

Lecture 12: Covariant derivatives, brackets, and geodesics

Let $M \subset \mathbb{R}^{n+1}$ be a regular hypersurface, with first fundamental form I and a local unit normal ν . For tangent vector fields X, Y , recall

$$\nabla_X Y = (D_X Y)^\top = D_X Y - \langle D_X Y, \nu \rangle \nu.$$

The **covariant derivative** keeps only the tangential part of the change of Y in the direction X .

If $c(0) = p$ and $\dot{c}(0) = X_p$, then

$$(\nabla_X Y)_p = \left(\frac{d}{dt} Y(c(t)) \Big|_{t=0} \right)^\top.$$

Only X_p is needed, but Y must be known near p . For a smooth function φ , the rules are

$$\nabla_{\varphi X} Y = \varphi \nabla_X Y, \quad \nabla_X(\varphi Y) = X(\varphi)Y + \varphi \nabla_X Y.$$

In particular, covariant differentiation is not bilinear over functions in both arguments.

The *Jacobi bracket*, or *Lie bracket*, is

$$[X, Y] = D_X Y - D_Y X = \nabla_X Y - \nabla_Y X.$$

The normal parts cancel because $II(X, Y) = II(Y, X)$. Thus the bracket is again tangent; the last equality is the torsion-free identity.

In coordinates $f : U \rightarrow M$, write

$$f_i = \partial_i f, \quad X = \sum_{i=1}^n \xi^i f_i, \quad Y = \sum_{j=1}^n \eta^j f_j.$$

Then

$$[X, Y] = \sum_{j=1}^n (X(\eta^j) - Y(\xi^j)) f_j, \quad X(\eta^j) = \sum_{i=1}^n \xi^i \partial_i \eta^j.$$

The terms involving f_{ij} cancel by equality of mixed partials. In particular, $[f_i, f_j] = 0$.

For given vector fields, the bracket requires no metric. It measures an asymmetry between their directional derivatives, not the normal bending of the hypersurface.

A vector field acts on scalar functions by differentiation. The bracket acts as the commutator:

$$[X, Y](\varphi) = X(Y(\varphi)) - Y(X(\varphi)).$$

Second derivatives of φ cancel, leaving exactly the coefficient formula on the preceding page.

The bracket is skew-symmetric and satisfies the ***Jacobi identity***:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

To prove this, compose differential operators, writing $XYZ(\varphi) = X(Y(Z(\varphi)))$. Then

$$[X, [Y, Z]] = XYZ - XZY - YZX + ZYX.$$

Add the two cyclic permutations. Each of the six possible compositions occurs once with each sign, so the sum is zero.

This proves the identity on every function. A vector field is determined by its action on functions, since applying it to coordinate functions recovers all its components.

Geometrically, the bracket measures the failure of two small motions to commute.

The *local flow* Φ_t^X follows the field X :

$$\frac{d}{dt}\Phi_t^X(p) = X_{\Phi_t^X(p)}, \quad \Phi_0^X(p) = p.$$

Starting at p , compare moving along X and then Y with moving in the reverse order:

$$P_\varepsilon = \Phi_\varepsilon^Y(\Phi_\varepsilon^X(p)), \quad Q_\varepsilon = \Phi_\varepsilon^X(\Phi_\varepsilon^Y(p)).$$

Taylor expansion in local coordinates gives

$$P_\varepsilon - Q_\varepsilon = \varepsilon^2[X, Y]_p + O(\varepsilon^3).$$

The first-order motions agree; the bracket records the leading discrepancy. The difference above is taken in coordinates, and its leading coefficient represents a tangent vector at p .

If $[X, Y] = 0$ throughout a neighborhood, the local flows commute there for sufficiently small times. Vanishing at just one point only removes the second-order discrepancy at that point.

On the plane, use the fields

$$X = \partial_x, \quad Y = x \partial_y.$$

Since $X(x) = 1$ and Y does not change the constant coefficients of X ,

$$[X, Y] = \partial_y.$$

Their flows are explicit:

$$\Phi_s^X(x, y) = (x + s, y), \quad \Phi_t^Y(x, y) = (x, y + tx).$$

Following X and then Y gives

$$\Phi_t^Y(\Phi_s^X(x, y)) = (x + s, y + tx + st).$$

Reversing the order gives

$$\Phi_s^X(\Phi_t^Y(x, y)) = (x + s, y + tx).$$

The difference is exactly $st \partial_y$.

Moving to the right first increases the upward speed prescribed by Y . This is the geometric source of the nonzero bracket.

For another example, compare dilation and translation:

$$X = x \partial_x + y \partial_y, \quad Y = \partial_x.$$

Here

$$D_X Y = 0, \quad D_Y X = \partial_x, \quad [X, Y] = -\partial_x.$$

The flow of X dilates from the origin; the flow of Y translates horizontally:

$$\Phi_s^X(x, y) = (e^s x, e^s y), \quad \Phi_t^Y(x, y) = (x + t, y).$$

Dilation followed by translation and translation followed by dilation have respective endpoints

$$(e^s x + t, e^s y), \quad (e^s(x + t), e^s y).$$

For $s = t = \varepsilon$, their difference is

$$(\varepsilon(1 - e^\varepsilon), 0) = -\varepsilon^2(1, 0) + O(\varepsilon^3).$$

Dilation after translation also enlarges the displacement created by the translation. The sign agrees with $[X, Y] = -\partial_x$.

In polar coordinates on a plane, the coordinate tangent fields commute:

$$[\partial_r, \partial_\theta] = 0.$$

Their lengths are 1 and r , respectively. The corresponding *unit* fields are

$$e_r = \partial_r, \quad e_\theta = \frac{1}{r}\partial_\theta.$$

Use the bracket product rule

$$[X, \varphi Y] = X(\varphi)Y + \varphi[X, Y].$$

It gives

$$[e_r, e_\theta] = \partial_r\left(\frac{1}{r}\right)\partial_\theta = -\frac{1}{r^2}\partial_\theta = -\frac{1}{r}e_\theta.$$

Moving radially outward decreases the angular speed needed to move at unit speed around a circle.

The orthonormal frame does not commute, even though the coordinate frame does. In particular, a nonzero bracket can occur on a flat plane: it is not a test for curvature.

Theorem (Kühnel 4.6). The covariant derivative is determined by the first fundamental form. In particular, it is intrinsic, even though we defined it using an ambient derivative and a projection.

Fix coordinates and write $g_{ij} = I(f_i, f_j)$. Introduce the scalar coefficients

$$\Gamma_{ij,k} = I(\nabla_{f_i} f_j, f_k).$$

Since the coordinate fields commute, torsion-freeness gives

$$\Gamma_{ij,k} = \Gamma_{ji,k}.$$

To prove the theorem, it is enough to recover these coefficients from g . Indeed, the inner products of a tangent vector with every basis vector determine that vector uniquely.

Then the product rule determines $\nabla_X Y$ for arbitrary fields X, Y from the derivatives $\nabla_{f_i} f_j$. The remaining step is to compute the $\Gamma_{ij,k}$.

Recall metric compatibility:

$$X(I(Y, Z)) = I(\nabla_X Y, Z) + I(Y, \nabla_X Z).$$

Applying it to coordinate fields supplies three equations:

$$\partial_i g_{jk} = \Gamma_{ij,k} + \Gamma_{ik,j},$$

$$\partial_j g_{ik} = \Gamma_{ji,k} + \Gamma_{jk,i},$$

$$\partial_k g_{ij} = \Gamma_{ki,j} + \Gamma_{kj,i}.$$

For example, the first follows by differentiating $g_{jk} = I(f_j, f_k)$ in direction f_i .

Add the first two equations and subtract the third. Using symmetry in the first two indices of Γ , the unwanted terms cancel:

$$2\Gamma_{ij,k} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}.$$

Therefore

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}).$$

Only the metric and its first derivatives appear. The second fundamental form and the normal vector have disappeared.

These coefficients are called the Christoffel symbols of the first kind (the terminology is introduced in 4.7).

Let $(g^{k\ell})$ be the inverse matrix of $(g_{k\ell})$, and write

$$\nabla_{f_i} f_j = \sum_{k=1}^n \Gamma_{ij}^k f_k.$$

Taking the inner product with f_ℓ gives $\Gamma_{ij,\ell} = \sum_k \Gamma_{ij}^k g_{k\ell}$, hence

$$\Gamma_{ij}^k = \sum_{\ell=1}^n g^{k\ell} \Gamma_{ij,\ell}.$$

Thus these coefficients, too, are determined by g . For $X = \sum_i \xi^i f_i$ and $Y = \sum_j \eta^j f_j$, the product rule now yields

$$\nabla_X Y = \sum_{k=1}^n \left(X(\eta^k) + \sum_{i,j=1}^n \xi^i \eta^j \Gamma_{ij}^k \right) f_k.$$

This proves Theorem 4.6. Knowing the metric as a function on a neighborhood, and the coefficient functions of X, Y , determines the covariant derivative.

The value of g at a single point is not enough: its first derivatives are needed. The conclusion concerns coordinate coefficients under the identification of tangent bases, not equality of ambient vectors on different hypersurfaces.

Compare the plane and the cylinder parametrizations

$$f(u, v) = (u, v, 0), \quad \widetilde{f}(u, v) = (\cos u, \sin u, v).$$

Restrict u to an interval of length less than 2π , so both are embedded patches. Both metric matrices are

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

All first derivatives of the metric vanish, so all $\Gamma_{ij,k}$ and Γ_{ij}^k vanish. On either patch, covariant differentiation simply differentiates the coefficients in the coordinate tangent basis.

Nevertheless, the ambient derivatives are different:

$$D_{f_u} f_u = 0, \quad D_{\widetilde{f}_u} \widetilde{f}_u = (-\cos u, -\sin u, 0).$$

Their tangential components both vanish.

The correspondence $f(u, v) \mapsto \widetilde{f}(u, v)$ preserves the first fundamental form and the covariant derivative, but not the second fundamental form. This is the distinction between intrinsic and extrinsic geometry.

The preceding proof gives both kinds of *Christoffel symbols*:

$$\Gamma_{ij,k} = I(\nabla_{f_i} f_j, f_k), \quad \nabla_{f_i} f_j = \sum_{k=1}^n \Gamma_{ij}^k f_k.$$

The first kind consists of inner products; the second kind consists of coefficients in the coordinate basis. They agree at a point if this basis is orthonormal there.

Writing $(g^{k\ell}) = (g_{k\ell})^{-1}$, the explicit formula is

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{\ell=1}^n g^{k\ell} (\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}).$$

Both kinds are symmetric in i, j , because coordinate fields commute and the connection is torsion-free.

In practice, compute g , its inverse, and its nonzero first derivatives; then use symmetry to reduce the work. These symbols describe the tangential change of the coordinate basis. They depend on coordinates, and can be nonzero even on a plane.

The *Gauss and Weingarten equations* describe the derivatives of the tangent basis and the normal.

The tangential-normal decomposition gives

$$f_{ij} = \sum_{k=1}^n \Gamma_{ij}^k f_k + h_{ij} \nu, \quad h_{ij} = \langle f_{ij}, \nu \rangle.$$

For the normal, differentiating $\langle \nu, f_j \rangle = 0$ gives

$$\langle \nu_i, f_j \rangle = -h_{ij}.$$

Also $\langle \nu_i, \nu \rangle = 0$, so ν_i is tangent. Solving for its tangent coefficients gives

$$\nu_i = - \sum_{j,k=1}^n h_{ij} g^{jk} f_k.$$

These are Kühnel's equations 4.8. They agree with $L(f_i) = -\nu_i$.

The Γ_{ij}^k describe the tangential part of the changing basis; h_{ij} describes its normal part. This is analogous to expressing derivatives of the Frenet frame in terms of the frame itself.

Consider polar coordinates on the Euclidean plane:

$$f(r, \theta) = (r \cos \theta, r \sin \theta, 0), \quad r > 0.$$

Work on an angular interval so that this is a chart. Then

$$g = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}.$$

The only nonzero derivative of a metric coefficient is $\partial_r g_{\theta\theta} = 2r$. Hence

$$\Gamma_{\theta\theta}^r = -r, \quad \Gamma_{r\theta}^\theta = \Gamma_{\theta r}^\theta = \frac{1}{r},$$

and all other symbols of the second kind vanish. For example,

$$\Gamma_{\theta\theta}^r = -\frac{1}{2}\partial_r(r^2) = -r.$$

Thus

$$\nabla_{f_\theta} f_\theta = -r f_r, \quad \nabla_{f_r} f_\theta = \frac{1}{r} f_\theta.$$

The plane is flat, but the polar coordinate basis varies. Its Christoffel symbols need not vanish.

Let $c(t) = f(u(t))$ be a curve in the patch, where

$$u(t) = (u^1(t), \dots, u^n(t)).$$

A ***tangent vector field along c*** assigns

$$Y(t) \in T_{c(t)}M$$

smoothly to each parameter value. It need not be defined at other points of the surface.

Define its ***covariant derivative along the curve*** by

$$\frac{\nabla Y}{dt} = \nabla_{\dot{c}} Y = \left(\frac{dY}{dt} \right)^\top.$$

If $\nu(t)$ is a unit normal along c , this means

$$\frac{\nabla Y}{dt} = \dot{Y} - \langle \dot{Y}, \nu \rangle \nu.$$

An extension of Y to a neighborhood on the surface is not needed. If such an extension exists, the chain rule recovers the definition from Lecture 11.

At a self-intersection, Y may take different values at different times, even though the point $c(t)$ is the same. A field along a curve is indexed by the parameter.

To compute a derivative along a curve, write

$$Y(t) = \sum_{j=1}^n \eta^j(t) f_j(u(t)), \quad \dot{c}(t) = \sum_{i=1}^n \dot{u}^i(t) f_i(u(t)).$$

Ordinary differentiation gives

$$\dot{Y} = \sum_{k=1}^n \dot{\eta}^k f_k + \sum_{i,j=1}^n \dot{u}^i \eta^j f_{ij}.$$

Project tangentially and use the Gauss formula:

$$\frac{\nabla Y}{dt} = \sum_{k=1}^n \left(\dot{\eta}^k + \sum_{i,j=1}^n \dot{u}^i \eta^j \Gamma_{ij}^k \right) f_k.$$

All basis vectors and symbols are evaluated at $u(t)$.

There are two contributions: the coefficients η^k change, and the coordinate basis changes. The Christoffel terms account for the latter.

In particular, constant coefficients η^k do not usually imply $\nabla_{\dot{c}} Y = 0$.

A tangent field $Y(t)$ is *parallel along* c if

$$\frac{\nabla Y}{dt} = 0.$$

Geometrically, its ambient change has no tangential component.

In coordinates, parallelism is the system

$$\dot{\eta}^k + \sum_{i,j=1}^n \dot{u}^i \eta^j \Gamma_{ij}^k(u(t)) = 0, \quad k = 1, \dots, n.$$

For the fixed curve $u(t)$, this is a linear system

$$\dot{\eta} = -A(t)\eta, \quad A_j^k(t) = \sum_{i=1}^n \dot{u}^i(t) \Gamma_{ij}^k(u(t)).$$

The unknowns are the components of Y , not the curve.

A field on an entire open patch is called parallel if $\nabla_X Y = 0$ for every tangent field X . This is a much stronger requirement. Parallel fields always exist along a given curve, but nonzero parallel fields on an entire open patch need not exist.

Theorem (Kühnel 4.10). Given a curve $c(t)$, an initial time t_0 , and $Y_0 \in T_{c(t_0)}M$, there is a unique parallel field along the curve with $Y(t_0) = Y_0$.

In a chart, the initial vector determines $\eta(t_0)$. The system

$$\dot{\eta} = -A(t)\eta$$

is linear with continuous coefficients, so it has a unique solution throughout the curve's parameter interval. Across charts, uniqueness makes the local solutions agree.

For a curve $c : [a, b] \rightarrow M$, the resulting map

$$P_c : T_{c(a)}M \longrightarrow T_{c(b)}M, \quad P_c(Y_0) = Y(b),$$

is called ***parallel transport***.

The map P_c is linear because the differential equation is linear. Transport along the reversed curve gives its inverse, by uniqueness. Transport along successive curves is the composition of the transport maps.

The same definition therefore works for piecewise smooth curves, by transporting along each piece.

Parallel transport preserves the first fundamental form. Metric compatibility along a curve says

$$\frac{d}{dt}I(Y, Z) = I\left(\frac{\nabla Y}{dt}, Z\right) + I\left(Y, \frac{\nabla Z}{dt}\right).$$

If Y and Z are parallel, the right side vanishes. Therefore

$$I_{c(b)}(P_c Y_0, P_c Z_0) = I_{c(a)}(Y_0, Z_0).$$

In particular, parallel transport preserves lengths and angles. This includes Kühnel's Corollary 4.11.

Thus P_c is a linear isometry between two tangent spaces. It supplies a way to compare vectors based at different points, without identifying them as ambient vectors.

Changing the speed of traversal does not change transport along the same oriented path: the derivative along the reparametrized curve acquires a common scalar factor.

Changing the *path* can change the answer, even if the endpoints stay fixed. We will see this on the sphere.

In the plane, a parallel vector is constant in ambient coordinates. Its polar coefficients need not be constant.

Along the circle

$$c(t) = (R \cos t, R \sin t, 0), \quad R > 0,$$

take the constant vector $Y(t) = (1, 0, 0)$. In the polar basis,

$$Y = \cos t f_r - \frac{\sin t}{R} f_\theta.$$

The parallel equations, using $r = R$ and $\theta = t$, are

$$\dot{\eta}^r - R\eta^\theta = 0, \quad \dot{\eta}^\theta + \frac{1}{R}\eta^r = 0.$$

The coefficients

$$\eta^r = \cos t, \quad \eta^\theta = -\frac{\sin t}{R}$$

satisfy both equations. Work locally in angular charts; the vector field itself is defined around the whole circle.

By contrast, the radial unit field f_r has constant coefficients $(1, 0)$ but is not parallel along this circle. Its direction rotates in the ambient plane.

A regular curve $c(t)$ is an *affinely parametrized geodesic* if its velocity is parallel along itself:

$$\nabla_{\dot{c}} \dot{c} = 0.$$

Since $D_{\dot{c}} \dot{c} = \ddot{c}$, this is equivalent to

$$(\ddot{c})^\top = 0.$$

Thus a geodesic has no tangential acceleration. Its ambient acceleration may be nonzero, but it must be normal to the surface.

The tangential-normal decomposition gives

$$\ddot{c} = \nabla_{\dot{c}} \dot{c} + II(\dot{c}, \dot{c})\nu.$$

Also, metric compatibility implies

$$\frac{d}{dt} \|\dot{c}\|^2 = 2I(\nabla_{\dot{c}} \dot{c}, \dot{c}) = 0.$$

Every affinely parametrized geodesic has constant speed. A non-constant one can be rescaled to unit speed; we then write its parameter as s and its velocity as $c'(s)$.

Substitute $Y = \dot{c}$, so that $\eta^k = \dot{u}^k$, into the formula for differentiation along a curve. The **geodesic equations** are

$$\ddot{u}^k + \sum_{i,j=1}^n \Gamma_{ij}^k(u) \dot{u}^i \dot{u}^j = 0, \quad k = 1, \dots, n.$$

Theorem (Kühnel 4.12). Given $p \in M$ and $v \in T_p M$, there is a unique local solution with $c(0) = p$, $\dot{c}(0) = v$. If $\|v\| = 1$, it is parametrized by arclength.

For the proof, introduce $w^k = \dot{u}^k$. The equations become

$$\dot{u}^k = w^k, \quad \dot{w}^k = - \sum_{i,j=1}^n \Gamma_{ij}^k(u) w^i w^j.$$

This is a smooth first-order system for (u, w) . The initial position and velocity determine its initial values, so the ordinary existence and uniqueness theorem applies. Constant speed proves the unit-speed assertion.

Unlike parallel transport along a prescribed curve, here the curve itself is unknown. Existence is initially local; global existence does not follow from this argument.

Parallel transport can depend on the path. On the unit sphere in $\mathbb{R}^3$, follow the three quarter great circles

$$e_1 \longrightarrow e_2 \longrightarrow e_3 \longrightarrow e_1,$$

where e_1, e_2, e_3 are the standard basis vectors. Start with the tangent vector $Y_0 = e_2$ at e_1 .

Along the first arc, Y is the unit tangent to the great circle. It is parallel because the arc is geodesic, and at e_2 it equals $-e_1$.

Along the second arc, the constant ambient vector $-e_1$ is tangent and parallel. It arrives at e_3 unchanged.

Parametrize the third arc by

$$c(s) = \sin s \, e_1 + \cos s \, e_3, \quad 0 \leq s \leq \pi/2.$$

Here the parallel field $-c'(s)$ starts at $-e_1$ and ends at e_3 . Therefore the complete transport sends

$$e_2 \longmapsto e_3 \quad \text{in the same tangent plane } T_{e_1} S^2.$$

The endpoint is unchanged, but the vector has turned through a right angle. Length and angle preservation do not imply path independence.