Question 1

Consider a two-dimensional Cartesian coordinate system (x, y) with the infinitesimal line element $ds^2 = dx^2 + dy^2$. We then introduce new coordinates u and v, defined by $u = (x + y)/2$ and $v = (x - y)/2$. Find the components of the metric tensor in the new coordinates (u, v) using the transformation rule for a $(0, 2)$ tensor, which states that $g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu}$. You should use this method exclusively, without relying on any alternative approaches.

In this question, we are examining a coordinate change in a two-dimensional space. We start from a standard Cartesian coordinate system (x, y) and move to a new coordinate system (u, v) defined by a linear transformation. The goal is to find the components of the metric tensor in the new coordinate system (u, v) using the transformation rule for a (0,2) tensor, i.e., a rank-2 covariant tensor. The transformation rule is given by $g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu}$, where $g_{\mu\nu}$ are the components of the metric tensor in the original coordinates (x, y), and $g_{\mu'\nu'}$ are the components of the metric tensor in the new coordinates (u, v). This formula tells us how the components of the metric tensor change when we change the coordinate system. It is important to note that we must exclusively use this formula for the solution, without using shortcuts or alternative methods.

Solution

(i) Metric tensor in the original Cartesian coordinates (x, y). In the (x, y) coordinates, the line element is

$$ds^2 = dx^2 + dy^2.$$

This formula represents the infinitesimal line element in two dimensions using the Cartesian coordinates x and y. In a flat (Euclidean) space, the infinitesimal distance squared, ds^2, is given by the sum of the squares of the infinitesimal differences of the coordinates. Here, dx^2 and dy^2 represent the squares of the infinitesimal variations along the x and y axes, respectively. Essentially, this is the Pythagorean theorem applied to infinitesimal distances. The formula implies that the space is Euclidean and that the coordinates x and y are orthogonal, i.e., there is no cross term like dxdy, which means there is no skew or tilt between the axes.

Hence, the metric tensor components $g_{\mu\nu}$ in these coordinates are:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

This is the metric tensor in Cartesian coordinates for a two-dimensional Euclidean space. The metric is diagonal with components $g_{xx} = 1$ and $g_{yy} = 1$, and off-diagonal components $g_{xy} = g_{yx} = 0$. This metric tells us that the x and y coordinates are orthogonal (because the off-diagonal terms are zero) and that the scale along each axis is unity (because the diagonal terms are one). Geometrically, this means we are using a standard, undistorted, orthogonal coordinate system.

where $g_{xx} = 1$, $g_{yy} = 1$, $g_{xy} = g_{yx} = 0$. These are the specific components of the metric tensor $g_{\mu\nu}$ in the Cartesian coordinates (x, y). $g_{xx} = 1$ indicates that the "length" or "scale" along the x-axis is unit, and similarly, $g_{yy} = 1$ indicates that the "length" or "scale" along the y-axis is unit. The terms $g_{xy} = g_{yx} = 0$ indicate that there is no correlation or "mixing" between the x and y coordinates, which is consistent with the fact that the Cartesian axes are orthogonal. In simpler terms, this tells us that we are using a standard, undistorted, orthogonal coordinate system.

(ii) Coordinates transformation to (u, v). We define:

$$u = \frac{x + y}{2}, \quad v = \frac{x - y}{2}.$$

Here we are defining the new coordinates u and v as linear combinations of the original coordinates x and y. The coordinate u is the average of x and y, while v is half the difference between x and y. This transformation corresponds to a 45-degree counterclockwise rotation, followed by a rescaling.

To apply the transformation rule for the metric, we need the inverse relations, which are:

$$x = u + v, \quad y = u - v.$$

These are the inverse transformations expressing the original coordinates x and y in terms of the new coordinates u and v. They were obtained by solving the previous system of equations for x and y. For example, adding the two equations gives $u + v = x$, and subtracting them gives $u - v = y$. These relations allow us to express the partial derivatives of x and y with respect to u and v, which are needed to apply the metric tensor transformation rule.

(iii) Calculating partial derivatives. We compute the partial derivatives of x and y with respect to the new coordinates (u, v):

$$\frac{\partial x}{\partial u} = 1, \quad \frac{\partial x}{\partial v} = 1, \quad \frac{\partial y}{\partial u} = 1, \quad \frac{\partial y}{\partial v} = -1.$$

These equations compute the partial derivatives of x and y with respect to the new coordinates u and v. For instance, $\frac{\partial x}{\partial u} = 1$ means that x increases by 1 unit when u increases by 1 unit, keeping v constant. Similarly, $\frac{\partial y}{\partial v} = -1$ means that y decreases by 1 unit when v increases by 1 unit, keeping u constant. These partial derivatives are constant because the transformation between (x, y) and (u, v) is linear.

(iv) Applying the $(0,2)$ tensor transformation rule. Recall the rule:

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

This is the transformation rule for a rank-2 covariant tensor, such as the metric tensor. It tells us how the components of the metric tensor transform when we move from one coordinate system to another. In this formula, $g_{\mu'\nu'}$ are the components of the metric tensor in the new coordinate system, $g_{\mu\nu}$ are the components in the old system, and $\frac{\partial x^\mu}{\partial x^{\mu'}}$ are the partial derivatives of the old coordinates with respect to the new ones. In practice, we multiply the components of the old metric tensor by the appropriate partial derivatives to get the components in the new system.

Let μ, ν denote the old coordinates (x or y) and μ', ν' the new ones (u or v). Since the old metric components are $g_{xx} = 1$, $g_{yy} = 1$, $g_{xy} = 0$, $g_{yx} = 0$, each new metric component is computed as follows:

- g_{uu}:

$$g_{uu} = \left(\frac{\partial x}{\partial u}\right)^2 g_{xx} + \left(\frac{\partial y}{\partial u}\right)^2 g_{yy} = 1^3 + 1^3 = 2.$$

  This computes the g_{uu} component of the metric tensor in the new coordinates. Using the transformation rule, we sum the products of the partial derivatives multiplied by the corresponding components of the original metric tensor. Since $g_{xy} = g_{yx} = 0$, the mixed terms vanish, leaving only the sum of the squares of the partial derivatives of x and y with respect to u, multiplied by g_{xx} and g_{yy} respectively. The result is $g_{uu} = 1^2 \cdot 1 + 1^2 \cdot 1 = 2$. This tells us that the "length" or "scale" along the u-axis is 2.

- g_{uv}:

$$g_{uv} = \left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right) g_{xx} + \left(\frac{\partial y}{\partial u}\right)\left(\frac{\partial y}{\partial v}\right) g_{yy} = (1)(1)(1) + (1)(-1)(1) = 0.$$

  This computes the g_{uv} component of the metric tensor. Again, we apply the transformation rule. The result is $g_{uv} = (1)(1) \cdot 1 + (1)(-1) \cdot 1 = 0$. This means that the u and v coordinates are orthogonal.

- g_{vu}:

$$g_{vu} = \left(\frac{\partial x}{\partial v}\right)\left(\frac{\partial x}{\partial u}\right) g_{xx} + \left(\frac{\partial y}{\partial v}\right)\left(\frac{\partial y}{\partial u}\right) g_{yy} = (1)(1)(1) + (-1)(1)(1) = 0.$$

  This computes the g_{vu} component. It is equal to g_{uv} due to the symmetry of the metric tensor, so it is also 0.

- g_{vv}:

$$g_{vv} = \left(\frac{\partial x}{\partial v}\right)^2 g_{xx} + \left(\frac{\partial y}{\partial v}\right)^2 g_{yy} = 1^3 + (-1)^2(1) = 2.$$

  This computes the g_{vv} component. Similar to g_{uu}, we find $g_{vv} = (1)^2 \cdot 1 + (-1)^2 \cdot 1 = 2$. This tells us that the "length" or "scale" along the v-axis is 2.

(v) Final components of the metric in (u, v).
Collecting these results, the new metric tensor is:

$$g_{\mu'\nu'} = \begin{pmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

This is the metric tensor in the new coordinates (u, v). It is still diagonal, which means that u and v are orthogonal, but now it has diagonal components equal to 2. This indicates that the space in the (u, v) coordinates is still flat (Euclidean), but distances are rescaled by a factor of $\sqrt{2}$ compared to the original Cartesian coordinates.

Physical interpretation: We see that the resulting metric is still diagonal (and represents the same flat space), but it is now scaled by a factor of 2 in both directions u and v. Hence, the line element in the new coordinates can be written as

$$ds^2 = 2\,du^2 + 2\,dv^2.$$

This is the physical interpretation of the metric tensor we have calculated. The fact that the metric tensor is still diagonal means that the coordinates u and v are orthogonal to each other. The factor of 2 in the diagonal components g_{uu} and g_{vv} indicates that distances measured in the u and v coordinates are rescaled by a factor of $\sqrt{2}$ relative to distances in the original Cartesian coordinates. This is consistent with the geometric interpretation of the coordinate transformation as a 45-degree rotation followed by a rescaling.

Conclusion: Using exclusively the tensor transformation rule, we have correctly derived the metric components in the (u, v) coordinates:

$$g_{uu} = 2, \quad g_{uv} = 0, \quad g_{vv} = 2.$$

In conclusion, we have computed the components of the metric tensor in the new coordinates (u, v) by applying the tensor transformation rule. We found that $g_{uu} = 2$, $g_{uv} = g_{vu} = 0$, and $g_{vv} = 2$. This result confirms that the u and v coordinates are orthogonal and that distances in this new coordinate system are scaled by a factor of $\sqrt{2}$ with respect to the original Cartesian coordinate system (x, y).

Question 2

Consider a two-dimensional plane in polar coordinates, where the infinitesimal line element is given by

$$ds^2 = dr^2 + r^2\, d\phi^2.$$

(i) How many independent Christoffel symbols are there in total in two dimensions?
(ii) How many independent and non-vanishing Christoffel symbols are there in this particular case?
(iii) Compute the explicit form of one non-vanishing Christoffel symbol of your choice.

This exercise focuses on the Christoffel symbols in a two-dimensional plane described by polar coordinates (r, ϕ). Even though the plane itself is geometrically flat, the choice of polar coordinates introduces nontrivial metric components: $g_{rr} = 1$ and $g_{\phi\phi} = r^2$. We will explore how these affect the connection coefficients.

Solution

(i) Total number of Christoffel symbols in 2D. In two dimensions, each index μ, ν, λ of $\Gamma^{\mu}_{\nu\lambda}$ can take 2 values (which we may denote by r and ϕ). If we ignore any symmetries, there are $2 \times 2 \times 2 = 8$ possible symbols.

However, for the Levi-Civita connection, we use the crucial symmetry

$$\Gamma^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\lambda\nu},$$

which states that interchanging the lower two indices does not produce a new or different symbol. Thus, the Γ are symmetric under $\nu \leftrightarrow \lambda$. Since we only consider distinct pairs (ν, λ) up to this symmetry, we effectively reduce the total count from 8 to 6. Hence, there are $\boxed{6}$ *independent Christoffel symbols in 2D.*

(ii) Non-vanishing Christoffel symbols in polar coordinates (r, ϕ). Given the 2D plane in polar coordinates:

$$ds^2 = dr^2 + r^2\, d\phi^2,$$

we read off the metric and its inverse:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad g^{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}.$$

Note that the metric is diagonal. In the formula for the Christoffel symbols,

$$\Gamma^{\mu}_{\nu\lambda} = \frac{1}{2}\, g^{\mu\rho}\Big(\partial_\nu g_{\rho\lambda} + \partial_\lambda g_{\rho\nu} - \partial_\rho g_{\nu\lambda}\Big),$$

a diagonal metric means that many terms vanish unless $\rho = \mu$. Concretely, if $\mu \neq \rho$, then $g^{\mu\rho}$ will be zero for a strictly diagonal metric. This helps eliminate many potential non-zero symbols.

In polar coordinates, the only non-zero partial derivative of the metric is:

$$\partial_r g_{\phi\phi} = 2\, r,$$

while $\partial_\phi g_{rr}$, $\partial_r g_{rr}$, and $\partial_\phi g_{\phi\phi}$ vanish. Consequently, any Christoffel symbol that does not involve $\partial_r g_{\phi\phi}$ will be zero. Checking each possible combination systematically, one finds that the only non-zero symbols are:

$$\Gamma^{r}_{\phi\phi} = -r, \quad \Gamma^{\phi}_{r\phi} = \Gamma^{\phi}_{\phi r} = \frac{1}{r}.$$

All others vanish.

Even though the plane is flat, the curvilinear coordinates (r, ϕ) introduce these non-zero Christoffel symbols. If we switched to Cartesian coordinates, we would get zero for all Christoffel symbols because the metric becomes constant and diagonal with no dependence on x or y.

(iii) Explicit calculation of $\Gamma^{r}_{\phi\phi}$. For illustration, let us compute $\Gamma^{r}_{\phi\phi}$. Substitute $\mu = r$, $\nu = \phi$, and $\lambda = \phi$ into the Levi-Civita connection formula:

$$\Gamma^{r}_{\phi\phi} = \frac{1}{2}\, g^{r\rho}\Big(\partial_\phi g_{\rho\phi} + \partial_\phi g_{\rho\phi} - \partial_\rho g_{\phi\phi}\Big).$$

Since $g^{rr} = 1$ and $g^{r\phi} = 0$, the only relevant piece comes from $\rho = r$. We use:

$$\partial_r g_{\phi\phi} = 2\, r, \quad \partial_\phi g_{\phi r} = 0,$$

thus

$$\Big(\partial_\phi g_{r\phi} + \partial_\phi g_{r\phi} - \partial_r g_{\phi\phi}\Big) = (0 + 0 - 2r) = -2r.$$

Multiplying by $g^{rr} = 1$ and then by $1/2$:

$$\Gamma^{r}_{\phi\phi} = \frac{1}{2}(-2r) = -r.$$

Hence, $\boxed{\Gamma^{r}_{\phi\phi} = -r}$.

This coefficient tells us how the basis vector in the r-direction changes when we vary ϕ. In a curvilinear coordinate system, such as polar coordinates, this accounts for the "circular arcs" nature of ϕ.

Question 3

Consider a two-dimensional spacetime where the infinitesimal line element is given by
$$ds^2 = -(1+x)^2 \, dt^2 + dx^2.$$

(i) How many independent Christoffel symbols are there in principle in two dimensions?
(ii) How many independent and non-vanishing Christoffel symbols are there for this example?
(iii) Compute the explicit form of Γ^t_{tx} for this example.

Here, we turn to a "(1+1)-dimensional" spacetime metric. The dependence of $g_{tt} = -(1+x)^2$ on the spatial coordinate x leads to interesting non-zero connection coefficients. This scenario illustrates how time can "flow differently" at different positions x.

Solution

(i) Number of independent Christoffel symbols in two dimensions. Just as in Question 2, we recognize that in 2D each index (μ, ν, λ) runs over t, x. Naively, there would be $2 \times 2 \times 2 = 8$ Christoffel symbols, but the symmetry
$$\Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{\lambda\nu}$$
cuts this count down to 6. Thus there are $\boxed{6}$ independent symbols.

The same logic applies: in 2D, for each upper index μ, there are 3 unique pairs (ν, λ) up to symmetry. Since μ can be t or x, we have $3 + 3 = 6$.

(ii) Number of independent and non-vanishing Christoffel symbols for this example. From
$$ds^2 = -(1+x)^2 \, dt^2 + dx^2,$$
we read off
$$g_{\mu\nu} = \begin{pmatrix} -(1+x)^2 & 0 \\ 0 & 1 \end{pmatrix}, \quad g^{\mu\nu} = \begin{pmatrix} -\frac{1}{(1+x)^2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Again, the metric is diagonal. Therefore, in the Christoffel symbol formula, many off-diagonal terms disappear. Moreover, note that $g_{tt} = -(1+x)^2$ depends on x, while $g_{xx} = 1$ is constant.

Recall the connection formula:
$$\Gamma^\mu_{\nu\lambda} = \frac{1}{2} g^{\mu\rho} \Big(\partial_\nu g_{\rho\lambda} + \partial_\lambda g_{\rho\nu} - \partial_\rho g_{\nu\lambda} \Big).$$

Since
$$\partial_x g_{tt} = -2(1+x), \quad \partial_x g_{xx} = 0, \quad \partial_t g_{\mu\nu} = 0,$$

the only non-zero derivative of the metric is $\partial_x g_{tt}$. Hence, any Christoffel symbol that does not involve $\partial_x g_{tt}$ in the sum will vanish. By matching indices in the connection formula, we find only three potential symbols could be non-zero: Γ^x_{tt}, $\Gamma^t_{tx} = \Gamma^t_{xt}$, and Γ^x_{xx}. A direct check confirms $\Gamma^x_{xx} = 0$. Thus the only non-zero Christoffel symbols are:
$$\Gamma^x_{tt} \quad \text{and} \quad \Gamma^t_{tx} = \Gamma^t_{xt}.$$

This precisely reflects the $(1+x)$-dependence in g_{tt}. If g_{tt} were constant, these symbols would vanish, signifying a trivial geometry.

(iii) Explicit form of Γ^t_{tx}. We compute:
$$\Gamma^t_{tx} = \frac{1}{2} g^{t\rho} \Big(\partial_t g_{x\rho} + \partial_x g_{t\rho} - \partial_\rho g_{tx} \Big).$$

Since $g^{tx} = 0$ (diagonal inverse metric) and $\partial_t g_{\alpha\beta} = 0$ (metric independent of t), the dominant term arises when $\rho = t$. That term involves $\partial_x g_{tt}$. Because the metric is diagonal, effectively we need $\mu = \rho$ in many sums to get a non-zero result.

Thus:
$$\Gamma^t_{tx} = \frac{1}{2} g^{tt} \, \partial_x g_{tt}.$$

We substitute:
$$g^{tt} = -\frac{1}{(1+x)^2}, \quad \partial_x g_{tt} = -2(1+x).$$

Hence:
$$\Gamma^t_{tx} = \frac{1}{2} \left(-\frac{1}{(1+x)^2} \right) \left(-2(1+x) \right) = \frac{1}{1+x}.$$

Interpretation: $\Gamma^t_{tx} = \frac{1}{1+x}$ shows how the time basis vector changes in the x-direction. Because $(1+x)$ appears in the time component of the metric, the rate of time flow depends on x. Moving along x effectively shifts how clocks "tick" in this spacetime.

Final Remarks for Question 3. *One can check $\Gamma^x_{tt} = 1+x$ and verify $\Gamma^x_{xx} = 0$. Together with $\Gamma^t_{tx} = \Gamma^t_{xt} = \frac{1}{1+x}$, these are the only non-zero symbols. They reflect the coordinate dependence of the metric and thus a non-trivial connection. In a 2D "spacetime" context, it tells us that an observer's notion of time changes with position x.*

4

Question 4

Consider a two-dimensional space whose infinitesimal line element is given by
$$ds^2 = (1 + x^2)\, dx^2 \; + \; (1 + y^2)\, dy^2.$$

We want to compute the Christoffel symbols Γ^x_{xx} and Γ^x_{yy}.
In this problem, we have a 2D space with metric:
$$ds^2 = (1 + x^2)\, dx^2 + (1 + y^2)\, dy^2.$$

The coordinates (x, y) are orthogonal, since there is no mixed term $dx\,dy$. Our goal is to compute two specific Christoffel symbols, Γ^x_{xx} and Γ^x_{yy}.

Solution (Question 4)

From the line element,
$$ds^2 \; = \; (1 + x^2)\, dx^2 \; + \; (1 + y^2)\, dy^2,$$

we identify the metric tensor in coordinates (x, y):
$$g_{\mu\nu} \; = \; \begin{pmatrix} 1 + x^2 & 0 \\ 0 & 1 + y^2 \end{pmatrix}.$$

This is a diagonal metric, with $g_{xx} = 1 + x^2$ and $g_{yy} = 1 + y^2$. Hence $g_{xy} = g_{yx} = 0$.
Its inverse is then
$$g^{\mu\nu} \; = \; \begin{pmatrix} \frac{1}{1+x^2} & 0 \\ 0 & \frac{1}{1+y^2} \end{pmatrix}.$$

Because the metric is diagonal, we simply invert each diagonal element: $g^{xx} = \frac{1}{1+x^2}$ and $g^{yy} = \frac{1}{1+y^2}$.

Christoffel Symbols. Recall the connection formula:
$$\Gamma^\mu_{\;\nu\lambda} = \frac{1}{2}\, g^{\mu\rho} \Big(\partial_\nu g_{\rho\lambda} \; + \; \partial_\lambda g_{\rho\nu} \; - \; \partial_\rho g_{\nu\lambda} \Big).$$

Only $g_{xx} = 1 + x^2$ depends on x, and only $g_{yy} = 1 + y^2$ depends on y. Therefore,
$$\partial_x g_{xx} = 2x, \quad \partial_y g_{yy} = 2y,$$

while all other partial derivatives of $g_{\mu\nu}$ vanish (for example, $\partial_x g_{yy} = 0$ and $\partial_y g_{xx} = 0$).

(i) Calculation of Γ^x_{xx}. We set $\mu = x$, $\nu = x$, $\lambda = x$ in the formula. Hence,
$$\Gamma^x_{\;xx} = \frac{1}{2}\, g^{x\rho} \Big(\partial_x g_{\rho x} \; + \; \partial_x g_{x\rho} \; - \; \partial_\rho g_{xx} \Big).$$

Since the metric is diagonal, $g_{x\rho}$ is nonzero only if $\rho = x$. Therefore, $g^{x\rho}$ is also nonzero only for $\rho = x$. This leaves us with:
$$\Gamma^x_{\;xx} = \frac{1}{2}\, g^{xx} \Big(\partial_x g_{xx} + \partial_x g_{xx} - \partial_x g_{xx} \Big) = \frac{1}{2}\, g^{xx} \big(\partial_x g_{xx} \big).$$

Since $\partial_x g_{xx} = 2x$ and $g^{xx} = \frac{1}{1+x^2}$, we obtain
$$\Gamma^x_{\;xx} = \frac{1}{2} \cdot \frac{1}{1 + x^2} \cdot (2x) = \frac{x}{1 + x^2}.$$

(ii) Calculation of Γ^x_{yy}. We now set $\mu = x$, $\nu = y$, $\lambda = y$:
$$\Gamma^x_{\;yy} = \frac{1}{2}\, g^{x\rho} \Big(\partial_y g_{\rho y} \; + \; \partial_y g_{y\rho} \; - \; \partial_\rho g_{yy} \Big).$$

Here, $\partial_y g_{yy} = 2y$ is the only derivative that might contribute. However, it will appear with $\rho = y$, in which case the factor outside becomes g^{xy}. Since $g^{xy} = 0$ (diagonal inverse), that term vanishes. Alternatively, if $\rho = x$, then $\partial_x g_{yy} = 0$.
Hence,
$$\Gamma^x_{\;yy} = 0.$$

Final Results
$$\boxed{\Gamma^x_{\;xx} = \frac{x}{1 + x^2} \quad \text{and} \quad \Gamma^x_{\;yy} = 0.}$$

Because g_{xx} depends on x, Γ^x_{xx} is non-zero. Meanwhile, g_{yy} does not depend on x, so Γ^x_{yy} vanishes.

Unified Perspective and Key Observations

- **Diagonal Metric Simplification:** *As in other 2D examples, the metric is diagonal, so $g^{\mu\nu}$ is also diagonal. This annihilates many terms in the Christoffel sum (since $g^{xy} = 0$, etc.), making computations much simpler.*

- **Coordinate Dependence vs. Curvature:** *A non-zero Γ^x_{xx} can arise either from genuine curvature or simply from the coordinate dependence of the metric. Here, $g_{xx} = 1 + x^2$ depends on x, producing a nontrivial connection coefficient. Meanwhile, Γ^x_{yy} vanishes because g_{yy} has no x-dependence.*

- **Symmetry in 2D:** *The property $\Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{\lambda\nu}$ cuts the naive 8 symbols down to 6 independent ones. From there, only those involving non-zero derivatives of $g_{\mu\nu}$ can survive.*

- **Summation Index Matching:** *For a diagonal metric, $g^{\mu\rho}$ is nonzero only if $\mu = \rho$. Thus, when we compute Γ^x_{xx}, only the $\rho = x$ term matters. Similarly, for Γ^x_{yy}, the $\rho = y$ term appears multiplied by g^{xy}, which is zero, forcing the entire expression to vanish.*

Question 5: Gravitational Time Dilation

This exercise explores gravitational time dilation, a consequence of Einstein's General Relativity. We will determine the difference in proper time measured by two clocks positioned at varying altitudes on Earth. We approximate Earth's gravitational field using the Schwarzschild metric and utilize the weak field approximation, neglecting effects due to Earth's rotation.

Metric Approximation:

A suitable approximation for the metric outside the Earth's surface (in a weak gravitational field) is:

$$ds^2 = -(1 + 2\Phi)\,dt^2 + (1 - 2\Phi)\,dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2,$$

where $\Phi = -\frac{GM_E}{r}$ is the Newtonian gravitational potential. We use units where the speed of light $c = 1$ unless otherwise specified.

Here, ds^2 is the spacetime interval, dt the coordinate time interval, dr the radial interval, and $d\theta$, $d\phi$ the angular intervals in spherical coordinates. Φ is the Newtonian gravitational potential, G the universal gravitational constant, M_E the Earth's mass, and r the radial distance from the Earth's center. The negative sign in the definition of Φ indicates that work is needed to move an object away from Earth's gravity.

Scenario:

Consider two clocks: one at the Earth's surface $(r = R_E)$ and another atop a building of height h $(r = R_E + h)$. We aim to calculate the proper time elapsed on each clock as a function of coordinate time t and then determine the ratio of these times in the limit $h \ll R_E$.

Proper time is the time measured by a clock along its path in spacetime, while coordinate time is the time measured by an observer at rest at infinity. We want to find the ratio of these proper times when the building's height h is significantly smaller than the Earth's radius R_E.

Solution

(i) Setup and Physical Context We use spherical coordinates (t, r, θ, ϕ). The provided metric is static, spherically symmetric, and valid outside the Earth, with $\Phi(r) = -\frac{GM_E}{r}$.

The metric is time-independent and spherically symmetric, a solution of Einstein's equations, applicable outside the Earth. $\Phi(r)$ is the Newtonian gravitational potential.

Neglecting Earth's rotation and assuming stationary clocks $(dr = d\theta = d\phi = 0)$, the relevant metric component is:

$$ds^2 = -d\tau^2 = -(1 + 2\Phi(r))\,dt^2.$$

Since clocks are stationary relative to Earth, only the time component of the metric matters.

In General Relativity, $ds^2 = -d\tau^2$ relates the spacetime interval between events to the proper time $d\tau$ measured by a clock moving between them. Here, the spacetime interval is purely temporal.

Thus,

$$d\tau = \sqrt{1 + 2\Phi(r)}dt.$$

This relates proper time $d\tau$ to coordinate time dt. Gravitational potential $\Phi(r)$ affects proper time: a more negative potential (closer to Earth) means slower proper time.

(ii) Proper Time Calculation Define:

$$\Phi_1 = \Phi(R_E) = -\frac{GM_E}{R_E}, \quad \Phi_2 = \Phi(R_E + h) = -\frac{GM_E}{R_E + h}.$$

We define gravitational potentials Φ_1 and Φ_2 at the Earth's surface and the building's top, respectively.

Clock 1 (Earth's Surface) At $r = R_E$:

$$d\tau_1 = \sqrt{1 + 2\Phi_1}dt = \sqrt{1 - \frac{2GM_E}{R_E}}dt.$$

Substituting $r = R_E$ into the proper time equation yields the relation between $d\tau_1$ (clock 1's proper time) and dt.

Integrating over a coordinate time interval t:

$$\tau_1 = \int d\tau_1 = \sqrt{1 - \frac{2GM_E}{R_E}} \times t.$$

Integrating over a coordinate time interval t (same for both clocks) gives the total proper time on clock 1.

Clock 2 (Building Top)   At $r = R_E + h$:

$$d\tau_2 = \sqrt{1 + 2\Phi_2}dt = \sqrt{1 - \frac{2GM_E}{R_E + h}}dt.$$

Substituting $r = R_E + h$ into the proper time equation yields the relation between $d\tau_2$ (clock 2's proper time) and dt.
    Integrating:

$$\tau_2 = \sqrt{1 - \frac{2GM_E}{R_E + h}} \times t.$$

Integrating gives the total proper time on clock 2.

(iii) Proper Time Ratio   The ratio of proper times is:

$$\frac{\tau_2}{\tau_1} = \frac{\sqrt{1 - \frac{2GM_E}{R_E + h}}}{\sqrt{1 - \frac{2GM_E}{R_E}}}.$$

We're interested in the limit $h \ll R_E$.
    We want the ratio when the building's height is much smaller than Earth's radius.

(iv) Approximation for $h \ll R_E$ We will use the binomial expansion to simplify the expression for the ratio of proper times, leveraging the fact that h is much smaller than R_E and we are in a weak gravitational field.
    Binomial Expansion:
    The binomial expansion states that for any real number n and $|x| < 1$:

$$(1 + x)^n = 1 + nx + \frac{n(n - 1)}{2!}x^2 + \frac{n(n - 1)(n - 2)}{3!}x^3 + \ldots$$

When x is very small ($|x| \ll 1$), we can approximate the expansion by neglecting higher-order terms:

$$(1 + x)^n \approx 1 + nx$$

Applying the Binomial Expansion:
We have two terms where we can apply the binomial expansion:
1. First Application:
    The terms inside the square roots in our ratio can be written in the form $\sqrt{1 - 2x}$, where $x = \frac{GM_E}{R_E + h}$ or $x = \frac{GM_E}{R_E}$. We can apply the binomial expansion because we are in a **weak gravitational field**, meaning $x = \frac{GM_E}{r}$ is very small when $r \geq R_E$. Specifically, for $r = R_E$, we have $x = \frac{GM_E}{R_E} \approx 6.957 \times 10^{-10}$ (as calculated later), which is much less than 1. Thus we can rewrite $\sqrt{1 - 2x}$ as $(1 - 2x)^{\frac{1}{2}}$ and apply the binomial expansion with $n = \frac{1}{2}$ and x replaced by $-2x$:

$$\sqrt{1 - 2x} = (1 - 2x)^{\frac{1}{2}} \approx 1 + \frac{1}{2}(-2x) = 1 - x$$

Applying this to our ratio, we get:

$$\frac{\tau_2}{\tau_1} \approx \frac{1 - \frac{GM_E}{R_E + h}}{1 - \frac{GM_E}{R_E}}$$

2. Second Application:
We have the term $\frac{1}{R_E + h}$ in the numerator. We can rewrite it as:

$$\frac{1}{R_E + h} = \frac{1}{R_E(1 + \frac{h}{R_E})} = \frac{1}{R_E} \cdot \frac{1}{(1 + \frac{h}{R_E})}$$

Here, we have a term of the form $\frac{1}{1+x}$, where $x = \frac{h}{R_E}$. Since $h \ll R_E$, we have $x \ll 1$. We can rewrite this as $(1 + x)^{-1}$ and apply the binomial expansion with $n = -1$:

$$\frac{1}{(1 + \frac{h}{R_E})} = (1 + \frac{h}{R_E})^{-1} \approx 1 - \frac{h}{R_E}$$

Therefore:

$$\frac{1}{R_E + h} \approx \frac{1}{R_E}\left(1 - \frac{h}{R_E}\right)$$

Substituting and Simplifying:

8

Substituting the second approximation into the first, we get:

$$\frac{\tau_2}{\tau_1} \approx \frac{1 - \frac{GM_E}{R_E}\left(1 - \frac{h}{R_E}\right)}{1 - \frac{GM_E}{R_E}} = \frac{1 - \frac{GM_E}{R_E} + \frac{GM_E h}{R_E^2}}{1 - \frac{GM_E}{R_E}}$$

Approximation in the Weak Field Limit Since we are in a weak gravitational field, the term $\frac{GM_E}{R_E}$ is very small (approximately 6.957×10^{-10} as calculated below). Therefore, we can approximate the denominator as:

$$1 - \frac{GM_E}{R_E} \approx 1$$

This simplifies our expression to:

$$\frac{\tau_2}{\tau_1} \approx 1 + \frac{GM_E h}{R_E^2}$$

We can further calculate $\frac{\tau_2}{\tau_1}$ explicitly using SI units:

$$G \approx 6.674 \times 10^{-11} \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2}$$

$$M_E \approx 5.972 \times 10^{24} \text{ kg}$$

$$R_E \approx 6.371 \times 10^6 \text{ m}$$

$$\frac{\tau_2}{\tau_1} \approx 1 + \frac{6.674 \times 10^{-11} \cdot 5.972 \times 10^{24}}{(6.371 \times 10^6)^2} h \approx 1 + (9.82\frac{\text{m}}{\text{s}^2})\frac{h}{c^2}$$

In natural units, $c = 1$, thus:

$$\frac{\tau_2}{\tau_1} \approx 1 + (9.82\frac{\text{m}}{\text{s}^2})h$$

Geometrized Units - Explanation In the calculations above, we used geometrized units for simplification. Let's explain what that means: To further simplify this expression, we introduce the concept of geometrized units (also known as natural units). In this system, we set fundamental constants like the speed of light c and the gravitational constant G to 1.

Setting $c = 1$: This implies that we are measuring distance and time in the same units. For example, we could measure distance in light-seconds or time in meters.

Setting $G = 1$: This implies that mass, length, and time are all measured in the same units.

When we set $c = 1$ and $G = 1$, the term $\frac{GM_E}{R_E}$ becomes dimensionless. Let's see why: In SI units, the dimensions of the terms are:

$$[G] = \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2}$$

$$[M_E] = \text{kg}$$

$$[R_E] = \text{m}$$

Therefore:

$$\left[\frac{GM_E}{R_E}\right] = \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2} \cdot \frac{\text{kg}}{\text{m}} = \frac{\text{m}^2}{\text{s}^2}$$

This is a velocity squared. In geometrized units, where $c = 1$ and $G = 1$, we have:

$$[c^2] = 1$$

$$[G] = 1$$

Since $c = 1$, then $\left[\frac{m^2}{s^2}\right] = 1$. Thus, $\frac{GM_E}{R_E}$ becomes dimensionless:

$$\left[\frac{GM_E}{R_E}\right] = 1$$

We can calculate its approximate numerical value in SI units and then convert it to geometrized units.

$$\frac{GM_E}{R_E c^2} \approx \frac{6.674 \times 10^{-11} \cdot 5.972 \times 10^{24}}{6.371 \times 10^6 \cdot (2.998 \times 10^8)^2} \approx 6.957 \times 10^{-10}$$

In natural units, $c = 1$ and $G = 1$. We can use the previous result to evaluate $\frac{GM_E}{R_E}$ by dropping the c^2 factor (since $c = 1$) in the denominator:

9

$$\frac{GM_E}{R_E} \approx 6.957 \times 10^{-10}$$

This is why, in geometrized units, we were able to simplify the denominator $1 - \frac{GM_E}{R_E}$ to just 1.

Final Result We found that:

$$\frac{\tau_2}{\tau_1} \approx 1 + \frac{GM_E}{R_E^2} h$$

In geometrized units. Recognizing that $\frac{GM_E}{R_E^2} = g$ (the acceleration due to gravity at the Earth's surface), and converting back to SI units by reintroducing c^2, we get the final result:

$$\frac{\tau_2}{\tau_1} \approx 1 + \frac{gh}{c^2}$$

where $g \approx 9.82 \, \text{m/s}^2$ is the acceleration due to gravity in SI units. We can neglect the term of the gravitational potential because we are in a weak gravitational field.

Question 6

Consider the metric for a two-dimensional sphere of unit radius, given by

$$ds^2 = d\theta^2 + \sin^2\theta \, d\phi^2.$$

This is the standard line element for a sphere in spherical coordinates, where θ is the polar angle (colatitude) and ϕ is the azimuthal angle (longitude). The line element ds^2 represents the infinitesimal squared distance between two nearby points on the sphere.

We label the coordinates as $x^\mu = (\theta, \phi)$. In this setup, the only non-vanishing Christoffel symbols on the sphere are:

$$\Gamma^\theta_{\phi\phi} = -\sin\theta \, \cos\theta \quad \text{and} \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \frac{\cos\theta}{\sin\theta} = \cot\theta.$$

The Christoffel symbols, denoted by Γ, appear in the geodesic equation. They represent the "connection coefficients" of the metric, describing how the basis vectors change from point to point. Christoffel symbols are computed from the metric tensor. In this case, the metric tensor is diagonal with $g_{\theta\theta} = 1$ and $g_{\phi\phi} = \sin^2\theta$. The provided Christoffel symbols follow from these metric components.

Task: Write down the geodesic equations for this metric and use them to show that

(i) lines of constant longitude ($\phi = $ const.) are geodesics,

(ii) the only geodesic at constant latitude ($\theta = $ const.) is the equator ($\theta = \frac{\pi}{2}$).

Solution

Geodesic equations. The geodesic equations in a 2D manifold, for coordinates $x^\mu = (\theta, \phi)$, read

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0,$$

where λ is an affine parameter along the curve.

This is the general form of the geodesic equation. A geodesic is the shortest path between two points in a given geometry (in this case, a curved geometry). The affine parameter λ parametrizes the trajectory. The geodesic equation is a second-order differential equation describing the "straightest possible lines" in a curved space. The first term represents the acceleration along the geodesic, while the second term, involving the Christoffel symbols, accounts for the curvature of the space.

Using the given Christoffel symbols, we obtain the explicit system:

$$\frac{d^2\theta}{d\lambda^2} - \sin\theta \, \cos\theta \left(\frac{d\phi}{d\lambda}\right)^2 = 0, \tag{1}$$

$$\frac{d^2\phi}{d\lambda^2} + 2 \cot\theta \frac{d\theta}{d\lambda} \frac{d\phi}{d\lambda} = 0. \tag{2}$$

These are the specific geodesic equations for the metric on the sphere, obtained by inserting the Christoffel symbols into the general geodesic equation. For $\mu = \theta$, there are contributions from $\Gamma^\theta_{\phi\phi}$, while for $\mu = \phi$, there are contributions from $\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta}$.

TEquation (1) describes how the θ coordinate changes along a geodesic when there is "motion" in the ϕ direction. The term $\sin\theta \, \cos\theta \left(\frac{d\phi}{d\lambda}\right)^2$ acts like a "force" due to the curvature of the sphere, specifically how the ϕ direction varies.

Equation (2) similarly shows how the ϕ coordinate is affected by the combined changes of θ and ϕ. The term $2 \cot\theta \frac{d\theta}{d\lambda} \frac{d\phi}{d\lambda}$ represents a "force" due to curvature, depending on how both θ and ϕ vary.

These equations describe how a free particle would move on the surface of the sphere (with no external forces), following the shortest path (a geodesic) between two points.

(i) Lines at constant longitude. A line at constant longitude implies

$$\phi(\lambda) = \text{const.} \quad \Longrightarrow \quad \frac{d\phi}{d\lambda} = 0, \quad \frac{d^2\phi}{d\lambda^2} = 0.$$

A line of constant longitude means that the coordinate ϕ remains fixed along the path. Therefore, the first and second derivatives of ϕ with respect to the affine parameter λ vanish.

Substitute these into the geodesic equations:

- From Eq. (1):

$$\frac{d^2\theta}{d\lambda^2} = 0.$$

The general solution is $\theta(\lambda) = a\,\lambda + b$. This describes a curve in which θ changes linearly with λ, i.e. a straight line in the θ-coordinate. Since $\frac{d\phi}{d\lambda} = 0$, the second term in equation (1) vanishes. We are left with $\frac{d^2\theta}{d\lambda^2} = 0$, which implies that the rate of change of θ with respect to λ is constant, so θ changes linearly with λ. This is consistent with motion along a meridian (a line of constant longitude).

- From Eq. (2):

$$\frac{d^2\phi}{d\lambda^2} + 2\cot\theta \,\frac{d\theta}{d\lambda}\,\underbrace{\frac{d\phi}{d\lambda}}_{=\,0} = 0 \implies 0 = 0,$$

which is trivially satisfied. Since $\frac{d\phi}{d\lambda} = 0$ and $\frac{d^2\phi}{d\lambda^2} = 0$, equation (2) is automatically satisfied, regardless of the value of θ and $\frac{d\theta}{d\lambda}$. This means that the geodesic equation for ϕ does not impose any additional constraints when ϕ is constant.

Geometrically, keeping ϕ constant means moving along a meridian (a great circle from the north pole to the south pole). The fact that the θ-equation reduces to a simple second-order ODE with constant coefficients confirms that meridians are geodesics. Indeed, a meridian is a line of constant longitude and, on a sphere, meridians are great circles passing through both poles. The fact that the geodesic equation for θ becomes $\frac{d^2\theta}{d\lambda^2} = 0$ shows that motion along a meridian is a geodesic, as it is the "straightest" possible path in the θ direction (no acceleration in θ).

(ii) Lines at constant latitude. A line at constant latitude implies

$$\theta(\lambda) = \text{const.} \implies \frac{d\theta}{d\lambda} = 0, \quad \frac{d^2\theta}{d\lambda^2} = 0.$$

A line of constant latitude means that the coordinate θ is fixed and does not vary along the path. Therefore, its first and second derivatives with respect to λ vanish.

Substitute into the geodesic equations:

- **Eq. (1):**

$$\frac{d^2\theta}{d\lambda^2} - \sin\theta\,\cos\theta \left(\frac{d\phi}{d\lambda}\right)^2 = 0.$$

Since $\frac{d^2\theta}{d\lambda^2} = 0$, we are left with

$$-\sin\theta\,\cos\theta\left(\frac{d\phi}{d\lambda}\right)^2 = 0.$$

Thus,

$$\sin\theta\,\cos\theta\left(\frac{d\phi}{d\lambda}\right)^2 = 0.$$

We conclude one of the following must hold:

- $\sin\theta = 0 \implies \theta = 0$ or $\theta = \pi$. These correspond to the north and south poles, which are single points (not full "circles"). If $\sin\theta = 0$, then θ is 0 or π, which correspond to the north and south poles, respectively. These are single points on the sphere, not lines of latitude.

- $\cos\theta = 0 \implies \theta = \frac{\pi}{2}$. This is precisely the equator. If $\cos\theta = 0$, then $\theta = \frac{\pi}{2}$, which corresponds to the equator. This is a line of constant latitude that is also a great circle.

- $\frac{d\phi}{d\lambda} = 0 \implies \phi = \text{const.}$, which again describes a meridian, not a parallel. If $\frac{d\phi}{d\lambda} = 0$, then ϕ is constant, indicating a meridian (a line of constant longitude), not a line of constant latitude.

- **Eq. (2):**

$$\frac{d^2\phi}{d\lambda^2} + 2\cot\theta\,\frac{d\theta}{d\lambda}\frac{d\phi}{d\lambda} = 0.$$

With $\frac{d\theta}{d\lambda} = 0$, the second term vanishes, so

$$\frac{d^2\phi}{d\lambda^2} = 0 \implies \phi(\lambda) = c\,\lambda + d,$$

meaning ϕ changes linearly with λ. Since $\frac{d\theta}{d\lambda} = 0$, the second term in equation (2) vanishes. We are left with the equation $\frac{d^2\phi}{d\lambda^2} = 0$, which implies that ϕ changes linearly with λ.

Hence, the only nontrivial parallel ($\theta = const$) that can be a geodesic is $\theta = \frac{\pi}{2}$, i.e. the equator. Indeed, the analysis shows that the only constant-latitude line satisfying both geodesic equations is the equator. For any other fixed value of θ (not equal to 0, $\pi/2$, or π), the geodesic equations cannot be simultaneously satisfied unless $\frac{d\phi}{d\lambda} = 0$, which describes a meridian (constant ϕ) rather than a parallel. Geometrically, the equator is a great circle (the largest possible circle on the sphere), whereas any circle at constant $\theta \neq \frac{\pi}{2}$ is not a great circle. Only great circles are geodesics on the sphere. In more intuitive terms, a great circle is the largest circumference one can draw on a sphere, having the same radius as the sphere itself. The equator is one such great circle, but other parallels are not, and thus only the equator is a geodesic among the lines of constant latitude.

Conclusion. We have shown that:

- Lines of constant longitude ($\phi = $ const.) solve the geodesic equations and hence are geodesics (they correspond to meridians). This result confirms that meridians, lines of constant longitude, are indeed geodesics on the sphere. This is consistent with the geometric intuition that meridians are great circles.

- The only latitude ($\theta = $ const.) that is a geodesic is $\theta = \frac{\pi}{2}$, the equator, which is a great circle. This result shows that among all lines of constant latitude, only the equator is a geodesic because it is the only parallel that is also a great circle.

Question 7

Consider the two-dimensional space whose infinitesimal line element is given by

$$ds^2 = d\theta^2 + \sin^2\theta \, d\phi^2,$$

with coordinates $x^\mu = (\theta, \phi)$. The non-vanishing Christoffel symbols are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta \, \cos\theta, \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.$$

Let $V^\mu = (V^\theta, V^\phi)$ be a vector field on this sphere. Compute the covariant derivatives $\nabla_\mu V^\nu$ and $\nabla_\mu V_\nu$, and then show explicitly that

$$\nabla_\mu (V^\nu V_\nu) = \partial_\mu (V^\nu V_\nu).$$

This exercise asks us to work with a two-dimensional space that represents the surface of a unit-radius sphere. The metric is given by the infinitesimal line element $ds^2 = d\theta^2 + \sin^2\theta \, d\phi^2$, where θ is the polar angle and ϕ is the azimuthal angle. We are given the non-zero Christoffel symbols and a vector field $V^\mu = (V^\theta, V^\phi)$ on this sphere. We must compute the covariant derivatives $\nabla_\mu V^\nu$ and $\nabla_\mu V_\nu$, and then demonstrate that the covariant derivative of the scalar product $V^\mu V_\mu$ equals its ordinary partial derivative.

Solution

(i) Metric, Coordinates, and Christoffel Symbols   We have

$$ds^2 = d\theta^2 + \sin^2\theta \, d\phi^2,$$

where θ is the polar angle $(0 \leq \theta \leq \pi)$ and ϕ is the azimuthal angle $(0 \leq \phi < 2\pi)$. *This is the infinitesimal line element for a unit sphere in spherical coordinates. $d\theta^2$ represents the contribution to the infinitesimal distance due to a variation in the polar angle θ, while $\sin^2\theta \, d\phi^2$ represents the contribution due to a variation in the azimuthal angle ϕ. The factor $\sin^2\theta$ accounts for the fact that the circumference of the circles of constant latitude decreases as we approach the poles.* The non-zero Christoffel symbols for this metric are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta \, \cos\theta, \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.$$

These are the non-vanishing Christoffel symbols for the unit-sphere metric. They describe how the basis vectors change from point to point on the sphere. $\Gamma^\theta_{\phi\phi}$ represents the variation of the θ-basis vector when moving along the ϕ-direction, whereas $\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta}$ represents the variation of the ϕ-basis vector when moving along the θ or ϕ directions. We note that these Christoffel symbols depend on θ, reflecting the curvature of the sphere. They arise when computing the connection on the two-sphere (intuitively, they reflect the curvature of the sphere).

(ii) Covariant Derivative of a Contravariant Vector   The covariant derivative of V^ν is

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda.$$

This is the formula for the covariant derivative of a contravariant vector V^ν. The covariant derivative takes into account not only the variation of the vector components but also the variation of the basis vectors themselves, encoded in the Christoffel symbols.

(a)   For $\mu = \theta$ and $\nu = \theta$:

$$\nabla_\theta V^\theta = \partial_\theta V^\theta + \Gamma^\theta_{\theta\lambda} V^\lambda.$$

Here we are computing the $\theta\theta$ component of the covariant derivative. We substitute $\mu = \theta$ and $\nu = \theta$ into the general formula. Since $\Gamma^\theta_{\theta\theta} = 0$ and $\Gamma^\theta_{\theta\phi} = 0$, Because the Christoffel symbols $\Gamma^\theta_{\theta\theta}$ and $\Gamma^\theta_{\theta\phi}$ are both zero for this metric,

$$\nabla_\theta V^\theta = \partial_\theta V^\theta.$$

the covariant derivative $\nabla_\theta V^\theta$ reduces to the ordinary partial derivative $\partial_\theta V^\theta$.

(b)   For $\mu = \theta$ and $\nu = \phi$:

$$\nabla_\theta V^\phi = \partial_\theta V^\phi + \Gamma^\phi_{\theta\lambda} V^\lambda.$$

Now we compute the $\theta\phi$ component of the covariant derivative. We substitute $\mu = \theta$ and $\nu = \phi$ into the general formula. Only $\Gamma^\phi_{\theta\phi} = \cot\theta$ is non-zero, thus The only non-zero Christoffel symbol with $\nu = \phi$ and $\mu = \theta$ is $\Gamma^\phi_{\theta\phi} = \cot\theta$, so

$$\nabla_\theta V^\phi = \partial_\theta V^\phi + \cot\theta \, V^\phi.$$

the covariant derivative $\nabla_\theta V^\phi$ includes the extra term $\cot\theta \, V^\phi$ in addition to the ordinary partial derivative.

(c) For $\mu = \phi$ and $\nu = \theta$:

$$\nabla_\phi V^\theta = \partial_\phi V^\theta + \Gamma^\theta_{\phi\lambda} V^\lambda.$$

We now compute the $\phi\theta$ component of the covariant derivative. Since $\Gamma^\theta_{\phi\phi} = -\sin\theta\,\cos\theta$, The only non-zero Christoffel symbol with $\nu = \theta$ and $\mu = \phi$ is $\Gamma^\theta_{\phi\phi} = -\sin\theta\,\cos\theta$, so

$$\nabla_\phi V^\theta = \partial_\phi V^\theta - \sin\theta\,\cos\theta\,V^\phi.$$

the covariant derivative $\nabla_\phi V^\theta$ includes an extra $-\sin\theta\,\cos\theta\,V^\phi$.

(d) For $\mu = \phi$ and $\nu = \phi$:

$$\nabla_\phi V^\phi = \partial_\phi V^\phi + \Gamma^\phi_{\phi\lambda} V^\lambda.$$

Finally, we compute the $\phi\phi$ component of the covariant derivative. With $\Gamma^\phi_{\phi\theta} = \cot\theta$, The only non-zero Christoffel symbol with $\nu = \phi$ and $\mu = \phi$ is $\Gamma^\phi_{\phi\theta} = \cot\theta$, so

$$\nabla_\phi V^\phi = \partial_\phi V^\phi + \cot\theta\,V^\theta.$$

the covariant derivative $\nabla_\phi V^\phi$ includes an extra $\cot\theta\,V^\theta$.

(iii) Covariant Derivative of a Covariant Vector   The covariant derivative of V_ν is

$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda.$$

This is the formula for the covariant derivative of a covariant vector V_ν. Note that the sign of the term involving the Christoffel symbols is opposite to that for a contravariant vector.

(a) For $\mu = \theta$ and $\nu = \theta$:

$$\nabla_\theta V_\theta = \partial_\theta V_\theta - \Gamma^\lambda_{\theta\theta} V_\lambda.$$

We compute the $\theta\theta$ component of the covariant derivative of a covariant vector. Since $\Gamma^\theta_{\theta\theta} = 0$ and $\Gamma^\phi_{\theta\theta} = 0$, Because the Christoffel symbols $\Gamma^\theta_{\theta\theta}$ and $\Gamma^\phi_{\theta\theta}$ are both zero,

$$\nabla_\theta V_\theta = \partial_\theta V_\theta.$$

the covariant derivative $\nabla_\theta V_\theta$ reduces to the ordinary partial derivative $\partial_\theta V_\theta$.

(b) For $\mu = \theta$ and $\nu = \phi$:

$$\nabla_\theta V_\phi = \partial_\theta V_\phi - \Gamma^\lambda_{\theta\phi} V_\lambda.$$

We compute the $\theta\phi$ component of the covariant derivative. Here $\Gamma^\phi_{\theta\phi} = \cot\theta$, The only non-zero Christoffel symbol with $\mu = \theta$ and $\nu = \phi$ is $\Gamma^\phi_{\theta\phi} = \cot\theta$, so

$$\nabla_\theta V_\phi = \partial_\theta V_\phi - \cot\theta\,V_\phi.$$

the covariant derivative $\nabla_\theta V_\phi$ includes an extra $-\cot\theta\,V_\phi$.

(c) For $\mu = \phi$ and $\nu = \theta$:

$$\nabla_\phi V_\theta = \partial_\phi V_\theta - \Gamma^\lambda_{\phi\theta} V_\lambda.$$

We compute the $\phi\theta$ component of the covariant derivative. Since $\Gamma^\phi_{\phi\theta} = \cot\theta$, The only non-zero Christoffel symbol with $\mu = \phi$ and $\nu = \theta$ is $\Gamma^\phi_{\phi\theta} = \cot\theta$, so

$$\nabla_\phi V_\theta = \partial_\phi V_\theta - \cot\theta\,V_\phi.$$

the covariant derivative $\nabla_\phi V_\theta$ includes an extra $-\cot\theta\,V_\phi$.

(d) For $\mu = \phi$ and $\nu = \phi$:

$$\nabla_\phi V_\phi = \partial_\phi V_\phi - \Gamma^\lambda_{\phi\phi} V_\lambda.$$

Finally, we compute the $\phi\phi$ component of the covariant derivative. With $\Gamma^\theta_{\phi\phi} = -\sin\theta\,\cos\theta$, The only non-zero Christoffel symbol with $\mu = \phi$ and $\nu = \phi$ is $\Gamma^\theta_{\phi\phi} = -\sin\theta\,\cos\theta$, so

$$\nabla_\phi V_\phi = \partial_\phi V_\phi + \sin\theta\,\cos\theta\,V_\theta.$$

the covariant derivative $\nabla_\phi V_\phi$ includes an extra $\sin\theta\,\cos\theta\,V_\theta$.

(iv) Checking $\nabla_\mu(V^\nu V_\nu) = \partial_\mu(V^\nu V_\nu)$ We want to show explicitly that the covariant derivative of the scalar $V^\nu V_\nu$ coincides with its ordinary partial derivative. Let us start from

$$\nabla_\mu(V^\nu V_\nu) = \nabla_\mu(g_{\alpha\beta} V^\alpha V^\beta) = V^\nu \nabla_\mu V_\nu + V_\nu \nabla_\mu V^\nu,$$

where we have used the product rule for the covariant derivative and the fact that the metric $g_{\alpha\beta}$ is covariantly constant, so $\nabla_\mu g_{\alpha\beta} = 0$.

(a) Substituting the definitions of $\nabla_\mu V_\nu$ and $\nabla_\mu V^\nu$ Recall that

$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda \quad \text{and} \quad \nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda.$$

Hence,

$$\nabla_\mu(V^\nu V_\nu) = V^\nu \left(\partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda \right) + V_\nu \left(\partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \right).$$

Expanding,

$$\nabla_\mu(V^\nu V_\nu) = \underbrace{V^\nu \partial_\mu V_\nu + V_\nu \partial_\mu V^\nu}_{\text{(A)}} + \underbrace{\left(-V^\nu \Gamma^\lambda_{\mu\nu} V_\lambda + V_\nu \Gamma^\nu_{\mu\lambda} V^\lambda \right)}_{\text{(B)}}.$$

Part (A) is exactly $\partial_\mu(V^\nu V_\nu)$. We now show that part (B) vanishes.

(b) Splitting part (B) into T_1 and T_2   Define

$$T_1 = -V^\nu \Gamma^\lambda_{\mu\nu} V_\lambda, \qquad T_2 = V_\nu \Gamma^\nu_{\mu\lambda} V^\lambda.$$

Then

$$\text{(B)} = T_1 + T_2.$$

(c) Renaming dummy indices and showing $T_1 + T_2 = 0$ We focus on

$$T_2 = V_\nu \Gamma^\nu_{\mu\lambda} V^\lambda.$$

Since ν and λ are dummy summation indices, we can swap their names:

$$\nu \longmapsto \lambda, \quad \lambda \longmapsto \nu.$$

Hence,

$$T_2 = V_\lambda \Gamma^\lambda_{\mu\nu} V^\nu$$

without changing the numerical value of the sum. Now compare with

$$T_1 = -V^\nu \Gamma^\lambda_{\mu\nu} V_\lambda.$$

Putting them together:

$$T_1 + T_2 = -\left[V^\nu \Gamma^\lambda_{\mu\nu} V_\lambda \right] + \left[V_\lambda \Gamma^\lambda_{\mu\nu} V^\nu \right].$$

Noting that V^ν and V_λ are just real numbers (for fixed indices), we can swap their order of multiplication. Therefore,

$$V_\lambda \Gamma^\lambda_{\mu\nu} V^\nu = V^\nu V_\lambda \Gamma^\lambda_{\mu\nu},$$

and thus

$$T_1 + T_2 = -\left[V^\nu V_\lambda \Gamma^\lambda_{\mu\nu} \right] + \left[V^\nu V_\lambda \Gamma^\lambda_{\mu\nu} \right] = 0.$$

(d) Final result   Since part (B) vanishes, we have

$$\nabla_\mu(V^\nu V_\nu) = \left(\partial_\mu V^\nu \right) V_\nu + V^\nu \left(\partial_\mu V_\nu \right) = \partial_\mu(V^\nu V_\nu).$$

Thus, the covariant derivative of the scalar $V^\nu V_\nu$ is equal to its ordinary partial derivative, as required. The key steps are renaming the summed indices in one of the two terms and recognizing that the components of the vectors are real scalars that commute.

Question 8

Given an arbitrary $(2,0)$ tensor $W^{\mu\nu}$, determine under what conditions the identity

$$\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} \;=\; 0$$

holds.

This exercise concerns the identity of the commutator of covariant derivatives applied to a tensor of type (2,0). The objective is to determine under what conditions such a commutator vanishes. This problem helps us to better understand the relationship between the geometry of spacetime (described by the Riemann tensor and the Ricci tensor) and the symmetry properties of the tensors.

We have an arbitrary tensor $W^{\mu\nu}$ of type (2,0), i.e., with two contravariant indices. We must find the conditions for which the commutator of the covariant derivatives ∇_μ and ∇_ν, applied to $W^{\mu\nu}$ with contracted indices, is equal to zero.

Solution

Before we start, we will make an important assumption: we will assume that we are working in a (pseudo) Riemannian manifold with zero torsion. This is a common assumption in General Relativity and simplifies the calculations considerably. The connection is therefore the Levi-Civita connection, which is the unique torsion-free metric connection. This assumption is not explicitly stated in the problem, but it is often implicitly assumed in this type of exercise. Without this assumption, the calculations would become significantly more complex due to the presence of torsion terms in the commutator of covariant derivatives. By making this assumption, we can focus on the core concepts related to the Riemann and Ricci tensors.

(i) Commutator of Covariant Derivatives   *Let's start by recalling the general expression for the commutator of two covariant derivatives acting on a (2,0) tensor. This expression involves the Riemann tensor, which describes the curvature of spacetime. The commutator essentially measures how much the order of covariant differentiation matters.* We start by recalling the general expression for the commutator of covariant derivatives acting on a $(2,0)$ tensor:

$$\left[\nabla_\alpha, \nabla_\beta\right] W^{\mu\nu} \;=\; R^\mu{}_{\sigma\alpha\beta} W^{\sigma\nu} \;+\; R^\nu{}_{\sigma\alpha\beta} W^{\mu\sigma},$$

where $R^\rho{}_{\sigma\mu\nu}$ is the Riemann curvature tensor. *Here, $[\nabla_\alpha, \nabla_\beta]$ represents the commutator of the covariant derivatives ∇_α and ∇_β, defined as $\nabla_\alpha\nabla_\beta - \nabla_\beta\nabla_\alpha$. This means we apply first ∇_β, then ∇_α, and subtract the result of applying them in the reverse order. $W^{\mu\nu}$ is a generic tensor of type (2,0), meaning it has two upper indices (contravariant). $R^\rho{}_{\sigma\mu\nu}$ is the Riemann curvature tensor, a fundamental object in differential geometry that quantifies the curvature of spacetime. It tells us how much a vector changes when parallel transported around an infinitesimal loop. The indices of the Riemann tensor have specific meanings: ρ is the index of the resulting vector after parallel transport, σ is the index of the original vector, and μ and ν define the plane of the infinitesimal loop.*

The first term on the right-hand side of the equation, $R^\mu{}_{\sigma\alpha\beta} W^{\sigma\nu}$, represents the effect of curvature on the μ component of the tensor. We contract (sum over) the index σ of the Riemann tensor with the first index of $W^{\sigma\nu}$. The second term, $R^\nu{}_{\sigma\alpha\beta} W^{\mu\sigma}$, represents the effect of curvature on the ν component of the tensor. Here, we contract the index σ of the Riemann tensor with the second index of $W^{\mu\sigma}$. In practice, this equation tells us how the parallel transport of a tensor around an infinitesimal parallelogram (defined by α and β) differs from the parallel transport in the opposite direction.

(ii) Specializing to $W^{\mu\nu}$ with Indices Contracted   *Now we apply the general formula of the commutator to our specific case, where we want the commutator to vanish when the indices of the tensor $W^{\mu\nu}$ are contracted with the indices of the covariant derivatives in the commutator. This means we are setting $\alpha = \mu$ and $\beta = \nu$ and summing over these indices.* We want the commutator to vanish when the resulting tensor is $W^{\mu\nu}$ contracted with the same pair of indices (μ, ν) as appear in the commutator:

$$\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} \;=\; 0.$$

This means that we are calculating $[\nabla_\mu, \nabla_\nu] W^{\mu\nu}$, where the first index of the commutator (μ) is contracted with the first index of the tensor (μ), and the second index of the commutator (ν) is contracted with the second index of the tensor (ν). We are summing over repeated indices, as per the Einstein summation convention. Substituting $\alpha = \mu$ and $\beta = \nu$ into the general formula, we get

$$\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} \;=\; R^\mu{}_{\sigma\mu\nu} W^{\sigma\nu} \;+\; R^\nu{}_{\sigma\mu\nu} W^{\mu\sigma}.$$

Substituting $\alpha = \mu$ and $\beta = \nu$ into the general formula, we obtain this specific expression for our case. Note that the indices μ and ν now appear both as covariant derivative indices and as indices of the Riemann tensor. We are now summing over μ and ν in the commutator and on the right-hand side.

(iii) Ricci Tensor and Antisymmetry *We now use the definition of the Ricci tensor and an important antisymmetry property of the Riemann tensor to simplify the expression. These are key properties that will allow us to manipulate the equation.* Next, we recall that the Ricci tensor $R_{\mu\nu}$ is defined as the contraction

$$R_{\mu\nu} \;=\; R^{\lambda}{}_{\mu\lambda\nu}.$$

The Ricci tensor $R_{\mu\nu}$ is obtained by contracting (summing over) the first and third indices of the Riemann tensor. This means we set the first index (ρ) and the third index (λ) of the Riemann tensor to be the same, λ, and sum over all possible values of λ. The Ricci tensor is a measure of the curvature of spacetime, and it plays a crucial role in Einstein's field equations. Hence, the first term can be written as

$$R^{\mu}{}_{\sigma\mu\nu}\, W^{\sigma\nu} \;=\; R_{\sigma\nu}\, W^{\sigma\nu}.$$

Using the definition of the Ricci tensor, we can rewrite the first term as a contraction between the Ricci tensor and the tensor $W^{\sigma\nu}$. More specifically we are using the definition of the Ricci tensor: $R_{\sigma\nu} = R^{\mu}{}_{\sigma\mu\nu}$, where we contracted the first and the third indices of the Riemann tensor. Since we are summing over μ, we can replace $R^{\mu}{}_{\sigma\mu\nu}$ with $R_{\sigma\nu}$.

For the second term, we use a key antisymmetry property of the Riemann tensor:

$$R^{\nu}{}_{\sigma\mu\nu} \;=\; -\,R^{\nu}{}_{\sigma\nu\mu}.$$

The Riemann tensor is antisymmetric in the last two indices. This means that by swapping the last two indices, the tensor changes sign. This is a fundamental property of the Riemann tensor, arising from its geometric interpretation in terms of parallel transport around an infinitesimal loop. Swapping the last two indices corresponds to reversing the direction of the loop. It follows that

$$R^{\nu}{}_{\sigma\mu\nu}\, W^{\mu\sigma} \;=\; -\,R^{\nu}{}_{\sigma\nu\mu}\, W^{\mu\sigma} \;=\; -\,R_{\sigma\mu}\, W^{\mu\sigma}.$$

Using this antisymmetry property, we can rewrite the second term, changing the sign and swapping the last two indices of the Riemann tensor. Then we can contract again the first and third index to obtain the Ricci tensor: $R^{\nu}{}_{\sigma\nu\mu} = R_{\sigma\mu}$. We are allowed to do this because we are summing over ν, so we can replace $R^{\nu}{}_{\sigma\nu\mu}$ with $R_{\sigma\mu}$.

(iv) Combining the Two Terms *Now we put the two terms back together and exploit the fact that dummy indices can be renamed. Dummy indices are indices that are summed over, and we can change their names without changing the meaning of the expression, as long as we do it consistently.* Putting these two contributions together, we have

$$\bigl[\nabla_{\mu}, \nabla_{\nu}\bigr] W^{\mu\nu} \;=\; R_{\sigma\nu}\, W^{\sigma\nu} \;-\; R_{\sigma\mu}\, W^{\mu\sigma}.$$

Observe that the two indices in the second term (σ, μ) are dummy indices. *The indices σ and μ in the second term are "dummy" indices, i.e., summation indices that can be renamed without changing the meaning of the expression. This is because we are summing over all possible values of these indices, so the specific letter used to denote them doesn't matter.*

By swapping and index in the second term (renaming $\mu \to \nu$), we get

$$R_{\sigma\nu}\, W^{\sigma\nu} \;-\; R_{\sigma\nu}\, W^{\nu\sigma} \;=\; R_{\sigma\nu}\bigl(W^{\sigma\nu} - W^{\nu\sigma}\bigr).$$

By renaming the dummy indices in the second term, we can rewrite the expression as a difference between two terms involving the Ricci tensor and the tensor W, but with the indices swapped. We renamed μ to ν, so the expression becomes $R_{\sigma\nu}W^{\nu\sigma}$. Then we can factor out the Ricci tensor, since it's now the same in both terms. Therefore,

$$\bigl[\nabla_{\mu}, \nabla_{\nu}\bigr] W^{\mu\nu} \;=\; R_{\sigma\nu}\bigl(W^{\sigma\nu} - W^{\nu\sigma}\bigr).$$

For this quantity to vanish for the *arbitrary* tensor $W^{\mu\nu}$, the factor in parentheses must vanish unless $R_{\sigma\nu} = 0$ everywhere. *For this expression to be equal to zero for an arbitrary tensor $W^{\mu\nu}$, the factor in parentheses must be zero, unless the Ricci tensor is identically zero. If the factor in parentheses is zero, the whole expression is zero, regardless of the value of the Ricci tensor. If the Ricci tensor is zero, the whole expression is zero, regardless of the value of $W^{\mu\nu}$.*

(v) Conditions for Vanishing Commutator and the Role of Torsion *We now analyze the conditions under which the commutator vanishes. We had initially identified two cases, but a closer examination, particularly in light of the tensor decomposition discussed below, reveals a more nuanced relationship between these cases and the underlying assumptions.*

Recall the result we derived:

$$\bigl[\nabla_{\mu}, \nabla_{\nu}\bigr] W^{\mu\nu} \;=\; R_{\sigma\nu}\bigl(W^{\sigma\nu} - W^{\nu\sigma}\bigr).$$

For this expression to be zero for arbitrary $W^{\mu\nu}$, we initially identified two distinct possibilities:

- **Case 1: $W^{\mu\nu}$ is symmetric.** If $W^{\mu\nu} = W^{\nu\mu}$, the term in the parentheses vanishes, leading to a zero commutator regardless of the value of the Ricci tensor $R_{\sigma\nu}$.

- **Case 2: $R_{\mu\nu} = 0$ (Ricci-flat spacetime).** If the Ricci tensor is zero, the entire expression vanishes, regardless of the symmetry properties of $W^{\mu\nu}$. This condition corresponds to a vacuum solution in General Relativity (in the absence of a cosmological constant), where $T_{\mu\nu} = 0$. *An example of this is the void case. The Ricci tensor describes how an infinitesimal volume of particles that moves along geodesics expands or contracts. $R_{\mu\nu} = 0$ means there is no volume variation, but you can still have curvature of spacetime as described by the full Riemann tensor.*

(vi) Physical Interpretation (Optional Remark)   *We give a brief physical interpretation of the two conditions. This helps to connect the mathematical results to their physical meaning in the context of General Relativity.* In the context of General Relativity, the condition $R_{\mu\nu} = 0$ (with zero cosmological constant) implies a vacuum solution of Einstein's field equations, indicating no matter or energy content. *In the context of General Relativity, the condition $R_{\mu\nu} = 0$ (in the absence of a cosmological constant) corresponds to a vacuum solution of Einstein's field equations, that is, a spacetime without matter or energy. This means that the spacetime is empty, and its curvature is entirely determined by the initial conditions.* On the other hand, requiring $W^{\mu\nu}$ to be symmetric imposes a restriction on the form of the tensor rather than on the geometry of the spacetime. *On the other hand, requiring that $W^{\mu\nu}$ be symmetric imposes a restriction on the form of the tensor, but not on the geometry of spacetime. It's a condition on the specific tensor we are considering, not on the underlying spacetime itself.*

(vii) Final Answer   *Let's summarize the final result.* Hence,

$$\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} \;=\; 0 \quad \Longleftrightarrow \quad \text{either } W^{\mu\nu} \text{ is symmetric, or } R_{\mu\nu} = 0.$$

Ultimately, the commutator $\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu}$ is equal to zero if and only if $W^{\mu\nu}$ is a symmetric tensor, or if the spacetime is Ricci-flat ($R_{\mu\nu} = 0$). These are the two possible conditions under which the given identity holds. However, read the following section to have a different view/approach to this problem.

Solution, different approach

We decompose the (2,0) tensor $W^{\mu\nu}$ into symmetric ($S^{\mu\nu}$) and antisymmetric ($A^{\mu\nu}$) parts:
$$W^{\mu\nu} = S^{\mu\nu} + A^{\mu\nu}$$
$$S^{\mu\nu} = \frac{1}{2}(W^{\mu\nu} + W^{\nu\mu}), \quad A^{\mu\nu} = \frac{1}{2}(W^{\mu\nu} - W^{\nu\mu})$$

Substituting into the commutator $\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu}$ yields:
$$\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} = R_{\sigma\nu}(S^{\sigma\nu} - S^{\nu\sigma}) + R_{\sigma\nu}(A^{\sigma\nu} - A^{\nu\sigma})$$

The symmetric part's contribution vanishes. The antisymmetric part contributes $2R_{\sigma\nu}A^{\sigma\nu}$.
Since the Ricci tensor is symmetric, its contraction with the antisymmetric part $A^{\sigma\nu}$ is always zero.

We assume a torsion-free connection (specifically, the Levi-Civita connection), so $T^\rho_{\mu\nu} = 0$.
With torsion, our analysis would be more complex.

Under zero torsion, we conclude:
The commutator $\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu}$ is always zero for any (2,0) tensor $W^{\mu\nu}$ in a torsion-free spacetime. This is because $W^{\mu\nu}$ can always be decomposed into symmetric and antisymmetric parts. The commutator vanishes for the symmetric part trivially, and for the antisymmetric part, it vanishes because the contraction of the symmetric Ricci tensor with an antisymmetric tensor is always zero.

Therefore, in a torsion-free spacetime, the identity $\left[\nabla_\mu, \nabla_\nu\right] W^{\mu\nu} = 0$ always holds for any (2,0) tensor.

Question 8: Theoretical Background: Understanding the Building Blocks

This appendix provides a brief overview of the fundamental concepts used in the exercise, including the Riemann curvature tensor, the Ricci tensor, and the covariant derivative. We will aim for an intuitive understanding, emphasizing the geometric meaning of these objects.

Covariant Derivative: Parallel Transport in Curved Spacetime   *In flat space, we are used to the idea that parallel lines never meet. However, in curved space, the notion of "parallel" becomes more subtle. The covariant derivative is a way to generalize differentiation to curved spaces, taking into account the curvature of the space. It allows us to define how a tensor changes as we move it along a curve while keeping it "as parallel as possible" to its initial direction.*

   Imagine you are walking on the surface of a sphere, carrying an arrow that always points in the same direction, tangent to the sphere. As you walk along a curved path, the direction of the arrow relative to your path will change, even though you are trying to keep it parallel. This change is due to the curvature of the sphere.

The covariant derivative, denoted by ∇_μ, captures this notion of change in curved space. When applied to a vector V^ν, it is defined as:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda$$

where ∂_μ is the ordinary partial derivative, and $\Gamma^\nu_{\mu\lambda}$ are the Christoffel symbols (also known as the connection coefficients). The Christoffel symbols encode information about the curvature of the space and can be computed from the metric tensor $g_{\mu\nu}$, which defines distances in the space.

Intuition: The first term, $\partial_\mu V^\nu$, represents the change in the vector as we would calculate it in flat space. The second term, $\Gamma^\nu_{\mu\lambda} V^\lambda$, is a "correction" term that accounts for the curvature of the space. It tells us how much the vector needs to be adjusted to keep it parallel as we move along the curve.

The covariant derivative can be similarly extended to tensors of higher rank. For example, for a (2,0) tensor $W^{\mu\nu}$, we have:

$$\nabla_\alpha W^{\mu\nu} = \partial_\alpha W^{\mu\nu} + \Gamma^\mu_{\alpha\sigma} W^{\sigma\nu} + \Gamma^\nu_{\alpha\sigma} W^{\mu\sigma}$$

Each upper index gets a correction term with a positive sign

The idea is that we are comparing the tensor at a given point with the tensor at a neighboring point, but we need to account for the fact that the coordinate system itself is changing due to the curvature. **Key takeaway: the covariant derivative allows for differentiation in curved space by keeping tensors parallel during transport, encapsulating the intrinsic curvature's effect on differentiation.**

Riemann Curvature Tensor: Measuring the Curvature *The Riemann curvature tensor, denoted by $R^\rho{}_{\sigma\mu\nu}$, is the central object that quantifies the curvature of a (pseudo) Riemannian manifold. It measures how much the geometry of the space deviates from flat Euclidean space.*

Geometric Interpretation: One way to understand the Riemann tensor is to consider parallel transporting a vector around a small closed loop in the space. If the space is flat, the vector will return to its original direction after being transported around the loop. However, if the space is curved, the vector will be rotated by an amount proportional to the area enclosed by the loop and the Riemann tensor.

More precisely, if we parallel transport a vector V^σ around an infinitesimal parallelogram defined by two vectors A^μ and B^ν, the change in the vector δV^ρ after one loop is given by:

$$\delta V^\rho \approx R^\rho{}_{\sigma\mu\nu} V^\sigma A^\mu B^\nu$$

This equation shows that the Riemann tensor relates the change in a vector under parallel transport to the area of the loop and the components of the vector itself.

Another way to interpret the Riemann tensor is through the commutator of covariant derivatives. As we saw in the exercise, for a (2,0) tensor:

$$\left[\nabla_\alpha, \nabla_\beta\right] W^{\mu\nu} = R^\mu{}_{\sigma\alpha\beta} W^{\sigma\nu} + R^\nu{}_{\sigma\alpha\beta} W^{\mu\sigma},$$

For a vector V^μ, the commutator is given by

$$[\nabla_\mu, \nabla_\nu] V^\rho = R^\rho{}_{\sigma\mu\nu} V^\sigma$$

This means the Riemann tensor also measures the non-commutativity of covariant derivatives. In flat space, the order of differentiation doesn't matter, so the commutator is zero. In curved space, the commutator is non-zero, and the Riemann tensor quantifies this non-commutativity.

Symmetries of the Riemann Tensor: The Riemann tensor has several important symmetries:

- $R^\rho{}_{\sigma\mu\nu} = -R^\rho{}_{\sigma\nu\mu}$: Antisymmetry in the last two indices. This reflects the fact that reversing the direction of the loop in the parallel transport picture reverses the rotation of the vector.

- $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$: Antisymmetry in the first two indices (when lowered with the metric).

- $R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$: Pair exchange symmetry.

- $R^\rho{}_{\sigma\mu\nu} + R^\rho{}_{\mu\nu\sigma} + R^\rho{}_{\nu\sigma\mu} = 0$: The first Bianchi identity.

These symmetries reduce the number of independent components of the Riemann tensor.

Key Takeaway: the Riemann curvature tensor is a fundamental geometric object capturing the essence of curvature by measuring the change in a vector after parallel transport around a closed loop and quantifying the non-commutativity of covariant derivatives.

Ricci Tensor: A Contraction of the Riemann Tensor *The Ricci tensor, denoted by $R_{\mu\nu}$, is obtained by contracting the first and third indices of the Riemann tensor:*

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}$$

In other words, we set $\rho = \lambda$ in $R^{\rho}{}_{\sigma\mu\nu}$ and sum over λ.

Geometric Interpretation: The Ricci tensor measures a particular aspect of the curvature: the tendency of volumes to change in curved space compared to flat space. More precisely, it describes how the volume of a small geodesic ball (a ball defined by geodesics emanating from a central point) differs from the volume of a ball of the same radius in Euclidean space.

If $R_{\mu\nu}$ is positive definite in a region, it means that volumes in that region tend to be smaller than in flat space. If $R_{\mu\nu}$ is negative definite, volumes tend to be larger. If $R_{\mu\nu} = 0$, volumes are locally the same as in flat space, at least to leading order.

In General Relativity: The Ricci tensor plays a crucial role in Einstein's field equations, which relate the geometry of spacetime to the distribution of mass and energy:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}$$

where R is the Ricci scalar (the trace of the Ricci tensor, $R = g^{\mu\nu}R_{\mu\nu}$), $g_{\mu\nu}$ is the metric tensor, G is the gravitational constant, and $T_{\mu\nu}$ is the stress-energy tensor. This equation tells us that the Ricci curvature is directly related to the presence of matter and energy.

Key Takeaway: The Ricci tensor provides a measure of how volumes change in curved space compared to flat space and is directly linked to the distribution of mass and energy in General Relativity.

Question 9

Consider a photon moving in a 2D spacetime described by the metric

$$ds^2 = -\left(1 + \frac{x^2}{\ell^2}\right) dt^2 + dx^2,$$

where ℓ is a constant. The photon follows a geodesic $x^\mu(\lambda)$ parametrized by an affine parameter λ, and its four-momentum is given by $p^\mu = \frac{dx^\mu}{d\lambda}$. Using the geodesic equation, derive all the conserved quantities for this system. Then use these results to express $\frac{dx}{d\lambda}$ in terms of the radial coordinate x and the conserved quantities.

Solution

1. The Metric and Photon Setup. *We are given the 2D metric:*

$$ds^2 = -\left(1 + \frac{x^2}{\ell^2}\right) dt^2 + dx^2.$$

A photon follows a geodesic $x^\mu(\lambda)$ parametrized by an affine parameter λ. Its four-momentum is defined as

$$p^\mu = \frac{dx^\mu}{d\lambda}.$$

2. Geodesic Equation. *The standard geodesic equation is:*

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0, \tag{3}$$

where $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols. They are computed from

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2} g^{\mu\sigma} \left(\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\sigma\alpha} - \partial_\sigma g_{\alpha\beta}\right).$$

Equation (3) tells us how the tangent vector to the geodesic, i.e., the four-momentum $p^\mu = \frac{dx^\mu}{d\lambda}$, changes along the geodesic itself. The Christoffel symbols $\Gamma^\mu_{\alpha\beta}$ represent the affine connection, and they encode how the spacetime is curved. They depend on the first derivatives of the metric.

3. Metric Components. *In our 2D case, the components of the metric $g_{\mu\nu}$ are:*

$$g_{tt} = -\left(1 + \frac{x^2}{\ell^2}\right), \qquad g_{xx} = 1,$$

and all other components are zero. The inverse metric is:

$$g^{tt} = -\frac{1}{1 + \frac{x^2}{\ell^2}}, \qquad g^{xx} = 1.$$

The inverse metric is obtained by calculating the inverse of the metric tensor matrix. We can verify that $g_{\mu\alpha} g^{\alpha\nu} = \delta^\nu_\mu$, as required.

4. Calculation of the Christoffel Symbols. *We only need the Christoffel symbols Γ^x_{tt} and Γ^t_{tx}. Given that the only non-zero derivative of the metric is $\partial_x g_{tt} = -\frac{2x}{\ell^2}$, we have:*

$$\Gamma^x_{tt} = \frac{1}{2} g^{xx} (\partial_t g_{xt} + \partial_t g_{tx} - \partial_x g_{tt}) = \frac{1}{2} \cdot 1 \cdot (0 + 0 - (-\frac{2x}{\ell^2})) = \frac{x}{\ell^2}$$

Here, we used the fact that $g^{xx} = 1$, and the only surviving term is $-\partial_x g_{tt}$ because $g_{tx} = g_{xt} = 0$.

$$\Gamma^t_{tx} = \Gamma^t_{xt} = \frac{1}{2} g^{tt} (\partial_t g_{tx} + \partial_x g_{tt} - \partial_t g_{tx}) = \frac{1}{2} \cdot \left(-\frac{1}{1 + \frac{x^2}{\ell^2}}\right) \cdot (-\frac{2x}{\ell^2}) = \frac{x}{\ell^2} \frac{1}{1 + \frac{x^2}{\ell^2}} = \frac{x}{x^2 + \ell^2}$$

Similarly, we used $g^{tt} = -\frac{1}{1 + x^2/\ell^2}$, and only $\partial_x g_{tt}$ is non-zero.

All other Christoffel symbols are zero. *This can be verified by plugging in the other indices into the formula for the Christoffel symbols.*

5. Applying the Geodesic Equation.

Let's write out the geodesic equation for $\mu = t$:

$$\frac{d^2t}{d\lambda^2} + \Gamma^t_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0$$

Since only $\Gamma^t_{tx} = \Gamma^t_{xt}$ is non-zero, this simplifies to:

$$\frac{d^2t}{d\lambda^2} + 2\Gamma^t_{tx}\frac{dt}{d\lambda}\frac{dx}{d\lambda} = 0$$

The factor of 2 comes from the fact that Γ^t_{tx} and Γ^t_{xt} are equal and both contribute.
Substituting the value of $\Gamma^t_{tx} = \frac{x}{x^2+\ell^2}$, we have:

$$\frac{d^2t}{d\lambda^2} + 2\frac{x}{x^2+\ell^2}\frac{dt}{d\lambda}\frac{dx}{d\lambda} = 0$$

Multiplying by $(1 + \frac{x^2}{\ell^2})$, we get:

$$(1 + \frac{x^2}{\ell^2})\frac{d^2t}{d\lambda^2} + \frac{2x}{\ell^2}\frac{x^2+\ell^2}{x^2+\ell^2}\frac{dt}{d\lambda}\frac{dx}{d\lambda} = 0$$

$$\left(1 + \frac{x^2}{\ell^2}\right)\frac{d^2t}{d\lambda^2} + \frac{2x}{\ell^2}\frac{dt}{d\lambda}\frac{dx}{d\lambda} = 0$$

Notice that this can be written as:

$$\frac{d}{d\lambda}\left((1 + \frac{x^2}{\ell^2})\frac{dt}{d\lambda}\right) = 0$$

Explicit passages: $\dfrac{d}{d\lambda}\left[\left(1 + \dfrac{x^2}{\ell^2}\right)\dfrac{dt}{d\lambda}\right]$

using the product rule: $\dfrac{d}{d\lambda}[f(\lambda)g(\lambda)] = f'(\lambda)g(\lambda) + f(\lambda)g'(\lambda),$

with $f(\lambda) = 1 + \dfrac{x^2}{\ell^2}$ *and* $g(\lambda) = \dfrac{dt}{d\lambda}.$

Since $\dfrac{d}{d\lambda}\left(1 + \dfrac{x^2}{\ell^2}\right) = \dfrac{1}{\ell^2}\dfrac{d}{d\lambda}(x^2) = \dfrac{2x}{\ell^2}\dfrac{dx}{d\lambda},$

we obtain $\dfrac{d}{d\lambda}\left[\left(1 + \dfrac{x^2}{\ell^2}\right)\dfrac{dt}{d\lambda}\right]$

$= \dfrac{2x}{\ell^2}\dfrac{dx}{d\lambda}\dfrac{dt}{d\lambda} + \left(1 + \dfrac{x^2}{\ell^2}\right)\dfrac{d^2t}{d\lambda^2}.$

This implies that $(1 + \frac{x^2}{\ell^2})\frac{dt}{d\lambda}$ is a conserved quantity. We call this constant of motion E, so that

$$\left(1 + \frac{x^2}{\ell^2}\right)\frac{dt}{d\lambda} = E.$$

This is our first conserved quantity: the photon's energy. We interpret E as the energy because the metric is time-translation invariant, meaning it doesn't depend explicitly on t.

6. Null Condition for the Photon and $\frac{dx}{d\lambda}$.

Since we are dealing with a photon (a massless particle), its path is a null geodesic. That is,

$$ds^2 = 0.$$

Thus, using our metric,

$$0 = -\left(1 + \frac{x^2}{\ell^2}\right)\left(\frac{dt}{d\lambda}\right)^2 + \left(\frac{dx}{d\lambda}\right)^2.$$

This equation reflects that the photon moves at the speed of light in this curved 2D geometry.
We can substitute the result from (),

$$\frac{dt}{d\lambda} = \frac{E}{1 + \frac{x^2}{\ell^2}},$$

into the null condition. Then we get:

$$0 = -\left(1 + \frac{x^2}{\ell^2}\right)\left(\frac{E}{1 + \frac{x^2}{\ell^2}}\right)^2 + \left(\frac{dx}{d\lambda}\right)^2.$$

23

Simplifying,

$$\left(\frac{dx}{d\lambda}\right)^2 = \frac{E^2}{1 + \frac{x^2}{\ell^2}}.$$

We simply moved the negative term to the right-hand side and canceled one factor of $(1 + \frac{x^2}{\ell^2})$.
   Taking the square root,

$$\frac{dx}{d\lambda} = \pm \frac{E}{\sqrt{1 + \frac{x^2}{\ell^2}}}.$$

The $\pm$ *sign indicates that the photon can move in either the positive or negative x direction.*

7. Final Expressions and Physical Interpretation.

- *From the geodesic equation we found* $\left(1 + \frac{x^2}{\ell^2}\right)\frac{dt}{d\lambda} = E$, *where E is a* conserved *quantity interpreted as the energy of the photon*

- *From this, we get* $\frac{dt}{d\lambda} = \frac{E}{1 + x^2/\ell^2}$. This tells us how coordinate time t changes with the affine parameter λ along the photon's path.

- *The null condition* $ds^2 = 0$ *then gives* $\frac{dx}{d\lambda} = \pm\frac{E}{\sqrt{1 + x^2/\ell^2}}$. This describes the spatial velocity of the photon in this geometry.

Thus, we see that the photon's motion in (t, x) *coordinates is completely described once the constant energy E is specified.*

Exercise 10

Starting from the expression of the Schwarzschild metric, compute the explicit expressions for the Christoffel symbols Γ^r_{rr} and $\Gamma^\phi_{r\phi}$.

Extended Solution with Additional Explanations

(i) Schwarzschild Metric. We consider the Schwarzschild metric in the coordinates (t, r, θ, ϕ):

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

   This metric describes the curved spacetime outside a spherically symmetric mass. A spherically symmetric mass means that its gravitational influence depends only on the distance from the center and not on directions. The term $2GM$ is called the Schwarzschild radius, and when r is at this value, one reaches the event horizon of a black hole solution. The Schwarzschild metric is a vacuum solution of Einstein's field equations, implying there is no matter or energy density outside the central mass. The coordinate t can be seen as the time measured by an observer far away from the mass (at infinity), and r can be understood as the radial coordinate, though at small distances it does not behave like the usual flat-space radius. The angles θ and ϕ are the usual spherical angles.

The nonzero components of the metric tensor $g_{\mu\nu}$ are:

$$g_{tt} = -\left(1 - \frac{2GM}{r}\right), \quad g_{rr} = \left(1 - \frac{2GM}{r}\right)^{-1}, \quad g_{\theta\theta} = r^2, \quad g_{\phi\phi} = r^2 \sin^2 \theta.$$

All other components vanish.

   We see that these components are arranged diagonally, reflecting spherical symmetry and no rotation (no off-diagonal terms). The negative sign in g_{tt} is a signature convention often used in General Relativity, and it encodes the fact that t is a timelike coordinate.

The inverse metric tensor $g^{\mu\nu}$ has the nonzero components:

$$g^{tt} = -\left(1 - \frac{2GM}{r}\right)^{-1}, \quad g^{rr} = \left(1 - \frac{2GM}{r}\right), \quad g^{\theta\theta} = \frac{1}{r^2}, \quad g^{\phi\phi} = \frac{1}{r^2 \sin^2 \theta}.$$

   Notice how each nonzero component of the inverse metric is simply the reciprocal of the corresponding component of $g_{\mu\nu}$ (up to a sign for the time component). This pattern is characteristic of a diagonal metric. The factor $\left(1 - \frac{2GM}{r}\right)$ is central to the Schwarzschild solution and it appears in many calculations of physical quantities, such as orbital speeds and gravitational redshift.

(ii) Christoffel Symbols: General Formula. The Christoffel symbols $\Gamma^\rho_{\mu\nu}$ are defined by:

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right).$$

   These symbols can be thought of as describing the "connection" in a curved space or spacetime, telling us how to take derivatives of vector or tensor fields when coordinates change from one point to another. Although they are not tensors themselves, they appear in important equations such as the geodesic equation. Recall that the geodesic equation

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

tells us how a free-falling particle (or photon) moves under gravity alone. For instance, setting the coordinate indices appropriately will reveal how a radial distance or an angular coordinate might change if there is no external force except gravity.

We are specifically interested in the components Γ^r_{rr} and $\Gamma^\phi_{r\phi}$.

(iii) Calculation of Γ^r_{rr}. We set $\rho = r$, $\mu = r$, and $\nu = r$. Then, by the definition:

$$\Gamma^r_{rr} = \frac{1}{2} g^{r\sigma} \left(\partial_r g_{\sigma r} + \partial_r g_{\sigma r} - \partial_\sigma g_{rr}\right).$$

Since $g^{r\sigma}$ is nonzero only when $\sigma = r$, we get:

$$\Gamma^r_{rr} = \frac{1}{2} g^{rr} \left(\partial_r g_{rr} + \partial_r g_{rr} - \partial_r g_{rr}\right) = \frac{1}{2} g^{rr} \partial_r g_{rr}.$$

   Observe that one of the three terms inside the parentheses is subtracted, so we effectively get a single partial derivative. This is a common simplification that happens when two of the terms combine to cancel out the third.

Recall that
$$g_{rr} = \left(1 - \tfrac{2GM}{r}\right)^{-1},$$
so its derivative with respect to r is (using the derivative of nested functions)
$$\partial_r \left(1 - \tfrac{2GM}{r}\right)^{-1} = -\left(1 - \tfrac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2}.$$

Also note that
$$g^{rr} = \left(1 - \tfrac{2GM}{r}\right).$$

It may help to do the derivative step by step. Let $f(r) = 1 - \tfrac{2GM}{r}$. Then $g_{rr} = f(r)^{-1}$. By the chain rule, $\frac{d}{dr}[f(r)^{-1}] = -f(r)^{-2}\frac{df}{dr}$. And $\frac{df}{dr} = \frac{2GM}{r^2}$. Combining these gives the expression above.

Putting everything together:
$$\Gamma^r_{rr} = \frac{1}{2}\left(1 - \tfrac{2GM}{r}\right)\left[-\left(1 - \tfrac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2}\right].$$

We can factor out the terms carefully:
$$\Gamma^r_{rr} = -\frac{1}{2}\left(1 - \tfrac{2GM}{r}\right)\left(1 - \tfrac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2} = -\frac{GM}{r^2}\left(1 - \tfrac{2GM}{r}\right)^{-1}.$$

The negative sign indicates that a free-falling object would accelerate in the negative r-direction, i.e. towards smaller r, which aligns with the idea of a gravitational pull inward. This Christoffel symbol is crucial in determining the radial part of the geodesic equation.

(iv) Calculation of $\Gamma^\phi_{r\phi}$. We set $\rho = \phi$, $\mu = r$, and $\nu = \phi$. From the definition, we have:
$$\Gamma^\phi_{r\phi} = \frac{1}{2}g^{\phi\sigma}\left(\partial_r g_{\sigma\phi} + \partial_\phi g_{\sigma r} - \partial_\sigma g_{r\phi}\right).$$

Since $g^{\phi\sigma}$ is nonzero only for $\sigma = \phi$, and since $g_{r\phi} = 0$ (there are no cross terms in the metric between r and ϕ), it follows that
$$\Gamma^\phi_{r\phi} = \frac{1}{2}g^{\phi\phi}\partial_r g_{\phi\phi}.$$

This simplification relies on the fact that the only relevant term is the derivative of $g_{\phi\phi}$ with respect to r. Again, cross terms do not appear in the Schwarzschild metric, keeping the expression neat.

Recall that
$$g_{\phi\phi} = r^2 \sin^2\theta,$$
and its derivative with respect to r is
$$\partial_r(r^2 \sin^2\theta) = 2r \sin^2\theta.$$

We also have
$$g^{\phi\phi} = \frac{1}{r^2 \sin^2\theta}.$$

Thus, $g_{\phi\phi}$ grows like r^2, reflecting the area of a sphere in spherical coordinates. The factor $\sin^2\theta$ ensures that the metric reduces properly on different slices of constant θ.

Putting these together:
$$\Gamma^\phi_{r\phi} = \frac{1}{2}\left(\frac{1}{r^2 \sin^2\theta}\right)(2r \sin^2\theta) = \frac{1}{r}.$$

This $1/r$ dependence is typical for spherical coordinates and reflects the fact that when you move radially outward, the circle of radius r (in the θ, ϕ plane) expands. The Christoffel symbol tells us precisely how the basis vectors for the ϕ direction change with r.

(v) Final Results. We conclude that the two Christoffel symbols requested in this problem are:
$$\Gamma^r_{rr} = -\frac{GM}{r^2}\left(1 - \tfrac{2GM}{r}\right)^{-1}, \qquad \Gamma^\phi_{r\phi} = \frac{1}{r}.$$

These components play an important role in understanding geodesics in Schwarzschild spacetime. The Γ^r_{rr} component affects the radial part of a freely falling particle's motion, while $\Gamma^\phi_{r\phi}$ influences the azimuthal coordinate ϕ when changing r. In short, they are key ingredients in describing the geometry around a non-rotating, spherically symmetric mass. They encapsulate how local coordinate directions "twist" and "turn" in a curved geometry, ensuring that we can write down the correct equations of motion in General Relativity.

Question 11

This exercise focuses on the orbits of massive particles in a Schwarzschild geometry, which describes the spacetime around a spherically symmetric and non-rotating object, such as a star or a black hole. In particular, the objective is to determine the minimum radius for a stable circular orbit.
For geodesic motion in the Schwarzschild metric, the following two quantities are conserved:

$$E = \left(1 - \frac{2GM}{r}\right)\frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda}.$$

These two conserved quantities, E and L, derive from the symmetries of the Schwarzschild metric. They are associated with the Killing vectors related to time-translation invariance and axisymmetry, respectively. E is the energy per unit mass of the particle, conserved due to the invariance under time translations (the metric does not explicitly depend on time t). L is the angular momentum per unit mass, conserved due to the spherical symmetry (the metric does not depend on the angle φ). λ is an affine parameter that parametrizes the geodesic. Determine the smallest allowed radius for a stable circular orbit for a massive object.

Solution

(i) Schwarzschild Metric and Conserved Quantities *We begin by writing the Schwarzschild metric and the conserved quantities associated with the motion of a massive particle in this metric.*
We consider the Schwarzschild metric:

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

This is the Schwarzschild metric in coordinates (t, r, θ, ϕ), where t is the coordinate time, r is the radial coordinate, and θ and φ are the angular coordinates. M is the mass of the central object (like a black hole or a star), and G is the universal gravitational constant. Because of the symmetries of this spacetime (time-translation invariance and spherical symmetry), two conserved quantities naturally arise for geodesic motion:

$$E = \left(1 - \frac{2GM}{r}\right)\frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda}.$$

In our case of a massive particle, we can use $d\tau$ instead of $d\lambda$, hence we get:

$$E = \left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau}, \quad L = r^2 \frac{d\phi}{d\tau}.$$

Here, E can be interpreted as the energy per unit mass of the particle, while L is the angular momentum per unit mass. *E is the energy per unit mass of the particle, conserved because of the invariance under time translations (the metric does not explicitly depend on time t). L is the angular momentum per unit mass, conserved because of the spherical symmetry (the metric does not depend on the angle φ). τ is the proper time that parametrizes the geodesic.*

(ii) Normalization Condition and Effective Potential *For a massive particle, its four-velocity $u^\mu = \frac{dx^\mu}{d\tau}$ is normalized. Using this normalization condition and the conserved quantities, we can define an effective potential that describes the radial motion of the particle.*
Theoretical Setup:
In General Relativity, the motion of a test particle is described by the geodesic equation, which dictates that the particle follows the shortest path (geodesic) through curved spacetime. To analyze this motion, we need tools to describe the particle's trajectory and the forces acting on it. This is where the four-velocity and its normalization condition come into play.

Four-Velocity The four-velocity is a four-vector that generalizes the concept of velocity to four-dimensional spacetime. For a massive particle, we use the proper time τ as the affine parameter. Proper time is the time measured by a clock moving along with the particle. The four-velocity is defined as:

$$u^\mu = \frac{dx^\mu}{d\tau}$$

where $x^\mu = (t, r, \theta, \phi)$ are the spacetime coordinates in Schwarzschild spacetime. Each component of the four-velocity represents the rate of change of the corresponding coordinate with respect to proper time.
Why use the four-velocity?
Covariance: The four-velocity is a four-vector, which means it transforms covariantly under coordinate transformations. This ensures that the equations of motion are independent of the chosen coordinate system.
Relativistic Invariance: The four-velocity incorporates both the particle's spatial velocity and the effects of time dilation, making

it a suitable quantity to describe motion in a relativistic context.

Normalization: The four-velocity has a specific normalization condition that provides a crucial constraint on the particle's motion, as explained below.

Normalization Condition   The four-velocity of a massive particle is normalized such that its scalar product with itself is equal to -1 (using the metric signature -, +, +, +). This arises from the definition of proper time, which is the time measured by a clock moving along with the particle. The proper time interval $d\tau$ is related to the spacetime interval ds by:

$$ds^2 = -d\tau^2$$

where we are using units where $c = 1$.

In four-vector notation, the spacetime interval is given by:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

Dividing both sides by $d\tau^2$, we get:

$$\frac{ds^2}{d\tau^2} = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = g_{\mu\nu}u^\mu u^\nu$$

Since $ds^2 = -d\tau^2$ (with $c = 1$), we have $\frac{ds^2}{d\tau^2} = -1$. Therefore, the normalization condition is:

$$g_{\mu\nu}u^\mu u^\nu = -1.$$

Why is this important?

Physical Constraint: This condition reflects the fact that massive particles travel along timelike geodesics, meaning their four-velocity is always timelike.

Connection to Energy: The normalization condition, combined with the conserved quantities (energy and angular momentum), allows us to derive an equation of motion for the radial coordinate that resembles the energy conservation equation in classical mechanics. This enables us to define an effective potential.

Equatorial Plane and Simplification   We further simplify the problem by restricting the motion to the equatorial plane, setting $\theta = \frac{\pi}{2}$.

Why can we do this?

Spherical Symmetry: The Schwarzschild spacetime is spherically symmetric. This means that if a particle starts moving in a particular plane, there are no forces to push it out of that plane. It will continue to move in that plane indefinitely. *Coordinate Choice:* Setting $\theta = \frac{\pi}{2}$ is simply a choice of coordinates that aligns our coordinate system with the plane of motion. This simplifies the equations without loss of generality.

With $d\theta = 0$ (since $\theta = \frac{\pi}{2}$ is constant) and using the definition $u^\mu = \frac{dx^\mu}{d\tau}$, the normalization condition becomes:

$$g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -1.$$

Substituting the components of the Schwarzschild metric and expanding the summation, we have:

$$g_{tt}\left(\frac{dt}{d\tau}\right)^2 + g_{rr}\left(\frac{dr}{d\tau}\right)^2 + g_{\phi\phi}\left(\frac{d\phi}{d\tau}\right)^2 = -1.$$

Plugging in the explicit form of the Schwarzschild metric components for the equatorial plane:

$$-\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\tau}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{d\phi}{d\tau}\right)^2 = -1.$$

Conserved Quantities and Effective Potential   We have the conserved quantities E and L defined as:

$$E = \left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau}, \quad L = r^2\frac{d\phi}{d\tau}.$$

These quantities are conserved due to the symmetries of the Schwarzschild spacetime (timelike and rotational Killing vectors). From these, we can isolate $\frac{dt}{d\tau}$ and $\frac{d\phi}{d\tau}$ as follows:

$$\frac{dt}{d\tau} = \frac{E}{1 - \frac{2GM}{r}}, \quad \frac{d\phi}{d\tau} = \frac{L}{r^2}.$$

Substituting these expressions into the normalization equation gives:

$$-\left(1 - \frac{2GM}{r}\right)\left(\frac{E}{1 - \frac{2GM}{r}}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{L}{r^2}\right)^2 = -1.$$

Simplifying the equation:

$$-\frac{E^2}{1 - \frac{2GM}{r}} + \left(1 - \frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2} = -1.$$

Multiply through by $\left(1 - \frac{2GM}{r}\right)$ to get rid of the denominators:

$$-E^2 + \left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2}\left(1 - \frac{2GM}{r}\right) = -\left(1 - \frac{2GM}{r}\right).$$

Expanding the last term on the left side:

$$-E^2 + \left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2} - \frac{2GML^2}{r^3} = -1 + \frac{2GM}{r}.$$

Rearranging to isolate $\left(\frac{dr}{d\tau}\right)^2$:

$$\left(\frac{dr}{d\tau}\right)^2 = E^2 - 1 + \frac{2GM}{r} - \frac{L^2}{r^2} + \frac{2GML^2}{r^3}.$$

Now, multiply the equation by $\frac{1}{2}$ and collect terms to match the form $\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + V_{\text{eff}}(r) = \frac{E^2}{2}$:

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 = \frac{E^2}{2} - \frac{1}{2} + \frac{GM}{r} - \frac{L^2}{2r^2} + \frac{GML^2}{r^3}.$$

Rearranging to isolate $\frac{E^2}{2}$ on the right side:

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + \left(-\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} + \frac{1}{2}\right) = \frac{E^2}{2}.$$

Comparing with the desired form, we identify the effective potential $V_{\text{eff}}(r)$ as:

$$V_{\text{eff}}(r) = -\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} + \frac{1}{2}.$$

This equation has the same form as the energy conservation equation in classical mechanics: Kinetic Energy + Potential Energy = Total Energy.

Here, $\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2$ represents the "kinetic energy" associated with radial motion, $V_{\text{eff}}(r)$ is the "effective potential energy", and $\frac{E^2}{2}$ is the "total energy".

This can also be written as:

$$V_{\text{eff}}(r) = \left(1 - \frac{2GM}{r}\right)\left(\frac{1}{2} + \frac{L^2}{2r^2}\right).$$

This equation has a form similar to that of energy in classical mechanics, where the first term represents the radial kinetic energy and the second term is the effective potential. $V_{\text{eff}}(r)$ is a function of the radial coordinate r and depends on the parameters G, M, and L. The effective potential combines the effects of gravitational attraction, centrifugal repulsion, and a relativistic correction. Expanding,

$$V_{\text{eff}}(r) = \frac{1}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3}.$$

Expanding the expression for the effective potential, we obtain a more explicit form that allows us to analyze its behavior. The first term is a constant, the second term is the Newtonian term (attractive), the third term is a centrifugal term (repulsive), and the fourth term is an additional term due to General Relativity (attractive). Intuitively, $V_{\text{eff}}(r)$ encodes the combined effects of the gravitational attraction (terms with GM/r) and the centrifugal repulsion (terms with L^2). Intuitively, $V_{\text{eff}}(r)$ takes into account both the gravitational attraction (terms with GM/r) and the centrifugal repulsion (terms with L^2). The effective potential represents the effective potential energy per unit mass of the particle as a function of the radial coordinate r.

(iii) Circular Orbits *Now we analyze the conditions for having circular orbits. A circular orbit is characterized by a constant radius, which implies that the radial velocity is zero.*

To have a circular orbit at radius r, the radial coordinate must remain constant, implying

$$\frac{dr}{d\tau} = 0.$$

For a circular orbit, the radial coordinate r must remain constant. This means that the derivative of r with respect to the proper time τ must be zero. This means we sit at an extremum of the effective potential (since the motion in r-direction is momentarily at rest and remains so):

$$\frac{dV_{\text{eff}}}{dr} = 0.$$

The condition $dr/d\tau = 0$ implies that we are at an extremum point (maximum or minimum) of the effective potential. At an extremum point, the derivative of the effective potential with respect to r is zero. Computing the derivative of $V_{\text{eff}}(r)$:

$$\frac{dV_{\text{eff}}}{dr} = \frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3GML^2}{r^4}.$$

We calculate the derivative of the effective potential with respect to r. Setting this to zero gives the condition for circular orbits:

$$\frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3GML^2}{r^4} = 0.$$

Setting the derivative of the effective potential equal to zero, we obtain the condition for the existence of circular orbits. Multiplying by r^4 yields a quadratic equation in r:

$$GMr^2 - L^2r + 3GML^2 = 0.$$

Multiplying the previous equation by r^4, we obtain a quadratic equation in r, the solution of which will give us the radii of the possible circular orbits. This equation will generally give up to two real solutions for r, corresponding to (at most) two possible circular orbit radii for a given L. *This quadratic equation can have up to two real solutions for r, which correspond to two possible radii of circular orbits for a given value of the angular momentum L.* Solving for r we get:

$$r = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM}.$$

(iv) Smallest Stable Radius *We now determine the minimum radius for a stable circular orbit. This corresponds to the point where the two solutions of the quadratic equation for r coincide.*

Solving

$$GMr^2 - L^2r + 3GML^2 = 0$$

for r gives

$$r = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM}.$$

We solve the quadratic equation for r using the quadratic formula for second-degree equations. Real solutions exist only if the discriminant is non-negative:

$$L^4 - 12G^2M^2L^2 \geq 0 \quad \implies \quad L^2\left(L^2 - 12G^2M^2\right) \geq 0 \quad \implies \quad L^2 \geq 12G^2M^2.$$

Real solutions exist only if the discriminant of the quadratic equation is non-negative. This gives us a condition on the minimum value of the angular momentum L for the existence of circular orbits. At the threshold $L^2 = 12G^2M^2$, the two solutions for r coincide. Substituting $L^2 = 12G^2M^2$ back into r, we get

$$r = \frac{L^2 \pm 0}{2GM} = \frac{12G^2M^2 \pm 0}{2GM} = 6GM.$$

At the threshold value $L^2 = 12G^2M^2$, the two solutions for r coincide. Substituting this value of L into the expression for r, we obtain the minimum radius for a circular orbit. This particular circular orbit at $r = 6GM$ corresponds to the *innermost stable circular orbit* (ISCO) for a massive particle in Schwarzschild geometry. Below this radius, any circular orbit would be unstable or not allowed by the geometry. *This particular circular orbit at $r = 6GM$ corresponds to the innermost stable circular orbit (ISCO) for a massive particle in the Schwarzschild geometry. Below this radius, any circular orbit would be unstable or not allowed by the geometry.*

Therefore, the smallest allowed radius for a stable circular orbit is $r = 6GM$.

Question 12

For geodesic motion in the Schwarzschild metric, the following two quantities are conserved:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \qquad L = r^2 \frac{d\phi}{d\lambda}.$$

These are the conserved energy E and angular momentum L per unit mass of a test particle moving along a geodesic in the Schwarzschild spacetime. Here, G is the gravitational constant, M is the mass of the central object, r is the radial coordinate, t is the time coordinate, ϕ is the azimuthal angle, and λ is an affine parameter along the geodesic.

Consider the circular orbits for a massive object and determine the radius of the stable orbit in the $L \to \infty$ limit. *We are asked to find the radius of the stable circular orbit for a massive object in the limit of large angular momentum.*

Compare this result with the corresponding Newtonian result.

Solution

(i) Schwarzschild Metric and Conserved Quantities We start by recalling that the Schwarzschild metric is:

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

This is the Schwarzschild metric, which describes the geometry of spacetime outside a spherically symmetric mass distribution. The metric is expressed in Schwarzschild coordinates (t, r, θ, ϕ), where t is the time coordinate, r is the radial coordinate, θ is the polar angle, and ϕ is the azimuthal angle. The term $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$ represents the metric on the 2-sphere.

For a test particle of unit mass, two quantities remain constant along its geodesic:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \qquad L = r^2 \frac{d\phi}{d\lambda}.$$

These conserved quantities derive from the symmetries of the Schwarzschild metric. In particular, the metric does not depend on t (time translation invariance) nor on ϕ (rotational symmetry around the z-axis).

These constants follow from, respectively, time-translation invariance (energy conservation) and spherical symmetry (angular momentum conservation).

(ii) Normalization Condition and Effective Potential For a massive particle, the four-velocity $u^\mu = \frac{dx^\mu}{d\lambda}$ must satisfy

$$g_{\mu\nu}\, u^\mu u^\nu = -1.$$

This is the normalization condition for the four-velocity of a massive particle. It indicates that the proper time along the particle's trajectory is measured by λ such that the square of the four-velocity is -1.

Restricting the motion to the equatorial plane $(\theta = \pi/2)$, the normalization condition becomes:

$$-\left(1 - \tfrac{2GM}{r}\right)\left(\tfrac{dt}{d\lambda}\right)^2 + \left(1 - \tfrac{2GM}{r}\right)^{-1}\left(\tfrac{dr}{d\lambda}\right)^2 + r^2\left(\tfrac{d\phi}{d\lambda}\right)^2 = -1.$$

We consider motion confined to the equatorial plane for simplicity, which is possible due to the spherical symmetry of the Schwarzschild metric.

Using

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda},$$

we rewrite the above equation as

$$-\frac{E^2}{\left(1 - \tfrac{2GM}{r}\right)} + \frac{1}{\left(1 - \tfrac{2GM}{r}\right)}\left(\tfrac{dr}{d\lambda}\right)^2 + \frac{L^2}{r^2} = -1.$$

We have substituted the expressions for the conserved quantities E and L into the normalization condition.

Rearranging leads to a form where we identify the *effective potential $V_{\text{eff}}(r)$*:

$$\tfrac{1}{2}\left(\tfrac{dr}{d\lambda}\right)^2 + \left(1 - \tfrac{2GM}{r}\right)\left(\tfrac{1}{2} + \tfrac{L^2}{2r^2}\right) = \tfrac{E^2}{2}.$$

Hence,

$$V_{\text{eff}}(r) = \left(1 - \tfrac{2GM}{r}\right)\left(\tfrac{1}{2} + \tfrac{L^2}{2r^2}\right) = \tfrac{1}{2} - \tfrac{GM}{r} + \tfrac{L^2}{2r^2} - \tfrac{GML^2}{r^3}.$$

The equation now resembles that of a particle with "energy" $\tfrac{E^2}{2}$ moving in an effective potential $V_{\text{eff}}(r)$. The effective potential includes terms corresponding to Newtonian gravity, a centrifugal barrier, and a relativistic correction term.

(Observe that the first two terms reproduce a Newtonian-like potential, while the last two terms encode relativistic corrections.)

(iii) Circular Orbits A circular orbit occurs at constant r, implying $\frac{dr}{d\lambda} = 0$. At this radius, the effective potential must be at an extremum:

$$\frac{dV_{\text{eff}}}{dr} = 0.$$

For a circular orbit, the radial distance r does not change with the affine parameter λ. This occurs at a minimum (stable) or maximum (unstable) of the effective potential.

Differentiating $V_{\text{eff}}(r)$ with respect to r:

$$\frac{dV_{\text{eff}}}{dr} = \frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3\,GM\,L^2}{r^4}.$$

Setting this to zero:

$$\frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3\,GM\,L^2}{r^4} = 0,$$

which, upon multiplying by r^4, becomes the quadratic in r:

$$GM\,r^2 - L^2\,r + 3\,GM\,L^2 = 0.$$

This is a quadratic equation for the radius r of a circular orbit. The solutions of this equation provide the radii for which circular orbits are possible.

(iv) Circular Orbit Radius in the $L \to \infty$ Limit The general solutions of the quadratic equation for r are:

$$r = \frac{L^2 \pm \sqrt{L^4 - 12\,G^2 M^2 L^2}}{2\,GM}.$$

For large L, we can approximate the term under the square root using the binomial expansion. We factor out L^4 from the square root:

$$\sqrt{L^4 - 12\,G^2 M^2 L^2} = \sqrt{L^4\Big(1 - \frac{12\,G^2 M^2}{L^2}\Big)} = L^2\sqrt{1 - \frac{12\,G^2 M^2}{L^2}}.$$

In the limit of large angular momentum L, we can use a binomial approximation to simplify the square root.

Now we apply the binomial expansion for $\sqrt{1+x}$ with $x = -\frac{12\,G^2 M^2}{L^2}$. The binomial expansion is given by:

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \dots$$

In our case, $n = \frac{1}{2}$ and $x = -\frac{12\,G^2 M^2}{L^2}$. Since L is large, x is small, and we can approximate the square root by taking the first two terms of the expansion:

$$\sqrt{1+x} \approx 1 + \frac{1}{2}x.$$

Thus,

$$\sqrt{1 - \frac{12\,G^2 M^2}{L^2}} \approx 1 + \frac{1}{2}\Big(-\frac{12\,G^2 M^2}{L^2}\Big) = 1 - \frac{6\,G^2 M^2}{L^2}.$$

Substituting this back into the expression for the square root, we get:

$$\sqrt{L^4 - 12\,G^2 M^2 L^2} \approx L^2\Big(1 - \frac{6\,G^2 M^2}{L^2}\Big).$$

Now we can substitute this back into the solutions for r:

$$r \approx \frac{L^2 \pm L^2\big(1 - \frac{6\,G^2 M^2}{L^2}\big)}{2\,GM} = \frac{L^2}{2GM} \pm \frac{L^2}{2GM}\Big(1 - \frac{6\,G^2 M^2}{L^2}\Big).$$

This yields two approximate solutions:

$$r_{(1)} \approx \frac{L^2}{2GM} + \frac{L^2}{2GM} - \frac{L^2}{2GM}\cdot\frac{6\,G^2 M^2}{L^2} = \frac{L^2}{GM} - 3\,GM,$$

$$r_{(2)} \approx \frac{L^2}{2GM} - \frac{L^2}{2GM} + \frac{L^2}{2GM}\cdot\frac{6\,G^2 M^2}{L^2} = 3\,GM.$$

The first solution, $r_{(1)} \approx \frac{L^2}{GM} - 3\,GM$, grows large for large L and describes stable circular orbits (the subleading constant term $-3GM$ becomes negligible for $L \to \infty$). *This solution corresponds to the stable circular orbit, and its radius increases as L increases.*

32

The second solution, $r_{(2)} \approx 3\,GM$, corresponds to an unstable circular orbit. *This solution corresponds to the unstable circular orbit at a fixed radius of $3GM$, independent of L.*

Therefore, in the limit $L \to \infty$, the stable circular orbit effectively satisfies

$$r \approx \frac{L^2}{GM}.$$

Detailed explanation:

We are interested in the stable orbit, which is given by $r_{(1)} \approx \frac{L^2}{GM} - 3\,GM$. As L approaches infinity, the term $\frac{L^2}{GM}$ grows without bound, while the term $3\,GM$ remains constant. Thus, the term $\frac{L^2}{GM}$ dominates the expression.

We can rewrite the expression for $r_{(1)}$ as:

$$r_{(1)} \approx \frac{L^2}{GM}\left(1 - \frac{3\,G^2 M^2}{L^2}\right).$$

As $L \to \infty$, the term $\frac{3\,G^2 M^2}{L^2}$ approaches zero. Therefore, the expression in the parenthesis approaches 1.

Hence, in the limit of large L, the radius of the stable circular orbit is approximately:

$$r \approx \frac{L^2}{GM}.$$

In the limit of large L, the radius of the stable circular orbit is approximately $\frac{L^2}{GM}$.

(v) Comparison with the Newtonian Result   In Newtonian gravity, the effective potential for a particle of mass m and angular momentum L in a central field GMm/r is

$$V_{\text{eff}}(r) = -\frac{GMm}{r} + \frac{L^2}{2\,m\,r^2}.$$

This is the Newtonian effective potential, which includes the gravitational potential energy and the centrifugal potential energy.

The circular-orbit condition $\frac{dV_{\text{eff}}}{dr} = 0$ gives

$$\frac{GMm}{r^2} - \frac{L^2}{m\,r^3} = 0 \quad \Longrightarrow \quad r = \frac{L^2}{GM\,m^2}.$$

If $m = 1$ (taking unit mass), we get the same leading-order form

$$r_{\text{Newton}} = \frac{L^2}{GM}.$$

Hence, at large L, the relativistic orbit radius $r \approx \frac{L^2}{GM}$ matches the Newtonian prediction up to relativistic corrections. *In the limit of large angular momentum, the radius of the stable circular orbit in general relativity approaches the Newtonian result.*

The difference is that in Schwarzschild geometry we treat a test particle (mass m is effectively canceled out), whereas in the Newtonian context r explicitly depends on m. In any case, the leading $1/GM$ scaling with L^2 is the same in both theories for large angular momentum.

Exercise 13

For geodesic motion in the Schwarzschild metric, the following two quantities are conserved:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \qquad L = r^2 \frac{d\phi}{d\lambda}.$$

Discuss all possible circular orbits for a massless object and determine their stability.

Solution

(i) Schwarzschild Metric and Conserved Quantities. Consider the Schwarzschild metric in (t, r, θ, ϕ) coordinates:

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

This is the Schwarzschild metric, describing the geometry of spacetime outside a spherically symmetric mass distribution. The metric is given in Schwarzschild coordinates (t, r, θ, ϕ), where t is the time coordinate, r is the radial coordinate, θ is the polar angle, and ϕ is the azimuthal angle. The term $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$ represents the metric on the 2-sphere.

Two quantities remain constant along any geodesic:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \tag{4}$$

$$L = r^2 \frac{d\phi}{d\lambda}. \tag{5}$$

These conserved quantities arise from the symmetries of the Schwarzschild metric. Specifically, the metric does not depend on t (time translation invariance) or ϕ (rotational symmetry about the z-axis). E represents the conserved energy and L the conserved angular momentum per unit mass of a test particle moving along a geodesic in the Schwarzschild spacetime. Here, G is the gravitational constant, M is the mass of the central object, r is the radial coordinate, t is the time coordinate, ϕ is the azimuthal angle, and λ is an affine parameter along the geodesic. Here, λ is an affine parameter along the geodesic.

(ii) Normalization Condition for a Massless Particle. For a massless object (e.g., a photon), its four-momentum $p^\mu = \frac{dx^\mu}{d\lambda}$ satisfies

$$g_{\mu\nu}\, p^\mu p^\nu = 0.$$

This is the normalization condition for the four-momentum of a massless particle. Unlike the massive case, where the condition is $g_{\mu\nu} u^\mu u^\nu = -1$, for massless particles, the condition is $g_{\mu\nu} p^\mu p^\nu = 0$. This reflects the fact that massless particles move at the speed of light, and their proper time cannot be used as an affine parameter. Instead, we use an affine parameter λ that is related to the coordinate time t and radial coordinate r through the conserved quantities E and L.

We can restrict the motion to the equatorial plane $\theta = \pi/2$, which simplifies the metric by eliminating the $d\theta^2$ term. *We consider the motion confined to the equatorial plane for simplicity, which is possible due to the spherical symmetry of the Schwarzschild metric. Setting $\theta = \pi/2$ implies $d\theta = 0$ and $\sin^2\theta = 1$. Thus,*

$$-\left(1 - \tfrac{2GM}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 + \left(1 - \tfrac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2 + r^2\left(\frac{d\phi}{d\lambda}\right)^2 = 0.$$

Substitute the conserved quantities $\frac{dt}{d\lambda} = \frac{E}{1 - \frac{2GM}{r}}$ and $\frac{d\phi}{d\lambda} = \frac{L}{r^2}$:

$$-\frac{E^2}{1 - \frac{2GM}{r}} + \left(1 - \tfrac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2 + \frac{L^2}{r^2} = 0.$$

We have substituted the expressions for the conserved quantities E and L into the normalization condition.
Rearrange to isolate $\left(\frac{dr}{d\lambda}\right)^2$:

$$\left(\frac{dr}{d\lambda}\right)^2 + \frac{L^2}{r^2}\left(1 - \tfrac{2GM}{r}\right) = E^2.$$

Define the effective potential for the massless particle:

$$V_{\text{eff}}(r) = \frac{L^2}{2r^2}\left(1 - \tfrac{2GM}{r}\right).$$

Then, our radial equation is $\left(\frac{dr}{d\lambda}\right)^2 + 2V_{\text{eff}}(r) = E^2$. *The equation now resembles that of a particle with "energy" E^2 moving in an effective potential $V_{\text{eff}}(r)$. The effective potential includes terms corresponding to a centrifugal barrier and a relativistic correction term. Unlike the massive case, there is no Newtonian-like term because massless particles do not experience Newtonian gravity.*

(iii) Circular Orbits for a Massless Object. A circular orbit requires a constant radius, hence $\frac{dr}{d\lambda} = 0$. In terms of the effective potential, this implies

$$E^2 = 2\,V_{\text{eff}}(r), \quad \text{and} \quad \frac{dV_{\text{eff}}}{dr} = 0.$$

For a circular orbit, the radial distance r does not change with the affine parameter λ. This occurs at an extremum (minimum or maximum) of the effective potential.

Differentiate $V_{\text{eff}}(r)$:

$$V_{\text{eff}}(r) = \frac{L^2}{2\,r^2}\left(1 - \frac{2GM}{r}\right).$$

$$\frac{dV_{\text{eff}}}{dr} = -\frac{L^2}{r^3}\left(1 - \frac{2GM}{r}\right) + \frac{L^2}{r^2}\frac{2GM}{r^2} = -\frac{L^2}{r^3} + \frac{3\,GM\,L^2}{r^4}.$$

Setting $\frac{dV_{\text{eff}}}{dr} = 0$ yields

$$-\frac{L^2}{r^3} + \frac{3\,GM\,L^2}{r^4} = 0 \quad \Longrightarrow \quad r = 3\,GM.$$

Thus, for a massless particle (photon), the only possible circular orbit is at $r = 3GM$. This is the well-known photon sphere. *This result shows that there is only one possible radius for a circular orbit for a massless particle in the Schwarzschild spacetime, and it is located at $r = 3GM$. This is known as the photon sphere.*

(iv) Stability of the Circular Orbit. To test stability, we look at the second derivative of V_{eff}. An extremum is stable if $\frac{d^2V_{\text{eff}}}{dr^2} > 0$ (minimum) and unstable if $\frac{d^2V_{\text{eff}}}{dr^2} < 0$ (maximum). We compute:

$$\frac{d^2V_{\text{eff}}}{dr^2} = \frac{3\,L^2}{r^4} - \frac{12\,GM\,L^2}{r^5}.$$

Evaluating at $r = 3GM$:

$$\left.\frac{d^2V_{\text{eff}}}{dr^2}\right|_{r=3GM} = \frac{3\,L^2}{(3GM)^4} - \frac{12\,GM\,L^2}{(3GM)^5} = \frac{3\,L^2}{81\,G^4M^4} - \frac{12\,L^2}{243\,G^4M^4}$$

$$= \frac{3\,L^2}{81\,G^4M^4} - \frac{4\,L^2}{81\,G^4M^4} = -\frac{L^2}{81\,G^4M^4} < 0. \tag{6}$$

Since this is negative, the orbit at $r = 3GM$ is a maximum of the effective potential and thus unstable. A small perturbation in r will drive the photon either inwards toward the black hole horizon at $r = 2GM$, or outwards to larger r. *The negative value of the second derivative of the effective potential at $r = 3GM$ indicates that this circular orbit is unstable. Any slight perturbation will cause the photon to either spiral inward towards the event horizon or escape to infinity.*

(v) Final Remarks. For a massless object (e.g. a photon) in Schwarzschild spacetime, there is exactly one circular orbit at $r = 3GM$. However, it is an unstable orbit: any tiny radial perturbation will send the particle plunging inward or escaping outward. This purely relativistic effect has no Newtonian analogue, since massless particles do not orbit under Newtonian gravity. *In conclusion, the photon sphere at $r = 3GM$ is a unique feature of the Schwarzschild spacetime. It is the only radius at which massless particles can orbit, but this orbit is unstable. This is a purely relativistic phenomenon with no counterpart in Newtonian gravity.*

Question 14

For geodesic motion in the Schwarzschild metric, the following two quantities are conserved:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda}.$$

These are the conserved energy E and angular momentum L per unit mass of a test particle moving along a geodesic in the Schwarzschild spacetime. Here, G is the gravitational constant, M is the mass of the central object, r is the radial coordinate, t is the time coordinate, ϕ is the azimuthal angle, and λ is an affine parameter along the geodesic. The conservation of these quantities is a consequence of the time translation invariance and rotational symmetry of the Schwarzschild metric.

In this problem, we set our units such that $GM = 1$ and consider the motion of a photon with angular momentum $L = 10$. We are choosing a system of units in which the product of the gravitational constant G and the mass of the central object M is 1. This simplifies the equations. We are considering a photon, which is a massless particle, with an angular momentum $L = 10$.

Find the minimum value of E such that the photon can reach the singularity at $r \to 0$.

Solution

(i) Schwarzschild Metric and Conserved Quantities   The Schwarzschild metric in coordinates $\{t, r, \theta, \phi\}$ is given by:

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

This is the Schwarzschild metric, which describes the geometry of spacetime outside a spherically symmetric mass distribution. The metric is expressed in Schwarzschild coordinates (t, r, θ, ϕ), where t is the time coordinate, r is the radial coordinate, θ is the polar angle, and ϕ is the azimuthal angle. The term $d\Omega^2 = d\theta^2 + \sin^2\theta \, d\phi^2$ represents the metric on a 2-sphere.

Since the metric is stationary (time translation symmetry) and spherically symmetric, two quantities are conserved along geodesics:

$$E = \left(1 - \tfrac{2GM}{r}\right)\frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda}.$$

The conserved quantity E is associated with the time translation symmetry of the metric (it does not depend on t), while L is associated with the rotational symmetry about the z-axis (the metric does not depend on ϕ).

Here, λ is an affine parameter along the photon's path.

(ii) Normalization Condition and Effective Potential   For a photon (massless particle), the four-momentum $p^\mu = \frac{dx^\mu}{d\lambda}$ satisfies

$$g_{\mu\nu} p^\mu p^\nu = 0.$$

This is the normalization condition for the four-momentum of a massless particle (photon). It expresses that the four-momentum is a null vector, i.e., its squared magnitude is zero.

Assume motion is confined to the equatorial plane ($\theta = \pi/2$), which is always possible by spherical symmetry. Then the condition becomes:

$$-\left(1 - \tfrac{2GM}{r}\right)\left(\tfrac{dt}{d\lambda}\right)^2 + \left(1 - \tfrac{2GM}{r}\right)^{-1}\left(\tfrac{dr}{d\lambda}\right)^2 + r^2\left(\tfrac{d\phi}{d\lambda}\right)^2 = 0.$$

We can choose the equatorial plane without loss of generality due to the spherical symmetry of the Schwarzschild metric. This simplifies the equations by setting $d\theta/d\lambda = 0$ and $\sin\theta = 1$.

Substitute the conserved quantities E and L to rewrite it in terms of $dr/d\lambda$:

$$-\frac{E^2}{1 - \tfrac{2GM}{r}} + \frac{1}{1 - \tfrac{2GM}{r}}\left(\tfrac{dr}{d\lambda}\right)^2 + \frac{L^2}{r^2} = 0.$$

We have substituted E and L into the normalization condition to eliminate $dt/d\lambda$ and $d\phi/d\lambda$.

Rearrange this into:

$$\tfrac{1}{2}\left(\tfrac{dr}{d\lambda}\right)^2 + \frac{L^2}{2\,r^2}\left(1 - \tfrac{2GM}{r}\right) = \frac{E^2}{2}.$$

We identify the *effective potential* as

$$V_{\text{eff}}(r) = \frac{L^2}{2\,r^2}\left(1 - \tfrac{2GM}{r}\right).$$

This equation has the form of an energy conservation equation, where the first term represents the "kinetic energy" and the second term the "potential energy". The effective potential $V_{\text{eff}}(r)$ determines the radial motion of the photon.

A photon can reach the singularity at $r = 0$ only if its total "energy" (in the effective potential sense) is large enough to overcome any potential barrier.

(iii) Setting $GM = 1$ and $L = 10$ We now set $GM = 1$. The angular momentum is given as $L = 10$.   Hence,

$$V_{\text{eff}}(r) = \frac{L^2}{2\,r^2}\left(1 - \frac{2}{r}\right) = \frac{100}{2\,r^2} - \frac{100}{r^3} = \frac{50}{r^2} - \frac{100}{r^3}.$$

We have substituted $GM = 1$ and $L = 10$ into the expression for the effective potential.

The radial equation becomes

$$\frac{1}{2}\left(\frac{dr}{d\lambda}\right)^2 + V_{\text{eff}}(r) = \frac{E^2}{2}.$$

This equation describes the radial motion of the photon in terms of the effective potential and the conserved energy E.

(iv) Condition for the Photon to Reach the Singularity   To find whether the photon can fall into $r = 0$, we look for the largest barrier in $V_{\text{eff}}(r)$. The maximum of V_{eff} occurs where

$$\frac{dV_{\text{eff}}}{dr} = -\frac{100}{r^3} + \frac{300}{r^4} = 0.$$

To find the maximum value of the effective potential, we calculate its derivative with respect to r and set it equal to zero.

Solving yields $r_{\text{max}} = 3$. Solving the equation $\frac{dV_{\text{eff}}}{dr} = 0$ gives $r_{\text{max}} = 3$ (in units where $GM = 1$).

At this point,

$$V_{\text{eff}}(3) = \frac{50}{3^2} - \frac{100}{3^3} = \frac{50}{9} - \frac{100}{27} = \frac{150 - 100}{27} = \frac{50}{27}.$$

We calculate the effective potential at $r_{\text{max}} = 3$ to find the height of the potential barrier.

The photon must have

$$\frac{E^2}{2} \geq V_{\text{eff}}(3), \quad \text{i.e.} \quad E^2 \geq 2\frac{50}{27} = \frac{100}{27}.$$

In order for the photon to reach the singularity at $r = 0$, its energy squared must be greater than or equal to the maximum value of the effective potential.

Thus,

$$E \geq \sqrt{\frac{100}{27}} = \frac{10}{\sqrt{27}} = \frac{10}{3\sqrt{3}}.$$

Hence the minimum energy for the photon to reach $r = 0$ is

$$E_{\text{min}} = \frac{10}{3\sqrt{3}}.$$

We find the minimum energy E_{min} by taking the square root of the minimum value of E^2.

(v) Additional Comments

- The radius $r = 3GM$ (or $r = 3$ in our units where $GM = 1$) is a special location in the Schwarzschild spacetime known as the photon sphere. At this radius, a photon with precisely the right energy and angular momentum can orbit the central mass in an unstable circular orbit. This is a purely relativistic effect with no Newtonian counterpart.

  To understand why this orbit is unstable, consider the effective potential

$$V_{\text{eff}}(r) = \frac{L^2}{2\,r^2}\left(1 - \frac{2GM}{r}\right).$$

  At the photon sphere ($r = 3GM$), the first derivative of the effective potential vanishes, $\frac{dV_{\text{eff}}}{dr}\big|_{r=3GM} = 0$, indicating a circular orbit. To check the stability, we need to compute the second derivative:

$$\frac{d^2V_{\text{eff}}}{dr^2} = \frac{d}{dr}\left(-\frac{L^2}{r^3} + \frac{3GML^2}{r^4}\right)$$

$$= \frac{3L^2}{r^4} - \frac{12GML^2}{r^5}$$

$$= \frac{L^2}{r^4}\left(3 - \frac{12GM}{r}\right).$$

  Evaluating this at $r = 3GM$, we get

$$\frac{d^2V_{\text{eff}}}{dr^2}\bigg|_{r=3GM} = \frac{L^2}{(3GM)^4}(3-4) = -\frac{L^2}{81(GM)^4} < 0.$$

37

The negative sign of the second derivative means that the effective potential has a maximum at $r = 3GM$. Therefore, the circular orbit at the photon sphere is unstable. A slight inward or outward radial perturbation will cause the photon to either spiral into the black hole or escape to infinity, respectively.

Recall that in the lectures, we discussed how the existence and properties of circular orbits are related to the shape of the effective potential. In particular, stable circular orbits correspond to minima of V_{eff}, while unstable ones correspond to maxima. The photon sphere is an example of the latter.

- When the photon's energy E is lower than the threshold value found in part (iv), i.e., $E < E_{\min} = \frac{10}{3\sqrt{3}}$ (in units where $GM = 1$ and $L = 10$), the photon will encounter a turning point at some radius $r_t > 2GM$ where its "radial kinetic energy" vanishes:

$$\frac{1}{2}\left(\frac{dr}{d\lambda}\right)^2 = \frac{E^2}{2} - V_{\text{eff}}(r) = 0.$$

At this point, the photon's radial velocity changes sign, and it cannot proceed further inwards. Instead, it will turn around and move back to larger radii, eventually escaping to infinity. The turning point can be found by solving the equation $E^2/2 = V_{\text{eff}}(r)$ for r.

This behavior is analogous to that of a classical particle moving in a potential with a barrier. If the particle's energy is less than the barrier height, it cannot cross the barrier and will be reflected back.

- The existence of a potential barrier for massless particles (photons) in the Schwarzschild spacetime is a remarkable feature of general relativity. In Newtonian gravity, there is no such barrier, and the path of light is unaffected by gravity (in the geometric optics approximation). The effective potential for photons in the Schwarzschild metric,

$$V_{\text{eff}}(r) \;=\; \frac{L^2}{2\,r^2}\left(1 - \frac{2GM}{r}\right),$$

arises from the curvature of spacetime around the central mass. The $1/r^3$ term, which is responsible for the barrier, has no Newtonian analogue. It represents a purely relativistic correction to the centrifugal potential $L^2/2r^2$.

This potential barrier is what allows for phenomena like the photon sphere and the deflection of light by massive objects (gravitational lensing), which we studied in the lectures. The fact that even massless particles are affected by gravity in this nontrivial way is a consequence of the fundamental principle of general relativity that all forms of energy and momentum contribute to the curvature of spacetime.

- In the full Schwarzschild geometry, the radius $r = 2GM$ defines the event horizon of the black hole, a surface of no return. Any photon that crosses the event horizon is trapped inside and cannot escape to infinity. In our linearized treatment, we have assumed weak gravity ($r > 2GM$) and a static situation, so this effect is not accurately captured. Nevertheless, the potential barrier and the photon sphere are real features that exist even in the full non-linear theory.

These additional comments aim to provide a deeper understanding of the physical and theoretical implications of the result obtained in the exercise, connecting it to the broader concepts discussed in the lectures.

Question 15

Consider the linear regime where $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $|h_{\mu\nu}| \ll 1$. First, define the vector $V_\mu = \partial_\nu h_\mu^\nu - \frac{1}{2}\partial_\mu h_\nu^\nu$. Consider the gauge transformation described by the function $\xi_\alpha(x^\mu)$. (i) What happens to the vector V_μ after such a gauge transformation? (ii) Is it possible to find a gauge where V_μ is vanishing?

Solution

(i) Definition of the Vector V_μ   We define the vector V_μ as:

$$V_\mu = \partial_\nu h_\mu^\nu - \frac{1}{2}\partial_\mu h_\nu^\nu \tag{7}$$

where $h_\nu^\nu = h = g^{\nu\alpha}h_{\alpha\nu} \approx \eta^{\nu\alpha}h_{\nu\alpha}$ represents the trace of $h_{\mu\nu}$.

(ii) Gauge Transformations   We consider an infinitesimal gauge transformation, which corresponds to a small coordinate transformation, described by the vector field $\xi_\alpha(x^\mu)$:

$$x^\mu \to x^\mu + \xi^\mu \tag{8}$$

Under this transformation, the metric perturbation $h_{\mu\nu}$ transforms as:

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \tag{9}$$

where $\xi_\mu = g_{\mu\alpha}\xi^\alpha \approx \eta_{\mu\alpha}\xi^\alpha$.

Physical explanation: These gauge transformations in linearized gravity are analogous to gauge transformations in electromagnetism. In electromagnetism, the potential vector A_μ can be changed by adding the gradient of a scalar function Λ without altering the physical electromagnetic field (determined by the Faraday tensor $F_{\mu\nu}$). Similarly, in linearized gravity, we can change $h_{\mu\nu}$ without altering the linearized Riemann tensor, which describes the curvature of spacetime. This means that the physics described by the metric remains invariant under these transformations.

(iii) Detailed derivation with indices ν, μ raised/lowered   We now provide a step-by-step, pedantic explanation of how the gauge transformation acts on the objects h_μ^ν and h_ν^ν.

1. Defining h_μ^ν and the trace h_ν^ν.   First, consider the metric of flat spacetime, $\eta_{\mu\nu}$, which is symmetric and invertible. In index notation, we may raise an index of $h_{\mu\nu}$ by contracting it with $\eta^{\nu\alpha}$. Explicitly:

$$h_\mu^\nu \;=\; \eta^{\nu\alpha}h_{\mu\alpha}.$$

Similarly, the *trace* is obtained by contracting h_μ^ν on its upper and lower indices:

$$h_\nu^\nu \;=\; \eta^{\nu\alpha}h_{\nu\alpha}.$$

From now on, we will **not** show $\eta^{\mu\nu}$ explicitly; we simply write "$h_{\mu\nu}$ with one index raised" as h_μ^ν, and "$h_{\mu\nu}$ fully contracted" as h_ν^ν.

2. Gauge transformations of h_μ^ν and h_ν^ν.   We impose the infinitesimal gauge transformation on $h_{\mu\nu}$:

$$h_{\mu\nu} \;\longrightarrow\; h_{\mu\nu} \;+\; \partial_\mu \xi_\nu \;+\; \partial_\nu \xi_\mu,$$

where ξ_μ is the small gauge vector. Once we have

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu,$$

we can "raise" one of its indices. By definition,

$$h_\mu^\nu \;=\; (\eta^{\nu\alpha}h_{\mu\alpha}),$$

so under the gauge transformation:

$$h_\mu^\nu \;\longrightarrow\; h_\mu^\nu \;+\; \partial_\mu \xi^\nu \;+\; \partial^\nu \xi_\mu,$$

where we silently use $\eta^{\nu\alpha}$ to convert $\xi_\alpha \to \xi^\nu$ and $\partial_\alpha \to \partial^\nu$.

Similarly, the trace

$$h_\nu^\nu \;=\; (\eta^{\nu\alpha}h_{\nu\alpha})$$

transforms as follows: 1. The old term h_ν^ν remains h_ν^ν.   2. Each of $\partial_\nu \xi^\nu$ and $\partial^\nu \xi_\nu$ appears upon contracting the new gauge piece; these two expressions coincide (thanks to symmetry and the fact that partial derivatives commute) and thus add up to $2\partial_\nu \xi^\nu$.

Hence,

$$h_\nu^\nu \;\longrightarrow\; h_\nu^\nu \;+\; 2\partial_\alpha \xi^\alpha.$$

3. The vector V_μ. Recall the definition:

$$V_\mu = \partial_\nu \left(h^\nu{}_\mu \right) - \tfrac{1}{2} \partial_\mu \left(h^\nu{}_\nu \right).$$

This object is constructed from the derivatives of $h^\nu{}_\mu$ and its trace $h^\nu{}_\nu$.

4. Gauge transformation of V_μ. We want to see how V_μ changes when

$$h^\nu{}_\mu \rightarrow h^\nu{}_\mu + \partial_\mu \xi^\nu + \partial^\nu \xi_\mu, \quad h^\nu{}_\nu \rightarrow h^\nu{}_\nu + 2 \partial_\alpha \xi^\alpha.$$

Substitute these into the definition of V_μ:

$$V_\mu \longrightarrow \partial_\nu \left(h^\nu{}_\mu + \partial_\mu \xi^\nu + \partial^\nu \xi_\mu \right) - \tfrac{1}{2} \partial_\mu \left(h^\nu{}_\nu + 2 \partial_\alpha \xi^\alpha \right)$$

$$= \underbrace{\partial_\nu \left(h^\nu{}_\mu \right)}_{\text{original } V_\mu \text{ piece}} + \partial_\nu \left(\partial_\mu \xi^\nu \right) + \partial_\nu \left(\partial^\nu \xi_\mu \right) - \tfrac{1}{2} \underbrace{\partial_\mu \left(h^\nu{}_\nu \right)}_{\text{original } V_\mu \text{ piece}} - \partial_\mu \left(\partial_\alpha \xi^\alpha \right). \tag{10}$$

In flat coordinates, partial derivatives commute, so combinations like $\partial_\nu \partial_\mu \xi^\nu - \partial_\mu \partial_\nu \xi^\nu$ cancel out. What remains is precisely

$$V_\mu + \Box \xi_\mu,$$

where $\Box = \partial^\alpha \partial_\alpha$ is the d'Alembertian operator (again, defined implicitly by raising an index on ∂_α).

5. Final result. Hence, under the gauge shift $h_{\mu\nu} \rightarrow h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$, the vector V_μ transforms as

$$\boxed{V_\mu \longrightarrow V_\mu + \Box \xi_\mu.}$$

This shows that V_μ is *not* gauge-invariant; however, by picking ξ_μ such that $\Box \xi_\mu = -V_\mu$, one can impose $V'_\mu = 0$ in the new gauge.

(iv) Existence of a Gauge where $V_\mu = 0$ Yes, it is always possible to find a gauge in which $V_\mu = 0$. To demonstrate this, we consider the transformation of V_μ:

$$V_\mu \rightarrow V_\mu + \Box \xi_\mu \tag{11}$$

If we choose ξ_μ such that it satisfies the following inhomogeneous wave equation:

$$\Box \xi_\mu = -\partial_t^2 \xi_\mu + \nabla^2 \xi_\mu = -V_\mu \tag{12}$$

then, in the new gauge, we will have:

$$V'_\mu = V_\mu + \Box \xi_\mu = V_\mu - V_\mu = 0 \tag{13}$$

This choice of gauge is known as the *Lorenz gauge* (or *harmonic gauge*) in linearized gravity.

Physical Explanation: The existence of a gauge in which $V_\mu = 0$ is a consequence of the redundancy in the description of the metric perturbation $h_{\mu\nu}$. This redundancy allows us to choose coordinates (gauge) that simplify the equations. The Lorenz gauge is particularly useful because it significantly simplifies the linearized Einstein equations, making them resemble wave equations.

This inhomogeneous wave equation admits a solution for ξ_μ, ensuring that we can always find a suitable gauge transformation to reach the Lorenz gauge. The specific solution will depend on the initial V_μ, but the existence of such a solution is guaranteed. This gauge choice does not uniquely define ξ_μ, as any ξ_μ satisfying the homogeneous wave equation $\Box \xi_\mu = 0$ can be added without changing the condition $V_\mu = 0$. This residual gauge freedom can be further exploited.

(v) Conclusion In conclusion, we have demonstrated that:

(i) Under an infinitesimal gauge transformation, the vector V_μ transforms as $V_\mu \rightarrow V_\mu + \Box \xi_\mu$.

(ii) It is always possible to choose a gauge (the Lorenz gauge) in which $V_\mu = 0$, by solving the equation $\Box \xi_\mu = -V_\mu$.

This solution leverages concepts and results derived in previous lectures, in particular the form of gauge transformations for the perturbed metric, the definition of V_μ, and the use of the d'Alembert operator.

Question 16

Consider a plane wave solution in the TT gauge $h_{\mu\nu}^{TT} = \text{Re}\left(C_{\mu\nu} \exp\left[ik_\sigma x^\sigma\right]\right)$ propagating along the positive direction of the z-axis.

(i) What are the explicit components of k^μ ?

(ii) What is the explicit form of $C_{\mu\nu}$ for the + polarization?

(iii) Take the plane wave solution for the + polarization and determine its form after a gauge transformation parameterized by the function $\xi^\mu = (\alpha(x^\mu), \beta(x^\mu), 0, 0)$ with $\alpha(x^\mu)$ and $\beta(x^\mu)$ generic functions of the spacetime coordinates.

(i) Explicit components of k^μ

We are given a plane wave solution propagating along the positive z-axis. We need to find the explicit components of the wave vector k^μ. The wave vector describes the direction and frequency of the wave. Since it's a gravitational wave, which is massless, it travels at the speed of light.
In general, for a plane wave, the wave vector k^μ can be written as:

$$k^\mu = (\omega, k_x, k_y, k_z) \tag{14}$$

where ω is the angular frequency and (k_x, k_y, k_z) is the spatial wave vector. The components represent the rate of change of phase with respect to time and spatial coordinates, respectively.
Since the wave is propagating along the positive z-axis, we have $k_x = k_y = 0$. This is because there is no change in phase along the x and y directions.
Also, for a gravitational wave (massless), the dispersion relation, which relates the frequency and the spatial wave vector, is:

$$k_\mu k^\mu = 0 \tag{15}$$

This relation comes from the fact that gravitational waves travel at the speed of light, implying a null wave vector. In other words, the four-vector is light-like. We can expand the scalar product

$$\eta_{\mu\nu} k^\mu k^\nu = -\left(k^0\right)^2 + \left(k^1\right)^2 + \left(k^2\right)^2 + \left(k^3\right)^2 = 0 \tag{16}$$

so, since $k^0 = \omega$, $k^1 = k_x = 0$, $k^2 = k_y = 0$

$$-\omega^2 + k_z^2 = 0 \implies k_z = \pm\omega \tag{17}$$

Since we are considering the positive z-axis, k_z is positive, so $k_z = \omega$.
Therefore, the explicit components of k^μ are:

$$k^\mu = (\omega, 0, 0, \omega) \tag{18}$$

with $\omega > 0$. This means the wave has a frequency ω and propagates in the positive z direction with a speed equal to the speed of light.

(ii) Explicit form of $C_{\mu\nu}$ for the + polarization

In this section, we aim to determine the explicit form of the polarization tensor $C_{\mu\nu}$ for a gravitational wave with "+" polarization in the TT gauge. The polarization tensor describes how the gravitational wave stretches and squeezes spacetime as it propagates. The TT gauge, or transverse-traceless gauge, is a specific coordinate choice that simplifies the description of gravitational waves.
Imagine spacetime as a fabric. A gravitational wave passing through this fabric will cause distortions. The polarization tensor $C_{\mu\nu}$ is a mathematical object that encodes the information about these distortions, telling us how the fabric is stretched and squeezed in different directions.
The TT gauge simplifies the description by ensuring that the gravitational wave only causes distortions that are:

1. *Transverse: Perpendicular to the direction of propagation.*

2. *Traceless: The distortions do not change the overall volume of spacetime, only its shape.*

Let's examine the two fundamental conditions of the TT gauge and how they apply to $C_{\mu\nu}$:
The polarization tensor $C_{\mu\nu}$ for a gravitational wave in the TT gauge must satisfy the following conditions:

1. Symmetry: $C_{\mu\nu} = C_{\nu\mu}$ This means that the tensor is symmetric, i.e., swapping the indices does not change its value. This is a general property of the metric perturbation tensor, which $C_{\mu\nu}$ is a part of. This symmetry will be important later when we analyze the components.

2. Transversality: $k^\mu C_{\mu\nu} = 0$ *This condition implies that the wave's polarization is perpendicular to its direction of propagation. In other words, if the wave is moving along the z-axis, the components of $C_{\mu\nu}$ that involve the z-direction or the time direction are related in a specific way, such that they don't contribute to any stretching or squeezing along the direction of propagation or along the time direction. The wave's effects are purely spatial and confined to the plane perpendicular to the propagation direction.* **It is crucial to understand that the transversality condition alone does not force the components $C_{0\nu}$ and $C_{3\nu}$ to be zero. Instead, it establishes a relationship between them.**

3. Tracelessness: $C_\mu^\mu = \eta^{\mu\nu} C_{\mu\nu} = 0$ *This condition means that the trace of the polarization tensor is zero. The trace is the sum of the diagonal elements. Physically, this ensures that there is no overall expansion or contraction of space associated with the wave; the wave only distorts the shape, not the volume.*

Given that $k^\mu = (\omega, 0, 0, \omega)$, which we found in part (i) and represents a wave propagating in the positive z-direction, the transversality condition implies:

$$k^\mu C_{\mu\nu} = 0 \tag{19}$$

$$\omega C_{0\nu} + \omega C_{3\nu} = 0 \implies C_{0\nu} = -C_{3\nu} \tag{20}$$

Since k^μ has only a time component and a z-component, when we contract it with $C_{\mu\nu}$, only the terms with $\mu = 0$ and $\mu = 3$ will survive. This leads to the conclusion that the time components of $C_{\mu\nu}$ are equal in magnitude and opposite in sign to the z-components. For example, if $\nu = 0$, we have $\omega C_{00} + \omega C_{30} = 0$, which simplifies to $C_{00} = -C_{30}$. Similarly, if $\nu = 1$, we get $C_{01} = -C_{31}$, and so on. **This demonstrates that the transversality condition links the time-like components to the z-components, without necessarily setting them to zero.**

The traceless condition implies:

$$\eta^{\mu\nu} C_{\mu\nu} = -C_{00} + C_{11} + C_{22} + C_{33} = 0 \tag{21}$$

Here, we are using the Minkowski metric to contract the indices. The Minkowski metric has -1 for the time component and +1 for the spatial components. This equation tells us that the sum of the spatial diagonal components must be equal to the time-time component. **Given the transversality condition $C_{00} = -C_{30} = -(-C_{33}) = C_{33}$, the traceless condition becomes** $-C_{00} + C_{11} + C_{22} + C_{33} = -C_{00} + C_{11} + C_{22} + C_{00} = C_{11} + C_{22} = 0$. **This is a crucial point: the tracelessness is what actually establishes the relationship between C_{11} and C_{22}.**

Now, let's talk about polarization. A gravitational wave can have different polarization states, just like light waves. The "+" polarization is one specific type of polarization.

For the "+" polarization, we have the following additional constraints based on standard conventions:

$$C_{11} = -C_{22} \tag{22}$$

$$C_{12} = C_{21} = 0 \tag{23}$$

These constraints define the "+" polarization. The first condition states that the xx-component is the negative of the yy-component. The second condition states that the xy-component (and its symmetric counterpart yx) is zero. In a specific choice of gauge within the TT gauge, often referred to as the "standard" TT gauge, we further set $C_{0\nu} = 0$. **This choice, along with the transversality condition $C_{0\nu} = -C_{3\nu}$, automatically implies that $C_{3\nu} = 0$ as well.**

With $C_{0\nu} = C_{3\nu} = 0$, and due to simmetry also $C_{\nu 0} = C_{\nu 3} = 0$, we are left with a 2×2 block in the spatial components $C_{11}, C_{12}, C_{21}, C_{22}$. The conditions for "+" polarization, $C_{11} = -C_{22}$ and $C_{12} = C_{21} = 0$, completely determine this block. The tracelessness condition is automatically satisfied because $C_{11} + C_{22} = 0$.

Therefore, we can see that the final form of $C_{\mu\nu}$ emerges from the combination of transversality, tracelessness, the specific gauge choice within the TT gauge, and the symmetry of the tensor.

Combining these conditions, we find that the explicit form of $C_{\mu\nu}$ for the "+" polarization is:

$$C_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_+ & 0 & 0 \\ 0 & 0 & -C_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{24}$$

where C_+ is a constant representing the amplitude of the "+" polarization. *This matrix is the explicit representation of the "+" polarization. The C_+ value determines the strength of the stretching and squeezing. Notice that only the C_{11} and C_{22} components are non-zero, and they are equal in magnitude but opposite in sign. This corresponds to the stretching and squeezing behavior in the x and y directions that we described earlier. The fact that all other components are zero is a consequence of the transversality and tracelessness conditions, along with the specific choice of the "+" polarization and the additional gauge fixing.*

In summary, the polarization tensor for a "+" polarized gravitational wave in the TT gauge, propagating in the z-direction, is a 4x4 matrix with only two non-zero elements: $C_{11} = -C_{22} = C_+$. This represents a wave that stretches and squeezes space in the x and y directions, perpendicular to its propagation direction along the z-axis. The specific form of this matrix is a direct result of the defining properties of the TT gauge and the "+" polarization. All the elements outside the 2×2 block in the x-y plane are zero because of the standard TT gauge choice, which sets to zero the temporal and longitudinal components, and the transversality condition.

(iii) Gauge transformation (Corrected)

In this part, we will explore how the metric perturbation $h_{\mu\nu}$, which represents the gravitational wave, changes under a specific coordinate transformation. This transformation is called a gauge transformation. You can think of it like changing your perspective on the same physical situation. The underlying physics doesn't change, but the way we describe it mathematically does. The specific gauge transformation we're considering is defined by the vector $\xi^\mu = (\alpha(x^\mu), \beta(x^\mu), 0, 0)$, where α and β are arbitrary functions of spacetime coordinates. This vector tells us how much we're shifting each of our spacetime coordinates.

We are given a gauge transformation parameterized by $\xi^\mu = (\alpha(x^\mu), \beta(x^\mu), 0, 0)$. This means we are shifting our time coordinate by an amount determined by the function α, and our x-coordinate by an amount determined by the function β. The y and z coordinates are not changed. The gauge transformation for the metric perturbation $h_{\mu\nu}$ is given by:

$$h_{\mu\nu} \to h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \tag{25}$$

This equation describes how the metric perturbation, which represents the gravitational wave, transforms under the infinitesimal coordinate transformation defined by ξ^μ. Essentially, we are adding to the original $h_{\mu\nu}$ a term that depends on the derivatives of the transformation vector ξ.

For the "+" polarization, the plane wave solution is:

$$h_{\mu\nu}^{TT} = \mathrm{Re}\left(C_{\mu\nu} \exp\left[ik_\sigma x^\sigma\right]\right) \tag{26}$$

with $C_{\mu\nu}$ as found in part (ii). *This represents a wave propagating in the z-direction with "+" polarization, as we discussed earlier. The term $\exp(ik_\sigma x^\sigma)$ is the oscillatory part of the wave, and $C_{\mu\nu}$ determines its amplitude and polarization.*
In our specific case, from part (ii), we know that $C_{\mu\nu}$ has only two non zero components: $C_{11} = -C_{22} = C_+$. Since we are considering only the real part, we can write explicitly the components of $h_{\mu\nu}$. With the "mostly plus" convention and a wave propagating in the positive z-direction, we have:

$$k_\sigma x^\sigma = -\omega(t - z) \tag{27}$$

Therefore, the plane wave solution becomes:

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_+ \cos(-\omega(t-z)) & 0 & 0 \\ 0 & 0 & -C_+ \cos(-\omega(t-z)) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{28}$$

Since cosine is an even function, $\cos(-\omega(t-z)) = \cos(\omega(t-z))$, we can further simplify this to:

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_+ \cos(\omega(t-z)) & 0 & 0 \\ 0 & 0 & -C_+ \cos(\omega(t-z)) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{29}$$

Let's compute the components of $\partial_\mu \xi_\nu + \partial_\nu \xi_\mu$. *We need to find the partial derivatives of the components of ξ_μ with respect to each coordinate. This is where the specific form of our gauge transformation comes into play.*

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu = \partial_\mu(\eta_{\nu\rho}\xi^\rho) + \partial_\nu(\eta_{\mu\rho}\xi^\rho)$$
$$= \eta_{\nu\rho}\partial_\mu\xi^\rho + \eta_{\mu\rho}\partial_\nu\xi^\rho$$

We lower indices of ξ using the Minkowski metric η, since $\xi_\nu = \eta_{\nu\rho}\xi^\rho$. Remember, $\xi^\rho = (\alpha, \beta, 0, 0)$, so we have $\xi_\nu = (-\alpha, \beta, 0, 0)$. The Minkowski metric is used to raise and lower indices, and it has components $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ in our convention. We have, by explicitly calculating the partial derivatives of the components of ξ_μ:

$$\xi^\rho = (\alpha, \beta, 0, 0) \implies \xi_\nu = \eta_{\nu\rho}\xi^\rho = (-\alpha, \beta, 0, 0)$$

Now we compute all the possible combinations of $\partial_\mu \xi_\nu$:

$$\partial_0 \xi_0 = -\partial_0 \alpha$$
$$\partial_0 \xi_1 = \partial_0 \beta$$
$$\partial_0 \xi_2 = 0$$
$$\partial_0 \xi_3 = 0$$
$$\partial_1 \xi_0 = -\partial_1 \alpha$$
$$\partial_1 \xi_1 = \partial_1 \beta$$
$$\partial_1 \xi_2 = 0$$
$$\partial_1 \xi_3 = 0$$
$$\partial_2 \xi_0 = -\partial_2 \alpha$$
$$\partial_2 \xi_1 = \partial_2 \beta$$
$$\partial_2 \xi_2 = 0$$
$$\partial_2 \xi_3 = 0$$
$$\partial_3 \xi_0 = -\partial_3 \alpha$$
$$\partial_3 \xi_1 = \partial_3 \beta$$
$$\partial_3 \xi_2 = 0$$
$$\partial_3 \xi_3 = 0$$

where we have explicitly used the components of ξ^μ and the fact that α and β are generic functions of all spacetime coordinates. For example, $\partial_0 \xi_0$ means the partial derivative of the 0-component of ξ_μ with respect to the 0-coordinate (time). Since $\xi_0 = -\alpha$, we have $\partial_0 \xi_0 = -\partial_0 \alpha$. Similarly for the other terms. Note how every derivative involving ξ_2 or ξ_3 is automatically zero because these components of ξ are identically zero.

Now we can compute the components of the transformed $h_{\mu\nu}$, which we will call $h'_{\mu\nu}$ to avoid confusion:

- $h'_{00} = h_{00} + \partial_0 \xi_0 + \partial_0 \xi_0 = 0 + (-\partial_0 \alpha) + (-\partial_0 \alpha) = -2\partial_0 \alpha$ *because $h_{00} = 0$ in TT gauge.*

- $h'_{01} = h_{01} + \partial_0 \xi_1 + \partial_1 \xi_0 = 0 + \partial_0 \beta - \partial_1 \alpha$ *because $h_{01} = 0$ in TT gauge*

- $h'_{02} = h_{02} + \partial_0 \xi_2 + \partial_2 \xi_0 = 0 + 0 - \partial_2 \alpha = -\partial_2 \alpha$ *because $h_{02} = 0$ in TT gauge.*

- $h'_{03} = h_{03} + \partial_0 \xi_3 + \partial_3 \xi_0 = 0 + 0 - \partial_3 \alpha = -\partial_3 \alpha$ *because $h_{03} = 0$ in TT gauge.*

- $h'_{10} = h_{10} + \partial_1 \xi_0 + \partial_0 \xi_1 = 0 - \partial_1 \alpha + \partial_0 \beta = \partial_0 \beta - \partial_1 \alpha$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{11} = h_{11} + \partial_1 \xi_1 + \partial_1 \xi_1 = C_+ \cos(\omega(t-z)) + \partial_1 \beta + \partial_1 \beta = C_+ \cos(\omega(t-z)) + 2\partial_1 \beta$ *because $h_{11} = C_+ \cos(\omega(t-z))$ in TT gauge for "+" polarization*

- $h'_{12} = h_{12} + \partial_1 \xi_2 + \partial_2 \xi_1 = 0 + 0 + \partial_2 \beta = \partial_2 \beta$ *because $h_{12} = 0$ in TT gauge*

- $h'_{13} = h_{13} + \partial_1 \xi_3 + \partial_3 \xi_1 = 0 + 0 + \partial_3 \beta = \partial_3 \beta$ *because $h_{13} = 0$ in TT gauge*

- $h'_{20} = h_{20} + \partial_2 \xi_0 + \partial_0 \xi_2 = 0 - \partial_2 \alpha + 0 = -\partial_2 \alpha$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{21} = h_{21} + \partial_2 \xi_1 + \partial_1 \xi_2 = 0 + \partial_2 \beta + 0 = \partial_2 \beta$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{22} = h_{22} + \partial_2 \xi_2 + \partial_2 \xi_2 = -C_+ \cos(\omega(t-z)) + 0 + 0 = -C_+ \cos(\omega(t-z))$ *because $h_{22} = -C_+ \cos(\omega(t-z))$ in TT gauge for "+" polarization*

- $h'_{23} = h_{23} + \partial_2 \xi_3 + \partial_3 \xi_2 = 0 + 0 + 0 = 0$ *because $h_{23} = 0$ in TT gauge*

- $h'_{30} = h_{30} + \partial_3 \xi_0 + \partial_0 \xi_3 = 0 - \partial_3 \alpha + 0 = -\partial_3 \alpha$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{31} = h_{31} + \partial_3 \xi_1 + \partial_1 \xi_3 = 0 + \partial_3 \beta + 0 = \partial_3 \beta$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{32} = h_{32} + \partial_3 \xi_2 + \partial_2 \xi_3 = 0 + 0 + 0 = 0$ *because of symmetry $h_{\mu\nu} = h_{\nu\mu}$*

- $h'_{33} = h_{33} + \partial_3 \xi_3 + \partial_3 \xi_3 = 0 + 0 + 0 = 0$ *because $h_{33} = 0$ in TT gauge*

In each of these calculations, we are simply adding the appropriate derivatives of ξ to the original components of $h_{\mu\nu}$. The derivatives of ξ are calculated based on the given form of ξ^μ. Notice how every term involving a derivative of ξ_2 or ξ_3 automatically vanishes because these components of ξ are identically zero.

Thus, the transformed metric perturbation $h'_{\mu\nu}$ is:

$$
h'_{\mu\nu} = \begin{pmatrix} -2\partial_0\alpha & \partial_0\beta - \partial_1\alpha & -\partial_2\alpha & -\partial_3\alpha \\ \partial_0\beta - \partial_1\alpha & C_+\cos(\omega(t-z)) + 2\partial_1\beta & \partial_2\beta & \partial_3\beta \\ -\partial_2\alpha & \partial_2\beta & -C_+\cos(\omega(t-z)) & 0 \\ -\partial_3\alpha & \partial_3\beta & 0 & 0 \end{pmatrix} \tag{30}
$$

This is the final form of the metric perturbation after the gauge transformation. We see that the transformation has introduced additional terms involving the derivatives of α and β. These terms represent the freedom in choosing the coordinates and do not affect the physical curvature of spacetime. The "+" polarized gravitational wave is still present, as indicated by the $C_+\cos(\omega(t-z))$ terms, but it is now described in a different coordinate system. The gauge transformation has mixed the components of $h_{\mu\nu}$, but the physics remains the same. The new terms with derivatives of α and β do not destroy the nature of the gravitational wave solution because they represent a coordinate artifact, not a physical change in the gravitational field.

Question 17

The general TT gauge plane wave solution results in
$$h_{\mu\nu}^{TT} = \mathrm{Re}\left(C_{\mu\nu}\, e^{i\, k_\sigma\, x^\sigma}\right).$$

We start by considering a Cartesian coordinate system (x, y, z) and
$$k^\mu = (\omega, 0, k, k).$$

(i) Relation between ω and k. Since gravitational waves propagate in vacuum, they travel at the speed of light, and their wave vector must be null:
$$k_\mu k^\mu = 0.$$

Using the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$ and
$$k^\mu = (\omega,\, 0,\, k,\, k), \quad k_\mu = (-\omega,\, 0,\, k,\, k),$$

we impose $k_\mu k^\mu = 0$:
$$-\omega^2 + k^2 + k^2 = 0 \implies \omega^2 = 2\, k^2 \implies \omega = \sqrt{2}\, k.$$

(ii) Explicit expression for $C_{\mu\nu}$ in the original system. We have
$$k^\mu = (\omega, 0, k, k), \quad \omega = \sqrt{2}\, k,$$

hence the phase is
$$k_\sigma\, x^\sigma = -\omega\, t + k\, y + k\, z = -\sqrt{2}\, k\, t + k\, y + k\, z.$$

We now define a new coordinate system $(x',\, y',\, z')$, where:
$$\begin{cases} x' = x, \\ z' = \frac{1}{\sqrt{2}}(y + z), \\ y' = \frac{1}{\sqrt{2}}(y - z). \end{cases}$$

This rotation in the (y, z)-plane makes $\hat{z}' \propto (0, 1, 1)$ the direction of propagation. In TT gauge, for a wave propagating along z', the only non-zero components in the *primed* frame are
$$C'_{x'x'} = h'_+, \quad C'_{y'y'} = -\, h'_+, \quad C'_{x'y'} = C'_{y'x'} = h'_\times, \quad C'_{t'\mu'} = 0, \quad C'_{z'\mu'} = 0.$$

The transformation from the new frame (t', x', y', z') back to (t, x, y, z) is given by the matrix
$$\Lambda^{\mu'}_{\ \mu} = \frac{\partial x^{\mu'}}{\partial x^\mu}, \quad \Lambda^{t'}_{\ t} = 1, \quad \Lambda^{x'}_{\ x} = 1, \quad \Lambda^{y'}_{\ y} = \frac{1}{\sqrt{2}}, \quad \Lambda^{y'}_{\ z} = -\frac{1}{\sqrt{2}}, \quad \Lambda^{z'}_{\ y} = \frac{1}{\sqrt{2}}, \quad \Lambda^{z'}_{\ z} = \frac{1}{\sqrt{2}}, \quad \text{(all other components 0)}.$$

The *amplitude tensor* in the old coordinates is then
$$C_{\mu\nu} = \Lambda^{\mu'}_{\ \mu}\, \Lambda^{\nu'}_{\ \nu}\, C'_{\mu'\nu'}.$$

Below, we list the *correct* final components (omitting lengthy intermediate sums). In particular, we stress:

$$\boxed{\begin{aligned} C_{xx} &= h'_+, \\ C_{xy} &= \frac{1}{\sqrt{2}}\, h'_\times, \qquad C_{xz} = -\frac{1}{\sqrt{2}}\, h'_\times, \\ C_{yy} &= -\tfrac{1}{2}\, h'_+, \qquad C_{yz} = +\tfrac{1}{2}\, h'_+, \qquad C_{zz} = -\tfrac{1}{2}\, h'_+, \\ &C_{\mu t} = C_{t\mu} = 0 \quad \text{(transverse)}, \quad \mathrm{Trace}[C_{ij}] = 0 \quad \text{(traceless)}. \end{aligned}}$$

Putting everything together in matrix form (with $\mu, \nu \in \{t, x, y, z\}$):

$$C_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h'_+ & \frac{1}{\sqrt{2}}\, h'_\times & -\frac{1}{\sqrt{2}}\, h'_\times \\ 0 & \frac{1}{\sqrt{2}}\, h'_\times & -\frac{1}{2}\, h'_+ & \frac{1}{2}\, h'_+ \\ 0 & -\frac{1}{\sqrt{2}}\, h'_\times & \frac{1}{2}\, h'_+ & -\frac{1}{2}\, h'_+ \end{pmatrix}.$$

This final $C_{\mu\nu}$ correctly reflects:

- Transversality: $C_{t\mu} = 0$.

- Tracelessness in the spatial block: $C_{xx} + C_{yy} + C_{zz} = 0$.

- Separation of polarizations: the off-diagonal (x, y) and (x, z) entries depend only on $h'_\times$, while (y, y), (y, z), (z, z) depend only on h'_+.

Question 18

Assume that the Friedmann equation $H^2 = 8\pi G\rho/3 - \kappa/a^2$ and the fluid equation $\dot{\rho} + 3H(\rho + p) = 0$ are valid. Without any further assumption, derive the result $\ddot{a}/a = -4\pi G(\rho + 3p)/3$.

Solution

The goal of this exercise is to derive the acceleration equation, starting from the Friedmann equation and the fluid equation. Let's start by recalling the relevant equations:

- **Friedmann Equation**: This equation describes the expansion rate of the Universe, relating the Hubble parameter (H) to the energy density (ρ) of the Universe, and the curvature parameter (κ) of the space, scaled by the scale factor(a).

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} \tag{31}$$

  where:

  - $H = \frac{\dot{a}}{a}$ is the Hubble parameter, representing the expansion rate of the Universe.
  - $\dot{a}$ is the time derivative of the scale factor, representing how fast the scale factor is changing.
  - a is the scale factor, a function of time that represents the relative size of the Universe.
  - ρ is the total energy density of the Universe.
  - κ is the curvature parameter, which determines the spatial geometry of the Universe.
  - G is the gravitational constant.

- **Fluid Equation**: This equation is derived from the conservation of energy-momentum tensor, $\nabla_\mu T^{\mu\nu} = 0$, when $\nu = 0$. It describes how the energy density of a perfect fluid changes as the Universe expands or contracts.

$$\dot{\rho} + 3H(\rho + p) = 0 \tag{32}$$

  where:

  - $\dot{\rho}$ is the time derivative of the energy density, representing how the energy density changes over time.
  - p is the pressure of the fluid.

Our goal is to derive the acceleration equation:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \tag{33}$$

Derivation Steps:
1. Differentiate the Friedmann Equation:
We start by taking the time derivative of the Friedmann equation (92):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{a}}{a}\right)^2 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{8\pi G}{3}\rho - \frac{\kappa}{a^2}\right) \tag{34}$$

Applying the chain rule on the left-hand side and the rules of differentiation on the right-hand side, we get:

$$2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{a}}{a}\right) = \frac{8\pi G}{3}\dot{\rho} + 2\frac{\kappa}{a^3}\dot{a} \tag{35}$$

2. Simplify using the definition of H:
Recall that $H = \frac{\dot{a}}{a}$. We can rewrite the equation as:

$$2H\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{a}}{a}\right) = \frac{8\pi G}{3}\dot{\rho} + 2\frac{\kappa}{a^2}\frac{\dot{a}}{a} \tag{36}$$

$$2H\left(\frac{\ddot{a}a - \dot{a}^2}{a^2}\right) = \frac{8\pi G}{3}\dot{\rho} + 2\frac{\kappa}{a^2}H \tag{37}$$

$$2H\left(\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2}\right) = \frac{8\pi G}{3}\dot{\rho} + 2\frac{\kappa}{a^2}H \tag{38}$$

$$2H\left(\frac{\ddot{a}}{a} - H^2\right) = \frac{8\pi G}{3}\dot{\rho} + 2\frac{\kappa}{a^2}H \tag{39}$$

Dividing both sides by $2H$, we get:

$$\frac{\ddot{a}}{a} - H^2 = \frac{4\pi G}{3}\dot{\rho} + \frac{\kappa}{a^2} \tag{40}$$

3. Substitute the Fluid Equation:
Now, we substitute the fluid equation (32) into equation (40). First, we solve for $\dot{\rho}$:

$$\dot{\rho} = -3H(\rho + p) \tag{41}$$

Substitute this into equation (40):

$$\frac{\ddot{a}}{a} - H^2 = \frac{4\pi G}{3}(-3H(\rho + p)) + \frac{\kappa}{a^2} \tag{42}$$

4. Rearrange and Substitute Friedmann Equation again:
Rearranging the terms, we get:

$$\frac{\ddot{a}}{a} = H^2 - 4\pi GH(\rho + p) + \frac{\kappa}{a^2} \tag{43}$$

Now, substitute the Friedmann equation (92) for H^2:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} - 4\pi GH(\rho + p) + \frac{\kappa}{a^2} \tag{44}$$

5. Simplify to Obtain the Acceleration Equation:
Notice that the $\frac{\kappa}{a^2}$ terms cancel out:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - 4\pi GH(\rho + p) \tag{45}$$

We can rewrite the equation as:

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho + H(\rho + p)\right) \tag{46}$$

Since the fluid equation, (32), can be rewritten as:

$$H(\rho + p) = -\frac{\dot{\rho}}{3}. \tag{47}$$

We can substitute the H expression inside the acceleration equation

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho - \frac{\dot{\rho}}{3}\right) \tag{48}$$

In order to find the acceleration equation we have to substitute the time derivative of the energy density, using the fluid equation expressed with respect to the time derivative of the scale factor:

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p) \tag{49}$$

So the acceleration equation becomes:

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho + \frac{3\dot{a}(\rho + p)}{3a}\right) = -4\pi G\left(-\frac{2}{3}\rho + \frac{\dot{a}}{a}(\rho + p)\right) \tag{50}$$

We can also substitute the Hubble parameter inside the equation, since it is equal to $H = \frac{\dot{a}}{a}$. So we get:

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho + H(\rho + p)\right) \tag{51}$$

At this point we can substitute H^2 from the Friedman equation, eq. 92:

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho + \sqrt{\frac{8\pi G}{3}\rho - \frac{\kappa}{a^2}}(\rho + p)\right) \tag{52}$$

Now we make an assumption, that is to consider a flat Universe, so that $\kappa = 0$

$$\frac{\ddot{a}}{a} = -4\pi G\left(-\frac{2}{3}\rho + \sqrt{\frac{8\pi G}{3}\rho}(\rho + p)\right) \tag{53}$$

Since this equation is not equal to the equation of acceleration, we made some mistakes along the way, so we will restart from eq. 40.

6. Substitute the Fluid Equation:

The fluid equation (32) can be rewritten as:

$$\dot{\rho} = -3H(\rho + p) \tag{54}$$

Substituting this into equation (40), we get:

$$\frac{\ddot{a}}{a} - H^2 = \frac{4\pi G}{3}(-3H(\rho + p)) + \frac{\kappa}{a^2} \tag{55}$$

7. Rearrange and Substitute the Friedmann Equation:

Rearranging the terms, we have:

$$\frac{\ddot{a}}{a} = H^2 - 4\pi G H(\rho + p) + \frac{\kappa}{a^2} \tag{56}$$

Substituting the Friedmann equation (92) for H^2, we get:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} - 4\pi G H(\rho + p) + \frac{\kappa}{a^2} \tag{57}$$

8. Simplify to Obtain the Acceleration Equation:

The $\frac{\kappa}{a^2}$ terms cancel out, leaving us with:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - 4\pi G H(\rho + p) \tag{58}$$

Substituting $H = \frac{\dot{a}}{a}$ and using the fluid equation in the form $\dot{\rho} = -3H(\rho + p)$, we can rewrite this as:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho + \frac{4\pi G}{3}\dot{\rho} \tag{59}$$

Substituting $\dot{\rho}$ from the fluid equation:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho + \frac{4\pi G}{3}\left(-3\frac{\dot{a}}{a}(\rho + p)\right) = \frac{8\pi G}{3}\rho - 4\pi G \frac{\dot{a}}{a}(\rho + p) \tag{60}$$

Substituting again $H = \frac{\dot{a}}{a}$ we get:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - 4\pi G H(\rho + p) \tag{61}$$

This equation can be further simplified by substituting $\dot{\rho}$ from the fluid equation into the equation:

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho - \frac{4\pi G}{3}(\dot{\rho} + 3H(\rho + p)) \tag{62}$$

Since $\dot{\rho} + 3H(\rho + p) = 0$ from the fluid equation, we finally obtain the acceleration equation:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \tag{63}$$

This is the desired result, which shows how the acceleration of the scale factor depends on the energy density and pressure of the Universe.

Question 19

Assume that the universe is spatially flat ($\kappa = 0$) and that is filled with a single fluid with equation of state $p = (2/3)\rho$. Let t_0 be the current time, $a_0 = 1$ the current value of the scale factor, ρ_0 the current value of the energy density. Derive the following quantities and provide the answers in terms of the variables defined in this exercise : (i) how the energy density $\rho(a)$ depends on the scale factor; (ii) how the scale factor $a(t)$ depends on time; (iii) how the Hubble parameter $H(t)$ depends on time; (iv) how the energy density $\rho(t)$ depends on time.

Solution

We are given a spatially flat universe ($\kappa = 0$) filled with a single fluid with an equation of state $p = \frac{2}{3}\rho$. We are also given that the present time is t_0, the present scale factor is $a_0 = 1$, and the present energy density is ρ_0.

(i) Energy density $\rho(a)$ as a function of the scale factor   From the conservation of four-momentum, $\nabla_\mu T^\mu_\nu = 0$, specifically for the temporal component ($\nu = 0$), we have (from **Lec 24**):

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \tag{64}$$

We are given the equation of state $p = w\rho$, where in this case $w = \frac{2}{3}$. Substituting this into the fluid equation (64), we get:

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}\left(\rho + \frac{2}{3}\rho\right) = 0 \tag{65}$$

$$\frac{\partial \rho}{\partial t} + 5\frac{\dot{a}}{a}\rho = 0 \tag{66}$$

Now, we can rewrite this equation as:

$$\frac{d\rho}{\rho} = -5\frac{da}{a} \tag{67}$$

Integrating both sides, we obtain:

$$\int_{\rho_0}^{\rho} \frac{d\rho'}{\rho'} = -5 \int_{a_0}^{a} \frac{da'}{a'} \tag{68}$$

$$\ln\left(\frac{\rho}{\rho_0}\right) = -5\ln\left(\frac{a}{a_0}\right) \tag{69}$$

$$\rho(a) = \rho_0 \left(\frac{a_0}{a}\right)^5 \tag{70}$$

Since $a_0 = 1$, we have:

$$\rho(a) = \rho_0 a^{-5} \tag{71}$$

This equation shows that the energy density ρ scales with the scale factor a as a^{-5} for a fluid with $w = \frac{2}{3}$.

(ii) Scale factor $a(t)$ as a function of time   For a spatially flat universe ($\kappa = 0$), the Friedmann equation (**Lec 24**) is:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho \tag{72}$$

Substituting $\rho(a) = \rho_0 a^{-5}$ into the Friedmann equation, we get:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho_0 a^{-5} \tag{73}$$

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi G\rho_0}{3}}a^{-5/2} \tag{74}$$

$$\dot{a} = \sqrt{\frac{8\pi G\rho_0}{3}}a^{-3/2} \tag{75}$$

$$\frac{da}{dt} = \sqrt{\frac{8\pi G\rho_0}{3}}a^{-3/2} \tag{76}$$

50

$$a^{3/2}da = \sqrt{\frac{8\pi G\rho_0}{3}}dt \tag{77}$$

Integrating both sides, we have:

$$\int_0^{a(t)} a'^{3/2}da' = \sqrt{\frac{8\pi G\rho_0}{3}}\int_0^t dt' \tag{78}$$

$$\frac{2}{5}a(t)^{5/2} = \sqrt{\frac{8\pi G\rho_0}{3}}t \tag{79}$$

$$a(t) = \left(\frac{5}{2}\sqrt{\frac{8\pi G\rho_0}{3}}t\right)^{2/5} \tag{80}$$

$$a(t) = \left(\frac{25}{4}\frac{8\pi G\rho_0}{3}\right)^{1/5}t^{2/5} \tag{81}$$

$$a(t) = \left(\frac{50\pi G\rho_0}{3}\right)^{1/5}t^{2/5} \tag{82}$$

Since $a(t_0) = a_0 = 1$, we can write:

$$a(t) = \left(\frac{t}{t_0}\right)^{2/5} \tag{83}$$

where t_0 is determined by the condition $a(t_0) = 1$:

$$1 = \left(\frac{50\pi G\rho_0}{3}\right)^{1/5}t_0^{2/5} \tag{84}$$

$$t_0 = \left(\frac{3}{50\pi G\rho_0}\right)^{1/2} \tag{85}$$

So the scale factor as a function of time is:

$$a(t) = \left(\frac{t}{t_0}\right)^{2/5} = \left(\sqrt{\frac{50\pi G\rho_0}{3}}t\right)^{2/5} \tag{86}$$

(iii) Hubble parameter $H(t)$ as a function of time The Hubble parameter is defined as $H(t) = \frac{\dot{a}(t)}{a(t)}$. Using the expression for $a(t)$ we derived above:

$$\dot{a}(t) = \frac{2}{5}\left(\frac{50\pi G\rho_0}{3}\right)^{1/5}t^{-3/5} = \frac{2}{5}\frac{a(t)}{t} \tag{87}$$

Therefore,

$$H(t) = \frac{\dot{a}(t)}{a(t)} = \frac{2}{5t} \tag{88}$$

(iv) Energy density $\rho(t)$ as a function of time We already found how the energy density depends on the scale factor: $\rho(a) = \rho_0 a^{-5}$. Now we can substitute the expression for $a(t)$:

$$\rho(t) = \rho_0 \left[\left(\frac{50\pi G\rho_0}{3}\right)^{1/5}t^{2/5}\right]^{-5} \tag{89}$$

$$\rho(t) = \rho_0 \left(\frac{3}{50\pi G\rho_0}\right)t^{-2} \tag{90}$$

$$\rho(t) = \frac{3}{50\pi Gt^2} \tag{91}$$

This shows that the energy density decreases with time as t^{-2}.

Question 20.

Consider a spatially flat universe ($\kappa = 0$) with a non-vanishing cosmological constant Λ and containing only non-relativistic matter with energy density $\rho_M(a)$. Let t_0 be the current time, $a_0 = 1$ the current value of the scale factor, ρ_0 the current value of the energy density, and H_0 the current value of the Hubble parameter. Define $\Omega_M(a) \equiv \rho_M(a)/\rho_{\text{cr}}(a)$, where $\rho_{\text{cr}}(a)$ is the critical energy density. (i) Derive an expression for $\Omega_M(a)$ as a function of the scale factor a in terms of the data of the problem. (ii) With the additional information $\Omega_0 = 0.3$, derive the value of the scale factor a_Λ that corresponds to the moment when the universe began accelerating.

Solution.

(i) First step We are given a spatially flat universe ($\kappa = 0$) with a cosmological constant Λ and non-relativistic matter. The Friedmann equation, which describes the evolution of the scale factor $a(t)$ in a homogeneous and isotropic universe, is given by (see *Lec 24*):

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} + \frac{\Lambda}{3} \tag{92}$$

where:

- $H(t) = \frac{\dot{a}(t)}{a(t)}$ is the Hubble parameter, representing the expansion rate of the universe.

- $a(t)$ is the scale factor, a function of time that describes the relative size of the universe.

- ρ is the total energy density of the universe.

- κ is the curvature parameter, which is 0 for a flat universe.

- Λ is the cosmological constant.

- G is the gravitational constant.

In our case, since the universe is spatially flat, $\kappa = 0$. The total energy density ρ is the sum of the matter energy density ρ_M and the cosmological constant energy density ρ_Λ. The energy density of non-relativistic matter scales as a^{-3} (see *Lec 24*), so we can write:

$$\rho_M(a) = \rho_{M,0}\left(\frac{a_0}{a}\right)^3 = \rho_0 a^{-3} \tag{93}$$

since $a_0 = 1$ by definition. The cosmological constant energy density ρ_Λ is related to Λ by (see *Lec 24*):

$$\rho_\Lambda = \frac{\Lambda}{8\pi G} \tag{94}$$

and it remains constant as the universe expands. The critical density ρ_{cr} is defined as the density required for a spatially flat universe without a cosmological constant. It is given by (see *Lec 24*):

$$\rho_{\text{cr}}(t) = \frac{3H^2(t)}{8\pi G} \tag{95}$$

The current value of the critical density is:

$$\rho_{\text{cr},0} = \frac{3H_0^2}{8\pi G} \tag{96}$$

We are given that the current value of the density parameter for matter is $\Omega_0 = 0.3$. This means:

$$\Omega_0 = \frac{\rho_0}{\rho_{\text{cr},0}} = 0.3 \tag{97}$$

Using the Friedmann equation (92) with $\kappa = 0$, we have:

$$H^2 = \frac{8\pi G}{3}(\rho_M + \rho_\Lambda) = \frac{8\pi G}{3}\rho_M + \frac{\Lambda}{3} \tag{98}$$

At the present time ($t = t_0$, $a = 1$, $H = H_0$, $\rho_M = \rho_0$):

$$H_0^2 = \frac{8\pi G}{3}\rho_0 + \frac{\Lambda}{3} \tag{99}$$

Dividing the Friedmann equation by H^2, we get:

$$1 = \frac{8\pi G}{3H^2}\rho_M + \frac{\Lambda}{3H^2} = \frac{\rho_M}{\rho_{\rm cr}} + \frac{\Lambda}{3H^2} \tag{100}$$

The density parameter for matter is defined as:

$$\Omega_M(a) = \frac{\rho_M(a)}{\rho_{\rm cr}(a)} \tag{101}$$

We want to express $\Omega_M(a)$ in terms of the given quantities. We know that $\rho_M(a) = \rho_0 a^{-3}$ and $\rho_{\rm cr}(a) = \frac{3H^2(a)}{8\pi G}$. Thus,

$$\Omega_M(a) = \frac{\rho_0 a^{-3}}{\frac{3H^2(a)}{8\pi G}} = \frac{8\pi G \rho_0}{3H^2(a)a^3} \tag{102}$$

We can rewrite this using the definition of $\rho_{\rm cr,0}$:

$$\Omega_M(a) = \frac{\rho_0}{\rho_{\rm cr,0}}\frac{\rho_{\rm cr,0}}{\rho_{\rm cr}(a)}\frac{1}{a^3} = \Omega_0 \frac{H_0^2}{H^2(a)}\frac{1}{a^3} \tag{103}$$

From the Friedmann equation, we have:

$$H^2(a) = \frac{8\pi G}{3}\rho_0 a^{-3} + \frac{\Lambda}{3} = H_0^2\Omega_0 a^{-3} + \frac{\Lambda}{3} \tag{104}$$

From $H_0^2 = \frac{8\pi G}{3}\rho_0 + \frac{\Lambda}{3}$, we can write $\frac{\Lambda}{3} = H_0^2 - \frac{8\pi G}{3}\rho_0 = H_0^2 - H_0^2\Omega_0 = H_0^2(1 - \Omega_0)$. Substituting this into the expression for $H^2(a)$, we get:

$$H^2(a) = H_0^2\Omega_0 a^{-3} + H_0^2(1 - \Omega_0) = H_0^2\left(\Omega_0 a^{-3} + 1 - \Omega_0\right) \tag{105}$$

Substituting this into the expression for $\Omega_M(a)$, we have:

$$\Omega_M(a) = \Omega_0 \frac{H_0^2}{H_0^2\left(\Omega_0 a^{-3} + 1 - \Omega_0\right)}\frac{1}{a^3} = \frac{\Omega_0 a^{-3}}{\Omega_0 a^{-3} + 1 - \Omega_0} \tag{106}$$

This is the desired expression for $\Omega_M(a)$ as a function of the scale factor a.

(ii) Second step   To find the value of the scale factor a_Λ when the universe began accelerating, we need to find when the acceleration $\ddot{a}$ changes sign. The acceleration equation is given by (see *Lec 24*):

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{107}$$

For non-relativistic matter, $p_M = 0$. For the cosmological constant, $p_\Lambda = -\rho_\Lambda$. Thus,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho_M - 2\rho_\Lambda) \tag{108}$$

The universe begins to accelerate when $\ddot{a} > 0$, which occurs when $\rho_M - 2\rho_\Lambda < 0$, or $\rho_M < 2\rho_\Lambda$. We know that $\rho_M(a) = \rho_0 a^{-3}$ and $\rho_\Lambda = \frac{\Lambda}{8\pi G}$. Also, $\Lambda = 3H_0^2(1 - \Omega_0)$. Thus,

$$\rho_\Lambda = \frac{3H_0^2(1 - \Omega_0)}{8\pi G} = \rho_{\rm cr,0}(1 - \Omega_0) = \frac{\rho_0}{\Omega_0}(1 - \Omega_0) \tag{109}$$

The condition for acceleration becomes:

$$\rho_0 a^{-3} < 2\frac{\rho_0}{\Omega_0}(1 - \Omega_0) \tag{110}$$

Solving for a, we get:

$$a^3 > \frac{\Omega_0}{2(1 - \Omega_0)} \tag{111}$$

Thus, the scale factor a_Λ at which the universe began accelerating is:

$$a_\Lambda = \left(\frac{\Omega_0}{2(1 - \Omega_0)}\right)^{1/3} \tag{112}$$

With $\Omega_0 = 0.3$, we have:

$$a_\Lambda = \left(\frac{0.3}{2(1 - 0.3)}\right)^{1/3} = \left(\frac{0.3}{1.4}\right)^{1/3} = \left(\frac{3}{14}\right)^{1/3} \approx 0.6 \tag{113}$$