Instructor Solutions Manual: *Modern General Relativity*

Mike Guidry

This document gives the solutions for all problems at the ends of chapters for the first edition of *Modern General Relativity: Black Holes, Gravitational Waves, and Cosmology* by Mike Guidry (Cambridge University Press, 2019). Unless otherwise indicated, literature references, equation numbers, figure references, table references, and section numbers refer to the print version of that book.

1 Introduction

1.1 From Eq. (1.2), the value of γ is infinite if $v = c$, so there is no Lorentz transformation to an inertial frame corresponding to a rest frame for light.

1.2 Since $E = m\gamma$, for a 7 TeV proton,

$$\gamma = \frac{E}{m} = \frac{7 \times 10^{12}\,\text{eV}}{938.3 \times 10^{6}\,\text{eV}} = 7460.$$

Then from the definition of γ,

$$\frac{v}{c} = \sqrt{1 - \frac{1}{\gamma^2}} = 0.999999991.$$

This is a speed that is only about 3 meters per second less than that of light.

1.3 This question is ambiguous, since it does not specify whether the curvature is that of the surface itself (which is called *intrinsic curvature*) or whether it is the apparent curvature of the surface seen embedded in a higher-dimensional euclidean space (which is called the *extrinsic curvature*). In general relativity the curvature of interest is usually intrinsic curvature. Then the sheet of paper can be laid out flat and is not curved, the cylinder is *also flat,* with no intrinsic curvature, because one can imagine cutting it longitudinally and rolling it out into a flat surface, but the sphere has finite intrinsic curvature because it cannot be cut and rolled out flat without distortion. The reason that the cylinder seems to be curved is because the 2D surface is being viewed embedded in 3D space, which gives a non-zero *extrinsic curvature,* but if attention is confined only to the 2D surface it has no *intrinsic curvature*. This is a rather qualitative discussion but in later chapters methods will be developed to quantify the amount of intrinsic curvature for a surface.

2 Coordinate Systems and Transformations

2.1 Utilizing Eq. (2.31) to integrate around the circumference of the circle,

$$C = \oint ds = \oint (dx^2 + dy^2)^{1/2} = 2\int_{-R}^{+R} dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2},$$

subject to the constraint $R^2 = x^2 + y^2$, where the factor of two and the limits are because x ranges from $-R$ to $+R$ over half a circle. The constraints yield $dy/dx = -(R^2 - x^2)^{-1/2}x$, which permits the integral to be written as

$$C = 2\int_{-R}^{R} dx \sqrt{\frac{R^2}{R^2 - x^2}}.$$

Introducing a new integration variable a through $a \equiv x/R$ then gives

$$C = 2R\int_{-1}^{+1} \frac{da}{\sqrt{1-a^2}} = 2\pi R,$$

since the integral is $\sin^{-1} a$. In plane polar coordinates the line element is given by Eq. (2.32) and proceeding as above the circumference is

$$\begin{aligned} C &= \oint ds = \oint (dr^2 + r^2 d\varphi^2)^{1/2} \\ &= \int_0^{2\pi} d\varphi \sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2} = R\int_0^{2\pi} d\varphi = 2\pi R, \end{aligned}$$

where $r = R$ has been used, implying that $dr/d\varphi = 0$.

2.2 Under a Galilean transformation $\boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{v}t$ and $t' = t$ it is clear that the acceleration $\boldsymbol{a}$ and the separation vector $\boldsymbol{r} = \Delta\boldsymbol{x}$ between two masses are unchanged. Thus the second law $\boldsymbol{F} = m\boldsymbol{a}$ and the gravitational law $\boldsymbol{F} = Gm_1m_2\hat{\mathbf{r}}/r^2$ are invariant under Galilean transformations.

2.3 Our solution follows Example 1.2.1 of Foster and Nightingale [88]. The tangent and dual basis vectors, and the products for $g_{ij} = g_{ji} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j$, were worked out in Example 2.3. The elements for $g^{ij} = g^{ji} = \boldsymbol{e}^i \cdot \boldsymbol{e}^j$ can be obtained in a similar fashion. For example,

$$g^{12} = g^{21} = (\tfrac{1}{2}\boldsymbol{i} + \tfrac{1}{2}\boldsymbol{j}) \cdot (\tfrac{1}{2}\boldsymbol{i} - \tfrac{1}{2}\boldsymbol{j}) = \tfrac{1}{4} - \tfrac{1}{4} = 0,$$

where the orthonormality of the cartesian basis vectors has been used. Summarizing the results,

$$g_{ij} = \begin{pmatrix} 4v^2+2 & 4uv & 2v \\ 4uv & 4u^2+2 & 2u \\ 2v & 2u & 1 \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} \frac{1}{2} & 0 & -v \\ 0 & \frac{1}{2} & -u \\ -v & -u & 2u^2+2v^2+1 \end{pmatrix}$$

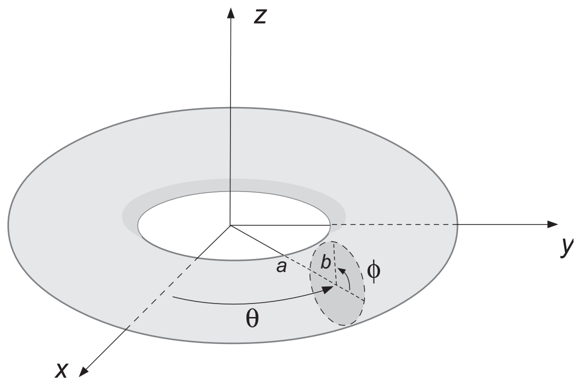


Fig. 2.1 Figure for Problem 2.5.

By direct multiplication the product of these two matrices is the unit matrix, verifying Eq. (2.26) explicitly for this case. Utilizing Eq. (2.29), the line element is

$$\begin{aligned} ds^2 &= g_{ij} du^i du^j \\ &= g_{uu} du^2 + 2g_{uv} du dv + 2g_{uw} du dw + g_{vv} dv^2 + 2g_{vw} dv dw + g_{ww} dw^2 \\ &= (4v^2+2) du^2 + 8uv du dv + 4v du dw + (4u^2+2) dv^2 + 4u dv dw + dw^2 \end{aligned}$$

where $g_{ij} = g_{ji}$ has been used and no summation is implied by repeated indices.

2.4 Using the spherical coordinates

$$u^1 = r \qquad u^2 = \theta \qquad u^3 = \varphi$$

defined through Eq. (2.2) and the results of Example 2.2,

$$\boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = 1 \qquad \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = r^2 \qquad \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 = r^2 \sin^2\theta,$$

while all non-diagonal components vanish. Thus the metric tensor is

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}.$$

The corresponding line element is

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2,$$

where Eq. (2.29) has been used.

2.5 This solution is based on Problem 1.2 in Ref. [88]. From the parameterization $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ with

$$x = (a + b\cos\varphi)\cos\theta \qquad y = (a + b\cos\varphi)\sin\theta \qquad z = b\sin\varphi,$$

where the radius of the doughnut a and radius of the circle b are defined in Fig. 2.1 [this document], the tangent basis vectors are

$$\boldsymbol{e}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = -\sin\theta(a+b\cos\varphi)\,\boldsymbol{i} + (a+b\cos\varphi)\cos\theta\,\boldsymbol{j}$$
$$\boldsymbol{e}_\varphi = \frac{\partial \boldsymbol{r}}{\partial \varphi} = -(b\sin\varphi\cos\theta)\,\boldsymbol{i} - (b\sin\varphi\sin\theta)\,\boldsymbol{j} + (b\cos\varphi)\,\boldsymbol{k}.$$

The corresponding elements of the metric tensor $g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j$ are

$$g_{\varphi\varphi} = b^2 \qquad g_{\varphi\theta} = g_{\theta\varphi} = 0 \qquad g_{\theta\theta} = (a+b\cos\varphi)^2.$$

2.6 The tangent basis vectors and metric tensor g_{ij} were given in Example 2.4. Since g^{ij} is the matrix inverse of g_{ij}, which is diagonal,

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \qquad \longrightarrow \qquad g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}$$

Then the dual basis may be obtained by raising indices with the metric tensor: $\boldsymbol{e}^i = g^{ij}\boldsymbol{e}_j$, giving

$$\boldsymbol{e}^1 = g^{11}\boldsymbol{e}_1 + g^{12}\boldsymbol{e}_2 = \boldsymbol{e}_1 \qquad \boldsymbol{e}^2 = g^{21}\boldsymbol{e}_1 + g^{22}\boldsymbol{e}_2 = \frac{1}{r^2}\boldsymbol{e}_2$$

for the elements of the dual basis.

2.7 For a constant displacement d in the x direction

$$x' = x - d \qquad y' = y \qquad z' = z.$$

Since d is constant

$$dx' = dx \qquad dy' = dy \qquad dz' = dz$$

and therefore $ds'^2 = ds^2$. From Eq. (2.41), a rotation in the $x-y$ plane may be written

$$x' = x\cos\theta + y\sin\theta \qquad y' = -x\sin\theta + y\cos\theta \qquad z' = z,$$

which gives the transformed line element

$$\begin{aligned} ds'^2 &= (dx')^2 + (dy')^2 + (dz')^2 \\ &= (\cos\theta dx + \sin\theta dy)^2 + (-\sin\theta dx + \cos\theta dy)^2 + dz^2 \\ &= (\cos^2\theta + \sin^2\theta)dx^2 + (\cos^2\theta + \sin^2\theta)dy^2 + dz^2 \\ &= dx^2 + dy^2 + dz^2 \\ &= ds^2. \end{aligned}$$

Therefore the euclidean spatial line element is invariant under displacements by a constant amount and under rotations.

2.8 Taking the scalar products using Eqs. (2.8), (2.9), and (2.20) gives

$$\boldsymbol{e}^i \cdot \boldsymbol{V} = \boldsymbol{e}^i \cdot (V^j \boldsymbol{e}_j) = V^j \boldsymbol{e}^i \cdot \boldsymbol{e}_j = V^j \delta^i_j = V^i,$$
$$\boldsymbol{e}_i \cdot \boldsymbol{V} = \boldsymbol{e}_i \cdot (V_j \boldsymbol{e}^j) = V_j \boldsymbol{e}_i \cdot \boldsymbol{e}^j = V_j \delta_i^j = V_i,$$

which is Eq. (2.22).

2.9 Utilizing that the angle θ between the basis vectors is determined by $\cos\theta = \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 / |\boldsymbol{e}_1||\boldsymbol{e}_2|$, the area of the parallelogram is

$$
\begin{aligned}
dA &= |\boldsymbol{e}_1||\boldsymbol{e}_2| \sin\theta \, dx^1 dx^2 \\
&= |\boldsymbol{e}_1||\boldsymbol{e}_2| (1 - \cos^2\theta)^{1/2} dx^1 dx^2 \\
&= \left(|\boldsymbol{e}_1|^2 |\boldsymbol{e}_2|^2 - (\boldsymbol{e}_1 \cdot \boldsymbol{e}_2)^2 \right)^{1/2} dx^1 dx^2 .
\end{aligned}
$$

The components of the metric tensor g_{ij} are

$$
\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = g_{12} = g_{21} \qquad |\boldsymbol{e}_1||\boldsymbol{e}_1| = \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = g_{11} \qquad |\boldsymbol{e}_2||\boldsymbol{e}_2| = \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = g_{22},
$$

so the area of the parallelogram may be expressed as

$$
dA = \left(g_{11} g_{22} - g_{12}^2 \right)^{1/2} dx^1 dx^2 = \sqrt{\det g} \, dx^1 dx^2 ,
$$

where $\det g$ is the determinant of the metric tensor. This is the 2D version of the invariant 4D volume element given in Eq. (3.48).

3 Tensors and Covariance

3.1 For the three cases

$$T'^{\mu\nu} = V'^{\mu}V'^{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}V^{\alpha}V^{\beta} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}T^{\alpha\beta}$$

$$T'_{\mu\nu} = V'_{\mu}V'_{\nu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}V_{\alpha}V_{\beta} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}T_{\alpha\beta}$$

$$T'^{\nu}_{\mu} = V'_{\mu}V'^{\nu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}V_{\alpha}V^{\beta} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}T_{\alpha}{}^{\beta}.$$

3.2 From Eqs. (3.50) and (3.51) with indices suitably relabeled

$$\begin{aligned}
A'_{\mu,\nu} - \Gamma'^{\lambda}_{\mu\nu}A'_{\lambda} &= A_{\alpha,\beta}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}} + A_{\alpha}\frac{\partial^2 x^{\alpha}}{\partial x'^{\nu}\partial x'^{\mu}} \\
&\quad - \left(\Gamma^{\kappa}_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x'^{\lambda}}{\partial x^{\kappa}} + \frac{\partial^2 x^{\alpha}}{\partial x'^{\mu}\partial x'^{\nu}}\frac{\partial x'^{\lambda}}{\partial x^{\alpha}}\right)\frac{\partial x^{\gamma}}{\partial x'^{\lambda}}A_{\gamma} \\
&= A_{\alpha,\beta}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}} + A_{\alpha}\frac{\partial^2 x^{\alpha}}{\partial x'^{\nu}\partial x'^{\mu}} \\
&\quad - \Gamma^{\kappa}_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x'^{\lambda}}{\partial x^{\kappa}}\frac{\partial x^{\gamma}}{\partial x'^{\lambda}}A_{\gamma} - \frac{\partial^2 x^{\alpha}}{\partial x'^{\mu}\partial x'^{\nu}}\frac{\partial x'^{\lambda}}{\partial x^{\alpha}}\frac{\partial x^{\gamma}}{\partial x'^{\lambda}}A_{\gamma} \\
&= A_{\alpha,\beta}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}} + A_{\alpha}\frac{\partial^2 x^{\alpha}}{\partial x'^{\nu}\partial x'^{\mu}} \\
&\quad - \Gamma^{\kappa}_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}A_{\kappa} - A_{\alpha}\frac{\partial^2 x^{\alpha}}{\partial x'^{\nu}\partial x'^{\mu}} \\
&= A_{\alpha,\beta}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}} - \Gamma^{\kappa}_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}A_{\kappa} \\
&= \left(A_{\alpha,\beta} - \Gamma^{\kappa}_{\alpha\beta}A_{\kappa}\right)\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}},
\end{aligned}$$

which is Eq. (3.52).

3.3 (a) Since δ^{ν}_{μ} is a rank-2 tensor with the same components in all coordinate systems (see Section 3.8), under a coordinate transformation $g_{\mu\alpha}g^{\alpha\nu} = \delta^{\nu}_{\mu}$ becomes $g'_{\mu\alpha}g'^{\alpha\nu} = \delta^{\nu}_{\mu}$. Since $g_{\mu\nu}$ is a tensor, if we assume $g^{\mu\nu}$ is also a tensor then

$$g'_{\mu\alpha} = \frac{\partial x^{\kappa}}{\partial x'^{\mu}}\frac{\partial x^{\eta}}{\partial x'^{\alpha}}g_{\kappa\eta}. \qquad g'^{\alpha\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}g^{\rho\sigma}.$$

Then evaluating $g'_{\mu\alpha}g'^{\alpha\nu}$,

$$g'_{\mu\alpha}\frac{\partial x'^{\alpha}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}g^{\rho\sigma} = \frac{\partial x^{\kappa}}{\partial x'^{\mu}}\frac{\partial x^{\eta}}{\partial x'^{\alpha}}g_{\kappa\eta}\frac{\partial x'^{\alpha}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}g^{\rho\sigma} = \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}} = \delta^{\nu}_{\mu},$$

where we have used

$$\frac{\partial x^{\eta}}{\partial x'^{\alpha}}\frac{\partial x'^{\alpha}}{\partial x^{\rho}} = \delta^{\eta}_{\rho} \qquad g_{\kappa\rho}g^{\rho\sigma} = \delta^{\sigma}_{\kappa}.$$

Comparing the result

$$g'_{\mu\alpha}\frac{\partial x'^{\alpha}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}g^{\rho\sigma} = \delta^{\nu}_{\mu}$$

with $g'_{\mu\alpha}g'^{\alpha\nu} = \delta^{\nu}_{\mu}$ requires that

$$g'^{\alpha\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}g^{\rho\sigma}.$$

which is the transformation law for a rank-2 contravariant tensor. Note that this result is an example of the quotient theorem described in Problem 3.13. Since $g_{\mu\alpha}g^{\alpha\nu} = \delta^{\nu}_{\mu}$ and $g_{\mu\nu}$ and δ^{ν}_{μ} are known to be tensors, $g^{\mu\nu}$ must also be a tensor.

(b) From Eq. (3.44) an arbitrary rank-2 tensor can be decomposed into a symmetric and antisymmetric part,

$$g_{\mu\nu} = \tfrac{1}{2}(g_{\mu\nu} + g_{\nu\mu}) + \tfrac{1}{2}(g_{\mu\nu} - g_{\nu\mu}).$$

Inserting this in the line element gives

$$\begin{aligned} ds^2 &= g_{\mu\nu}dx^{\mu}dx^{\nu} \\ &= \tfrac{1}{2}(g_{\mu\nu} + g_{\nu\mu})dx^{\mu}dx^{\nu} + \tfrac{1}{2}(g_{\mu\nu} - g_{\nu\mu})dx^{\mu}dx^{\nu} \\ &= \left[g_{\mu\nu} + \tfrac{1}{2}(g_{\nu\mu} - g_{\nu\mu})\right] dx^{\mu}dx^{\nu} \\ &= g_{\mu\nu}dx^{\mu}dx^{\nu}. \end{aligned}$$

Thus only the symmetric part of $g_{\mu\nu}$ contributes to the line element.

3.4 Under the transformation $x \to x'$,

$$\begin{aligned} T'^{\mu}{}_{\nu} = g'_{\nu\alpha}T'^{\mu\alpha} &= \frac{\partial x^{\alpha}}{\partial x'^{\nu}}\frac{\partial x^{\beta}}{\partial x'^{\alpha}}g_{\alpha\beta}\frac{\partial x'^{\mu}}{\partial x^{\gamma}}\frac{\partial x'^{\alpha}}{\partial x^{\sigma}}T^{\gamma\delta} \\ &= g_{\alpha\delta}T^{\gamma\delta}\frac{\partial x^{\alpha}}{\partial x'^{\nu}}\frac{\partial x'^{\mu}}{\partial x^{\gamma}} = T^{\gamma}{}_{\alpha}\frac{\partial x^{\alpha}}{\partial x'^{\nu}}\frac{\partial x'^{\mu}}{\partial x^{\gamma}}, \end{aligned}$$

where in going from the first line to the second line

$$\frac{\partial x^{\beta}}{\partial x'^{\alpha}}\frac{\partial x'^{\alpha}}{\partial x^{\delta}} = \delta^{\beta}_{\delta}$$

has been used. This is a tensor transformation law so it is valid in all frames. Proceeding in similar fashion,

$$\begin{aligned} T'_{\mu\nu} = g'_{\mu\alpha}g'_{\nu\beta}T'^{\alpha\beta} &= \frac{\partial x^{\varepsilon}}{\partial x'^{\mu}}\frac{\partial x^{\lambda}}{\partial x'^{\alpha}}g_{\varepsilon\lambda}\frac{\partial x^{\gamma}}{\partial x'^{\nu}}\frac{\partial x^{\delta}}{\partial x'^{\beta}}g_{\gamma\delta}\frac{\partial x'^{\alpha}}{\partial x^{\tau}}\frac{\partial x'^{\beta}}{\partial x^{\theta}}T^{\tau\theta} \\ &= \frac{\partial x^{\varepsilon}}{\partial x'^{\mu}}\frac{\partial x^{\gamma}}{\partial x'^{\nu}}T_{\varepsilon\gamma}, \end{aligned}$$

where in the last step

$$\frac{\partial x^{\lambda}}{\partial x'^{\alpha}}\frac{\partial x'^{\alpha}}{\partial x^{\tau}} = \delta^{\lambda}_{\tau} \qquad \frac{\partial x^{\delta}}{\partial x'^{\beta}}\frac{\partial x'^{\beta}}{\partial x^{\theta}} = \delta^{\delta}_{\theta} \qquad g_{\varepsilon\lambda}g_{\gamma\delta}T^{\lambda\delta} = T_{\varepsilon\gamma}$$

have been used. This is a tensor transformation law so it is valid in all frames.

3.5 (a) For example, consider a rank-4 tensor $T_\beta{}^{\mu\nu\alpha}$. Its transformation law is

$$T'^{\mu\nu\alpha}_{\beta} = \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}\frac{\partial x'^\alpha}{\partial x^\varepsilon}\frac{\partial x^\eta}{\partial x'^\beta}T_\eta{}^{\gamma\delta\varepsilon}.$$

Now set $\alpha = \beta$ for this tensor (implying a sum on this index). The resulting quantity must have two upper indices by the summation convention, so define it to be $T^{\mu\nu}$:

$$T^{\mu\nu} \equiv \delta^\beta_\alpha T_\beta{}^{\mu\nu\alpha} = T_\alpha{}^{\mu\nu\alpha}.$$

Is $T^{\mu\nu}$ a tensor? From the preceding equations, its transformation law is

$$\begin{aligned}
T'^{\mu\nu} &\equiv T'^{\mu\nu\alpha}_{\alpha} = \delta^\beta_\alpha T'^{\mu\nu\alpha}_{\beta} \\
&= \delta^\beta_\alpha \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}\frac{\partial x'^\alpha}{\partial x^\varepsilon}\frac{\partial x^\eta}{\partial x'^\beta}T_\eta{}^{\gamma\delta\varepsilon} = \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}\frac{\partial x'^\alpha}{\partial x^\varepsilon}\frac{\partial x^\eta}{\partial x'^\alpha}T_\eta{}^{\gamma\delta\varepsilon} \\
&= \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}\delta^\eta_\varepsilon T_\eta{}^{\gamma\delta\varepsilon} = \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}T_\eta{}^{\gamma\delta\eta} = \frac{\partial x'^\mu}{\partial x^\gamma}\frac{\partial x'^\nu}{\partial x^\delta}T^{\gamma\delta},
\end{aligned}$$

which is the transformation law for a contravariant rank-2 tensor. Similar proofs can be carried out for tensors of any order. Thus, setting an upper and lower index equal on a rank-N tensor and summing yields a tensor of rank $N-2$.

(b) For example, consider the linear combination of two rank-2 tensors, $T_\mu{}^\nu = aA_\mu{}^\nu + bB_\mu{}^\nu$. The transformation law is

$$\begin{aligned}
T'^{\nu}_{\mu} = aA'^{\nu}_{\mu} + bB'^{\nu}_{\mu} &= a\frac{\partial x'^\nu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\mu}A_\beta{}^\alpha + b\frac{\partial x'^\nu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\mu}B_\beta{}^\alpha \\
&= \frac{\partial x'^\nu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\mu}\left(aA_\beta{}^\alpha + bB_\beta{}^\alpha\right) = \frac{\partial x'^\nu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\mu}T_\beta{}^\alpha.
\end{aligned}$$

A similar proof holds for any such linear combination of tensors.

3.6 The line element is $ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\varphi^2$, so the non-zero components of the metric are

$$g_{00} = g_{tt} = -1 \quad g_{11} = g_{rr} = 1 \quad g_{22} = g_{\theta\theta} = r^2 \quad g_{33} = g_{\varphi\varphi} = r^2\sin^2\theta$$

and $\det g_{\mu\nu} = -r^4\sin^2\theta$. Then from Eq. (3.48) the invariant volume element is

$$dV = (-\det g_{\mu\nu})^{1/2} dr\, d\theta\, d\varphi = r^2 dr \sin\theta\, d\theta\, d\varphi,$$

which gives a volume

$$V = \int dV = \int_0^R r^2 dr \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\varphi = \tfrac{4}{3}\pi R^3,$$

as expected.

3.7 Since $A\cdot B = A_\mu B^\mu$ is a scalar it is unchanged by a coordinate transformation. Thus from the vector transformation law for B^μ

$$A'_\mu B'^\mu = A_\mu B^\mu = A_\nu \frac{\partial x^\nu}{\partial x'^\mu}B'^\mu \quad \longrightarrow \quad \left(A_\nu\frac{\partial x^\nu}{\partial x'^\mu} - A'_\mu\right)B'^\mu = 0.$$

But B'^{μ} is an arbitrary vector that does not generally vanish. Thus the quantity in parentheses must be equal to zero, implying that $A'_{\mu} = \left(\partial x^{\nu}/\partial x'^{\mu}\right) A_{\nu}$, which is the transformation law for a dual vector.

3.8 This problem is adapted from an example in Ref. [88]. From the transformation equations between spherical and cylindrical coordinates assuming $u = (r, \theta, \varphi)$ and $u' = (\rho, \varphi, z)$,

$$
\begin{aligned}
u'^1 &= \rho = r\sin\theta = u^1 \sin u^2 \\
u'^2 &= \varphi = u^3 \\
u'^3 &= z = r\cos\theta = u^1 \cos u^2
\end{aligned}
$$

and the inverse transformations are

$$
\begin{aligned}
u^1 &= r = \sqrt{\rho^2 + z^2} = \sqrt{(u'^1)^2 + (u'^3)^2} \\
u^2 &= \theta = \tan^{-1}\left(\frac{\rho}{z}\right) = \tan^{-1}\left(\frac{u'^1}{u'^3}\right) \\
u^3 &= \varphi = u'^2.
\end{aligned}
$$

From these the partial derivative entries in the matrices U and $\hat{U}$ defined in Example 3.7 may be computed directly. For example,

$$
\begin{aligned}
U^1_2 &= \frac{\partial u'^1}{\partial u^2} = \frac{\partial}{\partial u^2}(u^1 \sin u^2) = u^1 \cos u^2 = r\cos\theta \\
\hat{U}^2_1 &= \frac{\partial u^2}{\partial u'^1} = \frac{\partial}{\partial u'^1}\left[\tan^{-1}\left(\frac{u'^1}{u'^3}\right)\right] = \frac{u'^3}{(u'^1)^2 + (u'^3)^2} = \frac{\cos\theta}{r}.
\end{aligned}
$$

Computing all the derivatives and assembling them gives

$$
U = \begin{pmatrix} \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} \qquad \hat{U} = \begin{pmatrix} \sin\theta & 0 & \cos\theta \\ \dfrac{\cos\theta}{r} & 0 & -\dfrac{\sin\theta}{r} \\ 0 & 1 & 0 \end{pmatrix},
$$

and by explicit matrix multiplication, $\hat{U}U = I$.

3.9 From Eqs. (3.45) and (3.46),

$$
\begin{aligned}
T_{[\alpha\beta](\gamma\delta)} &= \frac{1}{2}\left(T_{\alpha\beta(\gamma\delta)} - T_{\beta\alpha(\gamma\delta)}\right) \\
&= \frac{1}{2}\left(\tfrac{1}{2}(T_{\alpha\beta\gamma\delta} + T_{\alpha\beta\delta\gamma}) - \tfrac{1}{2}(T_{\beta\alpha\gamma\delta} + T_{\beta\alpha\delta\gamma})\right) \\
&= \frac{1}{4}\left(T_{\alpha\beta\gamma\delta} + T_{\alpha\beta\delta\gamma} - T_{\beta\alpha\gamma\delta} - T_{\beta\alpha\delta\gamma}\right).
\end{aligned}
$$

3.10 (a) Use the symmetry properties and relabeling of dummy indices to write,

$$\begin{aligned} A^{\mu\nu}B_{\mu\nu} &= -A^{\nu\mu}B_{\mu\nu} && (A^{\mu\nu} \text{ is antisymmetic}) \\ &= -A^{\nu\mu}B_{\nu\mu} && (B_{\mu\nu} \text{ is symmetric}) \\ &= -A^{\mu\nu}B_{\mu\nu} && (\text{Interchange dummy indices } \mu \leftrightarrow \nu). \end{aligned}$$

But $A^{\mu\nu}B_{\mu\nu} = -A^{\mu\nu}B_{\mu\nu}$ can be true only if $A^{\mu\nu}B_{\mu\nu} = 0$.

(b) For example, if $A^{\mu\nu}$ is symmetric, $A^{\mu\nu} = A^{\nu\mu}$, then

$$A'^{\mu\nu} = \frac{\partial x'^{\mu}}{\partial x^{\gamma}}\frac{\partial x'^{\nu}}{\partial x^{\delta}}A^{\gamma\delta} = \frac{\partial x'^{\mu}}{\partial x^{\gamma}}\frac{\partial x'^{\nu}}{\partial x^{\delta}}A^{\delta\gamma} = A'^{\nu\mu},$$

with an analogous proof if $A^{\mu\nu} = -A^{\nu\mu}$.

3.11 Contracting δ^{μ}_{ν} with the components V^{ν} of an arbitrary vector gives

$$\delta^{\mu}_{\nu}V^{\nu} = V^{\mu} = g^{\mu\alpha}V_{\alpha} = g^{\mu\alpha}g_{\alpha\nu}V^{\nu}.$$

But V is arbitrary so $g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}_{\nu}$.

3.12 Multiply both sides of $T_{\mu\nu} = U_{\mu\nu}$ by $\partial x^{\mu}/\partial x'^{\alpha}$ and $\partial x^{\nu}/\partial x'^{\beta}$ and take the implied sums to give

$$\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}T_{\mu\nu} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}U_{\mu\nu}.$$

But from Eq. (3.36) this is just $T'_{\mu\nu} = U'_{\mu\nu}$.

3.13 In the scalar product expression $A \cdot B = g_{\mu\nu}A^{\mu}B^{\nu}$ of Eq. (3.43) the left side is a scalar and A and B on the right side are vectors. Since the quantities $g_{\mu\nu}$ contracted with tensors on the right side yield a tensor on the left side, by the quotient theorem $g_{\mu\nu}$ must define the components of a type $(0,2)$ tensor.

3.14 This solution is adapted from Example 1.8.1 in Ref. [88]. For an arbitrary contravariant vector V^{γ} the transformation law given in the problem is

$$T^{\alpha}{}_{\beta\gamma}V^{\gamma} = \frac{\partial x'^{\alpha}}{\partial x^{\delta}}\frac{\partial x^{\varepsilon}}{\partial x'^{\beta}}T^{\delta}{}_{\varepsilon\varphi}V^{\varphi},$$

indicating that $T^{\alpha}{}_{\beta\gamma}V^{\gamma}$ transforms as a $(1,1)$ tensor. By the quotient theorem then $T^{\alpha}{}_{\beta\gamma}$ must be a $(1,2)$ tensor. The proof follows from inserting $V^{\gamma} = (\partial x'^{\gamma}/\partial x^{\varphi})V^{\varphi}$ on the left side of the above equation and rearranging to give

$$\left(T^{\alpha}{}_{\beta\gamma}\frac{\partial x'^{\gamma}}{\partial x^{\varphi}} - \frac{\partial x'^{\alpha}}{\partial x^{\delta}}\frac{\partial x^{\varepsilon}}{\partial x'^{\beta}}T^{\delta}{}_{\varepsilon\varphi}\right)V^{\varphi} = 0.$$

This must be valid for any V^{φ} so choose $V_{\varphi} = \delta^{\varphi}_{\lambda}$ such that the quantity inside the parentheses is required to vanish, giving

$$T^{\alpha}{}_{\beta\gamma}\frac{\partial x'^{\gamma}}{\partial x^{\lambda}} = \frac{\partial x'^{\alpha}}{\partial x^{\delta}}\frac{\partial x^{\varepsilon}}{\partial x'^{\beta}}T^{\delta}{}_{\varepsilon\lambda}.$$

Multiply both sides of this expression by $\partial x^\lambda / \partial x'^\mu$ to give

$$T^{\alpha}_{\ \beta\gamma} \frac{\partial x'^\gamma}{\partial x^\lambda} \frac{\partial x^\lambda}{\partial x'^\mu} = \frac{\partial x'^\alpha}{\partial x^\delta} \frac{\partial x^\varepsilon}{\partial x'^\beta} \frac{\partial x^\lambda}{\partial x'^\mu} T^{\delta}_{\ \varepsilon\lambda}.$$

But on the left side

$$\frac{\partial x'^\gamma}{\partial x^\lambda} \frac{\partial x^\lambda}{\partial x'^\mu} = \delta^\gamma_\mu,$$

giving finally

$$T^{\alpha}_{\ \beta\mu} = \frac{\partial x'^\alpha}{\partial x^\delta} \frac{\partial x^\varepsilon}{\partial x'^\beta} \frac{\partial x^\lambda}{\partial x'^\mu} T^{\delta}_{\ \varepsilon\lambda},$$

which is the transformation law obeyed by a $(1,2)$ tensor.

3.15 (a) One may write

$$\delta^\nu_\mu \frac{\partial x'^\alpha}{\partial x^\nu} \frac{\partial x^\mu}{\partial x'^\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x^\mu}{\partial x'^\beta} = \frac{\partial x'^\alpha}{\partial x'^\beta} = \delta'^\alpha_\beta$$

which is the transformation law for a mixed, rank-2 tensor.

(b) In some coordinate system let $K^\nu_\mu = \delta^\nu_\mu = \mathrm{diag}\,(1,1,1,1)$. Then under an arbitrary coordinate transformation,

$$K'^\nu_\mu = \frac{\partial x'^\nu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x'^\mu} K^\alpha_\beta = \frac{\partial x'^\nu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x'^\mu} \delta^\alpha_\beta = \frac{\partial x'^\nu}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial x'^\mu} = \delta^\nu_\mu.$$

Thus $K^\nu_\mu = \delta^\nu_\mu$ is a $(1,1)$ tensor that has the same components (those of the unit matrix) in any coordinate system.

3.16 This is a particular example of a scalar product, so it must transform as a scalar. Explicitly,

$$\begin{aligned}
ds'^2 &= g_{\alpha\beta} \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} \frac{\partial x'^\mu}{\partial x^\gamma} ds^\gamma \frac{\partial x'^\nu}{\partial x^\delta} ds^\delta \\
&= g_{\alpha\beta} ds^\gamma ds^\delta \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x'^\mu}{\partial x^\gamma} \frac{\partial x^\beta}{\partial x'^\nu} \frac{\partial x'^\nu}{\partial x^\delta} \\
&= g_{\alpha\beta} ds^\gamma ds^\delta \frac{\partial x^\alpha}{\partial x^\gamma} \frac{\partial x^\beta}{\partial x^\delta} = g_{\alpha\beta} ds^\gamma ds^\delta \delta^\alpha_\gamma \delta^\beta_\delta \\
&= g_{\alpha\beta} ds^\alpha ds^\beta = ds^2
\end{aligned}$$

where Eq. (3.35) has been used. The squared line element (3.39) is clearly a scalar invariant and so it has the same value in all coordinate systems.

3.17 By the usual rank-2 tensor transformation law,

$$T'^{\mu\nu}(x') = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} T^{\alpha\beta}(x).$$

Upon differentiating Eq. (3.66),

$$\frac{\partial x'^\mu}{\partial x^\alpha} = \delta^\mu_\alpha + (\delta u)\,\partial_\alpha X^\mu(x),$$

which may be substituted into the first equation to give

$$
\begin{aligned}
T'^{\mu\nu}(x') &= \left(\delta^\mu_\alpha + (\delta u)\partial_\alpha X^\mu\right)\left(\delta^\nu_\beta + (\delta u)\partial_\beta X^\nu\right)T^{\alpha\beta}(x) \\
&= \left(\delta^\mu_\alpha\delta^\nu_\beta + \delta^\mu_\alpha(\delta u)\partial_\beta X^\nu + \delta^\nu_\beta(\delta u)\partial_\alpha X^\mu + \mathscr{O}\left(\delta u^2\right)\right)T^{\alpha\beta}(x) \\
&= T^{\mu\nu}(x) + \left[\partial_\beta X^\nu T^{\mu\beta} + \partial_\alpha X^\mu T^{\alpha\nu}\right]\delta u,
\end{aligned}
$$

where only terms first-order in δu have been retained.

3.18 The transformation law for dual vectors is given by Eq. (3.29). Using the expansion (3.66) to evaluate the partial derivative gives

$$
A'_\mu(x') = \frac{\partial x^\alpha}{\partial x'^\mu}A_\alpha(x) = \left(\delta^\alpha_\mu - \frac{\partial X^\alpha}{\partial x'^\mu}\right)\delta u A_\alpha(x) = A_\mu(x) - \partial_\mu X^\alpha(\delta u)A_\alpha(x).
$$

By analogy with Eq. (3.68) the Lie derivative is then

$$
\mathscr{L}_X A_\mu \equiv \operatorname*{Lim}_{\delta u\to 0}\left(\frac{A_\mu(x') - A'_\mu(x')}{\delta u}\right) = X^\alpha\partial_\alpha A_\mu + A_\alpha\partial_\mu X^\alpha,
$$

where a Taylor expansion as in Eq. (3.69) was used to evaluate $A_\mu(x')$.

3.19 Let $A_{\mu\nu} = U_u V_\nu$. Then by the Leibniz rule,

$$
\begin{aligned}
\mathscr{L}_X A_{\mu\nu} &= \mathscr{L}_X(U_\mu V_\nu) = (\mathscr{L}_X U_\mu)V_\nu + U_\mu(\mathscr{L}_X V_\nu) \\
&= X^\alpha\left[(\partial_\alpha U_\mu)V_\nu + U_\mu(\partial_\alpha V_\nu)\right] + U_\alpha(\partial_\mu X^\alpha)V_\nu + U_\mu V_\alpha(\partial_\nu X^\alpha) \\
&= X^\alpha\partial_\alpha A_{\mu\nu} + A_{\alpha\nu}\partial_\mu X^\alpha + A_{\mu\alpha}\partial_\nu X^\alpha,
\end{aligned}
$$

where in the second line Eq. (3.73) was used and in the third line $A_{\mu\nu} = U_u V_\nu$ and

$$
\partial_\alpha A_{\mu\nu} = \partial_\alpha(U_\mu V_\nu) = U_\mu(\partial_\alpha V^\nu) + (\partial_\alpha U_\mu)V_\nu
$$

were used.

3.20 Let $C = [A,B] = AB - BA$ and operate on an arbitrary function f,

$$
\begin{aligned}
Cf &= [A,B]f = ABf - BAf \\
&= A^\nu\partial_\nu(B^\mu\partial_\mu f) - B^\nu\partial_\nu(A^\mu\partial_\mu f) \\
&= A^\nu\partial_\nu B^\mu\partial_\mu f + A^\nu B^\mu\partial_\nu\partial_\mu f - B^\nu\partial_\nu A^\mu\partial_\mu f - B^\nu A^\mu\partial_\nu\partial_\mu f \\
&= (A^\nu\partial_\nu B^\mu)\partial_\mu f - (B^\nu\partial_\nu A^\mu)\partial_\mu f,
\end{aligned}
$$

where in the second line the vectors were expanded in the basis ∂_ν and the third line results from taking the partial derivative of the product. Since the function f is arbitrary, this implies the operator relation

$$
C = [A,B] = (A^\nu\partial_\nu B^\mu - B^\nu\partial_\nu A^\mu)\partial_\mu,
$$

and since ∂_μ is a vector basis, C is a vector with components

$$
C^\mu = [A,B]^\mu = A^\nu\partial_\nu B^\mu - B^\nu\partial_\nu A^\mu,
$$

which defines the *Lie bracket* $[A,B] = -[B,A]$ for the vectors A and B. Comparison with

Eq. (3.72) indicates that the Lie bracket is equivalent to a Lie derivative of a vector field: $[A,B]^\mu = \mathscr{L}_A B^\mu$. The Lie derivative of a tensor then may be viewed as a generalization of the Lie bracket for vectors.

3.21 (a) From Eqs. (3.15)–(3.17) and Example 3.4,

$$V(e^\mu) = V^\nu e_\nu(e^\mu) = \delta^\mu_\nu V^\nu = V^\mu$$
$$\omega(e_\mu) = \omega_\nu e^\nu(e_\mu) = \omega_\nu \delta^\nu_\mu = \omega_\mu,$$

which is Eq. (3.19).

(b) For vectors $V = V^\alpha e_\alpha$, by the chain rule under a coordinate transformation $x^\mu \to x'^\mu$ the basis vectors transform as

$$e_\alpha \to e'_\alpha = \frac{\partial x^\nu}{\partial x'^\alpha} e_\nu.$$

Thus, to keep V invariant under $x^\mu \to x'^\mu$ its components must transform as

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu,$$

which is equivalent to (3.31), since then

$$\begin{aligned} V \to V' = V'^\mu e'_\mu &= \frac{\partial x'^\mu}{\partial x^\nu} V^\nu \frac{\partial x^\alpha}{\partial x'^\mu} e_\alpha \\ &= \frac{\partial x'^\mu}{\partial x^\nu} \frac{\partial x^\alpha}{\partial x'^\mu} V^\nu e_\alpha \\ &= \delta^\alpha_\nu V^\nu e_\alpha \\ &= V^\alpha e_\alpha = V. \end{aligned}$$

3.22 The first two examples are trivial. Since two successive partial derivative operations commute,

$$\left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right] = 0 \qquad \left[\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right] = 0$$

and obviously these are coordinate bases. But for the third example

$$\begin{aligned} [\hat{e}_1, \hat{e}_2,] = \left[\frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{\partial \theta}\right] &= \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial \theta}\right) - \frac{1}{r}\frac{\partial}{\partial \theta}\frac{\partial}{\partial r} \\ &= -\frac{1}{r^2}\frac{\partial}{\partial \theta} + \frac{1}{r}\frac{\partial^2}{\partial r \partial \theta} - \frac{1}{r}\frac{\partial^2}{\partial r \partial \theta} \\ &= -\frac{1}{r^2}\frac{\partial}{\partial \theta} = -\frac{\hat{e}_2}{r} \neq 0. \end{aligned}$$

Thus $\hat{e}_1$ and $\hat{e}_2$ do not commute and they define a non-coordinate basis.

3.23 The Lie derivative for a vector is given by Eq. (3.72). Replacing the partial derivatives with covariant derivatives in this expression gives

$$\begin{aligned} \mathscr{L}_X A^\mu = X^\alpha \partial_\alpha A^\mu - A^\alpha \partial_\alpha X^\mu &\longrightarrow X^\alpha \nabla_\alpha A^\mu - A^\alpha \nabla_\alpha X^\mu \\ &= X^\alpha \left(\partial_\alpha A^\mu + \Gamma^\mu_{\beta\alpha} A^\beta\right) - A^\alpha \left(\partial_\alpha X^\mu + \Gamma^\mu_{\beta\alpha} X^\beta\right). \end{aligned}$$

The terms involving the connection coefficients cancel exactly:

$$\Gamma^\mu_{\beta\alpha} X^\alpha A^\beta - \Gamma^\mu_{\beta\alpha} X^\beta A^\alpha = \Gamma^\mu_{\alpha\beta} X^\beta A^\alpha - \Gamma^\mu_{\beta\alpha} X^\beta A^\alpha = 0,$$

where the dummy summation indices α and β have been interchanged in the first term and the symmetry of the connection coefficient in its lower indices has been invoked in the last step. Therefore,

$$\mathscr{L}_X A^\mu = X^\alpha \partial_\alpha A^\mu - A^\alpha \partial_\alpha X^\mu \;\leftrightarrow\; X^\alpha \nabla_\alpha A^\mu - A^\alpha \nabla_\alpha X^\mu.$$

It can be shown generally that for the Lie derivative of any tensor all partial derivatives may be replaced by covariant derivatives and vice versa on a manifold with a torsion-free connection because, as in the above example, the correction terms that convert a partial derivative to a covariant derivative vanish identically in the Lie derivative if the connection coefficient is symmetric in its lower indices.

3.24 The infinitesimal displacement ds must be invariant under coordinate transformation: $ds = ds'$. Expand both sides in the basis e_μ to give

$$ds = dx^\mu e_\mu = dx'^\mu e'_\mu.$$

But $dx^\mu = \dfrac{\partial x^\mu}{\partial x'^\nu} dx'^\nu$, so

$$\frac{\partial x^\mu}{\partial x'^\nu} dx'^\nu e_\mu = dx'^\mu e'_\mu.$$

This is true generally only if $e'_\mu = \dfrac{\partial x^\nu}{\partial x'^\mu} e_\nu$. By a similar proof, $e'^\mu = \dfrac{\partial x'^\mu}{\partial x^\nu} e^\nu$.

3.25 An arbitrary vector A can be expanded as

$$A = A^\mu e_\mu = A^\mu \frac{\partial}{\partial x^\mu} = A^\mu \partial_\mu,$$

using the basis (3.6). Require that a vector A be unchanged by a transformation to a primed coordinate system, $A' = A$, so

$$A'^\nu \frac{\partial}{\partial x'^\nu} = A^\nu \frac{\partial}{\partial x^\nu}.$$

Operate on x'^μ with both sides and invoke $\partial x'^\mu / \partial x'^\nu = \delta^\mu_\nu$ to give

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu,$$

which is the transformation law (3.31) for a vector.

3.26 (a) The components may be evaluated by inserting basis dual vectors as arguments:

$$(U \otimes V)(e_\mu, e_\nu) = U(e_\mu) V(e_\nu) = U_\mu V_\nu,$$

where Eq. (3.19) was used.

(b) Insert basis states $\{e^\mu, e^\nu, e_\lambda, e^\varepsilon\}$ as arguments, giving

$$(U \otimes V \otimes \Omega \otimes W)(e^\mu, e^\nu, e_\lambda e^\varepsilon) = U(e^\mu) V(e^\nu) \Omega(e_\lambda) W(e^\varepsilon) = U^\mu V^\nu \Omega_\lambda W^\varepsilon \equiv S^{\mu\nu}{}_\lambda{}^\varepsilon,$$

where Eq. (3.19) was used.

(c) Generalizing part (a), the tensor product is defined through

$$(T \otimes V)(A,B,C)) = T(A,B)V(C).$$

Inserting basis states as the arguments gives for the mixed-tensor components

$$T(e^{\mu}, e_{\nu})V(e_{\gamma}) = T^{\mu}{}_{\nu}V_{\gamma} \equiv S^{\mu}{}_{\nu\gamma}.$$

(d) The tensor product gives

$$(e_{\mu} \otimes e_{\nu})(e^{\alpha}, e^{\beta}) = e_{\mu}(e^{\alpha})e_{\nu}(e^{\beta}) = \delta_{\mu}^{\alpha}\,\delta_{\nu}^{\beta},$$

where Eqs. (3.19) and (3.17) were used. Hence

$$T^{\mu\nu}(e_{\mu} \otimes e_{\nu})(e^{\alpha}, e^{\beta}) = T^{\mu\nu}\delta_{\mu}^{\alpha}\delta_{\nu}^{\beta} = T^{\alpha\beta},$$

which are the contravariant components of T. Therefore we can expand T as

$$T = T^{\mu\nu}(e_{\mu} \otimes e_{\nu}),$$

and we see that $e_{\mu} \otimes e_{\nu}$ acts as a basis for $T = U \otimes V$. More generally, we can expand T in any of the forms

$$T = T^{\mu\nu}(e_{\mu} \otimes e_{\nu}) = T_{\mu\nu}(e^{\mu} \otimes e^{\nu}) = T_{\mu}{}^{\nu}(e^{\mu} \otimes e_{\nu}) = T^{\mu}{}_{\nu}(e_{\mu} \otimes e^{\nu})$$

by inserting different combinations of basis vectors or basis dual vectors in the preceding derivation.

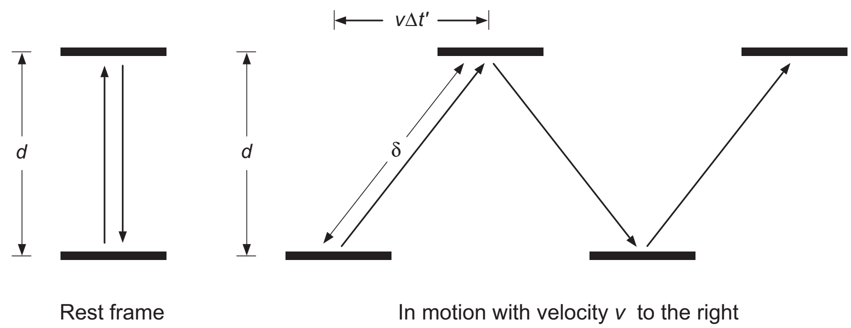
4 Lorentz Covariance and Special Relativity

4.1 After the transformation given by Eq. (4.20) the line element is

$$\begin{aligned} ds'^2 &= -c^2(dt')^2 + (dx')^2 + (dy')^2 + (dz')^2 \\ &= -(c\cosh\xi\, dt + \sinh\xi\, dx)^2 + (c\sinh\xi\, dt + \cosh\xi\, dx)^2 + dy^2 + dz^2 \\ &= -c^2dt^2 + dx^2 + dy^2 + dz^2 = ds^2, \end{aligned}$$

so it is invariant under the transformation.

4.2 Consider the following diagram:



The period in the rest frame of the clock is $\Delta t = d/c$ but the light for the moving observer is seen to travel a distance $\delta > d$ at a speed c in one tick. From the geometry of the diagram, the elapsed time observed for one tick of the moving clock is

$$\Delta t' = \frac{\delta}{c} = \frac{\sqrt{d^2 + (v\Delta t')^2}}{c}$$

Square both sides and solve for $\Delta t'$ to give

$$\Delta t' = \frac{d}{c\sqrt{1 - v^2/c^2}} = \gamma\frac{d}{c} = \gamma\Delta t,$$

which is the special relativistic time dilation formula: the observer in motion with respect to the clock sees the clock run more slowly an observer in the rest frame of the clock.

4.3 For infinitesimal displacements the Lorentz transformation (4.26) is

$$dt' = \gamma\big(dt - (v/c^2)dx\big) \qquad dx' = \gamma(dx - vdt) \qquad dy' = dy \qquad dz' = dz.$$

The velocity transformations are then obtained by evaluating derivatives of displacements with respect to time:

$$u'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma(dt - vdx/c^2)} = \frac{dx/dt - v}{1-(v/c^2)dx/dt} = \frac{u_x - v}{1 - vu_x/c^2}$$

$$u'_y = \frac{dy'}{dt'} = \frac{dy}{\gamma(dt - vdx/c^2)} = \frac{dy/dt}{\gamma(1-(v/c^2)dx/dt)} = \frac{u_y}{\gamma(1-(v/c^2)u_x)}$$

$$u'_z = \frac{dz'}{dt'} = \frac{dz}{\gamma(dt - vdx/c^2)} = \frac{dz/dt}{\gamma(1-(v/c^2)dx/dt)} = \frac{u_z}{\gamma(1-(v/c^2)u_x)}$$

For all motion along the x axis one obtains the standard velocity addition formula

$$u' = \frac{u-v}{1-uv/c^2}.$$

Setting $u = c$ gives $u' = (c-v)/(1-v/c) = c$, independent of v, implying a constant speed of light in all inertial frames, while for $v << 0$ one obtains $u' = u - v$, as expected from Galilean invariance [see Eq. (2.43)].

4.4 The Minkowski line interval is $(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2$, so the lengths of the sides in units of the grid spacing are

$$\overline{\mathrm{AB}} = (6^2)^{1/2} = 6 \qquad \overline{\mathrm{BC}} = (3^2)^{1/2} = 3 \qquad \overline{\mathrm{AC}} = (-3^2 + 6^2)^{1/2} = (27)^{1/2} = 5.2.$$

Thus AB is the longest side and BC is the shortest side. The distance by the path A $\rightarrow$ C is 5.2 grid spacings and the distance by the path A $\rightarrow$ B $\rightarrow$ C is $6+3 = 9$ grid spacings.

4.5 Two vectors A^μ and A^ν are orthogonal if their inner (scalar) product vanishes, $g_{\mu\nu}A^\mu A^\nu = 0$. But for lightlike (null) vectors X, $g_{\mu\nu}X^\mu X^\nu = X_\nu X^\nu = 0$. Therefore, a lightlike vector must be orthogonal to itself.

4.6 The variable ξ is a relativistic velocity parameter so the inverse transformation corresponds to $\xi \rightarrow -\xi$. Since $\sinh(-x) = \sinh x$ and $\cosh(-x) = \cosh x$, one has for the inverse of the transformation defined by Eq. (4.20),

$$\begin{pmatrix} cdt' \\ dx' \end{pmatrix} = \begin{pmatrix} \cosh\xi & -\sinh\xi \\ -\sinh\xi & \cosh\xi \end{pmatrix} \begin{pmatrix} cdt \\ dx \end{pmatrix}.$$

Because $i\sinh x = \sin(ix)$ and $\cosh x = \cos(ix)$, the Lorentz boost of Eq. (4.20) may be interpreted a a rotation through an imaginary angle.

4.7 A monochromatic source moves with respect to an observer with a radial velocity v. Let dt be a time interval measured in the rest frame of the source and dt' be a corresponding time interval measured by the observer. From Example 4.2,

$$dt = \frac{dt'}{\gamma} \qquad \gamma \equiv \frac{1}{\sqrt{1-v^2/c^2}}.$$

In dt' the distance between source and observer increases by vdt' (as seen by the observer). The time for light to travel this extra distance is $(v/c)dt'$. Thus the total time between wavecrests seen by the observer is

$$\Delta t = dt' + \frac{v}{c}dt' = \left(1+\frac{v}{c}\right)dt' = \left(1+\frac{v}{c}\right)\gamma dt = \frac{1+v/c}{\sqrt{1-v^2/c^2}}\,dt$$

Since frequency is inversely related to time, the ratio of the rest-frame frequency ν_0 and the observed frequency ν will be given by

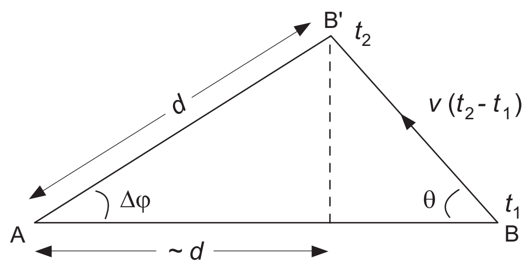
$$\frac{\nu_0}{\nu} = \frac{1/dt}{1/\Delta t} = \frac{\Delta t}{dt} = \frac{(1+v/c)dt}{\sqrt{1-v^2/c^2}\,dt}.$$

Thus the relativistic Doppler shift is given by

$$\frac{\nu_0}{\nu} = \frac{1+v/c}{\sqrt{1-v^2/c^2}} = \sqrt{\frac{1+v/c}{1-v/c}}.$$

If v^2/c^2 is small it can be ignored in the denominator and one recovers the usual nonrelativistic Doppler formula $\Delta\nu/\nu = v/c$.

4.8 From the diagram



(i) $\Delta\varphi \simeq v\delta t \sin\theta/d$. Observer A sees the light from B at t_1' and from B′ at t_2', with

$$t_1' = t_1 + \frac{d + v\delta t\cos\theta}{c} \qquad t_2' = t_2 + \frac{d}{c}.$$

The time measured at A for the source to move from B to B′ is

$$\begin{aligned}\Delta t = t_2' - t_1' &= t_2 + \frac{d}{c} - t_1 - \frac{d + v\delta t\cos\theta}{c}\\ &= t_2 - t_1 - \frac{v\delta t\cos\theta}{c} = \delta t(1-\beta\cos\theta),\end{aligned}$$

where $\beta \equiv v/c$ and $\delta t \equiv t_2 - t_1$. Then the apparent transverse velocity for the motion B to B′ observed at A is

$$\beta_{\mathrm{T}} \equiv \frac{v_{\mathrm{T}}}{c} = \frac{d}{c}\frac{\Delta\varphi}{\Delta t} = \frac{d}{c}\frac{v\delta t\sin\theta/d}{\delta t(1-\beta\cos\theta)} = \frac{\beta\sin\theta}{1-\beta\cos\theta}.$$

(ii) The maximum for β_{T} is found from the above formula by setting $\partial\beta_{\mathrm{T}}/\partial\theta = 0$. Taking the derivative, setting it equal to zero, and using $\sin(\cos^{-1}\beta) = (1-\beta^2)^{1/2}$, yields that the maximum value of β_{T} is

$$\beta_{\mathrm{T}}^{\max} = \frac{\beta}{1-\beta^2},$$

where β is the actual velocity (in units of c) and β_{T} is the apparent velocity. Thus, as β approaches its physical maximum of unity, the apparent transverse velocity grows without bound and it is possible to observe any transverse velocity, even those appearing to exceed the speed of light.

(iii) Setting $\theta = 10^\circ$ and $\beta = 0.995$, gives from the preceding formula $\beta_T^{\max} = 8.6$. Thus the apparent transverse velocity is observed to be 8.6 times that of light (superluminal), even though the actual transverse velocity is only $v = 0.995c$ (subluminal).

4.9 From the general transformation law for a rank-2 tensor, $\eta_{\nu\alpha} = \Lambda^{\mu}{}_{\nu}\Lambda^{\beta}{}_{\alpha}\eta_{\mu\beta}$, and from the index raising and lowering properties of the metric

$$\Lambda_{\mu}{}^{\lambda} = \eta_{\mu\beta}\eta^{\lambda\alpha}\Lambda^{\beta}{}_{\alpha},$$

where we define the matrices

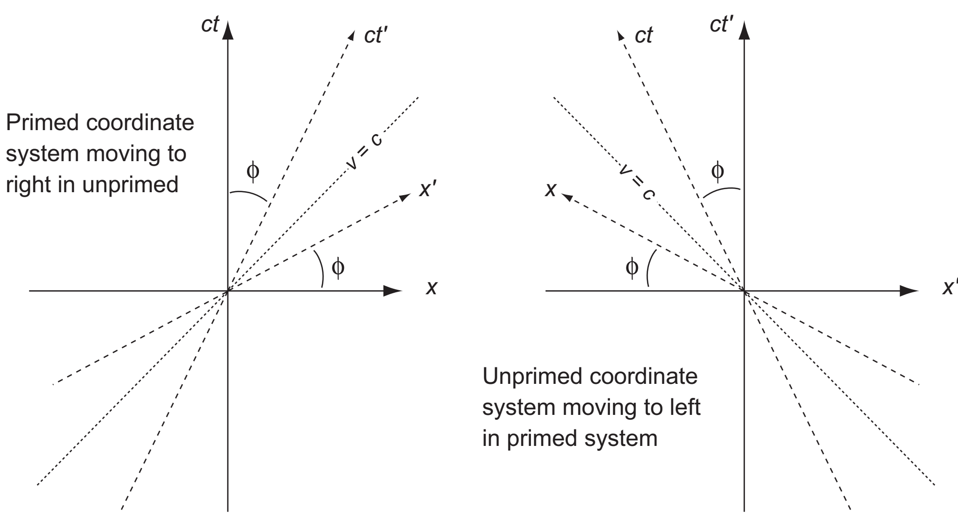
$$[\Lambda^{\mu}{}_{\nu}] \equiv \left[\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right] \qquad [\Lambda_{\mu}{}^{\nu}] \equiv \left[\frac{\partial x^{\nu}}{\partial x'^{\mu}}\right].$$

This may be used to show that

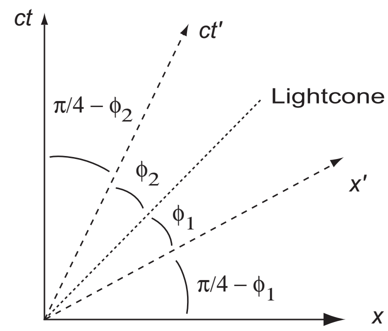
$$\Lambda^{\mu}{}_{\nu}\Lambda_{\mu}{}^{\lambda} = \Lambda^{\mu}{}_{\nu}\eta_{\mu\beta}\eta^{\lambda\alpha}\Lambda^{\beta}{}_{\alpha} = \Lambda^{\mu}{}_{\nu}\Lambda^{\beta}{}_{\alpha}\eta_{\mu\beta}\eta^{\lambda\alpha} = \eta_{\nu\alpha}\eta^{\lambda\alpha} = \delta^{\lambda}_{\nu},$$

implying that $\Lambda^{\mu}{}_{\nu}$ and $\Lambda_{\mu}{}^{\nu}$ are matrix inverses of each other.

4.10 From Fig. 4.6(a), if the unprimed coordinate system is moving to the right with velocity v along the x axis (with $\tan\varphi = v/c$), the primed axes plotted in the unprimed system are sketched in the left figure below. Then from the primed system the unprimed system moves to the left with velocity $-v$ along the x axis and the unprimed axes plotted in the primed coordinate system are given by the right figure below.



4.11 In the following diagram the unprimed axes are orthogonal, so their scalar product $x \cdot t = 0$. But if $\varphi_1 = \varphi_2$, the primed axes represent a Lorentz transform from the unprimed axes [see Fig. 4.6(b)]. Since the scalar product is invariant under Lorentz transformation, the primed axes also have vanishing scalar product and thus are orthogonal.



For the special case that $\varphi_1 = \varphi_2 = \pi/4$ in the above diagram, the two vectors become equivalent and lie on the lightcone. But by the above argument they must also be orthogonal to each other. Thus, a lightlike vector is orthogonal to itself.

4.12 Since the mass is macroscopic and the force is small, Newtonian mechanics describes the acceleration, but raising the temperature by ΔT adds an energy $\Delta E \sim k\Delta T$, where k is Boltzmann's constant. Thus, for equivalent forces of magnitude F the ratio of acceleration magnitudes is

$$\frac{a_1}{a_2} = \frac{F/m_1}{F/m_2} = \frac{m_2}{m_1} = \frac{m_1 + k\Delta T}{m_1} = 1 + \frac{k\Delta T}{m_1},$$

where mass 2 is assumed to be the hotter one. The effect is of course tiny in magnitude, because $k\Delta T/m_1$ is small.

4.13 In frame S the interval is $(\Delta s)^2 = (\Delta x)^2 - (c\Delta t)^2 = -c^2(\Delta t)^2$, since $\Delta x = 0$. But the interval is invariant under Lorentz transformations so in the S' frame

$$-c^2(\Delta t)^2 = (\Delta x')^2 - (c\Delta t')^2,$$

which may be solved to give $\Delta x' = c\left((\Delta t')^2 - (\Delta t)^2\right)^{1/2}$.

4.14 Since ds^2 is an invariant, in any inertial frame Eq. (4.5) gives the measured time interval between the two events as

$$c^2 dt^2 = -ds^2 + dx^2 + dy^2 + dz^2,$$

where $-ds^2$ is positive, since the interval is timelike. By definition, in the inertial frame where the spatial separation between two events is zero ($dx^2 + dy^2 + dz^2 = 0$), the measured time is the proper time,

$$dt^2 = -\frac{ds^2}{c^2} \equiv d\tau^2.$$

(Since the separation is timelike, there always is an inertial frame where the spatial separation between events is zero; see Fig. 4.7(b).) Thus, since $dx^2 + dy^2 + dz^2$ can never be negative, in any other inertial frame the measured time interval between the events will be larger than the proper time: for events with timelike separation, the minimum time interval that can be measured by any inertial observer is the proper time.

4.15 (a) From Eq. (4.50) the field tensor is defined by $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Under the gauge transformation $A^\mu \to A^\mu - \partial^\mu \chi$ given by (4.46), this transforms as

$$\begin{aligned} F'^{\mu\nu} &= \partial^\mu (A^\nu - \partial^\nu \chi) - \partial^\nu (A^\mu - \partial^\mu \chi) \\ &= \partial^\mu A^\nu - \partial^\nu A^\mu = F^{\mu\nu}, \end{aligned}$$

where $\partial^\mu \partial^\nu = \partial^\nu \partial^\mu$ was used. Thus $F^{\mu\nu}$ is gauge invariant. It is also Lorentz invariant, since it is by explicit construction a rank-2 Lorentz tensor [Eq. (4.50) is a tensor equation].

(b) Appeal to Eqs. (4.33) to construct the components of the electric and magnetic fields in terms of the potentials. For example, writing some components of Eq. (4.33) out explicitly,

$$E^1 = \partial^1 A^0 - \partial^0 A^1 = F^{10} = -F^{01} \qquad B^2 = \partial^1 A^3 - \partial^3 A^1 = F^{13} = -F^{31}.$$

Proceeding in this manner, one finds that the six independent components of $\boldsymbol{E}$ and $\boldsymbol{B}$ are elements of the antisymmetric rank-2 *electromagnetic field tensor*

$$F^{\mu\nu} = -F^{\nu\mu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$$

given by Eqs. (4.50) and (4.51).

(c) The equivalence may be established by multiplying the terms of Eqs. (4.53) and (4.54) out explicitly using Eqs. (4.51), (4.52), and (4.43). For example, setting $\nu = 0$ in Eq. (4.53) and using Eq. (4.51) gives

$$\partial_1 E^1 + \partial_2 E^2 + \partial_3 E^3 = j^0,$$

which is equivalent to Eq. (4.28).

5 Lorentz-invariant Dynamics

5.1 The variational principle $\delta \int ds = 0$ with line element $ds^2 = \eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu d\tau^2$ (where $\dot{x}^\mu \equiv dx^\mu/d\tau$) implies the Euler–Lagrange equation (5.18),

$$-\frac{d}{d\tau}\left(\frac{\partial L}{\partial(dx^\mu/d\tau)}\right) + \frac{\partial L}{\partial x^\mu} = 0 \qquad L \equiv \sqrt{-\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}.$$

But $\partial L/\partial x^\mu = 0$ and the preceding equation becomes

$$\frac{d}{d\tau}\left[\frac{\partial}{\partial \dot{x}^\mu}(-\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu)^{1/2}\right] = 0.$$

Using that $\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$ is independent of time, this reduces to $\eta_{\mu\nu}\frac{d}{d\tau}\dot{x}^\nu = \eta_{\mu\nu}\ddot{x}^\nu = 0$, where $\ddot{x}^\nu \equiv d^2x^\nu/d\tau^2$. In matrix form this is

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \ddot{t} \\ \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = 0,$$

implying that trajectories obey

$$\frac{d^2t}{d\tau^2} = 0 \qquad \frac{d^2x}{d\tau^2} = 0 \qquad \frac{d^2y}{d\tau^2} = 0 \qquad \frac{d^2z}{d\tau^2} = 0.$$

This corresponds to straight lines in Minkowski space.

5.2 If the inertial frame of the particle is chosen it is at rest and its 4-momentum is $p = (m, 0, 0, 0)$. Assuming motion of the observer along the x axis, the 4-velocity of the observer is $u = (\gamma, \gamma\boldsymbol{v}) = (\gamma, v\gamma, 0, 0) \equiv e_{\hat{0}}$. From Eq. (5.25) the observed energy is then

$$E \equiv -p \cdot u = -p \cdot e_0 = \gamma m.$$

This should not be surprising, since if the laboratory were at rest and the particle moving this would be the expected result in special relativity.

5.3 (a) Since $c = 1 = 3 \times 10^8\,\mathrm{m\,s^{-1}}$, one has $1\,\mathrm{s} = 3 \times 10^8\,\mathrm{m}$. Therefore,

$$1\,\mathrm{J} = 1\,\mathrm{kg}\,\frac{1\,\mathrm{m}^2}{(3\times 10^8\,\mathrm{m})^2} = 1.1 \times 10^{-17}\,\mathrm{kg}.$$

(b) Using $1\,\mathrm{s} = 3 \times 10^8\,\mathrm{m}$ again,

$$1\,\mathrm{atm} = 10^5\,\mathrm{kg}\,\frac{1\,\mathrm{m}^{-1}}{(3\times 10^8\,\mathrm{m})^2} = 1.1 \times 10^{-12}\,\mathrm{kg\,m^{-3}}.$$

5.4 (a) The SI unit of acceleration is $\mathrm{m\,s^{-2}}$. The required conversion factor is c^2 since

$$1\,\mathrm{m}^{-1} \times c^2 = (1\,\mathrm{m}^{-1})(3\times 10^8\,\mathrm{m\,s^{-1}})^2 = 9\times 10^{16}\,\mathrm{m\,s^{-2}}.$$

(b) The energy density has units of $\mathrm{J\,m^{-3}}$ in the SI system. The required conversion factor is c^2:

$$(2\,\mathrm{kg\,m^{-3}})(3\times 10^8\,\mathrm{m\,s^{-1}})^2 = 1.8\times 10^{17}\,\mathrm{J\,m^{-3}}.$$

5.5 If L does not depend on x^1 the Euler–Lagrange equation (5.18) implies that

$$-\frac{d}{d\sigma}\left(\frac{\partial L}{\partial(dx^1/d\sigma)}\right) = 0,$$

where

$$\frac{\partial L}{\partial(dx^1/d\sigma)} = -\frac{1}{2L}\left(g_{1\nu}(x)\frac{dx^\nu}{d\sigma} + g_{\mu 1}(x)\frac{dx^\mu}{d\sigma}\right).$$

But $g_{\mu\nu}$ is symmetric and μ and ν are dummy summation indices so the terms can be combined to give

$$\frac{\partial L}{\partial(dx^1/d\sigma)} = -\frac{1}{L}g_{1\mu}\frac{dx^\mu}{d\sigma}.$$

Then using $Ld\sigma = d\tau$ gives

$$\frac{\partial L}{\partial(dx^1/d\sigma)} = -g_{\alpha\mu}K^\alpha u^\mu = -K\cdot u,$$

where $K^\alpha = (0,1,0,0)$ is the Killing vector associated with the absence of x^1 dependence in the metric. Insertion of this result in the first equation above gives Eq. (5.29).

5.6 For low velocity $E = m\gamma = m(1-\boldsymbol{v}^2)^{-1/2} \simeq m(1+\frac{1}{2}\boldsymbol{v}^2)$. Restoring to normal units by $m \to mc^2$ and $\boldsymbol{v} \to \boldsymbol{v}/c$ gives $E \simeq mc^2 + \frac{1}{2}mv^2$, which is a sum of rest-mass and kinetic energies.

5.7 (a) $1\,\mathrm{dyne\,cm^{-2}} = 1\,\mathrm{g\,cm^{-1}\,s^{-2}}$. But from Box 5.1, $1\,\mathrm{g} = 7.4237\times 10^{-29}\,\mathrm{cm}$ and $1\,\mathrm{s} = 2.9979\times 10^{10}\,\mathrm{cm}$. Inserting these gives $1\,\mathrm{dyne\,cm^{-2}} = 8.26\times 10^{-50}\,\mathrm{cm^{-2}}$.

(b) $1\,\mathrm{MeV} = 1.6022\times 10^{-6}\,\mathrm{erg}$. Since $k = 8.617\times 10^{-5}\,\mathrm{eV\,K^{-1}}$, in $k=1$ units a kelvin degree is

$$K = 8.617\times 10^{-5}\,\mathrm{eV} = 1.3807\times 10^{-16}\,\mathrm{erg}.$$

Also, from Box 5.1, $1\,\mathrm{erg} = 8.2601\times 10^{-50}\,\mathrm{cm}$. Utilizing these conversions, it may be deduced that

$$1\,\mathrm{MeV} = 1.602\times 10^{-6}\,\mathrm{erg} = 1.160\times 10^{10}\,\mathrm{K} = 1.323\times 10^{-55}\,\mathrm{cm} = 1.783\times 10^{-27}\,\mathrm{g}.$$

where results from part (a) were used.

5.8 *By dimensional analysis:* For period P, total mass M, and reduced mass μ, the luminosity L in geometrized units is

$$L = \frac{128}{5}4^{2/3}M^{4/3}\mu^2\left(\frac{\pi}{P}\right)^{10/3},$$

Let the standard unit of length be $\mathscr{L}$, the standard unit of mass be $\mathscr{M}$, and the standard unit of time be $\mathscr{T}$. In standard units L has dimension $\mathscr{M}\mathscr{L}^2\mathscr{T}^{-3}$ (e.g., $\mathrm{erg\,s^{-1}}$). Since μ and M have units of $\mathscr{M}$ and P has units of $\mathscr{T}$, the right side of the equation for L must be multiplied by a factor of powers of G and c having the units of $\mathscr{M}^{-7/3}\mathscr{T}^{1/3}\mathscr{L}^2$ to be dimensionally correct. Because of the mass dependence of the equation G can enter only as the $7/3$ power so the required overall factor is $G^{7/3}/c^5$ and in standard units

$$L = \frac{128}{5}4^{2/3}\frac{G^{7/3}}{c^5}M^{4/3}\mu^2\left(\frac{\pi}{P}\right)^{10/3}.$$

Inserting G in units of $M_\odot{}^{-1}\mathrm{cm^3 s^{-2}}$ and c in units of $\mathrm{cm\,s^{-1}}$, the factor $G^{7/3}/c^5$ can be evaluated and units apportioned conveniently as

$$L = 2.3\times 10^{45}\left(\frac{M}{M_\odot}\right)^{4/3}\left(\frac{\mu}{M_\odot}\right)^{2}\left(\frac{1\,\mathrm{s}}{P}\right)^{10/3}\ \mathrm{erg\ s^{-1}}.$$

More automatically: Table B.1 indicates that conversion from geometrized units to standard units requires the replacements $M\to GM/c^2$, $\mu\to G\mu/c^2$, $P\to cP$, and $L\to (G/c^5)L$. Making these replacements and rearranging gives the same equations as above.

5.9 As for the derivation in Box 5.2, extremizing the classical action is equivalent to solving the corresponding Euler-Lagrange equation (5.18). Assuming a Lagrangian $L(x,\dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x)$,

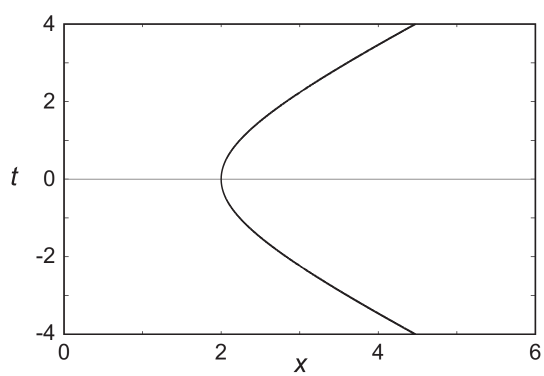
$$\frac{\partial L}{\partial \dot{x}} = m\dot{x} \qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x} \qquad \frac{\partial L}{\partial x} = -\frac{dV}{dx}.$$

Inserting this into the Euler–Lagrange equation (5.18) gives $m\ddot{x} = -\dfrac{dV}{dx}$, which is Newton's second law.

5.10 From the parameterized equations

$$x^2 - t^2 = a^2(\cosh^2\sigma - \sinh^2\sigma) = a^2,$$

so the equation is an hyperbola with $t = \pm(x^2-a^2)^{1/2}$, which has the following plot for $a = 2$.



The motion clearly is accelerating because of the curvature. From the Minkowski metric $d\tau^2 = -ds^2 = dt^2 - dx^2$. Evaluating $dt/d\sigma$ and $dx/d\sigma$ from the original parameterization gives

$$d\tau^2 = (ad\sigma)^2(\cosh^2\sigma - \sinh^2\sigma) = a^2 d\sigma^2.$$

Integrating and choosing the integration constant so that $\tau = 0$ when $\sigma = 0$ gives $\tau = a\sigma$, and thus

$$x(\tau) = a\cosh\left(\frac{\tau}{a}\right) \qquad t(\tau) = a\sinh\left(\frac{\tau}{a}\right).$$

The 4-velocity is given by

$$u = \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}\right) = \left(\cosh\left(\frac{\tau}{a}\right), \sinh\left(\frac{\tau}{a}\right)\right).$$

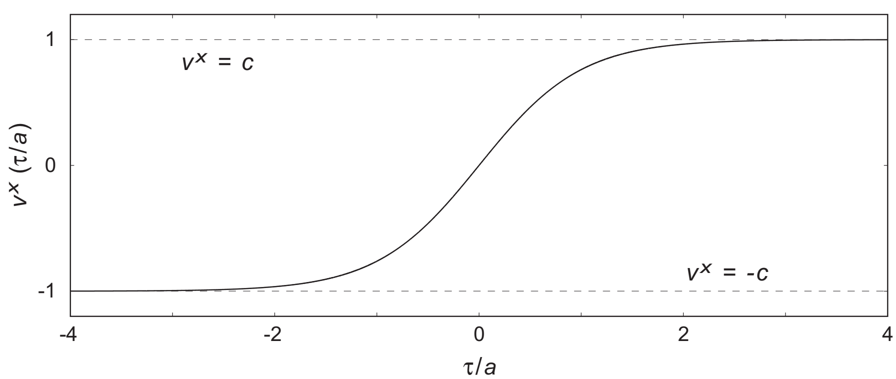
From the Minkowski metric with $g = \mathrm{diag}\,(-1,1)$,

$$u \cdot u = g_{ij}u^i u^j = g_{00}u^0u^0 + g_{11}u^1u^1 = \sinh^2\left(\frac{\tau}{a}\right) - \cosh^2\left(\frac{\tau}{a}\right) = -1,$$

which is the normalization given by Eq. (5.6). The 3-velocity is

$$v^x = \frac{dx}{dt} = \frac{dx/d\tau}{dt/d\tau} = \frac{\sinh(\tau/a)}{\cosh(\tau/a)} = \tanh\left(\frac{\tau}{a}\right),$$

which has the behavior



In these units $v^x = 1$ is the speed of light. Thus the 3-velocity is bounded asymptotically by c.

5.11 Substituting $x^\mu \to x'^\mu = x^\mu + \varepsilon K^\mu$ into

$$g_{\mu\nu}(x) = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\, g_{\alpha\beta}(x')$$

(in the following terms of order ε^2 are discarded) leads to

$$\begin{aligned} g_{\mu\nu}(x) &= \frac{\partial}{\partial x^\mu}(x^\alpha + \varepsilon K^\alpha)\frac{\partial}{\partial x^\nu}\left(x^\beta + \varepsilon K^\beta\right) g_{\alpha\beta}(x^\gamma + \varepsilon K^\gamma) \\ &= \left(\delta^\alpha_\mu + \varepsilon\partial_\mu K^\alpha\right)\left(\delta^\beta_\nu + \varepsilon\partial_\nu K^\beta\right) g_{\alpha\beta}(x'). \end{aligned}$$

Now expand $g_{\alpha\beta}(x')$ in a Taylor series around $g_{\alpha\beta}(x)$,

$$g_{\alpha\beta}(x') = g_{\alpha\beta}(x) + \left.\frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\right|_x \Delta x^\gamma = g_{\alpha\beta}(x) + \varepsilon K^\gamma \partial_\gamma g_{\alpha\beta}(x) + \mathscr{O}\left(\varepsilon^2\right),$$

where $\Delta x^\gamma = \varepsilon K^\gamma$. Combining the preceding two equations gives

$$\begin{aligned} g_{\mu\nu}(x) &= \left(\delta^\alpha_\mu + \varepsilon \partial_\mu K^\alpha\right)\left(\delta^\beta_\nu + \varepsilon \partial_\nu K^\beta\right)\left(g_{\alpha\beta}(x) + \varepsilon K^\gamma \partial_\gamma g_{\alpha\beta}(x)\right) \\ &= g_{\mu\nu}(x) + \varepsilon\left[K^\gamma \partial_\gamma g_{\mu\nu}(x) + \partial_\mu K^\beta g_{\beta\nu} + \partial_\nu K^\beta g_{\mu\beta} + \mathscr{O}\left(\varepsilon^2\right)\right] \end{aligned}$$

Subtracting $g_{\mu\nu}(x)$ from both sides and noting that ε is arbitrary, this can hold generally only if the quantity in square brackets vanishes. This leads to

$$g_{\mu\beta}\partial_\nu K^\beta + g_{\nu\beta}\partial_\mu K^\beta + K^\gamma \partial_\gamma g_{\mu\nu} = 0,$$

or equivalently,

$$\partial_\nu K_\mu + \partial_\mu K_\nu + K^\gamma \partial_\gamma g_{\mu\nu} = 0,$$

where contraction with the metric tensor was used to lower indices.

5.12 From Problem 5.11 or Box 5.3

$$\partial_\nu K_\mu + \partial_\mu K_\nu + K^\gamma \partial_\gamma g_{\mu\nu} = 0.$$

But from the formula for the Lie derivative given in Eq. (3.74)

$$\begin{aligned} \mathscr{L}_K g_{\mu\nu} &= K^\alpha \partial_\alpha g_{\mu\nu} + g_{\mu\alpha}\partial_\nu K^\alpha + g_{\nu\alpha}\partial_\mu K^\alpha \\ &= \partial_\nu K_\mu + \partial_\mu K_\nu + K^\gamma \partial_\gamma g_{\mu\nu}. \end{aligned}$$

For the Lie derivative in a metric space, partial and covariant derivative operations are interchangeable (see Section 3.13.5), so

$$\begin{aligned} \mathscr{L}_K g_{\mu\nu} &= \partial_\nu K_\mu + \partial_\mu K_\nu + K^\gamma \partial_\gamma g_{\mu\nu} \\ &= \nabla_\nu K_\mu + \nabla_\mu K_\nu + K^\gamma \nabla_\gamma g_{\mu\nu} = 0, \end{aligned}$$

and since $\nabla_\gamma g_{\mu\nu} = 0$ from Eq. (3.63),

$$\nabla_\nu K_\mu + \nabla_\mu K_\nu = \partial_\nu K_\mu + \partial_\mu K_\nu = 0,$$

which is Killing's equation.

5.13 Choose an inertial frame in which the star at rest and assume the emitted light to be monochromatic with frequency ω_0. The wavevector for the photon is $k^\mu = (\omega_0, \omega_0, 0, 0)$. For photons $E = \hbar\omega$, so from Eq. (5.25)

$$E = \hbar\omega = -p \cdot u = -\hbar k \cdot u,$$

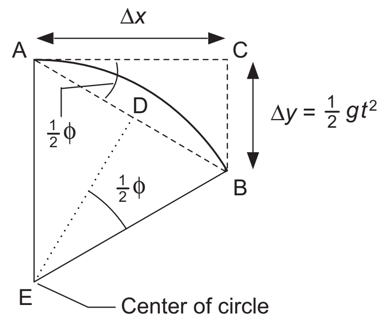
implying that $\omega = -k \cdot u$, where u is the velocity of the observer in the inertial frame of the star. Writing this scalar product out explicitly gives

$$\begin{aligned} \omega &= -k \cdot u = -\eta_{\mu\nu} k^\mu u^\nu = k^0 u^0 - k^1 u^1 \\ &= \omega_0\left(\cosh(\tau/a) - \sinh(\tau/a)\right) = \omega_0 e^{-\tau/a}, \end{aligned}$$

where τ is the proper time and $u = (\cosh(\tau/a), \sinh(\tau/a))$ from the solution of Problem 5.10 was used.

6 The Principle of Equivalence

6.1 Consider the following figure, where the arc approximates the path of the photon.



The flight time of the photon is $t \simeq \Delta x/c$ for a small arc. In the time t the elevator falls a distance $\Delta y = \frac{1}{2}gt^2$, where g is the gravitational acceleration. From the geometry, triangles EBD and ABC contain the same angles, so $\overline{BC}\,/\,\overline{AC} = \overline{DB}\,/\,\overline{DE}$. But from the figure one obtains

$$r_c \simeq \overline{DE} \qquad \overline{AC} = \Delta x \qquad \overline{BC} = \Delta y = \tfrac{1}{2}gt^2 \qquad \overline{DB} = \frac{\Delta x}{2\cos(\varphi/2)},$$

where r_c is the curvature radius and the last relation follows because $\overline{AD} = \overline{DB}$, implying that $\frac{1}{2}\Delta x/\overline{DB} = \cos(\varphi/2)$. Therefore,

$$\frac{\frac{1}{2}gt^2}{\Delta x} = \frac{\Delta x/[2\cos(\varphi/2)]}{r_c}.$$

Assuming a small deflection justifies the approximation $\cos(\varphi/2) \simeq 1$ and

$$r_c = \frac{c^2}{g} = \frac{c^2R^2}{GM},$$

where in the rightmost expression g has been evaluated at the surface of a sphere of mass M and radius R. Quantities calculated for the Earth, a white dwarf, and a neutron star are displayed in the following table.

Object	R(km)	M(kg)	ρ(g cm^{-3})	g(m s^{-2})	r_c(km)	R/r_c
Earth	6378	6×10^{24}	5.6	9.8	9.2×10^{12}	6.9×10^{-10}
White dwarf	5500	2.1×10^{30}	$\sim 10^6$	4.6×10^6	1.9×10^7	2.8×10^{-4}
Neutron star	10	2×10^{30}	$\sim 10^{14}$	1.3×10^{12}	67.5	0.15

The corresponding vertical deflection of the light is $\Delta y = \frac{1}{2}gt^2$, where t is the flight time for the light. For an elevator width of 2 meters, $t = 6.67 \times 10^{-9}$ seconds and the vertical deflection Δy is 2.2×10^{-16} m for Earth, 1.0×10^{-10} m for the white dwarf, and 3×10^{-5} m for the neutron star.

6.2 The particle created at z_2 has mass $m = h\nu/c^2$, where h is Planck's constant and ν is the frequency of the photon. Upon dropping to z_1 in the gravitational field, the energy is $mc^2 + mg(z_2 - z_1)$. Thus, the system creates spontaneously an energy $mg(z_2 - z_1)$ in each cycle, unless the photon loses an energy $h\nu g(z_2 - z_1)/c^2$ in moving from z_1 to z_2.

6.3 Apply Kepler's laws to the approximately circular orbit of period 12 hours, giving $r \simeq 2.7 \times 10^7$ m and $v \simeq 3.9\,\mathrm{km\,s^{-1}}$. Defining $\beta = v/c$, the special relativistic time dilation factor for the satellite is $\gamma = (1-\beta^2)^{-1/2} \simeq 1 + \frac{1}{2}\beta^2$, where the small effect of Earth's rotation has been neglected. The fractional change in frequency is determined by the second term,

$$\frac{\nu_\mathrm{s} - \nu_0}{\nu_0} = -\tfrac{1}{2}\beta^2 = -8.5 \times 10^{-11},$$

where the negative sign is because the time is dilated ($\nu_\mathrm{s} < \nu_0$) for the satellite viewed from Earth. For the general relativistic time dilation, integrating Eq. (6.7) gives

$$\int_{\nu_0}^{\nu_\mathrm{s}} \frac{d\nu}{\nu} = \int_R^{r_\mathrm{s}} \frac{GM}{r^2c^2} dr.$$

Evaluating the integrals on both sides yields (see the solution of Problem 6.7)

$$\frac{\nu_\mathrm{s}}{\nu_0} = \exp\left[-\frac{GM}{c^2}\left(\frac{1}{r_\mathrm{s}} - \frac{1}{R}\right)\right] \simeq 1 - \frac{GM}{c^2}\left(\frac{1}{r_\mathrm{s}} - \frac{1}{R}\right).$$

Solving this for the fractional shift in frequency gives

$$\frac{\nu_\mathrm{s} - \nu_0}{\nu_0} = \frac{GM}{c^2}\left(\frac{1}{R} - \frac{1}{r_\mathrm{s}}\right) = 5.3 \times 10^{-10}.$$

This is opposite in sign relative to the special relativistic effect and about six times larger. Thus, for every second of elapsed time

1. Special relativistic time dilation slows the satellite clock relative to the ground clock by about $8.5 \times 10^{-11} \times 1$ second $= 0.085$ ns.
2. Gravitational time dilation (general relativity) slows the ground clock relative to the satellite clock by about $5.3 \times 10^{-10} \times 1$ second $= 0.53$ ns.

The net effect is that for every second the satellite clock gains about $0.53 - 0.085 = 0.445$ ns relative to the ground clock because of relativistic corrections. Suppose that an accuracy of two meters is desired from the GPS system for locations on the ground. Light takes 6.7 ns to travel two meters. Thus, without the above corrections for special and general relativistic time dilation an error in timing that begins to compromise two-meter resolution will have accumulated after about 15 seconds.

6.4 From the gravitational redshift, $\Delta\nu = g\nu d/c^2$ for motion over a vertical distance d. The corresponding loss in energy for the redshifted light is $\Delta E = h\Delta\nu = -hg\nu d/c^2$. But a particle of mass m would lose an energy $\Delta E = -gmd$ in the same circumstances. Comparing the two expressions, one sees that the photon loses energy as if it had an effective mass $m = h\nu/c^2$ in the gravitational field. Of course a photon doesn't have a mass, but it behaves in some respects as if it did. Note that a similar argument is made in Problem 6.2 without assuming photons to be massive.

6.5 For a difference in height h, the change in the length of the time intervals is approximated by $\Delta\tau_1 = \Delta\tau_2(1 - gh/c^2)$, between points 1 and 2. Therefore, the fractional difference is

$$\frac{\Delta\tau_2 - \Delta\tau_1}{\Delta\tau_2} = \frac{gh}{c^2} = 1.77 \times 10^{-13}.$$

This implies a difference of 5.6×10^{-6} seconds per year between the two clocks, with the one at higher elevation running faster.

6.6 (a) For a pendulum with a string of length ℓ and negligible mass,

$$\ell m_{\text{inertial}} \frac{d^2\theta}{dt^2} = -m_{\text{grav}} g \sin\theta.$$

For small oscillations $\sin\theta \sim \theta$ and the solution is a harmonic oscillator with period

$$P = 2\pi\sqrt{\frac{\ell}{g}\left(\frac{m_{\text{inertial}}}{m_{\text{grav}}}\right)}.$$

Thus, if the gravitational and inertial masses are not equivalent, identical pendulums made from different materials having different ratios $m_{\text{inertial}}/m_{\text{grav}}$ should have different periods.

(b) For a block sliding along an inclined plane that makes an angle θ with the horizontal, the component of gravitational force along the inclined plane is $m_{\text{grav}} g \sin\theta$. Thus,

$$m_{\text{inertial}} a = m_{\text{grav}} g \sin\theta \quad \longrightarrow \quad a = g \sin\theta \frac{m_{\text{grav}}}{m_{\text{inertial}}}.$$

Therefore, if $m_{\text{grav}} \neq m_{\text{inertial}}$ one expects to observe material-dependent accelerations in inclined-plane experiments.

6.7 Integrating the gravitational time dilation formula (6.7) gives

$$\int_{\nu_0}^{\nu_{\text{s}}} \frac{d\nu}{\nu} = \int_R^{r_{\text{s}}} \frac{GM}{r^2 c^2}\, dr.$$

Evaluating the two integrals gives

$$\ln\nu_{\text{s}} - \ln\nu_0 = \ln(\nu_{\text{s}}/\nu_0) = \frac{GM}{c^2}\left(\frac{1}{R} - \frac{1}{r_{\text{s}}}\right)$$

and if both sides are exponentiated

$$\frac{\nu_{\text{s}}}{\nu_0} = \exp\left[\frac{GM}{c^2}\left(\frac{1}{R} - \frac{1}{r_{\text{s}}}\right)\right] \simeq 1 + \frac{GM}{c^2}\left(\frac{1}{R} - \frac{1}{r_{\text{s}}}\right).$$

Therefore,

$$\frac{\nu_{\text{s}}}{\nu_0} - 1 = \frac{\nu_{\text{s}} - \nu_0}{\nu_0} = \frac{GM}{c^2}\left(\frac{1}{R} - \frac{1}{r_{\text{s}}}\right).$$

6.8 (a) From Eq. (6.7),

$$\frac{\Delta \nu}{\nu} = \frac{gh}{c^2} = 2.45 \times 10^{-15},$$

where a height of $h = 22.5\,\mathrm{m}$ was used.

(b) To compensate for the gravitational redshift a blue Doppler shift $\Delta \nu_{\mathrm{Doppler}} = v/c$ is required. Setting

$$\frac{gh}{c^2} = \frac{v}{c}$$

and solving for v gives $v = gh/c = 7.35 \times 10^{-5}\,\mathrm{cm\,s^{-1}}$. Notice that this is just the blueshift invoked in Section 6.5.1 to compensate for the gravitational redshift in the falling-elevator thought experiment.

6.9 Consider two arbitrary masses 1 and 2 at points A and B, respectively, attracting each other gravitationally. Taking A as the origin of the coordinate system, the forces acting on the two masses are

$$\boldsymbol{F}_1 = G\frac{m_{\mathrm{p}}(1)m_{\mathrm{a}}(2)}{r^2}\,\hat{\boldsymbol{r}} \qquad \boldsymbol{F}_2 = -G\frac{m_{\mathrm{p}}(2)m_{\mathrm{a}}(1)}{r^2}\,\hat{\boldsymbol{r}},$$

where subscripts a and p denote active and passive masses, respectively. But by Newton's 3rd law $\boldsymbol{F}_1 = -\boldsymbol{F}_2$ and

$$\frac{m_{\mathrm{p}}(1)}{m_{\mathrm{a}}(1)} = \frac{m_{\mathrm{p}}(2)}{m_{\mathrm{a}}(2)} \equiv C,$$

where C must be a universal constant for all gravitational masses, since the two masses were arbitrary. Choosing units appropriately so that $C = 1$ gives $m_{\mathrm{p}} = m_{\mathrm{a}} \equiv m_{\mathrm{g}}$.

6.10 From Eq. (6.12) the gravitational redshift at the surface of a spherical gravitating body of radius R and mass M is

$$z = 2.12 \times 10^{-6}\left(\frac{M}{M_\odot}\right)\left(\frac{R_\odot}{R}\right).$$

The gravitational shift in spectral lines for the Sun is then only $\Delta\lambda/\lambda \sim 2 \times 10^{-6}$. If the gravitational redshift z is parameterized as a velocity giving rise to an equivalent Doppler shift through $v = cz$,

$$v = 0.636\left(\frac{M}{M_\odot}\right)\left(\frac{R_\odot}{R}\right)\,\mathrm{km\,s^{-1}}.$$

For the Sun this gives only $\sim 0.6\,\mathrm{km\,s^{-1}}$, which must be disentangled from much larger kinematic Doppler shifts caused by motion of the Sun relative to the Earth and motion of gas in the solar surface. The situation is similar for other main sequence stars since the gravitational redshift is determined to lowest order by M/R, which is the same to within a factor of 2–3 for main sequence stars.

7 Curved Spacetime and General Covariance

7.1 This problem is patterned after an example in Cheng [64].

(a) If the space can be parameterized so that the metric is globally independent of the coordinates, all derivatives in Eq. (7.2) vanish and $K = 0$.

(b) Taking plane polar coordinates $(x^1, x^2) = (r, \theta)$ gives the position-dependent metric

$$g_{11} = 1 \qquad g_{22} = r^2,$$

but insertion into Eq. (7.2) again gives $K = 0$.

(c) Choosing the spherical coordinates $(x^1, x^2) = (S, \varphi)$ defined in Fig. 7.1 gives

$$g_{11} = 1 \qquad g_{22} = R^2 \sin^2\left(\frac{x^1}{R}\right).$$

Insertion into Eq. (7.2) gives $K = R^{-2}$, which is constant.

(d) Choosing the cylindrical coordinates $(x^1, x^2) = (r, \varphi)$ defined in Fig. 7.1 gives

$$g_{11} = \frac{R^2}{R^2 - r^2} \qquad g_{22} = r^2,$$

and insertion into Eq. (7.2) again yields $K = R^{-2}$.

7.2 Since $g_{\mu\nu}$ is a rank-2 covariant tensor its transformation law is given by Eq. (3.56) as

$$g_{\mu\nu;\lambda} = g_{\mu\nu,\lambda} - \Gamma^{\alpha}_{\mu\lambda} g_{\alpha\nu} - \Gamma^{\alpha}_{\nu\lambda} g_{\mu\alpha}.$$

In a local inertial frame space is locally flat and $g_{\mu\nu,\lambda}$ vanishes. Likewise, since from Eq. (7.30) the affine connection is proportional to the derivative of the metric tensor, it also vanishes in the local inertial frame. Thus $g_{\mu\nu;\lambda} = 0$ for a local inertial frame. But this is a tensor equation, so it is valid in all reference frames. Thus the covariant derivative of the metric tensor vanishes in any reference frame. This may be verified by direct computation. For example, it follows from Eqs. (7.30) and (3.56) and some algebra. More elegantly, substituting Eq. (7.29) into Eq. (3.56) gives

$$\begin{aligned}
g_{\mu\nu;\lambda} &= g_{\mu\nu,\lambda} - \Gamma^{\alpha}_{\mu\lambda} g_{\alpha\nu} - \Gamma^{\alpha}_{\nu\lambda} g_{\mu\alpha} \\
&= \Gamma^{\rho}_{\lambda\mu} g_{\rho\nu} + \Gamma^{\rho}_{\lambda\nu} g_{\rho\mu} - \Gamma^{\alpha}_{\mu\lambda} g_{\alpha\nu} - \Gamma^{\alpha}_{\nu\lambda} g_{\mu\alpha} \\
&= \Gamma^{\alpha}_{\lambda\mu} g_{\alpha\nu} + \Gamma^{\alpha}_{\lambda\nu} g_{\alpha\mu} - \Gamma^{\alpha}_{\lambda\mu} g_{\alpha\nu} - \Gamma^{\alpha}_{\lambda\nu} g_{\alpha\mu} = 0
\end{aligned}$$

where the last step used that ρ and α are dummy indices and that $g_{\mu\nu}$ and $\Gamma^{\alpha}_{\lambda\mu}$ are symmetric in their lower indices.

7.3 In cylindrical polar coordinates (r, θ, z) the line element is $ds^2 = dr^2 + r^2 d\theta^2 + dz^2$, corresponding to metric components $g_{ij} = \mathrm{diag}\,(1, r^2, 1)$ and $g^{ij} = g_{ij}^{-1} = \mathrm{diag}\,(1, r^{-2}, 1)$, and obviously $\det g = r^2$. The geodesic equations are given by Eq. (7.21)

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0,$$

where the connection coefficients can be determined using Eq. (7.30),

$$\Gamma^\lambda_{\mu\nu} = \tfrac{1}{2}g^{\alpha\lambda}\left(\frac{\partial g_{\nu\alpha}}{\partial x^\mu} + \frac{\partial g_{\mu\alpha}}{\partial x^\nu} - \frac{\partial g_{\nu\mu}}{\partial x^\alpha}\right).$$

Because the metric is diagonal and only one entry is not constant, there are only three non-vanishing connection coefficients:

$$\Gamma^r_{\theta\theta} \equiv \Gamma^1_{22} = \frac{1}{2}g^{11}\left(-\frac{\partial g_{22}}{\partial r}\right) = -r \qquad \Gamma^\theta_{r\theta} \equiv \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2}g^{22}\left(\frac{\partial g_{22}}{\partial r}\right) = \frac{1}{r}.$$

Inserting these in Eq. (7.21) gives three equations of geodesic motion,

$$\frac{d^2r}{d\sigma^2} - r\left(\frac{d\theta}{d\sigma}\right)^2 = 0 \qquad \frac{d^2\theta}{d\sigma^2} + \frac{2}{r}\frac{dr}{d\sigma}\frac{d\theta}{d\sigma} = 0 \qquad \frac{d^2z}{d\sigma^2} = 0,$$

where σ parameterizes the position on a path.

7.4 For 2-dimensional polar coordinates (r, θ) the line element is $ds^2 = dr^2 + r^2 d\theta^2$, corresponding to a metric $g_{ij} = \mathrm{diag}\,(1, r^2)$. From Eq. (5.16) without the minus sign the Lagrangian is

$$L = \left[\left(\frac{dr}{d\sigma}\right)^2 + r^2\left(\frac{d\theta}{d\sigma}\right)^2,\right]^{1/2}$$

and from Eq. (5.18) the equations of motion are

$$\frac{d^2r}{d\tau^2} = r\left(\frac{d\theta}{d\tau}\right)^2 \qquad \frac{d}{d\tau}\left(r^2\frac{d\theta}{d\tau}\right) = \frac{d^2\theta}{d\tau^2} + \frac{2}{r}\frac{dr}{d\tau}\frac{d\theta}{d\tau} = 0,$$

where $\tau = \int_0^1 d\sigma L$, with σ parameterizing a path in the space. For the current 2-dimensional case the geodesic equation (7.21) reduces to

$$\frac{d^2r}{d\tau^2} = -\Gamma^0_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau} \qquad \frac{d^2\theta}{d\tau^2} = -\Gamma^1_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau}$$

Thus, by comparing

$$\frac{d^2r}{d\tau^2} = r\left(\frac{d\theta}{d\tau}\right)^2 \quad \longleftrightarrow \quad \frac{d^2r}{d\tau^2} = -\Gamma^0_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau}$$

term by term one deduces that

$$\Gamma^0_{11} = -r \qquad \Gamma^0_{00} = \Gamma^0_{10} = \Gamma^0_{01} = 0,$$

and a corresponding comparison of

$$\frac{d}{d\tau}\left(r^2\frac{d\theta}{d\tau}\right) = 0 \quad \longleftrightarrow \quad \frac{d^2\theta}{d\tau^2} = -\Gamma^1_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau}$$

yields

$$\Gamma^1_{01} = \Gamma^1_{10} = \frac{1}{r} \qquad \Gamma^1_{00} = \Gamma^1_{11} = 0.$$

It is easily checked that the same coefficients result from solution of Eq. (7.30). For example,

$$\Gamma^0_{11} = \frac{1}{2} g^{00} \left(\frac{\partial g_{10}}{\partial \theta} + \frac{\partial g_{10}}{\partial \theta} - \frac{\partial g_{11}}{\partial r} \right) = -r.$$

7.5 The connection coefficients $\Gamma^\alpha_{\mu\nu}$ may be constructed using either the method of comparing geodesic equations with the Euler–Lagrange equations of motion, as illustrated in Problem 7.4, or by direct solution of Eq. (7.30). Let's use the latter method. From the line element the metric is diagonal $g_{\mu\nu} = \text{diag}\left(-1, 1, \rho^2 + r^2, (\rho^2 + r^2)\sin^2\theta\right)$, implying that

$$g^{\mu\nu} = g^{-1}_{\mu\nu} = \text{diag}\left(-1, 1, (\rho^2 + r^2)^{-1}, (\rho^2 + r^2)^{-1}\sin^{-2}\theta\right).$$

From Eq. (7.30), for a diagonal metric the summation is restricted to a single term:

$$\Gamma^\sigma_{\lambda\mu} = \frac{1}{2} g^{\sigma\sigma} \left(\frac{\partial g_{\mu\sigma}}{\partial x^\lambda} + \frac{\partial g_{\lambda\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\sigma} \right) \qquad \text{(no sum)}.$$

By inspection of the metric, the only non-vanishing derivatives are

$$\frac{\partial g_{22}}{\partial x^1} = 2r \qquad \frac{\partial g_{33}}{\partial x^1} = 2r\sin^2\theta \qquad \frac{\partial g_{33}}{\partial x^2} = 2(\rho^2 + r^2)\sin\theta\cos\theta,$$

implying that

$$\begin{aligned}
\Gamma^1_{22} &= \tfrac{1}{2} g^{11} \left(-\frac{\partial g_{22}}{\partial x^1} \right) = -r \\
\Gamma^1_{33} &= \tfrac{1}{2} g^{11} \left(-\frac{\partial g_{33}}{\partial x^1} \right) = -r\sin^2\theta \\
\Gamma^2_{12} = \Gamma^2_{21} &= \tfrac{1}{2} g^{22} \left(\frac{\partial g_{22}}{\partial x^1} \right) = \frac{r}{\rho^2 + r^2} \\
\Gamma^2_{33} &= \tfrac{1}{2} g^{22} \left(-\frac{\partial g_{33}}{\partial x^2} \right) = -\sin\theta\cos\theta \\
\Gamma^3_{13} = \Gamma^3_{31} &= \tfrac{1}{2} g^{33} \left(\frac{\partial g_{33}}{\partial x^1} \right) = \frac{r}{\rho^2 + r^2} \\
\Gamma^3_{32} = \Gamma^3_{23} &= \tfrac{1}{2} g^{33} \left(\frac{\partial g_{33}}{\partial x^2} \right) = \cot\theta.
\end{aligned}$$

are the non-vanishing connection coefficients.

7.6 This example is worked out in Section 3.5 of Carroll [63], using the method illustrated in the first part of the Problem 7.4 solution above. The $\mu = 0$ component of the Euler–Lagrange equation is

$$\frac{d^2 t}{d\tau^2} + a\dot{a}\delta_{ij} \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} = 0.$$

Comparison with the geodesic equation (7.21) for $x^0 = t$,

$$\frac{d^2x^0}{d\tau^2} + \Gamma^0_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0,$$

then requires that

$$\Gamma^0_{00} = 0 \qquad \Gamma^0_{i0} = \Gamma^0_{0i} = 0 \qquad \Gamma^0_{ij} = a\dot{a}\delta_{ij}.$$

For any of the spatial coordinates the Euler–Lagrange equation is

$$\frac{d^2x^i}{d\tau^2} + 2\frac{\dot{a}}{a}\frac{dt}{d\tau}\frac{dx^i}{d\tau} = 0,$$

and comparison with the corresponding geodesic equation

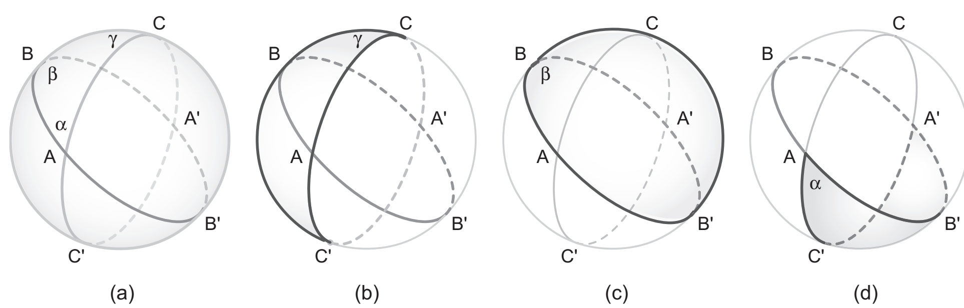
$$\frac{d^2x^i}{d\tau^2} + \Gamma^i_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0$$

requires that

$$\Gamma^i_{j0} = \Gamma^i_{0j} = \frac{\dot{a}}{a}\delta^i_j \qquad \Gamma^i_{00} = \Gamma^i_{jk} = 0$$

for the remaining connection coefficients.

7.7 This solution is adapted from an example in Cheng [64]. Consider 3 great circles $BCB'C'$, $BAB'A'$, and $ACA'C'$ on the sphere (a) in the following figure:



The area bounded by two great circles is called a lune. By spherical symmetry the two angles associated with a lune are equal. The area of a lune is given by $A = 2R^2\theta$, where θ is the angle between the great circles and R is the radius of the sphere. The three great circles define a total of six lunes (three pairs) having angles α, β, and γ respectively. Consider the three lunes marked by heavy lines and shading in figures (b), (c), and (d) above. The sum of the areas of these three lunes covers the entire facing hemisphere of the sphere, plus the area of spherical triangle ABC [since it is included in both the lune in (b) and the lune in (c)], plus the area of spherical triangle $A'B'C'$ [since it extends into the back hemisphere for the lune in (d)]. But spherical triangles ABC and $A'B'C'$ are congruent and the area of the front hemisphere may be expressed as

$$2\pi R^2 = 2R^2(\alpha + \beta + \gamma) - 2A_\Delta,$$

where A_Δ is the area of the spherical triangle ABC or $A'B'C'$. Solving the above equation for the angular excess $\varepsilon \equiv \alpha+\beta+\gamma-\pi$ of the spherical triangle ABC gives

$$\varepsilon = \alpha+\beta+\gamma-\pi = \frac{A_\Delta}{R^2} = KA_\Delta,$$

where $K = 1/R^2$ is the Gaussian curvature for a 2-sphere. This specific proof for spherical triangles on spheres may be generalized to arbitrary polygons on smooth curved surfaces, basically by noting that any spherical polygon can be decomposed into spherical triangles and any curved surface can be approximated locally by a spherical surface.

7.8 By the chain rule $d/d\sigma = (dx^\nu/d\sigma)\partial_\nu$, where $\partial_\nu \equiv \partial/\partial x^\nu$. Therefore,

$$\frac{d^2x^\mu}{d\sigma^2} = \frac{dx^\nu}{d\sigma}\partial_\nu \frac{dx^\mu}{d\sigma}.$$

Substitute the covariant derivative (3.55) for the partial derivative ∂_ν (with $A^\mu = dx^\mu/d\sigma$) to give

$$\frac{dx^\nu}{d\sigma}\partial_\nu \frac{dx^\mu}{d\sigma} \qquad \longrightarrow \qquad \frac{d^2x^\mu}{d\sigma^2} + \Gamma^\mu_{\nu\alpha}\frac{dx^\nu}{d\sigma}\frac{dx^\alpha}{d\sigma},$$

which is the left side of the geodesic equation (7.21). Thus $d^2x^\mu/d\sigma^2 = 0$ becomes the geodesic equation if partial derivatives are replaced by covariant derivatives.

7.9 From the general expression (7.30) for the connection coefficient in terms of the metric tensor and its derivatives,

$$\begin{aligned}\Gamma^\sigma_{\lambda\sigma} &= \tfrac{1}{2}g^{\nu\sigma}(g_{\sigma\nu,\lambda} + g_{\lambda\nu,\sigma} - g_{\sigma\lambda,\nu})\\ &= \tfrac{1}{2}g^{\nu\sigma}(g_{\lambda\nu,\sigma} - g_{\sigma\lambda,\nu}) + \tfrac{1}{2}g^{\nu\sigma}g_{\sigma\nu,\lambda},\end{aligned}$$

where the second line is just a rearrangement of the first. But the quantity on the second line inside the parentheses is antisymmetric under the exchange $\sigma \leftrightarrow \nu$, so when it is contracted with the symmetric $g^{\nu\sigma}$ it must vanish. More formally, using the symmetry of the metric tensor under exchange of indices and relabeling of dummy summation indices,

$$g^{\nu\sigma}g_{\lambda\nu,\sigma} = g^{\sigma\nu}g_{\lambda\nu,\sigma} = g^{\nu\sigma}g_{\lambda\sigma,\nu} = g^{\nu\sigma}g_{\sigma\lambda,\nu}.$$

Thus, $g^{\nu\sigma}(g_{\lambda\nu,\sigma} - g_{\sigma\lambda,\nu}) = 0$ and $\Gamma^\sigma_{\lambda\sigma} = \frac{1}{2}g^{\nu\sigma}g_{\nu\sigma,\lambda}$.

7.10 (a) From the definition (3.57),

$$\delta^\alpha_{\beta;\gamma} = \partial_\gamma\delta^\alpha_\beta + \Gamma^\alpha_{\varepsilon\gamma}\delta^\varepsilon_\beta - \Gamma^\varepsilon_{\beta\gamma}\delta^\alpha_\varepsilon = 0 + \Gamma^\alpha_{\beta\gamma} - \Gamma^\alpha_{\beta\gamma} = 0.$$

(b) Utilizing the result from part (a),

$$\delta^\alpha_{\beta;\gamma} = 0 = (g^{\alpha\varepsilon}g_{\varepsilon\beta})_{;\gamma} = g^{\alpha\varepsilon}_{;\gamma}g_{\varepsilon\beta} + g^{\alpha\varepsilon}g_{\varepsilon\beta;\gamma} = g^{\alpha\varepsilon}_{;\gamma}g_{\varepsilon\beta},$$

where we've used the product rule for covariant derivatives and that $g_{\varepsilon\beta;\gamma} = 0$ (metric connection). Thus $g^{\alpha\varepsilon}_{;\gamma}g_{\varepsilon\beta} = 0$ and multiplying by $g^{\beta\nu}$ gives $g^{\beta\nu}g^{\alpha\varepsilon}_{;\gamma}g_{\varepsilon\beta} = g^{\alpha\varepsilon}_{;\gamma}\delta^\nu_\varepsilon = 0$ and thus that $g^{\alpha\nu}_{;\gamma} = 0$.

7.11 Using Eq. (3.65), the condition (7.17) for parallel transport on a path parameterized by τ is

$$\frac{DA^\lambda}{D\tau} = \frac{dA^\lambda}{d\tau} + \Gamma^\lambda_{\mu\nu} A^\mu \frac{dx^\nu}{d\tau} = 0.$$

Setting A^μ equal to the vector $dx^\mu/d\tau$ tangent to the path gives immediately

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu} \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0,$$

which is the geodesic equation (7.21). Therefore, a particle obeying the geodesic equation follows a path that is the straightest possible in spacetime, in that the path may be constructed by parallel transport of its own tangent vector.

7.12 The inner product is $g_{\alpha\beta}A^\alpha B^\beta$ and its absolute derivative on a path parameterized by u is

$$\frac{D(g_{\alpha\beta}A^\alpha B^\beta)}{Du} = \frac{D(g_{\alpha\beta})}{Du} A^\alpha B^\beta + g_{\alpha\beta}\frac{DA^\alpha}{Du} B^\beta + g_{\alpha\beta}A^\alpha \frac{DB^\beta}{Du},$$

since the absolute derivative obeys the usual Leibniz rule for derivatives of products. But $DA^\alpha/Du = DB^\beta/Du = 0$ (definition of parallel transport of vectors) and $D(g_{\alpha\beta})/Du = 0$ (property of metric connection; see Section 7.8). Thus $D(g_{\alpha\beta}A^\alpha B^\beta)/Du = 0$ on the path and the inner product is unchanged by parallel transport if the connection is a metric connection.

7.13 From the general expression (7.30),

$$\Gamma^\sigma_{0\mu} = \frac{1}{2} g^{\nu\sigma}\left(\frac{\partial g_{\mu\nu}}{\partial x^0} + \frac{\partial g_{0\nu}}{\partial x^\mu} - \frac{\partial g_{\mu 0}}{\partial x^\nu}\right) \simeq \frac{1}{2} g^{\nu\sigma}\left(\frac{\partial g_{0\nu}}{\partial x^\mu} - \frac{\partial g_{0\mu}}{\partial x^\nu}\right),$$

where the first term has been neglected because the fields are assumed to vary slowly with time. But $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(x)$, so that $\partial g_{\mu\nu}/\partial x^\lambda = \partial h_{\mu\nu}/\partial x^\lambda$ and

$$\Gamma^\sigma_{0\mu} = \frac{1}{2}\eta^{\nu\sigma}\left(\frac{\partial h_{0\nu}}{\partial x^\mu} - \frac{\partial h_{0\mu}}{\partial x^\nu}\right),$$

where terms have been retained only to lowest order in $h_{\mu\nu}$. (We have been a a little cavalier here since we have assumed that since $h_{\mu\nu}$ is small its derivatives are also small.) Restricting this expression to $\mu = 0$ gives

$$\Gamma^\sigma_{00} = \frac{1}{2}\eta^{\nu\sigma}\left(\frac{\partial h_{0\nu}}{\partial x^0} - \frac{\partial h_{00}}{\partial x^\nu}\right) \simeq -\frac{1}{2}\eta^{\nu\sigma}\frac{\partial h_{00}}{\partial x^\nu},$$

where the first term has been dropped because the fields are assumed to vary slowly.

7.14 By the equivalence principle the local inertial frame corresponds to a locally flat space. From Eq. (7.30), the Christoffel symbols vanish because the first derivatives of the metric are all equal to zero in the local inertial frame. However, the second derivatives of the metric do not generally vanish so the first derivatives of (7.30) will not be zero.

7.15 The spherical and cartesian coordinates are related by $x = r\cos\theta$ and $y = r\sin\theta$.

From the transformation law $e'_i = (\partial x^j/\partial x'^i)e_j$ one obtains that the basis vectors in the (r,θ) coordinates are

$$\boldsymbol{e}_r = \cos\theta\boldsymbol{e}_x + \sin\theta\boldsymbol{e}_y \qquad \boldsymbol{e}_\theta = -r\sin\theta\boldsymbol{e}_x + r\cos\theta\boldsymbol{e}_y.$$

Since

$$\frac{\partial\boldsymbol{e}_i}{\partial x^j} = \Gamma^k_{ij}\boldsymbol{e}_k,$$

the partial derivatives $\partial\boldsymbol{e}_i/\partial x^j$ are required to evaluate the connection coefficients. From the above equations, these are

$$\begin{aligned}
\frac{\partial\boldsymbol{e}_r}{\partial r} &= 0\\
\frac{\partial\boldsymbol{e}_r}{\partial\theta} &= -\sin\theta\boldsymbol{e}_x + \cos\theta\boldsymbol{e}_y = \frac{1}{r}\boldsymbol{e}_\theta\\
\frac{\partial\boldsymbol{e}_\theta}{\partial r} &= -\sin\theta\boldsymbol{e}_x + \cos\theta\boldsymbol{e}_y = \frac{1}{r}\boldsymbol{e}_\theta\\
\frac{\partial\boldsymbol{e}_\theta}{\partial\theta} &= -r\cos\theta\boldsymbol{e}_x - r\sin\theta\boldsymbol{e}_y = -r\boldsymbol{e}_r.
\end{aligned}$$

Now the connection coefficients can be read off from the expansion. For example,

$$\frac{\partial\boldsymbol{e}_r}{\partial\theta} = \Gamma^r_{r\theta}\boldsymbol{e}_r + \Gamma^\theta_{r\theta}\boldsymbol{e}_\theta = \frac{1}{r}\boldsymbol{e}_\theta \qquad \longrightarrow \qquad \Gamma^r_{r\theta} = 0 \qquad \Gamma^\theta_{r\theta} = \frac{1}{r}.$$

Carrying this out for all eight index combinations yields

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r} \qquad \Gamma^r_{\theta\theta} = -r \qquad \Gamma^r_{rr} = \Gamma^\theta_{rr} = \Gamma^r_{r\theta} = \Gamma^r_{\theta r} = \Gamma^\theta_{\theta\theta} = 0.$$

This may be checked using Eq. (7.30), which can be written for this 2D case as

$$\Gamma^k_{ij} = \frac{1}{2}g^{mk}\left(\frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ji}}{\partial x^m}\right),$$

with i and j taking on the values r or θ. In polar coordinates the line element is $ds^2 = dr^2 + r^2d\theta^2$, so the metric takes the form

$$g_{ij} = \begin{pmatrix}1 & 0\\ 0 & r^2\end{pmatrix} \qquad g^{ij} = (g_{ij})^{-1} = \begin{pmatrix}1 & 0\\ 0 & 1/r^2\end{pmatrix}.$$

Thus, $g_{rr} = 1$, $g_{\theta\theta} = r^2$, $g_{r\theta} = g_{\theta r} = 0$, and $g^{rr} = 1$, $g^{\theta\theta} = 1/r^2$, $g^{\theta r} = g^{r\theta} = 0$. Since the metric is diagonal, only the $m = k$ terms will survive in the above summation for connection coefficient because of the factor g^{mk}, so for example

$$\Gamma^\theta_{r\theta} = \frac{1}{2}g^{\theta\theta}\left(\frac{\partial g_{\theta\theta}}{\partial r} + \frac{\partial g_{r\theta}}{\partial\theta} - \frac{\partial g_{\theta r}}{\partial\theta}\right) = \frac{1}{r},$$

which is the same result as obtained above. The other Γ^k_{ik} can be checked in the same way. (See also the related Problem 7.4.)

7.16 An alert reader may observe that the substitution $r = \sin\varphi$ in the line element $ds^2 = dr^2 + r^2d\theta^2$ for flat 2D space expressed in polar coordinates gives the line element $ds^2 =$

$\cos^2\varphi d\varphi^2 + \sin^2\varphi d\theta^2$ given in the problem. Thus, the parameterization obscures that this is just flat 2D space and the Gaussian curvature must be zero. Let's check this conclusion using Eq. (7.5) and $ds^2 = \cos^2\varphi d\varphi^2 + \sin^2\varphi d\theta^2$. Consider a line drawn from the point $(\varphi,\theta) = (0,0)$ to the point $(\lambda,0)$ at constant $\theta = 0$. The length of the line is

$$S = \int \sqrt{ds^2} = \int_0^\lambda \cos\varphi d\varphi = \sin\lambda.$$

The locus of points at constant $S = \sin\lambda$ for all values of θ then traces a circle of radius $S = \sin\lambda$. The circumference of this circle is given by the integral

$$C = \int ds = \sin\lambda \int_0^{2\pi} d\theta = 2\pi\sin\lambda.$$

Then from Eq. (7.5),

$$K = \operatorname*{Lim}_{\lambda\to 0} \frac{6}{S^2}\left(1 - \frac{C}{2\pi S}\right) = \operatorname*{Lim}_{\lambda\to 0} \frac{6}{\sin^2\lambda}\left(1 - \frac{2\pi\sin\lambda}{2\pi\sin\lambda}\right) = 0,$$

and the space indeed has zero intrinsic curvature.

7.17 In the coordinates (r,θ), construct a line segment from the origin $(0,0)$ to the point $(\lambda,0)$. From the metric $ds^2 = (1+\alpha r^2)dr^2 + r^2 d\theta^2$, the length of this segment is

$$S = \int \sqrt{ds^2} = \int_0^\lambda (1+\alpha r^2)^{1/2} dr \simeq \int_0^\lambda (1 + \tfrac{1}{2}\alpha r^2 + \ldots)\, dr = \lambda + \tfrac{1}{6}\alpha\lambda^3,$$

where $d\theta^2 = 0$ was used and the square root was expanded in a binomial series since our interest is in the limit $\lambda \to 0$. The points $(r = \lambda, \theta)$ as θ ranges 0 to 2π define a circle of radius $S = \lambda + \frac{1}{6}\alpha\lambda^3$, which has a circumference

$$C = \int \sqrt{ds^2} = (\lambda + \tfrac{1}{6}\alpha\lambda^3)\int_0^{2\pi} d\theta = 2\pi\lambda + \tfrac{1}{3}\pi\alpha\lambda^3.$$

Then from Eq. (7.5) the Gaussian curvature is

$$K = \operatorname*{Lim}_{\lambda\to 0} \frac{6}{\lambda^2}\left(1 - \frac{2\pi\lambda + \frac{1}{3}\pi\alpha\lambda^3}{2\pi\lambda}\right) = -\alpha,$$

where terms of order λ^3 have been ignored in denominators.

7.18 If the fluid is at rest $u = (u^0,0,0,0)$, and the normalization $u\cdot u = g_{\mu\nu}u^\mu u^\nu = -1$ gives $u_0 u^0 = -1$, since $u_\mu u^\nu = 0$ unless $\mu = \nu = 0$. Hence all non-diagonal elements of $T^\mu{}_\nu$ vanish and

$$T^0{}_0 = (\varepsilon + P)u_0 u^0 + P = -(\varepsilon + P) + P = -\varepsilon \qquad T^1{}_1 = T^2{}_2 = T^3{}_3 = P$$

in the rest frame of a perfect fluid.

7.19 Suppose a vector field $V(\lambda)$ defined only along a curve $x^\mu(\lambda)$ parameterized by λ in a manifold. Assuming that $V(\lambda) = V^\mu(\lambda)e_\mu(\lambda)$, where $e_\mu(\lambda)$ is a coordinate basis vector evaluated at the point on the curve labeled by λ, then

$$\begin{aligned}\frac{dV}{d\lambda} &= \frac{dV^\mu}{d\lambda} e_\mu + V^\mu \frac{de_\mu}{d\lambda} \\ &= \frac{dV^\mu}{d\lambda} e_\mu + V^\mu \frac{\partial e_\mu}{\partial x^\nu}\frac{dx^\nu}{d\lambda},\end{aligned}$$

where the chain rule was used in the last step. As in Section 2.3, expand the partial derivative factor in the vector basis,

$$\frac{\partial e_\mu}{\partial x^\nu} = \Gamma^\alpha_{\mu\nu} e_\alpha,$$

which gives

$$\begin{aligned}\frac{dV}{d\lambda} &= \frac{dV^\mu}{d\lambda} e_\mu + \Gamma^\alpha_{\mu\nu} V^\mu \frac{dx^\nu}{d\lambda} e_\alpha \\ &= \frac{dV^\mu}{d\lambda} e_\mu + \Gamma^\mu_{\alpha\nu} V^\alpha \frac{dx^\nu}{d\lambda} e_\mu \\ &= \left(\frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\alpha\nu} V^\alpha \frac{dx^\nu}{d\lambda}\right) e_\mu \\ &\equiv \frac{DV^\mu}{D\lambda} e_\mu,\end{aligned}$$

where in line two dummy summation indices in the second term were interchanged. Thus

$$\frac{DV^\mu}{D\lambda} = \frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\alpha\nu} V^\alpha \frac{dx^\nu}{d\lambda},$$

which is Eq. (3.65) for the absolute or intrinsic derivative.

7.20 (a) Differentiation of the 4-vector V gives

$$\begin{aligned}\frac{dV}{d\lambda} = \frac{d}{d\lambda}\left(V^\mu e_\mu\right) &= \frac{dV^\mu}{d\lambda} e_\mu + \frac{de_\mu}{d\lambda} V^\mu \\ &= \frac{\partial V^\mu}{\partial x^\nu}\frac{dx^\nu}{d\lambda} e_\mu + \frac{\partial e_\mu}{\partial x^\nu}\frac{dx^\nu}{d\lambda} V^\mu \\ &= \partial_\nu V^\mu u^\nu e_\mu + V^\mu u^\nu \partial_\nu e_\mu,\end{aligned}$$

where we have defined

$$u^\nu \equiv \frac{dx^\nu}{d\lambda} \qquad \partial_\nu \equiv \frac{\partial}{\partial x^\nu}.$$

Expanding $\partial_\nu e_\mu = \Gamma^\alpha_{\mu\nu} e_\alpha$ as in Section 2.3 gives

$$\begin{aligned}\frac{dV}{d\lambda} &= \partial_\nu V^\mu u^\nu e_\mu + \Gamma^\alpha_{\mu\nu} V^\mu u^\nu e_\alpha \\ &= \left(\partial_\nu V^\mu + \Gamma^\mu_{\alpha\nu} V^\alpha\right) u^\nu e_\mu,\end{aligned}$$

where dummy summation indices were switched in the second term. This is Eq. (7.14).

(b) From the first part

$$\frac{dV}{d\lambda} = \left(\frac{\partial V^\mu}{\partial x^\nu} + \Gamma^\mu_{\alpha\nu} V^\alpha\right)\frac{dx^\nu}{d\lambda} e_\mu.$$

Multiply both sides by $d\lambda/dx^\beta$ to give

$$\frac{dV}{dx^\beta} = \left(\frac{\partial V^\mu}{\partial x^\beta} + \Gamma^\mu_{\alpha\beta} V^\alpha\right) e_\mu.$$

Take the scalar product of both sides with e^η to give

$$\frac{dV^\mu}{dx^\nu} \equiv \nabla_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\nu\alpha} V^\alpha,$$

where some indices have been renamed. Comparison with Eq. (3.55) indicates that this defines the covariant derivative of the vector V.

8 The General Theory of Relativity

8.1 This problem was adapted from a similar one in Ref. [141]. From Eq. (7.30), the connection coefficients are

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{\ell k} \left(\frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right),$$

where i, j, k, ℓ take the values 0 or 1. From the line element the metric is

$$g_{ij} = \begin{pmatrix} -1 & 0 \\ 0 & f(t)^2 \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} -1 & 0 \\ 0 & f(t)^{-2} \end{pmatrix}.$$

The formula may now be used to compute the Γ^k_{ij}. For example,

$$\begin{aligned} \Gamma^1_{01} = \Gamma^1_{10} &= \tfrac{1}{2} g^{11} \left(\frac{\partial g_{11}}{\partial x^0} + \frac{\partial g_{01}}{\partial x^1} - \frac{\partial g_{10}}{\partial x^1} \right) \\ &= \tfrac{1}{2} g^{11} \frac{\partial g_{11}}{\partial x^0} = \frac{1}{2f^2}(2f\dot{f}) = \frac{\dot{f}}{f}, \end{aligned}$$

where $\dot{f} \equiv \partial f / \partial x^0$. Determining the other possibilities in a similar way gives

$$\Gamma^1_{01} = \Gamma^1_{10} = \dot{f}/f \qquad \Gamma^0_{11} = f\dot{f} \qquad \Gamma^0_{00} = \Gamma^1_{00} = \Gamma^1_{11} = \Gamma^0_{01} = \Gamma^0_{10} = 0.$$

Then from Eq. (8.14)

$$R^0{}_{101} = \frac{\partial \Gamma^0_{11}}{\partial x^0} - \Gamma^0_{11}\Gamma^1_{10} = \frac{\partial (f\dot{f})}{\partial x^0} - f\dot{f}\,\frac{\dot{f}}{f} = f(t)\ddot{f}(t).$$

for the curvature tensor component $R^0{}_{101}$. Since the space is two-dimensional, there is only one independent curvature component (see Box 8.1) and all other components must be zero or related to $R^0{}_{101}$ by the symmetries (8.15).

8.2 Take the line element in the form

$$ds^2 = -e^\sigma dt^2 + e^\lambda dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

where $\sigma = \sigma(r,t)$ and $\lambda = \lambda(r,t)$. The corresponding metric is

$$\begin{aligned} g_{\mu\nu} &= \operatorname{diag}(-e^\sigma, e^\lambda, r^2, r^2 \sin^2\theta) \\ g^{\mu\nu} &= g^{-1}_{\mu\nu} = \operatorname{diag}\left(-e^{-\sigma}, e^{-\lambda}.\frac{1}{r^2}, \frac{1}{r^2\sin^2\theta} \right) \end{aligned}$$

From Eq. (7.30) the non-vanishing Christoffel symbols are

$$\Gamma^0_{00} = \tfrac{1}{2}\dot{\sigma} \qquad \Gamma^0_{01} = \tfrac{1}{2}\sigma' \qquad \Gamma^0_{11} = \tfrac{1}{2}e^{\lambda-\sigma}\dot{\lambda} \qquad \Gamma^1_{00} = \tfrac{1}{2}e^{\sigma-\lambda}\sigma' \qquad \Gamma^1_{01} = \tfrac{1}{2}\dot{\lambda}$$
$$\Gamma^1_{11} = \tfrac{1}{2}\lambda' \qquad \Gamma^1_{22} = -re^{-\lambda} \qquad \Gamma^1_{33} = -re^{-\lambda}\sin^2\theta \qquad \Gamma^2_{12} = 1/r$$
$$\Gamma^2_{33} = -\sin\theta\cos\theta \qquad \Gamma^3_{13} = 1/r \qquad \Gamma^3_{23} = \cot\theta$$

where dots indicate time derivatives and primes indicate derivatives with respect to r.

The Riemann curvature tensor may then be calculated from Eq. (8.14); the non-zero components are

$$R_{0101} = \tfrac{1}{2}e^{\sigma}\sigma'' - \tfrac{1}{4}e^{\lambda}\dot{\lambda}^2 + \tfrac{1}{4}e^{\lambda}\dot{\sigma}\dot{\lambda} - \tfrac{1}{2}e^{\lambda}\ddot{\lambda} + \tfrac{1}{4}e^{\sigma}(\sigma')^2 - \tfrac{1}{4}e^{\sigma}\sigma'\lambda'$$
$$R_{0202} = \tfrac{1}{2}e^{\sigma-\lambda}\sigma' \qquad R_{0212} = \tfrac{1}{2}r\dot{\lambda} \qquad R_{0303} = \tfrac{1}{2}re^{\sigma-\lambda}\sigma'\sin^2\theta$$
$$R_{0313} = \tfrac{1}{2}r\dot{\lambda}\sin^2\theta \qquad R_{1212} = \tfrac{1}{2}r\lambda'$$
$$R_{1313} = \tfrac{1}{2}r\lambda'\sin^2\theta \qquad R_{2323} = r^2\sin^2\theta(1-e^{-\lambda}).$$

The Ricci tensor $R_{\mu\nu}$ then follows from Eq. (8.16), with non-vanishing components

$$R_{00} = -\tfrac{1}{2}e^{\sigma-\lambda}\sigma'' - \tfrac{1}{4}e^{\sigma-\lambda}(\sigma')^2 + \tfrac{1}{2}\ddot{\lambda} + \tfrac{1}{4}\dot{\lambda}^2 - \tfrac{1}{4}\dot{\sigma}\dot{\lambda} + \tfrac{1}{4}e^{\sigma-\lambda}\sigma'\lambda' - e^{\sigma-\lambda}\lambda'/r$$
$$R_{11} = \tfrac{1}{2}\sigma'' + \tfrac{1}{4}(\sigma')^2 - \tfrac{1}{2}e^{\lambda-\sigma}\ddot{\lambda} - \tfrac{1}{4}e^{\lambda-\sigma}\dot{\lambda}^2 + \tfrac{1}{4}e^{\lambda-\sigma}\dot{\sigma}\dot{\lambda} - \tfrac{1}{4}\sigma'\lambda' - \lambda'/r$$
$$R_{01} = -\dot{\lambda}/r \qquad R_{22} = \tfrac{1}{2}re^{-\lambda}\sigma' - \tfrac{1}{2}re^{-\lambda}\lambda' + e^{-\lambda} - 1 \qquad R_{33} = R_{22}\sin^2\theta.$$

The Ricci scalar R may then be constructed from Eq. (8.17),

$$R = 2\frac{e^{-\lambda}}{r^2} + 2\frac{e^{-\lambda}}{r}\sigma' - 2\frac{e^{-\lambda}}{r}\lambda' + e^{-\lambda}\sigma'' + \tfrac{1}{2}e^{-\lambda}(\sigma')^2$$
$$- e^{-\sigma}\ddot{\lambda} - \tfrac{1}{2}e^{-\sigma}\dot{\lambda}^2 - \frac{2}{r^2} + \tfrac{1}{2}e^{-\sigma}\dot{\sigma}\dot{\lambda} - \tfrac{1}{2}e^{-\lambda}\sigma'\lambda'.$$

and the Einstein tensor is

$$G_{00} = \frac{e^{\sigma}}{r^2} - \frac{e^{\sigma-\lambda}}{r^2} + \frac{e^{\sigma-\lambda}}{r}\lambda' \qquad G_{01} = -\frac{\dot{\lambda}}{r} \qquad G_{11} = \frac{e^{\lambda}}{r^2} - \frac{1}{r^2} - \frac{\sigma'}{r}$$
$$G_{22} = \tfrac{1}{2}r\lambda'e^{-\lambda} - \tfrac{1}{2}r\sigma'e^{-\lambda} - \tfrac{1}{2}r^2e^{-\lambda}\sigma'' - \tfrac{1}{4}r^2e^{-\lambda}(\sigma')^2 + \tfrac{1}{2}r^2e^{-\sigma}\ddot{\lambda}$$
$$+ \tfrac{1}{4}r^2e^{-\sigma}(\dot{\lambda})^2 - \tfrac{1}{4}r^2e^{-\sigma}\dot{\sigma}\dot{\lambda} + \tfrac{1}{4}r^2e^{-\lambda}\sigma'\lambda'$$
$$G_{33} = G_{22}\sin^2\theta,$$

where Eq. (8.20) was used. These results for a general spherical metric are summarized in Appendix C.

8.3 To calculate the commutator requires taking two successive covariant derivatives on an arbitrary vector, and then subtract from that the result of those two successive covariant differentiations taken in the opposite order. But in taking the covariant derivative and then taking the covariant derivative of the result it must be remembered that the covariant derivative of the vector yields a rank-2 tensor, so the second covariant derivative is not that of a vector but of a rank-2 tensor. Utilizing Eq. (3.53) in the first step and Eq. (3.56) on the result, and that the covariant derivative obeys the usual Leibniz rule for the derivative of products, the result of two successive covariant differentiations with respect to ν and then λ on an arbitrary vector V_μ is

$$\begin{aligned} V_{\mu;\nu\lambda} &= \partial_\lambda (V_{\mu;\nu}) - \Gamma^\alpha_{\mu\lambda} V_{\alpha;\nu} - \Gamma^\alpha_{\nu\lambda} V_{\mu;\alpha} \\ &= \partial_\lambda \partial_\nu V_\mu - (\partial_\lambda \Gamma^\sigma_{\mu\nu}) V_\sigma - \Gamma^\sigma_{\mu\nu} \partial_\lambda V_\sigma \\ &\quad - \Gamma^\alpha_{\mu\lambda} (\partial_\nu V_\alpha - \Gamma^\sigma_{\alpha\nu} V_\sigma) - \Gamma^\alpha_{\nu\lambda} (\partial_\alpha V_\mu - \Gamma^\sigma_{\mu\alpha} V_\sigma), \end{aligned}$$

where $\partial_\beta \equiv \partial / \partial x^\beta$. Form an analogous expression but with ν and λ interchanged and subtract it from the preceding expression to form the commutator $[\nabla_\nu, \nabla_\lambda] V_\mu \equiv V_{\mu;\nu\lambda} - V_{\mu;\lambda\nu}$. In the resulting expression various terms cancel because (1) repeated indices are dummy indices that can be replaced by arbitrary repeated indices, (2) $\partial_\nu \partial_\lambda = \partial_\lambda \partial_\nu$, and (3) the Christoffel symbols are symmetric in their lower indices. The surviving terms give

$$[\nabla_\nu, \nabla_\lambda] V_\mu = R^\sigma{}_{\mu\nu\lambda} V_\sigma,$$

where the definition (8.14) of the Riemann curvature tensor $R^\sigma{}_{\mu\nu\lambda}$ was used. Since the dual vector V_μ is arbitrary, this proves that the curvature tensor is the commutator of the covariant derivatives.

8.4 The deviation is given by the second term in Eq. (8.12):

$$h_{00} = 2\frac{GM}{rc^2} = 1.483 \times 10^{-28} \left(\frac{M}{1\text{g}} \right) \left(\frac{1\text{cm}}{r} \right)$$

Estimates for several objects are shown in the following table.

Object	Mass (g)	Radius (cm)	h_{00}
Sun	1.989×10^{33}	6.96×10^{10}	4.23×10^{-6}
Earth	5.976×10^{27}	6.378×10^{8}	1.4×10^{-9}
Proton	1.67×10^{-24}	10^{-13}	2.5×10^{-39}
Sirius B	2.09×10^{33}	5.5×10^{8}	$5.6 \times 10{-}4$
Neutron star	$\sim 2 \times 10^{33}$	10^{6}	0.30

Among these examples, only the surface of a neutron star exhibits substantial deviation from the flat metric. Therefore, Newtonian gravity should be a good approximation for all of these except the neutron star, for which deviations from Newtonian gravity may be substantial.

8.5 For the $\mu = 0$ part of Eq. (8.4) $d^2x^0/d\tau^2 = 0$. For the $\mu = i$ components, from (8.4) and (8.8),

$$\frac{d^2x^i}{d\tau^2} - \frac{1}{2}\frac{\partial h_{00}}{\partial x^i} \left(\frac{dx^0}{d\tau} \right)^2 = 0.$$

But for the first term

$$\begin{aligned}\frac{d^2x^i}{d\tau^2} &= \frac{d}{d\tau}\left(\frac{dx^i}{dx^0}\frac{dx^0}{d\tau}\right) = \frac{dx^0}{d\tau}\frac{d}{d\tau}\left(\frac{dx^i}{dx^0}\right) \\ &= \frac{dx^0}{d\tau}\frac{d}{dx^0}\frac{dx^0}{d\tau}\left(\frac{dx^i}{dx^0}\right) = \left(\frac{dx^0}{d\tau}\right)^2\frac{d}{dx^0}\left(\frac{dx^i}{dx^0}\right) \\ &= \frac{1}{c^2}\left(\frac{dx^0}{d\tau}\right)^2\frac{d}{dt}\left(\frac{dx^i}{dt}\right) = \frac{1}{c^2}\left(\frac{dx^0}{d\tau}\right)^2\frac{d^2x^i}{dt^2}\end{aligned}$$

where $dx^0 = cdt$ and $dx^0/d\tau =$ constant has been used. Inserting this result in the first equation above yields the second equation of Eq. (8.9).

8.6 In the rest frame of a clock falling freely special relativity is valid (equivalence principle). The special relativistic time dilation formula gives

$$d\tau = \left(1 - \frac{v^2}{c^2}\right)^{1/2} dx^0 \simeq \left(1 - \frac{v^2}{2c^2}\right) dx^0,$$

where the last step follows from the assumption of low velocity. By energy conservation, $\frac{1}{2}mv^2 = -m\varphi$, where φ is the gravitational potential, so $v^2 = -2\varphi$. Inserting this in the preceding equation gives

$$d\tau = \left(1 + \frac{\varphi}{c^2}\right) dx^0.$$

But from Eq. (8.10) in the weak-field limit,

$$(-g_{00})^{1/2} = \left(1 + \frac{2\varphi}{c^2}\right)^{1/2} \simeq 1 + \frac{\varphi}{c^2}.$$

Therefore, for objects falling slowly in weak gravitational fields the proper time τ and coordinate time x^0 are related by $d\tau = (-g_{00})^{1/2} dx^0$.

8.7 This solution is adapted from one given in Cheng [64]. The metric tensor can be expressed as

$$g_{ij} = \begin{pmatrix} g_{11} & 0 \\ 0 & g_{22} \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} 1/g_{11} & 0 \\ 0 & 1/g_{22} \end{pmatrix}$$

The non-trivial independent element of the Riemann curvature tensor (8.14) in two dimensions may be taken to be

$$\begin{aligned}R_{1212} &= g_{1k}R^k{}_{212} = g_{11}R^1{}_{212} \\ &= g_{11}\left(\frac{\partial\Gamma^1_{21}}{\partial x^2} - \frac{\partial\Gamma^1_{22}}{\partial x^1} - \Gamma^1_{11}\Gamma^1_{22} - \Gamma^1_{21}\Gamma^2_{22} + \Gamma^1_{12}\Gamma^1_{21} + \Gamma^1_{22}\Gamma^2_{21}\right).\end{aligned}$$

The required connection coefficients are given by Eq. (7.30). For example:

$$\Gamma^1_{11} = \frac{1}{2}g^{11}\left(\frac{\partial g_{11}}{\partial x^1} + \frac{\partial g_{11}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^1}\right) = \frac{1}{2}g^{11}\frac{\partial g_{11}}{\partial x^1} = \frac{1}{2g_{11}}\frac{\partial g_{11}}{\partial x^1}.$$

In a similar manner

$$\Gamma^1_{12} = \Gamma^1_{21} = \frac{1}{2g_{11}}\frac{\partial g_{11}}{\partial x^2} \qquad \Gamma^1_{22} = -\frac{1}{2g_{11}}\frac{\partial g_{22}}{\partial x^1}$$
$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial x^1} \qquad \Gamma^2_{22} = \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial x^2}.$$

Inserting these connection coefficients in the expression for R_{1212} gives

$$R_{1212} = \frac{1}{2}\left\{\frac{\partial^2 g_{22}}{\partial (x^1)^2} + \frac{\partial^2 g_{11}}{\partial (x^2)^2} - \frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x^1}\frac{\partial g_{22}}{\partial x^1} + \left(\frac{\partial g_{11}}{\partial x^2}\right)^2\right]\right.$$
$$\left. - \frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x^2}\frac{\partial g_{22}}{\partial x^2} + \left(\frac{\partial g_{22}}{\partial x^1}\right)^2\right]\right\},$$

which is Eq. (7.2), up to a normalization factor $-1/\det g$.

8.8 From the symmetries (8.15) it may be shown that all contractions of one upper index on the curvature tensor with one of its lower indices either gives $\pm R_{\mu\nu}$ defined in Eq. (8.16), or zero. For example

$$g^{\lambda\sigma}R_{\mu\lambda\sigma\nu} = R_\mu{}^\sigma{}_{\sigma\nu} \equiv R'_{\mu\nu}.$$

But also since $R_{\sigma\mu\nu\lambda} = -R_{\mu\sigma\nu\lambda}$,

$$g^{\lambda\sigma}R_{\mu\lambda\sigma\nu} = -g^{\lambda\sigma}R_{\lambda\mu\sigma\nu} = -R^\sigma{}_{\mu\sigma\nu} = -R_{\mu\nu}.$$

Thus, $R'_{\mu\nu} = -R_{\mu\nu}$. As another example, multiply $R_{\sigma\mu\nu\lambda} = -R_{\mu\sigma\nu\lambda}$, by $g^{\sigma\mu}$ to give

$$g^{\sigma\mu}R_{\sigma\mu\nu\lambda} = -g^{\sigma\mu}R_{\mu\sigma\nu\lambda} \quad\longrightarrow\quad R^\mu{}_{\mu\nu\lambda} = -R^\sigma{}_{\sigma\nu\lambda} \quad\longrightarrow\quad R^\mu{}_{\mu\nu\lambda} = -R^\mu{}_{\mu\nu\lambda}$$

where dummy (repeated) indices were relabeled. This can be true only if $R^\mu{}_{\mu\nu\lambda} = 0$.

8.9 By symmetry $R_{\beta\mu\alpha\nu} = R_{\alpha\nu\beta\mu}$. Therefore,

$$R_{\mu\nu} = g^{\alpha\beta}R_{\beta\mu\alpha\nu} = g^{\alpha\beta}R_{\alpha\nu\beta\mu} = R^\beta{}_{\nu\beta\mu} = R_{\nu\mu},$$

so $R_{\mu\nu}$ is symmetric. Alternatively, start from the cyclic identity in Eq. (8.15) and raise the first index by contracting with $g^{\alpha\sigma}$ to give

$$R^\alpha{}_{\mu\nu\lambda} + R^\alpha{}_{\lambda\mu\nu} + R^\alpha{}_{\nu\lambda\mu} = 0.$$

Now set $\alpha = \lambda$ to contract on that index:

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

But in Problem 8.8 it was shown that $R^\alpha{}_{\alpha\mu\nu} = 0$, and from Eq. (8.15) the first term in the above equation changes sign if the last two indices are switched. Thus $-R^\alpha{}_{\mu\alpha\nu} + 0 + R^\alpha{}_{\nu\alpha\mu} = 0$, which is equivalent to $R_{\mu\nu} = R_{\nu\mu}$.

8.10 Parameterize a point on a circle of latitude for a sphere by spherical coordinates

$$x^1 \equiv \theta \qquad x^2 \equiv \varphi = \frac{s}{R\sin\theta},$$

where R is the radius of the sphere, θ is the usual polar angle (so latitude is $\frac{\pi}{2}-\theta$), s is the arc length around a circle of latitude measured from $\varphi = 0$, and $R\sin\theta$ is the radius of a circle of latitude. The geodesic equation for this 2D space may then be written as

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0 \qquad (i, j, k = 1 \text{ or } 2).$$

The derivatives required for the geodesic equation are

$$\frac{d^2x^1}{ds^2} = \frac{dx^1}{ds} = 0 \qquad \frac{dx^2}{ds} = \frac{1}{R\sin\theta} \qquad \frac{d^2x^2}{ds^2} = 0.$$

Since all second derivatives vanish and the first derivative vanishes for x^1, only two geodesic equations are non-trivial:

$$\Gamma^1_{22}\frac{dx^2}{ds}\frac{dx^2}{ds} = \frac{\Gamma^1_{22}}{R^2\sin^2\theta} = 0 \qquad \Gamma^2_{22}\frac{dx^2}{ds}\frac{dx^2}{ds} = \frac{\Gamma^2_{22}}{R^2\sin^2\theta} = 0$$

From Example 8.2, all the connection coefficients are zero except for

$$\Gamma^1_{22} = -\sin\theta\cos\theta \qquad \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta.$$

Thus the second geodesic equation is satisfied identically because $\Gamma^2_{22} = 0$ and the first geodesic equation gives for a particular choice of θ,

$$\frac{\Gamma^1_{22}}{R^2\sin^2\theta} = \frac{-\sin\theta\cos\theta}{R^2\sin^2\theta} = \frac{-\cot\theta}{R^2} = 0.$$

But $\cot\theta = 0$ is satisfied only for $\theta = \frac{\pi}{2}$ in the interval $0-\pi$, so the equator is the only circle of latitude that is a geodesic on a sphere. Since the orientation of the spherical coordinate system is arbitrary, this also proves the well-known result that the geodesics of a sphere are the great circles that can be drawn on the sphere. This exercise also illustrates an important conceptual point: the geodesic equation describes the motion of a free particle in a gravitational field, so it may seem surprising from the usual perspective that it was used here to deduce a purely geometrical property of the sphere. But from the general relativity perspective *gravity is geometry,* so the geodesic equation is in fact a geometrical statement about the corresponding space.

8.11 Under most conditions where Newtonian gravity is valid the pressure contribution in the stress–energy tensor may be ignored and Eq. (7.9) may be written as $T_{\mu\nu} \simeq \varepsilon u_\mu u_\nu$. The metric is assumed to be of the form $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $h_{\mu\nu}$ is a small correction, so the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$ may be used to raise and lower indices. Contracting $T_{\mu\nu}$ with $\eta^{\lambda\nu}$ gives $T^\lambda_\mu = \varepsilon u_\mu u^\lambda$ and thus

$$\mathrm{Tr}\,T = T^\mu_\mu = \varepsilon u_\mu u^\mu = -\varepsilon = -\rho c^2,$$

where we have assumed the fluid to be nearly at rest so that the 4-velocities have only timelike components and $u_\mu u^\mu \sim -1$ (see the solution of Problem 7.18). Furthermore,

$$T_{\mu\nu} = \varepsilon u_\mu u_\nu = \varepsilon\eta_{\mu\nu}u_\mu u^\mu = -\varepsilon\eta_{\mu\nu} = -\rho c^2\eta_{\mu\nu},$$

so that $T_{00} = -\eta_{00}\rho c^2 = \rho c^2$. Taking the 00 component of the Einstein equations (8.23) and inserting these results for the stress–energy tensor gives

$$R_{00} = \frac{8\pi G}{c^4}\left(T_{00} - \tfrac{1}{2}g_{00}T^\lambda_\lambda\right) \simeq \frac{8\pi G}{c^4}\left(T_{00} - \tfrac{1}{2}\eta_{00}T^\lambda_\lambda\right) = \frac{4\pi G}{c^2}\rho.$$

In Eq. (8.16) one may assume the connection coefficients to be small in the weak-field limit and ignore the terms quadratic in $\Gamma^\lambda_{\mu\nu}$ to give a second expression for R_{00},

$$R_{00} \simeq \partial_\lambda \Gamma^\lambda_{00} - \partial_0 \Gamma^\lambda_{0\lambda} \simeq \partial_i \Gamma^i_{00} \quad (i = 1,2,3),$$

where the fields were assumed to be varying slowly in time, justifying neglect of all time-like derivatives. But from Eq. (8.8), $\Gamma^i_{00} = -\frac{1}{2}\partial_i h_{00}$ in the weak-field limit and thus

$$R_{00} = -\tfrac{1}{2}\partial_i\partial_i h_{00} = -\tfrac{1}{2}\nabla^2 h_{00}.$$

Equating the two expressions derived above for R_{00} gives

$$-\tfrac{1}{2}\nabla^2 h_{00} = \frac{4\pi G}{c^2}\rho,$$

and using Eq. (8.10) to eliminate h_{00} in favor of a scalar field $\varphi \equiv -\frac{1}{2}h_{00}c^2$ gives

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

This is the Poisson equation (8.1) that governs Newtonian gravitation.

8.12 Assume a general spherical metric with line element

$$ds^2 = -e^\sigma dt^2 + e^\lambda dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

where σ and λ are positive and independent of time. From Eqs. (7.10) and (3.57) the equation to be solved is

$$T^\mu_{\nu;\mu} = T^\mu_{\nu,\mu} + \Gamma^\mu_{\alpha\mu}T^\alpha_\nu - \Gamma^\alpha_{\nu\mu}T^\mu_\alpha = 0.$$

The non-zero connection coefficients required are given in the solution to Problem 8.2 and in Appendix C (but here all connection coefficients proportional to time derivatives are neglected, since the metric is static). The preceding equation corresponds to one equation for each of the four possible value of ν, with each equation involving implied sums over the repeated indices μ and α. When written out there are many terms, but a large number are identically zero because by inspection either the connection coefficient vanishes or the term is not diagonal in T^μ_ν. Collecting the terms that survive gives

$$\begin{aligned} P' &+ \Gamma^0_{10}T^1_\nu + \Gamma^1_{11}T^1_\nu + \Gamma^2_{12}T^1_\nu + \Gamma^3_{13}T^1_\nu + \Gamma^3_{23}T^2_\nu \\ &- \Gamma^0_{\nu 0}T^0_0 - \Gamma^1_{\nu 1}T^1_1 - \Gamma^2_{\nu 2}T^2_2 - \Gamma^3_{\nu 3}T^3_3 = 0, \end{aligned}$$

where a prime denotes a partial derivative with respect to r. This represents four separate equations for the respective choices $\nu = 0,1,2,3$. The only non-trivial result corresponds to setting $\nu = 1$, which gives

$$P' + (P+\rho)\frac{\sigma'}{2} = 0,$$

upon substituting the expressions for the connection coefficients and using $T_0^0 = -\rho$ and $T_1^1 = T_2^2 = T_3^3 = P$.

8.13 From the solution of Problem 8.3, the commutator of the covariant derivative is related to the curvature tensor by

$$[\nabla_\nu, \nabla_\lambda] V_\mu = R^\sigma{}_{\mu\nu\lambda} V_\sigma.$$

The covariant differentiation of vectors on the left side of this equation means that the left side is tensorial and thus the right side must also be a tensor. Hence, by the quotient theorem of Problem 3.13 the components $R^\sigma{}_{\mu\nu\lambda}$ that contract with the dual vector on the right side must be the elements of a tensor. This is a *much* faster proof than checking $R^\sigma{}_{\mu\nu\lambda}$ explicitly against the rank-4 tensor transformation law.

8.14 In a local inertial frame the connection coefficients vanish but not their first derivatives. Therefore, from Eq. (8.14) and Eq. (7.30), in this frame at an arbitrary point

$$R_{\sigma\mu\nu\lambda} = g_{\sigma\kappa} R^\kappa{}_{\mu\nu\lambda} = \tfrac{1}{2}(\partial_\lambda \partial_\sigma g_{\mu\nu} - \partial_\lambda \partial_\mu g_{\sigma\nu} + \partial_\nu \partial_\mu g_{\sigma\lambda} - \partial_\nu \partial_\sigma g_{\mu\lambda}),$$

where $\partial_\alpha = \partial/\partial x^\alpha$. The symmetries

$$R_{\sigma\mu\nu\lambda} = -R_{\mu\sigma\nu\lambda} = -R_{\sigma\mu\lambda\nu} \qquad R_{\sigma\mu\nu\lambda} = R_{\nu\lambda\sigma\mu}$$

of Eq. (8.15) follow immediately by exchanging indices and noting that the order of differentiation doesn't matter and the metric tensor is symmetric in its indices. These results were obtained in a special frame but since the relations are tensor relations they are valid in all frames.

8.15 From Eq. (8.18) the Bianchi identity is

$$\nabla_\lambda R_{\mu\nu\alpha\beta} + \nabla_\beta R_{\mu\nu\lambda\alpha} + \nabla_\alpha R_{\mu\nu\beta\lambda} = 0.$$

Contract this with $g^{\mu\alpha}$, remembering that since $\nabla_\mu g^{\mu\nu} = 0$, raising an index by contraction commutes with covariant differentiation,

$$\begin{aligned}
\nabla_\lambda g^{\mu\alpha} R_{\mu\nu\alpha\beta} + \nabla_\beta g^{\mu\alpha} R_{\mu\nu\lambda\alpha} + \nabla_\alpha g^{\mu\alpha} R_{\mu\nu\beta\lambda} &= 0 \\
\to \nabla_\lambda g^{\mu\alpha} R_{\mu\nu\alpha\beta} - \nabla_\beta g^{\mu\alpha} R_{\mu\nu\alpha\lambda} + \nabla_\alpha g^{\mu\alpha} R_{\mu\nu\beta\lambda} &= 0 \\
\to \nabla_\lambda R_{\nu\beta} - \nabla_\beta R_{\nu\lambda} + \nabla_\alpha g^{\mu\alpha} R_{\mu\nu\beta\lambda} &= 0,
\end{aligned}$$

where in the second line the last two indices were switched in the second term using Eq. (8.15) so that the contraction is consistent with the definition (8.16) of the Ricci tensor [see the footnote following Eq. (8.16) and Problem 8.8], and in the last line Eq. (8.16) was used. Now contract with $g^{\nu\beta}$ to give

$$\begin{aligned}
\nabla_\lambda g^{\nu\beta} R_{\nu\beta} - \nabla_\beta g^{\nu\beta} R_{\nu\lambda} + \nabla_\alpha g^{\mu\alpha} g^{\nu\beta} R_{\mu\nu\beta\lambda} &= 0 \\
\to \nabla_\lambda R - \nabla_\beta R^\beta{}_\lambda - \nabla_\alpha g^{\mu\alpha} g^{\nu\beta} R_{\nu\mu\beta\lambda} &= 0 \\
\to \nabla_\lambda R - \nabla_\beta R^\beta{}_\lambda - \nabla_\alpha g^{\mu\alpha} R_{\mu\lambda} &= 0 \\
\to \nabla_\lambda R - \nabla_\beta R^\beta{}_\lambda - \nabla_\alpha R^\alpha{}_\lambda &= 0 \\
\to \nabla_\lambda R - 2\nabla_\alpha R^\alpha{}_\lambda &= 0,
\end{aligned}$$

where in the second line the first two indices on R in the last term were switched to make the contraction compatible with (8.16) and the definition (8.17) was used, and in the last line the dummy summation indices were switched to the same variable so the last two terms could be added. Now contract with $g^{\mu\lambda}$ to raise the index on the last term,

$$\begin{aligned} \nabla_\lambda R g^{\mu\lambda} - 2\nabla_\alpha g^{\mu\lambda} R^\alpha{}_\lambda &= 0 \\ \to \nabla_\lambda R g^{\mu\lambda} - 2\nabla_\alpha R^{\mu\alpha} &= 0 \\ \to \nabla_\nu R g^{\mu\nu} - 2\nabla_\nu R^{\mu\nu} &= 0, \end{aligned}$$

where dummy summation indices have been switched. Finally, multiply both sides by $-\frac{1}{2}$ to give

$$\nabla_\nu (R^{\mu\nu} - \tfrac{1}{2} R g^{\mu\nu}) = 0,$$

which is $\nabla_\mu G^{\mu\nu} = 0$ for the symmetric Einstein tensor $G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu}$ defined in Eq. (8.20).

9 The Schwarzschild Spacetime

9.1 It is convenient to introduce an exponential parameterization $B(r) \equiv e^{\nu(r)}$ and $A(r) \equiv e^{\lambda(r)}$, so that the metric is

$$g_{\mu\nu} = \text{diag}\left(-e^{\nu}, e^{\lambda}, r^2, r^2 \sin^2\theta\right) \qquad g^{\mu\nu} = \text{diag}\left(-e^{-\nu}, e^{-\lambda}, \frac{1}{r^2}, \frac{1}{r^2 \sin^2\theta}\right).$$

The unknown functions $\nu(r)$ and $\lambda(r)$ may then be determined by requiring that the metric be consistent with the vacuum (vanishing stress–energy tensor) Einstein equation. However, for a vacuum solution it is not necessary to construct the full Einstein tensor $G_{\mu\nu}$ because it may be shown (see Problem 22.1) that solution of the vacuum Einstein equation is equivalent to solving $R_{\mu\nu} = 0$, where $R_{\mu\nu}$ is the Ricci tensor. The metric and Eq. (7.30) give the Christoffel symbols (affine connections) $\Gamma^{\gamma}_{\alpha\beta}$, and from these and Eqs. (8.14) and (8.16) the Ricci tensor may be constructed. Setting $R_{\mu\nu} = 0$ yields four equations, only three of which are independent. In particular,

$$\begin{aligned}
-R_{00} = e^{\nu-\lambda}\left(\tfrac{1}{2}\nu'' + \tfrac{1}{4}(\nu')^2 - \tfrac{1}{4}\nu'\lambda' + \frac{\nu'}{r}\right) &= 0 \\
R_{11} = \tfrac{1}{2}\nu'' + \tfrac{1}{4}(\nu')^2 - \tfrac{1}{4}\nu'\lambda' - \frac{\lambda'}{r} &= 0 \\
R_{22} = e^{-\lambda}\left(\tfrac{1}{2}(\nu' - \lambda')r + 1\right) - 1 &= 0,
\end{aligned}$$

where primes indicate derivatives with respect to r. The first two equations imply that $\nu' = -\lambda'$ and thus that $\nu = -\lambda + \text{constant}$. A time-independent solution is sought so the timescale may be shifted freely to make the constant zero and obtain $\nu = -\lambda$. Inserting this into the R_{22} equation gives $re^{\nu}\nu' + e^{\nu} = 1$, the left side of which is equivalent to $d(re^{\nu})/dr$, implying that $e^{\nu} = 1 - c/r$, where c is a constant. Choosing $c = 2M$, where M is another constant (that will be interpreted as the mass when compared with Newtonian gravity),

$$e^{\nu} = 1 - \frac{2M}{r} \qquad e^{\lambda} = e^{-\nu} = \left(1 - \frac{2M}{r}\right)^{-1},$$

which gives the Schwarzschild metric (9.5) when inserted into the equation for $g_{\mu\nu}$ above.

9.2 Use Eq. (9.30) to set $dV_{\text{eff}}/dr = 0$ and obtain

$$r_{\pm} = \frac{\ell^2}{2M} \pm \frac{1}{2}\sqrt{\frac{\ell^4}{M^2} - 12\,\ell^2},$$

from which the desired results follow.

9.3 The radius of the innermost stable circular orbit is $R_{\text{ISCO}} = 6M$. The second derivative of the effective potential (9.30) is

$$\frac{d^2V_{\rm eff}}{dr^2} = -\frac{2M}{r^3} + \frac{36M^2}{r^4} - \frac{144M^3}{r^5},$$

which gives zero when evaluated at $r = 6M$. Thus the ISCO is at a point of inflection in the potential and it is marginally stable.

9.4 From Eq. (9.43),

$$\begin{aligned}\delta\varphi = \frac{6\pi GM}{ac^2(1-e^2)} &= 1.398\times 10^{-27}\left(\frac{M}{\rm g}\right)\left(\frac{\rm cm}{a}\right)\frac{1}{1-e^2}\ \text{rad/orbit}\\ &= 1.858\times 10^{-7}\left(\frac{M}{M_\odot}\right)\left(\frac{\rm AU}{a}\right)\frac{1}{1-e^2}\ \text{rad/orbit},\end{aligned}$$

Some results calculated with this formula are given in the following table.

Object	Central mass ($M_\odot$)	a (AU)	e	Period	$\delta\varphi/\Delta t$
Mercury	1	0.387	0.206	88 d	43″/century
Earth	1	1	0.017	1 yr	3.8″/century
Binary Pulsar*	2.828	1.3×10^{-2}	0.617	7.75 hr	4.2°/yr

*Mass is total mass of binary; a is average separation

9.5 From Eq. (9.23), $\ell = r^2\,(d\varphi/d\tau)$ if $\theta = \frac{\pi}{2}$ is assumed, and for classical orbital motion

$$r^2\frac{d\varphi}{dt} = \sqrt{1-e^2}\left(\frac{2\pi}{P}\right)a^2.$$

Combining these equations assuming weak gravity gives

$$\ell^2 = \left(r^2\frac{d\varphi}{d\tau}\right)^2 \simeq \left(r^2\frac{d\varphi}{dt}\right)^2 = \frac{4\pi^2}{P^2}(1-e^2)a^4 = (1-e^2)GMa,$$

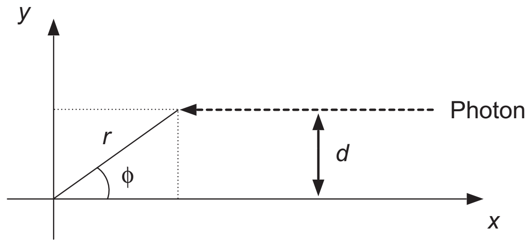
where in the last step $P^2 = (4\pi^2/GM)a^3$ (Kepler III) was used. Inserting this result in Eq. (9.42) then gives Eq. (9.43).

9.6 Differentiate the expression for $V_{\rm eff}$ in Eq. (9.55) with respect to r, set to zero, and solve to give $r = 3M = \frac{3}{2}r_{\rm s}$ for the maximum of the potential, implying an unstable circular orbit at this radius. Substitute this value of r to give

$$V_{\rm eff}(r=3M) = \frac{1}{27M^2}$$

for the height at the maximum.

9.7 For the following diagram,



define polar coordinates in the $\theta = \pi/2$ plane,

$$x = r\cos\varphi \qquad y = r\sin\varphi.$$

Then the parameter b is

$$\begin{aligned} b = \left|\frac{\ell}{\varepsilon}\right| &= \frac{r^2\sin^2\theta d\varphi/d\lambda}{(1-2M/r)dt/d\lambda} \\ &= \frac{r^2}{(1-2M/r)}\frac{d\varphi}{dt} \\ &= \frac{r^2}{(1-2M/r)}\frac{d\varphi}{dr}\frac{dr}{dt}, \end{aligned}$$

where Eqs. (9.53)–(9.55) have been used. But for $r \gg 2M$ in the Schwarzschild metric,

$$1 - \frac{2M}{r} \to 1 \qquad \varphi \to \frac{d}{r} \qquad \frac{dr}{dt} \to -1.$$

Therefore, $d\varphi/dr \simeq -d/r^2$ and

$$b = \frac{r^2}{(1-2M/r)}\frac{d\varphi}{dr}\frac{dr}{dt} = r^2\left(\frac{-d}{r^2}\right)(-1) = d,$$

so b may be interpreted as the impact parameter for the photon.

9.8 From the static Schwarzschild metric restricted to radial coordinates ($dt = d\theta = d\varphi = 0$), the coordinate distance dr is related to the proper distance ds by $dr = (1-2M/r)^{1/2}\,ds$. In geometrized units $M = 2.95$ km and $R = 7\times 10^5$ km for the Sun. Since the rod is very short the gravitational field may be assumed to be uniform over its length, giving the estimate

$$\Delta r \simeq \left(1 - \frac{2M}{r}\right)^{1/2}\Delta s \;\simeq\; 0.9999958\,\text{cm}$$

for the coordinate length of the rod at the surface of the Sun. A neutron star packs of order a solar mass into a radius of roughly 10 km. Thus for a neutron star $2M/r \simeq 0.59$ and $\Delta r \simeq 0.64$ cm, indicating a significant distortion of the metric from that of flat space at its surface.

9.9 Compare the line element

$$ds^2 = -(1-2M/r)dt^2 + \alpha(\rho)^2\left(d\rho^2 + \rho^2 d\theta^2 + \rho^2\sin^2\theta d\varphi^2\right),$$

term by term with the standard Schwarzschild line element (9.4). Requiring the coefficients of $d\theta^2+\sin^2\theta d\varphi^2$ to be equal in the two expressions gives $r^2=\alpha^2\rho^2$ and equating the radial terms requires $(1-2M/r)^{-1}dr^2=\alpha^2 d\rho^2$. Combining these last two expressions gives

$$\frac{d\rho}{\rho}=\pm\frac{dr}{\sqrt{r^2-2Mr}}.$$

Choose the positive sign by assuming $\rho\to\infty$ for $r\to\infty$ and integrate to give

$$\ln\rho=\ln\left(-M+r+\sqrt{r^2-2Mr}\right)+\ln b,$$

where $\ln b$ is an arbitrary integration constant. Choose the integration constant such that $\rho(r=2M)=M/2$ (see Problem 9.10 for a justification of this choice) to yield $r=\rho(1+M/2\rho)^2$ and thus $\alpha^2=(1+M/2\rho)^4$. Inserting these in the first expression above and a bit of algebra gives the isotropic form of the Schwarzschild line element,

$$ds^2=-\frac{(1-M/2\rho)^2}{(1+M/2\rho)^2}dt^2+(1+M/2\rho)^4\left(d\rho^2+\rho^2(d\theta^2+\sin^2\theta d\varphi^2)\right).$$

This form has the merit of making time slices $t=$ constant look as much as possible like euclidean space.

9.10 Solve the quadratic equation $r=\rho(1+M/2\rho)^2$ derived in Problem 9.9 for ρ to give

$$\rho=\frac{-(M-r)\pm\sqrt{(M-r)^2-M^2}}{2},$$

from which we may conclude

1. The variable ρ is complex if $0\le r<2M$, so the region inside the Schwarzschild radius $r_s=2M$ is not in the real domain of ρ.
2. If $r>2M$, each r corresponds to two real positive values of ρ, so this region of r-space is mapped out *twice* in ρ-space, once in the region $M/2\le\rho<\infty$ and once in the region $0\le\rho\le M/2$.

This second point is the reason for the choice of boundary condition $\rho(r=2M)=M/2$ in Problem 9.9.

9.11 This problem is based on an example discussed in *Invitation to Astrophysics Astrophysics*, T. Padmanabhan (World Scientific) 2006. The coordinate systems are related by $d\bar{t}=dT$ and $d\bar{\boldsymbol{r}}=d\boldsymbol{r}-\boldsymbol{v}dT$, which when substituted into the metric $ds^2=-d\bar{t}^2+d\bar{\boldsymbol{r}}^2$ expressed in the freely-falling coordinates gives

$$ds^2=-\left(1-\frac{2M}{r}\right)dT^2-2\sqrt{\frac{2M}{r}}\,dr\,dT+dr^2+r^2d\theta^2+r^2\sin^2\theta d\varphi^2.$$

Introducing a new time coordinate t through the linear transformation

$$dT=dt-\frac{\sqrt{2M/r}}{1-2M/r}dr$$

and a little algebra eliminates the non-diagonal term and yields the Schwarzschild line element of Eq. (9.4).

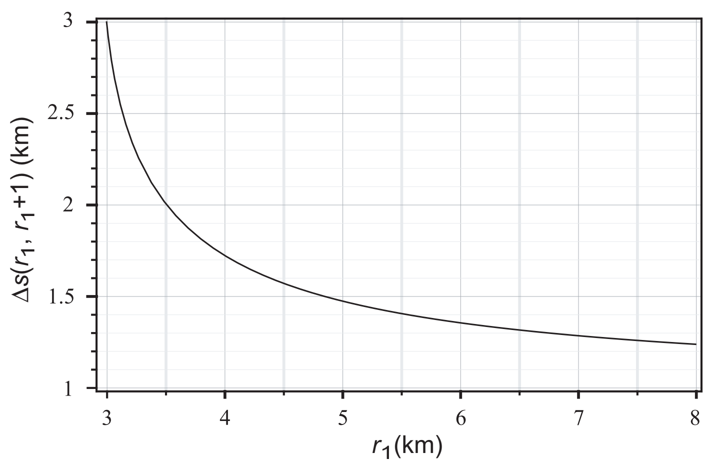
9.12 This problem is adapted from an example in Taylor and Wheeler [229]. Integrating the relationship (9.7) between the proper interval ds and the coordinate radial interval dr gives

$$\begin{aligned}\Delta s &= \int \left(1-\frac{2M}{r}\right)^{-1/2} dr = \int \frac{r^{1/2}dr}{\sqrt{r-2M}} = 2\int \frac{z^2 dz}{\sqrt{z^2-2M}} \\ &= z\sqrt{z^2-2M} + 2M\ln\left(z+\sqrt{z^2-2M}\right),\end{aligned}$$

where the substitution $r = z^2$ with $dr = 2zdz$ has been used. Rewriting in terms of r and evaluating between the two spherical shell radii gives finally

$$\Delta s = \left[\sqrt{r(r-2M)} + 2M\ln\left(\sqrt{r}+\sqrt{r-2M}\right)\right]\Big|_{r_1}^{r_2}.$$

For a 1 $M_\odot$ black hole $2M = 2.954$ km in geometrized units and for spherical shells with $r_1 = 5$ km and $r_2 = 6$ km, this formula yields $\Delta s = 1.474$ km for the physical distance between them. This is much larger than the coordinate difference of 1 km, indicating a large distortion of spacetime relative to a flat metric (see Fig. 9.1). The following figure illustrates the variation of the proper radial separation between spherical shells with radial coordinates r_1 and $r_2 = r_1 + 1$ km as a function of r_1 around a 1 solar mass black hole.



At large distance the proper separation approaches the coordinate separation of 1 km, but near the event horizon at $r \sim 2.954$ km the curvature is very large and the proper separation is much larger than the coordinate separation.

9.13 Write the integral of Eq. (9.49) in the form

$$t = t_0 - \frac{1}{\sqrt{2M}}\int \frac{r^{3/2}dr}{r-2M},$$

where t_0 is an integration constant. Then substitute $r = x^2$ and apply tabulated results for integrals of the form

$$\int \frac{x^m dx}{ax^2+bx+c}$$

iteratively to do the integration. Substitution of $x = \sqrt{r}$ and some algebra then gives Eq. (9.50).

9.14 For simplicity, take the circle to lie in the $\theta = \frac{\pi}{2}$ plane, centered on $r = 0$. Then $dt = dr = d\theta = 0$ and $\sin^2\theta = 1$. The Schwarzschild line element (9.4) then reduces to $ds^2 = r^2 d\varphi^2$ and the proper circumference C is

$$C = \int ds = \int_0^{2\pi} r d\varphi = 2\pi r.$$

This is formally the same result as for flat space. However C is a physical (proper) distance since it was computed from the metric, but r is a coordinate and *not* a physical distance. Conversely, in flat space both C and r in $C = 2\pi r$ can be interpreted as physical distances.

9.15 (a) Combine Eqs. (9.30) and (9.29) with the requirement that $V_{\text{eff}} = E$ to give Eq. (9.32).

(b) The angular velocity is given by

$$\Omega = \frac{d\varphi}{dt}.$$

Substitute Eqs. (9.22) and (9.23) to give Eq. (9.33).

(c) From Eqs. (9.25), (9.5), and (9.35),

$$-\left(1 - \frac{2M}{r}\right)(u^t)^2 + r^2(u^t)^2\Omega^2 = -1,$$

This may be solved using Eq. (9.34) to give,

$$u^t = \left(1 - \frac{3M}{r}\right)^{-1/2},$$

which is Eq. (9.36).

9.16 Since $u = (u^t, 0, 0, u^t\Omega)$ and $s \cdot u = 0$, utilizing the metric (9.5),

$$s \cdot u = g_{\mu\nu} s^\mu u^\nu = -\left(1 - \frac{2M}{R}\right) s^t u^t + R^2 s^\varphi u^\varphi = 0,$$

from which

$$s^t = \frac{R^2 s^\varphi u^\varphi}{(1 - 2M/R)u^t} = \frac{R^2\Omega}{1 - 2M/R} s^\varphi,$$

where Eqs. (9.35) and (9.36) was used in the last step. This is Eq. (9.61).

9.17 (i) Our solution follows that of Hartle [110], Ch. 14. Utilizing

$$u = (u^t, 0, 0, u^t\Omega) \qquad s = (s^t, s^r, 0, s^\varphi),$$

the only non-vanishing combinations for the sums in Eq. (9.62) are $\Gamma^r_{tt} s^t u^t$ and $\Gamma^r_{\varphi\varphi} s^\varphi u^\varphi$. From Appendix C, for a general spherical metric with line element

$$ds^2 = -e^\sigma dt^2 + e^\lambda dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

the required connection coefficients are

$$\Gamma^r_{tt} = \Gamma^1_{00} = \tfrac{1}{2} e^{\sigma-\lambda}\sigma' \qquad \Gamma^r_{\varphi\varphi} = \Gamma^1_{33} = -re^{-\lambda}\sin^2\theta,$$

where $\sigma' \equiv d\sigma/dr$. Specializing to the Schwarzschild metric with $\theta = \frac{\pi}{2}$ and letting $r = R$,

$$e^\sigma = \left(1 - \frac{2M}{R}\right) \qquad e^\lambda = \left(1 - \frac{2M}{R}\right)^{-1} \qquad \sigma = \ln\left(1 - \frac{2M}{R}\right)$$
$$\sigma' = \frac{d\sigma}{dr} = \frac{2M}{R^2}\left(1 - \frac{2M}{R}\right)^{-1} \qquad \sin^2\theta = 1,$$

and the connection coefficients evaluate to

$$\Gamma^r_{tt} = \frac{M}{R^2}\left(1 - \frac{2M}{R}\right) \qquad \Gamma^r_{\varphi\varphi} = -(R - 2M)$$

Therefore Eq. (9.62) becomes

$$\frac{ds^r}{d\tau} + \frac{M}{R^2}\left(1 - \frac{2M}{R}\right)s^t u^t - (R - 2M)s^\varphi u^\varphi = 0.$$

Utilizing $u^t = dt/d\tau$ and Eq. (9.61), this may be written as

$$\frac{ds^r}{d\tau} - (R - 3M)\Omega s^\varphi = 0.$$

(ii) Now consider Eq. (9.63). By similar considerations as above, the only nonvanishing contribution to the sum in the second term is $\Gamma^\varphi_{r\varphi} s^r u^\varphi$. From Appendix C the single required connection coefficient is $\Gamma^\varphi_{r\varphi} = r^{-1}$ and Eq. (9.63) becomes

$$\frac{ds^\varphi}{dt} + \frac{\Omega}{R}s^r = 0,$$

where $u^t = dt/d\tau$ and $u^\varphi = \Omega u^t$ from Eq. (9.35) were used (with $\Omega = d\varphi/dt$).

(iii) Thus the two equations

$$\frac{ds^r}{d\tau} - (R - 3M)\Omega s^\varphi = 0 \qquad \frac{ds^\varphi}{dt} + \frac{\Omega}{R}s^r = 0$$

must be solved simultaneously. Take d/dt of the second equation and plug in ds^r/dt from the first equation to give

$$\frac{d^2 s^\varphi}{dt^2} + \omega^2 s^\varphi = 0,$$

where $\omega \equiv (1 - 3M/R)^{1/2}\Omega$. This is the equation of a harmonic oscillator and the solutions are given by Eq. (9.64).

9.18 From Eq. (9.68) evaluated assuming a satellite in orbit around the Earth, the geodetic precession per orbit is

$$\Delta\varphi \simeq \frac{3\pi GM}{c^2 R} = 8.6\left(\frac{\text{km}}{R}\right) \text{arcsec orbit}^{-1}.$$

For GP-B, setting $R = a = 7027.4\,\mathrm{km}$ gives 1.22×10^{-3} arcsec orbit^{-1}. By Kepler's 3rd law, the period for a circular satellite orbit around the Earth is

$$P = \sqrt{\frac{4\pi^2 a^3}{GM_\oplus}} = 9.95 \times 10^{-3} \left(\frac{a}{\mathrm{km}}\right)^{3/2} \mathrm{s}.$$

For GP-B with $a = 7027.4\,\mathrm{km}$ this gives a period of 97.7 minutes, which translates to 5.38×10^3 orbits yr^{-1}. Therefore general relativity predicts that GP-B should exhibit a geodetic precession rate of

$$\frac{\Delta\varphi}{\Delta t} = (1.22 \times 10^{-3}\,\mathrm{arcsec\,orbit^{-1}}) \times (5.38 \times 10^3\,\mathrm{orbit\,yr^{-1}}) \simeq 6.6\,\mathrm{arcsec\,yr^{-1}}.$$

As discussed in Box 9.3, GP-B measured a geodetic precession rate for its gyroscopes that was within 0.07% of this value.

9.19 The precession rate is given by Eq. (9.76). The angular momentum of the Earth can be expressed as $J_\oplus = \mathscr{I}_\oplus \Omega_\oplus$, where $\mathscr{I}_\oplus$ is the Earth's moment of inertia for rotation about the polar axis and $\Omega_\oplus$ is the angular velocity of the Earth. In planetary science the moment of inertia is parameterized as $\mathscr{I} = kMR^2$, where M is the mass, R is the radius, and k indicates how much the interior mass distribution differs from uniform (for example, a completely uniform sphere has $k = \frac{2}{5} = 0.4$, but Saturn with a centrally-concentrated mass distribution has $k = 0.21$). For the Earth (and other terrestrial planets) $k \simeq 0.33$, so Earth's moment of inertia is

$$\mathscr{I}_\oplus = 0.33 M_\oplus R_\oplus^2 = 8.06 \times 10^{44}\,\mathrm{g\,cm^2},$$

the Earth's angular velocity is

$$\Omega_\oplus = \frac{2\pi}{24\,\mathrm{h}r} = 0.262\,\mathrm{rad\,hr^{-1}} = 7.27 \times 10^{-5}\,\mathrm{rad\,s^{-1}},$$

and thus the Earth's angular momentum is $J = \mathscr{I}_\oplus \Omega_\oplus = 5.86 \times 10^{40}\,\mathrm{g\,cm^2\,s^{-1}}$. Hence the Lense–Thirring precession rate for a gyroscope in free fall on Earth's rotation axis is

$$\Omega_{\mathrm{LT}} = \frac{2GJ}{c^2 z^3} = 5.65 \times 10^{10} \left(\frac{1\,\mathrm{km}}{z}\right)^3 \mathrm{arcsec\,yr^{-1}}.$$

For Gravity Probe B illustrated in Box 9.3, the semimajor axis of the nearly circular polar orbit was 7027.4 km. As the satellite passes over the North Pole it is in free fall with $z =$ 7027.4 km. Inserting this in the above equation gives a precession rate of 0.16 arcsec yr^{-1}. The Lense–Thirring precession rate depends on the latitude of a satellite in polar orbit, so the smaller general-relativistic prediction of 0.039 arcsec yr^{-1} per year shown in Box 9.3 represents an average over the satellite in polar orbit, which is less than our evaluation on the z axis. At any rate, this Lense–Thirring precession is a much smaller effect than the geodetic precession of more than 6 arcseconds per year.

10 Neutron Stars and Pulsars

10.1 We work in $c = G = 1$ units (Appendix B.1). The Schwarzschild radius is $r_s = 2M$, where M is the mass. Assuming gravity to pack neutrons down to a hardcore radius $r_0 \simeq 0.5 \times 10^{-13}$ cm, the radius of the neutron star will be $R \simeq r_0 A^{1/3}$, where A is the number of neutrons in the entire star (related to the total mass by $M \simeq Am$, where $m = 939\,\text{MeV} = 1.2 \times 10^{-52}$ cm is the neutron mass). Equating the neutron star radius to the Schwarzschild radius implies that $r_0 A^{1/3} = 2M$, from which M may be eliminated using $M = Am$ to give $A = (r_0/2m)^{3/2} \simeq 3 \times 10^{57}$ neutrons in a typical neutron star. Thus, by these simple considerations one estimates that the radius is $R = r_0 A^{1/3} \simeq 7$ km, the mass is $M = R/2 = 3.5\,\text{km} = 2.4 M_\odot$, and the average density is

$$\bar{\rho} = M/\tfrac{4}{3}\pi R^3 \simeq 0.0024\,\text{km}^{-2} = 3.2 \times 10^{15}\,\text{g}\,\text{cm}^{-3},$$

which is larger than nuclear matter density (about $2.5 \times 10^{14}\,\text{g}\,\text{cm}^{-3}$). More realistic estimates for actual neutron stars give similar numbers: a radius of ~ 10 km, a mass of about $\sim 1.5 M_\odot$, and an average density of $\sim 10^{15}\,\text{g}\,\text{cm}^{-3}$.

10.2 From Eq. (10.15),

$$M(r) = 4\pi \int_0^r \varepsilon(r) r^2\, dr \simeq \tfrac{4}{3}\pi \bar{\varepsilon} r^3,$$

where an average energy density $\bar{\varepsilon} = M(R)/\tfrac{4}{3}\pi R^3$ has been assumed. The gravitational energy is then

$$\begin{aligned} E_g &= \int_0^R \frac{M(r)}{r}\, dM(r) = \int_0^R \frac{M(r)(4\pi r^2 \bar{\varepsilon})}{r}\, dr \\ &= \frac{16}{3}\pi^2 \bar{\varepsilon}^2 \int_0^R r^4\, dr = \frac{3}{5}\frac{M^2}{R}, \end{aligned}$$

where $dM(r)/dr = 4\pi r^2 \bar{\varepsilon}$ was used to change integration variables and the preceding expressions for $M(r)$ and $\bar{\varepsilon}$ have been inserted. Assuming a neutron star with $M = 2M_\odot = 2.95$ km and $R = 10$ km (in geometrized units) gives

$$E_g = 0.52\,\text{km} = 6.3 \times 10^{53}\,\text{erg} = 4 \times 10^{59}\,\text{MeV} = 7 \times 10^{32}\,\text{g} = 0.35\, M_\odot.$$

The ratio of the gravitational energy to the rest mass energy is $E_g/M = (0.52\,\text{km}/2.95\,\text{km}) \simeq 0.18$, so the gravitational energy is a significant fraction of the rest mass energy and GR corrections to Newtonian gravity are expected to be important for neutron stars.

10.3 Neglecting the pressure term in the numerator and $2m$ relative to r in the denominator (weak gravity assumption) in Eq. (10.13) gives $d\varphi/dr \sim m/r^2$ in the Newtonian limit, where $2\varphi \equiv \sigma$. But this means that φ is just the Newtonian gravitational potential in $G = 1$

units (see Section 8.1). Thus, σ is proportional to the gravitational potential in the Newtonian limit. This interpretation is strengthened by substituting $\varphi = \frac{1}{2}\sigma$ and neglecting P relative to ρ in Eq. (10.8) to give

$$\frac{dP}{dr} = -\frac{1}{2}(P+\rho)\frac{d\sigma}{dr} = -(P+\rho)\frac{d\varphi}{dr} \simeq -\rho\frac{d\varphi}{dr} = -\frac{Gm\rho}{r^2},$$

where the last step follows from substituting $d\varphi/dr \sim Gm/r^2$. This is the pressure-balance equation in Newtonian hydrostatics for a gravitational potential $\varphi = \frac{1}{2}\sigma$.

10.4 Solving the differential form of Eq. (10.15) for dr, substitution in Eq. (10.14), and a little algebra gives Eq. (10.16).

10.5 From Eqs. (10.2) and (10.11) the radial metric component for the Oppenheimer–Volkov solution is given by

$$g_{11}(r) = \left(1 - \frac{2M(r)}{r}\right)^{-1},$$

which is unity at the center where $M(r) = 0$. Thus the ratio of g_{11} at the surface to that at the center is

$$\frac{g_{11}(R)}{g_{11}(0)} = \left(1 - \frac{2M(R)}{R}\right)^{-1} \simeq 2.5,$$

where we've assumed a neutron star of radius 10 km and mass $M = 2M_\odot \simeq 3\,\text{km}$. Since the average spacing between neutrons is $\sim 10^{-13}$ cm, this means that the metric changes by only of order one part in 10^{19} over the internucleonic spacing. Thus, on that distance scale the metric is very flat. See Glendenning [98], Section 4.4 for further discussion.

10.6 From Problem 10.2 the total gravitational binding energy is about 4×10^{59} MeV and from Problem 10.1 the total number of nucleons is about 3×10^{57}. Thus, from the ratio of these two numbers it may be estimated roughly that the gravitational binding energy per nucleon in a neutron star is about 133 MeV nucleon^{-1}. This is much larger than the nuclear binding energy per nucleon, which is about 8 MeV per nucleon for heavy nuclei and about 16 MeV per nucleon for extended symmetric nuclear matter.

10.7 Outside a spherical neutron star the metric should be well approximated by the Schwarzschild form. In the Schwarzschild metric the escape velocity is $v_{\rm esc} = (2M/R)^{1/2}$. The exact relationship between the mass and radius of a neutron star depends on the equation of state but if $M = 1M_\odot = 1.48\,\text{km}$ and $R = 12\,\text{km}$ one obtains $v_{\rm esc}/c = 0.5$, while if $M = 2M_\odot$ and $R = 10\,\text{km}$ then $v_{\rm esc}/c = 0.77$. Thus, any realistic choice of M and R for a neutron star will give an escape velocity that is a significant fraction of the speed of light. This is a signal that the gravitational field is very strong and general relativistic effects are significant. On the other hand, assuming for a white dwarf that $M = 1M_\odot$ and $R = 5000\,\text{km}$ gives an escape velocity of $v_{\rm esc}/c \simeq 0.024$ and general relativistic effects for a white dwarf are small (but not completely negligible).

10.8 From Eq. (6.5),

$$\varepsilon = \frac{GM}{Rc^2} = 7.416 \times 10^{-31}\left(\frac{M}{\text{kg}}\right)\left(\frac{\text{km}}{R}\right) = 1.475\left(\frac{M}{M_\odot}\right)\left(\frac{\text{km}}{R}\right),$$

where M is the mass producing the gravitational field and R is the characteristic distance over which it acts. The innermost white dwarf has a mass of $\sim 0.2M_\odot$, for which a radius estimate is $\sim 15,000\,\mathrm{km}$. For the neutron stars let's take a mass of $1.4M_\odot$ and radius of $\sim 10\,\mathrm{km}$. Then

$$\varepsilon_{\mathrm{WD}} \sim 2\times10^{-5} \qquad \varepsilon_{\mathrm{NS}} \sim 0.2 \qquad \varepsilon_{\mathrm{Earth}} \sim 7\times10^{-10} \qquad \varepsilon_{\mathrm{Moon}} \sim 3\times10^{-11},$$

confirming the assertion that any deviations from the strong equivalence principle would be greatly amplified in PSR J0337+1715 relative to the Solar System. See also the related Problem 25.2.

11 Spherical Black Holes

11.1 Begin with Eq. (11.9). For the $r > 2M$ case, $\ln|r/2M - 1| = \ln(r/2M - 1)$ and differentiating (11.9) gives

$$dt = dv - \left(1 - \frac{2M}{r}\right)^{-1} dr.$$

For the $r < 2M$ case, $\ln|r/2M - 1| = \ln(1 - r/2M)$ and differentiating (11.9) gives the same result as above. Therefore,

$$dt^2 = dv^2 - 2\left(1 - \frac{2M}{r}\right)^{-1} dv dr + \left(1 - \frac{2M}{r}\right)^{-2} dr^2.$$

Inserting this expression for dt^2 in the Schwarzschild line element (9.4) expressed in standard coordinates gives the Schwarzschild line element in Eddington–Finkelstein coordinates (11.10).

11.2 From Eq. (11.6), the proper time to fall from $r = 2M$ to $r = 0$ is $\tau = 4M/3$, in geometrized units. Restoring the G and c factors, $\tau \to c\tau$ and $M \to GM/c^2$, gives

$$\tau = \frac{4G}{3c^3}M = 6.6 \times 10^{-6}\left(\frac{M}{M_\odot}\right) \text{ seconds.}$$

Some times estimated from this formula to fall from the event horizon to the singularity of a spherical black hole are shown in the following table.

Type of black hole	Mass ($M_\odot$)	Time (s)
Typical stellar black hole	10	6.6×10^{-5}
GW150914 final black hole	62	4.1×10^{-4}
Milky Way central black hole	4.3×10^6	28.4
AGN central engine	10^9	6555

See also the related Problem 11.6.

11.3 For a Schwarzschild black hole of mass M the area of the horizon is $A = 4\pi(2M)^2 = 16\pi M^2$ [see Problem 11.9 and Eq. (13.11)]. Split the mass between two black holes $M \to \lambda M + (1-\lambda)M$, with λ varying from 0 to 1. Then the total horizon area for the two resulting black holes is $A' = 16\pi\left(\lambda^2 M^2 + (1-\lambda)^2 M^2\right)$ and

$$\frac{A'}{A} = \frac{16\pi\left(\lambda^2 M^2 + (1-\lambda)^2 M^2\right)}{16\pi M^2} = 2\lambda^2 - 2\lambda + 1,$$

which is less than one except for $\lambda = 0$ or $\lambda = 1$. Thus, any split into two finite black holes decreases the total horizon area and is forbidden by the area theorem.

11.4 No, because of the Hawking area theorem. The mass is dropped radially and is assumed uncharged, so no angular momentum or charge is added to the black hole and the resulting configuration must settle down eventually to a new Schwarzschild black hole. But for the original black hole the horizon area is $A = 16\pi M^2$ and this can never decrease in any (non-quantum) physical process. If gravitational waves were emitted they would carry away energy, and if this exceeded the amount of energy dropped in with the new mass the horizon area would have to decrease, violating the area theorem.

11.5 This problem is adapted from one in Ref. [64]. The Lagrangian is

$$L = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = \left(\frac{ds}{d\tau}\right)^2,$$

where $\dot{x}^\mu \equiv dx^\mu/d\tau$. Using the Schwarzschild metric defined in Eq. (9.5) and using the spherical symmetry to choose $\theta = \frac{\pi}{2}$ so that $d\theta = 0$ gives

$$L = -\left(1-\frac{2M}{r}\right)c^2\dot{t}^2 + \left(1-\frac{2M}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\varphi}^2 = -c^2,$$

where $L = -c^2$ [since $L = (ds/d\tau)^2$ and $ds^2 = -c^2 d\tau^2$] was used. A flat-space limit may be obtained by setting $2M/r = 0$, giving

$$L_{\text{flat}} = -c^2\left(\frac{dt}{d\tau}\right)^2 + \left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{d\varphi}{d\tau}\right)^2 = -c^2.$$

Utilizing that

$$dt = \gamma d\tau \qquad \gamma = \frac{1}{\sqrt{1-v^2/c^2}} \qquad v^2 = \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\varphi}{dt}\right)^2,$$

where $\boldsymbol{v}$ is the 3-velocity, this may be expressed as $-c^2\gamma^2 + \gamma^2 v^2 = -c^2$. Finally, multiply both sides by m^2c^2 and use that the relativistic energy is $E = \gamma mc^2$ and the relativistic momentum is $\boldsymbol{p} = \gamma m\boldsymbol{v}$ (see Section 5.2) to write this in the form $E^2 = p^2c^2 + m^2c^4$.

11.6 From Eq. (11.6), the proper time in free fall from the horizon to the singularity is (see the solution of Problem 11.2)

$$\tau = \frac{4G}{3c^3}M = 6.6\times 10^{-6}\left(\frac{M}{M_\odot}\right) \text{ seconds.}$$

Substituting $\tau = 1$ day and solving for M gives $M = 1.3\times 10^{10} M_\odot$.

11.7 The expression for Ω follows directly from Eqs. (9.22) and (9.23):

$$\varepsilon = \left(1-\frac{2M}{r}\right)\frac{dt}{d\tau} \qquad \ell = r^2\sin^2\theta\frac{d\varphi}{d\tau},$$

which may be combined assuming $\theta = \frac{\pi}{2}$ to give

$$\Omega = \frac{d\varphi}{dt} = \frac{d\varphi/d\tau}{dt/d\tau} = \frac{1}{r^2}\left(1-\frac{2M}{r}\right)\left(\frac{\ell}{\varepsilon}\right).$$

Note that this expression gives an angular velocity with respect to the Schwarzschild coordinate time, not with respect to the proper time (which is more complicated). For a circular orbit the variable E in Eq. (9.28) must be equal to the potential $V_{\text{eff}}(r)$ evaluated at the radial coordinate $r = r_0$ that gives the (single) minimum of the potential. Equations (9.29) and (9.30):

$$E = \frac{\varepsilon^2 - 1}{2} \qquad V_{\text{eff}}(r) = \frac{1}{2}\left[\left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} + 1\right) - 1\right],$$

then lead to the requirement

$$\varepsilon = \sqrt{\left(1 - \frac{2M}{r}\right)\left(1 + \frac{\ell^2}{r^2}\right)}$$

for a circular orbit.

11.8 The Schwarzschild radius is $r_{\text{S}} = 2.953(M/M_\odot)\,\text{km} = 8.86 \times 10^6$ km. From Problem 9.14, the coordinate radius r and proper circumference C are related by

$$r = \frac{C}{2\pi} = \frac{6.283 \times 10^8}{2\pi} = 10^8\,\text{km}.$$

If one now travels radially inward from $r = 10^8$ km to $r = 10^7$ km, the proper distance covered Δs is found in Problem 9.14 to be

$$\Delta s = \int_{r_1}^{r_2}\left(1 - \frac{r_{\text{S}}}{r}\right)^{-1/2} dr = \int_{10^7}^{10^8}\left(1 - \frac{8.86 \times 10^6}{r}\right)^{-1/2} dr = 1.0565 \times 10^8\,\text{km},$$

where the integral was done numerically (for example, with Wolfram Alpha online,[1] Maple, or Mathematica). If one then traverses a circle at this coordinate radius, the proper distance covered is $C = 2\pi r = 2\pi \times (10^7\,\text{km}) = 6.283 \times 10^7$ km.

11.9 From the discussion of invariant integration in Section 3.13.1, the area of the Schwarzschild horizon is

$$A = \int_0^{2\pi} d\varphi \int_0^{\pi} \sqrt{\det g}\, d\theta,$$

where the metric g for the horizon surface is 2-dimensional, corresponding to the Schwarzschild line element (9.4) evaluated at constant time and constant $r = r_{\text{s}}$:

$$ds^2 = r_{\text{s}}^2 d\theta^2 + r_{\text{s}}^2 \sin^2\theta\, d\varphi^2.$$

Thus the metric is specified by the diagonal 2×2 matrix $g = \text{diag}\,(r_{\text{s}}^2, r_{\text{s}}^2 \sin^2\theta)$, which has $\det g = r_{\text{s}}^4 \sin^2\theta$. Substituting in the above expression for A then gives a horizon area

$$A = r_{\text{s}}^2 \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta = 4\pi r_{\text{s}}^2 = 16\pi M^2,$$

where $r_{\text{s}} = 2M$ was used.

11.10 The Ricci curvature scalar R is the complete contraction of the Riemann curvature tensor for the corresponding space [see Eqs. (8.14)–(8.17)]. Assuming the surface of the

[1] www.wolframalpha.com/examples/mathematics/calculus-and-analysis/

Earth to have the metric of a 2-sphere, the 2-dimensional Riemann curvature tensor reduces to the Gaussian curvature $R = 2/r^2$, where r is the radius of the Earth (see Problem 7.1). The Ricci scalar is identically zero for the Schwarzschild metric, which is a solution of the vacuum Einstein equation, $R_{\mu\nu} = 0$. The more appropriate scalar measure of curvature is the Kretschmann scalar,

$$K \equiv R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta},$$

which is derived for general black hole solutions in Ref. [116]. The non-vanishing components of the Riemann tensor for the Schwarzschild metric evaluated in an orthonormal basis are (see Appendix B of [110])

$$R_{0101} = \frac{-2M}{r^3} \qquad R_{0202} = R_{0303} = \frac{M}{r^3} \qquad R_{2323} = \frac{2M}{r^3} \qquad R_{1212} = R_{1313} = \frac{-M}{r^3}$$

and evaluation of the Kretschmann curvature scalar for the Schwarzschild metric gives $K = 48M^2/r^6$ [116]. Thus for the poles of the Earth and at $r = 2M$ in the Schwarzschild geometry the components of the Riemann tensor and the corresponding curvature are finite and smoothly varying, suggesting that these singularities are coordinate and not physical singularities. Indeed in both cases explicit transformations to new coordinates are known that remove them. But at $r = 0$ in the Schwarzschild spacetime the Kretschmann curvature scalar goes to infinity. Since R is a scalar, it cannot be transformed to a finite value by a new choice of coordinates, suggesting that $r = 0$ is a physical, not coordinate, singularity in the Schwarzschild metric.

11.11 In cartesian coordinates the line element is $ds^2 = dx^2 + dy^2 + dz^2$ and the equation of the sphere is $x^2 + y^2 + z^2 = R^2$. Differentiate the equation of the sphere, solve for dz, and substitute into the metric to give

$$ds^2 = dx^2 + dy^2 + \frac{(xdx + ydy)^2}{R^2 - (x^2 + y^2)}$$

for the line element on the 2-dimensional spherical surface. Introduce new coordinates (r, φ) using $x = r\cos\varphi$ and $y = r\sin\varphi$ so that

$$dx = -r\sin\varphi d\varphi + \cos\varphi dr \qquad dy = r\cos\varphi d\varphi + \sin\varphi dr.$$

Substitution in the above equation for ds^2 and some algebra then gives

$$ds^2 = \frac{R^2 dr^2}{R^2 - r^2} + r^2 d\varphi^2$$

for the line element of the 2-sphere. This diverges at $r = R$, but the Gaussian curvature is constant over the whole sphere (Problem 7.1) and the sphere can be parameterized in other coordinates that eliminate this singularity, so this clearly is a coordinate singularity caused by a choice of variables with a restricted domain of validity. (Generally it may be shown that no single coordinate system is valid over the entire 2-sphere; a minimum of two overlapping coordinate patches are required to parameterize the entire surface.)

11.12 This result is indicated by the lightcone diagrams inside and outside the horizon shown in Chapter 11, but can be formalized concisely by requiring that for a stationary observer $dr = d\theta = d\varphi = 0$. Then from the metric

$$ds^2 = g_{00}dt^2 = -\left(1 - \frac{r_S}{r}\right) dt^2.$$

But g_{00} is positive inside the horizon r_S and negative outside it. Therefore ds^2 is positive (spacelike) inside the horizon and negative (timelike) outside the horizon. Since causality demands that the trajectories of particles be timelike, stationary observers are possible only outside the horizon.

11.13 This problem was adapted from one in Ref. [110]. The causal (light cone) structure may be examined by considering the motion of radial light rays, for which $d\theta = d\varphi = ds = 0$. Inserting these constraints in the line element equation gives the local equation for lightcones in this metric

$$\frac{dt}{dr} = \left(1 - \frac{r^2}{R^2}\right)^{-1}.$$

Thus at $r = 0$ one has $dt/dr = 1$ and the forward lightcone opens at a 45 degree angle upward, but as $r \to R$ this opening angle closes continuously to zero since $dt/dr \to \infty$. Hence for $r < R$ a photon emitted from some r can reach any other value of r in a finite time, but one emitted at $r = R$ remains at R for all time. From the line element the metric components g_{00} and g_{11} interchange signs at $r = R$. As was seen in the case of the Schwarzschild metric expressed in Schwarzschild coordinates, this implies that the lightcones rotate by 90 degrees for $r > R$. Thus no null or timelike trajectory originating at a point $r > R$ can ever cross the boundary $r = R$. Conclusion: $r = R$ defines an event horizon for observers *inside it:* once one passes from $r < R$ to $r > R$, it is impossible on causal grounds to return to or send a signal to the region $r < R$.

11.14 In flat spacetime or a local inertial frame, one would expect the 4-acceleration to be given by

$$a^\mu = \frac{du^\mu}{d\tau} = \frac{dx^\alpha}{d\tau}\frac{\partial u^\mu}{\partial x^\alpha} = u^\alpha \partial_\alpha u^\mu,$$

where u is the 4-velocity and τ is the proper time measured along the particle worldline. In curved spacetime the derivatives must be replaced by covariant derivatives, so the acceleration components are given by (see Problem 11.15) $a^\mu = u^\alpha \nabla_\alpha u^\mu$, where ∇_α is the covariant derivative (3.55). But from Eq. (9.18), for a stationary observer only the timelike component of the 4-velocity is non-vanishing:

$$u = (u^0, 0, 0, 0) = \left((1 - 2M/r)^{-1/2}, 0, 0, 0\right)$$

so from Eq. (3.55)

$$\begin{aligned} a^\mu &= u^\alpha \nabla_\alpha u^\mu = u^0 \nabla_0 u^\mu = u^0 \left(\frac{\partial u^\mu}{\partial t} + \Gamma^\mu_{0\alpha} u^\alpha\right) \\ &= u^0 \left(\frac{\partial u^\mu}{\partial t} + \Gamma^\mu_{00} u^0\right) = \Gamma^\mu_{00} (u^0)^2, \end{aligned}$$

where the result in the last line used that the components of u are independent of time and that only the timelike component of u is non-zero. The connection coefficients for the Schwarzschild metric may be deduced from the solution of Problem 8.2 (summarized in Appendix C). Since the metric is independent of time, the only non-vanishing connection coefficient of the form Γ^{μ}_{00} is

$$\Gamma^1_{00} = \frac{1}{2} e^{\sigma-\lambda} \frac{d\sigma}{dr} = \left(1 - \frac{2M}{r}\right)\left(\frac{M}{r^2}\right),$$

where for the Schwarzschild metric $e^{\sigma} = e^{-\lambda} = (1 - 2M/r)$. Thus, the acceleration of the stationary observer is directed radially with

$$a^1 = \Gamma^1_{00}(u^0)^2 = \frac{M}{r^2},$$

and the magnitude of the acceleration vector is

$$\sqrt{a \cdot a} = \sqrt{a^{\mu} g_{\mu\nu} a^{\nu}} = \sqrt{a^1 g_{11} a^1} = \left(1 - \frac{2M}{r}\right)^{-1/2} \frac{M}{r^2},$$

which diverges at the Schwarzschild radius $r = 2M$. Thus infinite acceleration is required to remain stationary at the event horizon.

11.15 In a local inertial frame,

$$a^{\mu} = \frac{du^{\mu}}{d\tau} = \frac{dx^{\alpha}}{d\tau} \frac{\partial}{\partial x^{\alpha}} u^{\mu} = u^{\alpha} \frac{\partial}{\partial x^{\alpha}} u^{\mu} = u^{\alpha} \partial_{\alpha} u^{\mu},$$

where $u^{\alpha} = dx^{\alpha}/d\tau$. To convert to general curvilinear coordinates, the partial derivatives ∂_{α} must be replaced with covariant derivatives ∇_{α}, giving

$$a^{\mu} = u^{\alpha} \nabla_{\alpha} u^{\mu} = u^{\alpha} \left(\frac{\partial u^{\mu}}{\partial x^{\alpha}} + \Gamma^{\mu}_{\alpha\gamma} u^{\gamma} \right),$$

where Eq. (3.55) was used.

11.16 The loss of light and energy from the Sun would be catastrophic, but there would be essentially no change in the Earth's orbit. To the degree that the Sun is spherical and its spin can be neglected, the spacetime *outside the Sun* is well described by a Schwarzschild metric, so the Earth would be orbiting in the same metric for the Sun or for the black hole. There would presumably be some tiny differences because the Sun spins slowly and is not exactly spherical, but those differences would be extremely small. The popular idea that black holes are unique gravitational predators sucking up everything in sight is a misconception, since the gravitational field produced by a black hole of mass M well outside the event horizon is just that expected for any mass M, whether it is a black hole or not.

12 Quantum Black Holes

12.1 Using the ideal gas law and assuming the Sun to be made of N ionized hydrogen atoms, the entropy can be estimated roughly as $S \simeq k_B N \simeq 10^{57} k_B$. For the Schwarzschild solution the entropy is given by Eq. (12.14) with constants G and c restored using $\hbar \to G\hbar/c^3$ and $M \to (G/c^2)M$ from Table B.1,

$$S = \frac{c^3 k_B A}{4G\hbar} = k_B \frac{A}{4\ell_p^2} = \frac{16\pi G}{4\hbar c} k_B M^2 \simeq 1.07 \times 10^{77} k_B \left(\frac{M}{M_\odot}\right)^2,$$

where $A = 16M^2$ is the area of the event horizon, the Planck length is $\ell_p = (\hbar G/c^3)^{1/2} = 1.62 \times 10^{-33}$ cm, and M is the mass. Thus, the entropy for the black hole would be about 10^{20} times larger than the present entropy of the Sun.

12.2 A reasonable estimate is to construct a density from the Planck mass divided by the cube of the Planck length:

$$\rho_P \simeq \frac{M_P}{\ell_P^3} = \frac{c^5}{\hbar G^2} \simeq 10^{94}\,\mathrm{g\,cm^{-3}},$$

which may be compared with the paltry $10^{14}\,\mathrm{g\,cm^{-3}}$ for a neutron star.

12.3 The average time for decay of the black hole and its temperature are given by Eqs. (12.11) and (12.9), respectively:

$$t_H = 8.4 \times 10^{-26} \left(\frac{M}{1\mathrm{g}}\right)^3 \mathrm{s} \qquad T = 6.2 \times 10^{-8} \left(\frac{M_\odot}{M}\right) \mathrm{K}.$$

Solving the first of these for the mass with $t_H = 500{,}000\,\mathrm{yr}$ gives $M = 5.7 \times 10^{12}\,\mathrm{g} = 2.9 \times 10^{-21} M_\odot$. Assuming a spherical event horizon, $r_s = 2M = 8.6 \times 10^{-14}$ cm, where $1M_\odot = 1.475$ km in geometrized units has been used. From the second equation for the temperature above and the mass computed above, $T \simeq 2.1 \times 10^{13}$ K.

12.4 From Eq. (12.8) a black hole with lifetime comparable to the age of the Universe has a mass of order 10^{14} g. One meter is very far outside its event horizon so using Newtonian gravity to estimate the gravitational acceleration gives $|g| = MG/r^2 \simeq 6.7\,\mathrm{m\,s^{-2}}$, which is comparable to the gravitational acceleration of the Earth at its surface.

12.5 From Eq. (12.10) with $M = 1M_\odot$, the black hole radiates a minuscule Hawking power of 9×10^{-29} W, which is about 55 orders of magnitude smaller than the actual solar power of 3.828×10^{26} W. From Eq. (12.9) the corresponding black hole temperature is a frigid 6.2×10^{-8} K.

12.6 From Eqs. (12.11) and (12.10) the mass is 2.3×10^5 kg and the power is 6.2×10^{21} W.

12.7 The CMB has a temperature of about 2.7 K. From Eq. (12.4) a Hawking black hole of mass 4.6×10^{22} kg would have the same temperature and thus be in equilibrium with the CMB (this mass is a little less than that of the Moon, or about 0.008 that of Earth). Black holes with less mass than this would have higher temperature than the CMB and thus could radiate more energy than they absorb; black holes with more mass than this would absorb more energy from the CMB than they could radiate by Hawking radiation.

12.8 From the Stefan–Boltzmann law for a blackbody radiator the power is

$$P = A\sigma T^4,$$

where A is the surface area, the Stefan–Boltzmann constant is

$$\sigma = \frac{\pi^2 k^4}{60\hbar^3 c^2},$$

and T is the temperature. For a Schwarzschild black hole the area of the event horizon is

$$A = 16\pi M^2 = \frac{16\pi G^2 M^2}{c^4}$$

(see results of Problem 11.9 with factors of c and G reinstated using Table B.1). Inserting these and the temperature given by Eq. (12.4) in $P = A\sigma T^4$ gives Eq. (12.5). The power radiated by the black hole comes at the expense of its mass so

$$P = -c^2 \frac{dM}{dt},$$

which leads to Eq. (12.6) when Eq. (12.5) is substituted for P.

12.9 This problem is adapted from a discussion in Perkins [181]. Using Newtonian gravity, the magnitude of the tidal force acting over a distance Δr can be estimated as

$$dF = \frac{dF}{dr}\,dr = \frac{2mMG}{r^3}\,dr \quad \rightarrow \quad \Delta F \simeq \frac{2mMG}{r^3}\,\Delta r,$$

where M is the mass of the black hole. The energy required to create the particle–hole pair from the vacuum is $E \simeq mc^2$. By the uncertainty principle the virtual pair can live for a time $\Delta t \sim \hbar/E$, and thus could separate a maximum distance

$$\Delta r \sim c\Delta t \sim \frac{\hbar c}{E}$$

in the time Δt. Requiring that the work done $\Delta F \cdot \Delta r$ be comparable to the rest mass energy,

$$\Delta F \cdot \Delta r = \frac{2mMG}{r^3}(\Delta r)^2 \sim E,$$

substituting $m \sim E/c^2$ and $\Delta r \sim \hbar c/E$, and solving for E gives

$$E \sim \sqrt{\frac{2\hbar^2 GM}{r^3}}.$$

Evaluating this at the Schwarzschild radius $r = 2MG/c^2$ gives

$$E \sim \frac{\hbar c^3}{GM},$$

where a factor of $\frac{1}{2}$ has been dropped since our approximations are crude. But up to numerical factors this is the result of Eq. (12.4) for the average energy $k_B T$ associated with the Hawking radiation.

13 Rotating Black Holes

13.1 This is straightforward but entails a considerable amount of algebra. For example, collecting the terms proportional to dt^2 in the alternative form of the metric,

$$\begin{aligned}
\frac{-\Delta}{\rho^2}dt^2 + \frac{\sin^2\theta}{\rho^2}a^2dt^2 &= \frac{1}{\rho^2}\left(a^2\sin^2\theta - \Delta\right)dt^2 \\
&= \frac{1}{\rho^2}\left(a^2\sin^2\theta - r^2 + 2Mr - a^2\right)dt^2 \\
&= \frac{1}{\rho^2}\left(a^2(\sin^2\theta - 1) - r^2 + 2Mr\right)dt^2 \\
&= \frac{1}{r^2 + a^2\cos^2\theta}\left(-r^2 - a^2\cos^2\theta + 2Mr\right)dt^2 \\
&= -\left(1 + \frac{2Mr}{r^2 + a^2\cos^2\theta}\right)dt^2 = -\left(1 + \frac{2Mr}{\rho^2}\right)dt^2,
\end{aligned}$$

which is the first term of Eq. (13.1).

13.2 From Eq. (13.10) the horizon area for a Kerr black hole may be written as

$$A_{\mathrm{K}} = 4\pi(r_{\mathrm{H}}^2 + a^2) \qquad r_{\mathrm{H}} \equiv M + \sqrt{M^2 - a^2}.$$

Because of entropy conservation (Hawking area theorem), the horizon area cannot decrease as angular momentum is extracted, and removing all the angular momentum leaves a Schwarzschild black hole. The maximum energy extraction will occur if the horizon remains constant in size, since if it grows the Schwarzschild black hole left behind will be more massive than if the horizon stays constant. Taking the $a = J/M = 0$ limit of the above expression,

$$A_{\mathrm{K}} = 4\pi r_{\mathrm{H}}^2 = 4\pi(2M_0)^2 = 16\pi M_0^2,$$

where $r_{\mathrm{H}} = 2M$ for the Schwarzschild black hole has been used, and where M_0 represents the mass left after all angular momentum has been extracted (sometimes termed the *irreducible mass*). Equating the above two expressions for A_{K},

$$16\pi M_0^2 = 4\pi\left[\left(M + \sqrt{M^2 - a^2}\right)^2 + a^2\right],$$

and solving for M_0 gives

$$M_0 = \frac{1}{2}\sqrt{\left[M + (M^2 - a^2)^{1/2}\right]^2 + a^2}.$$

Thus, the maximum total energy that can be extracted from the Kerr black hole by removing its angular momentum is

$$(M - M_0)c^2 = \left(M - \frac{1}{2}\sqrt{\left[M + (M^2 - a^2)^{1/2}\right]^2 + a^2} \right) c^2.$$

For an extremal Kerr black hole, $a = M$, which yields

$$E_{\max} = (M - M_0)c^2 = \left(1 - \frac{1}{\sqrt{2}}\right) Mc^2 = 0.29\,Mc^2,$$

for the maximum energy that could be extracted from the Kerr black hole by reducing it to a Schwarzschild black hole. In principle, this could be accomplished by a Penrose process, but they are difficult to arrange in practical astrophysics environments. A more plausible realization is the Blandford–Znajek mechanism [50], where compression of magnetic field lines near the horizon can cause the rotational energy to be emitted as a stream of e^+e^- pairs.

13.3 The steps are analogous to those for the equations of motion in the Schwarzschild metric derived in Section 9.3.

13.4 The distances are gotten by integrating the line element (13.8) around the corresponding curves. For the equator, $\theta = \frac{\pi}{2}$ so $d\theta = 0$ and only the second term contributes. The distance around the equator is

$$L = \oint \sqrt{ds^2} = 2M \int_0^{2\pi} d\varphi = 4\pi M,$$

where $\sin\theta = 1$ has been used. For a meridian through the poles take $\varphi = 0$ and $d\varphi = 0$, so only the first term contributes. The corresponding distance is

$$L' = \oint \sqrt{\rho_+}\, d\theta = 2\int_0^{\pi} \sqrt{r_+^2 + a^2\cos^2\theta}\, d\theta.$$

For the special case of an extremal black hole $a = M$ and $r_+ = M$, so that $\rho_+ = M(1 + \cos^2\theta)^{1/2}$ and

$$L' = 2M \int_0^{\pi} \sqrt{1 + \cos^2\theta}\, d\theta \simeq 7.6M,$$

where the definite integral has been evaluated numerically using Maple. Thus, the ratio of the equatorial to polar circumferences for the horizon of an extremal Kerr black hole is $L/L' = 4\pi M/7.6M = 1.65$. Although the horizon corresponds to a constant Boyer–Lindquist coordinate r_+, it does not have a spherical geometry.

13.5 Substitution of $a = J/M$ into Eq. (13.1), discarding terms higher than linear in J, and a little algebra gives

$$ds^2 \simeq (ds^2)_0 - \frac{4GJ}{c^3 r^2}\sin^2\theta (r d\varphi)(c dt),$$

where $(ds^2)_0$ is the Schwarzschild line element and factors of G and c have been restored using the conversion $J \to GJ/c^3$ from Table B.1. Utilizing the classical result that the angular momentum of a rotating body is given by $J \sim Mrv$, the dimensionless ratio (GJ/c^3r^2)

governing the strength of the rotational correction may be expressed as

$$\frac{4GJ}{c^3r^2} \simeq \left(\frac{v}{c}\right)\left(\frac{GM}{rc^2}\right).$$

The second factor GM/rc^2 is the measure of spacetime curvature produced by a non-rotating spherical mass given in Eq. (6.5). Thus the total spacetime curvature (which receives contributions from this correction plus that coming from the leading Schwarzschild term) depends both on the spherical mass and the rotational velocity. From Eq. (6.12) the gravitational redshift caused by the spherical mass is of order $1/c^2$, but the above effect is of order $1/c^3$.

13.6 From the Kerr line element with $\theta = \frac{\pi}{2}$ and $dt = d\varphi = d\theta = 0$

$$\frac{ds}{dr} = \sqrt{\frac{r^2}{r^2 - 2Mr + a^2}}.$$

If $a \to 0$ the Schwarzschild result (9.7) in geometrized units is recovered. If instead one sets $a = M$, then

$$\frac{ds}{dr} = \left(1 - \frac{2M}{r} + \frac{M^2}{r^2}\right)^{-1/2}$$

for an extremal Kerr black hole.

13.7 Let us calculate $d\varphi/dr$ for a particle dropped from rest ($\varepsilon = 1$) with zero angular momentum ($\ell = 0$) into a Kerr black hole. Since ε and ℓ are conserved they retain their initial values and from Eqs. (13.15) and (13.16) for an infalling trajectory,

$$\frac{d\varphi}{d\tau} = \frac{1}{\Delta}\left(\frac{2Ma}{r}\right) \qquad \frac{dr}{d\tau} = \pm\sqrt{\frac{2M}{r}\left(1 + \frac{a^2}{r^2}\right)},$$

where the negative sign should be chosen on the square root because the particle is infalling. These may be combined to give

$$\frac{d\varphi}{dr} = \frac{d\varphi/d\tau}{dr/d\tau} = -\frac{2Ma}{r\Delta}\left[\frac{2M}{r}\left(1 + \frac{a^2}{r^2}\right)\right]^{-1/2}.$$

Since in the general case $d\varphi/dr \neq 0$, the particle is dragged in φ as it falls radially inward, even though no forces act on it.

13.8 Assume $r \gg M$ and $r \gg a$, and drop terms quadratic in a. Then

$$\rho^2 = r^2 + a^2\cos\theta \simeq r^2 \qquad \Delta = r^2 - 2Mr + a^2 \simeq r^2 - 2Mr.$$

Substituting in the Kerr metric (13.1) gives

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 + \frac{2M}{r}\right)dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\varphi^2 - \frac{4Ma}{r}\sin^2\theta d\varphi dt,$$

where $(1 - 2M/r)^{-1} \sim (1 + 2M/r)$ was used. Taking the limit $r \to \infty$ at constant M and a gives

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

which is the flat-space Minkowski metric in spherical coordinates. Hence the Kerr spacetime is asymptotically flat.

14 Observational Evidence for Black Holes

14.1 From Kepler's third law,

$$a = \left(\frac{GMP^2}{4\pi^2}\right)^{1/3} = \left[7.495 \times 10^{-6} \left(\frac{M}{M_\odot}\right)\left(\frac{P}{1\,\text{day}}\right)^2\right]^{1/3} \text{AU},$$

where a is the semimajor axis, M is the total mass, and P is the period. Taking

$$M = 40M_\odot \qquad P = 5.6\,\text{days}$$

gives $a \simeq 0.2$ AU for the average separation. In one ms light can travel a distance

$$d \sim 2 \times 10^{-6}\,\text{AU} \sim 300\,\text{km}.$$

Thus, X-ray fluctuations on a millisecond scale set an upper limit of about 300 km for the diameter of the X-ray emitting region.

14.2 Insert $G = 6.6726 \times 10^{-8}\,\text{cm}^3\text{g}^{-1}\text{s}^{-2}$ and convert units to give

$$f(M) = \frac{PK^3}{2\pi G} = 1.036 \times 10^{-7} \left(\frac{P}{1\,\text{day}}\right)\left(\frac{K}{\text{km s}^{-1}}\right)^3 M_\odot.$$

Eyeballing from Fig. 14.2 $K \simeq 450$ km s^{-1} and using the period of 0.32 days from Table 14.1 gives $f \simeq 3.02M_\odot$. The discrepancy with the value of $3.18M_\odot$ quoted in Table 14.1 is because in the literature a value $K \simeq 457$ km s^{-1} is commonly used; see for example Refs. [150, 111]. These references may also be consulted for methods used to try to infer the actual mass of the black hole from this lower limit.

14.3 From standard conversion factors, 1 light-minute = 0.12 AU, 1 light-hour = 7.22 AU, and 1 light-day = 173.3 AU. Thus, from the quantities displayed in Fig. 14.5 the distance of closest approach is 17 light-hours = 122.7 AU (about 6 times larger than the orbit of Neptune). The event-horizon radius for a spherical black hole of mass $4.3 \times 10^6 M_\odot$ is

$$R = 2.95 \left(\frac{M}{M_\odot}\right) \text{km} \simeq 12.7 \times 10^6\,\text{km} = 0.085\,\text{AU}.$$

For a black hole with mass M and star with mass M_* and radius R_*, the tidal distortion radius for the star's orbit may be estimated as

$$r_\text{t} \simeq \left(\frac{M}{M_*}\right)^{1/3} R_*.$$

S0-2 is a $M_* = 15M_\odot$ main sequence star, which from stellar systematics implies that $R_* \sim 6.5R_\odot \sim 4.55 \times 10^6$ km. This gives $r_\text{t} \sim 3 \times 10^8$ km ~ 16.6 light-minutes ~ 2 AU. Thus S0-2 at closest approach is well outside the event horizon and the tidal distortion radius of the black hole. The average velocity of Earth on its orbit can be estimated by assuming the orbit to the circular. Then the distance around the orbit is $2\pi r$, where $r = 1$ AU. Dividing this by a period of 1 year gives an average velocity of 29.7 km s^{-1}, which is 168 times smaller than the 5000 km s^{-1} orbital velocity of S0-2 at closest approach to the black hole. The radius of the Sgr A* $4.3 \times 10^6\, M_\odot$ black hole is 0.085 AU, the radius of the Sun is 4.65×10^{-3} AU, and the semimajor axis of Mercury's orbit is 0.39 AU. Thus the black hole is 18.3 times larger than the Sun and 0.22 times the size of Mercury's orbit.

14.4 Evaluating the constants in Eq. (14.8) gives

$$M = 3.77 \times 10^{-11} \left(\frac{R}{\text{km}}\right)\left(\frac{\sigma}{\text{km/s}}\right)^2 M_\odot$$

(a) For M31 with $R = 0.8\,\text{pc} = 2.47 \times 10^{13}\,\text{km}$ and $\sigma = 240\,\text{km s}^{-1}$, the virial mass is

$$M = 3.77 \times 10^{-11}(2.47 \times 10^{13})(240)^2 M_\odot = 5.36 \times 10^7 M_\odot.$$

(b) For a typical Seyfert galaxy,

$$M = 3.77 \times 10^{-11}(3.09 \times 10^{15})(1000)^2 M_\odot = 1.16 \times 10^{11} M_\odot,$$

where $R = 100\,\text{pc} = 3.09 \times 10^{15}\,\text{km}$ and $\sigma = 1000\,\text{km s}^{-1}$ were assumed. (This suggests that 10% of the mass of the Seyfert galaxy is contained within a central region only 100 pc across.)

14.5 Let $m_1 = M$ and $m_2 = M_\text{c}$, and define the mass ratio $q \equiv m_2/m_1 = M_\text{c}/M$. Then Kepler's 3rd law is

$$m_1(1+q) = m_1 + m_2 = \frac{4\pi^2}{G}\frac{(a_1+a_2)}{P^2}$$

Utilizing $m_1 a_1 = m_2 a_2$ gives $(a_1+a_2)^3 = a_2^3(1+q)^3$ and thus

$$m_1(1+q) = \frac{4\pi^2}{G}\frac{a_2^3(1+q)^3}{P^2}.$$

For a spectroscopic binary the easily observable orbital quantities are the period P and the semiamplitude of the radial velocity $K \equiv v_2 \sin i$ for the visible star 2 (see Fig. 14.2), with

$$P = \frac{2\pi}{\omega} = \frac{2\pi a_2}{v_2} = \frac{2\pi a_2 \sin i}{v_2 \sin i} = \frac{2\pi a_2 \sin i}{K}.$$

Solving this expression for a_2 and inserting in the Kepler's law equation gives

$$m_1 = \frac{K^3 P}{2\pi G}\frac{(1+q)^2}{\sin^3 i}.$$

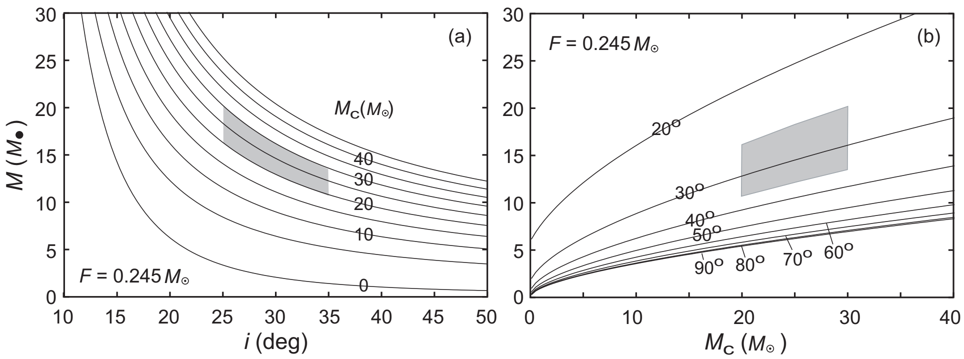
Rearranging and substituting $q = m_2/m_1$ gives

$$\frac{PK^3}{2\pi G} = \frac{m_1 \sin^3 i}{(1+q)^2} = \frac{m_1^3 \sin^3 i}{(m_1+m_2)^2},$$

which is Eq. (14.2) with $M = m_1$ and $M_\text{c} = m_2$.

14.6 From Fig. 2 of Ref. [171], $P = 5.6$ days and $K \sim 75\,\text{km s}^{-1}$, which gives from Eq.

(14.3) $F = PK^3/2\pi G = 0.245$. Inserting this in Eq. (14.6) and plotting M versus i and M versus M_c gives the following diagrams.



The gray areas indicate the ranges $i = 25$–35° and $M_c = 20$–$30 M_\odot$, which constrain the mass M of the unseen object to lie in the range ~ 10–$20 M_\odot$. A more diligent use of observational constraints described in Ref. [171] and Box 14.2 gives a value of $M = 14.8 \pm 1.0 M_\odot$. Since these mass estimates are far above the maximum mass expected for a neutron star, the unseen massive and compact object in Cygnus X-1 is almost certainly a black hole.

14.7 From Eq. (6.5), the intrinsic strength of gravity is governed by

$$\varepsilon = \frac{GM}{Rc^2} = 7.416 \times 10^{-31} \left(\frac{M}{\text{kg}}\right)\left(\frac{\text{km}}{R}\right) = 1.475 \left(\frac{M}{M_\odot}\right)\left(\frac{\text{km}}{R}\right),$$

where M is the mass producing the gravitational field and R is the characteristic distance over which it acts. For the Sun, taking the mass of the Sun as M and its radius as R gives $\varepsilon_\odot \sim 2.1 \times 10^{-6}$. For the Binary Pulsar, taking the separation at closest approach ($\sim 1.1 R_\odot$) as R and the mass of about $1.4 M_\odot$ for the other neutron star as M gives $\varepsilon_{\text{BP}} \sim 2.7 \times 10^{-6}$. For the star S0-2, taking the distance at closest approach of 17 light–hours as R and the mass $4.3 \times 10^6 M_\odot$ of the black hole as M gives $\varepsilon_{\text{S0-2}} = 3.45 \times 10^{-4}$ (which is comparable to the strength of gravity at the surface of a white dwarf). The star S0-102 comes even closer to the black hole so ε for it is larger but the same order of magnitude. Thus ε is about two orders of magnitude larger for stars in orbit around the black hole at Sgr A* than for gravity at the surface of the Sun or in the Binary Pulsar, and the orbits of stars like S0-2 and S0-102 can provide a test of general relativity in stronger gravity than for either Solar System measurements or binary pulsars.

15 Black holes as Central Engines

15.1 (a) From the masses before and after in nuclear reactions burning hydrogen to helium one finds a typical efficiency for mass to energy conversion of $\eta \sim 0.007$. Requiring that

$$\eta \dot{m} c^2 = 10^{47} \text{ erg s}^{-1} = 3.2 \times 10^{54} \text{ erg yr}^{-1}$$

yields that $\dot{m} \simeq 255 M_\odot \text{ yr}^{-1}$ if this luminosity is supplied by hydrogen fusion. It is difficult to conjecture a mechanism consistent with observations that could account for this.

(b) On the other hand, for black hole accretion the mass to energy conversion efficiency could lie in the range $\eta \simeq 0.1 - 0.4$. Then the luminosity could be sustained by accretion of $5 M_\odot - 20 M_\odot \text{ yr}^{-1}$, for which there are plausible mechanisms. The Eddington luminosity (maximum luminosity for which radiation pressure would not reverse accretion infall) is given by

$$L_{\text{Edd}} = 1.3 \times 10^{38} \left(\frac{M}{M_\odot} \right) \text{erg s}^{-1}.$$

Equating this to the observed $\sim 10^{47}$ erg s^{-1} and solving for the mass of the central object gives $M \simeq 7.7 \times 10^8 M_\odot$.

(c) Observed light variation on timescales of days argues by causality that the source has a maximum diameter of 1 light-day or about 170 AU (twice the diameter of the Solar System). Thus, it may be inferred on rather general grounds that the AGN central engine has of order $10^9 M_\odot$ concentrated in a region not much larger than the Solar System. The most plausible explanation is a supermassive black hole. The Schwarzschild radius of a $10^9 M_\odot$ black hole would be

$$R \simeq 2.95 \left(\frac{M}{M_\odot} \right) \text{km} \simeq 3 \times 10^9 \text{ km} \simeq 20 \text{ AU},$$

which is approximately the radius of the orbit of Uranus.

15.2 From Eq. (15.1), in Newtonian approximation the energy released through accretion is

$$\Delta E_{\text{acc}} = \frac{GMm}{R} = 1.327 \times 10^{21} \left(\frac{M}{M_\odot} \right) \left(\frac{m}{\text{g}} \right) \left(\frac{\text{km}}{R} \right) \text{erg},$$

where R is the radius and M the mass of the object onto which accretion is taking place, and m is the mass of the accreting material. The entries in Table 15.1 result from using in this formula 1 gram for m and the representative parameters

1. $M = 1 M_\odot$ and $R = 10$ km for a neutron star,
2. $M = 1 M_\odot$ and $R = 10^4$ km for a white dwarf,
3. solar parameters for a normal star, and

4. $M = 10M_\odot$ and $R = 3r_s$ for a black hole, where the Schwarzschild (event-horizon) radius is given by

$$r_s = 2.95\left(\frac{M}{M_\odot}\right)\,\text{km}$$

and $3r_s$ is the radius of the innermost stable circular orbit in a Schwarzschild spacetime; see Eq. (9.37).

The last column of the table is normalized to the thermonuclear burning of hydrogen to helium ("fusion"), which releases $\sim 6\times10^{18}$ erg for each gram of hydrogen burned. These numbers should be viewed as approximate for neutron stars and black holes since they have been estimated using Newtonian gravity. However, they show clearly that the accretion energy can be very large for highly compact objects.

15.3 From the geometry of the figure, the leading edge of the jet appears to move about 14 lightyears in two elapsed observing years. Thus the apparent transverse velocity is $v \sim 14/2 \sim 7c$.

15.4 From the inverse Compton boost factor given in Box 15.4, a visible photon of frequency 5×10^{14} Hz is boosted to a frequency

$$\nu = \gamma^2\nu_0 = \frac{\nu_0}{1-v^2/c^2} = \frac{5\times10^{14}\,\text{Hz}}{1-(0.9999995)^2} = 5\times10^{20}\,\text{Hz},$$

which lies in the γ-ray region of the spectrum.

15.5 Neglecting the spectral energy distribution of the emitted flux,

$$\frac{S_{\text{observed}}}{S_{\text{emitted}}} = \frac{1}{[\gamma(1-\beta\cos\theta)]^3}.$$

Thus, if one assumes the same γ and β for both jets and that the approaching jet makes an angle θ with the line of sight, the ratio of observed flux densities for light emitted from the approaching and receding jets will be

$$\frac{S_{\text{approach}}}{S_{\text{recede}}} = \frac{[\gamma(1-\beta\cos\theta)]^{-3}}{[\gamma(1-\beta\cos(\theta+\pi))]^{-3}} = \left(\frac{1+\beta\cos\theta}{1-\beta\cos\theta}\right)^3.$$

If, for example, $\beta = 0.98$ and $\theta = 10^\circ$, this ratio is $\sim 10^5$, implying that the counterjet will appear to be much fainter than the jet. If one takes into account the spectral energy distribution of the emitted flux, the above formula remains valid except that the exponent of 3 is replaced by a somewhat smaller exponent ~ 2.7 (see Section 8.3.3 of Ref. [204]).

15.6 By the Hubble law (Ch. 16), the distance is

$$d = \frac{v}{H_0} \sim \frac{cz}{H_0},$$

since $z \sim v/c$ for redshifts that are not too large. Using $H_0 = 72\,\text{km s}^{-1}\,\text{Mpc}^{-1}$,

$$d = \frac{cz}{H_0} = \frac{(3\times10^5\,\text{km s}^{-1})(0.158)}{72\,\text{km s}^{-1}\,\text{Mpc}^{-1}} \simeq 660\,\text{Mpc}.$$

The apparent magnitude m and the absolute magnitude M for a distant object are related by the distance modulus formula

$$M = m - 5\log\frac{d\,(\mathrm{pc})}{10},$$

where $d\,(\mathrm{pc})$ is the distance to the object in parsecs. Using the apparent visual magnitude $m = +12.9$ and $d = 660\times 10^6$ pc gives an absolute visual magnitude of $M = -26.2$ for 3C 273. Absolute magnitudes M and luminosities L for objects 1 and 2 are related by

$$\frac{L_1}{L_2} = 10^{0.4(M_2-M_1)}.$$

For the Sun the absolute visual magnitude is $+4.8$ and the above formula indicates that 3C 273 is 2.5×10^{12} times more luminous than the Sun at visual wavelengths. For M31, with absolute visual magnitude of about -21.5, the corresponding ratio is about 76, and for M87, with absolute visual magnitude of about -22, one obtains that 3C 273 is 48 times more luminous at visual wavelengths. Thus, 3C 273 is roughly 100 times more luminous at visual wavelengths than large galaxies. However, the quasar emits most of its light at nonvisual wavelengths. When the luminosities are integrated over all wavelengths 3C 273 is found to be about 1000 times more luminous than large normal galaxies.

15.7 For the thin disk radiating as a blackbody the radiation rate per unit area is σT^4, where σ is the Stephan–Boltzmann constant. Thus, if the disk has a radius R the luminosity L is

$$L = 2\pi R^2\sigma T^4,$$

where the factor of two comes from the disk having two sides. Assume the observed luminosity of the disk to be a fraction η of the Eddington luminosity (15.3). Then solving the preceding equation for T and using the expression (15.4) to approximate the Eddington luminosity gives for the temperature of the disk

$$T = \left(\frac{L}{2\pi R^2\sigma}\right)^{1/4} = \left(\frac{\eta L_{\mathrm{edd}}}{2\pi R^2\sigma}\right)^{1/4} = 7.72\times 10^7\left(\eta\frac{M/M_\odot}{(R/\,\mathrm{km})^2}\right)^{1/4}\ \mathrm{K},$$

where M is the gravitational mass responsible for the accretion (which may be approximated by the mass of the compact object, since this is much larger than the mass in the accretion disk). For a neutron star, assuming

$$R\sim 10\,\mathrm{km} \qquad M\sim 1M_\odot \qquad \eta\sim 1,$$

this formula yields $T\simeq 2.4\times 10^7$ K. By the Wien law, the corresponding blackbody spectrum peaks at a wavelength

$$\lambda_{\mathrm{peak}} = \frac{2.9\times 10^{-3}\,\mathrm{m\ K}}{T}\simeq 0.12\,\mathrm{nm},$$

which is in the X-ray portion of the spectrum. For a Schwarzschild (spherical) black hole, approximate R by the radius of the innermost stable circular orbit, which from Eq. (9.34) is given by

$$R = \frac{6GM}{c^2},$$

with factors of G and c restored. In convenient units, $G/c^2 = 1.475\,\mathrm{km}\,M_\odot{}^{-1}$ and the preceding equation for T may be written

$$T = 2.6 \times 10^7 \eta^{1/4} \left(\frac{M_\odot}{M}\right)^{1/4} \mathrm{K}.$$

Assuming radiation near the Eddington limit so that $\eta \sim 1$, for a 10 $M_\odot$ black hole this formula and the Wien law give

$$T = 1.5 \times 10^7\,\mathrm{K} \qquad \lambda_{\mathrm{peak}} = 0.20\,\mathrm{nm},$$

which is dominantly in the X-ray region of the spectrum. For a $10^8 M_\odot$ black hole, we find likewise that

$$T = 2.6 \times 10^5\,\mathrm{K} \qquad \lambda_{\mathrm{peak}} = 11\,\mathrm{nm},$$

which is dominantly in the UV portion of the spectrum. This has been a rather crude approximation to the physics of accretion disks (for more realistic descriptions, see Refs. [89, 183]), but it indicates correctly that accretion disks around neutron stars or stellar-size black holes are expected to radiate in the X-ray region, but the corresponding accretion disks around supermassive black holes should have lower temperatures and radiate at longer wavelengths, largely in the UV portion of the spectrum.

15.8 Evaluation of constants allows Eq. (15.15) to be written as

$$\begin{aligned}
\tau \simeq \frac{f\sigma_{\mathrm{T}} F D^2}{\delta t^2 m_{\mathrm{e}} c^4} &= 9.01 \times 10^{-40} f \left(\frac{D}{\mathrm{cm}}\right)^2 \left(\frac{F}{\mathrm{erg\,cm^{-2}}}\right) \left(\frac{\mathrm{s}}{\delta t}\right)^2 \\
&= 8.58 \times 10^{15} f \left(\frac{D}{\mathrm{Mpc}}\right)^2 \left(\frac{F}{\mathrm{erg\,cm^{-2}}}\right) \left(\frac{\mathrm{ms}}{\delta t}\right)^2 \\
&= 7.7 \times 10^{13} f \left(\frac{D}{3000\,\mathrm{Mpc}}\right)^2 \left(\frac{F}{10^{-7}\,\mathrm{erg\,cm^{-2}}}\right) \left(\frac{10\,\mathrm{ms}}{\delta t}\right)^2 .
\end{aligned}$$

Since f and the product of quantities in parentheses are of order one in the last expression for a typical gamma-ray burst, the resulting optical depth is huge ($\tau \sim 10^{14}$). This is inconsistent with the observed nonthermal spectrum for gamma-ray bursts, since a nonthermal spectrum typically requires a medium that is optically thin. The fallacy is that Eq. (15.15) is invalid for a gamma-ray burst because it must be modified to account for the ultrarelativistic kinematics of the burst. When that is done, as in Eq. (15.16), the above expression is multiplied by a factor approximately equal to $1/\gamma^{4+2\alpha}$, as discussed in Section 15.7.4. For a typical value $\alpha \sim 2$ this will yield optical depths smaller than one for γ of order 100 or larger.

16 The Hubble Expansion

16.1 This problem was adapted from a discussion in Peebles [178]. Assume a single luminosity bin i, with the number of galaxies in the bin given by $n_i \equiv n(L_i)$. Out to a distance r, the volume of space in 4π steradians is

$$V(4\pi\,\mathrm{sr}) = 4\pi \int_0^r r^2\,dr = \tfrac{4}{3}\pi r^3,$$

so the volume in 1 sr is

$$V(1\,\mathrm{sr}) = \frac{V(4\pi\,\mathrm{sr})}{4\pi} = \tfrac{1}{3}r^3.$$

Thus the mean number of galaxies per steradian brighter than $f = L/4\pi r^2$ (closer than r) is

$$n_i V = \tfrac{1}{3}n_i r^3 = \frac{n_i}{3}\left(\frac{L_i}{4\pi f}\right)^{3/2},$$

where $f = L/4\pi r^2 \to r^3 = (L/4\pi f)^{3/2}$ has been used. If the galaxies are distributed uniformly the preceding applies to each luminosity bin and the total is gotten by summing over all bins

$$N(f_>) = \sum_i n_i V = \tfrac{1}{3}\sum_i n_i \left(\frac{L_i}{4\pi f}\right)^{3/2},$$

which is proportional to $f^{-3/2}$. To convert to magnitudes m, use $m = -2.5\log f +$ constant to replace f with m, giving $N(f_>) \propto 10^{0.6m}$. More realistically space is not euclidean and not static, so the above discussion must be modified to reflect redshifts in the expanding Universe and the effect of galaxy evolution over time (since distant galaxies are observed at a younger age than more nearby galaxies). Nevertheless, the realization by Hubble that the galaxy observations available to him (out to about 800 Mpc, corresponding to a redshift $z \sim 0.2$ and to light that was emitted more than 2 billion years ago) were approximately consistent with the $f^{-3/2}$ power law was the first strong evidence that galaxies are distributed homogeneously on large scales.

16.2 From the solution of Problem 4.7, the relativistic Doppler frequency shift is given by

$$\frac{\nu_0}{\nu} = \frac{1+v/c}{\sqrt{1-v^2/c^2}} = \sqrt{\frac{1+v/c}{1-v/c}}.$$

Using $\nu = c/\lambda$ and $z \equiv (\lambda - \lambda_0)/\lambda_0$ gives

$$1+z = \frac{\lambda}{\lambda_0} = \frac{\nu_0}{\nu} = \frac{1+v/c}{\sqrt{1-v^2/c^2}}.$$

However, the cosmological redshift should not be interpreted as being due to a Doppler shift (though it commonly is, especially in popular-level discussions). It is actually caused by the expansion of space, as shown in Section 16.2.3.

16.3 Integration of the Hubble law $dr/dt = H_0 r$, assuming that H_0 is constant with time, gives $r(t) \sim e^{H_0 t}$. The volume of a spherical region is

$$V(t) = \tfrac{4}{3}\pi r^3 \simeq e^{3H_0 t}.$$

Since the volume expands but the density is assumed constant, matter must be created continuously to maintain the constant density. The total mass within a volume is $M = \rho V$, where ρ is the constant density. Then $\dot{M} = \rho \dot{V} = \rho \times 3H_0 V$ and the creation rate per unit volume is

$$\frac{\dot{M}}{V} = 3H_0 \rho \simeq 7 \times 10^{-48}\ \mathrm{g\,s^{-1}cm^{-3}},$$

where a matter density of $\rho \sim 10^{-30}\ \mathrm{g\,cm^{-3}}$ and a Hubble parameter $H_0 \simeq 72\ \ \mathrm{km\,s^{-1}Mpc^{-1}}$ were assumed. This is equivalent to the creation of about one hydrogen atom per cubic meter every 10 billion years.

16.4 (a) The most important point is that light from stars beyond a certain distance has not had time to reach us if the Universe is of finite age. A secondary point is that this light is redshifted to lower energies in an expanding universe. Thus big bang cosmology resolves Olber's paradox. Note that the sky actually *is* uniformly bright, not at visible wavelengths, but rather at millimeter wavelengths because of the 2.725 K cosmic microwave background radiation discussed in Section 20.4.

(b) The first alternative explanation is based on incorrect physical reasoning because intervening dust would absorb energy, equilibrate, and re-emit the energy that was originally absorbed. The second is inconsistent with the cosmological principle because it postulates a distribution of stars on large scales that is not uniform.

16.5 From Eq. (6.5), a general relativistic description is required if $GM/Rc^2 \simeq 1$. As a crude estimate take the Universe to be static and euclidean, with a radius given by the Hubble distance and a density comparable to the critical density. Then, treating the Universe as a spherical gravitating mass,

$$R \sim \frac{c}{H_0} \qquad M \sim \tfrac{4}{3}\pi R^3 \rho_{\mathrm{crit}} \qquad \rho_{\mathrm{crit}} \simeq \frac{3H_0^2}{8\pi G},$$

which implies that

$$\frac{GM}{Rc^2} = \frac{4}{3}\pi G \frac{R^2 \rho_{\mathrm{crit}}}{c^2} \simeq \frac{1}{2}.$$

Therefore, a correct cosmological description is expected to involve general relativity.

16.6 Taking a Hubble constant of $72\ \mathrm{km\,s^{-1}\ Mpc^{-1}}$, corresponding to $h = 0.72$, the recessional velocity by the Hubble law would be

$$v = H_0 d \simeq 100h \frac{\mathrm{km}}{\mathrm{s\,Mpc}} \times 16\,\mathrm{Mpc} = 1600h\,\mathrm{km\ \ s^{-1}} = 1152\,\mathrm{km\ \ s^{-1}}.$$

But the observed recessional velocity is only 985 $\mathrm{km\,s^{-1}}$, so the peculiar velocity is 167 $\mathrm{km\,s^{-1}}$, toward us.

16.7 The erroneous Hubble constant may be expressed as

$$H_0 = 550 \frac{\mathrm{km}}{\mathrm{s\,Mpc}} \times \frac{1\,\mathrm{Mpc}}{3.086 \times 10^{19}\,\mathrm{km}} = 1.78 \times 10^{-17}\,\mathrm{s}^{-1}.$$

Therefore, the Hubble time is $1/H_0 = 5.61 \times 10^{16}\,\mathrm{s} = 1.78 \times 10^9\,\mathrm{yr}$, which is about an order of magnitude smaller than the currently accepted age of the Universe.

16.8 The Lyman alpha spectral line normally at $\lambda_0 = 121.6\,\mathrm{nm}$, is observed to be shifted to $\lambda = 968.2\,\mathrm{nm}$. Thus,

$$z = \frac{\lambda}{\lambda_0} - 1 = \frac{968.2}{121.6} - 1 = 6.96.$$

The scale factor corresponding to this redshift is $a = (1+z)^{-1} \simeq 0.125$. The evolution of the scale factor with time according to the best current cosmological parameters is given by Fig. 19.10. From that plot, a redshift of $z \sim 7$ (or scale factor relative to today of $a(t) = 0.125$), corresponds to a time about 0.75 billion years after the big bang (defined by time when $a = 0$). A similar result may be obtained by using the cosmological calculator at Ref. [5]. Thus, in the standard cosmology $z \sim 7$ corresponds to light that was emitted about 750 million years after the big bang.

16.9 Take $v_\mathrm{p} = 200\,\mathrm{km\,s^{-1}}$ as a representative peculiar velocity. From the Hubble law with $H_0 = 72\,\mathrm{km\,s^{-1}Mpc^{-1}}$, a distance of $d = v_\mathrm{p}/H_0 = 2.8\,\mathrm{Mpc}$ corresponds to the distance where the cosmological recessional velocity is comparable to the average peculiar velocity in magnitude. Thus, a distance of 10 times that or about 30 Mpc is necessary before the peculiar velocities are only $\sim 10\%$ of the cosmological recessional velocities.

16.10 Supernova SN 2015A was a Type Ia supernova in the Sb spiral galaxy NGC 2995 with a redshift $z = 0.023329$ (corresponding to $cz = 6998.7\,\mathrm{km\,s^{-1}}$). The Hubble law is approximately valid and for small distances $v \sim cz$, so the distance to the supernova is

$$d \sim \frac{v}{H_0} \sim \frac{cz}{H_0} = 97.2\,\mathrm{Mpc},$$

where $H_0 = 72\,\mathrm{km\,s^{-1}Mpc^{-1}}$ has been assumed.

17 Energy and Matter in the Universe

17.1 The energy density in the CMB is given by the blackbody formula with $T = 2.725$ K,

$$\varepsilon_{\text{CMB}} = aT^4 = 4.17 \times 10^{-13}\,\text{erg}\,\text{cm}^{-3} = 2.6 \times 10^{-7}\,\text{MeV}\,\text{cm}^{-3}.$$

The corresponding density parameter is

$$\Omega_{\text{CMB}} = \frac{\varepsilon_{\text{CMB}}}{\varepsilon_{\text{crit}}} = \frac{2.6 \times 10^{-7}\,\text{MeV}\,\text{cm}^{-3}}{1.05 \times 10^{-2}h^2\,\text{MeV}\,\text{cm}^{-3}} = 4.8 \times 10^{-5},$$

assuming $h = 0.72$. For starlight we may estimate roughly,

$$\begin{aligned}\varepsilon_{\text{star}} &\simeq (2.6 \times 10^{-32}\,\text{erg}\,\text{s}^{-1}\text{cm}^{-3}) \times (14 \times 10^9\,\text{yr}) \times (3.16 \times 10^7\,\text{s}\,\text{yr}^{-1}) \\ &\simeq 1.15 \times 10^{-14}\,\text{erg}\,\text{cm}^{-3},\end{aligned}$$

if the Universe is 14 billion years old and therefore

$$\frac{\varepsilon_{\text{star}}}{\varepsilon_{\text{CMB}}} \simeq \frac{1.15 \times 10^{-14}\,\text{erg}\,\text{cm}^{-3}}{4.17 \times 10^{-13}\,\text{erg}\,\text{cm}^{-3}} = 0.028.$$

Therefore, the radiation density parameter for the present Universe,

$$\Omega_{\text{r}} \simeq \Omega_{\text{CMB}} \simeq \mathscr{O}(10^{-5}),$$

is dominated by the cosmic microwave background and is negligible compared with the matter density ($\Omega_{\text{m}} \simeq 0.3$) and the vacuum energy density ($\Omega_\Lambda \simeq 0.7$).

17.2 From Eqs. (17.15)–(17.16),

$$\dot{a}^2 = a_0^2 H_0^2 \left(1 + \frac{a_0}{a}\Omega_0 - \Omega_0\right).$$

Choose $\Omega_0 = 1$ (flat Universe), take the square root of both sides, and then integrate both sides to give

$$\int_0^a a^{1/2} da = a_0^{3/2} H_0 \int_0^t dt.$$

Performing the integrals and solving for a/a_0 gives

$$\frac{a(t)}{a_0} = \left(\frac{2}{3}\right)^{3/2} \left(\frac{t}{t_{\text{H}}}\right)^{2/3},$$

where $t_{\text{H}} = H_0^{-1}$ has been used. The redshift z is given by $a_0/a(t) = 1 + z$. Choosing by convention $a_0 = 1$ gives

$$\frac{t(z)}{t_{\text{H}}} = \tfrac{2}{3}(1+z)^{-3/2}.$$

Since $z = 0$ today, the age of a flat, dust-filled Universe is $t_0 \equiv t(z=0) = \frac{2}{3}t_{\text{H}}$.

17.3 From Eq. (17.18), for a closed universe

$$\frac{a}{a_0} = \frac{\Omega}{2(\Omega-1)}(1-\cos\psi),$$

and since $1+z = a_0/a$,

$$\cos\psi = 1 - \frac{2(\Omega-1)}{\Omega(1+z)}.$$

Also, from Eq. (17.19),

$$\psi - \sin\psi = \frac{2(\Omega-1)^{3/2}}{\Omega}\left(\frac{t}{t_{\rm H}}\right),$$

where $t_{\rm H} = 1/H_0$ was used. Combining these relations and a substantial amount of algebra then gives

$$\frac{t(z)}{t_{\rm H}} = \frac{\Omega}{2(\Omega-1)^{3/2}}\left[\cos^{-1}\left(\frac{\Omega z - \Omega + 2}{\Omega + \Omega z}\right) - \frac{2(\Omega^2 z + \Omega - \Omega z - 1)^{1/2}}{\Omega + \Omega z}\right],$$

which reduces to Eq. (17.25) upon setting $z = 0$. For an open universe, start from Eqs. (17.22) and (17.23) and proceed in a way analogous to above. After substantial algebra, one obtains

$$\frac{t(z)}{t_{\rm H}} = \frac{\Omega}{2(1-\Omega)^{3/2}}\left[\frac{2(-\Omega^2 z - \Omega + \Omega z + 1)^{1/2}}{\Omega + \Omega z} - \cosh^{-1}\left(\frac{\Omega z - \Omega + 2}{\Omega + \Omega z}\right)\right],$$

which reduces to Eq. (17.24) upon setting $z = 0$.

17.4 Calculate as for Example 17.2 using the formulas derived in Problems 17.2 and 17.3 (which are quoted in Example 17.1).

17.5 Balance of centripetal and gravitational forces for a circular orbit requires that $v^2(r) = GM(r)/r$. Thus, if the velocity is constant with increasing r the mass must be increasing as $M(r) \propto r$.

17.6 From the preceding problem, $v(r)^2 = GM(r)/r$. Thus,

$$M(r) = \frac{rv(r)^2}{G} = 2.32\times 10^5\left(\frac{r}{\rm kpc}\right)\left(\frac{v}{\rm km\,s^{-1}}\right)^2 M_\odot.$$

At the Sun's distance of about 8 kpc from the center the velocity $v \sim 220\,{\rm km\,s^{-1}}$ gives $M(8\,{\rm kpc}) \simeq 9\times 10^{10} M_\odot$. At 60,000 light years (18.4 kpc) from the center the measured velocity $v \sim 230\,{\rm km\,s^{-1}}$ gives $M(18.4\,{\rm kpc}) \simeq 2.3\times 10^{11} M_\odot$.

17.7 One has that

$$\langle v_i^2\rangle = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-v_i^2/2\sigma_i^2} v_i^2\, dv_i$$

Letting $x \equiv v_i/\sqrt{2}\sigma_i$ gives

$$\langle v_i^2\rangle = \frac{2\sigma_i^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty} x^2 e^{-x^2} dx = \sigma_i^2.$$

Therefore,

$$\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 \equiv \sigma^2$$

and $\langle v^2 \rangle = \sigma^2$.

17.8 From the Hubble law $v = H_0 r$ and for small v/c the redshift is $z \simeq v/c$. Thus, $z \simeq H_0 r/c$. Expand the scale factor to first order in time,

$$a(t) \simeq a_0 - \left.\frac{da}{dt}\right|_{t_0} (t_0 - t) \equiv a_0 - \dot{a}_0 \Delta t.$$

The redshift can also be written as

$$\begin{aligned} z &= \frac{\lambda_0}{\lambda} - 1 = \frac{a_0}{a} - 1 = \frac{a_0 - a}{a} \\ &\simeq \frac{a_0 - a_0 + \dot{a}_0 \Delta t}{a_0 - \dot{a}_0 \Delta t} \simeq \frac{\dot{a}_0}{a_0} \Delta t = \frac{\dot{a}_0}{a_0} \frac{r}{c}, \end{aligned}$$

where $r = c\Delta t$ was used in the last step. Comparing with the earlier expression $z \simeq H_0 r/c$ gives $\dot{a}_0 / a_0 = H_0$.

17.9 The coordinate distance is

$$d = c \int_t^{t_0} \frac{dt'}{a(t')} \simeq \frac{c}{a_0} \int_t^{t_0} \left(1 - H_0(t' - t_0)\right) dt',$$

where in the second step the expansion (17.36) was inserted, terms quadratic and higher in $t - t_0$ were dropped, and the integrand was expanded in a binomial series. Performing the integration gives

$$d = \frac{c}{a_0} \left((t_0 - t) + \tfrac{1}{2} H_0 (t_0 - t)^2\right),$$

which is Eq. (17.39) with $a_0 = 1$.

17.10 Equation (17.36) follows directly from inserting Eqs. (17.34)–(17.35) into Eq. (17.33). From $z \equiv a_0 a^{-1} - 1$, and (17.36) for a,

$$z = \left(1 + H_0(t - t_0) - \tfrac{1}{2} H_0^2 q_0 (t - t_0)^2\right)^{-1} - 1.$$

Then Eq. (17.37) follows from a binomial expansion $(1 + x)^{-1} \simeq 1 - x + x^2$, with terms higher than second order in $(t - t_0)$ discarded. Equation (17.37) is a quadratic equation in $(t_0 - t)$, which gives Eq. (17.38) when solved by the usual quadratic formula with the positive solution, and with the square root expanded according to $(1 + x)^{1/2} \simeq 1 + \frac{1}{2}x - \frac{1}{8}x^2$.

17.11 Inserting Eq. (17.38) in Eq. (17.39) and keeping terms of quadratic order or lower gives

$$\begin{aligned} d(t_0) &= \frac{c}{H_0} \left(z - (1 + \tfrac{1}{2} q_0) z^2 + \tfrac{1}{2} z^2\right) \\ &= \frac{cz}{H_0} \left(1 - \frac{1 + q_0}{2} z\right), \end{aligned}$$

which is Eq. (17.40).

17.12 From the identity $\ddot{a} = \frac{1}{2}d(\dot{a}^2)/da$ and Eq. (17.14),

$$\ddot{a} = \frac{1}{2}\frac{d}{da}\dot{a}^2 = \frac{-\frac{1}{2}H_0^2 a_0^3 \Omega}{a^2}.$$

Solving this for $d\dot{a}^2$ and integrating from the present time t_0 to a time t,

$$\int_{a_0}^{a} d\dot{a}^2 = -H_0^2 a_0^3 \Omega \int_{a_0}^{a} \frac{da}{a^2}.$$

Evaluating the integrals gives

$$\dot{a}^2 = \dot{a}_0^2 + H_0^2 a_0^3 \Omega \left(\frac{1}{a} - \frac{1}{a_0}\right),$$

and since $\dot{a}_0 = a_0 H_0$ (see Problem 17.8),

$$\dot{a}^2 = a_0^2 H_0^2 \left(1 + \Omega\frac{a_0}{a(t)} - \Omega,\right),$$

which is Eq. (17.15).

18 Friedmann Cosmologies

18.1 (a) The Friedmann equations for $P = 0$ and $k = 0$ are

$$3\dot{a}^2 = 8\pi G\rho a^2 \qquad \dot{\rho} + 3\rho\left(\frac{\dot{a}}{a}\right) = 0.$$

Rewrite the first equation as

$$da = \sqrt{\frac{8\pi G\rho}{3}}\, a\, dt$$

and the second is clearly satisfied by the condition $\rho(t)a(t)^3 = \rho_0$, where ρ_0 is the present density. Thus, upon substituting $\rho = \rho_0/a^3$ into the expression for da and solving for dt we obtain

$$dt = \sqrt{\frac{3}{8\pi G\rho_0}}\, a^{1/2} da.$$

Integrating this expression gives

$$t = \frac{2}{3}\sqrt{\frac{3}{8\pi G\rho_0}}\, a^{3/2} + \text{constant}.$$

Requiring that $a = 0$ when $t = 0$ implies that the constant is zero, so

$$a(t) = (6\pi G\rho_0)^{1/3} t^{2/3}$$

for the evolution of the scale factor.

(b) From this expression for $a(t)$, the Hubble parameter is $H = \dot{a}/a = 2/3t$, the deceleration parameter is

$$q_0 = -\frac{1}{H^2 a}\frac{d^2 a}{dt^2} = \frac{1}{2},$$

and taking $t = 0$ as the beginning, the age of the universe is just $t = 2/3H_0$, where H_0 is the current value of H.

18.2 The condition $T^{\mu\nu}{}_{;\nu} = 0$ with $T^{\mu\nu} = (\varepsilon + P)u^\mu u^\nu + Pg^{\mu\nu}$ implies that

$$\frac{\partial T^{\mu\nu}}{\partial x^\nu} + \Gamma^\mu_{\alpha\nu}T^{\alpha\nu} + \Gamma^\nu_{\alpha\nu}T^{\mu\alpha} = 0.$$

The non-zero components of $T_{\mu\nu}$ are obtained from Eq. (18.31) and the corresponding $T^{\mu\nu}$

can be obtained from these by contraction with the R–W metric tensor (18.16):

$$\begin{aligned}
T_{00} &= \varepsilon & T^{00} &= \varepsilon \\
T_{11} &= Pa^2/(1-kr^2) & T^{11} &= P(1-kr^2)/a^2 \\
T_{22} &= Pr^2a^2 & T^{22} &= P/(ra)^2 \\
T_{33} &= Pr^2a^2\sin^2\theta & T^{33} &= P/(ra\sin\theta)^2,
\end{aligned}$$

where a is the scale parameter. Consider the $\mu = 0$ component:

$$\frac{\partial T^{0\nu}}{\partial x^\nu} + \Gamma^0_{\alpha\nu}T^{\alpha\nu} + \Gamma^\nu_{0\nu}T^{00} = 0,$$

which is explicitly

$$\Gamma^0_{00}T^{00} + \Gamma^0_{11}T^{11} + \Gamma^0_{22}T^{22} + \Gamma^0_{33}T^{33} + \Gamma^0_{00}T^{00} \\ + \Gamma^1_{01}T^{00} + \Gamma^2_{02}T^{00} + \Gamma^3_{03}T^{00} = 0.$$

The required Christoffel symbols are given in Table 18.1,

$$\Gamma^0_{00} = 0 \qquad \Gamma^0_{11} = \frac{a\dot{a}}{1-kr^2} \qquad \Gamma^0_{22} = r^2a\dot{a} \qquad \Gamma^0_{33} = r^2\sin^2\theta a\dot{a}$$
$$\Gamma^1_{01} = \frac{\dot{a}}{a} \qquad \Gamma^2_{02} = \frac{\dot{a}}{a} \qquad \Gamma^3_{03} = \frac{\dot{a}}{a}$$

and from the preceding expressions for the stress–energy tensor components

$$\frac{\partial T^{0\nu}}{\partial x^\nu} = \frac{\partial T^{00}}{\partial x^0} = \dot{\varepsilon}.$$

Inserting these results into the previous expression for the $\mu = 0$ component and collecting terms gives

$$\dot{\varepsilon} + 3(\varepsilon + P)\frac{\dot{a}}{a} = 0,$$

which expresses conservation of mass–energy.

18.3 Let's use the Robertson–Walker metric in the form (18.19). For positive curvature the spatial line element is

$$d\ell^2 = a^2(d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\varphi^2)),$$

so the metric is diagonal with

$$g_{11} = a^2 \qquad g_{22} = a^2\sin^2\chi \qquad g_{33} = a^2\sin^2\chi\sin^2\theta$$

and

$$\sqrt{\det g} = a^3\sin^2\chi\sin\theta.$$

Then the volume is (see Section 3.13.1; here a positive sign is used under the square root because $\det g$ is positive for the spatial part of the metric)

$$\begin{aligned}V &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \int_0^{\pi} \sqrt{\det g}\, d\chi \\ &= a^3 \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \int_0^{\pi} \sin^2\chi d\chi \\ &= 2\pi^2 a^3.\end{aligned}$$

For negative curvature the line element is

$$d\ell^2 = a^2(d\chi^2 + \sinh^2\chi(d\theta^2 + \sin^2\theta d\varphi^2)),$$

so the metric is diagonal with

$$g_{11} = a^2 \qquad g_{22} = a^2\sinh^2\chi \qquad g_{33} = a^2\sinh^2\chi\sin^2\theta$$

and

$$\sqrt{\det g} = a^3\sinh^2\chi\sin\theta.$$

Therefore, for negative curvature the volume is

$$V = 4\pi a^3 \int_0^{\infty} \sinh^2\chi d\chi = \infty.$$

For a flat metric the line element is

$$d\ell^2 = a^2(d\chi^2 + \chi^2(d\theta^2 + \sin^2\theta d\varphi^2)),$$

and proceeding as above the volume for flat space is

$$V = 4\pi a^3 \int_0^{\infty} \chi d\chi = \infty.$$

Thus, the volume of a spatial slice described by a Robertson–Walker metric with positive curvature is finite but the volume of a spatial slice described by a R–W metric with negative curvature or no curvature is infinite.

18.4 This problem is suggested by a discussion in Padmanabhan [173]. Taking $a_0 = a(0) = 1$, the current horizon distance is given by

$$\ell_{\mathrm{h}} = \int_0^{t_0} \frac{cdt}{a(t)} = \int_0^{t_0} \frac{cdt}{t^n} = \frac{t_0}{1-n}.$$

The corresponding Hubble radius is

$$d_{\mathrm{H}} = \frac{1}{H_0} = \left(\frac{\dot{a}}{a}\right)^{-1}_{t=t_0} = \left(\frac{d(t^n)/dt}{t^n}\right)^{-1}_{t=t_0} = \frac{t_0}{n}.$$

Thus $\ell_{\mathrm{h}}/d_{\mathrm{H}} = n/(1-n) \sim 1$ for radiation-dominated ($n = \frac{1}{2}$) or matter-dominated ($n = \frac{2}{3}$) cosmologies. However, ℓ_{h} and d_{H} are conceptually rather different since d_{H} is determined only by the behavior near $t = t_0$ but ℓ_{h} is sensitive to the entire past history of the Universe. Hence, for cosmologies in which the scale factor is not of the form $a(t) = t^n$ with $\frac{1}{2} \le n \le \frac{2}{3}$ the Hubble distance could be a poor indicator of the horizon distance.

18.5 This is mostly a matter of substitution and a lot of algebra. For example, in the positive-curvature (closed universe) case Eqs. (18.17) may be used to write

$$\begin{aligned} dw &= -\sin\chi d\chi \\ dz &= -\sin\chi\sin\theta d\theta + \cos\chi\cos\theta d\chi \\ dx &= -\sin\chi\sin\theta\sin\varphi d\varphi + \sin\chi\cos\theta\cos\varphi d\theta + \cos\chi\sin\theta\cos\varphi d\chi \\ dy &= \sin\chi\sin\theta\cos\varphi d\varphi + \sin\chi\cos\theta\sin\varphi d\theta + \cos\chi\sin\theta\sin\varphi d\chi. \end{aligned}$$

Substitution of these derivatives in

$$d\ell^2 = dx^2 + dy^2 + dz^2 + dw^2,$$

and a substantial amount of algebra abetted by use of trigonometric identities leads to

$$d\ell^2 = d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\varphi^2),$$

which gives the first of Eqs. (18.19) when substituted in Eq. (18.2). The negative curvature and flat cases may be proved in a similar way. The equivalence of Eqs. (18.14) and (18.19) may be proved by using the change of variables in Eq. (18.20). For example, in the negative curvature case substitute

$$r = \sinh\chi \longrightarrow \sinh^2\chi = r^2 \qquad dr = \cosh\chi d\chi \longrightarrow d\chi^2 = \frac{dr^2}{\cosh^2\chi}$$

from Eq. (18.20) in the spatial part of the third equation in Eq. (18.19) to give

$$d\ell^2 = \left(\frac{dr^2}{\cosh^2\chi} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\right).$$

Upon substituting $\cosh^2\chi = 1 + \sinh^2\chi = 1 + r^2$, this becomes

$$d\ell^2 = \left(\frac{dr^2}{1+r^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\right),$$

which gives upon substitution in Eq. (18.2) the line element (18.14) with the curvature choice $k = -1$.

18.6 Writing the implied sum over ν in Eq. (7.30) out explicitly,

$$\Gamma^\sigma_{\lambda\mu} = \sum_{\nu=0}^{3} \tfrac{1}{2} g^{\nu\sigma}\left(\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu}\right).$$

But the metric is diagonal so only the term with $\nu = \sigma$ survives in the sum. Then, for example,

$$\begin{aligned} \Gamma^2_{02} &= \tfrac{1}{2} g^{22}\left(\frac{\partial g_{22}}{\partial x^0} + \frac{\partial g_{02}}{\partial x^2} - \frac{\partial g_{20}}{\partial x^2}\right) = \tfrac{1}{2} g^{22}\frac{\partial g_{22}}{\partial x^0} \\ &= \frac{1}{2}\frac{1}{a^2 r^2} r^2 \frac{\partial(a^2)}{\partial t} = \frac{\dot a}{a}, \end{aligned}$$

where the metric given in Eqs. (18.15) and (18.16) has been used and $(x^0, x^1, x^2, x^3) = (t, r, \theta, \varphi)$. The other $\Gamma^\sigma_{\lambda\mu}$ may be derived in similar fashion, with the results summarized in Table 18.1. The Ricci tensor may be constructed from the connection coefficients and Eq. (8.16),

$$R_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu,\lambda} - \Gamma^{\lambda}_{\mu\lambda,\nu} + \Gamma^{\lambda}_{\mu\nu}\Gamma^{\sigma}_{\lambda\sigma} - \Gamma^{\sigma}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma}.$$

For example, utilizing Table 18.1 and remembering that the only time dependence is in the scale factor a, the non-vanishing terms for $\mu = \nu = 1$ are

$$\begin{aligned} R_{11} = {} & \frac{\partial\Gamma^0_{11}}{\partial t} - \frac{\partial\Gamma^2_{12}}{\partial r} - \frac{\partial\Gamma^3_{13}}{\partial r} \\ & + \Gamma^0_{11}\Gamma^2_{02} + \Gamma^0_{11}\Gamma^3_{03} + \Gamma^1_{11}\Gamma^2_{12} + \Gamma^1_{11}\Gamma^3_{13} - \Gamma^0_{11}\Gamma^1_{10} - \Gamma^2_{12}\Gamma^2_{12} - \Gamma^3_{13}\Gamma^3_{13}. \end{aligned}$$

Inserting explicit forms for the connection coefficients from Table 18.1 and a little algebra then gives

$$R_{11} = \frac{a\ddot{a} + 2\dot{a}^2 + 2k}{1 - kr^2}.$$

The Ricci scalar (8.17) is obtained by contraction with the metric tensor, $R = g^{\mu\nu}R_{\mu\nu}$. The metric coefficients are given in Eq. (18.16) and the Ricci tensor components are given in Eq. (18.32). Inserting them and carrying out some algebra yields

$$R = g^{\mu\nu}R_{\mu\nu} = g^{00}R_{00} + g^{11}R_{11} + g^{22}R_{22} + g^{33}R_{33} = \frac{6(a\ddot{a} + \dot{a}^2 + k)}{a^2}$$

for the Ricci scalar.

18.7 This proof is completely analogous to that for Problem 8.11, except that one obtains from the Einstein equation an additional term for R_{00} depending on Λ:

$$R_{00} = \frac{4\pi G}{c^2}\rho + \Lambda g_{00} = \frac{4\pi G}{c^2}\rho - \Lambda.$$

Then carrying through exactly as in Problem 8.11 yields

$$\nabla^2\varphi = 4\pi G\rho - \Lambda c^2,$$

which is the Poisson equation modified by an additional term proportional to the cosmological constant Λ.

18.8 As shown in Problem 18.2, the constraint $T^{\mu\nu}_{;\nu} = 0$ evaluated for $\mu = 0$ gives

$$\dot{\varepsilon} + 3(\varepsilon + P)\frac{\dot{a}}{a} = 0.$$

This is equivalent to

$$\frac{d}{dt}(\varepsilon a^3) = -P\frac{d}{dt}a^3.$$

A comoving volume of space has a physical volume $V = \alpha a^3$, where α is a constant. The total energy contained in this volume is $U = \alpha\varepsilon a^3$. Substituting this in the above equation gives

$$\frac{d}{dt}\left(\frac{U}{\alpha}\right) = -P\frac{d}{dt}\left(\frac{V}{\alpha}\right) \quad \rightarrow \quad dU = -PdV.$$

The first law of thermodynamics is $dU = -PdV + \delta Q$, but by the cosmological principle there can be no heat differences δQ on large scales, so the preceding result is just the first law of thermodynamics in an expanding universe described by the Robertson–Walker metric.

18.9 Consider vectors u and v. The angle between them is given by

$$\cos\theta = \frac{v \cdot u}{|v||u|} = \frac{g_{\lambda\sigma} v^\lambda u^\sigma}{\sqrt{g_{\mu\nu} v^\mu v^\nu}\sqrt{g_{\alpha\beta} u^\alpha u^\beta}}.$$

Under a conformal transformation $g_{\gamma\delta} \to \Phi g_{\gamma\delta}$, where Φ is a scalar field (see Section 18.8.1). Clearly the above expression for θ is not modified by this transformation of the metric tensor components since the scalar field factors Φ cancel.

18.10 Introduce the conformal time η through $dt = a(t)d\eta$. The flat ($k = 0$) Robertson–Walker metric (18.14) is then given by

$$\begin{aligned} ds^2 &= -dt^2 + a^2(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2) \\ &= -a^2 d\eta^2 + a^2(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2) \\ &= a^2(-d\eta^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2), \end{aligned}$$

which is the same form as the metric for a uniformly-expanding Minkowski space. For radial light rays $d\theta = d\varphi = ds = 0$, so the line element implies that $a^2(-d\eta^2 + dr^2) = 0$ and hence $d\eta = \pm dr$. Thus, in the η–r plane light rays move at 45-degree angles at all times.

19 Evolution of the Universe

19.1 From the Friedmann equations,

$$\frac{\dot{\varepsilon}}{\varepsilon} + 3\left(1 + \frac{P}{\varepsilon}\right)\frac{\dot{a}}{a} = 0.$$

If an equation of state $P = w\varepsilon$ is assumed, this may be written as

$$\frac{d\varepsilon}{\varepsilon} = -(3+3w)\frac{da}{a}.$$

Assume w to be constant and integrate both sides from now ($a = a_0$ and $\varepsilon = \varepsilon_0$) until some scale factor a to give

$$\ln\frac{\varepsilon(a)}{\varepsilon_0} = \ln\left(\frac{a}{a_0}\right)^{-3(1+w)},$$

which implies that

$$\varepsilon(a) = \varepsilon_0\left(\frac{a}{a_0}\right)^{-3(1+w)}.$$

Thus matter ($w \simeq 0$), radiation ($w = \frac{1}{3}$), and vacuum energy ($w < -\frac{1}{3}$) have very different histories: for matter, $\varepsilon_{\mathrm{m}} \propto a^{-3}$; for radiation, $\varepsilon_{\mathrm{r}} \propto a^{-4}$; for vacuum energy, ε_Λ is constant, if it is assumed that $w = -1$ (as implied by a cosmological constant).

19.2 Take the Solar System to be a sphere with radius 40 AU, assume the energy density of radiation to be negligible, and assume the total mass to be approximately that of the Sun. The vacuum energy density is from observations about 70% of critical density, so

$$\varepsilon_{\mathrm{vac}} \simeq 0.7 \times 1.05 \times 10^{-2} h^2\,\mathrm{MeV\,cm^{-3}} \simeq 3.8 \times 10^{-3}\,\mathrm{MeV\,cm^{-3}},$$

where $h = 0.72$ was assumed. Multiplying this number by the volume of the Solar System gives $E_{\mathrm{vac}} \simeq 3.4 \times 10^{42}$ MeV for the total vacuum energy contained in the Solar System. The total energy of the matter in the Solar System is then estimated as $E_{\mathrm{matter}} \simeq (1M_\odot)c^2 \simeq 1.1 \times 10^{60}$ MeV. Thus, the ratio of vacuum energy to total energy in the Solar System is of order 10^{-18}. This suggests that a local experiment to measure the vacuum energy would be very difficult.

19.3 (a) Assume air to be an ideal gas of nitrogen molecules N_2 with mass $\mu \sim 2 \times 14 = 28$ amu at $T = 300$ K. From Eq. (19.10), $w = kT/\mu c^2$. Thus w is basically the ratio of the thermal energy to the rest mass energy, which will be small for a nonrelativistic gas. Inserting the numbers gives $w = 9.9 \times 10^{-13}$.

(b) For an ideal gas $\langle v^2 \rangle = 3kT/\mu$, implying that $w = kT/\mu c^2 = \langle v^2 \rangle / 3c^2$. Thus w may also be interpreted as the ratio of the average of the velocity squared to the speed of light squared, which is a very small number for nonrelativistic gases.

19.4 For a flat, matter-only universe Table 19.1 indicates that

$$a(t) = \left(\frac{t}{t_0}\right)^{2/3} \qquad t_0 = \frac{2}{3H_0},$$

and at time t the proper distance to a photon emitted from the present particle horizon at time $t = 0$ is given by

$$\begin{aligned}\ell(t) &= a(t)\int_t^{t_0} \frac{dt'}{a(t')} = a(t)\int_t^{t_0}\left(\frac{t'}{t_0}\right)^{-2/3} dt' \\ &= 3t_0^{2/3}a(t)\left(t_0^{1/3} - t^{1/3}\right) \\ &= 3t_0\left[\left(\frac{t}{t_0}\right)^{2/3} - \frac{t}{t_0}\right],\end{aligned}$$

where t_0 is the time today (age of Universe). Even though the coordinate distance of the photon from Earth decreases continuously, the proper distance first *increases* after emission, eventually reaches a maximum, and then decreases to zero as the photon reaches Earth. The maximum proper distance is obtained by requiring that $d\ell/dt = 0$, which gives that the time of maximum proper distance corresponds to $t(\ell_{\rm max}) = 0.296\,t_0$, and using that the current horizon distance is $2/H_0$ (Table 19.1) the maximum proper separation of the photon from Earth at $t = 0.296\,t_0$ is found to be 0.148 times the current horizon distance.

19.5 This is called the Milne universe. From Eqs. (19.57)–(19.59), the equations of motion are given by

$$dq = \sqrt{2(E-U)}\,d\tau \qquad U = -\frac{1}{2}\left(\frac{\Omega_{\rm r}}{q^2} + \frac{\Omega_{\rm m}}{q} + \Omega_\Lambda q^2\right) \qquad E = \frac{1}{2}(1-\Omega).$$

The Milne universe contains only curvature, so $U = 0$ and $\Omega \equiv \Omega_{\rm r} + \Omega_{\rm m} + \Omega_\Lambda = 0$ and $E = \frac{1}{2}$. Inserting these in the first equation above and integrating gives $q(t) - q(0) = \tau$. But $q \equiv a(t)/a_0$ so $q(0) = 0$ and $\tau \equiv H_0 t$, and from Eq. (19.61) the age of the Universe is $t_0 = H_0^{-1}$. Combining these results gives

$$\frac{a(t)}{a_0} = \frac{t}{t_0}.$$

Note that this solution implies negative curvature since from Eqs. (19.23)–(19.26)

$$\Omega_k = 1 - \Omega = 1 = \frac{-k}{a^2 H_0^2}$$

and thus $k = -a^2 H_0^2$. The proper distance at the time of detection is

$$\ell(t_0) = a_0 \int_{t_{\rm e}}^{t_0} \frac{dt}{a(t)} = t_0 \int_{t_{\rm e}}^{t_0} \frac{dt}{t} = t_0 \ln\left(\frac{t_0}{t_{\rm e}}\right),$$

where $a_0/a(t) = t_0/t$ was used and t_e is the time of emission. The redshift is

$$z = \frac{a_0}{a(t_e)} - 1 = \frac{t_0}{t_e} - 1.$$

The proper distance at the current time in terms of redshift is then

$$\ell(t_0) = t_0 \ln\left(\frac{t_0}{t_e}\right) = t_0 \ln(1+z) = \frac{1}{H_0}\ln(1+z),$$

and the proper distance at the time of emission is smaller by a factor $a(t_e)/a_0 = t_e/t = (1+z)^{-1}$, giving

$$\ell(t_e) = \frac{1}{1+z}\ell(t_0) = \frac{1}{H_0}\frac{\ln(1+z)}{1+z},$$

for the Milne universe.

19.6 From Eq. (19.12) generalized to multiple components (with factors of c restored)

$$\ddot{a} = -\frac{4\pi G a}{3c^2}\sum_i(\varepsilon_i + 3P_i) = -\frac{4\pi G a}{3c^2}\sum_i \varepsilon_i(1+3w_i),$$

where each component has an equation of state $P_i = w_i\varepsilon_i$. Multiply both sides by $-(aH^2)^{-1}$ and invoke the definition (17.34) of the deceleration parameter q_0 to write this as

$$q_0 \equiv \left(\frac{\ddot{a}}{aH^2}\right)_{t=t_0} = \frac{4\pi G}{3H_0^2c^2}\sum_i \varepsilon_i(1+3w_i),$$

where $H = \dot{a}/a$ was used. But from Eq. (17.6) the critical density is $\varepsilon_c = \rho_c c^2 = 3H_0^2c^2/8\pi G$, so

$$q_0 = \frac{1}{2}\sum_i \Omega_i(1+3w_i),$$

where $\Omega_i \equiv \varepsilon_i/\varepsilon_c$ has been used. Taking a 3-component Universe with matter ($w=0$), radiation ($w=\frac{1}{3}$), and vacuum energy ($w=-1$), this reduces to

$$q_0 = \tfrac{1}{2}\Omega_m + \Omega_r - \Omega_\Lambda.$$

This Universe can exhibit acceleration ($q_0 < 0$) only if it contains sufficient vacuum energy such that $\Omega_\Lambda > \frac{1}{2}\Omega_m + \Omega_r$. With the further assumption of a flat Universe (which requires that $\Omega_m + \Omega_r + \Omega_\Lambda = 1$),

$$q_0 = \tfrac{3}{2}\Omega_m + 2\Omega_r - 1.$$

Finally, taking for the standard cosmology $\Omega_m = 0.3$, negligible Ω_r, and $\Omega_\Lambda = 0.7$, one obtains $q_0 = -0.55$. Since the deceleration parameter is negative, the standard cosmology represents a Universe that is currently accelerating, with its density parameters satisfying $\Omega_m + \Omega_r + \Omega_\Lambda = 1$.

19.7 (a) Using Eq. (18.27) with factors of c restored and assuming $a(t) \propto t^n$ with $n < 1$, the proper horizon distance is

$$\ell_h = a(t)\int_0^t \frac{c\,dt'}{a(t')} = ct^n\int_0^t (t')^{-n}dt' = \frac{ct}{1-n}.$$

The ratio of this horizon distance to the scale factor is then

$$\frac{\ell_h(t)}{a(t)} = \frac{ct}{1-n}\cdot\frac{1}{t^n} = \left(\frac{c}{1-n}\right)t^{1-n},$$

which always grows with time since it has been assumed that $n < 1$.

(b) For a flat, matter-filled universe, Table 19.1 gives that $a(t) = (t/t_0)^{2/3}$, where $t_0 = 2/(3H_0)$ is the age. Thus, $n = \frac{2}{3}$ and from the preceding results, taking $t = t_0$ (time today),

$$\ell_h = \frac{ct_0}{1-n} = 3ct_0 = 3c\left(\frac{2}{3H_0}\right) = \frac{2c}{H_0}.$$

Assuming that $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, this yields a horizon distance of 8333 Mpc.

19.8 The result can be obtained immediately as a special case of the general solution in Problem 19.6 but we can obtain it directly from Eqs. (17.34), (17.6), and the Friedmann equation (19.12) for dust (assuming $P = 0$),

$$q_0 \equiv \left(-a\frac{\ddot{a}}{\dot{a}^2}\right)_{t=t_0} = -\frac{\ddot{a}_0}{a_0H_0^2} \qquad \rho_c = \frac{3H_0^2}{8\pi G} \qquad \frac{\ddot{a}_0}{a_0} = -\frac{4\pi G}{3}\rho.$$

Combining these gives

$$q_0 = \frac{1}{2}\frac{\rho}{\rho_c} = \frac{1}{2}\Omega,$$

where (17.8) was used.

19.9 Equation (19.40) may be written as

$$t_0 = \frac{2}{3(1+w)}\tau_H,$$

where $\tau_H = 1/H_0$ is the Hubble time. Hence in a single-component, flat universe the age of the universe is greater than τ_H if $w < -\frac{1}{3}$ and less than τ_H if $w > -\frac{1}{3}$.

19.10 For the angular part of the metric (18.19) displayed in Box 19.3, we have $\sqrt{\det g} = S_k^2 \sin\theta$ (where $a_0 \equiv 1$) and the spherical surface has a proper area $A_p(t_0)$ that differs from $4\pi r^2$ in curved space:

$$A_p(t_0) = \int_0^{2\pi} d\varphi \int_0^{\pi} \sqrt{\det g}\, d\theta = 4\pi S_k(r)^2.$$

This is a consequence of curvature, independent of whether the Universe is expanding. If the Universe is flat, $S_k = r$ and an area of $4\pi r^2$ is recovered.

20 The Big Bang

20.1 Assume the Universe to have been matter dominated for most of it history, so $a(t) \sim t^{2/3}$. The coordinate distance between the source and the observer is

$$r \simeq \int_{t_r}^{t_0} \frac{dt}{t^{2/3}} = 3\left(t_0^{1/3} - t_r^{1/3}\right),$$

where t_0 is the present time and t_r is the time of hydrogen recombination. The coordinate horizon distance at t_r is

$$r_H = \int_0^{t_r} \frac{dt}{t^{2/3}} = 3t_r^{1/3}$$

and the number of horizons separating the sources in opposite directions is

$$n_H \simeq \frac{2r}{r_H} \simeq 2\left[\left(\frac{t_0}{t_r}\right)^{1/3} - 1\right] \simeq 90,$$

assuming $t_0 = 10^{10}$ yr and $t_r = 10^5$ yr.

20.2 At a temperature of 10^9 K it is assumed that the neutron/proton ratio is 0.164 as nucleosynthesis begins. This means that out of 1000 nucleons 141 are neutrons and 859 are protons. How many ^{4}He nuclei can be made from these 1000 nucleons? That is just a matter of counting. Each ^{4}He consists of 2 neutrons and 2 protons. Since there are 141 neutrons available for each 1000 nucleons, 141/2 = 71 ^{4}He nuclei (rounded to an integer) can be made. To make these helium nuclei requires one proton for every neutron used, so the number of protons left over after all the neutrons are bound into helium nuclei is $859 - 141 = 718$ protons. Hence for every 1000 initial nucleons, the Universe is left after the initial period of big bang nucleosynthesis with 71 helium nuclei, 718 protons, no free neutrons, and a trace of other light nuclei. Since ^{4}He is about 4 times as massive as a proton, the mass fractions are $(4 \times 71)/1000 = 0.28$ for ^{4}He and $718/1000 = 0.72$ for hydrogen. Therefore, these very simple considerations suggest that after big bang nucleosynthesis the baryonic matter of the Universe should be about 28 percent ^{4}He by mass, with most of the rest hydrogen. A much more sophisticated treatment of the problem that is summarized in Fig. 20.5 obtains results that are not too different from these estimates.

20.3 The current baryon and radiation density parameters are $\Omega_b(t_0) \simeq 0.04$ and $\Omega_r(t_0) \simeq 8 \times 10^{-5}$ (Table 17.1), and these need to be extrapolated back to a redshift (that is, scale factor) where they become equal. Because the radiation density scales as a^{-4} and the baryonic density as a^{-3},

$$\frac{\Omega_b(t)}{\Omega_r(t)} = \frac{a(t)^{-3}\Omega_b(t_0)}{a(t)^{-4}\Omega_r(t_0)} = a(t)\frac{\Omega_b(t_0)}{\Omega_r(t_0)} = \left(\frac{1}{1+z}\right)\frac{\Omega_b(t_0)}{\Omega_r(t_0)}$$

Requiring that the ratio on the left side be unity then gives a redshift of

$$z = \frac{\Omega_{\rm b}(t_0)}{\Omega_{\rm r}(t_0)} - 1 \simeq 500.$$

To estimate the corresponding time, assume that the vacuum energy can be neglected in subsequent evolution and that the Universe evolved as matter dominated from the unknown time t until today (t_0). For a matter-dominated Universe, $a \sim t^{2/3}$, so

$$\frac{a_0}{a} = 1 + z = \left(\frac{t_0}{t}\right)^{2/3},$$

and the corresponding time is

$$t = \left(\frac{1}{1+z}\right)^{3/2} t_0.$$

Taking the present time to be $t_0 \sim 13.8 \times 10^9$ yr and using $z = 500$ from above gives $t = 1.2 \times 10^6$ yr after the big bang for equality of baryonic and radiation energy densities. This result is approximately correct but a better treatment would account for things like the influence of the neglected vacuum energy on evolution of the more recent Universe. A more sophisticated relationship of redshift to time gives $t \sim 1.35 \times 10^6$ yr for equality of baryonic and radiation densities [5]. Notice that what we have calculated is the redshift for equality of the radiation density with the *baryonic part* of the matter density. From Table 17.1 the total matter density parameter $\Omega_{\rm m}$ is about 7.5 times larger than the baryonic density parameter, so the above estimate applied to the total matter density would give a redshift $z = 500 \times 7.5 = 3750$ for matter–radiation equality. This may be compared with the measured value of $z = 3365$ from Table 20.2.

20.4 (a) For baryon density parameter $\Omega_{\rm b} \simeq 0.04$ the baryon energy density is

$$\varepsilon_{\rm b} = \Omega_{\rm b}\varepsilon_{\rm c} = 2.18 \times 10^{-4}\,\text{MeV cm}^{-3},$$

where Eq. (17.7) with $h = 0.72$ was used.

(b) The average baryon is a proton with a rest mass energy of $\bar{E}_{\rm b} \sim 939$ MeV, so the number density of baryons is

$$n_{\rm b} = \frac{\varepsilon_{\rm b}}{\bar{E}_{\rm b}} \simeq 2.32 \times 10^{-7}\,\text{cm}^{-3}.$$

(c) Most photons are in the CMB, which is a near perfect blackbody, so the energy density is

$$\varepsilon_\gamma = aT^4 = 2.61 \times 10^{-7}\,\text{MeV cm}^{-3},$$

where $T = 2.725$ K was used.

(d) For a blackbody the mean energy is about $2.7kT$ so

$$\bar{E}_\gamma \simeq 2.7kT = 6.34 \times 10^{-10}\,\text{MeV}$$

and therefore we may estimate

$$n_\gamma = \frac{\varepsilon_\gamma}{\bar{E}_\gamma} = 411\,\text{cm}^{-3},$$

for the number density of CMB photons.

(e) Combining the above results,

$$\eta \equiv \frac{n_\mathrm{b}}{n_\gamma} = \frac{2.32\times 10^{-7}\,\mathrm{cm}^{-3}}{411\,\mathrm{cm}^{-3}} \simeq 5.6\times 10^{-10}.$$

for the ratio of baryon number density to photon number density.

20.5 The energy density for matter scales as $\varepsilon \sim a^{-3} \sim (1+z)^3$, where $1+z = a_0/a$. Thus, for the energy density of (cold) dark matter

$$\varepsilon_\mathrm{dm}(z_{\ell s}) = \Omega_\mathrm{dm}(t_0)\varepsilon_\mathrm{c}(t_0)(1+z_{\ell s})^3$$

where t_0 denotes current values. Since $\Omega_\mathrm{m} \sim 0.3$ and $\Omega_\mathrm{b} \sim 0.04$, take $\Omega_\mathrm{dm} = 0.26$. Using $z_{\ell s} = 1100$ and taking ε_c from Eq. (17.7) with $h = 0.72$ gives for the dark matter at last scattering

$$\varepsilon_\mathrm{dm}(z_{\ell s}) = 1.9\times 10^6\,\mathrm{MeV\,cm}^{-3} \qquad \rho_\mathrm{dm}(z_{\ell s}) = \frac{\varepsilon_\mathrm{dm}(z_{\ell s})}{c^2} = 3.38\times 10^{-21}\,\mathrm{g\,cm}^{-3}.$$

The baryon energy density scales with a in the same manner as cold dark matter, so for baryonic matter at last scattering $\varepsilon_\mathrm{b}(z_{\ell s}) = (\Omega_\mathrm{b}(t_0)/\Omega_\mathrm{dm}(t_0))\varepsilon_\mathrm{dm}(z_{\ell s})$, which gives

$$\varepsilon_\mathrm{b}(z_{\ell s}) = 2.9\times 10^5\,\mathrm{MeV\,cm}^{-3} \qquad \rho_\mathrm{b}(z_{\ell s}) = \frac{\varepsilon_\mathrm{b}(z_{\ell s})}{c^2} = 5.1\times 10^{-22}\,\mathrm{g\,cm}^{-3}.$$

Photon energy densities scale as $(1+z)^{-4}$, so the photon energy density at last scattering is

$$\varepsilon_\gamma(z_{\ell s}) = \frac{\Omega_\gamma(t_0)}{\Omega_\mathrm{dm}(t_0)}(1+z_{\ell s})\varepsilon_\mathrm{dm}(z_{\ell s}) = 3.9\times 10^5\,\mathrm{MeV\,cm}^{-3},$$

where $\Omega_\gamma(t_0) \sim 4.8\times 10^{-5}$ has been used. Thus we estimate that $\varepsilon_\mathrm{dm} : \varepsilon_\gamma : \varepsilon_\mathrm{b} \sim 6.6 : 1.4 : 1$ at decoupling. A more sophisticated determination of these energy densities based on CMB observations of the Planck satellite is given in Problem 20.11.

20.6 Estimate the age of the Universe as $\tau_\mathrm{H} \simeq H_0^{-1} \simeq 13.6\times 10^9$ yr. For a matter dominated Universe the scale factor is $a = bt^{2/3}$, where b is a constant and t is the time since the big bang. Evaluate b assuming $a(0) = 0$ and $a(13.6\times 10^9\,\mathrm{yr}) = 1$ to give $t = 1.36\times 10^{10} a^{3/2}$ yr. Taking the redshift at decoupling to be $z \simeq 1100$, the scale factor at decoupling is $a_\mathrm{dec} = (1+z)^{-1} \simeq 9.08\times 10^{-4}$ and the time of decoupling is

$$t_\mathrm{dec} = (1.36\times 10^{10})(9.08\times 10^{-4})^{3/2} = 3.7\times 10^5\,\mathrm{yr}.$$

Using the speed of light $c = 3.067\times 10^{-7}\,\mathrm{Mpc\,yr}^{-1}$, the maximum distance light could have traveled from the big bang to decoupling is $ct_\mathrm{dec} \simeq 0.113$ Mpc. This distance would have been expanded by a factor that is the ratio of the scale factors today and at decoupling:

$$0.113\,\mathrm{Mpc} \times (a_0/a_\mathrm{dec}) = 0.113\,\mathrm{Mpc} \times (1/9.08\times 10^{-4}) \simeq 124\,\mathrm{Mpc}.$$

In the time τ_H light could have traveled $c\tau_\mathrm{H} \simeq 4171$ Mpc. Using this to approximate the distance to the last scattering surface, one may estimate that the maximum angular size on the sky of a causally-connected region in the CMB is $\theta \simeq 124\,\mathrm{Mpc}/4171\,\mathrm{Mpc} = 0.03\,\mathrm{rad} \simeq$

1.7 degrees. This illustrates the horizon problem: the observed CMB appears to be causally correlated over much larger angular regions than this. (See also Problems 20.1 and 21.3)

20.7 The entropy density is given by (see Section 20.2.1),

$$s = \frac{2\pi^2}{45} g_* T^3,$$

where g_* is defined in Eq. (20.14). The difference in photon and neutrino background temperatures is because when electron–positron annihilation falls out of equilibrium at a temperature of about 0.2 MeV the annihilation energy raises the temperature of the photons but not of the neutrinos, because the neutrinos have fallen out of equilibrium at a somewhat higher temperature (weak-interaction decoupling). The amount of photon reheating can be estimated using the following considerations:

1. Entropy is expected to be conserved across the transition, so $g_* T^3$ before and after must be equivalent.
2. From Eq. (20.14), the contribution to g_* is (a) $1\times$ states of polarization for bosons and (b) $7/8 \times$ (states of polarization) for fermions.
3. Before the transition (at higher temperature) the relevant species and their contributions to g_* are
   - For e^- there are two spin states and the contribution is $\frac{7}{8} \times 2 = \frac{7}{4}$.
   - For e^+ there are two spin states and the contribution is $\frac{7}{8} \times 2 = \frac{7}{4}$.
   - For photons there are two polarization states and the relative contribution is 2.

   Thus the total g_* before the transition is $g_*^+ = \frac{7}{4} + \frac{7}{4} + 2 = \frac{11}{2}$.
4. After the transition there are (essentially) only photons so $g_*^- = 2$.

The ratio of entropy densities before (+) and after (−) the transition must satisfy

$$\frac{s_+}{s_-} = \frac{g_*^+ T_+^3}{g_*^- T_-^3} = 1,$$

because of the entropy conservation assumption. Hence

$$T_- = \left(\frac{g_*^+}{g_*^-}\right)^{1/3} T_+ = \left(\frac{11/2}{2}\right)^{1/3} T_+ = \left(\frac{11}{4}\right)^{1/3} T_+$$

and the photons get reheated by a factor $T_- = (11/4)^{1/3} T_+ \sim 1.4 T_+$, but the neutrinos are not reheated because they have previously decoupled from equilibrium (see Example 19.1). Since the CMB temperature today is 2.725 K, the neutrino background radiation is expected to have a lower temperature $T_\nu = (4/11)^{1/3} \times (2.725\,\mathrm{K}) \simeq 1.95\,\mathrm{K}$.

20.8 For a blackbody spectrum one has (in $c = 1$ units) a number density distribution of the general form

$$n(\nu)d\nu = \frac{8\pi^2 \nu^2 d\nu}{e^{h\nu/k_B T} - 1}.$$

Assume at a time t' that the photons exhibit a blackbody spectrum. The number of photons per comoving volume $na(t)^3$ is approximately conserved. Thus, as the Universe evolves from the earlier time t' to a later time t,

$$n(\nu,t)d\nu = \left(\frac{a(t')}{a(t)}\right)^3 n(\nu',t')d\nu',$$

where ν is a frequency at time t and ν' is the corresponding frequency at time t'. But because of the expansion redshift,

$$\nu' = \frac{a(t)}{a(t')}\nu \qquad d\nu' = \frac{a(t)}{a(t')}d\nu.$$

Substituting the blackbody form for the number density and the preceding expressions gives

$$\begin{aligned} n(\nu,t)\,d\nu &= \left(\frac{a(t')}{a(t)}\right)^3 \frac{8\pi^2\nu'^2}{e^{h\nu'/k_\mathrm{B}T(t')}-1}\,d\nu' \\ &= \left(\frac{a(t')}{a(t)}\right)^3 \times \frac{8\pi^2\left(\dfrac{a(t)}{a(t')}\right)^2\nu^2}{e^{h\left[a(t)/a(t')\right]\nu/k_\mathrm{B}T(t')}-1} \times \frac{a(t)}{a(t')}\,d\nu \\ &= \frac{8\pi^2\nu^2 d\nu}{e^{h\nu/k_\mathrm{B}\left[a(t')/a(t)\right]T(t')}-1}. \end{aligned}$$

But as the radiation expands $\varepsilon \propto T^4$ and $\varepsilon \propto a^{-4}$, implying that $T \propto a^{-1}$. Thus,

$$\frac{a(t')}{a(t)}T(t') = T(t),$$

which may be substituted in the above to give

$$n(\nu)d\nu = \frac{8\pi^2\nu^2 d\nu}{e^{h\nu/k_\mathrm{B}T(t)}-1}.$$

The original blackbody spectrum evolves into a new blackbody spectrum at the new (lower) temperature $T(t) = [a(t')/a(t)]\,T(t')$ as the Universe expands from time t' to time t.

20.9 The bulk of the contribution for photons is from the CMB and for neutrinos from the cosmic neutrino background. For fermions the effective degeneracy parameter g_* is given by

$$g_* = \tfrac{7}{8}n_\mathrm{f}\cdot n_\mathrm{s}\cdot n_\mathrm{a},$$

where n_f is the number of flavors, n_s is the number of spin polarizations, and n_a is the number of particle–antiparticle species. There are a particle and antiparticle for each of three known neutrino flavors. Neutrinos are spin-$\frac{1}{2}$ particles with two spin (helicity) states. However, the phenomenology of the weak interactions indicates that only the left-handed neutrino and right-handed antineutrino participate in the weak interactions (maximal parity violation). Thus, below the electroweak symmetry-breaking scale there is only one effective spin degeneracy state and for the neutrinos and antineutrinos

$$g_*^\nu = \tfrac{7}{8}\times n_\mathrm{f}\cdot n_\mathrm{s}\cdot n_\mathrm{a} = \tfrac{7}{8}\times 3\times 1\times 2 = \tfrac{21}{4}.$$

(It is assumed here that the neutrinos are Dirac particles having distinct particle and antiparticle; if they are instead Majorana particles, for which a particle is its own antiparticle, the last factor of 2 above would be reduced to 1.) For the photons, there are two polarization states and $g^\gamma = 2$. Thus, the ratio of energy density parameters for neutrinos and photons is

$$\frac{\Omega_\nu}{\Omega_\gamma} = \frac{g_*^\nu}{g^\gamma}\left(\frac{T_\nu}{T_\gamma}\right)^4,$$

where the last factor is because energy densities scale as T^4 for relativistic particles. From Problem 20.7, the cosmic neutrino background temperature T_ν and CMB temperature T_γ are related by $T_\nu = (4/11)^{1/3} T_\gamma$. Therefore,

$$\frac{\Omega_\nu}{\Omega_\gamma} = \frac{g_*^\nu}{g^\gamma}\left(\frac{T_\nu}{T_\gamma}\right)^4 = \frac{21/4}{2}\left(\frac{4}{11}\right)^{4/3} \simeq 0.68$$

for the ratio of neutrino and photon energy densities.

20.10 Estimating the average energy as kT gives

$$T = \frac{E}{k} = \frac{126 \times 10^9\,\text{eV}}{8.617 \times 10^{-5}\,\text{eV K}^{-1}} \sim 1.5 \times 10^{15}\,\text{K}.$$

From Eq. (20.17), this corresponds to a time of $\sim 1.5 \times 10^{-11}$ seconds after the birth of the Universe, assuming that $g^* \sim 100$ (see Fig. 20.2). Check: read off from Fig. 20.4 that for a temperature $\sim 10^2$ GeV the time is $t \sim 10^{-11}$ s.

20.11 The scaling of the energy densities with expansion is

$$\varepsilon_{\text{dm}} \sim a^{-3} = (1+z)^3 \qquad \varepsilon_\gamma \sim a^{-4} = (1+z)^4 \qquad \varepsilon_{\text{b}} \sim a^{-3} = (1+z)^3$$

From Table 20.2 and Table 17.1, take the density parameters today to be

$$\Omega_{\text{m}} = 0.308 \qquad \Omega_{\text{r}} = 8 \times 10^{-5} \qquad \Omega_{\text{b}} = 0.048 \qquad \Omega_{\text{dm}} = \Omega_{\text{m}} - \Omega_{\text{b}} = 0.26$$

and from Eq. (17.7) the critical energy density is

$$\varepsilon_{\text{c}} = 1.05 \times 10^{-2} h^2\,\text{MeV cm}^{-3} = 4.83 \times 10^{-3}\,\text{MeV cm}^{-3},$$

where $h = 0.678$ inferred from Table 20.2 was used. Thus for dark matter at a redshift for last scattering z_{ls},

$$\varepsilon_{\text{dm}}(z_{\text{ls}}) = \Omega_{\text{dm}}\varepsilon_{\text{c}}(1+z_{\text{ls}})^3 = 1.6 \times 10^6\,\text{MeV cm}^{-3},$$

where $z_{\text{ls}} = 1080$ was used. Likewise, for photons

$$\varepsilon_\gamma(z_{\text{ls}}) = \Omega_\gamma\varepsilon_{\text{c}}(1+z_{\text{ls}})^4 = 5.3 \times 10^5\,\text{MeV cm}^{-3},$$

and for baryons

$$\varepsilon_{\text{b}}(z_{\text{ls}}) = \Omega_{\text{b}}\varepsilon_{\text{c}}(1+z_{\text{ls}})^3 = 2.9 \times 10^5\,\text{MeV cm}^{-3}.$$

The ratio of energy densities at decoupling was then $\varepsilon_{\text{dm}} : \varepsilon_\gamma : \varepsilon_{\text{b}} \sim 5.5 : 1.8 : 1$, and the Universe was dominated by dark matter at decoupling. See also related Problem 20.5.

20.12 Putting the parameters from Table 20.2 into the cosmological calculator at Ref. [5] gives an angular size distance $d_A = 12.9\,\text{Mpc}$ for $z = 1080$. From $D = d_A\Delta\theta$ in Box 20.2, on the last scattering surface assuming this cosmology

$$D_{\text{lss}} = d_A\Delta\theta = 12.9\,\text{Mpc}\left(\frac{\Delta\theta}{\text{rad}}\right) = 3.75\left(\frac{\Delta\theta}{\text{arcmin}}\right)\,\text{kpc}.$$

The smallest angular size resolved in the Planck data is about 5 arcmin, which gives ~ 19 kpc when inserted in the above formula for the size on the LSS. This would have scaled up to

$$D(t_0) = D_{\text{lss}}(1 + z_{\text{lss}}) \sim 20.5\,\text{Mpc},$$

as observed today because of the Hubble expansion, assuming $z_{\text{lss}} = 1080$. The current baryon mass density of the Universe is $\rho_b = \Omega_b\rho_c$, which gives $\rho_b \sim 4.15 \times 10^{-31}\,\text{g cm}^{-3}$ upon using $\Omega_b = 0.048$ and $h = 0.678$ from Table 20.2 in Eq. (17.6). Using this density a sphere of diameter $D = 20.5\,\text{Mpc}$ contains a total baryonic mass of $2.8 \times 10^{13} M_\odot$. This is comparable (given our crude estimates) to the baryonic mass of a cluster of galaxies, which typically have total masses of $10^{14} - 10^{15}\,M_\odot$. For example, the total mass of the (fairly rich) Virgo Cluster is estimated to be about $\sim 10^{15}\,M_\odot$, of which probably 5-10% is baryonic.

20.13 Before decoupling the barons and photons were strongly coupled, with the photons greatly outnumbering the baryons in the fluid. Thus the speed of sound was essentially that of a photon gas, which is $v_s = c/\sqrt{3}$. The proper distance that sound could travel from the big bang to the time of last scattering t_{ls} was then

$$\begin{aligned}\ell_s(t_{\text{ls}}) &= a(t_{\text{ls}})\int_0^{t_{\text{ls}}} \frac{v_s(t)dt}{a(t)} \simeq \frac{1}{\sqrt{3}}\,a(t_{\text{ls}})\int_0^{t_{\text{ls}}} \frac{cdt}{a(t)}\\ &= \frac{1}{\sqrt{3}}\,\ell_h = \frac{1}{\sqrt{3}}(0.25\,\text{Mpc}) = 0.144\,\text{Mpc}\end{aligned}$$

where we've used that at the time of last scattering for the CMB the distance to the horizon ℓ_h was approximately 0.25 Mpc. From Box 20.2 the corresponding angular size on the CMB as viewed today is

$$\theta_s = \frac{\ell_s(t_{\text{ls}})}{d_A} = \frac{0.144\,\text{Mpc}}{12.9\,\text{Mpc}} = 0.011\,\text{rad} = 0.6^\circ,$$

where an angular size distance $d_A = 12.9\,\text{Mpc}$ was computed for the parameters in Table 20.2 at a redshift $z = 1080$ [5]. The corresponding length scale as viewed today is stretched by a factor $1 + z_{\text{ls}}$:

$$\ell_s(t_0) = (1 + z_{\text{ls}})\ell_s(t_{\text{ls}}) = (1 + 1080)(0.144\,\text{Mpc}) = 156\,\text{Mpc}.$$

The present total mass density is given by ρ_c from Eq. (17.6) since the Universe is flat. The total matter density is $\rho_m = \Omega_m\rho_c = 2.66 \times 10^{-30}\,\text{g cm}^{-3}$, where $h = 0.678$ and $\Omega_m = 0.308$ were used. Then the mass contained within a volume of radius 156 Mpc is on average

$$M_s \sim \tfrac{4}{3}\pi(156\,\text{Mpc})^3\rho_m \simeq 6.3 \times 10^{17} M_\odot.$$

This is larger than the total mass for superclusters of galaxies and sets the minimal scale that must be analyzed to observe the effect of the baryon acoustic oscillations for clustering of visible matter.

20.14 From the last equation in Box 19.3 we have $d_{\rm L} = (1+z)\ell(t_0)$, where $\ell(t_0)$ is the proper distance at the present time t_0. Thus, from the last equation in Box 20.2

$$(1+z)d_{\rm A} = \frac{d_{\rm L}}{1+z} = \ell(t_0),$$

and using Eq. (16.12) to relate scale factors to redshift,

$$d_{\rm A} = \frac{\ell(t_0)}{1+z} = \ell(t_{\rm e}),$$

where $\ell(t_{\rm e})$ is proper distance at the time of emission $t_{\rm e}$.

21 Extending Classical Big Bang Theory

21.1 From Eq. (19.2),

$$\ddot{a} = -\frac{4\pi G}{3} a(\varepsilon + 3P).$$

Therefore, $\dot{a}$ decreases with time unless $\varepsilon + 3P \leq 0$. For standard cosmologies the horizon may be approximated as the inverse Hubble parameter, $H^{-1} = a/\dot{a}$, so

$$\frac{a}{H^{-1}} \simeq \dot{a}$$

and if $\dot{a}$ decreases with time the ratio of the distance scale $a(t)$ to the horizon $\ell_{\rm H}(t)$ also decreases with time:

$$\frac{\ell_{\rm H}(t)}{a(t)} \simeq \begin{cases} t^{1/2} & \text{(radiation dominated)} \\ t^{1/3} & \text{(matter dominated)} \end{cases}.$$

Therefore, in a standard cosmology the horizon grows more rapidly than the Universe expands and objects not presently in causal contact have never been in causal contact. In the theory of inflation $\varepsilon + 3P$ becomes negative, which permits objects to expand outside the horizon (outside of causal contact) during inflation, but then to come back inside the horizon during the later standard big bang evolution after inflation has stopped.

21.2 From Eqs. (19.5) and (17.6)

$$\frac{\Delta\rho}{\rho} = \frac{\rho - \rho_{\rm c}}{\rho} = \frac{3k}{8\pi G a^2 \rho}.$$

But $\rho \propto a^{-4}$ for radiation dominated evolution, in which case $\Delta\rho/\rho \simeq a^2 \simeq t$, so the deviation from flatness grows smaller as time is extrapolated backwards. Take $t_0 \sim 4 \times 10^{17}$ s for today. Then (assuming radiation dominance to make the estimate simple), at the Planck time of $t \sim 10^{-44}$ s

$$\left(\frac{\Delta\rho}{\rho}\right)_t \Bigg/ \left(\frac{\Delta\rho}{\rho}\right)_{t_0} \simeq \frac{10^{-44}\ \text{s}}{4 \times 10^{17}\ \text{s}} \simeq 10^{-62}.$$

So unless the flatness is tuned to this precision at the Planck scale, the Universe does not evolve into one that is flat today.

21.3 First we make an unsophisticated argument. At the time of decoupling $t_{\rm d} \sim 3 \times 10^5$ yr, corresponding to a redshift $z_{\rm d} \sim 1100$. The size of the horizon at this time is $\ell_{\rm d} \sim 2ct_{\rm d}$ for radiation dominated and $\ell_{\rm d} \sim 3ct_{\rm d}$ for matter dominated cosmologies. (Let's assume radiation dominated for simplicity in estimates.) The size ℓ corresponding to this horizon in the current Universe would be stretched by the expansion according to $\ell/\ell_{\rm d} = a_0/a_{\rm d}$. However, $z_{\rm d} = a_0/a_{\rm d} - 1$, so $\ell \simeq 2ct_{\rm d}(1 + z_{\rm d})$ is the size of a causally connected region at decoupling in the present Universe. Assuming a flat Universe, one may take the distance to the last scattering surface to be close to the present horizon and given approximately by $c(t_0 - t_\mathrm{d})$. Thus, the approximate angular size of causally connected regions on the last scattering surface is

$$\theta \simeq \frac{2ct_\mathrm{d}(1+z_\mathrm{d})}{c(t_0 - t_\mathrm{d})} \simeq 0.047\,\mathrm{rad}.$$

Hence regions in the sky separated by more than a degree or so should not have been causally connected at any time in the past (in standard big bang cosmology).

A more sophisticated argument can be made by using the angular diameter distance discussed in Box 20.2 in the form $d_\mathrm{A} = \ell_\mathrm{d}/\Delta\theta$. In the standard cosmology the angular size distance corresponding to $z = 1100$ can be computed to be about 12.9 Mpc [5], and the horizon size at decoupling is about $\ell_\mathrm{d} \sim 2ct_\mathrm{d} \sim 0.184\,\mathrm{Mpc}$, assuming radiation dominance. Then

$$\Delta\theta = \frac{\ell_\mathrm{d}}{d_\mathrm{A}} = \frac{0.184\,\mathrm{Mpc}}{12.9\,\mathrm{Mpc}} = 0.014\,\mathrm{rad} \sim 0.8^\circ.$$

Again we conclude that regions on the last scattering surface separated by more than a degree or so cannot have been in past causal contact in the standard cosmology, yet widely separated regions on the sky are observed to have the same cosmic microwave background temperature to one part in 10^5. This is the horizon problem.

21.4 By analogy with the solution of Problem 21.3, the physical horizon size at the GUT transition may be estimated as $r_\mathrm{h} \simeq 2ct_\mathrm{GUT} \simeq 6\times10^{-26}$ cm, if $t_\mathrm{GUT} = 10^{-36}$ s. Assuming one monopole per horizon volume, the number density of monopoles at the GUT transition is then

$$n_\mathrm{M} \sim (r_\mathrm{h})^{-3} = 4.6\times10^{75}\,\mathrm{cm}^{-3},$$

and the energy density of monopoles will be

$$\varepsilon_\mathrm{M} \simeq 10^{15}\,\mathrm{GeV}\times n_\mathrm{M} \simeq 4.6\times10^{90}\,\mathrm{GeV\,cm}^{-3},$$

where we've assumed the average mass of a monopole to be the GUT scale. The temperature at the GUT transition is about 10^{28} K, so the energy density of radiation is

$$\varepsilon_\mathrm{r} = aT_\mathrm{GUT}^4 \sim 4.7\times10^{100}\,\mathrm{GeV\,cm}^{-3}.$$

This is 10 orders of magnitude larger than the energy density of monopoles, so at the GUT scale the Universe is highly radiation dominated. However, the radiation energy density scales as a^{-4} and the massive monopole energy density as a^{-3}. Thus, after the Universe expands to a scale factor approximately 10^{10} times that at the GUT scale, the Universe will begin to be dominated by the monopole energy density. From Eqs. (20.17) and (20.5), in the early Universe $T \sim a^{-1}$ and $t \sim a^{-2}$, so this transition will occur when the temperature has fallen to $10^{28}\,\mathrm{K}\times10^{-10} \simeq 10^{18}$ K, at a time of $10^{36}\,\mathrm{s}\times(10^{-10})^2 \simeq 10^{-16}$ seconds after the Big Bang. Thus, the early Universe would have been strongly matter dominated, contradicting the observational evidence.

21.5 This problem is suggested by a discussion in Ryden [209]. For $n = 100$ e-foldings, the expansion during inflation is

$$\frac{a(t_\mathrm{f})}{a(t_\mathrm{i})} =\simeq e^n \simeq e^{100} \simeq 10^{43}.$$

The pure vacuum-energy solution is a de Sitter space. From Eqs. (19.5) and (19.14)–(19.15) with $T_{\mu\nu} = 0$, it is necessary to solve $\dot{a}/a = H^2$, with $H \equiv (\Lambda/3)^{1/2}$, which has a solution $a(t) \sim e^{Ht}$. The corresponding vacuum energy density is (with factors of c restored)

$$\varepsilon_\Lambda = \frac{c^2}{8\pi G}\Lambda = \frac{3c^2}{8\pi G}H^2.$$

For the present model of inflation $H \sim 10^{36}\,\mathrm{s}^{-1}$, which gives upon inserting into the above equation $\varepsilon_\Lambda \sim 10^{102}\,\mathrm{GeV\,cm^{-3}}$. On the other hand, the vacuum energy density corresponding to the present accelerated expansion of the Universe is about 70% of closure density. From Eq. (17.7) with $h = 0.72$, this is $\varepsilon_\Lambda \sim 3.8 \times 10^{-6}\,\mathrm{GeV\,cm^{-3}}$. It may be concluded that the present accelerated expansion is being driven by a vacuum energy density that is some 107 orders of magnitude smaller than that for the model assumed above for inflation in the early Universe.

22 Gravitational Waves

22.1 Contract both sides of the Einstein equation (8.21) with the metric tensor,

$$g^{\mu\nu}R_{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}g_{\mu\nu}R = \frac{8\pi G}{c^4}g^{\mu\nu}T_{\mu\nu}.$$

The first term on the left reduces to R and the second term to $2R$, and the term on the right reduces to $(8\pi G/c^4)T^{\nu}{}_{\nu}$. Solving for R then gives

$$R = -\frac{8\pi G}{c^4}T^{\nu}{}_{\nu}.$$

Inserting this back in the original Einstein equation gives

$$R_{\mu\nu} = \frac{8\pi G}{c^4}(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T^{\alpha}{}_{\alpha}).$$

The vacuum Einstein equation $R_{\mu\nu} = 0$ then results from setting $T_{\mu\nu}$ and $T^{\alpha}{}_{\alpha}$ to zero.

22.2 (a) In linearized gravity $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Under a Lorentz transformation of the metric,

$$g'_{\mu\nu} = \Lambda^{\alpha}_{\mu'}\Lambda^{\beta}_{\nu'}g_{\alpha\beta} = \Lambda^{\alpha}_{\mu'}\Lambda^{\beta}_{\nu'}(\eta_{\alpha\beta} + h_{\alpha\beta}) = \eta'_{\mu\nu} + h'_{\mu\nu},$$

where $h'_{\mu\nu} \equiv \Lambda^{\alpha}_{\mu'}\Lambda^{\beta}_{\nu'}h_{\alpha\beta}$. Thus the field defined by $h_{\mu\nu}$ behaves as a rank-2 tensor in Minkowski space, for which indices may be raised or lowered by contraction with the Minkowski metric tensor $\eta_{\mu\nu}$.

(b) Since $g_{\mu\nu}$ and $g^{\mu\nu}$ must be matrix inverses of each other, this requires that $g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$ so that

$$g_{\mu\alpha}g^{\alpha\nu} = (\eta_{\mu\alpha} + h_{\mu\alpha})(\eta^{\alpha\nu} - h^{\alpha\nu}) = \delta^{\nu}_{\mu} + \mathscr{O}\left(h^2\right),$$

to first order in h.

22.3 Substituting Eq. (22.6) in Eq. (22.7) and noting that $h_{\mu\nu}$ is symmetric under exchange of indices,

$$\begin{aligned}
\delta R_{\mu\nu} &= \frac{\partial}{\partial x^{\gamma}}\left[\frac{1}{2}\eta^{\gamma\delta}\left(\frac{\partial h_{\delta\mu}}{\partial x^{\nu}} + \frac{\partial h_{\delta\nu}}{\partial x^{\mu}} - \frac{\partial h_{\mu\nu}}{\partial x^{\delta}}\right)\right] - \frac{\partial}{\partial x^{\nu}}\left[\frac{1}{2}\eta^{\gamma\delta}\left(\frac{\partial h_{\delta\mu}}{\partial x^{\gamma}} + \frac{\partial h_{\delta\gamma}}{\partial x^{\mu}} - \frac{\partial h_{\mu\gamma}}{\partial x^{\delta}}\right)\right] \\
&= \frac{1}{2}\partial_{\gamma}\left[\eta^{\gamma\delta}\left(\partial_{\nu}h_{\delta\mu} + \partial_{\mu}h_{\delta\nu} - \partial_{\delta}h_{\mu\nu}\right)\right] - \frac{1}{2}\partial_{\nu}\left[\eta^{\gamma\delta}\left(\partial_{\gamma}h_{\delta\mu} + \partial_{\mu}h_{\delta\gamma} - \partial_{\delta}h_{\mu\gamma}\right)\right] \\
&= \frac{1}{2}\left(-\partial_{\gamma}\partial_{\delta}\eta^{\gamma\delta}h_{\mu\nu} + \partial_{\mu}\partial_{\gamma}\eta^{\gamma\delta}h_{\delta\nu} - \partial_{\mu}\partial_{\nu}\eta^{\gamma\delta}h_{\delta\gamma} + \partial_{\nu}\partial_{\delta}\eta^{\gamma\delta}h_{\mu\gamma}\right).
\end{aligned}$$

Introducing the definitions

$$\Box \equiv \eta^{\gamma\delta}\partial_{\gamma}\partial_{\delta} \qquad V_{\nu} \equiv \partial_{\gamma}h^{\gamma}_{\nu} - \tfrac{1}{2}\partial_{\nu}h^{\gamma}_{\gamma} = \partial_{\gamma}\eta^{\gamma\delta}h_{\delta\nu} - \tfrac{1}{2}\partial_{\nu}\eta^{\gamma\delta}h_{\delta\gamma}$$

and noting that, for example,

$$\partial_\mu V_\nu = \partial_\mu \partial_\gamma \eta^{\gamma\delta} h_{\delta\nu} - \tfrac{1}{2}\partial_\mu \partial_\nu \eta^{\gamma\delta} h_{\delta\gamma},$$

the Ricci tensor to first order in h becomes

$$\delta R_{\mu\nu} = \tfrac{1}{2}\left(-\Box h_{\mu\nu} + \partial_\mu V_\nu + \partial_\nu V_\mu\right).$$

Then to first order in the metric perturbation h the vacuum Einstein equation $\delta R_{\mu\nu} = 0$ is

$$\Box h_{\mu\nu} - \partial_\mu V_\nu - \partial_\nu V_\mu = 0,$$

which is Eq. (22.12).

22.4 Consider linearized gravity with the metric given by Eq. (22.2). Under the coordinate transformation (22.13), $x^\mu \to x'^\mu = x^\mu + \varepsilon^\mu(x)$, where it is assumed that ε^μ and $\partial\varepsilon^\mu/\partial x^\nu$ have magnitudes comparable to or smaller than $h_{\mu\nu}$, the metric tensor transforms as

$$\begin{aligned} g'_{\mu\nu} &= \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} \\ &\simeq \left(\delta^\alpha_\mu - \frac{\partial\varepsilon^\alpha}{\partial x^\mu}\right)\left(\delta^\beta_\nu - \frac{\partial\varepsilon^\beta}{\partial x^\nu}\right) g_{\alpha\beta} \\ &\simeq g_{\mu\nu} - g_{\mu\beta}\frac{\partial\varepsilon^\beta}{\partial x^\nu} - g_{\alpha\nu}\frac{\partial\varepsilon^\alpha}{\partial x^\mu}, \end{aligned}$$

where Eq. (22.13) was used, we have assumed that to first order $\partial\varepsilon^\mu/\partial x'^\nu = \partial\varepsilon^\mu/\partial x^\nu$, and terms higher-order in $\partial\varepsilon/\partial x$ have been neglected. Hence, from Eq. (22.2)

$$\begin{aligned} h'_{\mu\nu} &= g'_{\mu\nu} - \eta_{\mu\nu} \\ &= h_{\mu\nu} - g_{\mu\beta}\frac{\partial\varepsilon^\beta}{\partial x^\nu} - g_{\alpha\nu}\frac{\partial\varepsilon^\alpha}{\partial x^\mu} \\ &= h_{\mu\nu} - \partial_\mu\varepsilon_\nu - \partial_\nu\varepsilon_\mu, \end{aligned}$$

which is Eq. (22.15).

22.5 Transform to transverse traceless (TT) gauge by choosing

$$\bar h_{0i} = 0 \quad (i = 1,2,3) \qquad \mathrm{Tr}\,\bar h \equiv \bar h^\mu_\mu = 0,$$

implying from Eq. (22.19) that the polarization tensors in TT gauge satisfy

$$\alpha_{0i} = 0 \qquad \mathrm{Tr}\,\alpha = \alpha^\mu_\mu = 0.$$

(Note that we are now in TT gauge where $\bar h = h$. Hence the bars could be dropped on h if desired, but we will keep them for consistency.) From the gauge condition (22.17) and the above constraints, $\partial\bar h_{00}/\partial x^0 = 0$. Writing the gauge condition (22.17) for $\mu = 0$ out term by term subject to the preceding consraints also gives

$$\frac{\partial\bar h_{01}}{\partial x^1} + \frac{\partial\bar h_{02}}{\partial x^2} + \frac{\partial\bar h_{03}}{\partial x^3} = 0,$$

and evaluating the partial derivatives using Eq. (22.19),

$$\bar{h}_{\mu\nu} = \alpha_{\mu\nu} e^{ik\cdot x} = \alpha_{\mu\nu} e^{-i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})},$$

leads to the requirements that

$$\frac{\partial \bar{h}_{00}}{\partial x^0} = 0 \;\rightarrow\; i\omega\alpha_{00} e^{ik\cdot x} = 0 \qquad \frac{\partial \bar{h}_{01}}{\partial x^1} + \frac{\partial \bar{h}_{02}}{\partial x^2} + \frac{\partial \bar{h}_{03}}{\partial x^3} = 0 \;\rightarrow\; ik^i \alpha_{ij} e^{-ik\cdot x}.$$

The first can be satisfied generally only if $\alpha_{00} = 0$, so in TT gauge four of the polarization tensor components are identically zero: $\alpha_{0\mu} = \alpha_{\mu 0} = 0$, which is Eq. (22.23). The second can be satisfied generally only if $k^i \alpha_{ij} = 0$, which is the transversality condition (22.24).

22.6 The transversality condition $k^j \alpha_{ij} = 0$ from Eq. (22.24) may be written out explicitly as the set of equations

$$\begin{aligned} k^1\alpha_{11} + k^2\alpha_{12} + k^3\alpha_{13} &= 0 \\ k^1\alpha_{21} + k^2\alpha_{22} + k^3\alpha_{23} &= 0 \\ k^1\alpha_{31} + k^2\alpha_{32} + k^3\alpha_{33} &= 0. \end{aligned}$$

But from (22.26), $k^1 = k^2 = 0$, so $\alpha_{13} = \alpha_{23} = \alpha_{33} = 0$, and from (22.23), $\alpha_{0\mu} = 0$. Therefore, for the symmetric matrix $\alpha_{\mu\nu}$ the only nonvanishing components are α_{11}, $\alpha_{12} = \alpha_{21}$, and α_{22}, and these are further constrained by the trace requirement from Eq. (22.22), so $\alpha_{22} = -\alpha_{11}$.

22.7 For LISA take $L = 2.5 \times 10^9$ m, which gives

$$f_* = \frac{c}{2\pi L} = 1.9 \times 10^{-2}\,\text{Hz}.$$

For LIGO the laser effective path length is increased over the physical arm length by of order 100 because of multiple reflections. Taking an effective 200 reflections as an estimate, $L \sim 200 \times 4000\,\text{m} \sim 8 \times 10^5$ m and

$$f_* = \frac{c}{2\pi L} = 59.7\,\text{Hz}.$$

Comparing with Fig. 22.8, these are indeed the approximate optimal response frequencies for LISA and advanced LIGO.

22.8 From Eqs. (22.19) and (22.16)

$$\begin{aligned} \Box \bar{h}_{\mu\nu} &= \eta^{\lambda\sigma}\partial_\lambda\partial_\sigma \left(\alpha_{\mu\nu} e^{ik_\alpha x^\alpha}\right) \\ &= \eta^{\lambda\sigma} ik_\sigma \partial_\lambda \left(\alpha_{\mu\nu} e^{ik_\alpha x^\alpha}\right) \\ &= \eta^{\lambda\sigma} ik_\lambda ik_\sigma h_{\mu\nu} \\ &= -k_\lambda k^\lambda \bar{h}_{\mu\nu} = 0. \end{aligned}$$

But $\bar{h}_{\mu\nu}$ is not generally zero so a solution of the wave equation requires that k be a null vector, $k_\lambda k^\lambda = 0$.

22.9 The test particle is initially at rest with a 4-velocity $u^\mu = (c, 0, 0, 0)$. The geodesic equation (7.23) thus reduces to

$$\frac{du^\mu}{d\tau} = -\Gamma^\mu_{00}(u^0)^2 = -c^2\Gamma^\mu_{00}.$$

From Eq. (22.6), to first order in h,

$$\Gamma^\mu_{00} = \tfrac{1}{2}\eta^{\mu\nu}(\partial_0 h_{\nu 0} + \partial_0 h_{\nu 0} - \partial_\nu h_{00}),$$

but from Eqs. (22.23) and (22.19), $h_{\nu 0} = h_{00} = 0$, so $\Gamma^\mu_{00} = 0$ and the initial 4-acceleration vanishes, $du^\mu/d\tau = 0$. (Note that we are in TT gauge where $h = \bar{h}$.) Thus, in TT gauge the particle is stationary with respect to the coordinate system as the gravitational wave passes.

23 Weak Sources of Gravitational Waves

23.1 Use Eq. (23.18) to estimate $L \sim 8 \times 10^{-23}\,\mathrm{erg\,s^{-1}}$. Obviously detection of gravitational waves produced in the laboratory is not a practical experiment.

23.2 From Eq. (23.28) for circular orbits with c and G factors evaluated (see Problem 5.8),

$$L = 2.3 \times 10^{45} \left(\frac{M}{M_\odot}\right)^{4/3} \left(\frac{\mu}{M_\odot}\right)^{2} \left(\frac{1\text{ s}}{P}\right)^{10/3} \text{erg s}^{-1}.$$

Inserting $M \sim 2.8\ M_\odot$, $\mu = 0.7 M_\odot$, and $P = 7.75\,\mathrm{h}$ gives $6.8 \times 10^{30}\,\mathrm{erg\,s^{-1}}$. However, the orbit has eccentricity $e = 0.617$, which gives a correction factor $f(e) = 11.84$ from Eq. (23.30). Thus, from Eq. (23.29) the gravitational wave luminosity of the Binary Pulsar may be estimated as $L \sim (11.84) \times (6.8 \times 10^{30}\,\mathrm{erg\,s^{-1}}) \sim 8 \times 10^{31}\,\mathrm{erg\,s^{-1}}$. This is about 2% of the photon luminosity of the Sun but it is *much* harder to detect gravitational waves than photons so it isn't feasible to observe the gravitational-wave energy emitted by the Binary Pulsar directly. It can be inferred only indirectly from the observed decay of the orbit. See Problem 23.9 for an estimate of whether the gravitational wave *strain* for the Binary Pulsar is detectable from Earth.

23.3 From Eq. (23.7) and being careless about numerical factors,

$$\dddot{I}_{ij} \simeq \frac{MR^2}{P^3} = \frac{MR^3}{RP^3} \sim \frac{Mv^3}{R},$$

where Eq. (23.9) has been used. Inserting this equation into Eq. (23.16) and eliminating M in favor of the Schwarzschild radius $r_\mathrm{S} = 2M$ gives Eqs. (23.18)–(23.19) [a numerical factor has been omitted in (23.18)]. Utilizing Eq. (23.12), this also may be expressed as $L \simeq L_0 \left(r_\mathrm{S}/R\right)^5$. The total energy emitted in one period P is

$$\Delta E \simeq LP \simeq \left(L_0 \frac{r_\mathrm{S}^2}{R^2} \frac{v^6}{c^6}\right) P.$$

Utilizing $P = 2\pi R/v$, (23.12), (23.13), $r_\mathrm{s} = 2M$, and (23.19), this gives Eq. (23.20),

$$\Delta E \simeq Mc^2 \left(\frac{r_\mathrm{S}}{R}\right)^{7/2} = \varepsilon Mc^2,$$

where a factor of 2π has been dropped. Thus ε is a measure of the efficiency of converting mass to gravitational waves.

23.4 Defining the reduced mass $\mu \equiv m_1 m_2/(m_1 + m_2)$, separation $a \equiv a_1 + a_2$, and total mass $M \equiv m_1 + m_2$, and noting that by the definition of the center of mass $m_1 a_1 = m_2 a_2$,

$$\begin{aligned}\mu a &= \frac{m_1 m_2}{M} a_1 + \frac{m_1 m_2}{M} a_2 = \frac{m_2^2}{M} a_2 + \frac{m_1 m_2}{M} a_2 \\ &= \left(\frac{m_2^2}{M} + \frac{m_1 m_2}{M}\right) a_2 = m_2 a_2.\end{aligned}$$

By an analogous proof $\mu a = m_1 a_1$ and thus $\mu a = m_1 a_1 = m_2 a_2$. Therefore,

$$m_1 a_1^2 + m_2 a_2^2 = \mu a a_1 + \mu a a_2 = \mu a (a_1 + a_2) = \mu a^2.$$

Hence, from Eq. (23.21) for the second mass moments expressed in the coordinates (23.22),

$$\begin{aligned}I^{xx} = I^{11} &= m_1 x_1^2 + m_2 x_2^2 \\ &= m_1 a_1^2 \cos^2 \omega t + m_2 a_2^2 \cos^2 \omega t \\ &= (m_1 a_1^2 + m_2 a_2^2) \cos^2 \omega t \\ &= \mu a^2 \cos^2 \omega t \\ &= \tfrac{1}{2} \mu a^2 (1 + \cos 2\omega t),\end{aligned}$$

and by analogous proofs

$$\begin{aligned}I^{xy} = I^{yx} = I^{12} &= m_1 x_1 y_1 + m_2 x_2 y_2 \\ &= m_1 a_1^2 \cos \omega t \, \sin \omega t + m_2 a_2^2 \cos \omega t \, \sin \omega t \\ &= \mu a^2 \cos \omega t \, \sin \omega t = \tfrac{1}{2} \mu a^2 \sin 2\omega t \\ I^{yy} = I^{22} &= m_1 y_1^2 + m_2 y_2^2 \\ &= m_1 a_1^2 \sin^2 \omega t + m_2 a_2^2 \sin^2 \omega t \\ &= \mu a^2 \sin^2 \omega t = \tfrac{1}{2} \mu a^2 (1 - \cos 2\omega t),\end{aligned}$$

which are the results quoted in Eq. (23.23).

23.5 From Eq. (23.23) the non-zero components are

$$\begin{aligned}I^{11} &= I^{xx} = \mu a^2 \cos^2 \omega t = \tfrac{1}{2} \mu a^2 (1 + \cos 2\omega t), \\ I^{12} &= I^{xy} = \mu a^2 \cos \omega t \sin \omega t = \tfrac{1}{2} \mu a^2 \sin 2\omega t, \\ I^{22} &= I^{yy} = \mu a^2 \sin^2 \omega t = \tfrac{1}{2} \mu a^2 (1 - \cos 2\omega t).\end{aligned}$$

The trace-reversed amplitude is given by Eq. (23.4), which requires the second time derivatives. These are easily computed from the above equations. For example,

$$\begin{aligned}\dot{I}^{xx}(t) &= \frac{d}{dt} \left(\tfrac{1}{2} \mu a^2 (1 + \cos 2\omega t)\right) = -\mu a^2 \omega \sin 2\omega t, \\ \ddot{I}^{xx}(t) &= -2\omega^2 \mu a^2 \cos 2\omega t, \\ \bar{h}^{xx} &= \frac{2}{r} \ddot{I}^{xx}(t - r) = \frac{-4\omega^2 \mu a^2}{r} \cos 2\omega (t - r).\end{aligned}$$

Computing the second time derivatives for the other components in like manner gives

$$\bar{h}^{ij} = \frac{4\omega^2 \mu a^2}{r} \begin{pmatrix} -\cos 2\omega(t-r) & -\sin 2\omega(t-r) & 0 \\ -\sin 2\omega(t-r) & \cos 2\omega(t-r) & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which is Eq. (23.24). The triple time derivatives required to compute the gravitational wave power,

$$\begin{array}{ll} \dddot{I}^{xx}(t) = 4\omega^3\mu a^2 \sin 2\omega t & \dddot{I}^{yy}(t) = -4\omega^3\mu a^2 \sin 2\omega t \\ \multicolumn{2}{c}{\dddot{I}^{xy}(t) = \dddot{I}^{yx}(t) = -4\omega^3\mu a^2 \cos 2\omega t} \end{array},$$

may be found from the second derivatives computed above.

23.6 Differentiating both sides of Eq. (23.33) leads to

$$\frac{1}{P}\frac{dP}{dt} = \frac{3}{2a}\frac{da}{dt}.$$

Assume by energy conservation that the decay of the orbit causing the decrease in period results from emission of gravitational waves. The total energy of the binary orbital motion is given in Newtonian approximation by Eq. (23.32),

$$E = -\frac{Gm_1m_2}{2a},$$

from which

$$\frac{1}{a}\frac{da}{dt} = -\frac{1}{E}\frac{dE}{dt},$$

and combining the first and third equations from above gives

$$\frac{1}{P}\frac{dP}{dt} = -\frac{3}{2}\frac{1}{E}\frac{dE}{dt}.$$

Equating the change in orbital energy with the energy carried off by gravitational waves, $dE/dt = -L$, and using Eq. (23.27) to specify L gives

$$\frac{1}{E}\frac{dE}{dt} = \frac{64}{5}\frac{G^3}{c^5}\frac{M^2\mu}{a^4}.$$

Therefore,

$$\frac{dP}{dt} = -\frac{3}{2}\frac{1}{E}\frac{dE}{dt}P = -\frac{96}{5}\frac{G^3}{c^5}\frac{M^2\mu}{a^4}P,$$

which is Eq. (23.34). The period P and the separation a are related by Kepler's 3rd law $a^3 = (GM/4\pi^2)P^2$, which can be used to eliminate a, giving an expression depending only on the period and masses

$$\frac{dP}{dt} = -\frac{192\pi}{5}\frac{G^{5/3}}{c^5}\frac{m_1m_2}{M^{1/3}}\left(\frac{2\pi}{P}\right)^{5/3},$$

which is Eq. (23.35).

23.7 The mass of the system contributing to gravitational wave radiation is assumed to be $\sim 0.5M_\odot$ and the effective radius is taken to be $R \simeq 2R_\odot \simeq 14\times 10^5$ km. From Eq. (23.13)

$$\varepsilon^{2/7} = \frac{r_s}{R} = \frac{2(G/c^2)M}{R} = 2.95\left(\frac{M}{M_\odot}\right)\left(\frac{\text{km}}{R}\right),$$

which gives for 44 Boo an efficiency $\varepsilon^{2/7} \sim 10^{-6}$. Therefore, from Eq. (23.15) and the period of 6.4 hours the amplitude and frequency of the expected gravitational wave metric perturbation is

$$\bar{h} \simeq 7 \times 10^{-21} \qquad f \simeq 8.7 \times 10^{-5}\,\text{Hz},$$

since [see Eq. (23.23)], the gravitational wave frequency is twice the binary frequency of revolution. Consulting Fig. 22.8, the expected gravitational wave frequency is outside the favorable response range of LIGO or Virgo, but within the frequency window for LISA. Thus, it is possible that space-based arrays may be able to detect gravitational radiation from some galactic binaries.

23.8 Restore c and G factors in Eq. (23.8) to give

$$\bar{h}^{ij} = \frac{4G}{c^4}\frac{MR^2}{rP^2}$$

and use

$$R^2 = \left(\frac{GM}{4\pi^2}\right)^{2/3} P^{4/3}$$

from Kepler's 3rd law to eliminate R, giving after evaluation of constants

$$\bar{h}^{ij} = 1.47 \times 10^{-4}\left(\frac{M}{M_\odot}\right)^{5/3}\left(\frac{\text{s}}{P}\right)^{2/3}\left(\frac{\text{km}}{r}\right),$$

for masses given in solar masses, periods in seconds, and distances in kilometers.

23.9 From the information in Section 10.4.1, assume that the average separation of the neutron stars is $R \sim 2R_\odot$, the effective mass entering into generation of gravitational waves is $M \sim 1\,M_\odot$, and the distance is $r = 6.4$ kpc. Then from Eq. (23.13)

$$\varepsilon^{2/7} = \frac{r_\text{s}}{R} = \frac{2.95(M/M_\odot)}{2R_\odot} = 2.1 \times 10^{-6},$$

and from Eq. (23.15)

$$\bar{h} = 9.6 \times 10^{-17}\varepsilon^{2/7}\left(\frac{M}{M_\odot}\right)\left(\frac{\text{kpc}}{r}\right) \simeq 3.2 \times 10^{-23}.$$

The period of the binary is 7.75 hours, implying an orbital frequency $3.6 \times 10^{-5}\,\text{s}^{-1}$. The gravitational wave frequency is twice that, $f = 7.2 \times 10^{-5}\,\text{s}^{-1}$. From Fig. 22.8 this is roughly in the LISA frequency window but the strain is several orders of magnitude too small to be measurable by LISA.

24 Strong Sources of Gravitational Waves

24.1 The event corresponded to the merger of two black holes having a total mass of $\sim 70\ M_\odot$. The observed frequency near peak was ~ 150 Hz, implying a period for revolution of the binary of half that or 75 Hz. Assuming Kepler trajectories for a rough estimate, the separation between the objects was

$$
\begin{aligned}
r_1 + r_2 &= \left(\frac{G}{4\pi^2}(m_1+m_2)P^2 \right)^{1/3} \\
&= 1.5 \times 10^3 \left[\left(\frac{m_1+m_2}{M_\odot} \right) \left(\frac{P}{\mathrm{s}} \right)^2 \right]^{1/3} \ \mathrm{km}.
\end{aligned}
$$

For a period and total mass

$$
P = (75\,\mathrm{Hz})^{-1} = 1.33 \times 10^{-2}\,\mathrm{s} \qquad m_1 + m_2 = 70\,M_\odot,
$$

this gives 347 km for the separation. The Schwarzschild radius for a black hole of mass M is $r_\mathrm{s} = 2.95\,M$ km, which gives 207 km for the sum of Schwarzschild radii assuming $m_1 + m_2 = 70 M_\odot$.

24.2 The chirp waveform in the bottom panel of Fig. 24.4 indicates a binary merger. The theoretical chirp mass

$$
\mathscr{M} = \frac{\mu^{3/5}}{M^{2/5}}
$$

from Eq. (24.2) is plotted as a function of m_1 for different values of m_2 in Fig. 24.1 [this document]. The chirp mass $\mathscr{M} \sim 28 \pm 2\ M_\odot$ (Table 24.1; see also Problem 24.5) determined observationally from the frequency and its time derivative of the gravitational wave is indicated by the dashed horizontal line and gray uncertainty box. By summing m_1 and m_2 at the intersections of the curves with the $\mathscr{M} = 28$ line, one sees that the minimum total mass of the binary consistent with the chirp mass is around $65\,M_\odot - 70\,M_\odot$. From Problem 24.1 the separation of centers at the time of maximum frequency was about 350 km. Only black holes or neutron stars are compact enough to be consistent with that. Assuming neutron stars to have an upper mass limit of $\sim 2\ M_\odot$, two neutron stars would have far too little chirp mass to account for the data. From the $m_2 = 2\ M_\odot$ curve, for a neutron star and black hole to give the observed value of $\mathscr{M}$ the mass of the black hole would have to be huge, giving a very large total mass for the system that would lead to a much lower gravitational wave frequency than observed. Thus a black hole and neutron star binary is ruled out, leaving merging black holes with a total summed mass near 70 $M_\odot$ as the only plausible explanation.

24.3 Although the binary pulsar results of Section 23.2.3 are rather convincing, they are

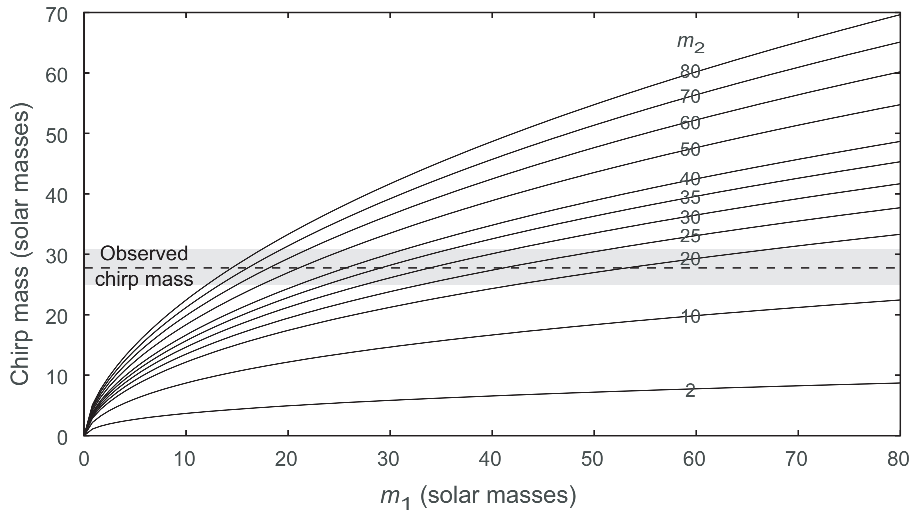


Fig. 24.1 Chirp mass for Problem 24.2 as a function of m_1 for different values of m_2.

not the same as a direct observation. Perhaps more important is that the GW150914 event is the first observation that tests general relativity in the strong-field limit; all previous tests have been for conditions in weak gravitational fields and thus have not been a full test. Finally, the astrophysical implications are potentially enormous, since GW150914 represents the first observation of radiation emitted from near the event horizon of a black hole, and the event suggests that there are extremely energetic things happening in the Universe that are not readily visible in electromagnetic waves.

24.4 The gravitational wave travels at $v = c$ so the frequency f and wavelength λ are related by $\lambda = c/f$. At peak strain the frequency of GW150914 was about 150 Hz, implying that $\lambda \sim 2000$ km. Estimate the characteristic source size as $4r_s$ for the black holes near merger. Assuming $30M_\odot$ for each black hole, this gives a source size of about 350 km. Thus $\lambda/d \sim 6$ and the quadrupole formula should be at least approximately correct. From Table 24.1 the redshift for the gravitational wave source was estimated as $z \sim 0.09$, so the corresponding frequency f_0 in the rest frame of the source at peak strain was

$$f_0 = (1+z)f = (1+0.09)150\,\mathrm{Hz} = 163.5\,\mathrm{Hz},$$

where f is the measured frequency. The rotational frequency ω of the binary is half the gravitational wave frequency, giving $\omega \sim 81.75\,\mathrm{s}^{-1}$. From Eq. (23.31) the luminosity is

$$L = 2.3 \times 10^{45} \left(\frac{M}{M_\odot}\right)^{4/3} \left(\frac{\mu}{M_\odot}\right)^{2} \left(\frac{1\,\mathrm{s}}{P}\right)^{10/3} \mathrm{erg\ s^{-1}},$$

where M is the total mass, μ is the reduced mass, and P is the period. Inserting $P = \omega^{-1} = 1.22 \times 10^{-2}$ s and assuming $m_1 = 36M_\odot$ and $m_2 = 29M_\odot$ yields a peak luminosity of $L = 3.7 \times 10^{56}\,\mathrm{erg\,s^{-1}}$, which is consistent with the value given in Table 24.1.

24.5 From Eq. (24.2) with constants evaluated

$$\mathscr{M} = \frac{c^3}{G}\left(\frac{5}{96}\right)^{3/5}\pi^{-8/5}f^{-11/5}\dot{f}^{3/5} = 5.53\times 10^3\left(\frac{f}{\mathrm{s}^{-1}}\right)^{-11/5}\left(\frac{\dot{f}}{\mathrm{s}^{-2}}\right)^{3/5} M_\odot.$$

Estimating from the bottom panel of Fig. 24.4 that at peak strain $f \sim 150\,\mathrm{s}^{-1}$ and $\dot{f} \sim \Delta f/\Delta t \sim 1.6\times 10^4\,\mathrm{s}^{-2}$ and inserting in the above equation gives $\mathscr{M} \sim 30 M_\odot$, which is consistent with the value quoted in Table 24.1.

24.6 Assuming the validity of Newtonian mechanics and Newtonian gravity, the total (kinetic plus potential) energy is

$$E = \tfrac{1}{2}m_1 v_2^2 + \tfrac{1}{2}m_2 v_2^2 - \frac{Gm_1m_2}{a} = \tfrac{1}{2}m_1 r_1^2\omega^2 + \tfrac{1}{2}m_2 r_2^2\omega^2 - \frac{Gm_1m_2}{a},$$

where $v = \omega r$ has been used with $\omega \equiv 2\pi/P$. But

$$r_1 = \frac{m_2}{M}a \qquad r_2 = \frac{m_1}{M}a \qquad M \equiv m_1 + m_2,$$

allowing the total energy to be written

$$\begin{aligned} E &= \tfrac{1}{2}m_1\left(\frac{m_2}{M}\right)^2 a^2\omega^2 + \tfrac{1}{2}m_2\left(\frac{m_1}{M}\right)^2 a^2\omega^2 - \frac{Gm_1m_2}{a} \\ &= \tfrac{1}{2}\mu a^2\omega^2 - \frac{Gm_1m_2}{a}, \end{aligned}$$

where $\mu \equiv m_1m_2/M$ is the reduced mass. Eliminating the frequency ω using Kepler's 3rd law in the form $a^3 = GM/\omega^2$ then gives for the total orbital energy $E = -Gm_1m_2/2a$.

25 Tests of General Relativity

25.1 Any viable relativistic gravitational theory should agree with the results of Newtonian gravity in the weak-field limit, as described in Section 8.1. There it was shown that the lowest-order relativistic correction to flat space modifies only the g_{00} component of the metric to $g_{00} = -(1 - 2GM/rc^2)$ [see Eq. (8.12)], with the other components unaltered to lowest order. Comparing with Eq. (25.1), agreement of general relativity with Newtonian gravity in the weak-field limit requires that to lowest order

$$A(r) = 1 - \frac{2GM}{rc^2} + \ldots \qquad B(r) = 1 + \ldots .$$

which is Eq. (25.3) to this order.

25.2 From Eq. (6.5), the strength of the gravitational field is measured by

$$\varepsilon = \frac{GM}{Rc^2} = 7.416 \times 10^{-31} \left(\frac{M}{\mathrm{kg}}\right)\left(\frac{\mathrm{km}}{R}\right) = 1.475 \left(\frac{M}{M_\odot}\right)\left(\frac{\mathrm{km}}{R}\right),$$

where M is the mass producing the gravitational field and R is the characteristic distance over which it acts.

1. For terrestrial experiments $R = R_\oplus$ and $M = M_\oplus$, which gives $\varepsilon \sim 7 \times 10^{-10}$.
2. For Mercury, take its average distance from the Sun of 5.7×10^7 km for R and $M = 1M_\odot$, which gives $\varepsilon \sim 2.6 \times 10^{-8}$.
3. For light deflection at the surface of the Sun take $R \sim R_\odot$ and $M = 1M_\odot$, leading to $\varepsilon \sim 2.1 \times 10^{-6}$.
4. For the Binary Pulsar, take $M \sim 1.4M_\odot$ and a smallest separation $R \sim 1.1R_\odot$, to give $\varepsilon \sim 2.7 \times 10^{-6}$.
5. For S0-2 take a closest approach to the black hole of $R = 17$ lighthours $= 1.8 \times 10^{10}$ km and $M = 4.3 \times 10^6 M_\odot$, which yields $\varepsilon = 3.5 \times 10^{-4}$.
6. For GW150914 estimate that at closest approach in the black hole merger $R \sim 100$ km and $M \sim 70M_\odot$, which gives $\varepsilon \sim 0.5$.

See also the related Problem 10.8.

26 Beyond Standard Models

26.1 This solution follows an example in Zwiebach [257]. For a square well in two variables $(x,y) \sim (x, y+2\pi R)$, the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial x^2}\right)=E\psi.$$

Substituting $\psi(x,y)=\psi(x)\varphi(y)$, the Schrödinger equation becomes

$$-\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2}-\frac{\hbar^2}{2m}\frac{1}{\varphi(y)}\frac{d^2\varphi(y)}{dy^2}=E.$$

The solutions of this equation are

$$\psi_k(x)=c_k\sin\left(\frac{k\pi x}{a}\right) \qquad \varphi_\ell(y)=a_\ell\sin\left(\frac{\ell y}{R}\right)+b_\ell\cos\left(\frac{\ell y}{R}\right)$$
$$E_{k,\ell}=\frac{\hbar^2}{2m}\left[\left(\frac{k\pi}{a}\right)^2+\left(\frac{\ell}{R}\right)^2\right] \qquad (k=1,2,3,\ldots\infty;\ \ \ell=0,1,2,\ldots\infty),$$

where $\ell=0$ is allowed because of the boundary conditions in the y direction. As shown in Problem 26.2, if R is small the new states introduced by the compactified y dimension will be very high in energy.

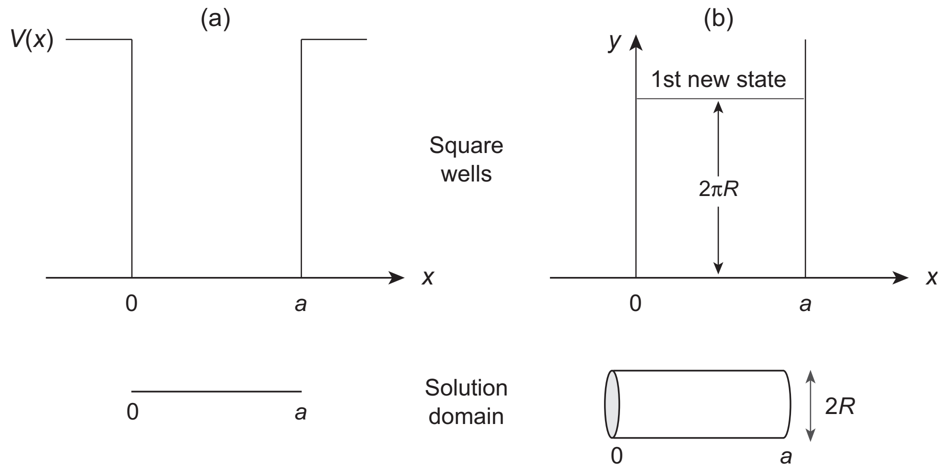
26.2 From the spectrum obtained in the solution of Problem 26.1, if $\ell=0$

$$E_{k,\ell}=\frac{\hbar^2}{2m}\left[\left(\frac{k\pi}{a}\right)^2+\left(\frac{\ell}{R}\right)^2\right] \quad \longrightarrow \quad \frac{\hbar^2}{2m}\left(\frac{k\pi}{a}\right)^2,$$

which is the spectrum of the 1-dimensional square well. Thus the states with $\ell=0$ are the old states of the 1-dimensional square well. The lowest-energy new state corresponds to $k=1$ and $\ell=1$, giving

$$E_{\text{min}}^{\text{new}}=E_{1,1}=\frac{\hbar^2}{2m}\left[\left(\frac{\pi}{a}\right)^2+\left(\frac{1}{R}\right)^2\right]\simeq\frac{\hbar^2}{2m}\left(\frac{1}{R}\right)^2,$$

where in the last step $R \ll a$ was assumed. This state has the energy of a state with $k=a/\pi R \gg 1$ in the original 1-dimensional spectrum, so it is very high in energy. The following figure illustrates schematically how the compactified dimension changes the spectrum.



In (a) the particle is restricted to a line between 0 and a. In (b) with the added compactified dimension the particle is restricted to the surface of a cylinder of length a and circumference $2\pi R$. If $R \ll a$, particle vibrations in the y direction are strongly restricted, which by uncertainty principle arguments means that they represent states having very high energy compared with those associated with motion in the x direction. Thus, if the radius of compactification is small compared with the characteristic length scales of a physical problem, the compactified dimensions have negligible influence on the low-energy spectrum.