



planetmath.org

Math for the people, by the people.

connection

Canonical name	Connection
Date of creation	2013-03-22 12:37:10
Last modified on	2013-03-22 12:37:10
Owner	rspuzio (6075)
Last modified by	rspuzio (6075)
Numerical id	43
Author	rspuzio (6075)
Entry type	Definition
Classification	msc 53B05
Related topic	Christoffel20symbols
Related topic	CurvatureOfACircle
Related topic	SphereTheoremFromGlobalDifferentialGeometry
Related topic	Geodesic
Related topic	RiemannCurvatureTensor
Defines	covariant derivative
Defines	torsion
Defines	curvature
Defines	torsionless
Defines	flat
Defines	Christoffel symbol

Note that this definition has come under new management and is still in the process of being edited and rewritten.

Intuitive geometric definition The notions of connection, parallel transport, and covariant derivative are closely related so, to prevent confusion, we will begin by explaining these notions intuitively before presenting formal definitions. Moreover, it helps to have a good grasp of the geometric notions involved before studying the more formal definitions.

In elementary vector analysis, one takes it for granted that vectors can be moved about freely. As long as one takes care not to change the magnitude or the direction of a vector, one can move the basepoint of the vector to any arbitrary location.

When one graduates to the study of vectors on curved spaces, however, it becomes apparent that one can no longer take this freedom of moving vectors about for granted. As defined, vectors are confined to their basepoint and the basic operations with vectors are only defined for vectors based at the same point.

To move a vector from one point to another, one needs to specify how this is to be done. A connection is a prescription for moving vectors based at one point of a space to another point. Intuitively speaking, a connection consists of a set of linear transformations which transform vectors based at a particular point into vectors based at infinitesimally nearby points. Unlike in elementary vector analysis where there is only one right way of moving a vector from one point to another, in differential geometry there are many ways of moving vectors around, so one needs to specify which connection one is using before one can move vectors from point to point.

This act of moving a vector from point to point is called *parallel transport* in analogy with the operation of elementary vector analysis which it generalizes. Not only can one speak of transporting a vector to a nearby point using a connection, but one can parallel transport a vector along a curve. To see how that works, imagine a curve as a sequence of points. Using the connection, we can transport a vector based at one point of a curve to the next point on the curve. Then we can use the connection to transport it to the point after that, and so on until we have transported it from one end of the curve to the other.

At this point, a striking difference between differential geometry and elementary vector analysis shows up. Typically, if we connect two points P and

Q by two or more curves and parallel transport a vector based at P to Q , we find that the result depends upon which curve we transported the vector along. In fact, in differential geometry, the definition of a curved space is a space in which there exist two distinct curves with the same endpoints such that parallel transport along one curve is not the same as parallel transport along the other curve.

Finally, there is the notion of covariant derivative. Suppose that one is given not just a single vector based at a certain point, but a whole vector field, i.e. a vector for each point of the manifold. Then one can try to compute the derivative of this vector field. To compute a derivative of a function, one subtracts the value of the function at a point from the value at a nearby point. But this is not possible for the vector field because we are only allowed to subtract vectors stationed at the same base point. However, we can use our connection to parallel transport the vector at a point to the nearby point, then subtract. This generalization of differentiation involving parallel transport is known as covariant differentiation.

Obviously, the above definitions leave much to be desired in the way of precision. They are not specific what a space is, how vectors are to be associated to the points of this space, and are based on vague notions of infinitesimally nearby points.

For the purpose of this article, we shall take our space to be a finite-dimensional manifold. To be sure, some of the definitions to be given apply to more general contexts, such as infinite-dimensional manifolds so one may speak of connections on these spaces as well. However, we shall not pursue this topic here since this exposition is intended to be accessible to newcomers to differential geometry who may not have the necessary background in Hilbert space theory, point set topology, and other subjects.

The definition of vector bundle makes precise the idea of "attaching vectors to points". The reader who is not comfortable with general vector bundles may take this bundle to be the tangent bundle.

There are about as many ways of framing a rigorous definition of connection as there are ways of formalizing differential geometry. Hence, under the headings below, we shall list various equivalent definitions.

Before proceeding to these definitions, a few words of warning may be in order. Since the notions of connection, parallel transport, and covariant derivative are so closely related, it is easy to translate propositions involving one of these terms into propositions involving a different one of three terms. In particular, propositions about connections are easily rewritten as propo-

sitions about covariant derivatives. In some formalisms, it is easier to define covariant derivative than to define connection. This leads to an abuse of terminology — some authors say things like “the connection ∇ ” instead of the more precise statement “the covariant differentiation operator ∇ ”. This can be disconcerting to the uninitiated, but once the principle involved has been grasped, this practise is harmless.

Preliminaries. Let M be a smooth, d -dimensional differential manifold. Let $\mathcal{F}(M)$ denote the ring of smooth, real-valued functions on M , and let $\mathcal{X}(M)$ denote the real vector space of smooth vector fields. Let B be a vector bundle over M whose structure group is the finite-dimensional Lie group G and whose fibers are isomorphic to the n -dimensional vector space V . Let $\mathcal{X}(B)$ denote the set of sections of B . Let G^M denote the set of smooth maps from M to G ; it forms a group under pointwise multiplication. (If f and g are two functions from M to G , then their product is h which is defined as $h(x) = f(x)g(x)$.) Likewise, let A^M denote the set of smooth maps from M to A .

For simplicity, we shall assume that $G = GL(V)$ for the time being. After stating the definitions of connection in this case, we shall describe how they can be modified to cover the case where $G \subset GL(V)$.

Recall that $\mathcal{F}(M)$ both acts and is acted upon by $\mathcal{X}(M)$. Given a function $f \in \mathcal{F}(M)$ and a vector field $X \in \mathcal{X}(M)$ we write $fX \in \mathcal{X}(M)$ for the vector field obtained by point-wise multiplying values of X by values of f , and write $X(f) \in \mathcal{F}(M)$ for the function obtained by taking the directional derivative of f with respect to X .

Coordinate definition. Let x^1, \dots, x^d be a set of coordinates on some neighborhood of M . These may be extended to coordinates on a subset of the bundle by augmenting them with coordinates x^{d+1}, \dots, x^{d+n} on the fiber. Since the fiber is a vector space, we will demand that the fiber coordinates be linear coordinantes. (This means that the coordinates of the sum of two vectors are the sum of the coordinates of the two vectors and the coordinantes of the scalar multiple of a vector are gotten by multiplying the coordinates of the original vecor by the scalar.) Adopt the convention that Latin indices run from 1 to d and that Greek indices run from $d+1$ to n .

In these coordinates, a connection will be represented by a three-index

field

$$C_{\nu i}^{\mu}(x^1, \dots, x^d)$$

on the manifold and the covariant derivative $\nabla_i S^{\mu}$ of a section $S^{\mu} \in \mathcal{X}(B)$ will be an element of $T(M) \otimes B$ given by the formula

$$\nabla_i S^{\mu} = \partial_i S^{\mu} + C_{\nu i}^{\mu} S^{\nu}$$

(Here ∂_i is short for $\partial/\partial x^i$ and the summation convention is in force. It might also be worth mentioning that sometimes the covariant derivative is defined with a minus sign instead of a plus sign (on rare occasions mostly occurring in high-energy physics theory, one even sees it defined with an imaginary unit i) so one needs to check which sign convention is in use.)

Before proceeding further, it might be helpful to present a warning. The notation $\nabla_i T^{\mu}$ can lead to some confusion, and this danger warrants an extra comment. The symbol ∇_i acting on a function, is customarily taken to mean the same thing as the corresponding partial derivative:

$$\nabla_i f = \partial_{x^i}(f) = \frac{\partial f}{\partial x^i}.$$

Thus, it is easy to make the mistake that $\nabla_i T^{\mu}$ is the result of applying an operator ∇_i to each component of T^{μ} . As can be seen from the definition, this is not the case. Rather, one should think of $\nabla_i T^{\mu}$ as if it were $(\nabla T)_i^{\mu}$, which is to say it denotes the components of a new tensor which was derived from T^{μ} by the operation of covariant differentiation.

The relation of these formulas to the naive picture is as follows: A connection is supposed to be a collection of linear maps from one tangent space to neighboring tangent space. Given a point $p \in M$, any vector w^{μ} in the fiber of B above p^i is transformed into the vector $w^{\mu} + C_{\nu i}^{\mu}(p)w^{\nu}dx^i$ in the fiber above the nearby point $p^i + dx^i$. (In this paragraph, I am using " dx^{μ} " in its naive sense of "infinitesimal displacement" rather than as a differential form.) Likewise, subtracting the value of $S^{\mu}(p^i + dx^i)$ from the parallel-transported value of $S^{\mu}(p)$ and dividing by dx , one obtains the formula for covariant derivative.

In order for a geometrical quantity to be defined properly by a coordinate expression, one must specify how the quantity transforms under change of coordinates. Under a change of coordinates

$$y^i = f^i(x^1, \dots, x^d)$$

$$y^\mu = \Lambda^\mu_\nu(x^1, \dots, x^d)y^\nu$$

the components of the connection transform as follows:

$$C^\mu_{\nu i}(y) = \frac{\partial f^j}{\partial f_i} \Lambda^\mu_\sigma (\Lambda^{-1})^\tau_\nu C^\sigma_{\tau j} + (\Lambda^{-1})^\mu_\kappa \frac{\partial \Lambda^\kappa_\nu}{\partial x^j}$$

Note that these rules imply that the components of a connection do not transform like the components of a tensor — the term involving the derivatives of Λ is not present in the transformation law of a tensor. However, if we have two connections on the same bundle, the difference of these connections will be a tensor because the extra terms cancel.

The reason for defining the transformation law in this way is so that the covariant derivative $\nabla_i S^\mu$ of a section of $S^\mu \in \mathcal{X}(B)$ will transform as an element of $T(m) \otimes B$ should. Furthermore, as one may check by transforming the various quantities that appear in the equation defining the covariant derivative, this is the only possible transformation law which will make $\nabla_i S^\mu$ transform properly. This property is the origin of the term “covariant derivative” — the covariant derivative maps tensor fields into quantities which transform in the same manner.

Alternative Notations There are many different systems of notations in differential geometry. (Indeed one humorous definition of differential geometry is “The study of invariants under change of notation”!) This section will discuss several notations for connections and covariant derivatives.

It is traditional to represent the components of the covariant derivative like this

$$Y^\mu_{;j} = \nabla_j Y^\mu$$

using the semi-colon to indicate that the extra index comes from covariant differentiation. Sometimes, as in the theory of embedded surfaces, there are two connections present so a semicolon is used to indicate covariant derivatives with respect to one connection and a vertical bar or a colon is used to indicate covariant derivatives with respect to the other connection. It might also be worth noting that commas are likewise used to indicate partial derivatives with respect to a given coordinate system. Using this notation, one might write the formula for covariant derivative as

$$T^\mu_{;i} = T^\mu_{,i} + C^\mu_{\nu i} T^\nu$$

Also, there are different ways of packaging the information contained in the connection components. One may collect the connection components into d matrices A_i :

$$A_i = \begin{pmatrix} C_{d+1\,i}^{d+1} & \cdots & C_{d+n\,i}^{d+1} \\ \vdots & \ddots & \vdots \\ C_{d+1\,i}^{d+n} & \cdots & C_{d+n\,i}^{d+n} \end{pmatrix}$$

Another common notational device is to collect the connection coefficients into the so-called “connection one-forms”

$$A_\nu^\mu = C_{\nu i}^\mu dx^i$$

When using this notation, the covariant derivative is written as a generalization D of the exterior derivative d :

$$DT^\mu = dT^\mu + A_\nu^\mu T^\nu$$

By combining the two devices and collecting the connection one-forms into a matrix A , one may do away with indices altogether. If one also collects the components T^μ into a column vector T , one may write

$$DT = dT + AT$$

A quantity like A is often referred to as a matrix-valued one-form.

Occasionally, one finds connection coefficients with only two indices instead of three. The reason is that the two indices referring to the bundle have been replaced by a single index referring to the Lie algebra. To relate this notation to the one discussed so far, we need to remember that the action of the structure group G on V defines a representation of the Lie algebra A on V , i.e. a map

$$\rho: A \rightarrow \text{Hom}(V, V)$$

If we choose linear coordinates y_1, \dots, y^m on the vector space A , this map may be expressed in components as

$$(y^1, \dots, y^m) \mapsto t_{\nu I}^\mu y^I$$

(Extend our conventions by agreeing that capital Latin indices run from 1 to m , where m is the dimension of the Lie algebra. In the case we are

considering, where $G = GL(V)$, we will have $m = n^2$.) To the two-index object A_i^I , we will associate the three-index object

$$C_{\nu i}^\mu = A_i^I t_{\nu I}^\mu$$

Therefore, one may also specify a connection in a coordinate system by giving an array indexed by an index referring to the Lie algebra and an index referring to the cotangent space of the manifold. This notation is useful in situations when one wants to emphasize the structure group rather than the manifold or when one is dealing with more than one bundle whose fibers are different representations of the same group.

Definition in terms of one-forms It is worth noting that one can define the connection directly in terms of the curvature one-forms. A noteworthy feature of such definition is that it does not make explicit reference to coordinate systems on the manifold, although it does make use of local neighborhoods. After the discussion of the last section, the relation of this definition to the preceding definition should be clear.

As in the last section, let $\rho: G \rightarrow Hom(V, V)$ denote the action of G on V .

Let (U, ϕ) be a local trivialization of the bundle B . Recall that U is an open set of M and that ϕ is a diffeomorphism between $\pi^{-1}(U) \in B$ and $U \times V$. To every local trivialization, associate an element A of v . In order for these elements to define a connection, they must transform properly under changes of local trivialization. Two local trivializations over the same set U are related by a transition function $g: U \rightarrow G$. The transformation law of an element $A \in T^*(U) \times Hom(V, V)$ is given by

$$A' = \rho(g)^{-1} A \rho(g) + \rho(g)^{-1} d\rho(g)$$

For this definition to be consistent, it must agree with the cocycle condition. The reason for this is that, if it didn't, one obtain different answers by transforming from one local trivialization to another in two different ways. That it is consistent is easily verified. Using the notation of the entry on fibre bundles,

$$\begin{aligned} \rho(g_{ij}g_{jk})^{-1} A \rho(g_{ij}g_{jk}) + \rho(g_{ij}g_{jk})^{-1} d\rho(g_{ij}g_{jk}) &= \\ \rho(g_{jk})^{-1} \rho(g_{ij})^{-1} A \rho(g_{ij}) \rho(g_{jk}) + \rho(g_{jk})^{-1} \rho(g_{ij})^{-1} d(\rho(g_{ij}) \rho(g_{jk})) &= \\ \rho(g_{jk})^{-1} (\rho(g_{ij})^{-1} A \rho(g_{ij}) + \rho(g_{ij})^{-1} d(\rho(g_{ij}))) \rho(g_{jk}) + (g_{jk})^{-1} d\rho(g_{jk}) \end{aligned}$$

Axiomatic definition of covariant differentiation In this definition, covariant differentiation is characterized axiomatically. As explained in the first section, it is not necessary to augment this with a separate definition of connection, since any statement about connections can be rephrased as a statement about covariant derivatives. An important feature of this definition which sets it apart from the previous two definitions is that it is global — there is no need to chop up the manifold or the bundle into patches, define the connection on each patch, then sew the patches back together again to make a complete manifold.

A covariant derivative ∇ is a mapping

$$\begin{aligned}\nabla : \mathcal{X}(M) \times \mathcal{X}(B) &\rightarrow \mathcal{X}(B) \\ (X, Y) &\mapsto \nabla_X Y, \quad X \in \mathcal{X}(M), Y \in \mathcal{X}(B)\end{aligned}$$

that for all $X, Y \in \mathcal{X}(M)$, all $Z, W \in \mathcal{X}(B)$, all $f \in \mathcal{F}(M)$, and all $\lambda \in A^H$ satisfies

1. $\nabla_{X+Y} Z = \nabla_X Z + \nabla_Y Z$
2. $\nabla_X (Z + W) = \nabla_X Z + \nabla_X W$
3. $\nabla_{fX} Z = f \nabla_X Z$
4. $\nabla_X (fZ) = X(f)Z + f \nabla_X Z$

Note that the lack of tensoriality in the second argument means that a connection *is not* a tensor field.

Also note that we can regard the connection as a mapping from $\mathcal{X}(M)$ to the space of type (1,1) tensor fields, i.e. for $Y \in \mathcal{X}(M)$ the object

$$\begin{aligned}\nabla Y : \mathcal{X}(M) &\rightarrow \mathcal{X}(M) \\ X &\mapsto \nabla_X Y, \quad X \in \mathcal{X}(M)\end{aligned}$$

is a type (1,1) tensor field called the *covariant derivative* of Y . In this capacity ∇ is often called the *covariant derivative operator*.

Recall that once a system of coordinates is chosen, a given vector field $Y \in \mathcal{X}(M)$ is represented by means of its components $Y^i \in \mathcal{F}(U)$ according to

$$Y = Y^i \partial_{x_i}.$$

The formula for the components follows directly from the defining properties of a connection and the definition of the Christoffel symbols. To wit:

$$Y^i_{;j} = Y^i_{,j} + \Gamma_{jk}^i Y^k$$

where the symbol with the comma

$$Y^i_{,j} = \partial_{x_j}(Y^i) = \frac{\partial Y^i}{\partial x^j}$$

denotes a derivate relative to the coordinate frame.

A related and frequently encountered notation is ∇_i , which indicates a covariant derivatives in direction ∂_{x_i} , i.e.

$$\nabla_i Y = \nabla_{\partial_{x_i}} Y, \quad Y \in \mathcal{X}(M).$$

This notation jibes with the point of view that the covariant derivative is a certain generalization of the ordinary directional derivative. The partials ∂_{x_i} are replaced by the covariant ∇_i , and the general directional derivative $V^i \partial_{x_i}$ relative to a vector-field V , is replaced by the covariant derivative operator $V^i \nabla_i$.

Group compatibility So far, we have been labouring under the assumption that $G = GL(V)$. The time has now come to remove this restriction. To do so, we need to come to grips with the issue of group compatibility. As usual, we shall begin by discussing the problem in intuitive terms, then formalize our intuition in various formalisms.

The structure group transforms vectors located at a point into each other whilst the connection transforms transforms vectors based at one point into vectors based at another point. To understand the problem of compatibility, let us focus attention on two nearby points P and Q of the manifold and the fibres above these points.

There are two ways to transform a vector $v \in V_P$. (Since it is crucial to remember that the fibers over different points are distinct vector spaces if one is to order to understand this discussion, we have indexed the copies of V which serve as fibers of the bundle over various points of the manifold with their basepoints. Likewise, we shall index the symbol ρ with a point of the manifold to indicate the action of the group on vectors based at that point.) The simplest way is to pick an element $g \in G$ and apply the transformation

$\rho_P(g)$ to v . Alternatively, one could first parallel transport v to $Tv \in V_Q$, apply the transform $\rho_Q(g)$ to the transported vector, then parallel transport the result back to P to obtain $T^{-1}\rho_Q(g)Tv$.

If the transform $T^{-1}\rho_Q(g)T$ does not equal $\rho_P(g')$ for any $g' \in G$, we are in trouble. By using the connection, we could generate a transformation of the fiber which is not described by the structure group of the bundle. To avoid this difficulty, we need to demand that the connection is compatible with the group. Group compatibility is the condition that for every map $T : V_P \rightarrow V_Q$ which parallel transports a vector from a point P to another point Q and for every $g \in G$, there exists a $g' \in G$ such that $\rho_P(g)T = T\rho_Q(g')$. In the language of representation theory, we would say that T intertwines the representations ρ_P and ρ_Q of G .

It is worth noting that, if we transport the vector v from P to Q by first transporting it to an intermediate point R , it is enough to check that the transport from P to R and the transport from R to Q are group compatible since, if they are, it will automatically follow that the transport from P to Q is group compatible. To verify this assertion, let T_1 be the matrix which transports vectors from V_P to V_R and let T_2 be the matrix which transports vectors from V_R to V_Q .

$$\nabla_X(\lambda Z) - \lambda \nabla_X Z = \mu Z \text{ for some } \mu \in A^M$$

Related Definitions. The torsion of a connection ∇ is a bilinear mapping

$$T : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$$

defined by

$$T(X, Y) = \nabla_X(Y) - \nabla_Y(X) - [X, Y],$$

where the last term denotes the Lie bracket of X and Y .

The curvature of a connection is a tri-linear mapping

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$$

defined by

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z, \quad X, Y, Z \in \mathcal{X}(M).$$

We note the following facts:

- The torsion and curvature are tensorial (i.e. $\mathcal{F}(M)$ -linear) with respect to their arguments, and therefore define, respectively, a type (1,2) and a type (1,3) tensor field on M . This follows from the defining properties of a connection and the derivation property of the Lie bracket.
- Both the torsion and the curvature are, quite evidently, anti-symmetric in their first two arguments.

A connection is called *torsionless* if the corresponding torsion tensor vanishes. If the corresponding curvature tensor vanishes, then the connection is called *flat*. A connection that is both torsionless and flat is locally Euclidean, meaning that there exist local coordinates for which all of the Christoffel symbols vanish.

Notes. The notion of connection is intimately related to the notion of parallel transport, and indeed one can regard the former as the infinitesimal version of the latter. To put it another way, when we integrate a connection we get parallel transport, and when we take the derivative of parallel transport we get a connection. Much more on this in the parallel transport entry.

As far as I know, we have Elie Cartan to thank for the word *connection*. With some trepidation at putting words into the master’s mouth, my guess is that Cartan would lodge a protest against the definition of *connection* given above. To Cartan, a connection was first and foremost a geometric notion that has to do with various ways of *connecting* nearby tangent spaces of a manifold. Cartan might have preferred to refer to ∇ as the covariant derivative operator, or at the very least to call ∇ an *affine* connection, in deference to the fact that there exist other types of connections (e.g. projective ones). This is no longer the mainstream view, and these days, when one wants to speak of such matters, one is obliged to use the term *Cartan connection*.

Indeed, many authors call ∇ an affine connection although they never explain the *affine* part.¹ One can also define connections and parallel transport in terms of principal fiber bundles. This approach is due to Ehresmann.

¹The silence is puzzling, and I must confess to wondering about the percentage of modern-day geometers who know exactly what is so affine about an affine connection. Has blind tradition taken over? Do we say “affine connection” because the previous person said “affine connection”? The meaning of “affine” is quite clearly explained by Cartan in his writings. There you go esteemed “everybody”: one more reason to go and read Cartan.

In this generalized setting an affine connection is just the type of connection that arises when working with a manifold's frame bundle.

Bibliography. [Exact references coming.]

Bishop and Goldberg (1968)

- Cartan's book on projective connection.
- Ehresmann's seminal mid-century papers.
- Kobayashi and Nomizu's books

Spivak (1965)

See also the bibliography for differential geometry.