Covariant differentiation

Vectors in curved spaces

In our everyday life—which spatially is largely restricted to a curved two-dimensional metric space: the surface of Earth—a vector is a small arrow pointing in a certain direction. For example, the arrow on a road-sign indicating the direction toward the nearest town.

The important property of the vector is that although it might point at different direction relative to your car (depending on your driving direction) it always points in the direction of the town. In other words a vector has certain useful transformational properties when one changes ones reference frames.

Geometrically an arrow can be described by specifying the coordinates of the nock and the head the arrow. Actually one only needs the coordinate differences between these two points as the absolute position of the road-sign is not important (as long as you can still see it from the car on the road).

The arrow must obviously be small compared to (at least) the relief structures on the Earth's surface. The smallness of the arrow in the mathematical language means that the coordinate difference between the nock and the head of the arrow is infinitesimal. In other words,

A vector—in the sense of a small pointing arrow—is a coordinate differential.

In physics we use vectors to describe many different physical quantities but they are all related to small arrows in the sense that under a change of the reference frame they all transform as coordinate differentials.

Coordinates and vectors

A set of four numbers used to specify the location of an event in space-time[1] is called *coordinates* and is denoted as $\{x^0, x^1, x^2, x^3\}$, or as x^a, or simply as x.

Under a general non-linear coordinate transformation $x \to x'(x)$ the coordinate differentials $dx \to dx'$ transform linearly,

$$dx'^a = \sum_{b=0}^{3} \frac{\partial x'^a}{\partial x^b} dx^b \ , \tag{1}$$

through the Jacobian matrix

$$J \equiv \frac{\partial x'^a}{\partial x^b} \ . \tag{2}$$

There exist another object that transforms linearly: the set of partial derivatives of a scalar function $\phi(x)$,

$$\frac{\partial \phi}{\partial x'^a} = \sum_{b=0}^{3} \frac{\partial \phi}{\partial x^b} \frac{\partial x^b}{\partial x'^a} \ , \tag{3}$$

which transforms through the inverse Jacobian matrix[2]

$$J^{-1} = \frac{\partial x^b}{\partial x'^a} \ . \tag{4}$$

[1] For example, *time*, *latitude*, *longitude*, and *height* above the sea level.

[2] Indeed,

$$JJ^{-1} = \sum_{b=0}^{3} \frac{\partial x^a}{\partial x'^b} \frac{\partial x'^b}{\partial x^c} = \frac{\partial x^a}{\partial x^c} = \delta^a_c \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ .$$

A set of four quantities $A^a \equiv \{A^0, A^1, A^2, A^3\}$, is called a *vector with index up*[3] if it transforms as coordinate differentials in equation (1),

$$A'^a = \sum_{b=0}^{3} \left(\frac{\partial x'^a}{\partial x^b}\right) A^b \ . \tag{5}$$

A set of four quantities A_a is called a *vector with index down* if it transforms as derivatives of the scalar in equation (3),

$$A'_a = \sum_{b=0}^{3} \left(\frac{\partial x^b}{\partial x'^a}\right) A_b \ . \tag{6}$$

In the following we shall use the *implicit summation notation* where we drop the summation sign and always assume a summation over the index that appears in the up and down positions in a term,

$$\boxed{A^a B_a \equiv \sum_{a=0}^{3} A^a B_a \ .} \tag{7}$$

Quantities which transform as vectors or their products—that is, linearly through Jacobi matrix or its inverse—are said to transform *covariantly.* These quantities are called *tensors.* The number of the indices is called the *rank* of the tensor.

The *contraction* $A^a B_a$ is apparently a scalar (tensor with rank zero), as it is invariant under coordinate transformations. Indeed,

$$A^a B_a = \frac{\partial x^a}{\partial x'^b} A'^b \frac{\partial x'^c}{\partial x^a} B'_c = A'^b B'_b \ . \tag{8}$$

Tensors

Objects which transform as products of vectors are called tensors. For example, the product $A^a B^b$ transforms as

$$A'^a B'^b = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} A^c B^d \ . \tag{9}$$

Now any quantitiy F^{ab} which transforms similarly,

$$F'^{ab} = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} F^{cd} \ . \tag{10}$$

is called a (rank-2, in this case) tensor.

Metric tensor

The *metric tensor* g_{ab} defines the metric—the invariant infinitesimal interval ds^2—between two close events in time-space separated by dx^a,

$$ds^2 = g_{ab} dx^a dx^b \ . \tag{11}$$

Since dx^a is an arbitrary index-up vector, the construction $g_{ab} dx^b$ transforms as an index-down vector and thus the metric tensor connects index-up and index-down vectors,

$$A_a = g_{ab} A^b \ . \tag{12}$$

[3]Also called contra-variant vector.

Covariant differential

In curvilinear coordinates the ordinary differential of a vector, dA^a, is not a covariant quantity, since generally

$$dA_a = d(g_{ab}A^b) = g_{ab}dA^b + dg_{ab}A^b \neq g_{ab}dA^b \ , \tag{13}$$

if $dg_{ab} \neq 0$, that is, if the metric tensor is not constant throughout time-space.

In order to write covariant differential equations we need a *covariant differential*, denoted as DA^a, which satisfies the covariance condition,

$$DA_a = g_{ab}DA^b = Dg_{ab}A^b. \tag{14}$$

The covariant differential has to contain dA^a and an additional term which would ensure covariance. As a contribution to differential the additional term must be linear in dx. Moreover, it must also be linear in A if we demand linearity of the operation of differentiation. The covariant differential can then be generally written as

$$DA^a = dA^a + \Gamma^a_{bc}A^b dx^c \ , \tag{15}$$

where the factor Γ^a_{bc} (apparently, not itself a covariant tensor) is called the *Christoffel symbol.*

Differential of a scalar $d(A_aB^a)$ is already a covariant quantity, therefore

$$D(A_aB^a) = d(A_aB^a) \tag{16}$$

for an arbitrary B^a. If we demand that covariant differential satisfies the Leibniz rule[4] we find[5]

$$DA_a = dA_a - \Gamma^b_{ac}A_b dx^c \ . \tag{17}$$

With the comma-notation for partial derivatives,

$$dA^a = \frac{\partial A^a}{\partial x^c}dx^c \equiv A^a_{,c}dx^c \ , \tag{18}$$

the covariant differentials take the form

$$DA^a = \left(A^a_{,c} + \Gamma^a_{bc}A^b\right)dx^c \ , \tag{19}$$

$$DA_a = \left(A_{a,c} - \Gamma^b_{ac}A_b\right)dx^c \ . \tag{20}$$

The expressions in the parentheses are tensors, since when multiplied by arbitrary vector dx^c they produce vectors. These tensors are called *covariant derivatives* and are denoted as

$$A^a_{;c} = A^a_{,c} + \Gamma^a_{bc}A^b \ , \tag{21}$$

$$A_{a;c} = A_{a,c} - \Gamma^b_{ac}A_b \ . \tag{22}$$

[4] The Leibniz rule states that

$$D(AB) = A\,DB + B\,DA \ .$$

[5] Applying the Leibniz rule to (16) gives

$$DA_aB^a + A_a(dB^a + \Gamma^a_{bc}B^b dx^c) = dA_aB^a + A_a dB^a \ .$$

Simplifying and exchanging $a \leftrightarrow b$ in the Γ-term gives

$$DA_aB^a = (dA_a - \Gamma^b_{ac}A_b dx^c)B^a \ .$$

Since B^a is arbitrary, $DA_a = dA_a - \Gamma^b_{ac}A_b dx^c$.

Christoffel symbols and the metric tensor

Although the Christoffel symbol itself is not a tensor, the linear combination

$$\Gamma^a_{bc} - \Gamma^a_{cb} \tag{23}$$

is a tensor. Indeed, consider the difference,

$$\varphi_{;c;b} - \varphi_{;b;c} = (\Gamma^a_{bc} - \Gamma^a_{cb})\varphi_{;a} \ , \tag{24}$$

where φ is an arbitrary scalar. The left-hand side is a tensor and $\varphi_{;a}$ is also a tensor, therefore $\Gamma^a_{bc} - \Gamma^a_{cb}$ is also a tensor. The latter is equal zero in a locally flat frame. Being a tensor it is than equal zero in all other frames. Therefore Christoffel symbol is symmetric[6] over exchange of the two lower indexes,

$$\Gamma^a_{bc} = \Gamma^a_{cb} \ . \tag{25}$$

The covariance condition (14) is fulfilled if the covariant derivative of the metric tensor[7] vanishes,

$$Dg_{ab} = 0 \ . \tag{26}$$

This condition together with the symmetry (25) determines[8] the Christoffel symbol through the derivatives of the metric tensor,

$$\Gamma_{abc} = \frac{1}{2}\left(g_{ab,c} - g_{bc,a} + g_{ac,b}\right) . \tag{27}$$

Exercises

1. Is our definition of vectors in curvilinear coordinates consistent with the definition of four-vectors in special relativity? With Euclidean three-vectors of Galilean relativity?
2. What is the metric tensor in the Galilean space of Newtonian kinematics? In Minkowski space of relativistic kinematics?
3. What is the Jacobian of a Galilean transformation? Of a Lorentz transformation?
4. Suppose a law of physics in a certain system of coordinates is given as[9]

$$A^a = B^a \ .$$

   How does this law look like after a coordinate transformation $x \to x'(x)$ into another system of coordinates?

[6] Spaces with this property are called *torsion-free*.

[7] The Leibniz rule $D(A^aB^b) = DA^aB^b + A^aDB^b$ gives the definition of the covariant differential of a tensor as

$$DF^{ab} = dF^{ab} + \Gamma^a_{cd}F^{cb}dx^d + \Gamma^b_{cd}F^{ac}dx^d.$$

[8] The (vanishing) Dg_{ab} is given as

$$Dg_{ab} = dg_{ab} - \Gamma^d_{ac}g_{db}dx^c - \Gamma^d_{bc}g_{ad}dx^c = 0 \ .$$

This can be written as

$$g_{ab,c} - \Gamma_{bac} - \Gamma_{abc} = 0 \ .$$

Exchanging indices gives two other equations,

$$-g_{bc,a} + \Gamma_{cba} + \Gamma_{bca} = 0 \ ,$$

$$g_{ca,b} - \Gamma_{acb} - \Gamma_{cab} = 0 \ .$$

Finally, adding the last three equations gives (27).

[9] For example, the Maxwell equation with sources,

$$F^{ab}_{;a} = 4\pi j^b$$

.

5. (extra) Let x, y be Cartesian coordinates in a flat two-dimensional space with the metric

$$dl^2 = dx^2 + dy^2 \ .$$

Consider the polar coordinates

$$x = r\cos\theta \ , \quad y = r\sin\theta \ .$$

   (a) Calculate the interval dl^2 in polar coordinates and find the corresponding metric tensor g_{ab} (the indexes a and b in polar coordinates take the values r, θ).

   (b) The line element $d\mathbf{l}$ is defined as

$$d\mathbf{l} = dx\mathbf{e}_x + dy\mathbf{e}_y \ ,$$

   where $\mathbf{e}_x$ and $\mathbf{e}_x$ are the Cartesian unit vectors,

$$\mathbf{e}_x \cdot \mathbf{e}_x = \mathbf{e}_y \cdot \mathbf{e}_y = 1 \ , \quad \mathbf{e}_x \cdot \mathbf{e}_y = 0 \ .$$

   Consider this line element in polar coordinates,

$$d\mathbf{l} = dr\mathbf{e}_r + d\theta\mathbf{e}_\theta \ ,$$

   and find the polar unit vectors $\mathbf{e}_a$, $a = r, \theta$.

   (c) Argue and check that $\mathbf{e}_a \cdot \mathbf{e}_b = g_{ab}$.

   (d) Find $g^{ab} = g_{ab}^{-1}$, $\mathbf{e}^a = g^{ab}\mathbf{e}_b$, and $\mathbf{e}^a \cdot \mathbf{e}_b$.

   (e) Consider a vector $\mathbf{A} = A^a\mathbf{e}_a = A_a\mathbf{e}^a$. Consider the relation between the actual differential of the vector, $DA^a = \mathbf{e}^a \cdot d\mathbf{A}$, and the apparent differential of the vector, $dA^a = d(\mathbf{e}^a\mathbf{A})$ and prove that[10]

$$DA^a = dA^a + (\mathbf{e}^a \cdot \mathbf{e}_{b,c})\, A^b dx^c,$$

$$DA_a = dA_a - \left(\mathbf{e}^b \cdot \mathbf{e}_{a,c}\right) A_b dx^c.$$

   Argue that the expression in parentheses is the Christoffel symbol.

6. The *Kronecker delta* symbol δ^a_b is defined as

$$\delta^a_b \doteq \begin{cases} 1 & \text{if} \quad a = b\,, \\ 0 & \text{if} \quad a \neq b\,. \end{cases}$$

   (a) Show that $\delta^a_b X^b = X^a$,

   (b) Show that $\delta^a_b X_a = X_b$,

   (c) Show that δ^a_b is a tensor.

   (d) Evaluate δ^a_a.

7. A tensor, F^{ab}, is an object that transforms as an outer product, $A^a B^b$, of two vectors A^a and B^b. Using the Leibniz rule, $D(A^a B^b) = DA^a B^b + A^a DB^b$, derive the expression for the covariant differential DF^{ab}.

8. Argue that Christoffel symbols are not tensors.

[10] comma-index denotes partial derivative,

$$\phi_{,c} \equiv \frac{\partial\phi}{\partial x^c} \ .$$

9. Prove[11] that the transformation rule for Christoffel symbols is given as

$$\Gamma^a_{bc} = \frac{\partial x^a}{\partial x'^e}\frac{\partial x'^f}{\partial x^b}\frac{\partial x'^d}{\partial x^c}\Gamma'^e_{fd} + \frac{\partial^2 x'^e}{\partial x^b \partial x^c}\frac{\partial x^a}{\partial x'^e}$$

10. Is the quantity $\partial\phi/\partial x^a$, where ϕ is a scalar (a tensor of rank 0), a tensor? Is the quantity $\partial^2\phi/\partial x^a \partial x^b$ a tensor?

11. Argue that if a covariant (that is, with indices down) rank 2 tensor is symmetric in one coordinate system, it is symmetric in all acceptable coordinate systems.

12. Argue that if A^a_{bc} and B^a_{bc} are tensors (with one index up and two down) then so is their sum and difference.

13. Argue that a product of two tensors is a tensor.

14. Argue that any rank 2 tensor with both indices up (or both down) can be written as a sum of a symmetric and an anti-symmetric tensor.

15. Argue that the metric tensor is indeed a tensor.

16. Argue that the quantity $u^a u_a = g_{ab} u^a u^b$ is a (covariant) scalar and calculate its value.

[11] Hints:

$$g_{ab} = \frac{\partial x'^e}{\partial x^a}\frac{\partial x'^f}{\partial x^b} g'_{ef}\ ,\quad g^{ab} = \frac{\partial x^a}{\partial x'^e}\frac{\partial x^b}{\partial x'^f} g'^{ef}\ ,\quad \Gamma^u_{bc} = \frac{1}{2} g^{ua}\left(g_{ab,c} - g_{bc,a} + g_{ca,b}\right)\ ,$$

$$g_{ab,c} = \frac{\partial}{\partial x^c}\left(\frac{\partial x'^e}{\partial x^a}\frac{\partial x'^f}{\partial x^b} g'_{ef}\right) = \left(\frac{\partial^2 x'^e}{\partial x^c \partial x^a}\frac{\partial x'^f}{\partial x^b} + \frac{\partial x'^e}{\partial x^a}\frac{\partial^2 x'^f}{\partial x^c \partial x^b}\right) g'_{ef} + \frac{\partial x'^e}{\partial x^a}\frac{\partial x'^f}{\partial x^b}\frac{\partial x'^d}{\partial x^c} g'_{ef,d}\ .$$