Chapter 5

The Equivalence Principle

There was a minute's pause perhaps. The Psychologist seemed about to speak to me, but changed his mind. Then the Time Traveller put forth his finger towards the lever. 'No,' he said suddenly. 'Lend me your hand.' And turning to the Psychologist, he took that individual's hand in his own and told him to put out his forefinger. So that it was the Psychologist himself who sent forth the model Time Machine on its interminable voyage. We all saw the lever turn. I am absolutely certain there was no trickery. There was a breath of wind, and the lamp flame jumped.

5.1 Inertial mass

"Gravitational mass is equivalent to inertial mass."

Newton's second law states that the force on an object is proportional to mass times acceleration. In this section, we will call the mass which appears in this law the *inertial mass*:

$$\overrightarrow{\mathbf{F}} = m_I \overrightarrow{\mathbf{a}}. \tag{5.1}$$

for example, consider the electro-static interaction between two particles with masses m_1, m_2, and charges q_1, q_2. Particle 2 feels a force

$$\overrightarrow{\mathbf{F}}_2 = \frac{q_1 q_2}{r_{12}^2} \hat{r}_{12}. \tag{5.2}$$

The acceleration felt by particle 2 can then be found by combining these two equations:

$$\overrightarrow{\mathbf{a}}_2 = \left(\frac{q_2}{m_{2\ I}} \right) \frac{q_1}{r_{12}^2} \hat{r}_{12}. \tag{5.3}$$

Thus, $\overrightarrow{\mathbf{a}}_2$ depends on the ratio of charge to inertial mass, $q_2/m_{2\ I}$.

Next, consider a particle falling to the ground. The Newtonian gravitational force resembles the electrostatic force (both are inverse square laws), with masses replacing electrical charges. We will simply call the gravitational charge m. Let $M_\oplus$ be the mass of the Earth, r the distance to the centre of the Earth, and $\hat{r}$ the upwards unit vector.

Then the force on an apple falling to the ground can be written

$$\overrightarrow{\mathbf{F}} = -\frac{CM_{\oplus}m}{r^2}\hat{r}, \tag{5.4}$$

where C is a constant.

Now, by Newton's 2nd law

$$\overrightarrow{\mathbf{F}} = m_I \overrightarrow{\mathbf{a}} \tag{5.5}$$

$$\Longrightarrow \overrightarrow{\mathbf{a}} = \left(\frac{m}{m_I}\right)\left(-C\frac{M_{\oplus}}{r^2}\right)\hat{r}. \tag{5.6}$$

Thus, $\overrightarrow{\mathbf{a}}$ depends on the ratio of gravitational mass to inertial mass, m/m_I. Galileo's experiments $\overrightarrow{\mathbf{a}}$ should be the same for *all* materials, once air resistance is neglected. This implies that m/m_I is the same constant for *all* matter.

Thus, we can combine the two constants C and m/m_I into a single constant

$$G = C\left(\frac{m}{m_I}\right) \tag{5.7}$$

$$\Longrightarrow \overrightarrow{\mathbf{a}} = -\frac{GM_{\oplus}}{r^2}\hat{r}. \tag{5.8}$$

We now see a fundamental difference between the electrical force and the gravitational force: the former depends on a charge-to-inertial-mass ratio, but the latter does not. The inertial forces (also known as fictitious forces) are similar to gravity in this respect – the acceleration has no dependence on mass.

5.2 Free Fall

Einstein had difficulty incorporating both gravity and inertial forces into special relativity. His great insight was to treat them together, using the principle of equivalence to eliminate the dependence on mass. Now, inertial forces can be eliminated by transferring to a non-accelerating frame. Einstein reasoned that gravitational forces can be removed in a similar way by transferring to a free-fall frame.

Example: Consider an object in the Earth's gravitational field near $r = R_{\oplus}$. Let z be the vertical direction.

$$\frac{d^2z}{dt^2} = -g$$

$$g = \frac{GM_{\oplus}}{R_{\oplus}}$$

$$\approx 10ms^{-1}$$

initial conditions:

$$\begin{aligned} z_0 &= z(t=0) = h \\ \dot{z}_0 &= V_0 \\ \Longrightarrow \quad \dot{z}(t) &= -gt + V_0 \\ z(t) &= h - \frac{1}{2}gt^2 + V_0 t \end{aligned}$$

Let's transform to new co-ordinates:

$$\xi(z,t) = z + \frac{1}{2}gt^2$$

Then:

$$\begin{aligned} \xi &= h + V_0 t \\ \dot{\xi} &= V_0 = \text{constant} \\ \ddot{\xi} &= 0 \end{aligned}$$

Since the acceleration is zero, there is no gravitational force.

5.2.1 Locally Inertial Frames

noindent**Definition** *Locally Inertial Frame (LIF)* An LIF is a reference frame with origin at space-time event P. An object at P is in free-fall (if there are no external forces). Near P, there are *no* gravitational forces, and special relativity holds.

In an LIF

a.
$$g_{ab}\Big|_P = \eta_{ab}, \tag{5.9}$$

b.
$$\partial_c g_{ab}\Big|_P = 0. \tag{5.10}$$

for all a, b, c. The second condition follows from the isotropy of space-time.

5.3 Geodesics

noindent**Definition** *Geodesic*

a. A geodesic is a path on a manifold M which is an extremum of length (i.e. max or min distance between two points).

b. A geodesic is a path which has zero covariant acceleration (to be defined later).

5.3.1 Examples

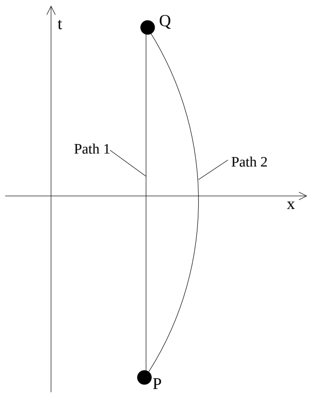
a. $\mathbb{R}^2$. Geodesics are straight lines.

b. S^2. Geodesics are the great circles, for example the equator.

c. Minkowski Space M^4.

In special relativity, "distance" becomes proper time, τ

$$d\tau^2 = ds^2 = dt^2 - dx^2 - dy^2 - dz^2$$



Path 1:

$$\tau_1 = \int d\tau = \int dt = \Delta t. \tag{5.11}$$

Path 2:

$$\tau_2 = \int \sqrt{dt^2 - dx^2 - dy^2 - dz^2} \tag{5.12}$$

$$< \int \sqrt{dt^2} = \Delta t. \tag{5.13}$$

Thus, $\tau_1 > \tau_2$. An object at rest has *maximum* proper time (and follows a geodesic in space-time).

5.3.2 The Geodesic Equation

In general relativity, the orbit of a satellite is a geodesic in a space-time distorted by the mass of the Earth. How can we find the orbit? We seek an expression for the acceleration of the satellite in ordinary coordinates fixed to the Earth. In the locally inertial (free-fall) frame, of course, the acceleration is easy: it is zero!

Let $(\xi^0, \xi^1, \xi^2, \xi^3)$ be coordinates in the satellite's LIF, while X are coordinates fixed to the Earth. In other words, near the spacetime event $(\xi^0, \xi^1, \xi^2, \xi^3) = (0, 0, 0, 0)$, the satellite is at rest in the LIF coordinates, and experiences no forces and no acceleration:

$$U^a_{LIF} = \frac{\mathrm{d}\xi^a}{\mathrm{d}\tau} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \tag{5.14}$$

$$\frac{\mathrm{d}U^a_{LIF}}{\mathrm{d}\tau} = \frac{\mathrm{d}^2\xi^a}{\mathrm{d}\tau^2} = 0. \tag{5.15}$$

Let $\overline{\mathbf{U}}$ be the 4-velocity in Earth coordinates X. We now transform $\overline{\mathbf{U}}_{LIF}$ to $\overline{\mathbf{U}}$:

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau}U^d_{LIF} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\xi^d}{\partial\mathsf{X}^b}U^b\right) \tag{5.16}$$

$$= \left(\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\xi^d}{\partial\mathsf{X}^b}\right)U^b + \frac{\partial\xi^d}{\partial\mathsf{X}^b}\frac{\mathrm{d}}{\mathrm{d}\tau}U^b \tag{5.17}$$

by the product rule. To understand the first term, we use the fact that

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{dt}{d\tau}\frac{\partial}{\partial t} + \frac{dx}{d\tau}\frac{\partial}{\partial x} + \dots \tag{5.18}$$

$$= U^0\partial_0 + U^1\partial_1 + \dots \tag{5.19}$$

$$= U^c\partial_c \tag{5.20}$$

to obtain

$$0 = U^b\left(U^c\partial_c\frac{\partial\xi^d}{\partial\mathsf{X}^b}\right) + \frac{\partial\xi^d}{\partial\mathsf{X}^b}\frac{\mathrm{d}}{\mathrm{d}\tau}U^b \tag{5.21}$$

$$= U^bU^c\frac{\partial^2\xi^d}{\partial\mathsf{X}^b\partial\mathsf{X}^c} + \frac{\partial\xi^d}{\partial\mathsf{X}^b}\frac{\mathrm{d}}{\mathrm{d}\tau}U^b. \tag{5.22}$$

The last term contains what we are searching for: the acceleration in the Earth frame $\mathrm{d}U^b/\mathrm{d}\tau$. However, we need to free this expression from the transformation matrix $\partial\xi^d/\partial\mathsf{X}^b$. To rid ourselves of this unwanted matrix, we multiply by its inverse $\partial\mathsf{X}^a/\partial\xi^d$:

$$\frac{\partial\mathsf{X}^a}{\partial\xi^d}\frac{\partial\xi^d}{\partial\mathsf{X}^b} = \delta^a{}_b. \tag{5.23}$$

This gives

$$0 = \frac{\partial\mathsf{X}^a}{\partial\xi^d}\left(U^bU^c\frac{\partial^2\xi^d}{\partial\mathsf{X}^b\mathsf{X}^c} + \frac{\partial\xi^d}{\partial\mathsf{X}^b}\frac{\mathrm{d}}{\mathrm{d}\tau}U^b\right) \tag{5.24}$$

$$= \frac{\partial\mathsf{X}^a}{\partial\xi^d}\frac{\partial^2\xi^d}{\partial\mathsf{X}^b\partial\mathsf{X}^c}U^bU^c + \delta^a{}_b\frac{\mathrm{d}}{\mathrm{d}\tau}U^b. \tag{5.25}$$

Now, $\delta^a{}_b\,\mathrm{d}U^b/\mathrm{d}\tau = \mathrm{d}U^a/\mathrm{d}\tau$. Rearranging terms, we finally obtain

$$\boxed{\frac{\mathrm{d}U^a}{\mathrm{d}\tau} + \frac{\partial\mathsf{X}^a}{\partial\xi^d}\frac{\partial^2\xi^d}{\partial\mathsf{X}^b\partial\mathsf{X}^c}U^bU^c = 0}. \tag{5.26}$$

Equation (5.26) can be written in the form

$$\frac{\mathrm{d}U^a}{\mathrm{d}\tau} + \Gamma^a{}_{bc} U^b U^c = 0, \tag{5.27}$$

where the *Christoffel symbols* $\Gamma^a{}_{bc}$ are given by

$$\Gamma^a{}_{bc} = \frac{\partial \mathsf{X}^a}{\partial \xi^d} \frac{\partial^2 \xi^d}{\partial \mathsf{X}^b \partial \mathsf{X}^c}. \tag{5.28}$$

Equation (5.26) is called the *geodesic equation*, and governs the motion of matter in the absence of forces. A more useful formula for the Christoffel symbols will be derived in the exercise below, and (in a different way) in the next chapter. Equation (5.26) has been introduced here because of its physical meaning. It computes the apparent acceleration $\mathrm{d}U^a / \mathrm{d}\tau$ of an object in one frame (X) in terms of transformations from the LIF. We experience gravitational forces only because we insist on viewing things from a non-inertial frame! Like any other fictitious or inertial force, gravitation arises from the acceleration of one frame of reference with respect to the inertial frames. We do feel the effects of weight, particularly after a long hike uphill, but we can now view this as the effect of forces coming from the ground under our feet, accelerating us away from our natural state – free-fall.

Exercise 5.1 *Here we derive an expression for the Christoffel symbols in terms of the metric. To simplify the notation, g_{ab} will denote the metric in the non-inertial frame, g_{Lab} the metric in the LIF, and $\partial_a \equiv \partial/\partial \mathsf{X}^a$ (i.e. the shorthand for partial derivatives applies only to the noninertial X frame).*

a. Show that

$$\partial_c g_{ab} = \left[(\partial_c \partial_a \xi^f)(\partial_b \xi^g) + (\partial_a \xi^f)(\partial_c \partial_b \xi^g)\right] g_{Lfg}. \tag{5.29}$$

b. Show that

$$\Gamma_{abc} = g_{ae} \Gamma^e{}_{bc} = (\partial_a \xi^f)(\partial_b \partial_c \xi^g) g_{Lfg}. \tag{5.30}$$

c. Show that

$$\Gamma_{abc} + \Gamma_{acb} = 2\partial_c g_{ab}. \tag{5.31}$$

d. Hence show that

$$\Gamma^a{}_{bc} = \frac{1}{2} g^{ad} \left(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}\right). \tag{5.32}$$

5.3.3 Covariant Acceleration

Theorem. If the components of a tensor vanish in one co-ordinate system, then they vanish in *all* frames.

Proof. This follows directly from the transformation laws for a tensor. For example, if in frame A we have $M^{ab}_{\mathsf{A}} = 0$ for all a, b, then in frame B

$$M^{cd}_{\mathsf{B}} = \frac{\partial \mathsf{B}^c}{\partial \mathsf{A}^a} \frac{\partial \mathsf{B}^d}{\partial \mathsf{A}^b} (0)^{ab} = 0. \tag{5.33}$$

Note that the acceleration term $dU^a/d\tau$ is zero in the LIF, but non-zero in the Earth frame. Thus it is not a tensor! To understand the motion of objects in general relativity further, we must learn how to differentiate vectors in a covariant way (i.e. so that the result is a tensor).

A vector $\overline{\mathbf{U}}$ involves not only its components U^a, but also the basis vectors of the coordinate system. If space-time is warped, or even if we are simply using non-Cartesian coordinates, these basis vectors will point in different directions at different points. In the next chapter, we will show how to differentiate vectors by including both components and basis vectors.

Chapter 6

Covariant Derivatives

One of the candles on the mantel was blown out, and the little machine suddenly swung round, became indistinct, was seen as a ghost for a second perhaps, as an eddy of faintly glittering brass and ivory; and it was gone–vanished! Save for the lamp the table was bare.

How do we differentiate vectors (& tensors)?

There are many examples in physics where derivatives need to be extended. For example, in fluid mechanics the Navier-Stokes force equation reads

$$\frac{D\overrightarrow{\mathbf{V}}}{Dt} = \nabla p + \nu \nabla^2 \overrightarrow{\mathbf{V}} \tag{6.1}$$

where D/Dt is the total Lagrangian derivative

$$D/Dt = \partial/\partial t + \overrightarrow{\mathbf{V}} \cdot \nabla. \tag{6.2}$$

In Quantum mechanics, Schrödinger's equation has the form

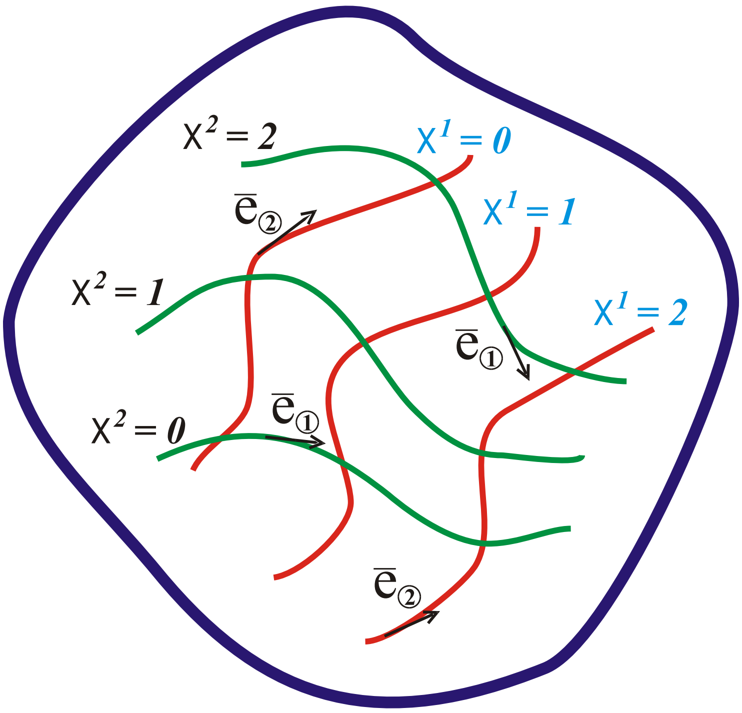
$$\left(\frac{1}{2m} \left(\frac{\hbar}{\mathrm{i}} \nabla \right)^2 + V \right) \psi = E\psi. \tag{6.3}$$

With an applied magnetic field the gradient ∇ is replaced by the 'gauge covariant derivative' $\nabla - \frac{ie}{m}\overrightarrow{\mathbf{A}}$, where $\nabla \times \overrightarrow{\mathbf{A}} = \overrightarrow{\mathbf{B}}$. In this context the vector potential $\overrightarrow{\mathbf{A}}$ is sometimes called the 'electromagnetic connection'.

6.1 Non-Euclidean Geometry

Basis Vectors A *Co-ordinate Line* is a line parameterized by one of the co-ordinates. A *basis vector* is a tangent vector to a co-ordinate line. Let $\overline{\mathbf{e}}_{ⓐ}$ be the basis vector tangent to the co-ordinate line following x^a .

Note A circled subscript appears because the subscript chooses between the vectors in a set – e.g. the set $\{\overline{\mathbf{e}}_{⓪}, \overline{\mathbf{e}}_{①}, \overline{\mathbf{e}}_{②}, \overline{\mathbf{e}}_{③}\}$ – ordinary subscripts choose the component of a single vector.



Thus, in component form (for co-ordinates X)

$$\bar{\mathbf{e}}_{\textcircled{1}} = \begin{pmatrix} \mathrm{d}\mathsf{X}^0/\mathsf{X}^1 \\ \mathrm{d}\mathsf{X}^1/\mathsf{X}^1 \\ \mathrm{d}\mathsf{X}^2/\mathsf{X}^1 \\ \mathrm{d}\mathsf{X}^3/\mathsf{X}^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{6.4}$$

$$\Longrightarrow \bar{\mathbf{e}}_{\textcircled{1}}{}^0 = 0, \bar{\mathbf{e}}_{\textcircled{1}}{}^1 = 1, \ldots \tag{6.5}$$

$$\text{or} \quad \bar{\mathbf{e}}_{\textcircled{1}}{}^a = \delta^a_1 \tag{6.6}$$

In general,

$$\boxed{\bar{\mathbf{e}}_{\textcircled{c}}{}^a = \delta^a_c}. \tag{6.7}$$

Note that the basis vectors need not be orthogonal or of unit size:

$$\bar{\mathbf{e}}_{\textcircled{b}} \cdot \bar{\mathbf{e}}_{\textcircled{c}} = g_{ad}\bar{\mathbf{e}}_{\textcircled{b}}{}^a \bar{\mathbf{e}}_{\textcircled{c}}{}^d \tag{6.8}$$

$$= g_{ad}\delta^a_b \delta^d_c \tag{6.9}$$

$$= g_{bc} \tag{6.10}$$

i.e. the scalar product of basis vectors $\bar{\mathbf{e}}_{\textcircled{b}}$ and $\bar{\mathbf{e}}_{\textcircled{c}}$ is equal to element g_{bc} of the metric.

All vectors can be written as sums of basis vectors:

$$\bar{\mathbf{V}} = \begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} = V^c \bar{\mathbf{e}}_{\textcircled{c}}. \tag{6.11}$$

6.2 The Covariant Derivative

Derivatives must satisfy the product rule. Thus for the derivative in the X^b direction

$$\nabla_b \overline{\mathbf{V}} = \nabla_b\big(V^c \overline{\mathbf{e}}_{\textcircled{c}}\big) = (\nabla_b V^c)\, \overline{\mathbf{e}}_{\textcircled{c}} + V^c(\nabla_b \overline{\mathbf{e}}_{\textcircled{c}}). \tag{6.12}$$

V^c is a number at each point (i.e. a function of position), so we can write

$$\begin{aligned} (\nabla_b V^c) &= \partial_b V^c = \frac{\partial V^c}{\partial \mathsf{X}^b} && (6.13)\\ \Longrightarrow \big(\nabla_b \overline{\mathbf{V}}\big) &= (\partial_b V^c)\, \overline{\mathbf{e}}_{\textcircled{c}} + V^c\left(\nabla_b \overline{\mathbf{e}}_{\textcircled{c}}\right). && (6.14)\end{aligned}$$

We need to define the last term in brackets. The object $\nabla_b \overline{\mathbf{e}}_{\textcircled{c}}$ is itself a vector. We will name its components with the capital Greek letter Γ:

$$\nabla_b \overline{\mathbf{e}}_{\textcircled{c}} = \begin{pmatrix} \Gamma^0{}_{bc} \\ \Gamma^1{}_{bc} \\ \Gamma^2{}_{bc} \\ \Gamma^3{}_{bc} \end{pmatrix}. \tag{6.15}$$

In terms of the basis vectors,

$$\nabla_b \overline{\mathbf{e}}_{\textcircled{c}} = \Gamma^0{}_{bc}\overline{\mathbf{e}}_{\textcircled{0}} + \Gamma^1{}_{bc}\overline{\mathbf{e}}_{\textcircled{1}} + \Gamma^2{}_{bc}\overline{\mathbf{e}}_{\textcircled{2}} + \Gamma^3{}_{bc}\overline{\mathbf{e}}_{\textcircled{3}}, \tag{6.16}$$

or

$$\boxed{\nabla_b \overline{\mathbf{e}}_{\textcircled{c}} = \Gamma^a{}_{bc}\overline{\mathbf{e}}_{\textcircled{a}}}\,. \tag{6.17}$$

The object $\Gamma^a{}_{bc}$ is called a *metric connection*, or alternatively a *Christoffel symbol.*

Let us go back to calculating the gradient of a vector:

$$\begin{aligned} \nabla_b \overline{\mathbf{V}} &= (\partial_b V^c)\, \overline{\mathbf{e}}_{\textcircled{c}} + V^c\left(\nabla_b \overline{\mathbf{e}}_{\textcircled{c}}\right) && (6.18)\\ &= (\partial_b V^c)\, \overline{\mathbf{e}}_{\textcircled{c}} + V^c\left(\Gamma^a{}_{bc}\overline{\mathbf{e}}_{\textcircled{a}}\right). && (6.19)\end{aligned}$$

Exchange $a \leftrightarrow c$ in the 1st term on the RHS:

$$\begin{aligned} \Longrightarrow \nabla_b \overline{\mathbf{V}} &= (\partial_b V^a)\, \overline{\mathbf{e}}_{\textcircled{a}} + V^c \Gamma^a{}_{bc}\overline{\mathbf{e}}_{\textcircled{a}} && (6.20)\\ &= (\partial_b V^a + V^c\Gamma^a{}_{bc})\, \overline{\mathbf{e}}_{\textcircled{a}}. && (6.21)\end{aligned}$$

This is a vector, with components

$$\boxed{\big(\nabla_b \overline{\mathbf{V}}\big)^a = (\partial_b V^a + \Gamma^a{}_{bc}V^c)}$$

<u>To recap:</u> The covariant derivative, ∇_b (derivatives in X^b direction)

a. Produces tensors.

b. Obeys the product rule.

c. For a scalar function f,

$$\nabla_b f = \partial_b f = \frac{\partial f}{\partial \mathsf{X}^b}. \tag{6.22}$$

d. There exists a set of numbers $\Gamma^a{}_{bc}$, where

$$\big(\nabla_b \overline{\mathbf{V}}\big)^a = \partial_b V^a + \Gamma^a{}_{bc}V^c.$$

6.3 Derivatives of Other Tensors

Use the properties listed above.

Example: Given a 2nd rank tensor, M_{cd}, find $(\nabla_b M)_{cd}$.

To do this, let V^c, W^d be arbitrary vectors. Let $f = M_{cd}V^cW^d$ be a scalar function. By the product rule for ∂_b,

$$\partial_b f = (\partial_b M_{cd}) V^c W^d + M_{cd} (\partial_b V^c) W^d + M_{cd} V^c \left(\partial_b W^d\right). \tag{6.23}$$

Also by the product rule for ∇_b,

$$\nabla_b f = (\nabla_b M)_{cd} V^c W^d + M_{cd} \left(\nabla_b \overline{\mathbf{V}}\right)^c W^d + M_{cd} V^c \left(\nabla_b \overline{\mathbf{W}}\right)^d. \tag{6.24}$$

Meanwhile, $\nabla_b f = \partial_b f$, so

$$\begin{aligned}(\nabla_b M)_{cd} V^c W^d &= (\partial_b M_{cd}) V^c W^d + M_{cd} \left(\partial_b V^c - \left(\nabla_b \overline{\mathbf{V}}\right)^c\right) W^d \\ &+ M_{cd} V^c \left(\partial_b W^d - \left(\nabla_b \overline{\mathbf{W}}\right)^d\right).\end{aligned}$$

Now apply rule(4):

$$(\nabla_b M)_{cd} V^c W^d = (\partial_b M_{cd}) V^c W^d + M_{cd} \left[\left(-\Gamma^c{}_{ba} V^a\right) W^d + V^c \left(-\Gamma^d{}_{ba} W^a\right)\right]. \tag{6.25}$$

Next we factor out V and W. To do this, we swap $a \leftrightarrow c$ in the second to last term, and $a \leftrightarrow d$ in the last term:

$$(\nabla_b M)_{cd} V^c W^d = (\partial_b M_{cd}) V^c W^d - \left(\Gamma^a{}_{bc} M_{ad} + \Gamma^a{}_{bd} M_{ca}\right) V^c W^d. \tag{6.26}$$

As the above is true for any values of V^c and W^d we can cancel $V^c W^d$ from both sides, with the final result

$$\boxed{(\nabla_b M)_{cd} = \partial_b M_{cd} - \Gamma^a{}_{bc} M_{ad} - \Gamma^a{}_{bd} M_{ca}}. \tag{6.27}$$

In general, we will obtain one Γ for each index of a tensor, with a minus sign for each subscript (form) index and a plus sign for each superscript (vector) index.

Exercise 6.1 *The* covariant derivative ∇_b *of a vector field* X^a *is*

$$\nabla_b X^a = \partial_b X^a + \Gamma^a_{bc} X^c. \tag{6.28}$$

Also, the covariant derivative of a scalar is the same as the partial derivative:

$$\nabla_a f = \partial_a f. \tag{6.29}$$

Use these equations to find an expression for the covariant derivative ∇_b *of a form* W_c*, i.e. find* $\nabla_b W_c$*.*

Exercise 6.2 *In general Relativity, the relation between the Faraday tensor and the vector potential* ϕ_a *is defined by*

$$F_{ab} = (\nabla_b \phi)_a - (\nabla_a \phi)_b. \tag{6.30}$$

Write out the right hand side in terms of Christoffel symbols. Show that the Christoffel terms cancel, leaving

$$F_{ab} = (\partial_b \phi)_a - (\partial_a \phi)_b. \tag{6.31}$$

Theorem: Suppose the ∇ operator (covariant derivative) satisfies

a.
$$\nabla g = 0 \qquad (g = \text{metric}) \tag{6.32}$$

b.
$$\Gamma^a{}_{bc} = \Gamma^a{}_{cb} \qquad (\text{“zero torsion”}) \tag{6.33}$$

then the Christoffel symbols are determined by

$$\boxed{\Gamma^a{}_{bc} = \frac{1}{2} g^{ad} \left(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc} \right)}\,. \tag{6.34}$$

Proof: g_{cd} is a 2nd order tensor with two lower indices, so

$$(\nabla_b g)_{cd} = \partial_b g_{cd} - \Gamma^a{}_{bc} g_{ad} - \Gamma^a{}_{bd} g_{ca}. \tag{6.35}$$

Define

$$\Gamma_{dbc} \equiv g_{ad} \Gamma^a{}_{bc}. \tag{6.36}$$

Then

$$\begin{aligned} (\nabla_b g)_{cd} &= \partial_b g_{cd} - \Gamma_{dbc} - \Gamma_{cbd} && (6.37)\\ &= 0 && (6.38)\\ \implies \partial_b g_{cd} &= \Gamma_{dbc} + \Gamma_{cbd}. && (6.39) \end{aligned}$$

Next apply assumption 2: $\Gamma_{cdb} = \Gamma_{cbd}$, so

$$\partial_b g_{cd} = \Gamma_{dbc} + \Gamma_{cdb}. \tag{6.40}$$

Obtain two more equations by cycling $b \to c$, $c \to d$, and $d \to b$:

$$\begin{aligned} \implies \partial_c g_{db} &= \Gamma_{bcd} + \Gamma_{dbc} && (6.41)\\ \partial_d g_{bc} &= \Gamma_{cdb} + \Gamma_{bcd}. && (6.42) \end{aligned}$$

Take the sum (equation (6.40) + equation (6.41) - equation (6.42)):

$$\begin{aligned} \partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc} &= 2\Gamma_{dbc} + 0 + 0 && (6.43)\\ \implies \Gamma_{dbc} &= \frac{1}{2} \left(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc} \right). && (6.44) \end{aligned}$$

Finally, use $g_{db} = g_{bd}$ in the second to last term and let

$$\Gamma^a{}_{bc} = g^{ad} \Gamma_{dbc} \tag{6.45}$$

to prove the theorem.

Exercise 6.3 *Consider a sphere of radius 1 as a 2-dimensional manifold with coordinates $\mathsf{X}^1 = \theta$ (colatitude) and $\mathsf{X}^2 = \phi$ (longitude). What is the metric g_{ab} and its inverse*

g^{bc}? Find the Christoffel symbols $\Gamma^a{}_{bc}$(there are 8 of these for a 2-dimensional manifold). Suppose a geodesic on the sphere is parameterized by λ. Use the geodesic equation

$$\frac{\mathrm{d}^2 \mathsf{X}^a}{\mathrm{d}\lambda^2} + \Gamma^a{}_{bc}\frac{\mathrm{d}\mathsf{X}^b}{\mathrm{d}\lambda}\frac{\mathrm{d}\mathsf{X}^c}{\mathrm{d}\lambda} = 0 \tag{6.46}$$

to find $\frac{\mathrm{d}^2\theta}{\mathrm{d}\lambda^2}$ and $\frac{\mathrm{d}^2\theta}{\mathrm{d}\lambda^2}$.

6.3.1 The gradient of the metric in General Relativity

For an object in free-fall (no external forces) the geodesic equation gives

$$\frac{\mathrm{d}U^a}{\mathrm{d}\tau} + \Gamma^a{}_{bc}U^bU^c = 0. \tag{6.47}$$

This is true in all coordinate frames. But in the LIF the object will appear to be at rest (or in uniform motion). Thus the components of $\overline{\mathbf{U}}$ in the LIF remain constant, i.e. $\mathrm{d}U^a/\mathrm{d}\tau = 0$. This implies that

$$\Gamma^a{}_{bc}U^bU^c = 0 \qquad \text{in LIF.} \tag{6.48}$$

Since this holds for arbitrary 4-velocities $\overline{\mathbf{U}}$, we must have the Christoffel symbols vanishing,

$$\Gamma^a{}_{bc} = 0 \qquad \text{in LIF.} \tag{6.49}$$

(Strictly speaking, only the part of $\Gamma^a{}_{bc}$ symmetric in the lower indices b and c need vanish. Einstein's theory employs the simplest assumption, that the antisymmetric part of $\Gamma^a{}_{bc}$ – called the torsion – is always zero. Some modified theories of gravity include a non-zero torsion.)

Now,

$$(\nabla_c g)_{ab} = \partial_c g_{ab} - \Gamma^d{}_{ca}g_{db} - \Gamma^d{}_{cb}g_{ad}. \tag{6.50}$$

In the LIF, however, $\Gamma^d{}_{ca} = \Gamma^d{}_{cb} = 0$. Also, by the definition of locally inertial frames, $\partial_c g_{ab} = 0$. Thus

$$\nabla g = 0. \tag{6.51}$$

And, since ∇g is a tensor, it must vanish in *all* frames. Thus in General Relativity equation (6.34) can be used to calculate the Christoffel symbols.

6.4 Covariant Directional Derivatives and Acceleration

Consider a curve γ parameterized by λ whose tangent vector is $\overline{\mathbf{V}}$ (see section 1.5.1). For a function f, the derivative in the direction along the curve is (Recall equation (1.76)):

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \overline{\mathbf{V}}\cdot\nabla f \tag{6.52}$$

$$= V^b\partial_b f. \tag{6.53}$$

For the directional derivative of a vector along the curve, the covariant gradient ∇ now includes Christoffel symbols. Thus for a vector $\overline{\mathbf{W}}$,

$$\frac{\mathrm{D}\overline{\mathbf{W}}}{\mathrm{D}\lambda} = \overline{\mathbf{V}} \cdot \nabla \overline{\mathbf{W}}, \tag{6.54}$$

or, expressed in components,

$$\begin{aligned} \left(\frac{\mathrm{D}\overline{\mathbf{W}}}{\mathrm{D}\lambda}\right)^a &= \left(\overline{\mathbf{V}} \cdot \nabla \overline{\mathbf{W}}\right)^a & (6.55) \\ &= V^b \left(\nabla_b \overline{\mathbf{W}}\right)^a & (6.56) \\ &= V^b \left(\partial_b W^a + \Gamma^a{}_{bc} W^c\right) & (6.57) \\ &= \frac{\mathrm{d}W^a}{\mathrm{d}\lambda} + \Gamma^a{}_{bc} V^b W^c. & (6.58) \end{aligned}$$

We can apply this to find the acceleration 4-vector. Let the world-line of an object be parameterized by its proper time τ, with tangent vector the 4-velocity $\overline{\mathbf{U}}$. Then the 4-acceleration is

$$\begin{aligned} \overline{a} &= \frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\tau} = \overline{\mathbf{U}} \cdot \nabla \overline{\mathbf{U}}; & (6.59) \\ a^a &= U^b \left(\nabla_b \overline{\mathbf{U}}\right)^a = U^b \left(\partial_b U^a + \Gamma^a{}_{bc} U^c\right) & (6.60) \\ &= \frac{\mathrm{d}U^a}{\mathrm{d}\tau} + \Gamma^a{}_{bc} U^b U^c. & (6.61) \end{aligned}$$

Note the similarity with the geodesic equation, equation (6.47): The geodesic equation can now be written in a very simple form:

$$\boxed{\overline{a} = 0}. \tag{6.62}$$

6.5 Newton's Law of motion

Newton's 2nd law becomes, for an external force f^a

$$f^a = m a^a = \boxed{m \left(\frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\tau}\right)^a}. \tag{6.63}$$

If $f^a = 0$, an object follows a geodesic.

We can now describe the apparent acceleration of component a of the 4-velocity, U^a:

$$\boxed{\frac{\mathrm{d}U^a}{\mathrm{d}\tau} = \frac{1}{m}\left(f^a - \Gamma^a{}_{bc} U^b U^c\right)}. \tag{6.64}$$

The first term on the RHS arises from external forces, while the second term arises from fictitious inertial forces, including gravity.

6.6 Twin Paradox

In the year 2000, one twin sets off for a distant planet, while the other twin stays home.

Flight Plan (in ship time = proper time):

a. 5 year acceleration $1g \approx 10\ ms^{-2}$, the surface acceleration of the Earth.

b. 5 year deceleration $1g$.

c. 1 year on planet.

d. 5 year acceleration $1g$.

e. 5 year deceleration $1g$.

In relativistic units $g \approx 1.03 yr^{-1}$.

When does the twin arrive back on Earth?

We need to compare Earth time t_{E} with the proper time τ. Conveniently, as measured in the Earth frame the zeroth component of the 4-velocity U^0_{E} is

$$U^0_{\mathsf{E}} = \frac{\mathrm{d}t_{\mathsf{E}}}{\mathrm{d}\tau} = \gamma. \tag{6.65}$$

So (setting $t_{\mathsf{E}} = \tau = 0$ at the start of the journey)

$$t_{\mathsf{E}}(\tau) = \int_0^{\tau} \gamma(\tau')\ \mathrm{d}\tau'. \tag{6.66}$$

Strategy: Compare U^a and a^a in both the Earth and the spaceship frame. We will ignore the y and z components, considering only the t and x components.

Ship: in the spaceship rest frame,

$$U^a_{\mathsf{S}}(\tau) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{6.67}$$

The astronauts feel a force of one Earth gravity, which allows them to walk around the spaceship rather than float about. This force is the normal force from the floor acting on the feet of the astronauts. For an astronaut of mass m, Newton's second law states that the normal force (in the forward $\mathsf{S}^1 = x$ direction) is

$$F = mg. \tag{6.68}$$

Let us write this in covariant form: the covariant Newton's 2nd law reads

$$F^a_{\mathsf{S}} = m \left(\frac{\mathrm{D}\overline{\mathbf{U}}_{\mathsf{S}}}{\mathrm{D}\tau} \right)^a = mg. \tag{6.69}$$

Expanding the covariant derivative in terms of ordinary derivatives and Christoffel symbols,

$$F^a_{\mathsf{S}} = m \left(\frac{\mathrm{d}U^a_{\mathsf{S}}}{\mathrm{d}\tau} + \Gamma^a_{\mathsf{S}bc} U^b_{\mathsf{S}} U^c_{\mathsf{S}} \right) = \frac{F^a_{\mathsf{S}}}{m}. \tag{6.70}$$

But in the ship's rest frame

$$U^a_{\mathsf{S}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \text{constant}, \tag{6.71}$$

$$\frac{\mathrm{d}U^a_{\mathsf{S}}}{\mathrm{d}\tau} = 0. \tag{6.72}$$

Thus

$$F^a_{\mathsf{S}} = m\Gamma^a_{\mathsf{S}00}. \tag{6.73}$$

In particular,

$$\Gamma^1_{\mathsf{S}00} = g \qquad (\approx 1.03\,\text{year}^{-1}). \tag{6.74}$$

Thus $\Gamma^1_{\mathsf{S}00}$ gives the fictional (inertial) force felt by the astronauts. Also, as $F^a_{\mathsf{S}} U_{\mathsf{S}a} = 0$, we must have $F^0_{\mathsf{S}} = 0$ and $\Gamma^0_{\mathsf{S}00} = 0$.

Earth:

Assume the Earth frame is an inertial frame, $\Gamma^a_{\mathsf{E}bc} = 0$, i.e. ignore the Earth's own gravity. Then

$$\left(\frac{\mathrm{D}\overline{\mathbf{U}}_{\mathsf{E}}}{\mathrm{D}\tau}\right)^a = \frac{dU^a_E}{d\tau} + 0. \tag{6.75}$$

Now,

$$U^a_{\mathsf{E}}(\tau) = \frac{\partial \mathsf{E}^a}{\partial \mathsf{S}^b} U^b_{\mathsf{S}}(\tau). \tag{6.76}$$

The co-ordinate transformation between the Earth frame and the spaceship frame depends on the speed and hence the position of the ship. The position of the ship is parameterized by the proper time τ. For a ship travelling at speed $V(\tau)$ the Lorentz boost formula gives

$$\frac{\partial \mathsf{E}^a}{\partial \mathsf{S}^b} = \begin{pmatrix} \gamma(\tau) & \gamma V(\tau) \\ \gamma V(\tau) & \gamma(\tau) \end{pmatrix}. \tag{6.77}$$

6.6.1 The rapidity

Let *the rapidity* ϕ be defined by $\phi = \tanh^{-1} V$. Thus the quantities V, $\gamma = (1 - V^2)^{-1/2}$, and γV are given by simple hyperbolic functions:

$$V = \tanh\phi, \tag{6.78}$$

$$\gamma = \cosh\phi, \tag{6.79}$$

$$\gamma V = \sinh\phi. \tag{6.80}$$

In terms of the rapidity, the transformation matrix has the simple form

$$\frac{\partial \mathsf{E}^a}{\partial \mathsf{S}^b} = \begin{pmatrix} \cosh\phi(\tau) & \sinh\phi(\tau) \\ \sinh\phi(\tau) & \cosh\phi(\tau) \end{pmatrix}. \tag{6.81}$$

Because both the velocity 4-vector and the covariant acceleration are tensors, they can be readily transformed to Earth coordinates from the spaceship frame as in equation (6.76):

$$\overline{\mathbf{U}}_{\mathsf{E}}(\tau) = \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{6.82}$$

$$= \begin{pmatrix} \cosh\phi(\tau) \\ \sinh\phi(\tau) \end{pmatrix}; \tag{6.83}$$

$$\frac{\mathrm{D}\overline{\mathbf{U}}_{\mathsf{E}}}{\mathrm{D}\tau} = \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix} \frac{\mathrm{D}\overline{\mathbf{U}}_{\mathsf{S}}}{\mathrm{D}\tau} \tag{6.84}$$

$$= \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix} \begin{pmatrix} 0 \\ g \end{pmatrix} \tag{6.85}$$

$$= g \begin{pmatrix} \sinh\phi(\tau) \\ \cosh\phi(\tau) \end{pmatrix}. \tag{6.86}$$

Meanwhile, the ordinary derivative of U^a_{E} is

$$\frac{\mathrm{d}U^a_{\mathsf{E}}}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau} \begin{pmatrix} \cosh\phi(\tau) \\ \sinh\phi(\tau) \end{pmatrix} \tag{6.87}$$

$$= \dot{\phi} \begin{pmatrix} \sinh\phi(\tau) \\ \cosh\phi(\tau) \end{pmatrix}, \tag{6.88}$$

where $\dot{\phi} = \mathrm{d}\phi/\mathrm{d}\tau$. Thus from equation (6.75), we find $\dot{\phi} = g$, which integrates to

$$\phi(\tau) = g\tau. \tag{6.89}$$

We can now go back to equation (6.66). As $U^0_{\mathsf{E}} = \gamma = \cosh(g\tau)$, we have

$$t_E(\tau) = \int_0^\tau \cosh(g\tau)d\tau' \tag{6.90}$$

$$\therefore t_E(\tau) = \frac{1}{g}\sinh(g\tau). \tag{6.91}$$

Now $g = 1.03\text{yr}^{-1}$, and so at $\tau = 5$ years,

$$t_{\mathsf{E}} = \sinh(5.15)/1.03 = 86.6\text{yr}. \tag{6.92}$$

At this point, Earth time is 2086, Ship time is 2005. Similarly,

- the 5 year deceleration takes 86.6 years on Earth;
- 1 year on alien planet takes 1 year on Earth;
- the return journey takes 2×86.6 years on Earth.

So the twin returns to Earth 21 years older, in the year 2347.

Note the asymmetry between the stay-at-home twin and the space-faring twin. Earth people see the spaceship moving away and returning. But people on the spaceship also see the Earth moving away and returning! However, there is no true symmetry here: only the spaceship resides in a non-inertial rest frame. This results in a true difference between the flow of time on the spaceship and on the Earth.

Chapter 7

Orbits

Everyone was silent for a minute. Then Filby said he was damned.

7.1 Noether's Theorem

For any continuous symmetry of a physical system, there is a conserved quantity.

This theorem is most often expressed in the context of Hamiltonian or Lagrangian mechanics, either quantum or classical. For example, if H is the Hamiltonian for a physical system, then:

a. If $\mathrm{d}H/\mathrm{d}t = 0$ (H symmetric to time translation), *energy* is conserved.

b. If $\partial H/\partial x = 0$ (H symmetric to translation in the x direction) then the x component of linear momentum is conserved.

c. If $\partial H/\partial \phi = 0$ (H symmetric to rotation), then angular momentum is conserved

d. In electromagnetism, if H is independent of gauge, then charge is conserved.

e. In particle theory, gauge symmetry can imply conservation of other kinds of 'charge'. For example SU(3) symmetry implies conservation of 'colour' (strong force) charge.

To employ symmetry arguments in the analysis of orbits, we first prove a variant of Noether's theorem applicable to geodesics.

Theorem:

$$\boxed{\frac{dp_a}{d\tau} = \frac{m}{2}\left(\partial_a g_{bc}\right) U^b U^c}. \tag{7.1}$$

Thus, for example, if $\partial_0 g_{bc} = 0$, then $\mathrm{d}p_0/\mathrm{d}\tau = 0$. In this case the energy $E = p_0$ will be conserved.

Proof 7.1

We derive the corresponding equation for the lowered form of the 4-velocity $\underline{\mathbf{U}}$. Here $\underline{\mathbf{U}} = \underline{\mathbf{p}}/m$ can be interpreted as energy-momentum per unit mass.

First, we write $U_a = g_{ae}U^e$ and apply the product rule:

$$\frac{dU_a}{d\tau} = \frac{d}{d\tau}\left(g_{ae}U^e\right) \tag{7.2}$$

$$= g_{ae}\frac{dU^e}{d\tau} + U^e\frac{dg_{ae}}{d\tau}. \tag{7.3}$$

Apply the geodesic equation to the first term, and write $\mathrm{d}/\,\mathrm{d}\tau = U^b\partial_b$ in the second term:

$$\frac{dU_a}{d\tau} = g_{ae}\left(-\Gamma^e{}_{bc}U^bU^c\right) + U^e\left(U^b\partial_b g_{ae}\right). \tag{7.4}$$

Next, we can change the dummy variable $e \to c$ in the last term, so that we can factor out U^bU^c:

$$\frac{dU_a}{d\tau} = g_{ae}\left(-\Gamma^e{}_{bc}U^bU^c\right) + U^c\left(U^b\partial_b g_{ac}\right) \tag{7.5}$$

$$= \left(\partial_b g_{ac} - g_{ae}\Gamma^e{}_{bc}\right)U^bU^c. \tag{7.6}$$

Note that g^{ed} is the inverse metric tensor, so by equation (6.34),

$$\Gamma_{abc} \equiv g_{ae}\Gamma^e{}_{bc} = \frac{1}{2}g_{ae}g^{ed}\left(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}\right) \tag{7.7}$$

$$= \frac{1}{2}\delta_a{}^d\left(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}\right) \tag{7.8}$$

$$= \frac{1}{2}\left(\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc}\right). \tag{7.9}$$

Thus

$$\frac{dU_a}{d\tau} = \left[\partial_b g_{ac} - \frac{1}{2}\left(\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc}\right)\right]U^bU^c \tag{7.10}$$

$$= \frac{1}{2}\left(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab}\right)U^bU^c, \tag{7.11}$$

using $g_{ca} = g_{ac}$ to combine the first two terms in equation (7.10).

Finally, the last two terms in equation (7.11) involve the factor $(\partial_b g_{ac} - \partial_c g_{ab})$, which is anti-symmetric in b and c. But this factor double-contracts with U^bU^c, which is symmetric in b and c. Contraction of symmetric and anti-symmetric tensors gives 0, i.e.

$$\left(\partial_b g_{ac} - \partial_c g_{ab}\right)U^bU^c = 0 \tag{7.12}$$

so we are left with

$$\frac{dU_a}{d\tau} = \frac{1}{2}\left(\partial_a g_{bc}\right)U^bU^c. \tag{7.13}$$

QED

7.2 The Schwarzschild Metric

Consider the space surrounding a planet or star or black hole of total mass M. We will assume that the central object is spherically symmetric (so we ignore rotation) and time independent (so we ignore time evolution). One can show from the Einstein field equations that the metric line element is

$$\boxed{d\tau^2 = \left(1 - \frac{r_s}{r}\right) dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta \, d\phi^2}. \tag{7.14}$$

where $r_s \equiv 2GM$ is called the *Schwarzschild Radius.*

Equivalently the metric tensor is

$$\boxed{g_{ab} = \begin{pmatrix} \left(1 - \frac{r_s}{r}\right) & 0 & 0 & 0 \\ 0 & -\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}}. \tag{7.15}$$

The Schwarzschild radius is quite small: for the Sun, $M = M_\odot$, $r_s = 3km$. The radius of the sun, however, is $R_\odot = 7 \times 10^5 km \gg r_s$. Thus for anything orbiting the sun, even in a very low orbit, $r > R_\odot$ so the ratio $r_s/r \ll 1$.

For the Earth, $M = M_\oplus$, $r_s = 0.886cm$. Again, for anything orbiting the Earth, $r_s/r \ll 1$.

Also note that if we neglect the r_s/r terms in the metric we get back to the Minkowski metric.

7.2.1 Symmetries and Conserved Quantities

To find a planetary orbit, we solve the geodesic equation for objects moving in the curved space described by the Schwarzschild metric. This is difficult to do directly. However we can take advantage of two symmetries in the problem.

The two symmetries are time invariance and rotational invariance:

$$\partial_0 g_{bc} = \partial_t g_{bc} = 0 \quad \text{for all } b, c = 0, 1, 2, 3 \tag{7.16}$$

$$\partial_3 g_{bc} = \partial_\phi g_{bc} = 0 \quad \text{for all } b, c = 0, 1, 2, 3. \tag{7.17}$$

The corresponding conserved quantities are energy E and angular momentum L:

$$\frac{dp_0}{d\tau} = \frac{dE}{d\tau} = 0; \tag{7.18}$$

$$\frac{dp_3}{d\tau} = \frac{dL}{d\tau} = 0. \tag{7.19}$$

For massive particles we can define the *energy per unit mass* $k = E/m$ and the *angular momentum per unit mass* $h = L/m$. For constant rest mass h and k will also be constant along a geodesic. We first consider k:

$$k = U_0 = g_{0b} U^b \tag{7.20}$$

$$= \left(1 - \frac{r_s}{r}\right) U^0. \tag{7.21}$$

As $U^0 = \mathrm{d}t/\mathrm{d}\tau$, we have

$$\boxed{\frac{\mathrm{d}t}{\mathrm{d}\tau} = k\left(1 - \frac{r_s}{r}\right)^{-1}}. \tag{7.22}$$

Thus the Noether symmetry arguments lead to an expression for how co-ordinate time t varies with proper time τ.

Next consider the angular momentum per unit mass h:

$$\begin{aligned} h = -U_3 &= -g_{3b}U^b && (7.23)\\ &= r^2 \sin^2\theta\, U^3. && (7.24) \end{aligned}$$

Now $U^3 = \mathrm{d}\phi/\mathrm{d}\tau$, so

$$\boxed{\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{h}{r^2\sin^2\theta}}. \tag{7.25}$$

Note how the velocity expressed as a vector $\overline{\mathbf{U}}$ with upper indices U^a has a different physical meaning from the form $\underline{\mathbf{U}}$ with lower indices U_a. The vector $\overline{\mathbf{U}}$ shows us where the object is going (as it represents the tangent to the world line). The form $\underline{\mathbf{U}}$, on the other hand, tells us how much energy and momentum (per unit mass) the object carries.

7.2.2 Orbits in the Equatorial Plane

Consider geodesics in the equatorial plane $\theta = \pi/2$. For the solar system this plane is called the *ecliptic*. (The constellations seen on the ecliptic are known as the zodiac.) For motion on this plane $\mathrm{d}\theta = 0$ and $\sin\theta = 1$, so the Schwarzschild metric line element simplifies to

$$\mathrm{d}\tau^2 = \left(1 - \frac{r_s}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\phi^2. \tag{7.26}$$

Let us find orbit equations in terms of h and k:

a. Divide the metric line element by $d\tau^2$ and use equations (7.22) and (7.25) (with $\sin\theta = 1$):

$$\begin{aligned} 1 &= k^2\left(1 - \frac{r_s}{r}\right)^{-1} - \left(1 - \frac{r_s}{r}\right)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - \frac{h^2}{r^2} && (7.27)\\ \implies \frac{\mathrm{d}r}{\mathrm{d}\tau} &= \sqrt{k^2 - \left(1 + \frac{h^2}{r^2}\right)\left(1 - \frac{r_s}{r}\right)}. && (7.28) \end{aligned}$$

This gives us a differential equation for $\mathrm{d}r/\mathrm{d}\tau$. Unfortunately, it is quite non-linear and difficult to solve in this form.

Exercise 7.1

(a) Starting with equation (7.28), derive an expression for $dr/d\tau$ in the form

$$\frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + V(r) = C$$

where C is a constant, and the effective potential is

$$V(r) = -\frac{1}{2}\left(\frac{r_s}{r} - \frac{h^2}{r^2} + \frac{r_s h^2}{r^3}\right).$$

What is the effective energy C?

(b) Find the radii r_1 and r_2, $r_1 < r_2$ where the effective potential has an extremum (maximum or minimum). Show that if $C = V(r_1)$ or $C = V(r_2)$ then the requirements for a circular orbit ($dr/d\tau = d^2r/d\tau^2 = 0$) are satisfied. Show that $h \geq \sqrt{3}\, r_s$ for these orbits. Also show that $r_2 \geq 3r_s$.

(c) Let $h = 2r_s$. What are r_1 and r_2? Show that for the outer orbit at r_2, $V''(r_2) > 0$ and hence that this orbit is stable. Is the inner orbit at r_1 stable?

b. To simplify the equation, we change the independent variable from $\tau \to \phi$, and find $r(\phi)$:

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right) = \left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)\left(\frac{h}{r^2}\right) \tag{7.29}$$

$$\Longrightarrow \left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 \frac{h^2}{r^4} = k^2 - \left(1 + \frac{h^2}{r^2}\right)\left(1 - \frac{r_s}{r}\right). \tag{7.30}$$

c. Next we change the dependent variable $r \to u = 1/r$. We will denote differentiation by ϕ with a prime, e.g. $u' = \mathrm{d}u/\,\mathrm{d}\phi$. Thus

$$r' = \frac{\mathrm{d}r}{\mathrm{d}\phi} = \frac{\mathrm{d}r}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}\phi} \tag{7.31}$$

$$= -\frac{1}{u^2}u' \tag{7.32}$$

$$\Longrightarrow \left(-\frac{u'}{u^2}\right)^2 u^4 h^2 = k^2 - \left(1 + h^2u^2\right)\left(1 - r_s u\right) \tag{7.33}$$

$$\Longrightarrow u'^2 = \frac{k^2}{h^2} - \frac{\left(1 + h^2u^2\right)\left(1 - r_s u\right)}{h^2} \tag{7.34}$$

We now have an equation for u' which at least has no terms in the denominator:

$$\boxed{u'^2 = \left(\frac{k^2 - 1}{h^2}\right) + \frac{r_s u}{h^2} - u^2 + r_s u^3}. \qquad \text{(Einstein)} \tag{7.35}$$

The corresponding Newtonian orbit equation leaves out the last term:

$$\boxed{u'^2 = k_N + \frac{r_s u}{h^2} - u^2}, \qquad \text{(Newton)} \tag{7.36}$$

where the Newtonian energy per unit mass is

$$k_N = \frac{V^2}{2} - \frac{GM}{r} = \frac{V^2}{2} - \frac{r_s}{2r}. \tag{7.37}$$

d. We can simplify further by differentiating with respect to ϕ:

$$2u'u'' = \frac{r_s}{h^2}u' - 2uu' + 3r_s u^2 u' \tag{7.38}$$

$$\implies u'' + u = \frac{r_s}{2h^2} + \frac{3r_s}{2}u^2. \tag{7.39}$$

Thus General Relativity predicts that orbits satisfy

$$\boxed{u'' + u = \frac{r_s}{2h^2} + \frac{3r_s}{2}u^2}. \qquad \text{(Einstein)} \tag{7.40}$$

In contrast, the Newtonian orbit equation is

$$\boxed{u'' + u = \frac{r_s}{2h^2}}. \qquad \text{(Newton)} \tag{7.41}$$

Compare the relativistic correction term (the last term in the Einstein version) to the linear term u:

$$\frac{3r_s u^2/2}{u} = \frac{3}{2}\frac{r_s}{r} = 3\frac{GM}{r}. \tag{7.42}$$

For planets orbiting the sun at a radius $r > 100R_\odot$,

$$3\frac{GM_\odot}{100R_\odot} \sim 10^{-7} \tag{7.43}$$

and so the relativistic correction term results in very small deviations from the Newtonian predictions. These deviations have, however, been observed!

Exercise 7.2

a. *Consider a sphere of radius 1 as a 2-dimensional manifold with coordinates $x_1 = \theta$, $x_2 = \phi$, and line-element*

$$ds^2 = \left(d\theta^2 + \sin^2\theta d\phi^2\right).$$

Show that geodesics have a conserved quantity (call it H).

b. *Using the metric line element, or otherwise, derive an equation for $d\theta/ds$ in terms of H.*

c. *Let $X = \cos\theta$. Show that X satisfies*

$$\left(\frac{dX}{ds}\right)^2 = 1 - X^2 - H^2.$$

d. *Obtain a second-order differential equation for $X(s)$ and write down its general solution. Suppose a geodesic starts at co-latitude θ_0 heading due East. Find $X(s)$ and hence $\theta(s)$ explicitly in terms of θ_0. Show that the total length of the geodesic (i.e. the length needed to go all the way around the sphere once) is independent of θ_0.*

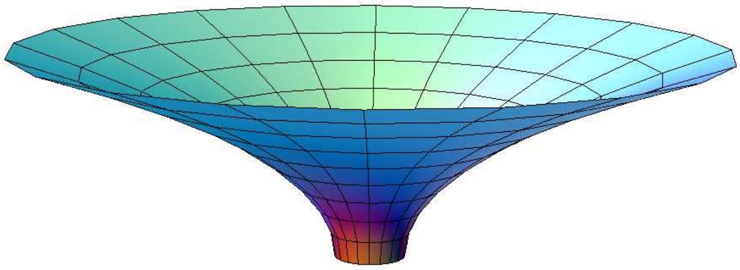


Figure 7.1: A visualization of the Schwarzschild Metric. More precisely, a 2-manifold imbedded in three dimensional space which has the same spatial metric as a constant time equatorial ($t = \text{constant}, \theta = \pi/2$) slice of the Schwarzschild metric (equation (7.26)).

Exercise 7.3

a. *Consider a surface embedded in 3 dimensional Euclidean space (e.g. the surface of a bowl). Using cylindrical coordinates (r, ϕ, z), the surface is specified by the function $z = Z(r)$. Let the two coordinates on the surface be $x^1 = r$ and $x^2 = \phi$. Show that the metric of the surface is given by*

$$g_{ab} = \begin{pmatrix} 1 + Z'^2 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{7.44}$$

b. *Next consider geodesics on the surface. Since the surface is purely spatial, we replace τ by arclength s in the geodesic equations. Show that the metric has a symmetry, and hence there exists a conserved quantity (call it h) along each geodesic. In other words, find a quantity h such that $dh/ds = 0$.*

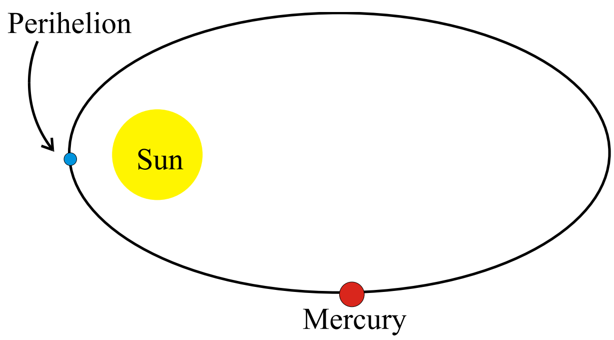
c. *Let $u = 1/r$ and let $u' = du/d\phi$. Derive the equation*

$$u'^2 = \left(\frac{1}{h^2} - u^2\right)\frac{1}{1 + Z'^2}.$$

d. *Now suppose $Z(r) = 2\sqrt{r-1}$. Show that this gives a surface whose metric is the same as the* spatial part *of the Schwarzschild metric, equation (7.26), (for $\theta = \pi/2$) in units where $r_s = 1$ (see figure 7.1). What is the equation for u'^2? How does this compare with the orbit equation derived from the full Schwarzschild metric?*

7.3 Precession of Mercury's Orbit

Perihelion = Closest approach to the Sun.



Observations show that the perihelion of Mercury precesses about 1000 arcsec / century. Influences of the other planets account for all but 43" of this. Einstein found the 43" could be explained by the extra term in the orbit equation. A numerical approach might be to go back to the first order equation, equation (7.35), and integrate directly:

$$\int \frac{\mathrm{d}u}{\sqrt{(k^2-1)/h^2 + r_s u/h^2 - u^2 + r_s u^3}} = \int \mathrm{d}\phi. \tag{7.45}$$

However, for planetary orbits we can exploit the fact that the relativistic correction to Newtonian theory is very small. This will enable us to find a simple analytic solution (approximate, but then so are numerical solutions!).

7.3.1 Method

The angle ϕ measures the net angle through which the planet has orbited the sun. During the first orbit $0 \le \phi \le 2\pi$. During the second orbit, $2\pi \le \phi \le 4\pi$ and so on. But each successive orbit does not exactly follow the previous orbit, so for example $u(2\pi) \neq u(0)$. Thus the function $u(\phi)$ cannot only contain terms periodic in ϕ like $\sin\phi$. There may be a linear term as well. We will solve for the function $u(\phi)$, then invert to obtain $\phi(u)$.

Now at perihelion $r = r_{\min}$, so $u = u_{\max}$. We will compare successive values of $\phi(u_{\max})$. Because the orbits are not exactly alike,

$$\underbrace{\phi(u_{\max})}_{\text{orbit 2}} - \underbrace{\phi(u_{\max})}_{\text{orbit 1}} = 2\pi + \delta\phi \tag{7.46}$$

for some angle $\delta\phi$. We will call $\delta\phi$ the precession.

But first we need to solve for $u(\phi)$. We will do this by writing this function as the sum of the Newtonian solution plus a small correction term (we derive the Newtonian solution below):

$$u(\phi) = u_0\left(\underbrace{(1+\epsilon\sin\phi)}_{\text{Newtonian Orbit}} + \underbrace{y(\phi)}_{\text{Correction}}\right); \tag{7.47}$$

$$u_0 \equiv \frac{r_s}{2h^2}. \tag{7.48}$$

We plug this into the orbit equation to obtain a new differential equation for $y(\phi)$. As $y(\phi)$ is small, we drop the non-linear terms (those in y^2 etc) to obtain a linear equation in y, which can then be readily solved.

7.3.2 Newtonian Solution

First, look at the solution to the Newtonian equations 7.36 and 7.41. The second order equation has sines and cosines as complementary functions and a constant as a particular solution, so the general solution can be written

$$u(\phi) = A\cos\phi + B\sin\phi + u_0. \tag{7.49}$$

Comment: Here there are two unknown constants of integration, as expected for a second order differential equation. These could be determined, for example, by setting initial conditions $u(\phi_0)$ and $u'(\phi_0)$ at some angle ϕ_0. However, we started with a first order equation, equation (7.36), which should only have one constant of integration. What has changed? When we differentiated to obtain the second order equation, we lost some information about the orbit. In particular, we lost the term $(k^2-1)/h^2$. This is the only term which tells us about the energy k. A solution of the first order equation may be characterized by just one of the initial conditions as well as k and h. The second order equation loses the explicit k dependence. Of course, the energy will be derivable as a function of the second initial condition, and vice-versa.

The perihelion occurs at minimum r, hence maximum u. Let us suppose this occurs at $\phi_0 = \pi/2$. Then one initial condition gives

$$u'(\frac{\pi}{2}) = 0 \Rightarrow A = 0. \tag{7.50}$$

Now plug into the first order Newtonian equation (7.36) (without the relativistic correction term $r_s u^3$). The result gives

$$B = \left(k_N + \frac{r_s^2}{4h^4}\right)^{1/2}. \tag{7.51}$$

Finally, define the *eccentricity* ϵ by

$$\epsilon = \frac{B}{u_0} \tag{7.52}$$

to yield

$$u = u_0\left(1 + \epsilon\sin\phi\right). \tag{7.53}$$

The eccentricity tells us the shape of the orbit. Thus $\epsilon = 0$ gives a circle, $0 < \epsilon < 1$ gives an ellipse, $\epsilon = 1$ gives a parabola, and $\epsilon > 1$ gives a hyperbola (the latter two are open orbits where an object comes in from ∞, is deflected, and escapes to ∞). The eccentricity of Mercury's orbit is about $\epsilon = 0.21$.

Note that for a circular orbit

$$u_{circ} = u_0 = \frac{r_s}{2h_{circ}^2} \tag{7.54}$$

$$\Rightarrow h_{circ} = \left(\frac{r_{circ}}{2r_s}\right)^{1/2} r_s. \tag{7.55}$$

Thus for planetary orbits $h_{circ} \gg r_s$.

7.3.3 Relativistic Correction

We now substitute equation (7.47) into the full relativistic orbit equation (7.40). This yields the following differential equation for $y(\phi)$:

$$y'' + y = \frac{3r_s u_0}{2} \left((1+\epsilon \sin\phi) + y\right)^2 \quad (7.56)$$

$$= \frac{3r_s u_0}{2} \left((1+\epsilon \sin\phi)^2 + 2(1+\epsilon\sin\phi)\, y + y^2\right). \quad (7.57)$$

We can ignore the y^2 term, as $y \ll 1$. Next, compare the terms linear in y. There are two terms: on the left hand side, with coefficient 1, and on the right, with coefficient $3r_s u_0 (1+\epsilon\sin\phi)$. The latter term is of order $r_s/r \ll 1$. Neglecting this term gives

$$y'' + y \approx \frac{3r_s u_0}{2} \left(1 + 2\epsilon\sin\phi + \epsilon^2 \sin^2\phi\right). \quad (7.58)$$

The terms on the right act as forcing functions for the harmonic oscillator on the left. The most interesting forcing function is the $2\epsilon\sin\phi$ term, as this is in resonance with the oscillator (the complementary functions include $\sin\phi$). This resonance drives the precession.

For initial conditions $y(\pi/2) = 0$, $y'(\pi/2) = 0$, the solution is

$$y(\phi) = \frac{3r_s u_0}{2} \left[(1-\sin\phi) + \epsilon\left(\frac{\pi}{2} - \phi\right)\cos\phi + \frac{\epsilon^2}{3}\left(2 - \sin^2\phi - \sin\phi\right)\right] \quad (7.59)$$

$$= \frac{3r_s u_0}{2}\, \epsilon \left(\frac{\pi}{2} - \phi\right)\cos\phi + \text{periodic terms}. \quad (7.60)$$

Recall the discussion leading to equation (7.46). The first orbit has perihelion at $\phi_1 = \pi/2$, while the second orbit has perihelion at $\phi_2 = 5\pi/2 + \delta\phi$. Solving $u'(\phi_2) = 0$ gives

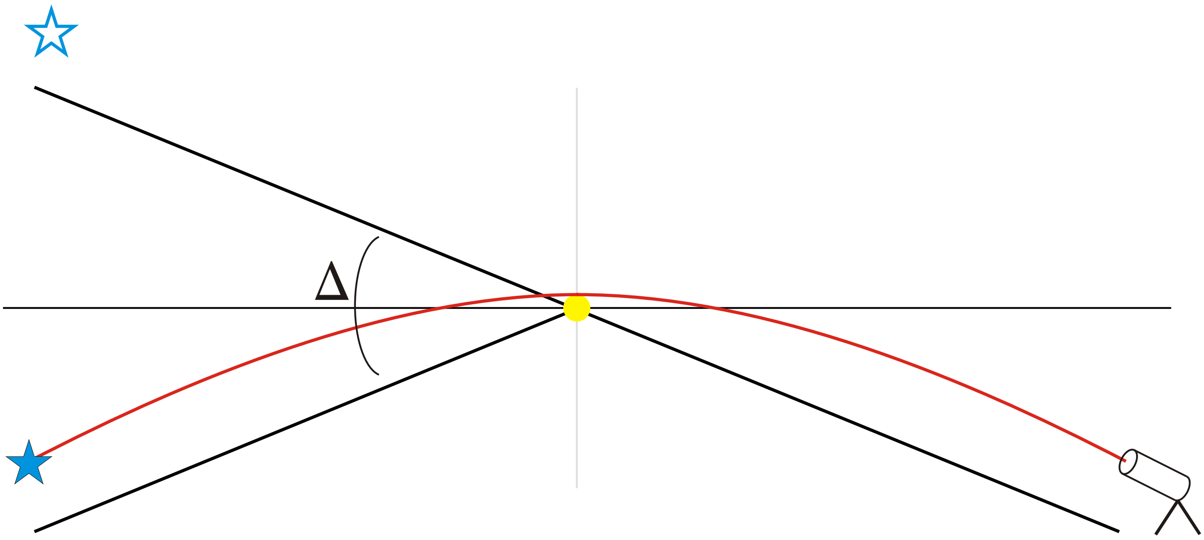
$$\delta\phi \approx 3\pi r_s u_0. \quad (7.61)$$

The mean radius is $\bar{r} \approx u_0^{-1} = 58 \times 10^6$ km. Thus

$$\delta\phi \approx 0.1\, \text{arcseconds/orbit}. \quad (7.62)$$

The Mercury year is 88 earth days, making the precession 1 arcsecond every 880 days, or 43 arcseconds per century.

7.4 Deflection of Starlight



As light travels from a star, its path is distorted by the curvature of space-time. Thus photons passing near a massive object like the sun will curve around the object. When the light is viewed, it will appear in the wrong place on the sky. This effect has been seen not only for starlight in the sun's gravitational field, but also for light from distant galaxies or quasars passing by nearer galaxies on the way to Earth.

The effect on starlight is best seen during a solar eclipse. During an eclipse we can see stars in the sky close to the sun without being swamped by daylight. Photons from these stars receive the maximum deflection.

Figure 7.4 shows the path of the starlight. The photon starts at the star, with

$$u_\star \approx 0; \qquad \phi_\star = \pi + \frac{\Delta}{2}. \tag{7.63}$$

Because of the gravitational deflection, we observe the photon coming in from the angle $\phi = \pi - \Delta/2$. Thus the angle of deflection is Δ.

First we look at the Newtonian prediction, then derive the relativistic deflection.

7.4.1 Newtonian Theory

The Newtonian orbit for $m \neq 0$ is, from equation (7.47)

$$u(\phi) = u_0\,(1 + \epsilon \sin\phi)\,; \qquad u_0 = \frac{r_s}{2h^2} = \frac{GM_\odot}{h^2}. \tag{7.64}$$

The angular momentum per unit mass h is

$$h = r^2\dot{\phi} = rV_\phi. \tag{7.65}$$

In figure 7.4, the photon travelling from the star to our telescope has its closest approach (perihelion) at $\phi = \pi/2$. At perihelion $dr/dt = 0$, so $V_r = 0$ and thus $|V| = V_\phi$. But for a photon $|V| = c = 1$ in relativistic units. So

$$h = r_{\min} \approx R_\odot \tag{7.66}$$

if the starlight just grazes the surface of the sun on its way to our telescope. Thus

$$u(\phi) = \frac{GM_\odot}{R_\odot^2}(1 + \epsilon \sin\phi)\,. \tag{7.67}$$

At the star, equation (7.63) gives

$$0 \approx \frac{GM_\odot}{R_\odot^2}\left(1 + \epsilon \sin\left(\pi + \frac{\Delta}{2}\right)\right) \tag{7.68}$$

$$\approx \frac{GM_\odot}{R_\odot^2}\left(1 - \epsilon\left(\frac{\Delta}{2}\right)\right), \tag{7.69}$$

as $\sin(\pi + x) \approx -x$ for small x. Thus $\epsilon \approx 2/\Delta$. We now have

$$u(\phi) \approx \frac{GM_\odot}{R_\odot^2}\left(1 + \frac{2}{\Delta}\sin\phi\right). \tag{7.70}$$

Finally, at $\phi = \pi/2$ the starlight reaches $r = R_\odot$, so

$$\frac{1}{R_\odot} \approx \frac{GM_\odot}{R_\odot^2}\left(1 + \frac{2}{\Delta}\right), \tag{7.71}$$

$$\approx \frac{GM_\odot}{R_\odot^2}\left(\frac{2}{\Delta}\right), \tag{7.72}$$

using $2/\Delta \gg 1$. Thus

$$\boxed{\Delta \approx \frac{2GM_\odot}{R_\odot}}\,. \tag{7.73}$$

7.4.2 Relativistic Theory

As photons are massless, we consider the orbit equations in the limit $m \to 0$. Photons do have energy and momentum (and angular momentum), so we hold energy $E = mk$ and angular momentum $L = mh$ constant. The first order equation (equation (7.35)) becomes

$$u'^2 = \left(\frac{E^2 - m}{L^2}\right) - \frac{r_s u m^2}{L^2} - u^2 + r_s u^3. \tag{7.74}$$

In the limit $m \to 0$, then,

$$\boxed{u'^2 = \frac{E^2}{L^2} - u^2 + r_s u^3}\,. \tag{7.75}$$

Differentiation gives the second order equation

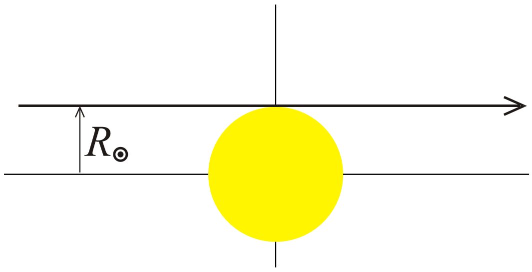
$$\boxed{u'' + u = \frac{3}{2} r_s u^2}\,. \tag{7.76}$$

First, consider the $r_s = 0$ case. Here $M = 0$ and space-time is flat:

$$u'' + u = 0 \tag{7.77}$$
$$\implies u = A\cos\phi + B\sin\phi \tag{7.78}$$
$$\tag{7.79}$$

Apply boundary conditions that the light reaches $r = \infty$ at $\phi = 0$, and (again) that the perihelion of the light path occurs at $(r, \phi) = (R_\odot, \pi/2)$. As a result, $A = 0$ and $B = R_\odot^{-1}$, i.e.

$$u = \frac{1}{R_\odot}\sin\phi. \tag{7.80}$$



This is the equation of a straight horizontal line ($y = R_\odot$) in polar coordinates. Light passing near a massive object, however, will be perturbed by the missing r_s term. As in the analysis of the precession of Mercury we try a small non-linear correction to u:

$$u(\phi) = \frac{1}{R_\odot}(\sin\phi + y(\phi)) \tag{7.81}$$

where $y(\phi) << 1$. Then equation (7.76) gives

$$\frac{1}{R_\odot}(-\sin\phi + y'') + \frac{1}{R_\odot}(\sin\phi + y) = \frac{3r_s}{2R_\odot^2}(\sin\phi + y)^2 \tag{7.82}$$
$$\implies y'' + y = \frac{3r_s}{2R_\odot}(\sin\phi + y)^2\,. \tag{7.83}$$

Include only the terms of lowest order in y and r_s (as y, r_s are both small):

$$y'' + y \approx \frac{3r_s}{2R_\odot}\sin^2\phi = \frac{3r_s}{4R_\odot}(1 - \cos 2\phi). \tag{7.84}$$

Combining complementary functions and particular integrals gives

$$y = A\cos\phi + B\sin\phi + C + D\cos 2\phi + E\sin 2\phi. \tag{7.85}$$

We can determine the constants as follows:

a. Plugging equation (7.85) into equation (7.84) gives $C = 3D$, $D = r_s/4R_\odot$ and $E = 0$.

b. At perihelion, $y(\pi/2) = 0$, so $B = -2D$.

c. Also, as seen in the figure, the photon path is symmetric about $\phi = \pi/2$. Hence, for example, $y(0) = y(\pi)$, which implies $A = 0$.

We now have

$$y = \frac{r_s}{4R_\odot}(3 - 2\sin\phi + \cos 2\phi). \tag{7.86}$$

Finally, we can apply the boundary conditions at the star, equation (7.4). These give

$$\begin{aligned}
0 &= R_\odot\, u\,(\pi + \Delta/2) && (7.87)\\
&= \sin(\pi + \Delta/2) + \frac{r_s}{4R_\odot}\,(3 - 2\sin(\pi + \Delta/2) + \cos(2\pi + \Delta)) && (7.88)\\
&= -\sin(\Delta/2) + \frac{r_s}{4R_\odot}(3 + 2\sin(\Delta/2) + \cos 2\Delta) && (7.89)\\
&\approx -\frac{\Delta}{2} + \frac{r_s}{4R_\odot}(4 + \mathrm{Order}(\Delta)). && (7.90)
\end{aligned}$$

We can ignore the terms of order $(r_s/R_\odot)\Delta$, giving

$$\boxed{\Delta \approx \frac{2r_s}{R_\odot} = \frac{4GM_\odot}{R_\odot}} \approx 1.74\,\text{arcseconds}. \tag{7.91}$$

Thus relativity predicts a deflection almost exactly double the Newtonian result.

7.5 Energy Conservation on Geodesics

According to Newtonian theory, an object moving in a central gravitational field has energy

$$E = \frac{1}{2}mV^2 + m\Phi(r); \qquad \Phi(r) = -\frac{GM}{r}, \tag{7.92}$$

where $\Phi(r)$ is the potential energy. We now show that for weak gravitational fields ($r \gg r_s$) and small velocities ($V \ll 1$), the relativistic energy is approximately the same as the Newtonian energy, apart from the contribution from rest mass.

Recall the Schwarzschild metric line element in the equatorial plane, equation (7.26). In terms of Φ,

$$\mathrm{d}\tau^2 = (1 + 2\Phi)\ \mathrm{d}t^2 - (1 + 2\Phi)^{-1}\ \mathrm{d}r^2 - r^2\ \mathrm{d}\phi^2. \tag{7.93}$$

Next, divide by $\mathrm{d}t^2$, using equation (7.21) for $\mathrm{d}\tau/\,\mathrm{d}t$:

$$\begin{aligned}
\left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^2 &= \frac{1}{k^2}\,(1 + 2\Phi)^2 && (7.94)\\
&= (1 + 2\Phi) - (1 + 2\Phi)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 - r^2\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2 && (7.95)\\
&= (1 + 2\Phi) - (1 + 2\Phi)^{-1}\,V_r^2 - V_\phi^2. && (7.96)
\end{aligned}$$

Now $\Phi \ll 1$ and $V^2 \ll 1$ so we can expand the right hand side, ignoring terms like Φ^2 or ΦV^2:

$$\begin{aligned}
\frac{1}{k^2}\,(1 + 2\Phi)^2 &\approx (1 + 2\Phi) - (1 - 2\Phi)\,V_r^2 - V_\phi^2 && (7.97)\\
&\approx (1 + 2\Phi) - (V_r^2 + V_\phi^2) && (7.98)\\
&= \left(1 + 2\Phi - V^2\right). && (7.99)
\end{aligned}$$

Thus

$$k^2 \approx \frac{(1+2\Phi)^2}{1+2\Phi-V^2} \tag{7.100}$$

$$\approx (1+4\Phi)(1-2\Phi+V^2) \tag{7.101}$$

$$\approx 1+V^2+2\Phi. \tag{7.102}$$

Finally, we take the square root of this approximation, using $(1+x)^{1/2} \approx 1+x/2$. With $E = mk$ the Newtonian correspondence becomes clear:

$$E \approx m + \frac{1}{2}mV^2 + m\Phi. \tag{7.103}$$