Covariant differentiation of vectors in curved spaces

Introduction

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

The important property of the vector is that although it might seem to point at different direction relative to your car—depending on your driving direction—it actually always points in the direction of the town. In mathematical words, vectors have certain useful transformational properties when one changes ones reference frame.

Geometrically an arrow can be described by specifying the coordinates of the nock and the head the arrow. Actually, one only needs to specify the coordinate differences between these two points. Indeed the absolute position of the road-sign is not so important as long as you can still see it from your 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—a small arrow—is, mathematically speaking, 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 the same way as do coordinate differentials.

Coordinates, vectors, tensors

Coordinates and their differentials

A set of four numbers used to specify the location of an event in space-time is called *coordinates* and is denoted as $\{x^0, x^1, x^2, x^3\}$, or as x^a, or simply as x. For example, on Earth one could use *time*, *latitude*, *longitude*, and *height* above the sea level.

The coordinates themselves transform obviously non-linearly under a general non-linear coordinate trasformation. However, coordinate differentials—our archetypal vectors—do transform linearly even under a non-linear coordinate trasformation. Indeed, the set of coordinate differentials dx^a under a general coordinate transformation $x^a \to x'^a(x^b)$ transforms linearly,

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

via the *Jacobian matrix*,

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

Apart from coordinate differentials there exist another object that transforms linearly under a general coordinate transformation: 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[1]

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

[1]Indeed,

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

Vectors

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

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

Similarly, a set of four quantities A_a is called a *vector with index down* if it transforms similar to the set of partial derivatives of a 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, for example,

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

Contraction (generalized scalar product)

The *contraction* $A^a B_a$ is apparently a scalar (an object that transforms identically), 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 that under a general coordinate transformation transform similar to vectors and their products—that is, linearly via Jacobi matrix or its inverse—are said to transform *covariantly.* These objects are called *tensors.* The number of indices is called the *rank* of the tensor. For example, a scalar is a tensor of rank-0, and a vector is a tensor of rank-1. The general covariance principle (which states that the laws of physics must have the same form in all frames) postulates that the laws of physics must be formulated in terms of tensors.

An example of a tensor of rank-2 is the product $A^a B^b$ of the components of two vectors. Indeed it transforms linearly via Jacobian matrices 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 quantity 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,

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

[2] Also called contra-variant vector.

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,

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

Covariant differentials in curvilinear coordinates

The laws of physics (for example, the Lorentz force equation or the Maxwell equations) are formulated mathematically as (covariant) differential equations with tensors. A differential of a scalar is always a tensor and can therefore be directly used in such an equation. However the differential of a vector generally does not transform covariantly in curvilinear coordinates[3] and therefore is not a tensor.

We thus need to introduce covariant differentials of tensor which can be used to formulate laws of physics in curvilinear coordinates (and, via the equivalence principle, also in genuinely curved spaces with gravitational fields).

Covariant differential in flat spaces with curvilinear coordinates

Let us consider—as an example—a Euclidean space that contains a vector field $\mathbf{A}$. Let us introduce some arbitrary curvilinear coordinates (spherical, for example) with the metric tensor g_{ab} and with certain unit vectors $\mathbf{e}^a$ and $\mathbf{e}_a$ such that $\mathbf{e}^a \cdot \mathbf{e}_b = \delta^a_b$, $\mathbf{e}_a \cdot \mathbf{e}_b = g_{ab}$, $\mathbf{e}_a = g_{ab}\mathbf{e}^b$. The components A^a of the vector in these coordinates are then given as

$$A^a \doteq \mathbf{e}^a \cdot \mathbf{A} \ . \tag{13}$$

Unfortunately, the component differentials, dA^a, where

$$dA^a = \mathbf{e}^a \cdot d\mathbf{A} + d\mathbf{e}^a \cdot \mathbf{A} \tag{14}$$

do not transform covariantly if $d\mathbf{e}^a \neq 0$. The first term does actually transform covariantly, as it contains the actual change of the vector $d\mathbf{A}$. Let us call it the covariant differential, DA^a,

$$DA^a = \mathbf{e}^a \cdot d\mathbf{A} \tag{15}$$

It indeed transforms covariantly,

$$DA'^a = \mathbf{e}'^a \cdot d\mathbf{A} = (\mathbf{e}'^a \cdot \mathbf{e}_b)(\mathbf{e}^b \cdot d\mathbf{A}) = J^a_b \, DA^b \ , \tag{16}$$

where J^a_b is the Jacobian matrix[4]. And it is exactly the term which appears when we project a differential vector equation, say $d\mathbf{A} = d\mathbf{B}$, on our unit vectors,

$$d\mathbf{A} = d\mathbf{B} \ \Rightarrow \ \mathbf{e}^a d\mathbf{A} = \mathbf{e}^a d\mathbf{B} \ \Rightarrow \ DA^a = DB^a \ . \tag{17}$$

That is, a differential vector equation

$$d\mathbf{A} = d\mathbf{B}$$

in curvilinear coordinates becomes

$$DA^a = DB^a \ .$$

We can get the component expression for the covariant differential DA^a from (14),

$$DA^a = dA^a - d\mathbf{e}^a\mathbf{A} = dA^a - (d\mathbf{e}^a\mathbf{e}_b)(\mathbf{e}^b\mathbf{A}) = dA^a + (\mathbf{e}^a\mathbf{e}_{b,c})A^b dx^c \ . \tag{18}$$

[3] Curvilinear coordinates are coordinates where the metric tensor is not everywhere pseudo-Euclidean.

[4]
$$dx'^a = \mathbf{e}'^a d\mathbf{x} = \mathbf{e}'^a \mathbf{e}_b dx^b \doteq J^a_b dx^b$$

where we have introduced the comma notation for partial derivatives,

$$f_{,c} \equiv \frac{\partial f}{\partial x^c} \; . \tag{19}$$

The expression in parentheses is called the *Christoffel symbol*,

$$(\mathbf{e}^a \mathbf{e}_{b,c}) \doteq \Gamma^a{}_{bc} \; . \tag{20}$$

Christoffel symbols (in this case) are symmetric under exchange of the last indices,[5]

$$\Gamma^a{}_{bc} = \Gamma^a{}_{cb} \; . \tag{24}$$

Using the equation $g_{ab} = \mathbf{e}_a \mathbf{e}_b$ and the symmetry property we can express Christoffel symbols in terms of the metric tensor as[6],

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

The covariant differential in curvilinear coordinates is then given as[7]

$$DA^a = dA^a + \Gamma^a{}_{bc} A^b dx^c \; . \tag{28}$$

Now, using the comma-notation for partial derivatives,

$$dA^a = A^a_{,c} dx^c \; , \tag{29}$$

gives

$$DA^a = (A^a_{,c} + \Gamma^a{}_{bc} A^b) dx^c \; . \tag{30}$$

The expression in parentheses is called covariant derivative and is denoted as $A^a_{;c}$,

$$A^a_{;c} \doteq A^a_{,c} + \Gamma^a{}_{bc} A^b \; . \tag{31}$$

It is apparently a tensor, unlike the partial derivative.

Now, the principle of equivalence tells us that in genuinely curved spaces the formulae should be the same.

[5] Let us denote the (constant) Cartesian unit vectors as $\mathbf{E}_a$ and the corresponding Cartesian coordinates as ξ^a. The infinitesimal vector $d\mathbf{l}$ connecting two close points is given as

$$d\mathbf{l} = \mathbf{E}_a d\xi^a = \mathbf{E}_a \frac{\partial \xi^a}{\partial x^b} dx^b = \mathbf{e}_b dx^b \; , \tag{21}$$

from which we get

$$\mathbf{e}_b = \mathbf{E}_a \frac{\partial \xi^a}{\partial x^b} \; , \; \mathbf{e}_c = \mathbf{E}_a \frac{\partial \xi^a}{\partial x^c} \; , \tag{22}$$

Now,

$$\mathbf{e}_{b,c} = \mathbf{E}_a \frac{\partial^2 \xi^a}{\partial x^b \partial x^c} \; , \; \mathbf{e}_{c,b} = \mathbf{E}_a \frac{\partial^2 \xi^a}{\partial x^c \partial x^b} \; . \tag{23}$$

These two expressions are equal (assuming double differentiability of the coordinate transformation functions).

[6] The partial derivative of the metric tensor is given as

$$g_{ab,c} = (\mathbf{e}_a \mathbf{e}_b)_{,c} = \Gamma_{bac} + \Gamma_{abc} \; . \tag{25}$$

Writing this expression two more times with cyclic permutation of indices and then making a linear combination gives (taking into account the symmetry property)

$$g_{ab,c} - g_{bc,a} + g_{ca,b} = 2\Gamma_{abc} \; . \tag{26}$$

[7] In mathematics our covariant differential is referred to as *metric torsion-free connection* or *Levi-Civita connection.*

Covariant differential in metric spaces

Let us repeat the derivations without a reference to the underlying flat space.

In a curved space the differential of vector component, dA^a, is not a covariant quantity, since generally

$$dA_a = d(g_{ab}A^b) = g_{ab}dA^b + dg_{ab}A^b \neq g_{ab}dA^b \ , \tag{32}$$

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

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

$$DA_a = g_{ab}DA^b = Dg_{ab}A^b. \tag{33}$$

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

$$DA^a = dA^a + \Gamma^a{}_{bc}A^b dx^c \ , \tag{34}$$

where the factor $\Gamma^a{}_{bc}$ (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

$$D(A_aB^a) = d(A_aB^a) \tag{35}$$

for an arbitrary B^a. If we demand that covariant differential satisfies the Leibniz product rule[8] we find[9]

$$DA_a = dA_a - \Gamma^b{}_{ac}A_b dx^c \ . \tag{36}$$

With the comma-notation for partial derivatives,

$$dA^a = \frac{\partial A^a}{\partial x^c}dx^c \equiv A^a_{,c}dx^c \ , \tag{37}$$

the covariant differentials take the form

$$DA^a = \left(A^a_{,c} + \Gamma^a{}_{bc}A^b\right)dx^c \ , \tag{38}$$

$$DA_a = \left(A_{a,c} - \Gamma^b{}_{ac}A_b\right)dx^c \ . \tag{39}$$

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

$$A^a_{;c} = A^a_{,c} + \Gamma^a{}_{bc}A^b \ , \tag{40}$$

$$A_{a;c} = A_{a,c} - \Gamma^b{}_{ac}A_b \ . \tag{41}$$

[8] The Leibniz product rule states that

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

[9] Applying the Leibniz product rule to (35) 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

$$\Gamma^a_{bc} - \Gamma^a{}_{cb} \tag{42}$$

is a tensor (called *torsion tensor*). Indeed, consider the difference,

$$\varphi_{;c;b} - \varphi_{;b;c} = (\Gamma^a{}_{bc} - \Gamma^a{}_{cb})\varphi_{;a} \ , \tag{43}$$

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 over exchange of the two lower indexes[10],

$$\Gamma^a{}_{bc} = \Gamma^a{}_{cb} \ . \tag{44}$$

The covariance condition (33) is fulfilled if the covariant derivative of the metric tensor[11] vanishes,

$$Dg_{ab} = 0 \ . \tag{45}$$

This condition together with the symmetry (44) determines[12] the Christoffel symbol through the derivatives of the metric tensor,

$$\Gamma_{abc} = \frac{1}{2}\left(g_{ab,c} - g_{bc,a} + g_{ac,b}\right) . \tag{46}$$

Exercises

1. Prove that matrix $\partial x^a/\partial x'^b$ is inverse to matrix $\partial x'^c/\partial x^d$.
2. Is our definition of vectors in curvilinear coordinates consistent with
   - the Euclidean three-vectors of Galilean relativity?
   - the Minkowski four-vectors of special relativity?

   Think about velocity, momentum, (potential) force, and whatever other vectors you can think of.
3. What is the metric tensor in the Galilean space of Newtonian kinematics? In Minkowski space of relativistic kinematics?
4. What is the Jacobian matrix of the Galilean transformation? Of the Lorentz transformation?

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

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

$$DF^{ab} = dF^{ab} + \Gamma^a{}_{cd}F^{cb}dx^d + \Gamma^b{}_{cd}F^{ac}dx^d .$$

[12] 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 (46).

5. Suppose a law of physics in a certain system of coordinates is given as[13]

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

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

6. In a 2D Euclidean (flat) space with Cartesian coordinates $\{x, y\}$ consider a coordinate transformation

$$x = r\cos\phi,\ y = r\sin\phi$$

to polar coordinates $\{r, \phi\}$.

- Find the metrix tensor in polar coordinates. Is it *conformally flat*?[14]
- Find the unit vectors $\mathbf{e}_a$ and $\mathbf{e}^a$ where $a = r, \phi$.
- Check that $\mathbf{e}_a\mathbf{e}_b = g_{ab}$.
- Calculate the Christoffel symbols.

7. Prove that in a flat space with curvilinear coordinates the Christoffel symbols satisfy the equation

$$\mathbf{e}^a\mathbf{e}_{b,c} = \frac{1}{2}\left(g_{ab,c} - g_{bc,a} + g_{ca,b}\right) \ .$$

8. Using $A^a_{;b} = A^a_{,b} + \Gamma^a{}_{bc}A^c$ and the Leibniz poduct rule prove that $A_{a;b} = A_{a,b} + \Gamma^c{}_{ab}A^c$.

9. Using our formalism of covariant differentiation of vectors write down the 3D Cartesian equation

$$\nabla\mathbf{E} = 4\pi\rho$$

(the Gauss law) in cylindrical and spherical coordinates. Compare with the expressions in your textbook (or in Wikipedia). Hint: you might want first to rewrite the equation in a generally covariant form by substituting $_{,a}$ with $_{;a}$.

10. 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.

(e) Evaluate $D\delta^a_b$.

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

12. Argue that Christoffel symbols are not tensors.

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

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

.

[14] Let us define *conformal* as differing only in a factor that is a smooth function of coordinates.

13. (extra) Prove[15] that the transformation rule for Christoffel symbols is given as

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

14. 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?

15. Argue that if a rank-2 tensor is symmetric, $T^{ab} = T^{ba}$, in one coordinate system, it is symmetric in all acceptable coordinate systems.

16. Argue that if A^a_{bc} and B^a_{bc} are tensors then so is their sum and difference.

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

18. 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.

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

20. Argue that the quantity $u^a u_a = g_{ab} u^a u^b$ (where $u^a \doteq dx^a/ds$ is the four-velocity) is a (covariant) scalar and calculate its value.

[15] Hints:

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

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