

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 $dx dy$, 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} :

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

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} :

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

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 $\Gamma_{\nu\lambda}^\mu$ 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

$$\Gamma_{\nu\lambda}^\mu = \Gamma_{\lambda\nu}^\mu,$$

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 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,

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

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:

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

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:

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

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 $\Gamma_{\phi\phi}^r$. For illustration, let us compute $\Gamma_{\phi\phi}^r$. Substitute $\mu = r$, $\nu = \phi$, and $\lambda = \phi$ into the Levi-Civita connection formula:

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

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

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

thus

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

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

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

Hence, $\Gamma_{\phi\phi}^r = -r$.

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_{tx}^t 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

$$\Gamma_{\nu\lambda}^\mu = \Gamma_{\lambda\nu}^\mu$$

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:

$$\Gamma_{\nu\lambda}^\mu = \frac{1}{2} g^{\mu\rho} (\partial_\nu g_{\rho\lambda} + \partial_\lambda g_{\rho\nu} - \partial_\rho g_{\nu\lambda}).$$

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_{tt}^x , $\Gamma_{tx}^t = \Gamma_{xt}^t$, and Γ_{xx}^x . A direct check confirms $\Gamma_{xx}^x = 0$. Thus the only non-zero Christoffel symbols are:

$$\Gamma_{tt}^x \quad \text{and} \quad \Gamma_{tx}^t = \Gamma_{xt}^t.$$

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_{tx}^t . We compute:

$$\Gamma_{tx}^t = \frac{1}{2} g^{t\rho} (\partial_t g_{x\rho} + \partial_x g_{t\rho} - \partial_\rho g_{tx}).$$

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:

$$\Gamma_{tx}^t = \frac{1}{2} g^{tt} \partial_x g_{tt}.$$

We substitute:

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

Hence:

$$\Gamma_{tx}^t = \frac{1}{2} \left(-\frac{1}{(1+x)^2} \right) (-2(1+x)) = \frac{1}{1+x}.$$

Interpretation: $\Gamma_{tx}^t = \frac{1}{1+x}$ 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 $\Gamma_{tt}^x = 1+x$ and verify $\Gamma_{xx}^x = 0$. Together with $\Gamma_{tx}^t = \Gamma_{xt}^t = \frac{1}{1+x}$, 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 .

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_{xx}^x and Γ_{yy}^x .

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_{xx}^x and Γ_{yy}^x .

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:

$$\Gamma_{\nu\lambda}^\mu = \frac{1}{2} g^{\mu\rho} (\partial_\nu g_{\rho\lambda} + \partial_\lambda g_{\rho\nu} - \partial_\rho g_{\nu\lambda}).$$

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_{xx}^x . We set $\mu = x$, $\nu = x$, $\lambda = x$ in the formula. Hence,

$$\Gamma_{xx}^x = \frac{1}{2} g^{x\rho} (\partial_x g_{\rho x} + \partial_x g_{x\rho} - \partial_\rho g_{xx}).$$

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:

$$\Gamma_{xx}^x = \frac{1}{2} g^{xx} (\partial_x g_{xx} + \partial_x g_{xx} - \partial_x g_{xx}) = \frac{1}{2} g^{xx} (\partial_x g_{xx}).$$

Since $\partial_x g_{xx} = 2x$ and $g^{xx} = \frac{1}{1+x^2}$, we obtain

$$\Gamma_{xx}^x = \frac{1}{2} \cdot \frac{1}{1+x^2} \cdot (2x) = \frac{x}{1+x^2}.$$

(ii) Calculation of Γ_{yy}^x . We now set $\mu = x$, $\nu = y$, $\lambda = y$:

$$\Gamma_{yy}^x = \frac{1}{2} g^{x\rho} (\partial_y g_{\rho y} + \partial_y g_{y\rho} - \partial_\rho g_{yy}).$$

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,

$$\Gamma_{yy}^x = 0.$$

Final Results

$$\boxed{\Gamma_{xx}^x = \frac{x}{1+x^2} \quad \text{and} \quad \Gamma_{yy}^x = 0.}$$

Because g_{xx} depends on x , Γ_{xx}^x is non-zero. Meanwhile, g_{yy} does not depend on x , so Γ_{yy}^x 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 $|x| < 1$:

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

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:

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:

$$\begin{aligned} G &\approx 6.674 \times 10^{-11} \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2} \\ M_E &\approx 5.972 \times 10^{24} \text{ kg} \\ R_E &\approx 6.371 \times 10^6 \text{ m} \\ \frac{\tau_2}{\tau_1} &\approx 1 + \frac{6.674 \times 10^{-11} \cdot 5.972 \times 10^{24}}{(6.371 \times 10^6)^2} h \approx 1 + (9.82 \frac{\text{m}}{\text{s}^2}) \frac{h}{c^2} \end{aligned}$$

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:

$$\begin{aligned} [G] &= \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2} \\ [M_E] &= \text{kg} \\ [R_E] &= \text{m} \end{aligned}$$

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:

$$\begin{aligned} [c^2] &= 1 \\ [G] &= 1 \end{aligned}$$

Since $c = 1$, then $\left[\frac{\text{m}^2}{\text{s}^2} \right] = 1$. 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:

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

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:

$$\Gamma_{\phi\phi}^\theta = -\sin \theta \cos \theta \quad \text{and} \quad \Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi = \frac{\cos \theta}{\sin \theta} = \cot \theta.$$

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 ($\phi = \text{const.}$) are geodesics,
- (ii) the only geodesic at constant latitude ($\theta = \text{const.}$) is the equator ($\theta = \frac{\pi}{2}$).

Solution

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

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

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 $\Gamma_{\phi\phi}^\theta$, while for $\mu = \phi$, there are contributions from $\Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi$.

Equation (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

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

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 ($\theta = \text{const}$) 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 ($\phi = \text{const.}$) 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 ($\theta = \text{const.}$) 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

$$\Gamma_{\phi\phi}^\theta = -\sin \theta \cos \theta, \quad \Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi = \cot \theta.$$

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

$$\Gamma_{\phi\phi}^\theta = -\sin \theta \cos \theta, \quad \Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi = \cot \theta.$$

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. $\Gamma_{\phi\phi}^\theta$ represents the variation of the θ -basis vector when moving along the ϕ -direction, whereas $\Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi$ 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

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

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 $\nu = \theta$:

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

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 $\Gamma_{\theta\phi}^\theta = 0$, Because the Christoffel symbols $\Gamma_{\theta\theta}^\theta$ and $\Gamma_{\theta\phi}^\theta$ 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 $\nu = \phi$:

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

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

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

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

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

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

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

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 $\nu = \theta$:

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

We compute the $\theta\theta$ component of the covariant derivative of a covariant vector. 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,

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

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

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

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

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

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

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

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

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

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

Hence,

$$\nabla_\mu(V^\nu V_\nu) = V^\nu (\partial_\mu V_\nu - \Gamma_{\mu\nu}^\lambda V_\lambda) + V_\nu (\partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda).$$

Expanding,

$$\nabla_\mu(V^\nu V_\nu) = \underbrace{V^\nu \partial_\mu V_\nu + V_\nu \partial_\mu V^\nu}_{(A)} + \underbrace{(-V^\nu \Gamma_{\mu\nu}^\lambda V_\lambda + V_\nu \Gamma_{\mu\lambda}^\nu V^\lambda)}_{(B)}.$$

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

$$T_1 = -V^\nu \Gamma_{\mu\nu}^\lambda V_\lambda, \quad T_2 = V_\nu \Gamma_{\mu\lambda}^\nu V^\lambda.$$

Then

$$(B) = T_1 + T_2.$$

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

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

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

$$\nu \mapsto \lambda, \quad \lambda \mapsto \nu.$$

Hence,

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

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

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

Putting them together:

$$T_1 + T_2 = -[V^\nu \Gamma_{\mu\nu}^\lambda V_\lambda] + [V_\lambda \Gamma_{\mu\nu}^\lambda V^\nu].$$

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

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

and thus

$$T_1 + T_2 = -[V^\nu V_\lambda \Gamma_{\mu\nu}^\lambda] + [V^\nu V_\lambda \Gamma_{\mu\nu}^\lambda] = 0.$$

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

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

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

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

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:

$$[\nabla_\alpha, \nabla_\beta] W^{\mu\nu} = R^\mu_{\sigma\alpha\beta} W^{\sigma\nu} + R^\nu_{\sigma\alpha\beta} W^{\mu\sigma},$$

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:

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

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

$$[\nabla_\mu, \nabla_\nu] W^{\mu\nu} = R^\mu_{\sigma\mu\nu} W^{\sigma\nu} + R^\nu_{\sigma\mu\nu} W^{\mu\sigma}.$$

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:

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

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

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

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

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

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 $\mu \rightarrow \nu$), we get

$$R_{\sigma\nu} W^{\sigma\nu} - R_{\sigma\nu} W^{\nu\sigma} = R_{\sigma\nu} (W^{\sigma\nu} - W^{\nu\sigma}).$$

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,

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

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 We conclude by deriving the two possible conditions under which the commutator vanishes. These conditions arise from the fact that the product of two factors is zero if and only if at least one of the factors is zero. Thus, there are two distinct possibilities for

$$[\nabla_{\mu}, \nabla_{\nu}] W^{\mu\nu} = 0:$$

- **Case 1: $W^{\mu\nu}$ is symmetric.** If $W^{\mu\nu}$ is symmetric, then $W^{\sigma\nu} - W^{\nu\sigma} = 0$, and consequently

$$R_{\sigma\nu} (W^{\sigma\nu} - W^{\nu\sigma}) = 0$$

for any Ricci tensor $R_{\sigma\nu}$, making the whole expression vanish. *Case 1: If $W^{\mu\nu}$ is symmetric (i.e., $W^{\mu\nu} = W^{\nu\mu}$), then the difference $W^{\sigma\nu} - W^{\nu\sigma}$ is zero, and therefore the whole expression vanishes, regardless of the value of the Ricci tensor. A tensor is symmetric if swapping its indices doesn't change its value. This means that if we swap the indices σ and ν , the tensor remains the same.*

- **Case 2:** $R_{\mu\nu} = 0$. If the spacetime is Ricci-flat, meaning $R_{\mu\nu} = 0$, then

$$R_{\sigma\nu} (W^{\sigma\nu} - W^{\nu\sigma}) = 0$$

regardless of whether $W^{\mu\nu}$ is symmetric or not. In this scenario, the commutator also vanishes for *any* $(2,0)$ tensor. *Case 2: If the spacetime is Ricci-flat (i.e., $R_{\mu\nu} = 0$), then the whole expression vanishes, regardless of the symmetry or antisymmetry of $W^{\mu\nu}$. In this case, the commutator vanishes for any tensor of type $(2,0)$. Ricci-flat means that the Ricci tensor is zero everywhere. This is a strong condition on the geometry of spacetime.*

(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,

$$[\nabla_\mu, \nabla_\nu] W^{\mu\nu} = 0 \iff \text{either } W^{\mu\nu} \text{ is symmetric, or } R_{\mu\nu} = 0.$$

Ultimately, the commutator $[\nabla_\mu, \nabla_\nu] W^{\mu\nu}$ 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.

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:

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

where ∂_μ is the ordinary partial derivative, and $\Gamma_{\mu\lambda}^\nu$ 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, $\Gamma_{\mu\lambda}^\nu V^\lambda$, 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:

$$\nabla_\alpha W^{\mu\nu} = \partial_\alpha W^{\mu\nu} + \Gamma_{\alpha\sigma}^\mu W^{\sigma\nu} + \Gamma_{\alpha\sigma}^\nu W^{\mu\sigma}$$

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:

$$[\nabla_\alpha, \nabla_\beta] W^{\mu\nu} = R^\mu_{\sigma\alpha\beta} W^{\sigma\nu} + R^\nu_{\sigma\alpha\beta} W^{\mu\sigma},$$

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:

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

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 in Contravariant Form. The standard geodesic equation is:

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

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

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

Equation () 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 $\Gamma_{\alpha\beta}^\mu$ represent the affine connection, and they encode how the spacetime is curved. They depend on the first derivatives of the metric.

Substituting $p^\mu = \frac{dx^\mu}{d\lambda}$ into (), we get the contravariant form of the four-momentum evolution:

$$\frac{dp^\mu}{d\lambda} + \Gamma_{\alpha\beta}^\mu p^\alpha p^\beta = 0.$$

This equation tells us how the contravariant four-momentum changes along the geodesic.

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

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

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

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

The inverse metric is obtained by calculating the inverse of the metric tensor matrix.

4. Conserved Quantities and Symmetries. A crucial point in General Relativity is that if the metric $g_{\alpha\beta}$ does not explicitly depend on a coordinate x^σ , then the corresponding component of the covariant four-momentum p_σ is conserved.

Why?

- Multiplying the geodesic equation () by $g_{\mu\nu}$, we get:

$$g_{\mu\nu} \frac{d^2 x^\mu}{d\lambda^2} + g_{\mu\nu} \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0.$$

Recalling that $p_\nu = g_{\mu\nu} p^\mu$ and $p^\mu = \frac{dx^\mu}{d\lambda}$, we can rewrite the first term as $\frac{dp_\nu}{d\lambda}$. For the second term, we use the definition of the Christoffel symbols () and, for simplicity, consider only the part containing the metric derivatives:

$$g_{\mu\nu} \Gamma_{\alpha\beta}^\mu = g_{\mu\nu} \frac{1}{2} g^{\mu\sigma} (\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\sigma\alpha} - \partial_\sigma g_{\alpha\beta}) = \frac{1}{2} \delta_\nu^\sigma (\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\sigma\alpha} - \partial_\sigma g_{\alpha\beta})$$

where we used the fact that $g_{\mu\nu}g^{\mu\sigma} = \delta_\nu^\sigma$. Simplifying further, we obtain:

$$g_{\mu\nu}\Gamma_{\alpha\beta}^\mu = \frac{1}{2}(\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta})$$

Substituting into the geodesic equation multiplied by $g_{\mu\nu}$, we get:

$$\frac{dp_\nu}{d\lambda} = \frac{1}{2}\partial_\nu g_{\alpha\beta} p^\alpha p^\beta.$$

This is the covariant form of the geodesic equation.

In fact, explicitly writing out the steps, we have:

$$\begin{aligned} g_{\mu\nu}\frac{d^2x^\mu}{d\lambda^2} + g_{\mu\nu}\Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} &= 0 \\ \frac{d}{d\lambda} \left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \right) - \frac{dg_{\mu\nu}}{d\lambda} \frac{dx^\mu}{d\lambda} + g_{\mu\nu}\Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} &= 0 \\ \frac{dp_\nu}{d\lambda} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \frac{dx^\rho}{d\lambda} \frac{dx^\mu}{d\lambda} + g_{\mu\nu} \frac{1}{2} g^{\mu\sigma} (\partial_\alpha g_{\sigma\beta} + \partial_\beta g_{\sigma\alpha} - \partial_\sigma g_{\alpha\beta}) \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} &= 0 \\ \frac{dp_\nu}{d\lambda} - \partial_\rho g_{\mu\nu} p^\rho p^\mu + \frac{1}{2} (\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}) p^\alpha p^\beta &= 0 \\ \frac{dp_\nu}{d\lambda} - \partial_\alpha g_{\beta\nu} p^\alpha p^\beta + \frac{1}{2} \partial_\alpha g_{\nu\beta} p^\alpha p^\beta + \frac{1}{2} \partial_\beta g_{\nu\alpha} p^\alpha p^\beta - \frac{1}{2} \partial_\nu g_{\alpha\beta} p^\alpha p^\beta &= 0 \\ \frac{dp_\nu}{d\lambda} - \frac{1}{2} \partial_\alpha g_{\beta\nu} p^\alpha p^\beta - \frac{1}{2} \partial_\beta g_{\alpha\nu} p^\alpha p^\beta + \frac{1}{2} \partial_\alpha g_{\nu\beta} p^\alpha p^\beta + \frac{1}{2} \partial_\beta g_{\nu\alpha} p^\alpha p^\beta - \frac{1}{2} \partial_\nu g_{\alpha\beta} p^\alpha p^\beta &= 0 \end{aligned}$$

where the first 4 terms cancel in pairs, since α and β are dummy indices, and we finally obtain

$$\frac{dp_\nu}{d\lambda} = \frac{1}{2} \partial_\nu g_{\alpha\beta} p^\alpha p^\beta$$

If $\partial_\nu g_{\alpha\beta} = 0$ for some index ν , this immediately implies that $\frac{dp_\nu}{d\lambda} = 0$. Hence, p_ν is constant along the geodesic.

- Physically, this corresponds to the existence of a Killing vector (a symmetry) along that coordinate. In simpler terms, $\partial_t g_{\alpha\beta} = 0$ reflects the invariance of the metric under time translations, yielding a conserved energy.

In our specific problem, the metric does not depend on the time coordinate t . Thus, $\partial_t g_{\alpha\beta} = 0$, and therefore p_t is conserved. We call this constant of motion

$$p_t = -E,$$

where E is interpreted as the energy of the photon.

Note on Notation: Writing $p_t = -E$ is a common convention. The minus sign is conventional so that E is considered a positive energy, consistent with the sign of g_{tt} .

5. Explicit Calculation of p_t . We can calculate p_t explicitly:

$$p_t = g_{t\alpha} p^\alpha = g_{tt} p^t = - \left(1 + \frac{x^2}{\ell^2} \right) \frac{dt}{d\lambda}.$$

Since $p_t = -E$ is constant along the photon geodesic, we have

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

This is our first conserved quantity: the photon's energy.

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:

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

Simplifying,

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

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.

- $p_t = -E$ is the conserved energy component, due to the time translation symmetry ($\partial_t g_{\alpha\beta} = 0$).
- From p_t , we get $\frac{dt}{d\lambda} = \frac{E}{1 + x^2/\ell^2}$.
- The null condition $ds^2 = 0$ then gives $\frac{dx}{d\lambda} = \pm \frac{E}{\sqrt{1 + x^2/\ell^2}}$.

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

Why We Used the Geodesic Equation for p_μ Conservation. One last remark: to show that p_t is constant, one needs to analyze how p_μ evolves along λ . The geodesic equation (in its covariant form) states:

$$\frac{dp_\nu}{d\lambda} = -\frac{1}{2} \partial_\nu g_{\alpha\beta} p^\alpha p^\beta.$$

If the metric does not depend on t , then $\partial_t g_{\alpha\beta} = 0$, which directly implies that $\frac{dp_t}{d\lambda} = 0$. Thus, p_t is a conserved quantity. This is the rigorous link between the symmetry of the metric ($\partial_t = 0$) and the constancy of p_t .

Exercise 10

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

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 $\Gamma_{\mu\nu}^\rho$ are defined by:

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

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

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

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_{rr}^r and $\Gamma_{r\phi}^\phi$.

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

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

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

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

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

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

so its derivative with respect to r is (using the derivative of nested functions)

$$\partial_r \left(1 - \frac{2GM}{r}\right)^{-1} = - \left(1 - \frac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2}.$$

Also note that

$$g^{rr} = \left(1 - \frac{2GM}{r}\right).$$

It may help to do the derivative step by step. Let $f(r) = 1 - \frac{2GM}{r}$. 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:

$$\Gamma_{rr}^r = \frac{1}{2} \left(1 - \frac{2GM}{r}\right) \left[- \left(1 - \frac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2} \right].$$

We can factor out the terms carefully:

$$\Gamma_{rr}^r = -\frac{1}{2} \left(1 - \frac{2GM}{r}\right) \left(1 - \frac{2GM}{r}\right)^{-2} \cdot \frac{2GM}{r^2} = -\frac{GM}{r^2} \left(1 - \frac{2GM}{r}\right)^{-1}.$$

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 $\Gamma_{r\phi}^\phi$. We set $\rho = \phi$, $\mu = r$, and $\nu = \phi$. From the definition, we have:

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

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

$$\Gamma_{r\phi}^\phi = \frac{1}{2} g^{\phi\phi} \partial_r g_{\phi\phi}.$$

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:

$$\Gamma_{r\phi}^\phi = \frac{1}{2} \left(\frac{1}{r^2 \sin^2 \theta} \right) (2r \sin^2 \theta) = \frac{1}{r}.$$

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:

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

These components play an important role in understanding geodesics in Schwarzschild spacetime. The Γ_{rr}^r component affects the radial part of a freely falling particle's motion, while $\Gamma_{r\phi}^\phi$ 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:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \quad (3)$$

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:

$$E = \left(1 - \frac{2GM}{r}\right) \frac{dt}{d\lambda}, \quad L = r^2 \frac{d\phi}{d\lambda}. \quad (4)$$

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

$$E = \left(1 - \frac{2GM}{r}\right) \frac{dt}{d\tau}, \quad L = r^2 \frac{d\phi}{d\tau}. \quad (5)$$

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:

$$u^\mu = \frac{dx^\mu}{d\tau} \quad (6)$$

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:

$$ds^2 = -d\tau^2 \quad (7)$$

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

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

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (8)$$

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

$$\frac{ds^2}{d\tau^2} = g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = g_{\mu\nu} u^\mu u^\nu \quad (9)$$

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

$$g_{\mu\nu} u^\mu u^\nu = -1. \quad (10)$$

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:

$$g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = -1. \quad (11)$$

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

$$g_{tt} \left(\frac{dt}{d\tau} \right)^2 + g_{rr} \left(\frac{dr}{d\tau} \right)^2 + g_{\phi\phi} \left(\frac{d\phi}{d\tau} \right)^2 = -1. \quad (12)$$

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

$$- \left(1 - \frac{2GM}{r} \right) \left(\frac{dt}{d\tau} \right)^2 + \left(1 - \frac{2GM}{r} \right)^{-1} \left(\frac{dr}{d\tau} \right)^2 + r^2 \left(\frac{d\phi}{d\tau} \right)^2 = -1. \quad (13)$$

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

$$E = \left(1 - \frac{2GM}{r} \right) \frac{dt}{d\tau}, \quad L = r^2 \frac{d\phi}{d\tau}. \quad (14)$$

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:

$$\frac{dt}{d\tau} = \frac{E}{1 - \frac{2GM}{r}}, \quad \frac{d\phi}{d\tau} = \frac{L}{r^2}. \quad (15)$$

Substituting these expressions into the normalization equation gives:

$$-\left(1 - \frac{2GM}{r}\right) \left(\frac{E}{1 - \frac{2GM}{r}}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1} \left(\frac{dr}{d\tau}\right)^2 + r^2 \left(\frac{L}{r^2}\right)^2 = -1. \quad (16)$$

Simplifying the equation:

$$-\frac{E^2}{1 - \frac{2GM}{r}} + \left(1 - \frac{2GM}{r}\right)^{-1} \left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2} = -1. \quad (17)$$

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

$$-E^2 + \left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2} \left(1 - \frac{2GM}{r}\right) = -\left(1 - \frac{2GM}{r}\right). \quad (18)$$

Expanding the last term on the left side:

$$-E^2 + \left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{r^2} - \frac{2GML^2}{r^3} = -1 + \frac{2GM}{r}. \quad (19)$$

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

$$\left(\frac{dr}{d\tau}\right)^2 = E^2 - 1 + \frac{2GM}{r} - \frac{L^2}{r^2} + \frac{2GML^2}{r^3}. \quad (20)$$

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

$$\frac{1}{2} \left(\frac{dr}{d\tau}\right)^2 = \frac{E^2}{2} - \frac{1}{2} + \frac{GM}{r} - \frac{L^2}{2r^2} + \frac{GML^2}{r^3}. \quad (21)$$

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

$$\frac{1}{2} \left(\frac{dr}{d\tau}\right)^2 + \left(-\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} + \frac{1}{2}\right) = \frac{E^2}{2}. \quad (22)$$

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

$$V_{\text{eff}}(r) = -\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} + \frac{1}{2}. \quad (23)$$

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:

$$V_{\text{eff}}(r) = \left(1 - \frac{2GM}{r}\right) \left(\frac{1}{2} + \frac{L^2}{2r^2}\right). \quad (24)$$

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,

$$V_{\text{eff}}(r) = \frac{1}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3}. \quad (25)$$

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

$$\frac{dr}{d\tau} = 0. \quad (26)$$

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):

$$\frac{dV_{\text{eff}}}{dr} = 0. \quad (27)$$

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

$$\frac{dV_{\text{eff}}}{dr} = \frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3GML^2}{r^4}. \quad (28)$$

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

$$\frac{GM}{r^2} - \frac{L^2}{r^3} + \frac{3GML^2}{r^4} = 0. \quad (29)$$

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 :

$$GMr^2 - L^2r + 3GML^2 = 0. \quad (30)$$

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:

$$r = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM}. \quad (31)$$

(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:

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

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 $r = 6GM$.