Spacetime curvature

Edoardo Milotti

October 17, 2023

1 The equivalence principle and the necessity of curved spacetime

Later in the course we shall explore in detail the meaning of the equivalence principle. For now, it suffices to say that it establishes the equivalence of inertial and gravitational mass. If this is so, mass simplifies in the equations of motion of different bodies that feel no forces other than gravity, therefore the equations do not depend on mass and all bodies "fall" in the same way in the gravitational field if launched from the same position with the same initial velocity: indeed this is the essence of the **geodesic hypothesis** (a geodesic is a path that has the shortest length between any two points), *a free particle follows a geodesic in spacetime.* Geodesics in spacetime are determined by the distribution of mass energy in the universe, as we can easily see by example, considering for instance the forces that act on an astronaut (Fig. 1), or the tides on Earth (Fig. 2).

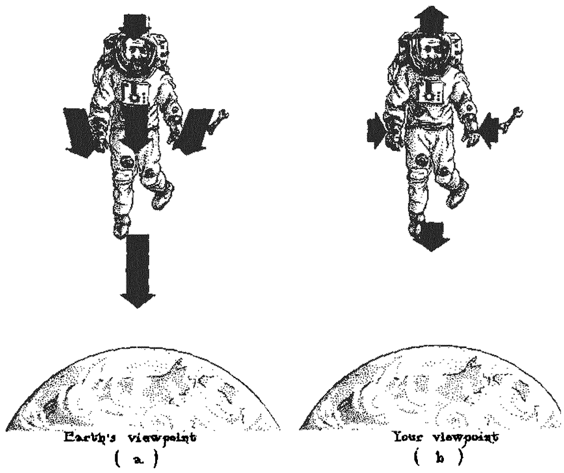

Figure 1: As you fall toward Earth, tidal gravitational forces stretch you from head to foot and squeeze you from the sides (taken from [4]).

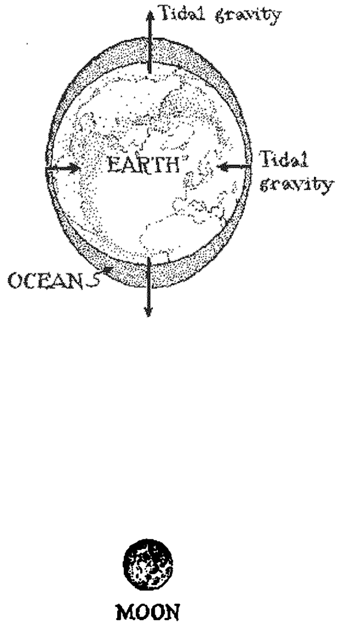

Figure 2: On the side of the Earth nearest the Moon, the lunar gravity is stronger than at the Earth's center, so it pulls the oceans toward the Moon more strongly than it pulls the solid Earth, and the oceans in response stretch outward a bit toward the Moon. On the side farthest from the Moon, the lunar gravity is weaker, so it pulls the oceans toward the Moon less strongly than it pulls the solid Earth, and the oceans in response stretch out away from the Moon. On the left side of the Earth, the Moon's gravitational pull, which points toward the Moon's center, has a slight rightward component, and on the right side it has a slight leftward component; and these components squeeze the oceans inward. This pattern of oceanic stretch and squeeze produces two high tides and two low tides each day, as the Earth rotates (taken from [4]).

Fig. 3(a) shows the same situation in a more schematic way: the freely-falling lab frame is represented by a box with four floating balls. The frame falls towards the Earth, and because of the variable intensity and direction of gravity on each ball, the balls appear to move with respect to the box's center.

When we trace the trajectories of the balls in Fig. 3(a) in a spacetime diagram, we find that they are curved, see Fig. 4

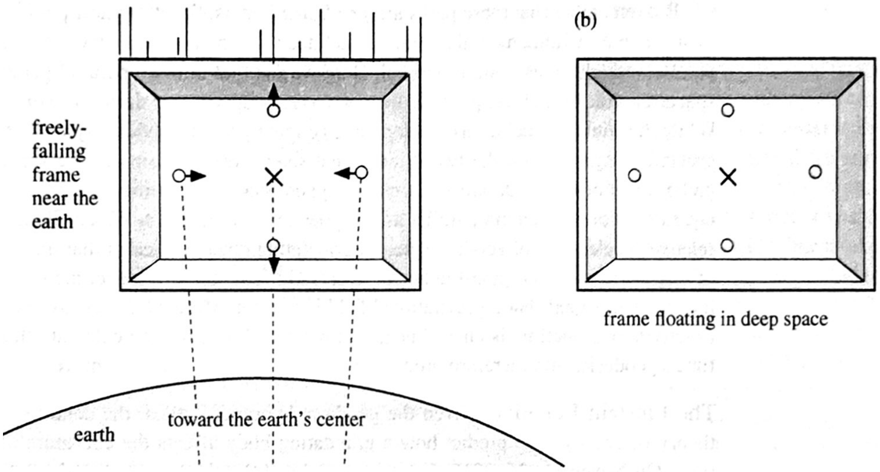

Figure 3: (a) Because the gravitational field of the earth (indeed, any gravitating object) is non-uniform, off-center floating balls in a freely falling frame will experience small accelerations relative to the frame's center of mass. (b) Such accelerations are not observed in a frame floating in deep space: floating balls initially at rest remain truly at rest (taken from [3]).

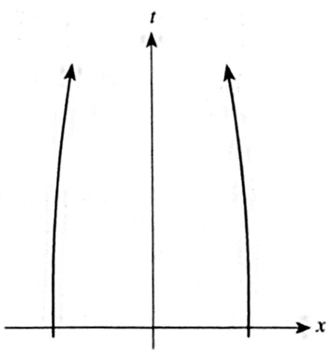

Figure 4: Plotted in spacetime, the geodesics that the balls in figure 3(a) follow (as measured in the freely-falling frame) are initially parallel (the balls have initially constant separation), but gradually bend toward each other (because their separation eventually decreases). This bending of initial parallel lines signals that the underlying spacetime is curved (taken from [3]).

2 Riemann manifolds

We found that the equivalence principle (in essence, the indistinguishability between gravitational pull and dynamical acceleration) implies that the effects of gravity must be associated with *a curvature of space-time*. Mathematically, this means that we must turn to the manifolds studied by Riemann. Again, I take some material from [1] ch. 4 (grayed text), and [2] ch. 1 (blue text).

In a nutshell:

- A Riemann manifold can be loosely described as a smoothly curved space that is locally flat.
- A manifold is n-dimensional when the position of a point is specified by n coordinates.
- Not all continuous spaces are manifolds. E.g., a one-dimensional line emerging from a plane is not a manifold; two cones joined at the apex are not a manifold (parts of these objects are not locally Euclidean).
- A manifold can be embedded in a larger space and display **extrinsic curvature**.
- A manifold can exist without any embedding at all and display **intrinsic curvature**.
- A manifold can display both extrinsic and intrinsic curvature.
- Manifolds exist that have extrinsic curvature and no intrinsic curvature.
- Riemann discovered that the *metric tensor* $g_{\mu\nu} = \mathbf{e}_\mu \cdot \mathbf{e}_\nu$ contains all the information necessary to describe a manifold.
- The metric tensor is symmetric, i.e., $g_{\mu\nu} = g_{\nu\mu}$.
- Riemann manifolds have positive-definite metrics; the metric of space-time is not positive-definite, and it is described by a pseudo-Riemann manifold.
- The line element of a Riemann manifold is given by $d\ell^2 = g_{ij} dx^i dx^j$.
- The description of Riemann manifolds requires tensor calculus.
- The manifolds of General Relativity are pseudo-Riemann because the metric is not positive-definite.

2.1 Manifolds

The model for the spacetime of general relativity makes use of a certain kind of four-dimensional manifold, so I need to explain what this involves. In doing this, I shall not give a precise mathematical definition, but rather explain and describe the properties of an N-dimensional manifold. I assume that this manifold is endowed with a *metric tensor field* (which is not a general requirement of manifolds) and explain how this is used to define and handle metric properties. I shall be guided by the notation and terminology developed when considering arbitrary curvilinear coordinates in Euclidean space and parameterized surfaces. What makes a manifold N-dimensional is that points in it can be labeled by a system of N real coordinates $x^1, x^2, \ldots, x^N$, in such a way that the correspondence between the points and the labels is one-to-one. We do not require that the whole of the manifold M should be covered by one system of coordinates, nor do we regard any one system as in some way preferred. The general situation is that we have a collection of coordinate systems, each covering some part of M, and all these are on an equal footing. Where two coordinate systems overlap, there are sets of equations giving each coordinate of one system as a function of the coordinates of the other. So if the coordinates x^i cover the region U and the coordinates x'^i cover the region U', where these are overlapping regions, then the coordinates of points in the overlap are related by equations of the form

$$x'^i = x'^i(x^1, \ldots, x^N) \quad (i = 1, \ldots, N) \tag{1}$$

giving each x'^i as a function of the coordinates x^i, and these have inverses of the form

$$x^i = x^i(x'^1, \ldots, x'^N) \quad (i = 1, \ldots, N) \tag{2}$$

giving each x^i as a function of the coordinates x'^i. We shall assume that the functions involved are differentiable so that the partial derivatives

$$\frac{\partial x'^i}{\partial x^j}, \quad \frac{\partial x^i}{\partial x'^j} \tag{3}$$

exist. This means that the manifold M is a *differentiable manifold* (mathematicians allow manifolds in which the coordinate-transformation functions are merely continuous and call them *topological manifolds*.)

3 The metric tensor

The distinction between contravariant and covariant transformations was established earlier, when I also observed that it only became meaningful when working with coordinate systems that are not Cartesian. Now I examine relationships that can systematize the use of more general metric spaces (the Riemannian spaces). For illustration purposes, I start with three dimensions.

Letting q^i denote coordinates in a general coordinate system, writing the index as a superscript to reflect the fact that coordinates transform contravariantly, I define covariant basis vectors $\boldsymbol{\varepsilon}_i$ that describe the displacement (in Euclidean space) per unit change in q^i , keeping the other q^j constant. For the situations of interest here, both the direction and magnitude of $\boldsymbol{\varepsilon}_i$ may be functions of position, so it is defined as the derivative

$$\boldsymbol{\varepsilon}_i = \frac{\partial x}{\partial q^i}\hat{\mathbf{e}}_x + \frac{\partial y}{\partial q^i}\hat{\mathbf{e}}_y + \frac{\partial z}{\partial q^i}\hat{\mathbf{e}}_z \tag{4}$$

An arbitrary vector $\mathbf{A}$ can now be formed as a linear combination of the basis vectors, multiplied by coefficients:

$$\mathbf{A} = A^1\boldsymbol{\varepsilon}_1 + A^2\boldsymbol{\varepsilon}_2 + A^3\boldsymbol{\varepsilon}_3 \tag{5}$$

At this point we have a linguistic ambiguity: $\mathbf{A}$ is a fixed object (usually called a vector) that may be described in various coordinate systems. But it is also customary to call the collection of coefficients A^i a vector (more specifically, a contravariant vector), while we have already called $\boldsymbol{\varepsilon}_i$ a covariant basis vector. The important thing to observe here is that $\mathbf{A}$ is a fixed object that is not changed by our transformations, while its representation (the A^i) and the basis used for the representation (the $\boldsymbol{\varepsilon}_i$) change in mutually inverse ways (as the coordinate system is changed) so as to keep $\mathbf{A}$ fixed.

Given our basis vectors, we can compute the displacement (change in position) associated with changes in the q^i. Because the basis vectors depend on position, our computation needs to be for small (infinitesimal) displacements ds. We have

$$ds^2 = (dq^i\boldsymbol{\varepsilon}_i)\cdot(dq^j\boldsymbol{\varepsilon}_j) = g_{ij}dq^idq^j \tag{6}$$

where we have used the *metric tensor*

$$g_{ij} = \boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_j \tag{7}$$

Since ds^2 is an invariant under rotational (and reflection) transformations, it is a scalar, and the quotient rule permits us to identify g_{ij} as a covariant tensor. Because of its role in defining displacement, g_{ij} is called the **covariant metric tensor**.

Note that the basis vectors can be defined by their Cartesian components, but they are, in general, neither unit vectors nor mutually orthogonal. Because they are often not unit vectors we have identified them by the symbol $\boldsymbol{\varepsilon}$, not $\hat{\mathbf{e}}$. The lack of both a normalization and an orthogonality requirement means that g_{ij}, though manifestly symmetric, is not required to be diagonal, and its elements (including those on the diagonal) may be of either sign. It is convenient to define a **contravariant metric tensor** that satisfies

$$g_{ik}g^{kj} = g^{jk}g_{ki} = \delta_i^j \tag{8}$$

and is therefore the inverse of the covariant metric tensor. We will use g_{ij} and g^{ij} to make conversions between contravariant and covariant vectors that we then regard as related. Thus, we write

$$g_{ij}F^j = F_i, \quad g^{ij}F_j = F^i \tag{9}$$

Following these rules, we can manipulate the representation of a vector

$$\mathbf{A} = A^i \boldsymbol{\varepsilon}_i = A^i \delta_i^k \boldsymbol{\varepsilon}_k = A^i g_{ij} g^{jk} \boldsymbol{\varepsilon}_k = A_j \boldsymbol{\varepsilon}^j, \tag{10}$$

showing that the same vector can be represented either by contravariant or covariant components.

4 Covariant and contravariant bases

Above, we have already met the contravariant basis vectors

$$\boldsymbol{\varepsilon}^i = \frac{\partial q^i}{\partial x}\hat{\mathbf{e}}_x + \frac{\partial q^i}{\partial y}\hat{\mathbf{e}}_y + \frac{\partial q^i}{\partial z}\hat{\mathbf{e}}_z, \tag{11}$$

giving them this name in anticipation of the fact that we can prove them to be the contravariant versions of the $\boldsymbol{\varepsilon}_i$. Our first step in this direction is to verify that

$$\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}_j = \frac{\partial x}{\partial q^j}\frac{\partial q^i}{\partial x} + \frac{\partial y}{\partial q^j}\frac{\partial q^i}{\partial y} + \frac{\partial z}{\partial q^j}\frac{\partial q^i}{\partial z} = \frac{\partial q^i}{\partial q^j} = \delta_j^i, \tag{12}$$

a consequence of the chain rule and the fact that q^i and q^j are independent variables. Similarly, we can easily show that

$$(\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}^k)(\boldsymbol{\varepsilon}_k \cdot \boldsymbol{\varepsilon}^j) = g^{ik} g_{kj} = \delta_j^i. \tag{13}$$

The relation between the covariant and contravariant basis vectors is useful for writing relationships between vectors. Let **A** and **B** be vectors with contravariant representations A^i and B^i. We may convert the representation of **B** to $B_i = g_{ij}B^j$, after which the scalar product $\mathbf{A} \cdot \mathbf{B}$ takes the form

$$\mathbf{A} \cdot \mathbf{B} = A^i \boldsymbol{\varepsilon}_i \cdot B_j \boldsymbol{\varepsilon}^j = A^i B_j \boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}^j = A^i B_i \tag{14}$$

Another application is in writing the gradient in general coordinates. If a function ψ is given in a general coordinate system q^i, its gradient $\nabla\psi$ is a vector with Cartesian components

$$(\nabla\psi)_j = \frac{\partial\psi}{\partial q^i}\frac{\partial q^i}{\partial x^j} \tag{15}$$

or, in vector notation

$$\nabla\psi = \frac{\partial\psi}{\partial q^i}\boldsymbol{\varepsilon}^i \tag{16}$$

showing that the covariant representation of $\nabla\psi$ is the set of derivatives $\partial\psi/\partial q^i$. If we have reason to use a contravariant representation of the gradient, we can convert its components using the metric tensor.

5 Covariant derivatives

Moving on to the derivatives of a vector, we find that the situation is much more complicated because the basis vectors $\boldsymbol{\varepsilon}_i$ are in general not constant, and the derivative will not be a tensor whose components are the derivatives of the vector components. Starting from the transformation rule for a contravariant vector,

$$V'^i = \frac{\partial x^i}{\partial q^k} V^k \tag{17}$$

and differentiating with respect to q^j, we get (for each i)

$$\frac{\partial V'^i}{\partial q^j} = \frac{\partial x^i}{\partial q^k}\frac{\partial V^k}{\partial q^j} + \frac{\partial^2 x^i}{\partial q^j \partial q^k} V^k \tag{18}$$

which appears to differ from the transformation law for a second-rank tensor because it contains a second derivative. To see what to do next, let's write Eq. (18) as a single vector equation in the x^i coordinates, which we take to be Cartesian. The result is

$$\frac{\partial \mathbf{V}'}{\partial q^j} = \frac{\partial V^k}{\partial q^j}\boldsymbol{\varepsilon}_k + V^k \frac{\partial \boldsymbol{\varepsilon}_k}{\partial q^j} \tag{19}$$

We now recognize that $\partial \boldsymbol{\varepsilon}_k / \partial q^j$ must be some vector in the space spanned by the set of all $\boldsymbol{\varepsilon}_i$ and we therefore write

$$\frac{\partial \boldsymbol{\varepsilon}_k}{\partial q^j} = \Gamma^{\mu}_{jk}\boldsymbol{\varepsilon}_\mu \tag{20}$$

The quantities Γ^{μ}_{jk} are known as **Christoffel symbols of the second kind** (those of the first kind will be encountered shortly) or **connection coefficients**. Using the orthogonality property of the $\boldsymbol{\varepsilon}$, Eq. (12), we can solve Eq. (20) by taking its dot product with any $\boldsymbol{\varepsilon}^m$, finding

$$\boldsymbol{\varepsilon}^m \cdot \frac{\partial \boldsymbol{\varepsilon}_k}{\partial q^j} = \boldsymbol{\varepsilon}^m \cdot \Gamma^{\mu}_{jk}\boldsymbol{\varepsilon}_\mu = \delta^m_\mu \Gamma^{\mu}_{jk} = \Gamma^m_{jk} \tag{21}$$

i.e.,

$$\Gamma^m_{jk} = \boldsymbol{\varepsilon}^m \cdot \frac{\partial \boldsymbol{\varepsilon}_k}{\partial q^j}. \tag{22}$$

Moreover, we note that $\Gamma^m_{jk} = \Gamma^m_{kj}$, which can be demonstrated from the expression

$$\frac{\partial \boldsymbol{\varepsilon}_i}{\partial q^j} = \frac{\partial^2 x}{\partial q^i \partial q^j}\hat{\mathbf{e}}_x + \frac{\partial^2 y}{\partial q^i \partial q^j}\hat{\mathbf{e}}_y + \frac{\partial^2 z}{\partial q^i \partial q^j}\hat{\mathbf{e}}_z \tag{23}$$

which is clearly symmetric with respect to the exchange of i and j.

Returning now to Eq. (19) and inserting the expression (20), we initially get

$$\frac{\partial \mathbf{V}'}{\partial q^j} = \frac{\partial V^k}{\partial q^j}\boldsymbol{\varepsilon}_k + V^k \Gamma^{\mu}_{jk}\boldsymbol{\varepsilon}_\mu = \frac{\partial V^k}{\partial q^j}\boldsymbol{\varepsilon}_k + V^\mu \Gamma^{k}_{j\mu}\boldsymbol{\varepsilon}_k = \left(\frac{\partial V^k}{\partial q^j} + V^\mu \Gamma^k_{j\mu}\right)\boldsymbol{\varepsilon}_k \tag{24}$$

The parenthesized quantity in Eq. (24) is known as the **covariant derivative** – sometimes also called **absolute gradient** – of V , and it has (unfortunately) become standard to identify it by the awkward notation

$$V^k_{;j} = \frac{\partial V^k}{\partial q^j} + V^\mu \Gamma^k_{j\mu} \quad \frac{\partial \mathbf{V}'}{\partial q^j} = V^k_{;j} \varepsilon_k \tag{25}$$

An alternative symbol for the covariant derivative is $\nabla_j V^k = V^k_{;j}$ (preferred in these handouts).

If we rewrite Eq. (24) as a differential

$$d\mathbf{V}' = \left(dq^j \ \nabla_j V^k\right) \varepsilon_k \tag{26}$$

and take note that dq^j is a contravariant vector, while ε_k is covariant, we see that the covariant derivative, $\nabla_j V^k$ is a mixed second-rank tensor. **However, it is important to realize that although they bristle with indices, neither $\partial V^k \partial q^j$ nor $\Gamma^k_{j\mu}$ have individually the correct transformation properties to be tensors**. It is only the combination in Eq. (25) that has the requisite transformational attributes.

What about the covariant derivative of a covariant vector? We can determine its expression by taking the derivative of a scalar quantity

$$\nabla_k(A^i B_i) = (\nabla_k A^i)B_i + A^i(\nabla_k B_i) = \left(\frac{\partial A^i}{\partial q^k} + A^m \Gamma^i_{km}\right) B_i + A^i(\nabla_k B_i), \tag{27}$$

and remarking that since $A^i B_i$ is a scalar quantity with a representation which does not depend on the local basis vectors, and therefore the covariant derivative is the same as the gradient, i.e.,

$$\nabla_k(A^i B_i) = \partial_k(A^i B_i) = \frac{\partial A^i}{\partial q^k} B_i + A^i \frac{\partial B_i}{\partial q^k}, \tag{28}$$

where ∂_k is shorthand for the derivative with respect to q^k. Then, we find

$$\left(\frac{\partial A^i}{\partial q^k} + A^m \Gamma^i_{km}\right) B_i + A^i(\nabla_k B_i) = \frac{\partial A^i}{\partial q^k} B_i + A^i \frac{\partial B_i}{\partial q^k}. \tag{29}$$

After simplifying the term common to both sides, we find

$$A^i \Gamma^m_{ki} B_m + A^i \nabla_k B_i = A^i \frac{\partial B_i}{\partial q^k}, \tag{30}$$

where the two dummy indices in the first term have been exchanged. This means that

$$A^i \left(\Gamma^m_{ki} B_m + \nabla_k B_i - \frac{\partial B_i}{\partial q^k}\right) = 0. \tag{31}$$

Since A^i is an arbitrary vector, the expression in parenthesis must vanish, and therefore

$$\nabla_k B_i = \frac{\partial B_i}{\partial q^k} - B_m \Gamma^m_{ki}; \tag{32}$$

the expression of the covariant derivative is the same as that of a contravariant vector, except for the sign of the additional term. Like $\nabla_j V^i$, $\nabla_j V_i$ is also a second-rank tensor.

This proof can be extended to show that the covariant derivative of a tensor contains as many additional positive terms with the Christoffel symbol as there are contravariant indices and as many additional terms as there are covariant indices. For instance,

$$\nabla_i T^{jk}_\ell = \frac{\partial T^{jk}_\ell}{\partial q^i} + \Gamma^j_{im} T^{mk}_\ell + \Gamma^k_{im} T^{jm}_\ell - \Gamma^m_{i\ell} T^{jk}_m \tag{33}$$

The physical importance of the covariant derivative is that it includes the changes in the basis vectors pursuant to a general dq^j, and is therefore more appropriate for describing physical phenomena than a formulation that considers only the changes in the coefficients multiplying the basis vectors.

6 Evaluation of the Christoffel symbols

Although a quantity with three indices in 4-dimensional space can have as may as 64 independent components, the connection coefficients have symmetrical lower indices, and therefore, for a fixed upper index, they have 10 components. Finally, taking into account the 4 possibile values of the upper index, there are 40 independent values of the connection coefficients in 4-dimensional space. It is easy to see that the number of independent components reduces to 6 in 2-dimensional space and to 18 in 3-dimensional space.

It may be more convenient to evaluate the Christoffel symbols by relating them to the metric tensor than simply to use Eq. (22). As an initial step in this direction, we define the **Christoffel symbol of the first kind** $[ij,k]$ by

$$[ij,k] = g_{mk} \Gamma^m_{ij} \tag{34}$$

from which the symmetry $[ij,k] = [ji,k]$ follows. Again, this $[ij,k]$ is not a third-rank tensor. Inserting Eq. (22) and applying the index-lowering transformation, we have

$$[ij,k] = g_{mk} \boldsymbol{\varepsilon}^m \cdot \frac{\partial \boldsymbol{\varepsilon}_i}{\partial q^j} = \boldsymbol{\varepsilon}_k \cdot \frac{\partial \boldsymbol{\varepsilon}_i}{\partial q^j} \tag{35}$$

Next, we write $g_{ij} = \boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_j$ as we did earlier and differentiate it, identifying the result with the aid of Eq. (35):

$$\frac{\partial g_{ij}}{\partial q^k} = \frac{\partial \boldsymbol{\varepsilon}_i}{\partial q^k} \cdot \boldsymbol{\varepsilon}_j + \boldsymbol{\varepsilon}_i \cdot \frac{\partial \boldsymbol{\varepsilon}_j}{\partial q^k} = [ik,j] + [jk,i] \tag{36}$$

Changing the index set, we find

$$\frac{\partial g_{ik}}{\partial q^j} = [ij,k] + [jk,i] \tag{37a}$$

$$\frac{\partial g_{jk}}{\partial q^i} = [ik,j] + [ij,k] \tag{37b}$$

Therefore, combining the last three equations, we find the identity

$$[ij,k] = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}\right) \tag{38}$$

Finally, using the definition of $[ij,k] = g_{mk}\Gamma^m ij$, we obtain

$$\Gamma^n_{ij} = g^{nk}[ij,k] = \frac{1}{2}g^{nk}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}\right) \tag{39}$$

$$= \frac{1}{2}g^{nk}\left(\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij}\right) \tag{40}$$

where the last line contains the shorthand notation for the partial derivative.

The apparatus of this subsection becomes unnecessary in Cartesian coordinates, because the basis vectors have vanishing derivatives, and the covariant and ordinary partial derivatives then coincide.

References

[1] GB Arfken, HJ Weber, and FE Harris. *Mathematical Methods for Physicists: A Comprehensive Guide. 7th Edition.* Academic Press, 2012.

[2] James Foster and J David Nightingale. *A short course in General Relativity.* Springer, 1995.

[3] Thomas Andrew Moore. *A general relativity workbook.* University Science Books Mill Valley, 2013.

[4] Kip Thorne. *Black Holes & Time Warps: Einstein's Outrageous Legacy (Commonwealth Fund Book Program).* WW Norton & Company, 1995.