4 Differentiation and Integration

In this section we discuss differentiation and integration in curved spacetime. These might seem like delicate subjects but, given the tensor algebra that we have developed, tensor calculus is straightforward.

4.1 Gradient of a scalar

Consider first the gradient of a scalar field \( f_X \). We have already shown in Section 2 that the gradient operator \( \nabla \) is a one-form (an object that is invariant under coordinate transformations) and that, in a coordinate basis, its components are simply the partial derivatives with respect to the coordinates:

\[
\nabla f = (\partial_\mu f) \partial^\mu = (\nabla_\mu f) \partial^\mu ,
\]

where \( \partial_\mu = (\partial/\partial x^\mu) \). We have introduced a second notation, \( \nabla_\mu \), called the covariant derivative with respect to \( x^\mu \). By definition, the covariant derivative behaves like the component of a one-form. But, from equation (55), this is also true of the partial derivative operator \( \partial_\mu \). Why have we introduced a new symbol?

Before answering this question, let us first note that the gradient, because it behaves like a tensor of rank \((0,1)\) (a one-form), changes the rank of a tensor field from \((m,n)\) to \((m,n+1)\). (This is obviously true for the gradient of a scalar field, with \( m = n = 0 \).) That is, application of the gradient is like taking the tensor product with a one-form. The difference is that the components are not the product of the components, because \( \nabla_\mu \) is not a number. Nevertheless, the resulting object must be a tensor of rank \((m,n+1)\); e.g., its components must transform like components of a \((m,n+1)\) tensor. The gradient of a scalar field \( f \) is a \((0,1)\) tensor with components \((\partial_\mu f)\).
4.2 Gradient of a vector: covariant derivative

The reason that we have introduced a new symbol for the derivative will become clear when we take the gradient of a vector field \( \vec{A}_x = A^\mu x^\mu \).

In general, the basis vectors are functions of position as are the vector components! So, the gradient must act on both. In a coordinate basis, we have

\[
\nabla \vec{A} = \nabla (A^\nu e_\nu) = \bar{e}^\mu \partial_\mu (A^\nu e_\nu) = (\partial_\mu A^\nu) \bar{e}^\mu e_\nu + A^\nu \bar{e}^\mu (\partial_\mu e_\nu) = (\nabla_\mu A^\nu) \bar{e}^\mu e_\nu .
\]

(56)

We have dropped the tensor product symbol \( \otimes \) for notational convenience although it is still implied. Note that we must be careful to preserve the ordering of the vectors and tensors and we must not confuse subscripts and superscripts. Otherwise, taking the gradient of a vector is straightforward. The result is a \((1,1)\) tensor with components \( \nabla_\mu A^\nu \). But now \( \nabla_\mu \neq \partial_\mu \)! This is why we have introduced a new derivative symbol. We reserve the covariant derivative notation \( \nabla_\mu \) for the actual components of the gradient of a tensor. We note that the alternative notation \( A^\nu{}_{\mu} = \nabla_\mu A^\nu \) is often used, replacing the comma of a partial derivative \( A^\nu{}_{\mu} = \partial_\mu A^\nu \) with a semicolon for the covariant derivative. The difference seems mysterious only when we ignore basis vectors and stick entirely to components. As equation (56) shows, vector notation makes it clear why there is a difference.

Equation (56) by itself does not help us evaluate the gradient of a vector because we do not yet know what the gradients of the basis vectors are. However, they are straightforward to determine in a coordinate basis. First we note that, geometrically,
\[ \partial_\nu \bar{e}_\mu \text{ is a vector at } x: \text{ it is the difference of two vectors at infinitesimally close points, divided by a coordinate interval. (The easiest way to tell that } \partial_\nu \bar{e}_\mu \text{ is a vector is to note that it has one arrow!)} \text{ So, like all vectors, it must be a linear combination of basis vectors at } x. \text{ We can write the most general possible linear combination as} \]

\[ \partial_\nu \bar{e}_\mu x \equiv \Gamma^\lambda_{\mu\nu} x \bar{e}_\lambda x. \]  

### 4.3 Christoffel symbols

We have introduced in equation (57) a set of coefficients, \( \Gamma^\lambda_{\mu\nu} \), called the connection coefficients or Christoffel symbols. (Technically, the term Christoffel symbols is reserved for a coordinate basis.) It should be noted at the outset that, despite their appearance, the Christoffel symbols are not the components of a \((1, 2)\) tensor. Rather, they may be considered as a set of four \((1, 1)\) tensors, one for each basis vector \( \bar{e}_\mu \), because \( \nabla \bar{e}_\mu = \Gamma^\lambda_{\mu\nu} \bar{e}^\nu \bar{e}_\lambda. \) However, it is not useful to think of the Christoffel symbols as tensor components for fixed \( \nu \) because, under a change of basis, the basis vectors \( \bar{e}_\nu \) themselves change and therefore the four \((1, 1)\) tensors must also change. So, forget about the Christoffel symbols defining a tensor. They are simply a set of coefficients telling us how to differentiate basis vectors. Whatever their values, the components of the gradient of \( A \), known also as the covariant derivative of \( A^\nu \), are, from equations (56) and (57),

\[ \nabla_\mu A^\nu = \partial_\mu A^\nu + \Gamma^\nu_{\mu\lambda} A^\lambda. \] 

How does one determine the values of the Christoffel symbols? That is, how does one evaluate the gradients of the basis vectors? One way is to express the basis vectors in terms of another set whose gradients are known. For example, consider polar coordinates \((\rho, \theta)\) in the Cartesian plane as discussed in Section 2. The polar coordinate basis vectors were given in terms of the Cartesian basis vectors in equation (51). We know that the gradients of the Cartesian basis vectors vanish and we know how to transform from Cartesian to polar coordinates. It is a straightforward and instructive exercise from this to compute the gradients of the polar basis vectors:

\[ \nabla \bar{e}_\rho = \frac{1}{\rho} \bar{e}^\theta \otimes \bar{e}_\theta, \quad \nabla \bar{e}_\theta = \frac{1}{\rho} \bar{e}^\rho \otimes \bar{e}_\theta - \rho \bar{e}^\theta \otimes \bar{e}_\rho. \]  

(We have restored the tensor product symbol as a reminder of the tensor nature of the objects in eq. 59.) From equations (57) and (59) we conclude that the nonvanishing Christoffel symbols are

\[ \Gamma^\theta_{\theta\rho} = \Gamma^\theta_{\rho\theta} = \rho^{-1}, \quad \Gamma^\rho_{\theta\theta} = -\rho. \] 

It is instructive to extend this example further. Suppose that we add the third dimension, with coordinate \( z \), to get a three-dimensional Euclidean space with cylindrical
coordinates \((\rho, \theta, z)\). The line element (cf. eq. 50) now becomes \(ds^2 = d\rho^2 + \rho^2 d\theta^2 + dz^2\). Because \(\vec{e}_\rho\) and \(\vec{e}_\theta\) are independent of \(z\) and \(\vec{e}_z\) is itself constant, no new non-vanishing Christoffel symbols appear. Now consider a related but different manifold: a cylinder. A cylinder is simply a surface of constant \(\rho\) in our three-dimensional Euclidean space. This two-dimensional space is mapped by coordinates \((\theta, z)\), with basis vectors \(\vec{e}_\theta\) and \(\vec{e}_z\). What are the gradients of these basis vectors? They vanish! But, how can that be? From equation (59), \(\partial_\theta \vec{e}_\theta = -\rho \vec{e}_\rho\). Have we forgotten about the \(\vec{e}_\rho\) direction?

This example illustrates an important lesson. We cannot project tensors into basis vectors that do not exist in our manifold, whether it is a two-dimensional cylinder or a four-dimensional spacetime. A cylinder exists as a two-dimensional mathematical surface whether or not we choose to embed it in a three-dimensional Euclidean space. If it happens that we can embed our manifold into a simpler higher-dimensional space, we do so only as a matter of calculational convenience. If the result of a calculation is a vector normal to our manifold, we must discard this result because this direction does not exist in our manifold. If this conclusion is troubling, consider a cylinder as seen by a two-dimensional ant crawling on its surface. If the ant goes around in circles about the \(z\)-axis it is moving in the \(\vec{e}_\theta\) direction. The ant would say that its direction is not changing as it moves along the circle. We conclude that the Christoffel symbols indeed all vanish for a cylinder described by coordinates \((\theta, z)\).

### 4.4 Gradients of one-forms and tensors

Later we will return to the question of how to evaluate the Christoffel symbols in general. First we investigate the gradient of one-forms and of general tensor fields. Consider a one-form field \(\bar{A}_X = A_\mu X^\mu \bar{e}_X\). Its gradient in a coordinate basis is

\[ \nabla \bar{A} = \nabla(A_\mu \bar{e}^\nu) = \bar{e}^\nu \partial_\mu (A_\nu \bar{e}^\nu) = \left( \partial_\mu A_\nu \right) \bar{e}^\nu \bar{e}^\mu + A_\nu \bar{e}^\mu \left( \partial_\mu \bar{e}^\nu \right) = (\nabla_\nu A_\nu) \bar{e}^\mu \bar{e}^\nu. \]  

Again we have defined the covariant derivative operator to give the components of the gradient, this time of the one-form. We cannot assume that \(\nabla_\mu\) has the same form here as in equation (58). However, we can proceed as we did before to determine its relation, if any, to the Christoffel symbols. We note that the partial derivative of a one-form in equation (61) must be a linear combination of one-forms:

\[ \partial_\mu \bar{e}^\nu \equiv \Pi^\nu_{\lambda \mu} \bar{e}_X \bar{e}^\lambda_X, \]  

for some set of coefficients \(\Pi^\nu_{\lambda \mu}\) analogous to the Christoffel symbols. In fact, these coefficients are simply related to the Christoffel symbols, as we may see by differentiating the scalar product of dual basis one-forms and vectors:

\[ 0 = \partial_\mu \langle \bar{e}^\kappa, \bar{e}_\lambda \rangle = \Pi^\nu_{\kappa \mu} \langle \bar{e}^\kappa, \bar{e}_\lambda \rangle + \Gamma^\nu_{\lambda \mu} \langle \bar{e}^\kappa, \bar{e}_\kappa \rangle = \Pi^\nu_{\lambda \mu} + \Gamma^\nu_{\lambda \mu}. \]
We have used equation (13) plus the linearity of the scalar product. The result is
\[ \Pi_{\nu \lambda} = -\Gamma_{\nu \lambda}^\mu, \]
so that equation (62) becomes, simply,
\[ \partial_\mu \mathbf{e}_\nu = -\Gamma_{\nu \lambda}^\mu \mathbf{e}_\lambda. \quad (64) \]

Consequently, the components of the gradient of a one-form \( \tilde{A} \), also known as the covariant derivative of \( A_\nu \), are
\[ \nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma_{\nu \lambda}^\mu A_\lambda. \quad (65) \]

This expression is similar to equation (58) for the covariant derivative of a vector except for the sign change and the exchange of the indices \( \nu \) and \( \lambda \) on the Christoffel symbol (obviously necessary for consistency with tensor index notation). Although we still don’t know the values of the Christoffel symbols in general, at least we have introduced no more unknown quantities.

We leave it as an exercise for the reader to show that extending the covariant derivative to higher-rank tensors is straightforward. First, the partial derivative of the components is taken. Then, one term with a Christoffel symbol is added for every index on the tensor component, with a positive sign for contravariant indices and a minus sign for covariant indices. That is, for a \((m, n)\) tensor, there are \( m \) positive terms and \( n \) negative terms. The placement of labels on the Christoffel symbols is a straightforward extension of equations (58) and (65). We illustrate this with the gradients of the \((0, 2)\) metric tensor, the \((1, 1)\) identity tensor and the \((2, 0)\) inverse metric tensor:
\[ \nabla \mathbf{g} = \left( \nabla_{\lambda} g_{\mu \nu} \right) \mathbf{e}_\lambda \otimes \mathbf{e}_\mu \otimes \mathbf{e}_\nu, \quad \nabla_{\lambda} g_{\mu \nu} = \partial_{\lambda} g_{\mu \nu} - \Gamma_\mu^{\kappa \lambda} g_{\kappa \nu} - \Gamma_\nu^{\kappa \lambda} g_{\mu \kappa}, \quad (66) \]
\[ \nabla \mathbf{l} = \left( \nabla_{\lambda} \delta_{\mu \nu} \right) \mathbf{e}_\lambda \otimes \mathbf{e}_\mu \otimes \mathbf{e}_\nu, \quad \nabla_{\lambda} \delta_{\mu \nu} = \partial_{\lambda} \delta_{\mu \nu} + \Gamma_{\mu \kappa}^{\lambda \nu} \delta_{\kappa \nu} - \Gamma_{\nu \kappa}^{\lambda \nu} \delta_{\mu \kappa}, \quad (67) \]
and
\[ \nabla \mathbf{g}^{-1} = \left( \nabla_{\lambda} g^{\mu \nu} \right) \mathbf{e}_\lambda \otimes \mathbf{e}_\mu \otimes \mathbf{e}_\nu, \quad \nabla_{\lambda} g^{\mu \nu} = \partial_{\lambda} g^{\mu \nu} + \Gamma_{\mu \kappa}^{\lambda \nu} g^{\kappa \nu} + \Gamma_{\nu \kappa}^{\lambda \nu} g^{\mu \kappa}. \quad (68) \]

Examination of equation (67) shows that the gradient of the identity tensor vanishes identically. While this result is not surprising, it does have important implications. Recall from Section 2 the isomorphism between \( \mathbf{g} \), \( \mathbf{l} \) and \( \mathbf{g}^{-1} \) (eq. 48). As a result of this isomorphism, we would expect that all three tensors have vanishing gradient. Is this really so?

For a smooth (differentiable) manifold the gradient of the metric tensor (and the inverse metric tensor) indeed vanishes. The proof is sketched as follows. At a given point \( \mathbf{x} \) in a smooth manifold, we may construct a locally flat orthonormal (Cartesian) coordinate system. We define a \(\textit{locally flat coordinate system}\) to be one whose coordinate basis vectors satisfy the following conditions in a finite neighborhood around \( \mathbf{x} \): \( \mathbf{e}_\mu \mathbf{x} \cdot \mathbf{e}_\nu \mathbf{x} = 0 \) for \( \mu \neq \nu \) and \( \mathbf{e}_\mu \mathbf{x} \cdot \mathbf{e}_\nu \mathbf{x} = \pm 1 \), (with no implied summation).
The existence of a locally flat coordinate system may be taken as the definition of a smooth manifold. For example, on a two-sphere we may erect a Cartesian coordinate system $x^\tilde{\mu}$, with orthonormal basis vectors $\tilde{e}_\tilde{\mu}$, applying over a small region around $x$. (We use a bar to indicate the locally flat coordinates.) While these coordinates cannot, in general, be extended over the whole manifold, they are satisfactory for measuring distances in the neighborhood of $x$ using equation (42) with $g_{\tilde{\mu}\tilde{\nu}} = \eta_{\tilde{\mu}\tilde{\nu}} = \delta^{\tilde{\mu}\tilde{\nu}}$, where $\eta_{\tilde{\mu}\tilde{\nu}}$ is the metric of a flat space or spacetime with orthonormal coordinates (the Kronecker delta or the Minkowski metric as the case may be). The key point is that this statement is true not only at $x$ but also in a small neighborhood around it. (This argument relies on the absence of curvature singularities in the manifold and would fail, for example, if it were applied at the tip of a cone.) Consequently, the metric must have vanishing first derivative at $x$ in the locally flat coordinates: $\partial_\lambda g_{\tilde{\mu}\tilde{\nu}} = 0$. The gradient of the metric (and the inverse metric) vanishes in the locally flat coordinate basis. But, the gradient of the metric is a tensor and tensor equations are true in any basis. Therefore, for any smooth manifold,

$$\tilde{\nabla} g = \tilde{\nabla} g^{-1} = 0 . \tag{69}$$

### 4.5 Evaluating the Christoffel symbols

We can extend the argument made above to prove the symmetry of the Christoffel symbols: $\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}$, for any coordinate basis. At point $x$, the basis vectors corresponding to our locally flat coordinate system have vanishing derivatives: $\partial_\mu \tilde{e}_\nu = 0$. From equation (57), this implies that the Christoffel symbols vanish at a point in a locally flat coordinate basis. Now let us transform to any other set of coordinates $x^\mu$. The Jacobian of this transformation is $\Lambda^\mu_{\tilde{\mu}} = \partial x^\mu / \partial x^{\tilde{\mu}}$ (eq. 36). Our basis vectors transform (eq. 28) according to $\tilde{e}_\mu = \Lambda^\mu_{\tilde{\mu}} \tilde{e}_\nu$. We now evaluate $\partial_\mu \tilde{e}_\nu = 0$ using the new basis vectors, being careful to use equation (57) for their partial derivatives (which do not vanish in non-flat coordinates):

$$0 = \partial_\mu \tilde{e}_\nu = \frac{\partial^2 x^\kappa}{\partial x^{\tilde{\mu}} \partial x^{\tilde{\nu}}} \tilde{e}_\kappa + \frac{\partial x^\kappa}{\partial x^{\tilde{\mu}}} \frac{\partial x^\lambda}{\partial x^{\tilde{\nu}}} \Gamma^\sigma_{\kappa\lambda} \tilde{e}_\sigma = 0 . \tag{70}$$

Exchanging $\tilde{\mu}$ and $\tilde{\nu}$ we see that

$$\Gamma^\sigma_{\kappa\lambda} = \Gamma^\sigma_{\lambda\kappa} \quad \text{in a coordinate basis} , \tag{71}$$

implying that our connection is torsion-free (Wald 1984).

We can now use equations (66), (69) and (71) to evaluate the Christoffel symbols in terms of partial derivatives of the metric coefficients in any coordinate basis. We write $\nabla_\lambda g_{\mu\nu} = 0$ and permute the indices twice, combining the results with one minus sign and using the inverse metric at the end. The result is

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\kappa} (\partial_\mu g_{\kappa\nu} + \partial_\nu g_{\kappa\mu} - \partial_\kappa g_{\mu\nu}) \quad \text{in a coordinate basis} . \tag{72}$$
Although the Christoffel symbols vanish at a point in a locally flat coordinate basis, they do not vanish in general. This confirms that the Christoffel symbols are not tensor components: If the components of a tensor vanish in one basis they must vanish in all bases.

We can now summarize the conditions defining a locally flat coordinate system $x^\mu$ about point $x_0$: $g_{\mu\nu}x_0 = \eta_{\mu\nu}$ and $\Gamma^\mu_{\alpha\beta}x_0 = 0$ or, equivalently, $\partial_\mu g_{\nu\lambda}x_0 = 0$.

### 4.6 Transformation to locally flat coordinates

We have derived an expression for the Christoffel symbols beginning from a locally flat coordinate system. The problem may be turned around to determine a locally flat coordinate system at point $x_0$, given the metric and Christoffel symbols in any coordinate system. The coordinate transformation is found by expanding the components $g_{\mu\nu}x$ of the metric in the non-flat coordinates $x^\mu$ in a Taylor series about $x_0$ and relating them to the metric components $\eta_{\mu\nu}$ in the locally flat coordinates $x^\mu$ using equation (34):

$$g_{\mu\nu}x = g_{\mu\nu}x_0 + (x^\lambda - x_0^\lambda)\partial_\lambda g_{\mu\nu}x_0 + O(x - x_0)^2 = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x_\mu}\frac{\partial x^\nu}{\partial x_\nu} + O(x - x_0)^2. \quad (73)$$

Note that the partial derivatives of $\eta_{\mu\nu}$ vanish as do those of any correction terms to the metric in the locally flat coordinates at $x^\mu = x_0^\mu$. Equation (73) imposes the two conditions required for a locally flat coordinate system: $g_{\mu\nu}x_0 = \eta_{\mu\nu}$ and $\partial_\mu g_{\nu\lambda}x_0 = 0$. However, the second partial derivatives of the metric do not necessarily vanish, implying that we cannot necessarily make the derivatives of the Christoffel symbols vanish at $x_0$. Quadratic corrections to the flat metric are a manifestation of curvature. In fact, we will see that all the information about the curvature and global geometry of our manifold is contained in the first and second derivatives of the metric. But first we must see whether general coordinates $x^\mu$ can be transformed so that the zeroth and first derivatives of the metric at $x_0$ match the conditions implied by equation (73).

We expand the desired locally flat coordinates $x^\mu$ in terms of the general coordinates $x^\mu$ in a Taylor series about the point $x_0$:

$$x^\mu = x_0^\mu + A^\mu_\kappa(x^\kappa - x_0^\kappa) + B^\mu_{\kappa\lambda}(x^\kappa - x_0^\kappa)(x^\lambda - x_0^\lambda) + O(x - x_0)^3, \quad (74)$$

where $x_0^\mu$, $A^\mu_\kappa$, and $B^\mu_{\kappa\lambda}$ are all constants. We leave it as an exercise for the reader to show, by substituting equations (74) into equations (73), that $A^\mu_\kappa$ and $B^\mu_{\kappa\lambda}$ must satisfy the following constraints:

$$g_{\kappa\lambda}x_0 = \eta_{\mu\nu}A^\mu_\kappa A^\nu_{\lambda}, \quad B^\mu_{\kappa\lambda} = \frac{1}{2}A^\mu_\mu \Gamma^\mu_{\kappa\lambda}x_0. \quad (75)$$

If these constraints are satisfied then we have found a transformation to a locally flat coordinate system. It is possible to satisfy these constraints provided that the metric and
the Christoffel symbols are finite at $x_0$. This proves the consistency of the assumption underlying equation (69), at least away from singularities. (One should not expect to find a locally flat coordinate system centered on a black hole.)

From equation (75), we see that for a given matrix $A^{\mu}_{\kappa}$, $B^{\mu}_{\kappa\lambda}$ is completely fixed by the Christoffel symbols in our nonflat coordinates. So, the Christoffel symbols determine the quadratic corrections to the coordinates relative to a locally flat coordinate system. As for the $A^{\mu}_{\kappa}$ matrix giving the linear transformation to flat coordinates, it has 16 independent coefficients in a four-dimensional spacetime. The metric tensor has only 10 independent coefficients (because it is symmetric). From equation (75), we see that we are left with 6 degrees of freedom for any transformation to locally flat spacetime coordinates. Could these 6 have any special significance? Yes! Given any locally flat coordinates in spacetime, we may rotate the spatial coordinates by any amount (labeled by one angle) about any direction (labeled by two angles), accounting for three degrees of freedom. The other three degrees of freedom correspond to a rotation of one of the space coordinates with the time coordinate, i.e., a Lorentz boost! This is exactly the freedom we would expect in defining an inertial frame in special relativity. Indeed, in a locally inertial frame general relativity reduces to special relativity by the Equivalence Principle.