^2} P_{,\mu} u^\mu u^\nu - P_{,\mu} \eta^{\mu\nu} \tag{3.3}$$

Contracting equation (3.3) with u_ν and applying (3.1-9) yields equation (3.5) :

$$\begin{aligned} 0 &= \rho\, u^\mu{}_{,\mu}\, c^2 + \rho\, u^\nu{}_{,\mu}\, u^\mu\, u_\nu + \frac{P}{c^2} u^\mu{}_{,\mu}\, c^2 + \frac{P}{c^2} u^\mu\, u_\nu\, u^\nu{}_{,\mu} + \frac{1}{c^2} P_{,\mu}\, u^\mu\, c^2 - P_{,\mu}\, u_\nu\, \eta^{\mu\nu} \\ &= c^2 \left(\rho\, u^\mu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu}\right) + u^\mu\, u_\nu\, u^\nu{}_{,\mu} \left(\rho + \frac{P}{c^2}\right) + P_{,\mu} \left(u^\mu - u^\mu\right) \\ &\overset{(3.4)}{=} c^2 \left(\rho\, u^\mu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu}\right) \qquad ✓ \end{aligned}$$

Applying equation (3.5) to equation (3.3) yields equation (3.6). ✓
Conversely, if equations (3.5) and (3.6) hold, plugging them into equation (3.2) yields $T^{\mu\nu}{}_{,\mu} = 0$. ■

In an instantaneous rest system the coordinate divergence is zero, so the equation $T^{\mu\nu}{}_{,\mu} = 0$ is equivalent to setting its zeroth component to zero: $T^{\mu 0}{}_{,\mu} = 0$:

For a particle at rest,

$$u^\mu \overset{(3.1-8)}{=} (c, \mathbf{0}) \;\Rightarrow\; u_\nu \overset{(\text{Ex } A.0\times.1)}{=} u_0 = c \text{ and } u_i = 0 \text{ for } i = 1,2,3.$$

So,

$$0 = T^{\mu\nu}{}_{,\mu}\, u_\nu = T^{\mu 0}{}_{,\mu}\, u_o + T^{\mu i}{}_{,\mu}\, u_i = c\, T^{\mu 0}{}_{,\mu} \quad \Leftrightarrow \quad T^{\mu 0}{}_{,\mu} = 0 \quad ✓$$

Equations (3.7) and (3.9), below, were derived in Section 2.6 as equation (2.6-12) and (2.6-15), respectively. The first represents conservation of mass in fluid flow; the second is the fluid equation of motion, $\boldsymbol{F} = m\boldsymbol{a}$, where force on a fluid is represented by pressure and mass is represented by density, and it and includes both internal and external forces.

Definition The **classical continuity equation for a perfect fluid** is

$$\boxed{\partial\rho/\partial t + \nabla\cdot(\rho\boldsymbol{v}) = 0} \tag{3.7}$$

where (in the classical case) ρ is both the proper density (i.e., global flat spacetime coordinates) and the coordinate density (i.e., local coordinates), and

$$\nabla\cdot(\rho\boldsymbol{v}) = \left(\frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k}\right)\cdot(\rho v^1\boldsymbol{i} + \rho v^2\boldsymbol{j} + \rho v^3\boldsymbol{k}) = \frac{\partial}{\partial x^i}(\rho\, v^i) \tag{3.1-13}$$

Euler's **classical equation of motion for a perfect fluid** is

$$\boxed{\rho\left(\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\nabla\right)\boldsymbol{v} = -\nabla P} \tag{3.9}$$

where

$$\boldsymbol{v}\cdot\nabla = v^i\frac{\partial}{\partial x^i}. \tag{3.1-14}$$

We now show that the relativistic continuity and motion equations (3.5) and (3.6) reduce, respectively, to equations (3.7) and (3.9) in the classical limit of a slowly moving fluid ($v \ll c$) and small pressure ($P \ll c^2$). In this case,

$$\gamma = 1,\ \ u^\mu = v^\mu,\ \ \frac{v^0}{c} = 1,\ \ \frac{v^i}{c} = 0,\ \text{and}\ \ \frac{P}{c^2} = 0. \tag{3.1-15}$$

Equation (3.5), $(\rho\, u^\mu)_{,\mu} + \frac{P}{c^2}u^\mu{}_{,\mu} = 0$, reduces to

$$(\rho\, c)_{,0} + \left(\rho\, v^i\right)_{,i} = 0: \tag{3.1-16}$$

$$(\rho\, u^\mu)_{,\mu} + \frac{P}{c^2}u^\mu{}_{,\mu} \overset{(3.1-15)}{=} (\rho\, u^\mu)_{,\mu} \overset{(3.1-6)}{=} (\rho\,\gamma\, c)_{,0} + \left(\rho\ \gamma\, v^i\right)_{,i}$$

$$\overset{(3.1-15)}{=} (\rho\, c)_{,0} + \left(\rho\ v^i\right)_{,i}\ \ ✓$$

which in 3-vector notation is equation (3.7):

$$0 \overset{(3.1-16)}{=} \frac{\partial(\rho\, c)}{\partial(c\, t)} + \frac{\partial}{\partial x^i}(\rho\, v^i) \overset{(3.1-13)}{=} \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v}) \qquad ✓ \qquad \blacksquare$$

Equation (3.6), $(\rho + \frac{P}{c^2})\, u^\nu{}_{,\mu}\, u^\mu = (\eta^{\mu\nu} - \frac{1}{c^2} u^\mu u^\nu)\, P_{,\mu}$, reduces to

$$\rho\, v^\nu{}_{,\mu} v^\mu = (\eta^{\mu\nu} - \frac{1}{c^2} v^\mu v^\nu)\, P_{,\mu} \qquad (3.8)$$

because $\frac{P}{c^2} \overset{(3.1\text{–}15)}{=} 0$ and $u^\mu \overset{(3.1\text{–}15)}{=} v^\mu$.

The zeroth component of equation (3.8) is 0 = 0:

$$\text{LHS} = \rho\, v^0{}_{,\mu} v^\mu = \rho \frac{\partial v^0}{\partial x^\mu} v^\mu \overset{(3.1\text{–}4)}{=} \rho \frac{\partial c}{\partial x^\mu} v^\mu = 0 \qquad ✓$$

$$\text{RHS} = (\eta^{\mu 0} - \frac{1}{c^2} v^\mu v^0)\, P_{,\mu} = [\eta^{00} - \frac{1}{c^2} (v^0)^2\,]\, P_{,0} + [\eta^{i\,0} - \frac{1}{c^2} v^i v^0]\;\; P_{,i}$$

$$\overset{(3.1\text{–}4)}{=} [1 - \frac{c^2}{c^2}]\, P_{,0} + [0 - \frac{1}{c^2} v^i v^0]\;\; P_{,i} \overset{(3.1\text{–}15)}{=} 0 + [0 - 0] = 0 \qquad ✓$$

The non-zero components are

$$\rho\, v^i{}_{,\mu} v^\mu \overset{(3.8)}{=} (\eta^{\mu i} - \frac{1}{c^2} v^\mu v^i)\, P_{,\mu} = (\eta^{0\,i} - \frac{1}{c^2} v^0 v^i)\, P_{,0} + (\eta^{j\,i} - \frac{1}{c^2} v^j v^i)\, P_{,j}$$

$$\overset{(3.1\text{–}4)}{=} (0 - \frac{v^i}{c})\, P_{,0} + (-\,\delta^{j\,i} - \frac{1}{c^2} v^j v^i)\, P_{,j} \overset{(3.1\text{–}15)}{=} (0\text{–}\,0)\, P_{,0} - \delta^{j\,i}\, P_{,j}$$

$$= -\, P_{,i}$$

But,

$$\rho\, v^i{}_{,\mu} v^\mu = \rho\, (v^i{}_{,0} v^0 + v^i{}_{,j} v^j) \overset{(3.1\text{–}4)}{=} \rho\, (\frac{\partial v^i}{\partial (c\,t)}\; c + v^i{}_{,j} v^j) = \rho\, (\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j).$$

So, the ith component of equation (3.8) can be expressed as

$$\rho\, (\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j) = -\, P_{,i}\;,$$

which in 3-vector form is equation (3.9), $\rho \left(\frac{\partial}{\partial t} + \boldsymbol{v} \bullet \nabla \right) \boldsymbol{v} = -\nabla P$:

$\frac{\partial \boldsymbol{v}}{\partial t}$ has ith component $\frac{\partial v^i}{\partial t}$.

$(\boldsymbol{v} \bullet \nabla)\boldsymbol{v} \overset{(3.1\text{–}14)}{=} v^j\, \frac{\partial \boldsymbol{v}}{\partial x^j}$ has ith component $v^j\, \frac{\partial v^i}{\partial x^j} = v^i{}_{,j} v^j$.

So, LHS of equation (3.9) has ith component $\rho\, (\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j)$. ✓

RHS has ith component $-\,\frac{\partial P}{\partial x^i} = -\, P_{,i}$ ✓ ■

The relativistic continuity equation (3.5) contains pressure while the classical version, equation (3.7), does not. That makes sense because energy rather than mass is conserved, and a fluid under pressure contributes to energy.

The relativistic equation of motion for a fluid (3.6) can be expressed as

$$(\rho + \frac{P}{c^2})\frac{d^2 x^\nu}{d\tau^2} = (\eta^{\mu\nu} - \frac{1}{c^2} u^\mu u^\nu)\, P_{,\mu} :$$

$$\frac{d^2 x^\nu}{d\tau^2} = \frac{d}{d\tau}\frac{dx^\nu}{d\tau} = (\frac{d}{d\tau}\delta^\nu_\mu)\frac{dx^\mu}{d\tau} = (\frac{d}{d\tau}\frac{\partial x^\nu}{\partial x^\mu})\frac{dx^\mu}{d\tau} = (\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\tau})\frac{dx^\mu}{d\tau}$$

$$= \frac{\partial u^\nu}{\partial x^\mu} u^\mu = u^\nu{}_{,\mu}\, u^\mu$$

Thus, LHS of equation (3.6) can be written as $(\rho + \frac{P}{c^2})\frac{d^2 x^\nu}{d\tau^2}$ ✓

In this form, equation (3.6) looks more like an equation of motion, $F = ma$, pushing fluid particles off of geodesics (i.e., $\frac{d^2 x^\nu}{d\tau^2} = 0$) by the pressure gradient $P_{,\mu}$.

This finishes our derivation for flat spacetime. Using Principles 1 and 2 of Section 2.5, we can generalize these special relativity results to general relativity. This involves replacing $\eta^{\mu\nu}$ by $g^{\mu\nu}$ and replacing the partial derivative by the covariant derivative.

Definitions The **continuity equation** is

$$\boxed{(\rho\, u^\mu)_{;\mu} + \frac{P}{c^2}\, u^\mu{}_{;\mu} = 0} \,. \tag{3.1-17}$$

The **equation of motion of a perfect fluid** is

$$\boxed{\left(\rho + \frac{P}{c^2}\right) u^\nu{}_{;\mu}\, u^\mu = \left(g^{\mu\nu} - \frac{1}{c^2}\, u^\mu\, u^\nu\right) P_{;\mu}} \,. \tag{3.1-18}$$

The **stress tensor** for a perfect fluid is defined as

$$\boxed{T^{\mu\nu} \equiv \left(\rho + \frac{P}{c^2}\right) u^\mu\, u^\nu - P g^{\mu\nu}} \tag{3.10}$$

and the vanishing of its divergence is expressed as

$$T^{\mu\nu}{}_{;\mu} = 0 \,. \tag{3.11}$$

Recall Theorem 3.1.1, which easily generalizes to say that equation (3.11) holds iff both the continuity equation (3.1-17) and equation of motion equation (3.1-18) hold. Also, since energy and mass are equivalent in relativity, the relativistic continuity equation (3.1-17) now represents conservation of energy, not just conservation of fluid mass like its classical counterpart equation (3.7).

Associated energy-momentum tensors are

$$T_{\mu\nu} = (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\mu\nu} \quad \text{and} \quad T_\mu{}^\nu = (\rho + \frac{P}{c^2})\, u_\mu u^\nu - P\delta_\mu^\nu : \qquad (3.1\text{-}19)$$

$$T_\mu{}^\nu = g_{\mu\sigma} T^{\sigma\nu} = (\rho + \frac{P}{c^2})\, g_{\mu\sigma} u^\sigma u^\nu - P\, g_{\mu\sigma} g^{\sigma\nu} = (\rho + \frac{P}{c^2})\, u_\mu u^\nu - P\delta_\mu^\nu \quad ✓$$

$$T_{\mu\nu} = g_{\nu\rho} T_\mu{}^\rho = (\rho + \frac{P}{c^2})\, g_{\nu\rho} u_\mu u^\rho - P g_{\nu\rho} g_\mu^\rho = (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\nu\mu} \quad ✓$$

With suitable definition of $T^{\mu\nu}$, equation (3.11) is valid for all fluids and fields. The stress tensor will soon be shown to be the source of the gravitational field, but first we must discuss curvature. We should also point out that we define the stress to include all forces that exert pressure on the perfect fluid. That is, curved spacetime will result not just from gravitational force but from all forces.

3.2 The curvature tensor and related tensors

Unlike with partial derivatives, the order of differentiation matters with covariant derivatives. In this section, we develop the curvature tensor, *R,* from conditions that govern when the order matters, then we explore how *R* relates to spacetime curvature, and we conclude with introduction of the associated Ricci and Einstein tensors. During this development, we only assume an *N*-dimensional manifold and, so, we use indices *a*, *b*, *c*, etc. instead of μ, ν, ρ. These results, however, apply to general relativity with $N = 4$.

The covariant derivative of a vector λ_a is

$$\lambda_{a\,;\,b} \overset{(2.56)}{=} \partial_b \lambda_a - \Gamma^d_{a\,b} \lambda_d \tag{3.2-1}$$

and its second derivative is

$$\lambda_{a\,;\,b\,c} = \partial_c(\lambda_{a\,;b}) - \Gamma^e_{a\,c} \lambda_{e\,;\,b} - \Gamma^e_{b\,c} \lambda_{a\,;\,e} : \tag{3.2-2}$$

Let $\tau_{a\,b} \equiv \lambda_{a\,;\,b}$. Then $\lambda_{a\,;\,b\,c} = \tau_{a\,b\,;\,c} \overset{(2.58)}{=} \partial_c \tau_{a\,b} - \Gamma^e_{a\,c} \tau_{e\,b} - \Gamma^e_{b\,c} \tau_{a\,e} = \text{RHS}$ ✓

Applying (3.2-1) to each term of (3.2-2) yields

$$\lambda_{a\,;\,b\,c} = \partial_c \partial_b \lambda_a - (\partial_c \Gamma^d_{a\,b})\, \lambda_d - \Gamma^d_{a\,b}\, \partial_c \lambda_d - \Gamma^e_{a\,c}\, [\partial_b \lambda_e - \Gamma^d_{e\,b} \lambda_d] - \Gamma^e_{b\,c}\, [\partial_e \lambda_a - \Gamma^d_{a\,e} \lambda_d]\,.$$

Interchanging *b* and *c* gives

$$\lambda_{a\,;\,c\,b} = \partial_b \partial_c \lambda_a - (\partial_b \Gamma^d_{a\,c})\, \lambda_d - \Gamma^d_{a\,c}\, \partial_b \lambda_d - \Gamma^e_{a\,b}\, [\partial_c \lambda_e - \Gamma^d_{e\,c} \lambda_d] - \Gamma^e_{c\,b}\, [\partial_e \lambda_a - \Gamma^d_{a\,e} \lambda_d]\,.$$

Subtracting gives

$$\lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b} = (\partial_b \Gamma^d_{a\,c} - \partial_c \Gamma^d_{a\,b} + \Gamma^e_{a\,c}\, \Gamma^d_{e\,b} - \Gamma^e_{a\,b}\, \Gamma^d_{e\,c})\, \lambda_d$$

since $\Gamma^e_{b\,c} = \Gamma^e_{c\,b}$. Define

$$\boxed{R^d{}_{a\,b\,c} \equiv \partial_b \Gamma^d_{a\,c} - \partial_c \Gamma^d_{a\,b} + \Gamma^e_{a\,c} \Gamma^d_{e\,b} - \Gamma^e_{a\,b} \Gamma^d_{e\,c}}\,. \tag{3.13}$$

Then $$\lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b} = R^d{}_{a\,b\,c}\, \lambda_d \tag{3.12}$$

The following theorem is an immediate consequence of equation (3.12).

Theorem 3.2.1 $\lambda_{a\,;\,b\,c} = \lambda_{a\,;\,c\,b}$ iff $R^d{}_{a\,b\,c} = 0$ for all *a, b, c, d*. That is, a necessary and sufficient condition for interchanging the order of covariant differentiation of covariant vector fields is that $R^d{}_{a\,b\,c} = 0$.

Theorem 3.2.2 $R^d{}_{a\,b\,c}$ is a tensor.

Proof. $\lambda_{a\,;\,b\,c}$ is a tensor, so $R^d{}_{a\,b\,c}\,\lambda_d = \lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b}$ is a tensor for arbitrary vectors λ_d. Thus, $R^d{}_{a\,b\,c}$ satisfies the hypothesis of the Quotient Theorem 1.8.1, that contraction with an arbitrary vector generates a tensor. Thus, $R^d{}_{a\,b\,c}$ is a tensor. ■

Since $R^d{}_{a\,b\,c}$ is defined in terms of the metric tensor and its derivatives, we can expect it to be related to curvature. This is motivation for the following definition.

Definition $\boldsymbol{R^d{}_{a\,b\,c}}$ is called the **curvature tensor** (or **Riemann-Christoffel tensor** or **Riemann tensor**).

The next theorem generalizes Theorem 3.2.1 from vectors to tensors.

Theorem 3.2.3 (Exercise 3.2.1) A necessary and sufficient condition for interchanging the order of covariant differentiation of tensor fields is that $R^d{}_{a\,b\,c} = 0$.

In the solution to Exercise 3.2.1, we learned that the difference between the two 2nd order partial derivatives of a general tensor can be expressed

$$\boxed{\begin{aligned}&\tau^{a_1\ldots a_r}{}_{b_1\ldots b_s\,;c\,d} - \tau^{a_1\ldots a_r}{}_{b_1\ldots b_s\,;d\,c}\\ &= -\textstyle\sum_{k=1}^{r} R^{a_k}_{e\,c\,d}\,\tau^{a_1\ldots a_{k-1}\,e\,a_{k+1}\ldots a_r}_{b_1\ldots b_s} + \textstyle\sum_{k=1}^{s} R^{e}_{b_k\,c\,d}\,\tau^{a_1\ldots a_r}_{b_1\ldots b_{k-1}\,e\,b_{k+1}\ldots b_s}\end{aligned}} \tag{3.2-3}$$

Special cases that more easily reveal the pattern are

$$\lambda^a{}_{;\,b\,c} - \lambda^a{}_{;\,c\,b} = -\,R^a_{d\,b\,c}\,\lambda^d$$

$$\tau^{a\,b}{}_{;\,c\,d} - \tau^{a\,b}{}_{;\,d\,c} = -\,R^a_{e\,c\,d}\,\tau^{e\,b} - R^b_{e\,c\,d}\,\tau^{a\,e}$$

$$\tau^{a\,b}{}_{c\,;\,d\,e} - \tau^{a\,b}{}_{c\,;\,e\,d} = -\,R^a_{f\,d\,e}\,\tau^{f\,b}{}_c - R^b_{f\,d\,e}\,\tau^{a\,f}{}_c + R^f_{c\,d\,e}\,\tau^{a\,b}{}_f$$

$$\lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b} = R^d_{a\,b\,c}\,\lambda_d$$

$$\tau_{a\,b\,;\,c\,d} - \tau_{a\,b\,;\,d\,c} = R^e_{a\,c\,d}\,\tau_{e\,b} + R^e_{b\,c\,d}\,\tau_{a\,e}$$

In flat Euclidean space using the Cartesian coordinate system, $\Gamma^a_{b\,c} = 0\ \ \forall\ a, b, c$. This implies that $R^a{}_{b\,c\,d} = 0$. A sphere, though embedded in flat Euclidean space, has non-zero $R^a{}_{b\,c\,d}$ in spherical coordinates. Based on such examples, we make the following definition.

Definition A **manifold is flat** if $R^a{}_{bcd} = 0$ at every point; otherwise, the **manifold is curved**.

By the Theorem 3.2.3, a manifold is flat iff we can interchange the order of the 1st and 2nd partial differentials of tensors.

Theorem 3.2.4 In any flat region of a manifold it is possible to introduce a Cartesian coordinate system (i.e., $g_{ab} = \eta_{ab}$).

At first glance, $R^a{}_{bcd}$ has N^4 components. However, $R^a{}_{bcd}$ possesses a number of symmetries that can be shown to cut the number down to $\frac{1}{12}N^2(N^2\text{-}1)$. For spacetime, this is a reduction from 256 unknowns to 20.

Definition The equation

$$R^a{}_{bcd} + R^a{}_{cdb} + R^a{}_{dbc} = 0 \tag{3.14}$$

is known as the **cyclic identity**. (Exercise 3.2.2)

We saw in Section 1.8 that the metric tensors are used to create *associated tensors*.

Two **associated curvature tensors** are

$$R_{abcd} \equiv g_{ae}\, R^e{}_{bcd} \quad \text{and} \quad R^{ab}{}_{cd} \equiv g^{be}\, R^a{}_{ecd} \tag{3.2-4}$$

After extensive manipulation,

$$\begin{aligned} R_{abcd} = & \tfrac{1}{2}\,[\partial_d\partial_a g_{bc} - \partial_d\partial_b g_{ac} + \partial_c\partial_b g_{ad} - \partial_c\partial_a g_{bd}] \\ & - g^{ef}\,[\Gamma_{eac}\,\Gamma_{fbd} - \Gamma_{ead}\,\Gamma_{fbc}] \end{aligned} \tag{3.15}$$

From equation (3.15) it is simple to check the following symmetry properties:

$$R_{abcd} = - R_{bacd} \tag{3.16}$$
$$R_{abcd} = - R_{abdc} \tag{3.17}$$
$$R_{abcd} = R_{badc} \tag{3.2-5}$$
$$R_{abcd} = R_{cdab} \tag{3.18}$$

For example,

$$\begin{aligned} - R_{bacd} &\overset{(3.15)}{=} -\tfrac{1}{2}\,[\partial_d \partial_b g_{ac} - \partial_d \partial_a g_{bc} + \partial_c \partial_a g_{bd} - \partial_c \partial_b g_{ac}] \\ &+ \; g^{ef}\,[\Gamma_{ebc}\,\Gamma_{fad} - \Gamma_{ebd}\,\Gamma_{fac}] \\ &= \tfrac{1}{2}\,[-\partial_d \partial_b g_{ac} + \partial_d \partial_a g_{bc} - \partial_c \partial_a g_{bd} + \partial_c \partial_b g_{ac}] \\ &\quad - \; g^{ef}\,[-\,\Gamma_{fbc}\,\Gamma_{ead} + \Gamma_{fbd}\,\Gamma_{eac}] \\ &\overset{(3.15)}{=} R_{abcd} \qquad \checkmark \end{aligned}$$

Note: Take care that the curvature and other associated tensors may not necessarily mimic the behaviors (3.16 – 3.18 and 3.2-5):

To show this, we use the fact that $g^{ae}\, R_{ebcd} = R^a{}_{bcd}$ *raises b* and turns it into *a*. It is true that, for example, that

$$\begin{aligned} R^{ab}{}_{cd} &\overset{\text{(raise)}}{=} g^{be} R^a{}_{ecd} \overset{\text{(raise)}}{=} g^{be} g^{af}\, R_{fecd} \overset{(3.2\text{–}5)}{=} g^{af}\, g^{be}\, R_{efdc} \\ &\overset{\text{(raise)}}{=} g^{af} R^b{}_{fdc} \overset{\text{(raise)}}{=} R^{ba}{}_{dc} \end{aligned} \tag{3.2-6}$$

and

$$R^a{}_{bcd} \overset{\text{(raise)}}{=} g^{ae}\, R_{ebcd} \overset{(3.17)}{=} -\, R_{ebdc} \overset{\text{(raise)}}{=} -\, R^a{}_{bdc} \tag{3.2-7}$$

but we run into possible difficulty with property (3.16) for the curvature tensor:

$$R^a{}_{bcd} \neq R_a{}^b{}_{cd} \overset{\text{(raise)}}{=} g^{be}\, R_{aecd} \overset{(3.16)}{=} -\; g^{be}\, R_{eacd} \overset{\text{(raise)}}{=} -\, R^b{}_{acd}\,.$$

Claim: $R^a{}_{acd} = 0$: (3.19)

$$R^a{}_{acd} = g^{ba}\, R_{bacd} \;\text{ and }\; R^a{}_{acd} = g^{ab}\, R_{abcd} \overset{(3.16)}{=} -\, g^{ab}\, R_{bacd}$$

$$\Rightarrow \quad 2\, R^a{}_{acd} = g^{ba}\, R_{bacd} - g^{ab}\, R_{bacd} = (g^{ba} - g^{ab})\, R_{bacd} = 0 \qquad \checkmark$$

Problem 3.1 In a 2-dimensional Riemannian manifold

$R_{1212} = R_{2121} = -R_{1221} = -R_{2112}$ and $R_{ABCD} = 0$ otherwise.

Theorem 3.2.5 The covariant derivatives $R^a{}_{acd;e}$ satisfy the **Bianchi identity**

$$R^a{}_{bcd;e} + R^a{}_{bde;c} + R^a{}_{bec;d} = 0 : \tag{3.20}$$

Proof.

First, $R^a{}_{bcd} \overset{(3.13)}{=} \partial_c \Gamma^a_{bd} - \partial_d \Gamma^a_{bc} + \Gamma^e_{bd}\, \Gamma^a_{ec} - \Gamma^e_{bc}\, \Gamma^a_{ed}$.

So,

$$\begin{aligned} R^a{}_{bcd;e} &= \partial_e \partial_c \Gamma^a_{bd} - \partial_e \partial_d \Gamma^a_{bc} \\ &+ (\partial_e \Gamma^e_{bd})\, \Gamma^a_{ec} + \Gamma^e_{bd}\, \partial_e \Gamma^a_{ec} - (\partial_e \Gamma^e_{bc})\, \Gamma^a_{ed} - \Gamma^e_{bc}\, \partial_e \Gamma^a_{ed}\,. \end{aligned} \tag{3.2-8}$$

At any point P we can construct a coordinate system with $(\Gamma^a_{bc})_P = 0 \;\; \forall\, a, b, c$. However, this does not also mean that $(\partial_e \Gamma^a_{bc})_P = 0$. Hence,

$$R^a{}_{bcd;e} \overset{(3.2\text{–}8)}{=} (\partial_e \partial_c \Gamma^a_{bd} - \partial_e \partial_d \Gamma^a_{bc})_P \,.$$

Cyclically permuting c, d, and e yields equation (3.20). Since P is arbitrary, equation (3.2) holds everywhere. Since a tensor equation that equals zero in one coordinate system equals zero in every coordinate system, the Bianchi identity holds in all coordinate systems. ■

Even though $R^a{}_{acd} \overset{(3.19)}{=} 0$, in general $R^c{}_{abc} \neq 0$. Thus, we make the following definition.

Definition The **Ricci tensor** is the contraction

$$\boxed{R_{ab} \equiv R^c{}_{abc}}\,. \tag{3.21}$$

The Ricci tensor is symmetric in a and b (Exercise 3.2.4) :

$$R_{ab} = R_{ba} \tag{3.2-9}$$

$$0 \overset{(3.14)}{=} R^a{}_{bca} + R^a{}_{cab} + R^a{}_{abc} \overset{(3.16)}{=} R^a{}_{bca} + R^a{}_{cab} \overset{(3.21)}{=} R_{bc} + R^a{}_{cab}$$

$$\overset{(3.2\text{–}7)}{=} R_{bc} - R^a{}_{cba} \overset{(3.21)}{=} R_{bc} - R_{cb} \qquad ✓$$

Thus, $R^a{}_b = g^{ac} R_{cb} = g^{ac} R_{bc} = R_b{}^a$, and we define the **associated Ricci tensor**

$$R^a_b \equiv R^a{}_b = R_b{}^a \tag{3.2-10}$$

Another **associated Ricci tensor** is

$$R^{ab} \equiv g^{bc} R^a_c. \tag{3.2-11}$$

The (1,1) Ricci tensor can also be contracted from the (2,2) curvature tensor.

$$R^a_b = R^{ca}{}_{bc} : \tag{3.2-12}$$

$$R^{ca}{}_{bc} \overset{(3.2\text{–}4)}{=} g^{ae} R^c{}_{ebc} \overset{(3.21)}{=} g^{ae} R_{eb} = R^a_b \;\; ✓$$

Definition The **curvature scalar** is $\boxed{R \equiv R^a_a = g^{ab} R_{ba}}$ (3.22)

Problem 3.1 (continued from above) In a Riemannian manifold, using spherical coordinates for a sphere of radius a,

$$R_{1212} = a^2 sin^2\theta \quad \text{and} \quad R = -\frac{2}{a^2} .$$

Definition The **Einstein tensor** is $\boxed{G_{ab} \equiv R_{ab} - \frac{1}{2} R\, g_{ab}}$ (3.23)

The Einstein tensor is important because it is the LHS of Einstein's field equations (sans cosmological constant). It is symmetric since R and g are. The **associated Einstein tensors** are

$$\boxed{G^a_b = R^a_b - \frac{1}{2} R\, \delta^a_b} \tag{3.2-13}$$

and

$$\boxed{G^{ab} = G^a_c\, g^{cb} = R^{ab} - \frac{1}{2} R\, g^{ab}} : \tag{3.2-14}$$

$$G^a_b = g^{ac}\, G_{cb} = g^{ac}\,(R_{cb} - \tfrac{1}{2} R\, g_{cb}) = R^a_b - \tfrac{1}{2} R\, \delta^a_b \quad ✓$$
$$G^{ab} = G^a_c\, g^{cb} \;✓\; \text{and}\; G^a_c\, g^{cb} = (R^a_c - \tfrac{1}{2} R\, \delta^a_c)\, g^{cb} = R^{ab} - \tfrac{1}{2} R\, g^{ab} \quad ✓$$

Because G_{ab} is symmetric, it has only one divergence, $G^{ab}{}_{;a}$. Divergence is the magnitude of "fluid" flow across the boundary of a region. Because the RHS of Einstein's field equations have zero divergence, the following theorem provides support for Einstein's invention of G for the LHS.

Theorem 3.2.6 $G^{ab}{}_{;a} = 0$.

Proof. Contracting a with d (i.e., subscript $d \mapsto a$) in the Bianchi identity (3.20) gives

$$R^a{}_{bca;e} + R^a{}_{bae;c} + R^a{}_{bec;a} = 0, \text{ or} \tag{3.2-15}$$

We next generate expressions for the product of $g^{b\,e}$ with each of the three terms being differentiated in equation (3.2-15):

$$R^{a}{}_{b\,c\,a} \overset{(3.21)}{=} R_{b\,c} \quad \text{and} \quad R^{a}{}_{b\,a\,e} \overset{(3.2-7)}{=} - R^{a}{}_{b\,e\,a} \overset{(3.21)}{=} - R_{b\,e} \tag{3.2-16}$$

$$g^{b\,e}\, R^{a}{}_{b\,c\,a} \overset{(3.2-16)}{=} g^{b\,e}\, R_{b\,c} = R^{e}_{c} \tag{3.2-17}$$

$$g^{b\,e}\, R^{a}{}_{b\,a\,e} \overset{(3.2-16)}{=} - g^{e\,b}\, R_{b\,e} \overset{(3.22)}{=} - \mathrm{R} \tag{3.2-18}$$

$$g^{b\,e}\, R^{a}{}_{b\,e\,c} \overset{(3.2-4)}{=} R^{a\,e}{}_{e\,c} \overset{(3.2-6)}{=} R^{e\,a}{}_{c\,e} \overset{(3.2-12)}{=} R^{a}_{c} \tag{3.2-19}$$

In general, a variable like $g^{b\,e}$ cannot be moved in and out of a partial derivative like $R^{a}{}_{b\,c\,a\,;\,e}$. However, we can assume we are using a coordinate system at a point P with $\Gamma^{a}_{b\,c} = 0 \;\; \forall\, a, b, c$. So, then $g^{b\,e} = \eta^{b\,e}$, a constant (i.e., +1, -1, or 0), and we are free, below, to move it in and out.

That is,

$$0 \overset{(3.2-15)}{=} g^{b\,e}\,(R^{a}{}_{b\,c\,a\,;\,e} + R^{a}{}_{b\,a\,e\,;\,c} + R^{a}{}_{b\,e\,c\,;\,a}) \overset{(3.2-17,\,3.2-18,\,3.2-19)}{=} R^{e}_{c\,;\,e} - R_{;\,c} + R^{a}_{c\,;\,a}$$

$$= R^{b}_{c\,;\,b} - R_{;\,c} + R^{b}_{c\,;\,b} = 2R^{b}_{c\,;\,b} - \left(R\,\delta^{b}_{c}\right)_{;\,b} = (2R^{b}_{c} - R\,\delta^{b}_{c})_{;\,b}\,,$$

$$0 = (R^{b}_{c} - \tfrac{1}{2} R\,\delta^{b}_{c})_{;\,b} \overset{(3.2-13)}{=} G^{b}_{c\,;\,b}\,, \tag{3.2-20}$$

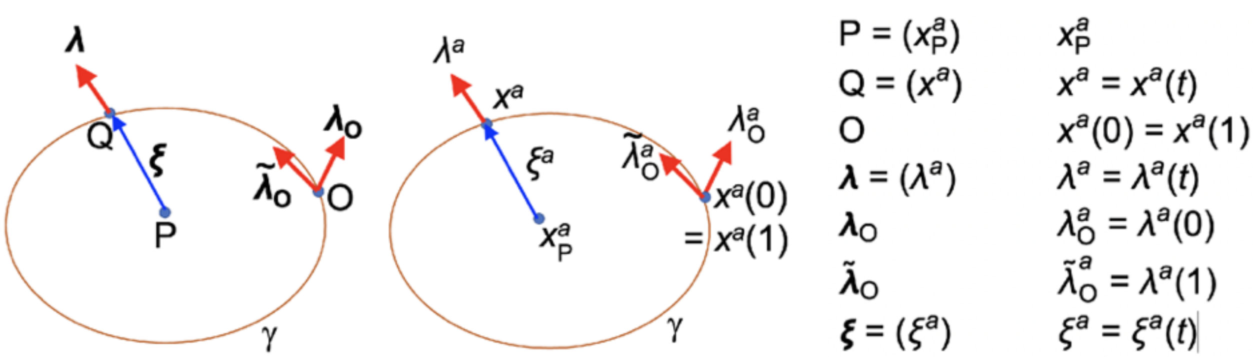
and

$$G^{b\,a}{}_{;\,b} = g^{\mathrm{ca}}\, G^{b}_{c\,;\,b} \overset{(3.2-20)}{=} 0. \tag{3.2-21}$$

Since P can be any point, divergence is zero everywhere. Finally, since divergence equaling zero is a tensor equation, it holds in every coordinate system. ■

3.3 Curvature and parallel transport

Parallel transport provides a method to compute the curvature components $R^a{}_{b\,c\,d}$. Specifically, in this section we develop an expression, equation (3.30), for $\Delta\lambda^a$ in terms of the curvature tensor $(R^a{}_{b\,c\,d})_{\mathrm{P}}$, where $\Delta\lambda^a$ is the change that results from parallelly transporting a vector λ^a around a small closed loop γ located about a point P in an N-manifold, $a = 1, \ldots, N$. This shows that $R^a{}_{b\,c\,d}$ is directly related to curvature effects and provides a justification for the name "curvature" tensor. We then illustrate how this equation can be used to find the components $R^a{}_{b\,c\,d}$.



Points and Vectors **Components** **Notation**

Let the curve γ be parameterized by t: $0 \le t \le 1$, let $\lambda^a_{\mathrm{O}} = \lambda^a(0)$ be the initial vector, anchored at point O, and let $\tilde{\lambda}^a_{\mathrm{O}}$ be the vector after it is moved around the curve and back to point O by parallel transport. By definition (2.23) of parallel transport of λ^a_{O},

$$\frac{d\lambda^a}{dt} = -\Gamma^a_{b\,c}\,\lambda^b\,\frac{dx^c}{dt}\ . \qquad (3.24)$$

By the Fundamental Theorem of Calculus, integrating equation (3.24) around the curve from point O to point Q gives

$$\lambda^a - \lambda^a_{\mathrm{O}} = \int_{\mathrm{O}}^{\mathrm{Q}} d\lambda^a \overset{(3.24)}{=} -\int_{\widehat{\mathrm{OQ}}} \Gamma^a_{b\,c}\,\lambda^b\,dx^c,$$

or

$$\lambda^a = \lambda^a_{\mathrm{O}} - \int_{\widehat{\mathrm{OQ}}} \Gamma^a_{b\,c}\,\lambda^b\,dx^c\ . \qquad (3.25)$$

Let $\boldsymbol{\xi}$ be a small vector extending from P to Q. That is,

$$\xi^a \equiv x^a - x^a_{\mathrm{P}} \quad \text{or} \quad x^a = x^a_{\mathrm{P}} + \xi^a. \qquad (3.3\text{-}1)$$

Then $dx^a \overset{(3.3-1)}{=} d\xi^a$ and, so,

$$\lambda^a \overset{(3.25)}{=} \lambda^a_O - \int_{\widehat{OQ}} \Gamma^a_{b\,c}\, \lambda^b\, d\xi^c \,. \tag{3.26}$$

Moving Q around the curve, from O around to O, gives

$$\tilde{\lambda}^a_O = \lambda^a_O - \oint \Gamma^a_{b\,c}\, \lambda^b\, d\xi^c \tag{3.3-2}$$

where $\tilde{\lambda}^a_O = \lambda^a(1)$, the vector at O after completing a circuit around γ.

Equation (3.3-2) provides a means for computing $\Delta\lambda^a = \tilde{\lambda}^a_O - \lambda^a_O$, but, unfortunately, λ^b is also on RHS of the equation. However, for small loops, $\lambda^b \approx \lambda^b_O$. This is a zeroth order approximation. We plug in λ^b_O for λ^b on RHS of equation (3.26) to get a first order approximation; then plug in that result for λ^b to get the next approximation, and so on. We will only need a second order approximation.

<u>Zeroth order approximation:</u> $\lambda^a = \lambda^a_O$

<u>First order approximation:</u>

$$\lambda^a = \lambda^a_O - \int_{\widehat{OQ}} \Gamma^a_{b\,c}\, \lambda^b_O\, d\xi^c = \lambda^a_O - \lambda^b_O \int_{\widehat{OQ}} \Gamma^a_{b\,c}\, d\xi^c$$

$$\Rightarrow \quad \lambda^b = \lambda^b_O - \lambda^d_O \int_{\widehat{OQ}} \Gamma^b_{d\,e}\, d\xi^e \tag{3.3-3}$$

<u>Second order approximation:</u>

Replacing λ^b in equation (3.3-2) by λ^b from equation (3.3-3) yields:

$$\tilde{\lambda}^a_O = \lambda^a_O - \oint \Gamma^a_{b\,c} \left[\, \lambda^b_O - \lambda^d_O \int_{\widehat{OQ}} \Gamma^b_{d\,e}\, d\xi^e \,\right] d\xi^c$$

$$\tilde{\lambda}^a_O = \lambda^a_O - \lambda^b_O \oint \Gamma^a_{b\,c}\, d\xi^c + \lambda^d_O \oint \Gamma^a_{b\,c} \int_{\widehat{OQ}} \Gamma^b_{d\,e}\, d\xi^e\, d\xi^c \tag{3.27}$$

$\Gamma^a_{b\,c}$ appears multiple time in equation (3.27), and we can express it as a Taylor series in order to generate 0th and 1st order approximations for it that we will then use to modify equation (3.27). To do this, recall from definition (2.13) that, though not obvious from the notation, Γ^a_{bc} is a function of x^a, x^b, and x^c. By freezing b and c, Γ^a_{bc} becomes a function of x^a: $f(x^a) \equiv \Gamma^a_{bc}$. Keep in mind that Γ^a_{bc} and its partial derivatives are defined on the manifold, not along the curve γ.

The Taylor's series we develop will be expanded about the point P = (x^a_P). Recall the Taylor series formula for a function $f(x^a) = f(x^1, ..., x^N)$ of several variables:

$$f(x^a) = f(x^a_P) + \frac{\partial}{\partial x^d} f(x^a_P)\,(x^d - x^d_P) + \frac{1}{2!}\frac{\partial^2}{\partial x^d\,\partial x^e} f(x^a_P)\,(x^d - x^d_P)\,(x^e - x^e_P) + \ldots .$$

Since $f(x^a) = \Gamma^a_{bc}$, then $f(x^a_P) = (\Gamma^a_{bc})_P$. Recalling that $\xi^d \overset{(3.3\text{–}1)}{=} x^d - x^d_P$, the Taylor series becomes

$$\Gamma^a_{bc} = (\Gamma^a_{bc})_P + \frac{\partial}{\partial x^d}(\Gamma^a_{bc})_P\,\xi^d + \frac{1}{2!}\frac{\partial^2}{\partial x^d\,\partial x^e}(\Gamma^a_{bc})_P\,\xi^d\,\xi^e + \ldots \qquad (3.3\text{-}4)$$

So, to the zeroth order, $\quad \Gamma^a_{bc} = (\Gamma^a_{bc})_P \qquad (3.3\text{-}5)$

and to the first order in ξ^d, $\quad \Gamma^a_{bc} = (\Gamma^a_{bc})_P + \frac{\partial}{\partial x^d}(\Gamma^a_{bc})_P\,\xi^d\ . \qquad (3.3\text{-}6)$

Examples, in general, of 1st order approximations are to approximate a circle by a line or a sphere by a plane. If we integrate a 1st order approximation over the circle or sphere, we ***improve***, achieving 2nd order approximations.

Before we substitute approximations for Γ^a_{bc} in equation (3.27), we make a couple of observations. Applying the Fundamental Theorem of Calculus to $\boldsymbol{\xi}$ over the full loop yields

$$\oint d\xi^c = \xi^c(1) - \xi^c(0) = 0. \qquad (3.3\text{-}7)$$

Applying the Fundamental Theorem over just the arc $\overset{\frown}{OQ}$ yields

$$\int_{\overset{\frown}{OQ}} d\xi^e = \xi^e(t) - \xi^e(0) = \xi^e - \xi^e(0)$$

$$\Rightarrow \quad \oint \int_{\overset{\frown}{OQ}} d\xi^e\, d\xi^c = \oint [\xi^e - \xi^e(0)]\, d\xi^c = [\, \oint \xi^e\, d\xi^c\,] - [\, \xi^e(0) \oint d\xi^c\,]$$

$$\overset{(3.3\text{–}7)}{=} \oint \xi^e\, d\xi^c \qquad (3.3\text{-}8)$$

Equation (3.27) is a 2nd order approximation. If we replace $\Gamma^a_{b\,c}$ in the 2nd term on RHS by a 1st order approximation, the 2nd term will remain part of a 2nd order approximation because the path integral “improves” the order of approximation. If we replace $\Gamma^a_{b\,c}$ in the 3rd term on RHS by a 0th order approximation, the 3rd term will remain part of a 2nd order approximation because of the double path integral. Specifically, for the second term,

$$\oint \Gamma^a_{b\,c}\, d\xi^c \;\mapsto\; (\Gamma^a_{b\,c})_P \oint d\xi^c + \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c \overset{(3.3\text{–}7)}{=} \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c$$

and for the third term,

$$\oint \Gamma^a_{b\,c} \int_{\overparen{OQ}} \Gamma^b_{d\,e}\, d\xi^e\, d\xi^c \;\mapsto\; (\Gamma^a_{b\,c})_P\, (\Gamma^b_{d\,e})_P \oint \int_{\overparen{OQ}} d\xi^e\, d\xi^c \overset{(3.3\text{–}8)}{=} (\Gamma^a_{b\,c})_P\, (\Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c.$$

Thus, the 2nd order approximation, equation (3.27), becomes

$$\tilde{\lambda}^a_O = \lambda^a_O - \lambda^b_O \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c + \lambda^d_O\, (\Gamma^a_{b\,c}\, \Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c,$$

and, to a 2nd order approximation,

$$\Delta\lambda^a = -\,\lambda^b_O\, [\, \partial_c(\Gamma^a_{b\,d})_P - (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P\,] \oint \xi^c\, d\xi^d : \qquad (3.28)$$

$$\begin{aligned}
\Delta\lambda^a = \tilde{\lambda}^a_O - \lambda^a_O &= -\,\lambda^b_O\, \partial_d(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c + \lambda^d_O\, (\Gamma^a_{b\,c}\, \Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c \\
&= -\,\lambda^b_O\, \partial_c(\Gamma^a_{b\,d})_P \oint \xi^c\, d\xi^d + \lambda^b_O\, (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P \oint \xi^c\, d\xi^d \\
&= -\,\lambda^b_O\, [\, \partial_c(\Gamma^a_{b\,d})_P - (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P\,] \oint \xi^c\, d\xi^d . \quad \checkmark
\end{aligned}$$

Claim $\oint \xi^c\, d\xi^d = \frac{1}{2} \oint [\, \xi^c\, d\xi^d - \xi^d\, d\xi^c\,]$: (3.3-9)

(Exercise 3.3.1, part 1) By the Fundamental Theorem of Calculus,

$$\oint d(\xi^c \xi^d) = [\xi^c \xi^d]\,(1) - [\xi^c \xi^d]\,(0) = \xi^c(1)\, \xi^d(1) - \xi^c(0)\, \xi^d(0) = 0.$$

So,

$$0 = \oint d(\xi^c \xi^d) = \oint \xi^d\, d\xi^c + \oint \xi^c\, d\xi^d \qquad (3.3\text{-}10)$$

$$\Rightarrow \qquad 2 \oint \xi^c\, d\xi^d = \oint [\, \xi^c\, d\xi^d - \xi^d\, d\xi^c\,] \qquad \checkmark$$

Define a tensor $f^{c\,d}$ by

$$\boxed{f^{c\,d} \equiv \oint \xi^c\, d\,\xi^d \overset{(3.3\text{–}9)}{=} \frac{1}{2} \oint \left[\xi^c\, d\,\xi^d - \xi^d\, d\,\xi^c\right]} . \qquad (3.29)$$

Observe that f^{cd} is antisymmetric; i.e., $f^{cd} = -f^{dc}$. (3.3-11)

Claim that the 2nd order approximation equation (3.28) can now be expressed as

$$\boxed{\Delta\lambda^a = -\tfrac{1}{2}\,(R^a{}_{bcd})_P\,\lambda^b_O\,f^{cd}} \tag{3.30}$$

where f^{cd} is defined in equation (3.29):

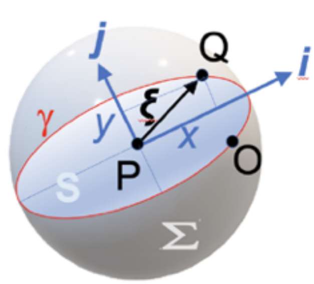
$$-\,(\partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd}$$
$$\overset{(3.3\text{-}11)}{=}\ -\frac{1}{2}(\partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd} + \frac{1}{2}(\partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{dc}$$
$$\overset{(c\leftrightarrow d)}{=}\ -\frac{1}{2}(\partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd} + \frac{1}{2}(\partial_d\Gamma^a_{bc} - \Gamma^a_{ec}\,\Gamma^e_{bd})_P\,f^{cd}$$

$$-\,(\partial_c\Gamma^a_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd} = -\frac{1}{2}(\partial_c\Gamma^a_{bd} - \partial_d\Gamma^a_{bc} + \Gamma^a_{ec}\,\Gamma^e_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd} \tag{3.3-12}$$

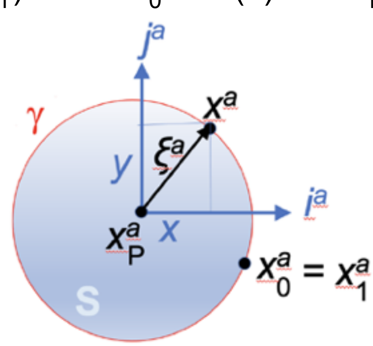
$$\Delta\lambda^a \overset{(3.28,\,3.29)}{=} -\,\lambda^b_O\,[\,\partial_c(\Gamma^a_{bd})_P - (\Gamma^a_{ed}\,\Gamma^e_{bc})_P\,]\,f^{cd}$$
$$\overset{(3.3\text{-}12)}{=}\ -\,\lambda^b_O\left(\tfrac{1}{2}\right)(\partial_c\Gamma^a_{bd} - \partial_d\Gamma^a_{bc} + \Gamma^a_{ec}\,\Gamma^e_{bd} - \Gamma^a_{ed}\,\Gamma^e_{bc})_P\,f^{cd}$$
$$\overset{(3.13)}{=}\ -\frac{1}{2}\,(R^a{}_{bcd})_P\,\lambda^b_O\,f^{cd} \qquad \checkmark$$

This concludes development of equation (3.30), which we now show can be exploited under certain conditions to solve for the components of the curvature tensor $R^a{}_{bcd}$. We show this by developing equation (3.32), a geometric variation of equation (3.30), and then computing the components, Example 3.3.1 below.

Let P = (x^a_P) be a point in an N-manifold M. Consider a small sphere Σ of radius ρ about P, shown below. Form a loop γ by intersecting a plane containing P with Σ. Let the loop be parameterized by t, $0 \le t \le 1$, so that points of γ are expressed as $(x^a(t))$. Let the loop begin and end at point O = (x^a_0) = (x^a_1) where $x^a_0 = x^a(0)$ and $x^a_1 = x^a(1)$.



Points and Vectors



Components

Consider a pair of orthogonal unit vectors {***i***, ***j***} in the disk S enclosed by γ and anchored at P . If the manifold M is flat, then {***i***, ***j***} constitutes a Cartesian coordinate system for S. If M is curved, then {***i***, ***j***} is approximately a Cartesian coordinate system for nearby points. A Cartesian coordinate system is a 1st order (linear) approximation. In particular, since γ is the boundary of S, points of γ can be expressed approximately to 1st order by

$$Q \approx P + x\,\boldsymbol{i} + y\,\boldsymbol{j}$$

where $x^2 + y^2 = \rho^2$. Note that x and y are functions of t since Q is a function of t. Also ***i*** and ***j*** are fixed vectors. We express $\boldsymbol{i} = (\imath^a)$ and $\boldsymbol{j} = (j^a)$, with $\imath$ in a different font to clearly distinguish it from j). The equation above can now be expressed as

$$(x^a) \approx (x^a_{\mathrm{P}}) + x\,(\imath^a) + y\,(j^a).$$

In terms of components, the curve can be expressed

$$x^a \approx x^a_{\mathrm{P}} + x\,\imath^a + y\,j^a$$

where $\imath^a$ and j^a are constants.

Let $\boldsymbol{\xi}$ be a (small) vector from P to a point Q = (x^a) of the curve γ. In terms of components,

$$\xi^a = x^a - x^a_{\mathrm{P}} \approx x\,\imath^a + y\,j^a\,.$$

From equation (3.29),

$$f^{c\,d} = \tfrac{1}{2}\oint[\xi^c \mathrm{d}\xi^d - \xi^d \mathrm{d}\xi^c].$$

Claim $\xi^c \mathrm{d}\xi^d - \xi^d \mathrm{d}\xi^c = (x\,\mathrm{d}y - y\,\mathrm{d}x)\,(\imath^c j^d - \imath^d j^c)$:

(Exercise 3.3.1, part 2)

$$\begin{aligned}
\text{LHS} &= (x\,\imath^c + y\,j^c)\,(\mathrm{d}x\,\imath^d + \mathrm{d}y\,j^d) - (x\,\imath^d + y\,j^d)\,(\mathrm{d}x\,\imath^c + \mathrm{d}y\,j^c)\\
&= x\,\mathrm{d}x\,(\imath^c\imath^d - \imath^d\imath^c) + y\,\mathrm{d}y\,(j^c j^d - j^d j^c)\\
&\quad + x\,\mathrm{d}y\,(\imath^c j^d - \imath^d j^c) - y\,\mathrm{d}x\,(j^d\imath^c - j^c\imath^d)\\
&= (x\,\mathrm{d}y - y\,\mathrm{d}x)\,(\imath^c j^d - j^c\imath^d) \quad \checkmark
\end{aligned}$$

Since the 1st order Cartesian approximation undergoes a path integral, it "improves" to a 2nd order approximation:

$$f^{c\,d} \overset{(3.29)}{=} \tfrac{1}{2}\oint[\,(x\,\mathrm{d}y - y\,\mathrm{d}x)\,(\imath^c j^d - \imath^d j^c)\,]$$

$$f^{c\,d} = \tfrac{1}{2}\,(\imath^c j^d - \imath^d j^c)\oint(x\,\mathrm{d}y - y\,\mathrm{d}x) \tag{3.31}$$

Now recall Green's Theorem (for Cartesian coordinates):

$$\oint_\gamma L\,dx + M\,dy = \int_S \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y}\right) dA$$

where S is the disc enclosed by the loop γ, A is the area of S, and L and M are continuously differentiable functions of x and y.

For $M = x$ and $L = 0$, $\oint_\gamma x\,dy = \int_S dA = A$

For $M = 0$ and $L = y$, $\oint_\gamma y\,dx = -\int_S dA = -A$

$$\Rightarrow \quad \tfrac{1}{2} \oint (x\,dy - y\,dx) = A \qquad (3.3\text{-}13)$$

On the (curved) manifold M, equation (3.3-13) is a 2nd order approximation (because it is a locally Euclidean, 1st order, approximation within a path integral). So, to the 2nd order we have

$$f^{c\,d} \overset{(3.31)}{=} A\,(i^c j^d - i^d j^c).$$

Since

$$\Delta\lambda^a \overset{(3.30)}{=} -\tfrac{1}{2}\,(R^a{}_{b\,c\,d})_P\,\lambda^b_O\,f^{c\,d}$$

to the 2nd order, then replacing $f^{c\,d}$ with a 2nd order approximation still yields a 2nd order approximation:

$$\frac{\Delta\lambda^a}{A} = -\,(R^a{}_{b\,c\,d})_P\,\lambda^b_O\,i^c j^d : \qquad (3.32)$$

$$\begin{aligned}
\Delta\lambda^a &= -\tfrac{1}{2}\,(R^a{}_{b\,c\,d})_P\,\lambda^b_O\,(i^c j^d - i^d j^c) \\
&= -\tfrac{1}{2}\,\lambda^b_O\,[\,(R^a{}_{b\,c\,d})_P\,i^c j^d - (R^a{}_{b\,c\,d})_P\,i^d j^c\,] \\
&= -\tfrac{1}{2}\,\lambda^b_O\,[\,(R^a{}_{b\,c\,d})_P\,i^c j^d - (R^a{}_{b\,d\,c})_P\,i^c j^d\,] \\
&\overset{(3.17)}{=} -\tfrac{1}{2}\,\lambda^b_O\,[\,(R^a{}_{b\,c\,d})_P\,i^c j^d - (R^a{}_{b\,c\,d})_P\,i^d j^c\,] \\
&= -\tfrac{1}{2}\,\lambda^b_O\,(R^a{}_{b\,c\,d})_P\,[\,i^c j^d + i^c j^d\,] \\
&= -\,(R^a{}_{b\,c\,d})_P\,\lambda^b_O\,i^c j^d \qquad \checkmark
\end{aligned}$$

In Problem 3.1 of Section 3.2 we showed how to compute the components of the curvature tensor R_{abcd} using the raw definitions, with the result that all non-zero components equal $\pm R_{1212}$. The next example shows another technique for calculating R_{abcd} that uses equation (3.32) along with judicious choices for i^a, j^a, and the parallel transport vector λ_O^a.

Example 3.3.1 Let P be the point with coordinates (θ_0,ϕ_0) on a sphere of radius a, and let $\boldsymbol{\lambda}$ be a vector at P pointing south. We calculate $\Delta\boldsymbol{\lambda}$, the change in $\boldsymbol{\lambda}$ from parallelly transporting it around a small, closed loop about P.

Let the loop γ be the sections of the latitude circles $\theta = \theta_0 \pm \epsilon$ and longitude circles $\phi = \phi_0 \pm \epsilon$ as shown in Figure 3.2.

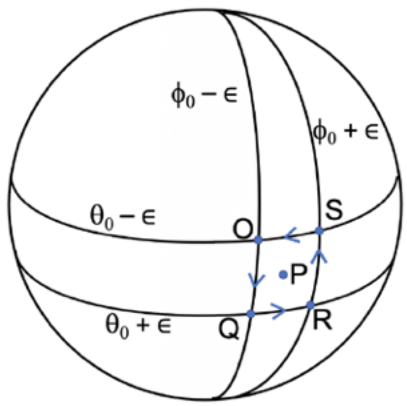


Figure 3.2

In Example 2.21 of Section 2.1 we learned that if a vector, located at a point (θ,ϕ), points α radians east of south, then the vector has components

$$\boxed{\lambda^1 = \frac{\cos\alpha}{a} \text{ and } \lambda^2 = \frac{\sin\alpha}{a\sin\theta}}. \tag{2.25}$$

We also learned that if such a vector is parallely transported along a latitude arc from (θ,ϕ) to $(\theta, \phi + \Delta\phi)$, then the final vector will have components

$$\boxed{\lambda^1 = \frac{\cos(\alpha-\omega)}{a} \text{ and } \lambda^2 = \frac{\sin(\alpha-\omega)}{a\sin\theta}}, \text{ where } \omega = \Delta\phi\cos\theta. \tag{2.2-10}$$

Of course, parallely transported vectors along longitude arcs don't change because longitudes are great circles, geodesics.

Start at point O with a unit vector $\boldsymbol{\lambda}_O$ pointing south and parallelly transport it from O to Q, R, S, and back to O. The coordinates of the initial south-pointing vector at O are

$$\lambda_O^1 = \frac{1}{a} \text{ and } \lambda_O^2 = 0,$$

$\boldsymbol{\lambda}$ arrives at, and then exits, point Q = (θ_0-ϵ, ϕ_0-ϵ) still pointing south. So,

$$\lambda_Q^1 = \frac{1}{a} \text{ and } \lambda_Q^2 = 0.$$

The inputs to equation (2.2-10) at point R = (θ_0+ϵ, ϕ_0+ϵ) are $\alpha = 0$, $\theta = \theta_0+\epsilon$, $\Delta\phi = 2\epsilon$, and $\omega = 2\epsilon \cos(\theta_0+\epsilon)$. So, the vector arrives at, and departs, R with

$$\lambda_R^1 = \frac{\cos[-2\,\epsilon \cos(\theta_0+\epsilon)]}{a} \text{ and } \lambda_R^2 = \frac{\sin[-2\,\epsilon \cos(\theta_0+\epsilon)]}{a \sin(\theta_0+\epsilon)}.$$

Next we transport along longitude arc $\widehat{RS}$, where the vector does not change:

$$\lambda_S^1 = \frac{\cos[-2\,\epsilon \cos(\theta_0+\epsilon)]}{a} \text{ and } \lambda_S^2 = \frac{\sin[-2\,\epsilon \cos(\theta_0+\epsilon)]}{a \sin(\theta_0-\epsilon)}.$$

We return to point O = (θ_0–ϵ, ϕ_0–ϵ) where the inputs to equation (2.2-10) are $\alpha = -2\,\epsilon \cos(\theta_0+\epsilon)$, $\theta = \theta_0 - \epsilon$, $\Delta\phi = -2\epsilon$, and $\omega = -2\epsilon \cos(\theta_0 - \epsilon)$. So,

$$\begin{aligned}\alpha - \omega &= -2\,\epsilon \cos(\theta_0+\epsilon) + 2\epsilon \cos(\theta_0-\epsilon)\\ &= -2\epsilon\,(\cos\theta_0 \cos\epsilon - \sin\theta_0 \sin\epsilon) + 2\epsilon\,(\cos\theta_0 \cos\epsilon + \sin\theta_0 \sin\epsilon)\\ &= 4\epsilon \sin\theta_0 \sin\epsilon,\end{aligned}$$

and the final vector is

$$\tilde{\lambda}_O^1 = \frac{\cos(4\,\epsilon \sin\theta_0 \sin\epsilon)}{a} \text{ and } \tilde{\lambda}_O^2 = \frac{\sin(4\,\epsilon \sin\theta_0 \sin\epsilon)}{a \sin(\theta_0-\epsilon)}.$$

To use equation (3.32) to approximate the curvature tensor, we need a 2nd order approximation for $\Delta\boldsymbol{\lambda}_P$ as well as for the area A enclosed by the 4 arcs, and we will need to choose a pair of orthogonal unit vectors {$\boldsymbol{i}$, $\boldsymbol{j}$}.

Recall that the 2nd order Taylor series approximations for sine and cosine are

$$\sin\theta = \theta \text{ and } \cos\theta = 1 - \tfrac{1}{2}\theta^2.$$

From this we get additional 2nd order approximations:

$sin(\theta_0 - \epsilon) = \theta_0 - \epsilon\ = \theta_0, \qquad 4\epsilon\ sin\ \theta_0\ sin\ \epsilon = 4\epsilon^2\theta_0,$

$sin(4\epsilon\ sin\theta_0\ sin\ \epsilon)\ = sin(4\epsilon^2\theta_0) = 4\epsilon^2\theta_0,$

$cos(4\epsilon\ sin\theta_0\ sin\ \epsilon) = cos(4\epsilon^2\theta_0) = 1 + \frac{1}{2}\ (16\epsilon^4)\ \theta_0^2 = 1.$

Thus, to the 2nd order in ϵ, (Exercise 3.3.3)

$$\Delta\lambda^1 \approx \Delta\lambda_O^1 = \tilde{\lambda}_O^1 - \lambda_O^1 = \frac{cos\,(4\,\epsilon\,sin\,\theta_0\,sin\,\epsilon\,)}{a} - \frac{1}{a} = \frac{1}{a} - \frac{1}{a} = 0,$$

$$\Delta\lambda^2 \approx \Delta\lambda_O^2 = \tilde{\lambda}_O^2 - \lambda_O^2 = \frac{sin\,(4\,\epsilon\,sin\,\theta_0\,sin\,\epsilon\,)}{a\,sin\,(\theta_0 - \epsilon)} - 0 = \frac{4\,\epsilon^2\,\theta_0}{a\,\theta_0} = \frac{4\,\epsilon^2}{a};$$

i.e.,

$$\Delta\lambda^a = \frac{4\,\epsilon^2}{a}\,\delta_2^a. \tag{3.3-14}$$

We next generate a 2nd order approximation for the area A enclosed by the 4 arcs. While the equator has radius a, a circle at latitude $\hat{\theta} \neq \frac{\pi}{2}$ has a smaller radius, namely r = a $sin\hat{\theta}$. For latitude arc $\overset{\frown}{\mathrm{QR}}$, r = a $sin(\theta_0+\epsilon)$ and the angle between the center of the sphere and points Q and R is $\theta \approx 2\epsilon$. Then $s \equiv r\theta = 2\epsilon\ a\ sin(\theta_0+\epsilon)$. For latitude arc $\overset{\frown}{\mathrm{OS}}$, $s = 2\epsilon\ a\ sin(\theta_0 - \epsilon)$. So the lengths of the top and bottom of the "square" enclosed by the four arcs is $s = 2\epsilon\ a\ sin\ \theta_0$ to a 1st order approximation.

Every longitude circle has radius a, so the sides of the "square" have length $s = 2\epsilon a$, also to a 1st order approximation. Consequently, the area is

$$\mathrm{A} = (2\epsilon\ a\ sin\ \theta_0)\ (2\epsilon a) = 4\epsilon^2 a^2 sin\ \theta_0, \tag{3.3-15}$$

a 2nd order approximation in ϵ.

Finally, a simple choice for ***i*** and ***j*** are the unit vectors in the tangent directions $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$. For spherical coordinates, from Example 1.1.4,

$$\mathbf{e}_\theta = a\ cos\theta\ cos\phi\ \boldsymbol{i} + a\ cos\theta\ sin\phi\ \boldsymbol{j} - a\ sin\theta\ \boldsymbol{k} \tag{3.3-16}$$

$$\mathbf{e}_\phi = -\ a\ sin\theta\ sin\phi\ \boldsymbol{i} + a\ sin\theta\ cos\phi\ \boldsymbol{j}\ . \tag{3.3-17}$$

So, $|\mathbf{e}_\theta| = a$ and $|\mathbf{e}_\phi| = a\ sin\ \theta_0$. We select orthonormal unit vectors

$$i^1 = \frac{1}{a}, \quad i^2 = 0, \quad j^1 = 0, \quad \text{and} \quad j^2 = \frac{1}{a\,sin\,\theta_0},$$

or

$$i^a = \frac{1}{a}\delta_1^a \quad \text{and} \quad j^a = \frac{1}{a\,sin\,\theta_0}\delta_2^a. \tag{3.3-18}$$

We now have all the pieces we need to use equation (3.32) to find $(R^a{}_{b\,c\,d})_P$.

$$\frac{\Delta\lambda^a}{A} \overset{(3.3-14,\,3.3-15)}{=} \frac{4\,\epsilon^2}{a}\delta_2^a \frac{1}{4\,\epsilon^2\,a^2\,sin\,\theta_0} = \delta_2^a \frac{1}{a^3\,sin\,\theta_0}. \tag{3.3-19}$$

The vector, $\boldsymbol{\lambda}$, we have been transporting started at point O and points south. So, $\lambda_O^b = \frac{1}{a}\delta_1^b$, and

$$- (R^a{}_{b\,c\,d})_P\,\lambda_O^b\,i^c j^d \overset{(3.3-18)}{=} - (R^a{}_{b\,c\,d})_P\,\frac{1}{a}\delta_1^b\,\frac{1}{a}\delta_1^c\,\frac{1}{a\,sin\,\theta_0}\delta_2^d$$

$$= - (R^a{}_{112})_P\,\frac{1}{a^3\,sin\,\theta_0} \tag{3.3-20}$$

$$\delta_2^a\,\frac{1}{a^3\,sin\,\theta_0} \overset{(3.3-19)}{=} \frac{\Delta\lambda^a}{A} \overset{(3.32)}{=} - (R^a{}_{b\,c\,d})_P\,\lambda_O^b\,i^c j^d \overset{(3.3-20)}{=} - (R^a{}_{112})_P\,\frac{1}{a^3\,sin\,\theta_0}$$

$$\Rightarrow \quad (R^a{}_{112})_P = -\,\delta_2^a \tag{3.3-21}$$

$$\Leftrightarrow \quad (R^1{}_{112})_P = 0 \;\; \text{and} \quad (R^2{}_{112})_P = -1 \tag{3.3-22}$$

This enables us to compute all of the Curvature tensor components:

From 3.3-16 and 3.3-17:

$$g_{11} = \mathbf{e}_{\theta_0} \bullet \mathbf{e}_{\theta_0} = a^2 \tag{3.3-23}$$

$$g_{12} = g_{21} = \mathbf{e}_{\theta_0} \bullet \mathbf{e}_{\phi_0} = 0 \tag{3.3-24}$$

$$g_{22} = \mathbf{e}_{\phi_0} \bullet \mathbf{e}_{\phi_0} = a^2\,sin^2\,\theta_0 \tag{3.3-25}$$

$$(R_{a\,112})_P = g_{a\,e}\,(R^e{}_{112})_P \overset{(3.3-21)}{=} - g_{a\,e}\,\delta_2^e = - g_{a\,2}$$

For $a = 2$, $(R_{1212})_P = - (R_{2112})_P = g_{22} \overset{(3.3-25)}{=} a^2\,sin^2\,\theta_0$.

By Problem 3.1, $(R_{1212})_P = (R_{2121})_P = -(R_{2112})_P = -(R_{1221})_P = g_{22}$, and the remaining components are zero.

For $a = 1$, $(R_{1112})_P = g_{21} = 0$, which is in agreement with the remaining components being zero. ■

The following observations are not found in the book. $R_{1212} = a^2\,sin^2\,\theta_0$ is interesting because in classical physics for a sphere of radius a, the (scalar) curvature is $\kappa = \frac{1}{a}$. So, R_{1212} appears to be related. Pursuing this further:

$$\mathbf{e}^{\theta} = \frac{\cos\theta\ \cos\phi}{a}\ \boldsymbol{i} + \frac{\cos\theta\ \sin\phi}{a}\ \boldsymbol{j} - \frac{\sin\theta}{a}\ \boldsymbol{k}$$

$$\mathbf{e}^{\phi} = -\frac{\sin\phi}{a\sin\theta}\ \boldsymbol{i} + \frac{\cos\phi}{a\sin\theta}\ \boldsymbol{j}$$

$$g^{11} = \mathbf{e}^{\theta_0} \bullet \mathbf{e}^{\theta_0} = \frac{1}{a^2}$$

$$g^{12} = g^{21} = \mathbf{e}^{\theta_0} \bullet \mathbf{e}^{\phi_0} = 0$$

$$g^{22} = \mathbf{e}^{\phi_0} \bullet \mathbf{e}^{\phi_0} = \frac{1}{a^2 \sin^2 \theta_0}$$

$$R_{a\,b} \overset{(3.21)}{=} R^{c}_{\ a\,b\,c} = g^{e\,c}\, R_{e\,a\,b\,c}$$

$$\Rightarrow \quad R_{11} = g^{e\,c} R_{e11c}$$

The only non-zero value of R_{e11c} occurs when $e = c = 2$:

$$R_{11} = g^{22} R_{2112} = \frac{1}{a^2 \sin^2 \theta_0}\ (-\, a^2 \sin^2 \theta_0) = -\,1$$

The only non-zero value of R_{e12c} occurs when $e = 2$ and $c = 1$:

$$R_{12} = g^{e\,c} R_{e12c} = g^{21} R_{2121} = (0)\ (a^2 \sin^2 \theta_0) = 0 \ \text{ and } \ R_{21} = 0$$

The only non-zero value of R_{e22c} occurs when $e = c = 1$:

$$R_{22} = g^{e\,c} R_{e22c} = g^{11} R_{1221} = \frac{1}{a^2}\ (-\, a^2 \sin^2 \theta_0) = -\sin^2 \theta_0$$

$$R \overset{(3.22)}{=} g^{a\,b}\, R_{a\,b} = g^{11}\, R_{11} + g^{22}\, R_{22} = \frac{1}{a^2}\ (-1) + \frac{1}{a^2 \sin^2 \theta_0}\ (-\sin^2 \theta_0) = -\,\frac{2}{a^2}.$$

So, though they may be related, κ and R are not the same. Moreover, it is useful to have an example where we have values of all the curvature tensors and associated tensors. To this end we include the following tensor values:

$R^{a\,b} = g^{b\,c}\, R^{a}_{c} = g^{a\,d} g^{b\,c}\, R_{d\,c}$:

$R_{12} = R_{21} = 0$, so we need only consider terms with R_{11} and R_{22}.

$$R^{11} = g^{1\,d} g^{2\,c}\, R_{d\,c} = g^{11} g^{21}\, R_{11} + g^{12} g^{22}\, R_{22} = \frac{1}{a^4}(-1) + 0 = -\frac{1}{a^4}$$

$$R^{12} = g^{1\,d} g^{1\,c}\, R_{d\,c} = g^{11} g^{11}\, R_{11} + g^{12} g^{12}\, R_{22} = 0 \ \text{ and } \ R^{21} = 0$$

$$R^{22} = g^{2\,d} g^{2\,c}\, R_{d\,c} = g^{21} g^{21}\, R_{11} + g^{22} g^{22}\, R_{22} = 0 + \frac{1}{a^4 \sin^4 \theta_0}(-\sin^2 \theta_0)$$

$$= -\frac{1}{a^4 \sin^2 \theta_0}$$

$R^a_b = g^{a\,c} R_{c\,b}$:

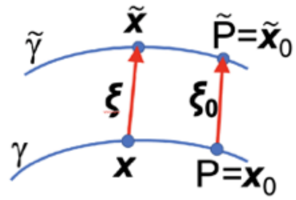
Again, we only need to consider terms with R_{11} and R_{22}.

$R^1_1 = g^{11} R_{11} = \frac{1}{a^2}\,(-1) = -\frac{1}{a^2}$

$R^1_2 = g^{12} R_{22} = 0$ and $R^2_1 = g^{21} R_{11} = 0$

$R^2_2 = g^{22} R_{22} = \frac{1}{a^2 \sin^2 \theta_0}\,(-\sin^2 \theta_0) = -\frac{1}{a^2}$

3.4 Geodesic Deviation



Another layer of support for Einstein's field equations come from the equation of geodesic deviation, 3.35. We derive the equation and then show that it reduces to its Euclidean counterpart.

Let P be a point in a manifold M. Since M is locally Euclidean, there is an open neighborhood U about P that is homeomorphic to Euclidean space. Thus, P has a Euclidean coordinate system (x^a) defined on U. Let $\tilde{\text{P}}$ be a different point in U. Then $\tilde{\text{P}}$ also has a Euclidean coordinate system, $(\tilde{x}^a)$, defined on U.

Let γ be a geodesic curve through P that is contained in U, and let $\tilde{\gamma}$ be a geodesic curve through $\tilde{\text{P}}$ that is also contained in U. Let both curves be parameterized by u such that P = $\boldsymbol{x}_0(0)$ and $\tilde{\text{P}} = \tilde{\boldsymbol{x}}_0(0)$. Denote $\boldsymbol{x}_0 = \boldsymbol{x}_0(0)$ and $\tilde{\boldsymbol{x}}_0 = \tilde{\boldsymbol{x}}_0(0)$. Typical points on γ and $\tilde{\gamma}$ are denoted $\boldsymbol{x} = \boldsymbol{x}(u)$ and $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}(u)$, respectively. One of the curves may encompass a larger range for u than the other, so we restrict ourselves to a range of u that is valid for both curves.

Define a "vector" $\boldsymbol{\xi} = \boldsymbol{\xi}(u) = \tilde{\boldsymbol{x}} - \boldsymbol{x}$ and denote $\boldsymbol{\xi}_0 = \tilde{\boldsymbol{x}}_0 - \boldsymbol{x}_0$. ($\boldsymbol{\xi}$ is not actually a vector because it does not lie in the tangent plane at either $\boldsymbol{x}$ or $\tilde{\boldsymbol{x}}$.) Recall that curve γ is null if $g_{ab}\,\dot{x}^a\,\dot{x}^b = 0$ at any point, where "dot" refers to differentiation by u. If neither γ nor $\tilde{\gamma}$ is null, then by Exercise 2.1.3 we can let u be an affine parameter $u = As + B$, where s is arc length. If γ is null then for some x^a, then $ds \overset{(1.83)}{=} \sqrt{\left|\,g_{ab}\,\dot{x}^a\,\dot{x}^b\,\right|} = 0$

$\Rightarrow$ s is a constant, not suitable as a parameter on γ.

Since γ and $\tilde{\gamma}$ are geodesics, they satisfy the geodesic equation (2.72):

$$\frac{d^2\,\tilde{x}^a}{du^2} + \tilde{\Gamma}^a_{b\,c}\,\frac{d\,\tilde{x}^b}{du}\,\frac{d\,\tilde{x}^c}{du} = 0 \quad \text{or} \quad \ddot{\tilde{x}}^a + \tilde{\Gamma}^a_{b\,c}\,\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c = 0 \qquad (3.33)$$

$$\frac{d^2\,x^a}{du^2} + \Gamma^a_{b\,c}\,\frac{dx^b}{du}\,\frac{dx^c}{du} = 0 \quad \text{or} \quad \ddot{x}^a + \Gamma^a_{b\,c}\,\dot{x}^b\,\dot{x}^c = 0 \qquad (3.34)$$

where

$$\tilde{\Gamma}^{a}_{b\,c} \equiv (\Gamma^{a}_{b\,c})_{\tilde{x}} \quad \text{and} \quad \Gamma^{a}_{b\,c} \equiv (\Gamma^{a}_{b\,c})_{\mathbf{x}}.$$

We next derive equations labeled (3.4-2) – (3.4-5), and (3.35). These derivations constitute Exercise 3.4.1 in the book.

For $\xi^a = \tilde{x}^a - x^a$, it was shown in Section 3.3 that to the 1st order in ξ^a,

$$\tilde{\Gamma}^{a}_{b\,c} \overset{(3.3-6)}{=} \Gamma^{a}_{b\,c} + \left(\frac{\partial}{\partial x^d}\Gamma^{a}_{b\,c}\right) \xi^d = \Gamma^{a}_{b\,c} + \Gamma^{a}_{b\,c,\,d}\, \xi^d.$$

So, equation (3.33) can be rewritten as

$$\ddot{\tilde{x}}^a + \left(\Gamma^{a}_{b\,c} + \Gamma^{a}_{b\,c,\,d}\, \xi^d\right) \dot{\tilde{x}}^b\, \dot{\tilde{x}}^c = 0 \qquad (3.4\text{-}1)$$

Also,

$$\tilde{x}^a = x^a + \xi^a, \quad \dot{\tilde{x}}^a = \dot{x}^a + \dot{\xi}^a\,, \ \text{and} \quad \ddot{\tilde{x}}^a = \ddot{x}^a + \ddot{\xi}^a.$$

Subtracting (3.34) from (3.4-1) yields

$$\ddot{\tilde{x}}^a - \ddot{x}^a + \Gamma^{a}_{b\,c}\,(\dot{\tilde{x}}^b\, \dot{\tilde{x}}^c - \dot{x}^b\, \dot{x}^c) + \Gamma^{a}_{b\,c,\,d}\, \xi^d\, \dot{\tilde{x}}^b\, \dot{\tilde{x}}^c = 0. \qquad (3.4\text{-}2)$$

Claim: equation (3.4-2) can be expressed to the 1st order in ξ^a as

$$\ddot{\xi}^a + \Gamma^{a}_{b\,c,\,d}\, \dot{x}^b\, \dot{x}^c\, \xi^d + \Gamma^{a}_{b\,c}\,(\dot{x}^b \dot{\xi}^c + \dot{\xi}^b \dot{x}^c) = 0: \qquad (3.4\text{-}3)$$

$$\ddot{\tilde{x}}^a - \ddot{x}^a = \ddot{\xi}^a$$

$$\Gamma^{a}_{b\,c}\,(\dot{\tilde{x}}^b\, \dot{\tilde{x}}^c - \dot{x}^b\, \dot{x}^c) = \Gamma^{a}_{b\,c}\,[\,(\dot{x}^b + \dot{\xi}^b)\,(\dot{x}^c + \dot{\xi}^c) - \dot{x}^b\, \dot{x}^c] = \Gamma^{a}_{b\,c}\,(\dot{x}^b \dot{\xi}^c + \dot{\xi}^b\, \dot{x}^c)$$

$$\begin{aligned}\Gamma^{a}_{b\,c,\,d}\, \xi^d\, \dot{\tilde{x}}^b\, \dot{\tilde{x}}^c &= \Gamma^{a}_{b\,c,\,d}\, \xi^d\,(\dot{x}^b + \dot{\xi}^b)\,(\dot{x}^c + \dot{\xi}^c)\\ &= \Gamma^{a}_{b\,c,\,d}\, \xi^d\,(\dot{x}^b\, \dot{x}^c + \dot{x}^b \dot{\xi}^c + \dot{\xi}^b \dot{x}^c + \dot{\xi}^b \dot{\xi}^c)\\ &= \Gamma^{a}_{b\,c,\,d}\, \dot{x}^b\, \dot{x}^c\, \xi^d\end{aligned}$$

to the 1st order since $\xi^d\,(\dot{x}^b \dot{\xi}^c + \dot{\xi}^b \dot{x}^c\,)$ is 2nd order in ξ^a. I think this is 2nd order because $\xi^d \xi^c$ is 2nd order, and so $\xi^d \dot{\xi}^c$ is at least 2nd order. ✓

Claim: equation (3.4-3) can be rewritten as

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{bc}\,\xi^b \dot{x}^c\right) - \Gamma^a_{bc,d}\,\xi^b \dot{x}^c \dot{x}^d - \Gamma^a_{bc}\,\xi^b \ddot{x}^c + \Gamma^a_{bc,d}\,\dot{x}^b \dot{x}^c \xi^d + \Gamma^a_{bc}\,\dot{x}^b \dot{\xi}^c = 0 : \qquad (3.4\text{-}4)$$

$$\frac{d}{du}\Gamma^a_{bc} = \frac{\partial}{\partial x^d}\Gamma^a_{bc}\,\frac{dx^d}{du} = \Gamma^a_{bc,d}\,\dot{x}^d$$

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{bc}\,\xi^b \dot{x}^c\right) = \ddot{\xi}^a + \frac{d}{du}\Gamma^a_{bc}\,\xi^b \dot{x}^c + \Gamma^a_{bc}\,\dot{\xi}^b \dot{x}^c + \Gamma^a_{bc}\,\xi^b \ddot{x}^c$$

$$= \ddot{\xi}^a + \Gamma^a_{bc,d}\,\xi^b \dot{x}^c \dot{x}^d + \Gamma^a_{bc}\,\dot{\xi}^b \dot{x}^c + \Gamma^a_{bc}\,\xi^b \ddot{x}^c$$

Plugging the expression above into equation (3.4-4) yields equation (3.4-3):

$$0 = \ddot{\xi}^a + \Gamma^a_{bc,d}\,\xi^b \dot{x}^c \dot{x}^d + \Gamma^a_{bc}\,\dot{\xi}^b \dot{x}^c + \Gamma^a_{bc}\,\xi^b \ddot{x}^c$$

$$- \Gamma^a_{bc,d}\,\xi^b \dot{x}^c \dot{x}^d - \Gamma^a_{bc}\,\xi^b \ddot{x}^c + \Gamma^a_{bc,d}\,\dot{x}^b \dot{x}^c \xi^d + \Gamma^a_{bc}\,\dot{x}^b \dot{\xi}^c$$

$$= \ddot{\xi}^a + \Gamma^a_{bc,d}\,\dot{x}^b \dot{x}^c \xi^d + \Gamma^a_{bc}\left(\dot{x}^b \dot{\xi}^c + \dot{\xi}^b \dot{x}^c\right) \quad ✓$$

Claim: equation (3.4-4) can be written

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{bc}\,\xi^b \dot{x}^c\right) + \Gamma^a_{de}\left(\dot{\xi}^d + \Gamma^d_{bc}\,\xi^b \dot{x}^c\right)\dot{x}^e - \Gamma^a_{de}\,\Gamma^d_{bc}\,\xi^b \dot{x}^c \dot{x}^e$$

$$- \Gamma^a_{bc,d}\,\xi^b \dot{x}^c \dot{x}^d + \Gamma^a_{bc}\,\xi^b\,\Gamma^c_{de}\,\dot{x}^d \dot{x}^e + \Gamma^a_{bc,d}\,\dot{x}^b \dot{x}^c \xi^d = 0 : \qquad (3.4\text{-}5)$$

Equation (3.34) can be expressed as $\ddot{x}^c = -\,\Gamma^c_{de}\,\dot{x}^d \dot{x}^e$. So, substituting the two expressions below into equation (3.4-4) yields equation (3.4-5):

$$-\,\Gamma^a_{bc}\,\xi^b \ddot{x}^c \overset{(3.34)}{=} \Gamma^a_{bc}\,\xi^b\,\Gamma^c_{de}\,\dot{x}^d \dot{x}^e \quad ✓$$

$$\Gamma^a_{bc}\,\dot{x}^b \dot{\xi}^c = \Gamma^a_{ed}\,\dot{x}^e \dot{\xi}^d = \Gamma^a_{de}\dot{\xi}^d \dot{x}^e$$

$$= \Gamma^a_{de}\dot{\xi}^d \dot{x}^e + \Gamma^a_{de}\,\Gamma^d_{bc}\,\xi^b \dot{x}^c \dot{x}^e - \Gamma^a_{de}\,\Gamma^d_{bc}\,\xi^b \dot{x}^c \dot{x}^e$$

$$= \Gamma^a_{de}\left(\dot{\xi}^d + \Gamma^d_{bc}\,\xi^b \dot{x}^c\right)\dot{x}^e - \Gamma^a_{de}\,\Gamma^d_{bc}\,\xi^b \dot{x}^c \dot{x}^e \quad ✓$$

From $\frac{D\xi^a}{du} \overset{(2.45)}{=} \dot{\xi}^a + \Gamma^a_{bc}\,\xi^b \dot{x}^c$ we see that the first two terms of equation (3.4-5) can be expressed as

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{bc}\,\xi^b \dot{x}^c\right) + \Gamma^a_{de}\left(\dot{\xi}^d + \Gamma^d_{bc}\,\xi^b \dot{x}^c\right)\dot{x}^e = \frac{D^2\,\xi^a}{du^2} \quad ✓$$

The remaining terms can be relabeled so that they all end in $\xi^b \dot{x}^c \dot{x}^d$:

$$- \Gamma^a_{d\,e}\,\Gamma^d_{b\,c}\,\xi^b \dot{x}^c \dot{x}^e - \Gamma^a_{b\,c,\,d}\,\xi^b \dot{x}^c \dot{x}^d + \Gamma^a_{b\,c}\,\xi^b\,\Gamma^c_{d\,e}\,\dot{x}^d \dot{x}^e + \Gamma^a_{b\,c,\,d}\,\dot{x}^b \dot{x}^c \xi^d$$

$$= (- \Gamma^a_{e\,d}\,\Gamma^e_{b\,c} - \Gamma^a_{b\,c,\,d} + \Gamma^a_{b\,e}\,\Gamma^e_{c\,d} + \Gamma^a_{c\,d,\,b})\,\xi^b \dot{x}^c \dot{x}^d$$

$$\overset{(3.13)}{=} R^a{}_{c\,b\,d}\,\xi^b \dot{x}^c \dot{x}^d \qquad ✓$$

Thus, equation (3.4-5) can be simplified to read

$$\boxed{\frac{\mathrm{D}^2 \xi^a}{\mathrm{d}u^2} + R^a{}_{c\,b\,d}\,\xi^b\,\dot{x}^c\,\dot{x}^d = 0} \,. \tag{3.35}$$

This is the **equation of geodesic deviation**.

For example, in a flat manifold, $R^a{}_{b\,c\,d} = 0$. In Euclidean space, $\dfrac{\mathrm{D}^2 \xi^a}{\mathrm{d}u^2} = \dfrac{\mathrm{d}^2 \xi^a}{\mathrm{d}u^2}$. Thus, the equation of geodesic deviation for flat Euclidean space is $\ddot{\xi}^a = 0$ and, hence, a parametric separation vector between two geodesics (i.e., lines) varies according to the formula $\xi^a = A u + B$. The two Euclidean geodesics deviate linearly with u, and with s if γ is non-null.

In a curved manifold, on the other hand, the separation between two geodesics accelerates with u. If acceleration is negative, the distance decreases until, possibly, the geodesics cross and then the distance increases, perhaps forever or perhaps until an inflection point is reached, upon which the distance decreases again.

A final point. How does one reconcile that gravity is a force in Newtonian physics with it not being a force, but curvature, in general relativity? The answer is that gravity is a “fictitious” force, just like the centrifugal, Euler, and Coriolis forces. Newton transformed away the latter forces by choosing a non-rotating reference frame centered at the particle. Had he chosen a reference frame that was freely falling as well, he could have transformed away the gravitational force.

3.5 Einstein's Field Equations

In 1914-15 Einstein made many attempts to find an equation that would relate matter and energy to spacetime curvature. Consider

$$g^{\mu\nu} = \kappa\, T^{\mu\nu} \tag{3.36}$$

where κ is a constant. The metric tensor $g^{\mu\nu}$ is tied to curvature from the geodesic equation. The energy-momentum tensor $T^{\mu\nu}$ is defined in terms of the momentum density and pressure of an ideal fluid and, thus, involves all of the mass and energy information of the fluid. Both tensors are symmetric (i.e., they agree), and both tensors have zero divergence:

$$g^{\mu\nu}{}_{;\mu} \overset{(2.60)}{=} 0$$

and

$T^{\mu\nu}{}_{;\mu} = 0$ iff the relativity continuity equation (3.5) for conservation of energy and the relativity equation of motion ($f^{\mu} = \frac{\mathrm{d}p^{\mu}}{\mathrm{d}\tau}$) for a perfect fluid (3.6) are satisfied (by Theorem 3.1.1).

However, equation (3.36) does not reduce to **Poisson's equation**, the Newtonian field equation:

$$\nabla^2 V = 4\pi G\rho \tag{3.5-1}$$

where V is the gravitational potential, G is Newton's gravitational constant, and ρ is the mass density of a point (x^i) under consideration.

In order to reduce to equation (3.5-1), the curvature tensor on LHS of equation (3.36) must have terms involving the 2nd derivative of V. To that end, in equation (2.83) of Section 2.7 we saw that

$$g_{00} = \eta_{00} + \frac{2\,V}{c^2},$$

meaning that g_{00} has the same partial derivatives (up to a constant) as the gravitational potential V in Euclidean space. In a more general curved manifold, the potential would get dispersed throughout $g_{\mu\nu}$. So, a tensor involving the 2nd derivative of $g^{\mu\nu}$ would be appropriate. One of Einstein's later attempts at the field equations was to use the Ricci tensor instead of $g^{\mu\nu}$:

$$R^{\mu\nu} = \kappa\, T^{\mu\nu}. \tag{3.37}$$

We observe that $R^{\mu\nu}$ involves 2nd derivatives of $g^{\mu\nu}$ because:

$$R^{\mu\nu} = g^{\mu\sigma} g^{\nu\rho} R_{\sigma\rho} \overset{(3.21)}{=} g^{\mu\sigma} g^{\nu\rho} R^{\chi}{}_{\sigma\rho\chi}$$

$$\overset{(3.13)}{=} g^{\mu\sigma} g^{\nu\rho} [\partial_\rho \Gamma^{\chi}_{\sigma\chi} - \partial_\chi \Gamma^{\chi}_{\sigma\rho} + \Gamma^{\zeta}_{\sigma\chi} \Gamma^{\chi}_{\zeta\rho} - \Gamma^{\zeta}_{\sigma\rho} \Gamma^{\chi}_{\zeta\chi}] ,$$

and

$$\Gamma^{\sigma}_{\mu\nu} \overset{(2.13)}{=} \frac{1}{2} g^{\sigma\rho} (\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

so that $R^{\mu\nu}$ involves second derivatives of $g_{\mu\nu}$ and, consequently, of $g^{\mu\nu}$ since it is the inverse of $g_{\mu\nu}$.

$R^{\mu\nu}$ is also symmetric. However, $R^{\mu\nu}{}_{;\mu} \neq 0$.

Einstein later proposed

$$\boxed{G^{\mu\nu} \overset{(3.23)}{=} R^{\mu\nu} - \frac{1}{2} R\, g^{\mu\nu} = \kappa\, T^{\mu\nu}}, \tag{3.38}$$

which satisfies all of the requirements: it is symmetric, involves the metric tensor, reduces to Poisson's equation (as will be shown in the next section), and $G^{\mu\nu}{}_{;\mu} \overset{(\text{Th } 3.2.6)}{=} 0$. In reducing equations (3.38) to Poisson's equation, the constant κ turns out to have the value - $8\pi \frac{G}{c^4}$. Observe that (3.38) consist of 10 equations:

$4\mu \times 4\nu = 16$ total cases. 6 duplicate cases $G_{\mu\nu} = G_{\nu\mu}$ ✓

Recapping what we have so far. If we are given a metric tensor $g^{\mu\nu}$, the corresonding Christoffel coefficients can be computed from it, and the Ricci tensor can be computed from the metic tensor and the Christoffel coeficients. So, G is completely determined by a metric tensor or, equivalently, a line element.

Definition A **solution to Einstein's field equations** is a line element that satisfies equation (3.38).

Metric tensors are coordinate-dependent. If we change the coordinate system, the metric tensor might change. So, for example, the Schwarzschild solution, sometimes called the Schwarzschild geometry, is a coordinate system and a line element in terms of that coordinate system.

By Exercise (3.5.1), an alternate form for field equations (3.38) is

$$R^{\mu\nu} = \kappa \left(T^{\mu\nu} - \tfrac{1}{2} T\, g^{\mu\nu}\right) \tag{3.39}$$

where $T = T^{\mu}_{\mu}$. This makes it unclear whether there is a preference for the middle term to belong on the left (curvature) or on the right (mass-energy). If it is on the left, then both sides have zero divergence. If it is on the right side, then although neither side has zero divergence, they are still in agreement.

After the equations were discovered to show that the universe could not be static but must be either expanding or shrinking, Einstein added a **cosmological constant** term Λ, and later disavowed the term when the universe was found to be expanding. This is expressed in Chapter 6 as

$$R^{\mu\nu} - \tfrac{1}{2} R\, g^{\mu\nu} + \Lambda g^{\mu\nu} = \kappa T^{\mu\nu}\ . \tag{6.50}$$

However, $\Lambda g^{\mu\nu}$ allows equation (6.50) to be fine-tuned in order to match the rate of expansion (or contraction) found in cosmological measurements. If it is vacuum energy that drives the accelerated expansion of the universe, then Λ may represent the vacuum energy. Whatever it represents, the new term is symmetric, has zero divergence, and disappears when reducing to Newtonian physics, so the equations still satisfy all of the requirements for general relativity field equations.

$T^{\mu\nu}$ contains all forms of energy and momentum. A region of spacetime in which $T^{\mu\nu} = 0$ is called **empty**. It is devoid not only of matter but of radiative energy and momentum. From equation (3.39) the **empty spacetime field equations** are

$$R^{\mu\nu} = 0\ . \tag{3.40}$$

Further support for the correctness of the field equations comes from comparing the equation of geodesic deviation with its Newtonian counterpart. With proper time τ as an affine parameter, equation (3.35) of geodesic deviation takes the form

$$\frac{D^2 \xi^\mu}{d\tau^2} + R^{\mu}{}_{\sigma\nu\rho}\, \xi^\nu \dot{x}^\sigma \dot{x}^\rho = 0 \tag{3.41}$$

where $\xi^\mu(\tau)$ is the small “vector” connecting corresponding points on neighboring geodesics. For comparison with its Newtonian counterpart, let us write this as

$$\frac{D^2 \xi^\mu}{d\tau^2} = - K^{\mu}_{\nu}\, \xi^\nu \tag{3.42}$$

where

$$K^{\mu}{}_{\nu} = R^{\mu}{}_{\sigma\nu\rho}\,\dot{x}^{\sigma}\dot{x}^{\rho} \overset{(3.2\text{–}7)}{=} -R^{\mu}{}_{\sigma\rho\nu}\,\dot{x}^{\sigma}\dot{x}^{\rho}\,. \tag{3.43}$$

The corresponding situation in Newtonian theory is two particles moving under gravity on nearby paths $\tilde{x}^i(t)$ and $x^i(t)$. Equation (2.6-2) expresses acceleration in terms of potential:

$$a^i = -\partial_i V = -\delta^{ik}\,\partial_k V.$$

So, the equations of motion for the two particles are

$$\frac{d^2\tilde{x}^i}{dt^2} = -\delta^{ik}\,\tilde{\partial}_k V \quad \text{and} \quad \frac{d^2 x^i}{dt^2} = -\delta^{ik}\,\partial_k V \tag{3.5-2}$$

where

$\tilde{\partial}_k V$ is the gradient of the gravitational potential V evaluated at $\tilde{x}^i(t)$.

Define “vectors” between the two paths:

$$\xi^i(t) \equiv \tilde{x}^i(t) - x^i(t). \tag{3.5-3}$$

The 1st order Taylor series approximation of $\tilde{\partial}_k V$ for small ξ^i is

$$\tilde{\partial}_k V \equiv f(\tilde{x}^k) \approx f(x^k) + \frac{\partial}{\partial x^j} f(\tilde{x}^k)\,(\tilde{x}^k - x^k) = \partial_k V + (\partial_j\partial_k V)\,\xi^j\,. \tag{3.5-4}$$

So,

$$\frac{d^2\xi^i}{dt^2} = -K^i_j\,\xi^j \tag{3.44}$$

where

$$K^i_j \equiv \delta^{ik}\,\partial_j\partial_k V: \tag{3.45}$$

$$\frac{d^2\xi^i}{dt^2} \overset{(3.5\text{–}3)}{=} \frac{d^2\tilde{x}^i}{dt^2} - \frac{d^2 x^i}{dt^2} \overset{(3.5\text{–}2)}{=} -\delta^{ik}(\tilde{\partial}_k V - \partial_k V) \overset{(3.5\text{–}4)}{=} -\delta^{ik}\,\partial_j\partial_k V\,\xi^j = -K^i_j\,\xi^j \;\checkmark$$

Equation (3.44) is the Newtonian counterpart of geodesic deviation equation (3.42), which shows that

$$K^{\mu}{}_{\nu} = -R^{\mu}{}_{\sigma\rho\nu}\,\dot{x}^{\sigma}\dot{x}^{\rho} \qquad \longleftrightarrow \qquad K^i_j = \delta^{ik}\,\partial_j\partial_k V\,.$$

The **empty space field equation of Newtonian gravitation** is $\nabla^2 V \overset{(2.6\text{–}4)}{=} 0$. This is equivalent to $K^i_i = 0$:

$$\nabla^2 V \equiv (\nabla\bullet\nabla) V = \frac{\partial^2 V}{(\partial x^1)^2} + \frac{\partial^2 V}{(\partial x^2)^2} + \frac{\partial^2 V}{(\partial x^3)^2} = \sum \partial_i^2 V$$

$$K^i_i = K^i_\ell\, \delta^\ell_i = \delta^\ell_i\, \delta^{i\,k}\, \partial_\ell \partial_k V = \delta^{i\,k}\, \partial_i \partial_k V = \sum \partial_i \partial_i V = \sum \partial_i^2 V$$

$$\Rightarrow \quad \nabla^2 V = K^i_i \qquad \checkmark$$

Thus, the general relativity empty spacetime field equations (3.40) are the counterpart to the Newtonian empty space field equation:

Newtonian empty space field equation $\Leftrightarrow$ $\nabla^2 V = 0$ $\Leftrightarrow$ $K^i_i = 0$

$\longleftrightarrow$ $K^\mu{}_\mu = 0$ $\Leftrightarrow$ $R_{\sigma\rho} = R^\mu{}_{\sigma\rho\mu} = 0$ $\Leftrightarrow$ $R^{\mu\nu} = g^{\mu\sigma} g^{\nu\rho} R_{\sigma\rho} = 0$

$\Leftrightarrow$ general relativity empty spacetime field equations $\checkmark$

This lends support for for the empty spacetime field equation. Support for the nonempty spacetime field equations (3.38) or (3.39) is given in the next section.

3.6 Einstein's Equation compared with Poisson's equation

The field equation of Newtonian gravitation is Poisson's equation:

$$\nabla^2 V = 4\pi G\rho \,. \tag{2.6.4}$$

This represents one equation in the one unknown, V. By equation (2.6.1),

$$V = -\int_{-\infty}^{x} g\, dx.$$

So, solving for V in equation (2.6.4) is equivalent to solving for g.

The covariant version of Einstein's gravitational field equations is

$$R_{\mu\nu} = \kappa\,(T_{\mu\nu} - \tfrac{1}{2} T\, g_{\mu\nu}) :$$

$$R_{\mu\nu} = g_{\mu\sigma}\, g_{\nu\rho}\, R^{\sigma\rho} \overset{(3.39)}{=} g_{\mu\sigma}\, g_{\nu\rho}\, \kappa\,(T^{\sigma\rho} - \tfrac{1}{2} T\, g^{\sigma\rho}) = \kappa\,(T_{\mu\nu} - \tfrac{1}{2} T\, g_{\mu\nu}) \checkmark$$

Because $g_{\mu\nu} = g_{\nu\mu}$, we can restrict our attention to $\mu \le \nu$, reducing the 16 gravitational field equations in 16 unknowns down to 10 equations in 10 unknowns, as compared to the Newtonian equation (2.6.4) that is only 1 equation in 1 unknown.

To show that Einstein's field equations reduce to Poisson's equation, we will show that R_{00} reduces to Poisson's equation under approximately Euclidean conditions.

$$R_{00} = \kappa\,(T_{00} - \tfrac{1}{2} T\, g_{00}) \,. \tag{3.46}$$

Not mentioned in the book, we must, in addition, show that the other 9 simultatneous equations are redundant. To that end, in the next section, Section 3.7, when we develop the Schwarzsfield solution to Einstein's field equations, we will show that solving for the non-zero connection coefficients reduces the ten field equations down to four equations (3-54 to 3-57) in two unknowns, A and B. We then show that in asymtotically flat space that $B \propto 1/A$, reducing this to four equations in one unknown, which are then shown to be consistent under the approximately Newtonian conditions assumed by Schwarszchild. Thus, the ten simultaneous field equations reduce to the single equation (3.46) $\Leftrightarrow$ equation (3-54) as Newtonian conditions are approached.

To show that R_{00} reduces to equation (3.46), as in Section 2.7, we assume we have an almost Cartesian coordinate system in which:

- $g_{\mu\nu} \overset{(2.74)}{=} \eta_{\mu\nu} + h_{\mu\nu}$ where $h_{\mu\nu}$ are small (i.e., $h_{\mu\nu} \approx 0$ to 1st order) (3.6-1)

- Products
  $h_{\mu\nu}\, h_{\sigma\rho} \approx h^{\mu\nu}\, h^{\sigma\rho} \approx h^{\mu\nu} h_{\sigma\rho} \approx 0$ to 2nd order (3.6-2)

- Weak field products
  $\partial_\alpha h_{\mu\nu}\, \partial_\beta h_{\sigma\rho} \approx 0$ to 2nd order (3.6-3)

- The gravitational field $h_{\mu\nu}$ is quasi-static:
  $|\partial_0 h_{\mu\nu}| \overset{(2.7-7)}{\ll} |\partial_i h_{\sigma\rho}|$ and $|\partial_0 \Gamma^\mu_{\nu\alpha}| \ll |\partial_\mathrm{i} \Gamma^\sigma_{\rho\beta}|$ for all $i, \alpha, \beta, \mu, \nu, \sigma, \rho$. (3.6-4)

The quasi-static condition (3.6-4) is a much stronger version than previously given as equation (2.7-7). In thermodynamics, quasi-static means that changes happen slowly so that equilibrium is maintained. In this section we need not only that 1st derivatives change slowly (1st order) but also that 2nd derivatives change slowly (2nd order). The second part of equations (3.6-4) includes 2nd derivative terms like $\partial_0 \partial_\gamma h_{\nu\alpha} \ll \partial_i \partial_\delta h_{\rho\beta}$. In fact, requiring that $\Gamma^\mu_{\nu\alpha}$ change slowly is an even more stringent requirement than just that 2nd derivatives change slowly.

We also assume

- Particles (in this coordinate system) have speeds $v \ll c$ so that
  $\gamma \approx 1$ and $\left|\dfrac{\mathrm{dx}^i}{\mathrm{dt}}\right| \overset{(2.7-6)}{\ll} c$. (3.6-5)

From Appendix A.0, Example 1,

$u_0 \approx c$ and $|u_i| \ll c$: (3.6-6)

$$u^\mu = \gamma\,(c, \boldsymbol{v}) \;\Rightarrow\; u_0 = u^0 \overset{(3.6-5)}{\approx} c$$

$$|u_i| = |-u^i| = |\gamma v^i| \overset{(3.6-5)}{\approx} |v^i| \overset{(A.23)}{=} \left|\frac{\mathrm{dx}^i}{\mathrm{dt}}\right| \overset{(3.6-5)}{\ll} c. \checkmark$$

Finally, we also assume

- $\dfrac{P}{c^2} \ll \rho$, or $P \ll \rho c^2$ (3.6-7)
  as is true for most classical distributions (Sun, water, high pressure gas).

Under these conditions, the covariant version of the energy-momentum stress tensor becomes $T_{\mu\nu} \approx \rho\, u_\mu u_\nu - P g_{\mu\nu}$:

$$T_{\mu\nu} \overset{(3.1-19)}{=} (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\mu\nu} \overset{(3.6-7)}{\approx} \rho\, u_\mu u_\nu - P g_{\mu\nu}$$

Thus, $T_{00} \approx \rho c^2$ and $T_{i\,i} \approx \rho\,(u_i)^2 + P$: (3.6-8)

$$T_{00} \approx \rho\,(u_0)^2 - P\,\eta_{00} = \rho\,(u_0)^2 - P \overset{(3.6-6)}{\approx} \rho c^2 - P \overset{(3.6-7)}{\approx} \rho c^2$$
$$T_{i\,i} \approx \rho\,(u_i)^2 - P\,\eta_{i\,i} = \rho\,(u_i)^2 + P$$

Claim $T \approx \rho c^2$: (3.6-9)

$$T = T^\mu_\mu = g^{\mu\nu} T_{\mu\nu} = g^{00} T_{00} + g^{i\,i} T_{i\,i} \overset{(3.6-8)}{\approx} (1)\,\rho c^2 + (-1)\ [\rho\,(u_i)^2 + P]$$
$$= \rho\, c^2 - \rho\, u_i^2 - P \overset{(3.6-7)}{\approx} \rho c^2$$

Claim $R_{00} \approx \frac{1}{2} \kappa \rho c^2$: (3.47)

$$R_{00} \overset{(3.46)}{=} \kappa\,(T_{00} - \tfrac{1}{2} T\, g_{00}) \overset{(3.6-8,\ 3.6-9)}{\approx} \kappa\rho\,({u_0}^2 - \tfrac{1}{2}\, c^2\, g_{00})$$
$$\overset{(3.6-6)}{\approx} \kappa\rho\,[c^2 - \tfrac{1}{2} c^2 (1)] = \tfrac{1}{2}\, \kappa\rho c^2 \quad ✓$$

Also,

$$R_{00} \overset{(3.21)}{=} R^\mu_{00\,\mu} \overset{(3.13)}{=} \partial_0 \Gamma^\mu_{0\,\mu} - \partial_\mu \Gamma^\mu_{00} + \Gamma^\nu_{0\,\mu}\, \Gamma^\mu_{\nu 0} - \Gamma^\nu_{00}\, \Gamma^\mu_{\nu\mu} \qquad (3.48)$$

Claim $\Gamma^\mu_{\nu\alpha}\, \Gamma^\sigma_{\rho\beta} \approx 0$: (3.6-10)

From Section 2.7 we have that

$$\Gamma^\mu_{\nu\alpha}\, \Gamma^\sigma_{\rho\beta} \overset{(2.78)}{\approx} \frac{1}{4}\, \eta^{\mu\gamma} \eta^{\nu\delta}\, (\partial_\nu h_{\gamma\alpha} + \partial_\alpha h_{\nu\gamma} - \partial_\gamma h_{\nu\alpha})\, (\partial_\rho h_{\delta\beta} + \partial_\beta h_{\rho\delta} - \partial_\delta h_{\rho\beta}) \overset{(3.6-3)}{\approx} 0 \quad ✓$$

By (3.6-4), $|\partial_0 \Gamma^\mu_{\nu\alpha}| \ll |\partial_\mathrm{i} \Gamma^\sigma_{\rho\beta}|$ for all i, α, β, μ, ν, σ, ρ . Thus equation (3.48) simplifies to

$$R_{00} = -\, \partial_\mathrm{i} \Gamma^\mathrm{i}_{00}\ . \qquad (3.6-11)$$

Claim $R_{00} \approx -\frac{1}{c^2}\,\nabla^2 V$: (3.6-12)

$$R_{00} \overset{(3.6-11)}{=} -\partial_i \Gamma^i_{00} = -(\partial_1\Gamma^1_{00} + \partial_2\Gamma^2_{00} + \partial_3\Gamma^3_{00})$$

$$\overset{(2.7-8)}{=} -\frac{1}{2}(\partial_1\partial_1 h_{00} + \partial_2\partial_2 h_{00} + \partial_3\partial_3 h_{00})$$

$$= -\frac{1}{2}\,\nabla^2 h_{00} \overset{(2.83)}{=} -\frac{1}{2}\,\nabla^2\left(\frac{2V}{c^2}\right)$$

$$= -\frac{1}{c^2}\,\nabla^2 V \quad ✓$$

$\therefore\ \nabla^2 V \approx -\frac{1}{2}\kappa\rho c^4$: (3.49)

$$\nabla^2 V \overset{(3.6-12)}{\approx} -c^2 R_{00} \overset{(3.47)}{\approx} -\tfrac{1}{2}\kappa\rho c^4 ✓$$

Equation (3.49) is precisely Poisson's equation (2.6.4) if

$$\boxed{\kappa = -\frac{8\pi G}{c^4}} : \quad (3.6\text{-}13)$$

$$\nabla^2 V \approx 4\pi G\rho \qquad ✓ \quad (3.50)$$

This shows that $c^2\ R_{00} \approx 4\pi G\rho$; that is, the 1st of the ten $R_{\mu\nu} = \kappa\,(T_{\mu\nu} - \frac{1}{2}T\,g_{\mu\nu})$ field equations is precisely Poisson's equation.

3.7 The Schwarzschild Solution

Schwarzschild examined the special case of a spherically symmetric gravitational field in the empty spacetime surrounding a massive spherical object like a star. His assumptions are:

(a) The metric tensor field is static
(b) The metric tensor field is spherically symmetric
(c) Spacetime is empty
(d) Spacetime is asymptotically flat
(e) Spacetime can be coordinatized by (t,r,θ,ϕ) where **coordinate *t* is timelike**, meaning the tangent vector to the coordinate curve is timelike, θ and ϕ are the usual spherical angles, and r is a radial coordinate.

He postulated a line element

$$c^2\,\mathrm{d}\tau^2 = A(r)\,\mathrm{d}t^2 - B(r)\,\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,\sin^2\theta\,\mathrm{d}\phi^2 \qquad (3.51)$$

which can also be expressed as

$$c^2\,\mathrm{d}\tau^2 = \frac{1}{c^2}A(r)\,\mathrm{d}(c\,t)^2 - B(r)\,\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,\sin^2\theta\,\mathrm{d}\phi^2\,.$$

Then

$$(g_{\mu\nu}) = \begin{pmatrix} \frac{1}{c^2}A(r) & & & \\ & -B(r) & & \\ & & -r^2 & \\ & & & -r^2\sin^2\theta \end{pmatrix} \quad \forall\, r, \qquad (3.7\text{-}1)$$

and

$$(g^{\mu\nu}) = \begin{pmatrix} c^2/A(r) & & & \\ & -1/B(r) & & \\ & & -1/r^2 & \\ & & & -1/r^2\sin^2\theta \end{pmatrix},\ r > 0. \qquad (3.7\text{-}2)$$

Assumption (a) is satisfied because all $g_{\mu\nu}$ are independent of t. ✓

If we freeze r and t, we get a surface Σ with the geometry of a sphere of radius r, parameterized by θ and ϕ, and, by Exercise 1.6.2 (a), whose line element is

$$\mathrm{d}s^2 = r^2 \mathrm{d}\theta^2 + r^2 \sin^2\theta \, \mathrm{d}\phi^2 \tag{3.52}$$

This expresses assumption (b). ✓

Assumption (c) means that $A(r)$ and $B(r)$ are to be solved using the empty spacetime field equations (3.40): $R_{\mu\nu} = 0$.

Assumption (d) gives the boundary conditions on A and B, namely

$$A(r) \to c^2 \quad \text{and} \quad B(r) \to 1 \quad \text{as} \quad r \to \infty. \tag{3.53}$$

Here is an outline of the steps we take to solve the empty spacetime field equations to determine $A(r)$ and $B(r)$. We first note that there are only four non-zero metric tensor components, and that results in only thirteen non-zero connection coefficients (left table, below). That, in turn, leads to only four non-zero Ricci tensors $R_{\mu\nu}$, (3.54) – (3.57) below. Thus, the original ten "$R_{\mu\nu} = 0$" empty spacetime field equations equations are reduced to only four equations, and in only two unknowns, $A(r)$ and $B(r)$. The first two differential equations (in the two unknowns) are solved simultaneously, resulting in $A(r)$ and $B(r)$ being functions of three new unknowns (constants of integration): k_1, k_2, and k_3. Assumption (d) is used to reduce this to the single unknown, k_1. We then solve for $B(r)$ in terms of $A(r)$ by using a change of coordinates between Euclidean and Schwarzschild spacetime coordinates, and the result is four equations in the one unknown, k_1. Three of the equations are shown to be redundant. The net result is 1 equation in 1 unknown, which is solved. This result was referenced in Section 3.6 where it was shown that Einstein's ten field equations reduce to Poisson's equation in Euclidean space.

For easier readability while solving the field equations, we shorten $A(r)$ and $B(r)$ to A and B, respectively, and we use prime to denote the derivatives with respect to r:

From equations (3.21) and (3.13) we get

$$R_{\mu\nu} \overset{(3.21)}{=} R^{\sigma}{}_{\mu\nu\sigma} \overset{(3.13)}{\equiv} \partial_\nu \Gamma^{\sigma}_{\mu\sigma} - \partial_\sigma \Gamma^{\sigma}_{\mu\nu} + \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\rho\nu} - \Gamma^{\rho}_{\mu\nu}\Gamma^{\sigma}_{\rho\sigma}.$$

Using Mathematica, we compute the non-zero connection coefficients $\Gamma^{\mu}_{\nu\gamma}$ and then the non-zero Ricci tensors $R_{\mu\nu}$ (for $r > 0$):

μ	ν	γ	$\Gamma^{\mu}_{\nu\gamma}$
0	0	1	$\frac{A'}{2A}$
0	1	0	$\frac{A'}{2A}$
1	0	0	$\frac{A'}{2B}$
1	1	1	$\frac{B'}{2B}$
1	2	2	$-\frac{r}{B}$
1	3	3	$-\frac{r\,\mathrm{Sin}[\theta]^2}{B}$
2	1	2	$\frac{1}{r}$
2	2	1	$\frac{1}{r}$
2	3	3	$-\mathrm{Cos}[\theta]\ \mathrm{Sin}[\theta]$
3	1	3	$\frac{1}{r}$
3	2	3	$\mathrm{Cot}[\theta]$
3	3	1	$\frac{1}{r}$
3	3	2	$\mathrm{Cot}[\theta]$

μ	ν	$R_{\mu\nu}$	
0	0	$-\frac{A'}{B\,r}+\frac{(A')^2}{4\,A\,B}+\frac{A'\,B'}{4\,B^2}-\frac{A''}{2\,B}$	(3.54)
1	1	$-\frac{(A')^2}{4\,A^2}-\frac{B'}{B\,r}-\frac{A'\,B'}{4\,A\,B}+\frac{A''}{2\,A}$	(3.55)
2	2	$-1+\frac{1}{B}+\frac{r\,A'}{2\,A\,B}-\frac{r\,B'}{2\,B^2}$	(3.56)
3	3	$-\mathrm{Sin}[\theta]^2+\frac{\mathrm{Sin}[\theta]^2}{B}+\frac{r\,\mathrm{Sin}[\theta]^2\,A'}{2\,A\,B}-\frac{r\,\mathrm{Sin}[\theta]^2\,B'}{2\,B^2}$	(3.57)

(Generation of $\Gamma^{\mu}_{\nu\gamma}$ constitutes Problem 2.7. Generation of (3.54 – 3.57) constitutes Exercise 3.7.1.) Because of Assumption (c), we wish to solve $R_{\mu\nu}$ = 0. We observe that $R_{33} = \sin^2\theta\, R_{22}$, which makes R_{33} = 0 redundant, a multiple of R_{22} = 0. We further observe that R_{00} = 0 and R_{11} = 0 constitute two differential equations in the two unknowns $A(r)$ and $B(r)$. Using Mathematica to DSolve them yields a solution with three constants (due to integration), k_1, k_2, and k_3:

$$A = k_2 - \frac{k_1}{r} = -\frac{k_1 - k_2\,r}{r} \quad \text{and} \quad B = \frac{k_3\,r}{k_1 - k_2\,r} = -\frac{k_3}{A}. \tag{3.7-3}$$

We next find the values of two of the constants:

$$c^2 \overset{(3.53)}{=} \lim_{r\to\infty} A \overset{(3.7-3)}{=} k_2 \quad \text{and} \quad 1 \overset{(3.53)}{=} \lim_{r\to\infty} B \overset{(3.7-3)}{=} -\frac{k_3}{k_2} \quad \Rightarrow \quad k_3 = -k_2 = -c^2$$

$$\Rightarrow \quad A = c^2 - \frac{k_1}{r} = \frac{c^2\,r - k_1}{r} \quad \text{and} \quad B = \frac{c^2}{A}. \tag{3.7-4}$$

We have not yet used R_{22} = 0, which appears at first glance to be a single equation in the single unknown k_1. But, plugging in A and B yields R_{22} identically zero; i.e., we get 0 = 0. Thus, equation (3.56) is redundant, and we can't use it to solve for k_1. This is fortunate because $A = \frac{c^2 r - k_1}{r}$ needs to reflects mass or gravity, and so k_1 should include Newton's gravitational constant G. We find the value of k_1 by imposing condition (d), that spacetime be asymptotically flat, and by using the afore-mentioned change of coordinates.

Substituting the expressions (3.7-4) for $A(r)$ and $B(r)$ into line element equation (3.51) yields

$$c^2\,\tau^2 = (c^2 - \frac{k_1}{r})\,\mathrm{d}t^2 - c^2\left(c^2 - \frac{k_1}{r}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,sin^2\theta\,\mathrm{d}\phi^2 \tag{3.7-5}$$

or, alternatively,

$$c^2\,\tau^2 = (1 - \frac{k_1}{c^2\,r})\,\mathrm{d}(c\,t)^2 - \left(1 - \frac{k_1}{c^2\,r}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,sin^2\theta\,\mathrm{d}\phi^2 \tag{3.7-6}$$

As $r \to \infty$, the line element approaches

$$c^2\,\tau^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,sin^2\theta\,\mathrm{d}\phi^2, \tag{3.7-7}$$

or

$$c^2\,\tau^2 = \mathrm{d}(c\,t)^2 - \mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,sin^2\theta\,\mathrm{d}\phi^2, \tag{3.7-8}$$

which is the line element for spherical coordinates in flat spacetime. We choose lower case for Schwarzschild index notation:

$$x^0 = c\,t, \quad x^1 = r, \quad x^2 = \theta, \quad x^3 = \phi, \tag{3.7-9}$$

where, in flat space,

$$x = r\sin\theta\cos\phi, \;\; y = r\sin\theta\sin\phi, \;\; z = r\cos\theta. \tag{3.7-10}$$

For flat spacetime Cartesian coordinates, we choose upper case for the index notation:

$$X^0 = c\,t, \quad X^1 = x, \quad X^2 = y, \quad X^3 = z. \tag{3.7-11}$$

Recall

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \text{arc cos}\,\frac{z}{r}, \quad \phi = \text{arc tan}\,\frac{y}{x}. \tag{3.7-12}$$

From equation (3.7-6) we capture the four non-zero Schwarzschild-coordinates covariant metric tensors $g_{\mu\nu}$ and then express them in Cartesian spacetime coordinates:

$$g_{00} = 1 - \frac{k_1}{c^2 r} = 1 - \frac{k_1}{c^2\sqrt{x^2+y^2+z^2}}$$

$$g_{11} = -\left(1-\frac{k_1}{c^2 r}\right)^{-1} = -\left(1-\frac{k_1}{c^2\sqrt{x^2+y^2+z^2}}\right)^{-1} \qquad (3.7\text{-}13)$$

$$g_{22} = -r^2 = -(x^2+y^2+z^2)$$

$$g_{33} = -r^2 \sin^2\theta = -r^2(1-\cos^2\theta) = -r^2 + z^2 = -(x^2+y^2)$$

To find the Cartesian metric tensors G_{MN} , we must perform a change of coordinates operation. Both Schwarzsfield and Euclidean coordinate systems represent inertial frames because they each satisfy the inertial-frame definition, that their respective coordinates are orthogonal. For example, $\frac{\partial r}{\partial\theta} = 0$ and $\frac{\partial x}{\partial t} = 0$. Thus, Equation (A.8), $G_{MN} = \Lambda_M^\sigma \Lambda_N^\rho g_{\sigma\rho}$, the Lorentz transformation between inertial frames, is the formula to use for changing Schwarzschild coordinates to Cartesian coordinates.

To use this equation, we must find an expression for Λ_M^σ. Recall from equation (A1.2) that the homogeneous Lorentz transformation $\tilde{\Lambda}_M^\sigma$ has the formula $\tilde{\Lambda}_M^\sigma = \frac{\partial x^\sigma}{dX^M}$, and that the (general) Lorentz transformation $\Lambda_M^\sigma = \tilde{\Lambda}_M^\sigma$ iff the offset $a^\mu = 0$. Since Schwarzschild coordinates have the same origin as Cartesian spacetime coordinates, the offset is zero.

So, we have

$$G_{MN} = X_M^\sigma X_N^\rho g_{\sigma\rho} = \frac{\partial x^\sigma}{\partial X^M}\frac{\partial x^\rho}{\partial X^N} g_{\sigma\rho} . \qquad (3.7\text{-}14)$$

Since $g_{\sigma\rho} = 0$ unless $\sigma = \rho$, this simplifies to $G_{MN} = \Lambda_M^\sigma \Lambda_N^\sigma g_{\sigma\sigma}$, and expanding this out yields

$$G_{MN} = \Lambda_M^0 \Lambda_N^0 g_{00} + \Lambda_M^1 \Lambda_N^1 g_{11} + \Lambda_M^2 \Lambda_N^2 g_{22} + \Lambda_M^3 \Lambda_N^3 g_{33} \qquad (3.7\text{-}15)$$

We write break this into three pieces:

$$G_{00} = \frac{\partial(c\,t)}{\partial(c\,t)}\frac{\partial(c\,t)}{\partial(c\,t)} g_{00} + 0 + 0 + 0 = (1)\,(1)\,(1 - \frac{k_1}{c^2 r}) = 1 - \frac{k_1}{c^2 r} , \qquad (3.7\text{-}16)$$

$$G_{0I} = \frac{\partial(c\,t)}{\partial(c\,t)}\frac{\partial(c\,t)}{\partial x^I}\,g_{00} + \frac{\partial x^1}{\partial(\mathrm{ct})}\frac{\partial x^1}{\partial x^I}\,g_{11} + \frac{\partial x^2}{\partial(\mathrm{ct})}\frac{\partial x^2}{\partial x^I}\,g_{22} + \frac{\partial x^3}{\partial(\mathrm{ct})}\frac{\partial x^3}{\partial x^I}\,g_{33}$$
$$= 0 \quad \left(\text{because } \frac{\partial(c\,t)}{\partial x^I} = 0 \text{ and } \frac{\partial x^i}{\partial(\mathrm{ct})} = 0\right) \tag{3.7-17}$$

$$G_{IJ} = \frac{\partial x^1}{\partial x^I}\frac{\partial x^1}{\partial x^J}\,g_{11} + \frac{\partial x^2}{\partial x^I}\frac{\partial x^2}{\partial x^J}\,g_{22} + \frac{\partial x^3}{\partial x^I}\frac{\partial x^3}{\partial x^J}\,g_{33} \tag{3.7-18}$$

We used Mathematica to compute G_{IJ} from the change of coordinates equation (3.7-18). The table below summarizes the results for all G_{MN} (for $r > 0$).

σ	M	$\Lambda^\sigma_M = \frac{\partial x^\sigma}{\partial x^M}$
0	0	1
0	2	0
0	2	0
0	3	0
1	0	0
1	1	$\frac{x}{r}$
1	2	$\frac{y}{r}$
1	3	$\frac{z}{r}$
2	0	0
2	1	$\frac{x\,z}{r^2\sqrt{x^2+y^2}}$
2	2	$\frac{y\,z}{r^2\sqrt{x^2+y^2}}$
2	3	$-\frac{\sqrt{x^2+y^2}}{r^2}$
3	0	0
3	1	$-\frac{y}{x^2+y^2}$
3	2	$\frac{x}{x^2+y^2}$
3	3	0

M	N	$G_{MN} = \Lambda^\sigma_M \Lambda^\sigma_N g_{\sigma\sigma}$
0	0	$1 - \frac{k_1}{c^2\,r}$
0	2	0
0	2	0
0	3	0
1	0	0
1	1	$-1 + \frac{x^2\,k_1}{r^2\,(-c^2\,r+k_1)}$
1	2	$-\frac{x\,y\,k_1}{r^2\,(c^2\,r-k_1)}$
1	3	$-\frac{x\,z\,k_1}{r^2\,(c^2\,r-k_1)}$
2	0	0
2	1	$-\frac{x\,y\,k_1}{r^2\,(c^2\,r-k_1)}$
2	2	$-1 + \frac{y^2\,k_1}{r^2\,(-c^2\,r+k_1)}$
2	3	$-\frac{y\,z\,k_1}{r^2\,(c^2\,r-k_1)}$
3	0	0
3	1	$-\frac{x\,z\,k_1}{r^2\,(c^2\,r-k_1)}$
3	2	$-\frac{y\,z\,k_1}{r^2\,(c^2\,r-k_1)}$
3	3	$-1 + \frac{z^2\,k_1}{r^2\,(-c^2\,r+k_1)}$

(3.7-19)

Table 3.7-19 constitutes Exercise 3.7.2. The book's solution found a constant k, not k_1, but they are related by $k_1 = -c^2 k$ and they generate the same results

As $r \to \infty$, the Schwarzschild geometry approaches flat spacetime, and the Cartesian metric tensor G approaches $\eta + h$, where $h = 0$ to the 2nd order. That is,

$G_{MN} = \eta_{MN} + h_{MN}$:

$$G_{MN} = \begin{pmatrix} 1-\frac{k_1}{c^2 r} & 0 & 0 & 0 \\ 0 & -1+\frac{x^2 k_1}{r^2(-c^2 r+k_1)} & -\frac{x y k_1}{r^2(c^2 r-k_1)} & -\frac{x z k_1}{r^2(c^2 r-k_1)} \\ 0 & -\frac{x y k_1}{r^2(c^2 r-k_1)} & -1+\frac{y^2 k_1}{r^2(-c^2 r+k_1)} & -\frac{y z k_1}{r^2(c^2 r-k_1)} \\ 0 & -\frac{x z k_1}{r^2(c^2 r-k_1)} & -\frac{y z k_1}{r^2(c^2 r-k_1)} & -1+\frac{z^2 k_1}{r^2(-c^2 r+k_1)} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} -\frac{k_1}{c^2 r} & 0 & 0 & 0 \\ 0 & \frac{x^2 k_1}{r^2(-c^2 r+k_1)} & -\frac{x y k_1}{r^2(c^2 r-k_1)} & -\frac{x z k_1}{r^2(c^2 r-k_1)} \\ 0 & -\frac{x y k_1}{r^2(c^2 r-k_1)} & \frac{y^2 k_1}{r^2(-c^2 r+k_1)} & -\frac{y z k_1}{r^2(c^2 r-k_1)} \\ 0 & -\frac{x z k_1}{r^2(c^2 r-k_1)} & -\frac{y z k_1}{r^2(c^2 r-k_1)} & \frac{z^2 k_1}{r^2(-c^2 r+k_1)} \end{pmatrix} \qquad (3.7\text{-}20)$$

This allows us to solve for the constant k_1 (for $r > 0$):

$k_1 = 2\,MG$: (3.7-21)

$$-\frac{k_1}{c^2\, r} \overset{(3.7-20)}{=} h_{00} \overset{(2.83)}{=} \frac{2\,V}{c^2} \overset{(2.8-2)}{=} -\frac{2}{c^2}\,\frac{GM}{r} \qquad ✓$$

Finally, we replace k_1 by $2\,MG$ in the line element (3.7-6) to generate the Schwarzschild line element and corresponding metric tensor:

$$\boxed{c^2\,d\tau^2 \overset{(3.7-21)}{=} \left(1-\frac{2\,MG}{c^2\,r}\right) d(c\,t)^2 - \left(1-\frac{2\,MG}{c^2\,r}\right)^{-1} dr^2 - r^2\,d\theta^2 - r^2\,sin^2\,\theta\,d\phi^2} \qquad (3.59)$$

$$(g_{\mu\nu}) \overset{(3.59)}{=} \begin{pmatrix} 1 - \frac{2\,MG}{c^2\,r} & & & \\ & -\frac{1}{1-\frac{2\,MG}{c^2\,r}} & & \\ & & -r^2 & \\ & & & -r^2\,sin^2\,\theta \end{pmatrix} \qquad (3.7\text{-}22)$$

where M is the mass of the body creating the gravitational field, and G is the gravitational constant. This finishes the derivation of the Schwarzschild metric tensor.

Note 1 With $M = 0$, equation (3.59) reduces to spherical coordinates in flat spacetime. As M grows, it eventually represents a black hole, and equation (3.59) descrbes spacetime near a black hole.

Note 2 The result (3.7-22) is a static metric; it does not depend upon t. Nothing was required of the body of mass M except that it generate a metric tensor field that is (a) static and (b) spherically symmetric. The body itself need not be static; it could be a collapsing star, as long as the collapse is symmetric.

Note 3 There are no concave-convex “ripples” in “Schwarzschild geometry. The deviation of $g_{\mu\nu}$ from flat spacetime $\eta_{\mu\nu}$ is concave outside a radius of $\frac{2\,G\,M}{c^2}$ and convex inside that radius:

$M \geq 0$ and $G > 0$.

$$r > \frac{2\,G\,M}{c^2} \;\Rightarrow\; g_{\mu\mu} - \eta_{\mu\mu} \overset{(3.7-22)}{\leq} 0 \quad \text{and} \quad r < \frac{2\,G\,M}{c^2} \;\Rightarrow\; g_{\mu\mu} - \eta_{\mu\mu} \overset{(3.7-22)}{\geq} 0 \qquad \checkmark$$

In Section 4.8, the radius $\frac{2\,G\,M}{c^2}$ will be seen to be the Schwarzschild radius for a black hole.

Note 4 When $r \neq \frac{2\,G\,M}{c^2}$, we cannot assume that r is radial distance. All we know about r for now is that the surface area of a space sphere Σ of radius r (see comment preceding equation 3.52) is $4\pi\,r^2$. Section 4.1 will explain why we can no longer regard the coordinates t and r as having the same simple meaning as their Euclidean counterparts.

Chapter 4 Physics in the vicinity of a massive object

4.0 Introduction

The Schwarzschild solution provides the line element for measuring distances between points that are exterior to some object with mass. The metric tensor in the Schwarzschild solution is a function of the mass. This allows us to explore spacetime from flat (where $m = 0$) all the way to the extreme curvature that occurs just outside a black hole (where m is large).

Section 4.1 investigates the physical meaning of coordinates like t, r, θ, and ϕ in curved spacetime.

Sections 4.2 and 4.3 develop equations of physics for things like the parallax of Venus that enable general relativity to make predictions.

Sections 4.4 – 4.6 explore the geodesics of light and other massless particles.

Section 4.7 develops parallel transport in spacetime for particles and includes the geodesic effect that that due to curvature, parallel transport around a circle can now experience a change in the tangent angle.

Section 4.8 develops Eddington-Finkelstein coordinates, an alternate Schwarzschild coordinate system that is valid inside the Schwarzschild radius.

Section 4.9 develops isotropic coordinates, an additional Schwarzschild coordinate system that reduces to Cartesian coordinates rather than spherical coordinates. It also develops Kruskal-Szekeres coordinates in which null geodesics are lines sloped at 45°.

Section 4.10 develops the Kerr solution, a generalization of the Schwarzschild solution, that applies to rotating black holes in otherwise empty spacetime. It shows that for rotating black holes there are two event horizons, not one.

4.1 Length and time

Setting $m = \frac{GM}{c^2}$ allows us explore Schwarzschild's solution using more compact notation:

$$c^2\, d\tau^2 \overset{(3.59)}{=} c^2\,(1 - \frac{2\,m}{r})\, dt^2 - \frac{1}{1-\frac{2\,m}{r}}\, dr^2 - r^2\, d\theta^2 - r^2\, sin^2\theta\, d\phi^2 \tag{4.1}$$

where $r > 0$, $m \geq 0$.

We note that m has units of distance, important because the Schwarzschild radius will be define in Section 4.8 to be $2m$.

The line element (4.1) is an expression in terms of the metric tensor:

$$c^2\, d\tau^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu \tag{4.1-1}$$

where

$$g_{00} = (1 - \frac{2\,m}{r}),\quad g_{11} = -\frac{1}{1-\frac{2\,m}{r}},\quad g_{22} = -\,r^2,\ \text{and}\ \ g_{33} = -\,r^2\, sin^2\theta \tag{4.1-2}$$

and

$$x^0 = ct,\ \ x^1 = r,\ \ x^2 = \theta,\ \text{and}\ \ x^3 = \phi. \tag{4.1-3}$$

Define a metric tensor $\tilde{g}_{\mu\nu}$ for the 3-dimensional manifold S that results from freezing *t:*

$$\tilde{g}_{i\,j} \equiv -\,g_{i\,j}. \tag{4.1-4}$$

The line element for S is

$$ds^2 = \tilde{g}_{i\,j}\, dx^i\, dx^j = \frac{1}{1-\frac{2\,m}{r}}\, dr^2 + r^2\, d\theta^2 + r^2\, sin^2\theta\, d\phi^2. \tag{4.2}$$

$\tilde{g}_{i\,j}$ is a positive definite metric tensor:

Positive definite means that

$\tilde{g}_{i\,j} x^i x^j \geq 0\ \forall\ i, j$ and $\tilde{g}_{i\,j} x^i x^j = 0$ only if $x^i = 0$ for all i.

First, $\tilde{g}_{i\,i} > 0$ for $i = 1–3$:

$$1 - \frac{2\,m}{r} \overset{(3.7-24)}{>} 0 \quad \Rightarrow \quad g_{11} \overset{(4.1-2)}{<} 0 \quad \Rightarrow \quad \tilde{g}_{11} \overset{(4.1-4)}{>} 0.$$

Since $\tilde{g}_{ij}$ is not defined on the z-axis or at the origin, we also have

$\tilde{g}_{22} = r^2 > 0$ and $\tilde{g}_{33} = r^2 sin^2\theta > 0$. ✓

Next, $x^1 = r > 0$, $x^2 = \theta > 0$, and $x^3 = \phi \geq 0$.

Thus, $\tilde{g}_{ij} x^i x^j > 0$:

$\tilde{g}_{ij} x^i x^j = \tilde{g}_{11} x^1 x^1 + \tilde{g}_{22} x^2 x^2 + \tilde{g}_{33} x^3 x^3$

The third term is nonnegative but each of the first two terms is positive. ✓

This makes $\tilde{g}_{ij}$ positive definite. ✓

$\tilde{g}_{ij}$ being positive definite means that the 3-dimensional manifold S is a slice of space rather than spacetime. Moreover, none of the $\tilde{g}_{ij}$ depend on t, so we can refer to events with the same (r,θ,ϕ) but different t as occurring at the same point in space. We call these **fixed points of space**, and this process of separating spacetime into space and time is called a **spacetime split.** This is not a general feature of spacetime. However, when it happens, it is much easier to solve problems because we are working in static space rather than in non-static spacetime.

When $M = 0$, the line element (4.1) becomes that of flat spacetime in spherical coordinates, while the line element (4.2) becomes that of Euclidean space in spherical coordinates. As M grows, it causes increasing distortion in both spacetime and space so that neither is flat. Distortion is measured by the dimensionless quantity $\frac{2m}{r}$ and is maximized at the boundary of the object, the smallest the radius r can achieve.

The table below shows how much distortion is caused by objects of various radius, r_B, and mass, M. Distortion around our sun is small enough to ignore for most calculations but is large enough to show up in precision measurements of the perihelions of Mercury, Venus, and Earth and also in the deflection of light. Treating the distortion is necessary for white dwarfs, and of central relevance for a neutron star. The Schwarzschild equations must be modified in order to be relevant for black holes.

Table 4.1-1

Object	r_B (meters)	M (kg)	$2m$	$2m \,/\, r_B$	Distortion
Flat spacetime	-	-	-	0	None
Proton	$8\text{x}10^{-16}$	$2\text{x}10^{-27}$	$2\text{x}10^{-54}$	10^{-39}	Infinitessimal
Earth	$6\text{x}10^{6}$	$6\text{x}10^{24}$	$9\text{x}10^{-3}$	10^{-9}	Very tiny
Sun	$7\text{x}10^{8}$	$2\text{x}10^{30}$	$3\text{x}10^{4}$	10^{-6}	Tiny
White Dwarf	$7\text{x}10^{6}$	$2\text{x}10^{30}$	$2\text{x}10^{3}$	10^{-4}	Important
Neutron Star	$1\text{x}10^{4}$	$4\text{x}10^{30}$	$6\text{x}10^{3}$	.6	Huge
Black Hole	0	Up to 10^{40}	Up to 10^{13}	Undefined	Undefined

In flat spacetime, r is simply the distance from the origin, but for $M > 0$, r has a positive lower bound, the radius r_B, and the meaning of r is not so clear. What does it represent?

For $M > 0$, the space manifold S is curved. We consider two different submanifolds of S, a sphere and a line. Let Σ be a sphere in space having radius r. Since dr = 0 on Σ, equation (4.2) reduces to the same equation as the sphere Σ from Section 3.7 whose line element is

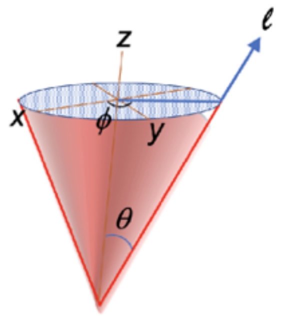
$$dL^2 \overset{(3.52)}{=} r^2\,(d\theta^2 + sin^2\theta\, d\phi^2). \tag{4.3}$$

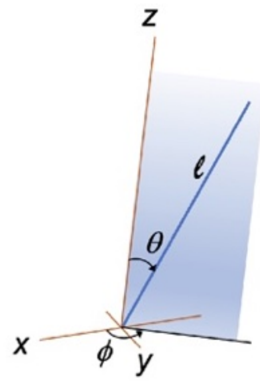
It follows that Σ has the 2-dimensional geometry of a sphere of radius r in flat Euclidean space, and, so, infinitesimal (tangential) distances between points on S are given by

$$dL = r\sqrt{d\theta^2 + sin^2\theta\, d\phi^2}\,. \tag{4.4}$$

That is, the distance between points is the familiar Euclidean distance.

However... let ℓ be the radial line generated by constant ϕ and constant θ. See the first figure below where ℓ is shown as the line at angle ϕ in the cone of angle θ, or the second figure that shows ℓ as the line at angle θ in the plane of constant ϕ, and also in the figure in Example 1.1.4.





The line element for ℓ can be obtained from equation (4.2) by setting dθ = dϕ = 0:

$$dR^2 = \frac{1}{1-\frac{2m}{r}}\, dr^2$$

and, so, the infinitesimal (tangential) distance between points is

$$\boxed{dR = \frac{1}{\sqrt{1-\frac{2m}{r}}}\, dr}. \tag{4.5}$$

Distance between points is obtained by integrating dR over a geodesic between the points. That is, R measures the distance between points. Because dR > dr, r no longer measures coordinate distance.

We have just learned that length can be Euclidean in one submanifold (the sphere) yet be non-Euclidean in another submanifold (the line) of the same space S. This conflicting story is explained by curvature.

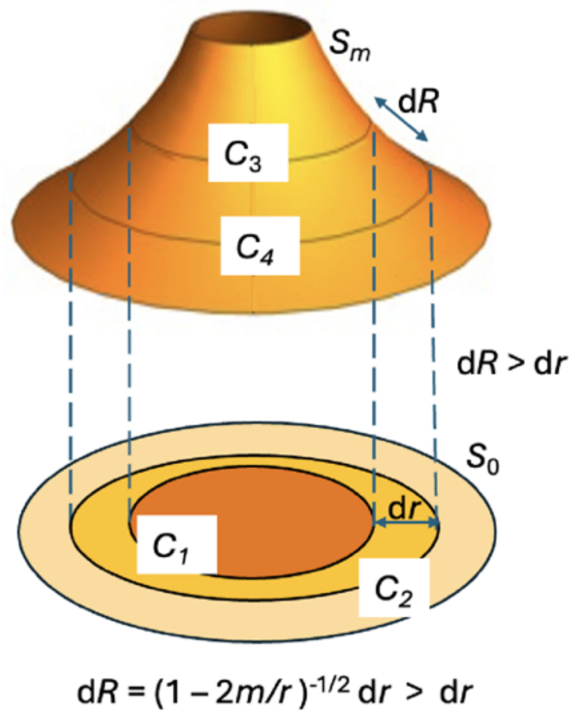


Figure 4.1 Radial Distance in the Schwarzschild Geometry

Figure 4.1 shows Euclidean spheres of radii r and $r + \mathrm{d}r$ embedded in both a flat space S_0 and a curved space S_m. The distance between the spheres is dr in flat Euclidean space but is $\mathrm{d}R > \mathrm{d}r$ in curved space, where dR is given in equation (4.5). The range of r is $(2m, \infty)$ and we see that $\lim_{r\to\infty} \mathrm{d}R = \mathrm{d}r$ and $\lim_{r\to 2m} \mathrm{d}R = \infty$. For a stellar object like a planet or a star, r cannot reach the lower limit $2m$ because it reaches r_B, the radius of the object, first. That is, $r > r_B$ and $r_B > 2m$.

We conclude that in the manifold S, r can measure length in some cases but be smaller than length in some other cases, distorted by a factor of $\frac{1}{\sqrt{1-\frac{2m}{r}}}$.

Switching topics now to time, recall that in special relativity, clocks record proper time τ along their world lines. For a stationary clock at location (r,θ,ϕ), we can find the relation between t *and* τ by setting $\mathrm{d}r = \mathrm{d}\theta = \mathrm{d}\phi = 0$ in equation (4.1):

$$\mathrm{d}\tau = \sqrt{1 - \frac{2m}{r}}\ \mathrm{d}t. \qquad (4.6)$$

Thus, in flat spacetime, where $m = 0$, $d\tau = dt$ records coordinate time. However, in curved spacetime, where $m > 0$, a stationary clock's measurement of coordinate time is distorted by a factor of $\sqrt{1 - \frac{2m}{r}}$. Notice that as $r \to \infty$, $d\tau \to dt$, and the distortion disappears.

4.8 Black Holes

The Schwarzschild solution uses coordinates (t,r,θ,ϕ). The metric tensor $g_{\mu\nu}$ does not depend on t, so the solution is static and we can assume t is a fixed finite number, possibly negative. Reviewing Table 4.1-1, we see that the value of r can decrease from infinity until it reaches either the boundary of the object, r_B, or the lower limit of $2m$, where $m = \frac{\text{GM}}{c^2}$.

Definition An object for which $r_B \leq 2m$ is called a **black hole,** and the distance $\frac{2\,\text{GM}}{c^2}$ is known as the **Schwarzschild radius**.

Recall the Schwarzschild line element, equation (4.1):

$$c^2\, \mathrm{d}\tau^2 \overset{(3.59)}{=} c^2\, (1 - \frac{2\,m}{r})\, \mathrm{d}t^2 - \frac{1}{1-\frac{2\,m}{r}}\, \mathrm{d}r^2 - r^2\, \mathrm{d}\theta^2 - r^2\, sin^2\theta\, \mathrm{d}\phi^2$$

We see that $g_{11} \to -\infty$ as $r \to 2m$. However, metric tensor singularities are coordinate-dependent. We might refer to these as “coordinate singularities”. For example, in Euclidean 3-space $\mathbb{E}^3$ we saw that the spherical coordinates dual basis line element exhibits singularities in the metric tensor at $r = 0$ and also at $\phi = 0$. But, these same points have finite metric tensor values in other coordinate systems like Euclidean coordinates (which has no singularities). We develop alternate Schwarszchild geometry coordinate systems below in which the metric tensor has a finite value at $r = 2\text{m}$, showing itis not a real singularity. We define a **singularity** as a coordinate singularity that holds for all coordinate system. (This can be difficult to show. For example, the Schwarzschild singularity at $r = 0$ is a real singularity.)

Because of the coordinate singularity, Schwarzschild coordinates (t,r,θ,ϕ) are inadequate for treating $r \leq 2m$. To explore null trajectories (i.e., photons) near the lower limit $2m$, various alternate coordinate systems for Schwarzschild geometry have been developed that do not have coordinate singularities at $r = 2m$, including Eddington-Finkelstein coordinates (the actual equations were developed by Roger Penrose), Kruskai-Szekeres coordinates, Lemaitre coordinates, and Gullistrand-Painleve coordinates. Vaidya has developed a generalized version of the Schwarzschild solution by extending the Schwarzschild line element to allow the mass M to vary as a function of the coordinates.

We develop the Eddington-Finkelstein coordinate systems. There are two coordinate systems, Out-going ("retarded") and In-going ("Advanced"). In the In-going coordinate system, we replace t by

$$v = ct + r + 2m \ln \left| \frac{r}{2m} - 1 \right|. \tag{4.60}$$

This system allows us to watch a photon fall into a black hole. The absolute value sign handles the situation when $r_B < r < 2m$.

In the Out-going coordinate system, we replace t by

$$u = ct - r - 2m \ln \left| \frac{r}{2m} - 1 \right|.$$

This system allows us to analyze null trajectories (photons) inside a black hole.

The Schwarzschild line element (4.1) for the In-going coordinate system becomes (Exercise 4.8.1)

$$c^2\, d\tau^2 = \left(1 - \frac{2m}{r}\right) dv^2 - 2\, dv\, dr - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2 \tag{4.61}$$

Notice that the coordinate singularity as $r \to 2m$ has disappeared.

Definition (v, r, θ, ϕ), where $r > r_B$, are called **Eddington-Finkelstein coordinates**.

These coordinates are defined for all v, θ, and ϕ, and for $r > r_B$.

Since $c^2\,(d\tau)^2 \equiv g_{\mu\nu}\, dx^\mu\, dx^\nu$ where $x^0 = v$, the metric tensor is

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{2m}{r} & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

When $r = 2m$, none of the metric tensor components becomes infinite, so unlike Schwarzschild coordinates, we can analyze photons that cross the Schwarzschild radius by setting $r_B < 2m$.

Photon trajectories are null geodesics ($d\tau = 0$).

Equation (4.61) $\Rightarrow d\theta = d\phi = 0$

$\Rightarrow \theta$ and ϕ are constants.

This simply means that an inward photon follows a radial path to the center of the black hole. The path is the line of intersection of the cone, θ = constant, and the plane, ϕ = constant.

The other part of dτ = 0 involves $\frac{dv}{dr}$:

$$0 = c^2 \left(\frac{d\tau}{dr}\right)^2 \overset{(4.61)}{=} (1 - \frac{2m}{r}) \left(\frac{dv}{dr}\right)^2 - 2\,\frac{dv}{dr} = \frac{dv}{dr}\,[\,(1 - \frac{2m}{r})\,\frac{dv}{dr} - 2\,]$$

$$\Rightarrow \quad \frac{dv}{dr} = 0 \qquad (4.8\text{-}1)$$

or

$$\frac{dv}{dr} = \frac{2}{1 - \frac{2m}{r}} \qquad (4.8\text{-}2)$$

$$\frac{dv}{dr} \overset{(4.60)}{=} c\,\frac{dt}{dr} + 1 + \frac{2m\,\frac{1}{2m}}{\frac{r}{2m} - 1} = c\,\frac{dt}{dr} + 1 + \frac{2m}{r-2m} = c\,\frac{dt}{dr} + \frac{r-2m+2m}{r-2m}$$

$$= c\,\frac{dt}{dr} + \frac{r}{r-2m} = c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2m}{r}} \qquad (4.8\text{-}3)$$

$$\frac{dv}{dr} = 0$$

$$\overset{(4.8-3)}{\Rightarrow} \; c\,\frac{dt}{dr} = -\frac{1}{1 - \frac{2m}{r}} < 0$$ if $r > 2m$ (i.e., for an in-going photon outside of $2m$)

$$\frac{dv}{dr} = \frac{2}{1 - \frac{2m}{r}}$$

$$\overset{(4.8-3)}{\Rightarrow} \; c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2m}{r}} = \frac{2}{1 - \frac{2m}{r}} = c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2m}{r}} > 0$$

if $r > 2m$ (i.e., for an out-going photon outside of $2m$)

Integrating equation (4.8-1) yields

$v = A$ (where A is a constant) (In-going) (4.8-4)

Observe that equation (4.8-2) can be written $\frac{dv}{dr} = 2 + \frac{4m}{r-2m}$:

$$\frac{dv}{dr} = \frac{2}{1-\frac{2m}{r}} = \frac{2r}{r-2m} = \frac{2r-4m}{r-2m} + \frac{4m}{r-2m} = 2 + \frac{4m}{r-2m} \quad \checkmark$$

So, integrating equation (4.8-2) yields

$v = 2r + 4\text{m} \ln|r - 2m| + B$ (where B is a constant) (Out-going) (4.8-5)|

We observe that for radial null rays (4.8-1) and (4.8-2), as $r \to \infty$, dv/dr → 0 and ±2, respectively, not ±1 as one might expect if one regarded v as "time".

Equations (4.8-4) and (4.8-5) enable us to draw spacetime plots having axes v and r. Just as we did in flat spacetime, we draw null radials $v = A$ at 45°. This includes the r-axis, which is $v = 0$. Figure 4.1.3, to the right of the vertical line $r = 2m$, shows the in-going and out-going light rays. The out-going null radials approach -∞ at $r = 2m$.

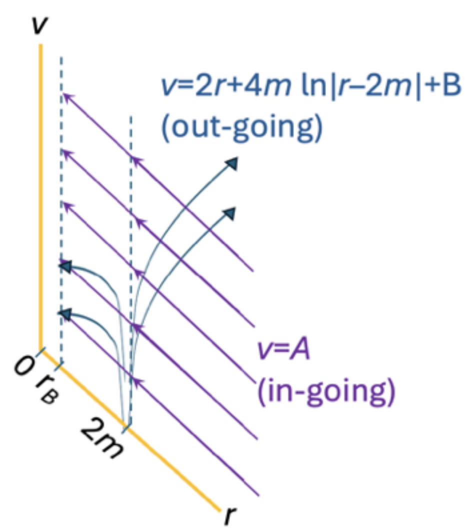


Figure 4.1.3 Eddington-Finkelstein In-going and Out-going Null Geodesics

In addition, equations (4.8-4) and (4.8-5) can be valid to the left of the vertical line $r = 2m$ by letting $r_B < 2m$ since both equations are well-defined for $0 \le r_B < r < 2m$. In both cases, $\frac{dv}{dr} < 0$ when $r < 2m$, so Figure 4.8-1 appropriately plots both in-going and out-going trajectories as in-going. This illustrates what is meant when one says that light cannot escape from inside a black hole. Because light cannot escape from a sphere of radius $2m$, an outside observer cannot see such events, and we make the following definition.

Definition The hyperspace “sphere” of “radius” *r* in spacetime is called an **Event Horizon**.