

Chapter 2

Derivatives

This chapter introduces several important notions of derivatives of tensors. In chapters 5 and 6 we also introduce partial derivatives of functions into Riemannian manifolds.

The main goal is the construction of the connection and its use as covariant differentiation. We give a motivation of this concept that depends on exterior and Lie derivatives. Covariant differentiation, in turn, allows for nice formulas for exterior derivatives, Lie derivatives, divergence and much more. It is also crucial in the development of curvature which is the central construction in Riemannian geometry.

Surprisingly, the idea of a connection postdates Riemann's introduction of the curvature tensor. Riemann discovered the Riemannian curvature tensor as a second-order term in a Taylor expansion of a Riemannian metric at a point with respect to a suitably chosen coordinate system. Lipschitz, Killing, and Christoffel introduced the connection in various ways as an intermediate step in computing the curvature. After this early work by the above-mentioned German mathematicians, an Italian school around Levi-Civita, Ricci, Bianchi et al. began systematically to study Riemannian metrics and tensor analysis. They eventually defined parallel translation and through that clarified the use of the connection. Hence the name Levi-Civita connection for the Riemannian connection. Most of their work was still local in nature and mainly centered on developing tensor analysis as a tool for describing physical phenomena such as stress, torque, and divergence. At the beginning of the twentieth century Minkowski started developing the geometry of space-time as a mathematical model for Einstein's new special relativity theory. It was this work that eventually enabled Einstein to give a geometric formulation of general relativity theory. Since then, tensor calculus, connections, and curvature have become an indispensable language for many theoretical physicists.

Much of what we do in this chapter carries over to the pseudo-Riemannian setting as long as we keep in mind how to calculate traces in this context.

2.1 Lie Derivatives

2.1.1 Directional Derivatives

There are many ways of denoting the *directional derivative* of a function on a manifold. Given a function $f : M \rightarrow \mathbb{R}$ and a vector field Y on M we will use the following ways of writing the directional derivative of f in the direction of Y

$$\nabla_Y f = D_Y f = L_Y f = df(Y) = Y(f).$$

If we have a function $f : M \rightarrow \mathbb{R}$ on a manifold, then the differential $df : TM \rightarrow \mathbb{R}$ measures the change in the function. In local coordinates, $df = \partial_i(f)dx^i$. If, in addition, M is equipped with a Riemannian metric g , then we also have the *gradient* of f , denoted by $\text{grad}f = \nabla f$, defined as the vector field satisfying $g(v, \nabla f) = df(v)$ for all $v \in TM$. In local coordinates this reads, $\nabla f = g^{ij}\partial_i(f)\partial_j$, where g^{ij} is the inverse of the matrix g_{ij} (see also section 1.5.1). Defined in this way, the gradient clearly depends on the metric.

But is there a way of defining a gradient vector field of a function without using Riemannian metrics? The answer is no and can be understood as follows. On \mathbb{R}^n the gradient is defined as

$$\nabla f = \delta^{ij}\partial_i(f)\partial_j = \sum_{i=1}^n \partial_i(f)\partial_i.$$

But this formula depends on the fact that we used Cartesian coordinates. If instead we use polar coordinates on \mathbb{R}^2 , say, then

$$\nabla f = \partial_x(f)\partial_x + \partial_y(f)\partial_y \neq \partial_r(f)\partial_r + \partial_\theta(f)\partial_\theta,$$

One rule of thumb for items that are invariantly defined is that they should satisfy the Einstein summation convention. Thus, $df = \partial_i(f)dx^i$ is invariantly defined, while $\nabla f = \partial_i(f)\partial_i$ is not. The metric $g = g_{ij}dx^i dx^j$ and gradient $\nabla f = g^{ij}\partial_i(f)\partial_j$ are invariant expressions that also depend on our choice of metric.

2.1.2 Lie Derivatives

Let X be a vector field and F^t the corresponding locally defined flow on a smooth manifold M . Thus $F^t(p)$ is defined for small t and the curve $t \mapsto F^t(p)$ is the integral curve for X that goes through p at $t = 0$. The *Lie derivative* of a tensor in the direction of X is defined as the first-order term in a suitable Taylor expansion of the tensor when it is moved by the flow of X . The precise formula, however, depends on what type of tensor we use.

If $f : M \rightarrow \mathbb{R}$ is a function, then

$$f(F^t(p)) = f(p) + t(L_X f)(p) + o(t),$$

or

$$(L_X f)(p) = \lim_{t \rightarrow 0} \frac{f(F^t(p)) - f(p)}{t}.$$

Thus the Lie derivative $L_X f$ is simply the directional derivative $D_X f = df(X)$. Without specifying p we can also write

$$f \circ F^t = f + tL_X f + o(t) \text{ and } L_X f = D_X f = df(X).$$

When we have a vector field Y things get a little more complicated as $Y|_{F^t}$ can't be compared directly to Y since the vectors live in different tangent spaces. Thus we consider the curve $t \mapsto DF^{-t}(Y|_{F^t(p)})$ that lies in $T_p M$. When this is expanded in t near 0 we obtain an expression

$$DF^{-t}(Y|_{F^t(p)}) = Y|_p + t(L_X Y)|_p + o(t)$$

for some vector $(L_X Y)|_p \in T_p M$. In other words we define

$$(L_X Y)|_p = \lim_{t \rightarrow 0} \frac{DF^{-t}(Y|_{F^t(p)}) - Y|_p}{t}.$$

This Lie derivative turns out to be the Lie bracket.

Proposition 2.1.1. *If X, Y are vector fields on M , then $L_X Y = [X, Y]$.*

Proof. While Lie derivatives are defined as a limit of suitable difference quotients it is generally far more convenient to work with their implicit definition through the first-order Taylor expansion.

The Lie derivative comes from

$$DF^{-t}(Y|_{F^t}) = Y + tL_X Y + o(t)$$

or equivalently

$$Y|_{F^t} - DF^t(Y) = tDF^t(L_X Y) + o(t).$$

Consider the directional derivative of a function f in the direction of $Y|_{F^t} - DF^t(Y)$

$$\begin{aligned} D_{Y|_{F^t} - DF^t(Y)} f &= D_{Y|_{F^t}} f - D_{DF^t(Y)} f \\ &= (D_Y f) \circ F^t - D_Y (f \circ F^t) \end{aligned}$$

$$\begin{aligned}
&= D_Y f + t D_X D_Y f + o(t) \\
&\quad - D_Y (f + t D_X f + o(t)) \\
&= t (D_X D_Y f - D_Y D_X f) + o(t) \\
&= t D_{[X, Y]} f + o(t).
\end{aligned}$$

This shows that

$$\begin{aligned}
L_X Y &= \lim_{t \rightarrow 0} \frac{Y|_{F^t} - DF^t(Y)}{t} \\
&= [X, Y].
\end{aligned}$$

□

We are now ready to define the Lie derivative of a $(0, k)$ -tensor T and also give an algebraic formula for this derivative. Define

$$(F^t)^* T = T + t (L_X T) + o(t)$$

or with variables included

$$\begin{aligned}
((F^t)^* T)(Y_1, \dots, Y_k) &= T(DF^t(Y_1), \dots, DF^t(Y_k)) \\
&= T(Y_1, \dots, Y_k) + t (L_X T)(Y_1, \dots, Y_k) + o(t).
\end{aligned}$$

As a difference quotient this means

$$(L_X T)(Y_1, \dots, Y_k) = \lim_{t \rightarrow 0} \frac{(F^t)^* T - T}{t}.$$

Proposition 2.1.2. *If X is a vector field and T a $(0, k)$ -tensor on M , then*

$$(L_X T)(Y_1, \dots, Y_k) = D_X(T(Y_1, \dots, Y_k)) - \sum_{i=1}^k T(Y_1, \dots, L_X Y_i, \dots, Y_k).$$

Proof. We restrict attention to the case where $k = 1$. The general case is similar but requires more notation. Using that

$$Y|_{F^t} = DF^t(Y) + t DF^t(L_X Y) + o(t)$$

we get

$$((F^t)^* T)(Y) = T(DF^t(Y))$$

$$\begin{aligned}
&= T(Y|_{F^t} - tDF^t(L_X Y)) + o(t) \\
&= T(Y) \circ F^t - tT(DF^t(L_X Y)) + o(t) \\
&= T(Y) + tD_X(T(Y)) - tT(DF^t(L_X Y)) + o(t).
\end{aligned}$$

Thus

$$\begin{aligned}
(L_X T)(Y) &= \lim_{t \rightarrow 0} \frac{((F^t)^* T)(Y) - T(Y)}{t} \\
&= \lim_{t \rightarrow 0} (D_X(T(Y)) - T(DF^t(L_X Y))) \\
&= D_X(T(Y)) - T(L_X Y).
\end{aligned}$$

□

Finally, we have that Lie derivatives satisfy all possible product rules, i.e., they are *derivations*. From the above propositions this is already obvious when multiplying functions with vector fields or $(0, k)$ -tensors.

Proposition 2.1.3. *If T_1 and T_2 be $(0, k_i)$ -tensors, then*

$$L_X(T_1 \cdot T_2) = (L_X T_1) \cdot T_2 + T_1 \cdot (L_X T_2).$$

Proof. Recall that for 1-forms and more general $(0, k)$ -tensors we define the product as

$$T_1 \cdot T_2(X_1, \dots, X_{k_1}, Y_1, \dots, Y_{k_2}) = T_1(X_1, \dots, X_{k_1}) \cdot T_2(Y_1, \dots, Y_{k_2}).$$

The proposition is then a simple consequence of the previous proposition and the product rule for derivatives of functions. □

Proposition 2.1.4. *If T is a $(0, k)$ -tensor and $f : M \rightarrow \mathbb{R}$ a function, then*

$$L_{fX}T(Y_1, \dots, Y_k) = fL_XT(Y_1, \dots, Y_k) + \sum_{i=1}^k (LY_i f) T(Y_1, \dots, X, \dots, Y_k).$$

Proof. We have that

$$\begin{aligned}
L_{fX}T(Y_1, \dots, Y_k) &= D_{fX}(T(Y_1, \dots, Y_k)) - \sum_{i=1}^k T(Y_1, \dots, L_{fX}Y_i, \dots, Y_k) \\
&= fD_X(T(Y_1, \dots, Y_k)) - \sum_{i=1}^k T(Y_1, \dots, [fX, Y_i], \dots, Y_k)
\end{aligned}$$

$$\begin{aligned}
&= fD_X(T(Y_1, \dots, Y_p)) - f \sum_{i=1}^k T(Y_1, \dots, [X, Y_i], \dots, Y_k) \\
&\quad + \sum_{i=1}^k (L_{Y_i} f) T(Y_1, \dots, X, \dots, Y_k).
\end{aligned}$$

□

The case where $X|_p = 0$ is of special interest when computing Lie derivatives. We note that $F^t(p) = p$ for all t . Thus $DF^t : T_p M \rightarrow T_p M$ and

$$\begin{aligned}
L_X Y|_p &= \lim_{t \rightarrow 0} \frac{DF^{-t}(Y|_p) - Y|_p}{t} \\
&= \frac{d}{dt} (DF^{-t})|_{t=0} (Y|_p).
\end{aligned}$$

This shows that $L_X = \frac{d}{dt} (DF^{-t})|_{t=0}$ when $X|_p = 0$. From this we see that if θ is a 1-form then $L_X \theta = -\theta \circ L_X$ at points p where $X|_p = 0$. This is a general phenomenon.

Lemma 2.1.5. *If a vector field X vanishes at p , then the Lie derivative $L_X T$ at p depends only on the value of T at p .*

Proof. We have that

$$(L_X T)(Y_1, \dots, Y_k) = D_X(T(Y_1, \dots, Y_k)) - \sum_{i=1}^k T(Y_1, \dots, L_X Y_i, \dots, Y_k).$$

So if X vanishes at p , then

$$(L_X T)(Y_1, \dots, Y_k)|_p = - \sum_{i=1}^k T(Y_1, \dots, L_X Y_i, \dots, Y_k)|_p.$$

□

It is also possible to define Lie derivatives of more general tensors and even multilinear maps on vector fields. An important instance of this is the Lie derivative of the Lie bracket $[Y, Z]$ or even the Lie derivative of the Lie derivative $L_Y T$. This is algebraically defined as

$$\begin{aligned}
(L_X L_Y)T &= L_X(L_Y T) - L_{L_X Y} T - L_Y(L_X T) \\
&= [L_X, L_Y]T - L_{[X, Y]}T.
\end{aligned}$$

Proposition 2.1.6 (The Generalized Jacobi Identity). *For all vector fields X, Y and tensors T*

$$(L_X L)_Y T = 0.$$

Proof. When T is a function this follows from the definition of the Lie bracket:

$$\begin{aligned} (L_X L)_Y f &= [L_X, L_Y]f - L_{[X, Y]}f \\ &= [D_X, D_Y]f - D_{[X, Y]}f \\ &= 0. \end{aligned}$$

When $T = Z$ is a vector field it is the usual Jacobi identity:

$$\begin{aligned} (L_X L)_Y Z &= [L_X, L_Y]Z - L_{[X, Y]}Z \\ &= [X, [Y, Z]] - [Y, [X, Z]] - [[X, Y], Z] \\ &= [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] \\ &= 0. \end{aligned}$$

When $T = \omega$ is a one-form it follows automatically from those two observations provided we know that

$$([L_X, L_Y] \omega)(Z) = [L_X, L_Y](\omega(Z)) - \omega([L_X, L_Y]Z)$$

since we then have

$$\begin{aligned} ((L_X L)_Y \omega)(Z) &= ([L_X, L_Y] \omega)(Z) - (L_{[X, Y]} \omega)(Z) \\ &= [L_X, L_Y](\omega(Z)) - \omega([L_X, L_Y]Z) \\ &\quad - L_{[X, Y]}(\omega(Z)) + \omega(L_{[X, Y]}Z) \\ &= 0. \end{aligned}$$

A few cancellations must occur for the first identity to hold. Note that

$$\begin{aligned} ([L_X, L_Y] \omega)(Z) &= (L_X (L_Y \omega))(Z) - (L_Y (L_X \omega))(Z), \\ (L_X (L_Y \omega))(Z) &= L_X ((L_Y \omega)(Z)) - (L_Y \omega)(L_X Z) \\ &= L_X (L_Y (\omega(Z))) - L_X (\omega(L_Y Z)) \\ &\quad - L_Y (\omega(L_X Z)) + \omega(L_Y L_X Z), \end{aligned}$$

and similarly

$$(L_Y (L_X \omega)) (Z) = L_Y (L_X (\omega (Z))) - L_Y (\omega (L_X Z)) - L_X (\omega (L_Y Z)) + \omega (L_X L_Y Z) .$$

This shows that

$$([L_X, L_Y] \omega) (Z) = [L_X, L_Y] (\omega (Z)) - \omega ([L_X, L_Y] Z) .$$

The proof for general tensors now follows by observing that these are tensor products of the above three simple types of tensors and that Lie derivatives act as derivations. \square

The Lie derivative can also be used to give a formula for the exterior derivative of a k -form

$$\begin{aligned} d\omega (X_0, X_1, \dots, X_k) &= \frac{1}{2} \sum_{i=0}^k (-1)^i (L_{X_i} \omega) (X_0, \dots, \widehat{X}_i, \dots, X_k) . \\ &\quad + \frac{1}{2} \sum_{i=0}^k (-1)^i L_{X_i} \left(\omega (X_0, \dots, \widehat{X}_i, \dots, X_k) \right) \end{aligned}$$

For a 1-form this gives us the usual definition

$$d\omega (X, Y) = D_X (\omega (Y)) - D_Y (\omega (X)) - \omega ([X, Y]) .$$

2.1.3 Lie Derivatives and the Metric

The Lie derivative allows us to define the *Hessian* of a function on a Riemannian manifold as a $(0, 2)$ -tensor:

$$\text{Hess} f (X, Y) = \frac{1}{2} (L_{\nabla f} g) (X, Y) .$$

At a critical point for f this gives the expected answer. To see this, select coordinates x^i around p such that the metric coefficients satisfy $g_{ij}|_p = \delta_{ij}$. If $df|_p = 0$, then $\nabla f|_p = 0$ and it follows that

$$\begin{aligned} L_{\nabla f} (g_{ij} dx^i dx^j) |_p &= L_{\nabla f} (g_{ij}) |_p + \delta_{ij} L_{\nabla f} (dx^i) dx^j + \delta_{ij} dx^i L_{\nabla f} (dx^j) \\ &= \delta_{ij} L_{\nabla f} (dx^i) dx^j + \delta_{ij} dx^i L_{\nabla f} (dx^j) \\ &= L_{\nabla f} (\delta_{ij} dx^i dx^j) |_p . \end{aligned}$$

Thus $\text{Hess}f|_p$ is the same if we compute it using g and the Euclidean metric in the fixed coordinate system.

It is perhaps still not clear why the Lie derivative formula for the Hessian is reasonable. The idea is that the Hessian measures how the metric changes as we flow along the gradient field. To justify this better let us define the *divergence* of a vector field X as the function $\text{div } X$ that measures how the volume form changes along the flow for X :

$$L_X \text{vol} = (\text{div } X) \text{vol}.$$

Note that the form $L_X \text{vol}$ is always exact as

$$L_X \text{vol} = di_X \text{vol},$$

where $i_X T$ evaluates T on X in the first variable.

The *Laplacian* of a function is defined as in vector calculus by

$$\Delta f = \text{div } \nabla f$$

and we claim that it is also given as the trace of the Hessian. To see this select a positively oriented orthonormal frame E_i and note that

$$\begin{aligned} \text{div } X &= (L_X \text{vol})(E_1, \dots, E_n) \\ &= L_X (\text{vol}(E_1, \dots, E_n)) \\ &\quad - \sum \text{vol}(E_1, \dots, L_X E_i, \dots, E_n) \\ &= - \sum g(L_X E_i, E_i) \\ &= \frac{1}{2} \sum (L_X (g(E_i, E_i)) - g(L_X E_i, E_i) - g(E_i, L_X E_i)) \\ &= \sum \frac{1}{2} (L_X g)(E_i, E_i). \end{aligned}$$

We can also show that the Hessian defined in this way gives us back the usual Hessian of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with the canonical metric on Euclidean space:

$$\begin{aligned} L_{\nabla f} (\delta_{ij} dx^i dx^j) &= L_{\sum \partial_j f \partial_j} \sum dx^i dx^j \\ &= \sum L_{\partial_j f \partial_j} dx^i dx^j \\ &= \sum (L_{\partial_j f \partial_j} dx^i) dx^j + \sum dx^i (L_{\partial_j f \partial_j} dx^j) \\ &= \sum \partial_j f (L_{\partial_j} dx^i) dx^j + \sum \partial_j f dx^i (L_{\partial_j} dx^j) \end{aligned}$$

$$\begin{aligned}
&= + \sum d(\partial_j f) dx^j (\partial_j) dx^i + \sum d(\partial_j f) dx^i dx^j (\partial_j) \\
&= 2 \sum d(\partial_j f) dx^j \\
&= 2 \sum \partial_{ji} f dx^j dx^i \\
&= 2 \text{Hess} f.
\end{aligned}$$

2.1.4 Lie Groups

Lie derivatives as might be expected also come in handy when working with Lie groups. For a Lie group G we have the inner automorphism $\text{Ad}_h : x \mapsto h x h^{-1}$ and its differential at $x = e$ denoted by the same letters $\text{Ad}_h : \mathfrak{g} \rightarrow \mathfrak{g}$.

Lemma 2.1.7. *The differential of $h \mapsto \text{Ad}_h$ is given by $U \mapsto \text{ad}_U(X) = [U, X]$.*

Proof. If we write $\text{Ad}_h(x) = R_{h^{-1}} L_h(x)$, then its differential at $x = e$ is given by $\text{Ad}_h = DR_{h^{-1}} DL_h$. Now let F^t be the flow for U . Then $F^t(x) = x F^t(e) = L_x(F^t(e))$ as both curves go through x at $t = 0$ and have U as tangent everywhere since U is a left-invariant vector field. This also shows that $DF^t = DR_{F^t(e)}$. Thus

$$\begin{aligned}
\text{ad}_U(X)|_e &= \frac{d}{dt} DR_{F^{-t}(e)} DL_{F^t(e)}(X|_e)|_{t=0} \\
&= \frac{d}{dt} DR_{F^{-t}(e)}(X|_{F^t(e)})|_{t=0} \\
&= \frac{d}{dt} DF^{-t}(X|_{F^t(e)})|_{t=0} \\
&= L_U X = [U, X].
\end{aligned}$$

□

This is used in the next lemma.

Lemma 2.1.8. *Let $G = \text{GL}(V)$ be the Lie group of invertible matrices on V . The Lie bracket structure on the Lie algebra $\mathfrak{gl}(V)$ of left-invariant vector fields on $\text{GL}(V)$ is given by commutation of linear maps. i.e., if $X, Y \in T_l \text{GL}(V)$, then*

$$[X, Y]|_l = XY - YX.$$

Proof. Since $x \mapsto h x h^{-1}$ is a linear map on the space $\text{Hom}(V, V)$ we see that $\text{Ad}_h(X) = h X h^{-1}$. The flow of U is given by $F^t(g) = g(I + tU + o(t))$ so we have

$$\begin{aligned}
[U, X] &= \frac{d}{dt} (F^t(I) X F^{-t}(I))|_{t=0} \\
&= \frac{d}{dt} ((I + tU + o(t)) X (I - tU + o(t)))|_{t=0} \\
&= \frac{d}{dt} (X + tUX - tXU + o(t))|_{t=0} \\
&= UX - XU.
\end{aligned}$$

□

2.2 Connections

2.2.1 Covariant Differentiation

We now come to the question of attaching a meaning to the change of a vector field. The Lie derivative is one possibility, but it is not a strong enough concept as it doesn't characterize the Cartesian coordinate fields in \mathbb{R}^n as having zero derivative. A better strategy for \mathbb{R}^n is to write $X = X^i \partial_i$, where ∂_i are the Cartesian coordinate fields. If we want the coordinate vector fields to have zero derivative, then it is natural to define the *covariant derivative* of X in the direction of Y as

$$\nabla_Y X = (\nabla_Y X^i) \partial_i = d(X^i)(Y) \partial_i.$$

Thus we measure the change in X by measuring how the coefficients change. Therefore, a vector field with constant coefficients does not change. This formula clearly depends on the fact that we used Cartesian coordinates and is not invariant under change of coordinates. If we take the coordinate vector fields

$$\partial_r = \frac{1}{r} (x\partial_x + y\partial_y), \quad \partial_\theta = -y\partial_x + x\partial_y$$

that come from polar coordinates in \mathbb{R}^2 , then we see that they are not constant.

In order to better understand such derivatives we need to find a coordinate independent definition. This is done most easily by splitting the problem of defining the change in a vector field X into two problems.

First, we can measure the change in X by asking whether or not X is a gradient field. If $i_X g = \theta_X$ is the 1-form dual to X , i.e., $(i_X g)(Y) = g(X, Y)$, then we know that X is locally the gradient of a function if and only if $d\theta_X = 0$. In general, the 2-form $d\theta_X$ then measures the extent to which X is a gradient field.

Second, we can measure how a vector field X changes the metric via the Lie derivative $L_X g$. This is a symmetric $(0, 2)$ -tensor as opposed to the skew-symmetric $(0, 2)$ -tensor $d\theta_X$. If F^t is the local flow for X , then we see that $L_X g = 0$ if and only if F^t are isometries (see also section 8.1). When this happens we say that X is a *Killing field*.

In case $X = \nabla f$ is a gradient field we saw that the expression $\frac{1}{2}L_{\nabla f}g$ is the Hessian of f . From that calculation we can also quickly see what the Killing fields on \mathbb{R}^n should be: If $X = X^i \partial_i$, then X is a Killing field if and only if $\partial_k X^i + \partial_i X^k = 0$. This implies that

$$\begin{aligned} \partial_j \partial_k X^i &= -\partial_j \partial_i X^k \\ &= -\partial_i \partial_j X^k \\ &= \partial_i \partial_k X^j \\ &= \partial_k \partial_i X^j \\ &= -\partial_k \partial_j X^i \\ &= -\partial_j \partial_k X^i. \end{aligned}$$

Thus we have $\partial_j \partial_k X^i = 0$ and hence

$$X^i = \alpha_j^i x^j + \beta^i$$

with the extra conditions that

$$\alpha_j^i = \partial_j X^i = -\partial_i X^j = -\alpha_i^j.$$

In particular, the angular field ∂_θ is a Killing field. This also follows from the fact that the corresponding flow is matrix multiplication by the orthogonal matrix

$$\begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}.$$

More generally, one can show that the flow of the Killing field X is

$$F^t(x) = \exp(At)x + t\beta, \quad A = [\alpha_j^i], \quad \beta = [\beta^i].$$

In this way we see that a vector field on \mathbb{R}^n is constant if and only if it is both a Killing field and a gradient field.

Finally we make the important observation.

Proposition 2.2.1. *The covariant derivative in \mathbb{R}^n is given by the implicit formula:*

$$2g(\nabla_Y X, Z) = (L_X g)(Y, Z) + (d\theta_X)(Y, Z).$$

Proof. Since both sides are tensorial in Y and Z it suffices to check the formula on the Cartesian coordinate vector fields. Write $X = a^i \partial_i$ and calculate the right-hand side

$$\begin{aligned}
(L_X g)(\partial_k, \partial_l) + (d\theta_X)(\partial_k, \partial_l) &= D_X \delta_{kl} - g(L_X \partial_k, \partial_l) - g(\partial_k, L_X \partial_l) \\
&\quad + \partial_k g(X, \partial_l) - \partial_l g(X, \partial_k) - g(X, [\partial_k, \partial_l]) \\
&= -g(L_{a^i \partial_i} \partial_k, \partial_l) - g(\partial_k, L_{a^i \partial_i} \partial_l) \\
&\quad + \partial_k a^l - \partial_l a^k \\
&= -g(-(\partial_k a^i) \partial_i, \partial_l) - g(\partial_k, -(\partial_l a^i) \partial_i) \\
&\quad + \partial_k a^l - \partial_l a^k \\
&= +\partial_k a^l + \partial_l a^k + \partial_k a^l - \partial_l a^k \\
&= 2\partial_k a^l \\
&= 2g((\partial_k a^i) \partial_i, \partial_l) \\
&= 2g(\nabla_{\partial_k} X, \partial_l).
\end{aligned}$$

□

Since the right-hand side in the formula for $\nabla_Y X$ makes sense on any Riemannian manifold we can use this to give an implicit definition of the *covariant derivative* of X in the direction of Y . This covariant derivative turns out to be uniquely determined by the following properties.

Theorem 2.2.2 (The Fundamental Theorem of Riemannian Geometry). *The assignment $X \mapsto \nabla X$ on (M, g) is uniquely defined by the following properties:*

(1) $Y \mapsto \nabla_Y X$ is a $(1, 1)$ -tensor, i.e., it is well-defined for tangent vectors and linear

$$\nabla_{\alpha v + \beta w} X = \alpha \nabla_v X + \beta \nabla_w X.$$

(2) $X \mapsto \nabla_Y X$ is a derivation:

$$\begin{aligned}
\nabla_Y (X_1 + X_2) &= \nabla_Y X_1 + \nabla_Y X_2, \\
\nabla_Y (fX) &= (D_Y f) X + f \nabla_Y X
\end{aligned}$$

for functions $f : M \rightarrow \mathbb{R}$.

(3) Covariant differentiation is torsion free:

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

(4) Covariant differentiation is metric:

$$D_Z g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y).$$

Proof. We have already established (1) by using that

$$(L_X g)(Y, Z) + (d\theta_X)(Y, Z)$$

is tensorial in Y and Z . This also shows that the expression is linear in X . To check the derivation rule we observe that

$$\begin{aligned} L_{fX}g + d\theta_{fX} &= fL_Xg + df \cdot \theta_X + \theta_X \cdot df + d(f\theta_X) \\ &= fL_Xg + df \cdot \theta_X + \theta_X \cdot df + df \wedge \theta_X + fd\theta_X \\ &= f(L_Xg + d\theta_X) + df \cdot \theta_X + \theta_X \cdot df + df \cdot \theta_X - \theta_X \cdot df \\ &= f(L_Xg + d\theta_X) + 2df \cdot \theta_X. \end{aligned}$$

Thus

$$\begin{aligned} 2g(\nabla_Y(fX), Z) &= f2g(\nabla_YX, Z) + 2df(Y)g(X, Z) \\ &= 2g(f\nabla_YX + df(Y)X, Z) \\ &= 2g(f\nabla_YX + (D_Yf)X, Z). \end{aligned}$$

To establish the next two claims it is convenient to create the following expansion also known as *Koszul's formula*.

$$\begin{aligned} 2g(\nabla_YX, Z) &= (L_Xg)(Y, Z) + (d\theta_X)(Y, Z) \\ &= D_Xg(Y, Z) - g([X, Y], Z) - g(Y, [X, Z]) \\ &\quad + D_Y\theta_X(Z) - D_Z\theta_X(Y) - \theta_X([Y, Z]) \\ &= D_Xg(Y, Z) - g([X, Y], Z) - g(Y, [X, Z]) \\ &\quad + D_Yg(X, Z) - D_Zg(X, Y) - g(X, [Y, Z]) \\ &= D_Xg(Y, Z) + D_Yg(Z, X) - D_Zg(X, Y) \\ &\quad - g([X, Y], Z) - g([Y, Z], X) + g([Z, X], Y). \end{aligned}$$

We then see that (3) follows from

$$\begin{aligned} 2g(\nabla_XY - \nabla_YX, Z) &= D_Yg(X, Z) + D_Xg(Z, Y) - D_Zg(Y, X) \\ &\quad - g([Y, X], Z) - g([X, Z], Y) + g([Z, Y], X) \\ &\quad - D_Xg(Y, Z) - D_Yg(Z, X) + D_Zg(X, Y) \\ &\quad + g([X, Y], Z) + g([Y, Z], X) - g([Z, X], Y) \\ &= 2g([X, Y], Z). \end{aligned}$$

And (4) from

$$\begin{aligned}
 2g(\nabla_Z X, Y) + 2g(X, \nabla_Z Y) &= D_X g(Z, Y) + D_Z g(Y, X) - D_Y g(X, Z) \\
 &\quad -g([X, Z], Y) - g([Z, Y], X) + g([Y, X], Z) \\
 &\quad + D_Y g(Z, X) + D_Z g(X, Y) - D_X g(Y, Z) \\
 &\quad -g([Y, Z], X) - g([Z, X], Y) + g([X, Y], Z) \\
 &= 2D_Z g(X, Y).
 \end{aligned}$$

Conversely, if we have a covariant derivative $\bar{\nabla}_Y X$ with these four properties, then

$$\begin{aligned}
 2g(\nabla_Y X, Z) &= (L_X g)(Y, Z) + (d\theta_X)(Y, Z) \\
 &= D_X g(Y, Z) + D_Y g(Z, X) - D_Z g(X, Y) \\
 &\quad -g([X, Y], Z) - g([Y, Z], X) + g([Z, X], Y) \\
 &= g(\bar{\nabla}_X Y, Z) + g(Y, \bar{\nabla}_X Z) + g(\bar{\nabla}_Y Z, X) + g(Z, \bar{\nabla}_Y X) \\
 &\quad -g(\bar{\nabla}_Z X, Y) - g(X, \bar{\nabla}_Z Y) + g(\bar{\nabla}_Z Y, X) - g(\bar{\nabla}_X Z, Y) \\
 &\quad -g(\bar{\nabla}_X Y, Z) + g(\bar{\nabla}_Y X, Z) - g(\bar{\nabla}_Y Z, X) + g(\bar{\nabla}_Z Y, X) \\
 &= 2g(\bar{\nabla}_Y X, Z)
 \end{aligned}$$

showing that $\nabla_Y X = \bar{\nabla}_Y X$. □

Any assignment on a manifold that satisfies (1) and (2) is called an *affine connection*. If (M, g) is a Riemannian manifold and we have a connection that in addition also satisfies (3) and (4), then we call it a *Riemannian connection*. As we just saw, this connection is uniquely defined by these four properties and is given implicitly through the formula

$$\begin{aligned}
 2g(\nabla_Y X, Z) &= (L_X g)(Y, Z) + (d\theta_X)(Y, Z) \\
 &= D_X g(Y, Z) + D_Y g(Z, X) - D_Z g(X, Y) \\
 &\quad -g([X, Y], Z) - g([Y, Z], X) + g([Z, X], Y).
 \end{aligned}$$

Before proceeding we need to discuss how $\nabla_Y X$ depends on X and Y . Since $\nabla_Y X$ is tensorial in Y , we see that the value of $\nabla_Y X$ at $p \in M$ depends only on $Y|_p$. But in what way does it depend on X ? Since $X \mapsto \nabla_Y X$ is a derivation, it is definitely not tensorial in X . Therefore, we cannot expect $(\nabla_Y X)|_p$ to depend only on $X|_p$ and $Y|_p$. The next two lemmas explore how $(\nabla_Y X)|_p$ depends on X .

Lemma 2.2.3. *Let M be a manifold and ∇ an affine connection on M . If $p \in M$, $v \in T_p M$, and X, Y are vector fields on M such that $X = Y$ in a neighborhood $U \ni p$, then $\nabla_v X = \nabla_v Y$.*

Proof. Choose $\lambda : M \rightarrow \mathbb{R}$ such that $\lambda \equiv 0$ on $M - U$ and $\lambda \equiv 1$ in a neighborhood of p . Then $\lambda X = \lambda Y$ on M . Thus at p

$$\nabla_v \lambda X = \lambda(p) \nabla_v X + d\lambda(v) \cdot X(p) = \nabla_v X$$

since $d\lambda|_p = 0$ and $\lambda(p) = 1$. In particular,

$$\nabla_v X = \nabla_v \lambda X = \nabla_v \lambda Y = \nabla_v Y.$$

□

For a Riemannian connection we could also have used the Koszul formula to prove this since the right-hand side of that formula can be localized. This lemma tells us an important thing. Namely, if a vector field X is defined only on an open subset of M , then ∇X still makes sense on this subset. Therefore, we can use coordinate vector fields or more generally frames to compute ∇ locally.

Lemma 2.2.4. *Let M be a manifold and ∇ an affine connection on M . If X is a vector field on M and $c : I \rightarrow M$ a smooth curve with $\dot{c}(0) = v \in T_p M$, then $\nabla_v X$ depends only on the values of X along c , i.e., if $X \circ c = Y \circ c$, then $\nabla_{\dot{c}} X = \nabla_{\dot{c}} Y$.*

Proof. Choose a frame E_1, \dots, E_n in a neighborhood of p and write $Y = \sum Y^i E_i$, $X = \sum X^i E_i$ on this neighborhood. From the assumption that $X \circ c = Y \circ c$ we get that $X^i \circ c = Y^i \circ c$. Thus,

$$\begin{aligned} \nabla_v Y &= \nabla_v (Y^i E_i) \\ &= Y^i(p) \nabla_v E_i + E_i(p) dY^i(v) \\ &= X^i(p) \nabla_v E_i + E_i(p) dX^i(v) \\ &= \nabla_v X. \end{aligned}$$

□

This shows that $\nabla_v X$ makes sense as long as X is prescribed along some curve (or submanifold) that has v as a tangent.

It will occasionally be convenient to use coordinates or orthonormal frames with certain nice properties. We say that a coordinate system is *normal* at p if $g_{ij}|_p = \delta_{ij}$ and $\partial_k g_{ij}|_p = 0$. An orthonormal frame E_i is *normal* at $p \in M$ if $\nabla_v E_i(p) = 0$ for all $i = 1, \dots, n$ and $v \in T_p M$. It is not hard to show that such coordinates and frames always exist (see exercises 2.5.20 and 2.5.19).

2.2.2 Covariant Derivatives of Tensors

The connection, as we shall see, is also useful in generalizing many of the well-known concepts (such as Hessian, Laplacian, divergence) from multivariable calculus to the Riemannian setting (see also section 2.1.3).

If S is a (s, t) -tensor field, then we can define a *covariant derivative* ∇S that we interpret as an $(s, t + 1)$ -tensor field. Recall that a vector field X is a $(1, 0)$ -tensor field and ∇X is a $(1, 1)$ -tensor field. The main idea is to make sure that Leibniz' rule holds. So for a $(1, 1)$ -tensor S we should have

$$\nabla_X (S(Y)) = (\nabla_X S)(Y) + S(\nabla_X Y).$$

Therefore, it seems reasonable to define ∇S as

$$\begin{aligned} \nabla S(X, Y) &= (\nabla_X S)(Y) \\ &= \nabla_X (S(Y)) - S(\nabla_X Y). \end{aligned}$$

In other words

$$\nabla_X S = [\nabla_X, S].$$

It is easily checked that $\nabla_X S$ is still tensorial in Y .

More generally, when $s = 0, 1$ we obtain

$$\begin{aligned} \nabla S(X, Y_1, \dots, Y_r) &= (\nabla_X S)(Y_1, \dots, Y_r) \\ &= \nabla_X (S(Y_1, \dots, Y_r)) - \sum_{i=1}^r S(Y_1, \dots, \nabla_X Y_i, \dots, Y_r). \end{aligned}$$

Here ∇_X is interpreted as the directional derivative when applied to a function and covariant differentiation on vector fields. This also makes sense when $s \geq 2$, if we make sense of defining covariant derivatives of, say, tensor products of vector fields. This can also be done using the product rule:

$$\nabla_X (X_1 \otimes X_2) = (\nabla_X X_1) \otimes X_2 + X_1 \otimes (\nabla_X X_2).$$

A tensor is said to be *parallel* if $\nabla S \equiv 0$. In Euclidean space one can easily show that if a tensor is written in Cartesian coordinates, then it is parallel if and only if it has constant coefficients. Thus $\nabla X \equiv 0$ for constant vector fields. On a Riemannian manifold (M, g) the metric and volume forms are always parallel.

Proposition 2.2.5. *On a Riemannian n -manifold (M, g)*

$$\begin{aligned}\nabla g &= 0, \\ \nabla \text{vol} &= 0.\end{aligned}$$

Proof. The metric is parallel due to property (4):

$$(\nabla g)(X, Y_1, Y_2) = \nabla_X (g(Y_1, Y_2)) - g(\nabla_X Y_1, Y_2) - g(Y_1, \nabla_X Y_2) = 0.$$

To check that the volume form is parallel we evaluate the covariant derivative on an orthonormal frame E_1, \dots, E_n :

$$\begin{aligned}(\nabla_X \text{vol})(E_1, \dots, E_n) &= \nabla_X \text{vol}(E_1, \dots, E_n) \\ &\quad - \sum \text{vol}(E_1, \dots, \nabla_X E_i, \dots, E_n) \\ &= - \sum g(E_i, \nabla_X E_i) \\ &= -\frac{1}{2} \sum D_X (g(E_i, E_i)) \\ &= 0.\end{aligned}$$

□

The covariant derivative gives us a different way of calculating the Hessian of a function.

Proposition 2.2.6. *If $f : (M, g) \rightarrow \mathbb{R}$, then*

$$(\nabla_X df)(Y) = g(\nabla_X \nabla f, Y) = \text{Hess} f(X, Y).$$

Proof. First observe that

$$\begin{aligned}(\nabla df)(X, Y) &= (\nabla_X df)(Y) \\ &= D_X D_Y f - df(\nabla_X Y) \\ &= D_X D_Y f - D_{\nabla_X Y} f.\end{aligned}$$

This shows that

$$(\nabla_X df)(Y) - (\nabla_Y df)(X) = [D_X, D_Y]f - D_{[X, Y]}f = 0.$$

Thus $(\nabla_X df)(Y)$ is symmetric. This can be used to establish the formulas

$$\begin{aligned}
 (\nabla df)(X, Y) &= (\nabla_X df)(Y) \\
 &= D_X g(\nabla f, Y) - g(\nabla f, \nabla_X Y) \\
 &= g(\nabla_X \nabla f, Y) \\
 &= \frac{1}{2} g(\nabla_X \nabla f, Y) + \frac{1}{2} g(X, \nabla_Y \nabla f) \\
 &= \frac{1}{2} (\nabla_{\nabla f} g)(X, Y) + \frac{1}{2} g(\nabla_X \nabla f, Y) + \frac{1}{2} g(X, \nabla_Y \nabla f) \\
 &= \frac{1}{2} D_{\nabla f} g(X, Y) - \frac{1}{2} g([\nabla f, X], Y) - \frac{1}{2} g(X, [\nabla f, Y]) \\
 &= \frac{1}{2} (L_{\nabla f} g)(X, Y).
 \end{aligned}$$

□

2.2.2.1 The Adjoint of the Covariant Derivative

The *adjoint* to the covariant derivative on (s, t) -tensors with $t > 0$ is defined as

$$(\nabla^* S)(X_2, \dots, X_r) = - \sum (\nabla_{E_i} S)(E_i, X_2, \dots, X_r),$$

where E_1, \dots, E_n is an orthonormal frame. This means that while the covariant derivative adds a variable, the adjoint eliminates one. The adjoint is related to the divergence of a vector field (see section 2.1.3) by

Proposition 2.2.7. *If X is a vector field and θ_X the corresponding 1-form, then*

$$\operatorname{div} X = -\nabla^* \theta_X.$$

Proof. See section 2.1.3 for the definition of divergence. Select an orthonormal frame E_i , then

$$\begin{aligned}
 -\nabla^* \theta_X &= \sum (\nabla_{E_i} \theta_X)(E_i) \\
 &= \sum D_{E_i} g(X, E_i) - \sum g(X, \nabla_{E_i} E_i) \\
 &= \sum g(\nabla_{E_i} X, E_i) \\
 &= \sum \frac{1}{2} (L_X g)(E_i, E_i) \\
 &= \operatorname{div} X.
 \end{aligned}$$

□

The adjoint really is the adjoint of the covariant derivative with respect to the integrated inner product.

Proposition 2.2.8. *If S is a compactly supported (s, t) -tensor and T a compactly supported $(s, t + 1)$ -tensor, then*

$$\int g(\nabla S, T) \text{vol} = \int g(S, \nabla^* T) \text{vol}.$$

Proof. Define a 1-form by $\omega(X) = g(i_X T, S)$. To calculate its divergence more easily, select an orthonormal frame E_i such that $\nabla_v E_i = 0$ for all $v \in T_p M$. To further simplify things a bit assume that $s = t = 1$, then

$$\begin{aligned} -\nabla^* \omega &= (\nabla_{E_i} \omega)(E_i) \\ &= \nabla_{E_i} g(T(E_i, E_j), S(E_j)) \\ &= g(\nabla_{E_i} T(E_i, E_j), S(E_j)) + g(T(E_i, E_j), \nabla_{E_i} S(E_j)) \\ &= -g(\nabla^* T, S) + g(T, \nabla S). \end{aligned}$$

So the result follows by the divergence theorem or Stokes' theorem:

$$\int \text{div } X \text{vol} = \int di_X \text{vol} = 0,$$

where X is any compactly supported vector field. □

2.2.2.2 Exterior Derivatives

The covariant derivative gives us a very nice formula for exterior derivatives of forms as the skew-symmetrized covariant derivative:

$$(d\omega)(X_0, \dots, X_k) = \sum (-1)^i (\nabla_{X_i} \omega)(X_0, \dots, \hat{X}_i, \dots, X_k).$$

While the covariant derivative clearly depends on the metric this formula shows that for forms we can still obtain derivatives that do not depend on the metric. It will also allow us to define exterior derivatives of more complicated tensors. Suppose we have a $(1, k)$ -tensor T that is skew-symmetric in the k variables. Then we can define the $(1, k + 1)$ -tensor

$$(d^\nabla T)(X_0, \dots, X_k) = \sum (-1)^i (\nabla_{X_i} T)(X_0, \dots, \hat{X}_i, \dots, X_k).$$

In case $k = 0$ the tensor $T = Y$ is a vector field and we obtain the $(1, 1)$ -tensor:

$$(d^\nabla Y)(X) = \nabla_X Y.$$

When $k = 1$ we have a $(1, 1)$ -tensor and obtain the $(1, 2)$ -tensor:

$$\begin{aligned} (d^\nabla T)(X, Y) &= (\nabla_X T)(Y) - (\nabla_Y T)(X) \\ &= \nabla_X(T(Y)) - \nabla_Y(T(X)) - T[X, Y]. \end{aligned}$$

2.2.2.3 The Second Covariant Derivative

For a (s, t) -tensor field S we define the *second covariant derivative* $\nabla^2 S$ as the $(s, t + 2)$ -tensor field

$$\begin{aligned} (\nabla_{X_1, X_2}^2 S)(Y_1, \dots, Y_r) &= (\nabla_{X_1}(\nabla S))(X_2, Y_1, \dots, Y_r) \\ &= (\nabla_{X_1}(\nabla_{X_2} S))(Y_1, \dots, Y_r) - (\nabla_{\nabla_{X_1} X_2} S)(Y_1, \dots, Y_r). \end{aligned}$$

With this we obtain another definition for the $(0, 2)$ version of the Hessian of a function:

$$\begin{aligned} \nabla_{X, Y}^2 f &= \nabla_X \nabla_Y f - \nabla_{\nabla_X Y} f \\ &= \nabla_X df(Y) - df(\nabla_X Y) \\ &= (\nabla_X df)(Y) \\ &= \text{Hess} f(X, Y). \end{aligned}$$

The second covariant derivative on functions is symmetric in X and Y . For more general tensors, however, this will not be the case. The defect in the second covariant derivative not being symmetric is a central feature in Riemannian geometry and is at the heart of the difference between Euclidean geometry and all other Riemannian geometries.

From the new formula for the Hessian we see that the Laplacian can be written as

$$\Delta f = -\nabla^* \nabla f = \sum_{i=1}^n \nabla_{E_i, E_i}^2 f.$$

2.2.2.4 The Lie Derivative of the Covariant Derivative

We can define the Lie derivative of the connection in a way similar to the Lie derivative of the Lie bracket

$$\begin{aligned} (L_X \nabla)_U V &= (L_X \nabla)(U, V) \\ &= L_X(\nabla_U V) - \nabla_{L_X U} V - \nabla_U L_X V \\ &= [X, \nabla_U V] - \nabla_{[X, U]} V - \nabla_U [X, V]. \end{aligned}$$

Since $[U, V] = \nabla_U V - \nabla_V U$ it follows that

$$(L_X \nabla)(U, V) - (L_X \nabla)(V, U) = L_X L_U V = 0.$$

Moreover as $\nabla_U V$ is tensorial in U the Lie derivative $(L_X \nabla)_U V$ will also be tensorial in U . The fact that it is also symmetric shows that it is tensorial in both variables.

2.2.2.5 The Covariant Derivative of the Covariant Derivative

We can also define the covariant derivative of the covariant derivative

$$(\nabla_X \nabla)_Y T = \nabla_X (\nabla_Y T) - \nabla_{\nabla_X Y} T - \nabla_Y (\nabla_X T).$$

Note however, that this is not tensorial in X !

It is related to the *second covariant derivative* of T by

$$\nabla_{X,Y}^2 T = (\nabla_X \nabla)_Y T + \nabla_Y (\nabla_X T).$$

2.3 Natural Derivations

We've seen that there are many natural derivations on tensors coming from various combinations of derivatives. We shall attempt to tie these together in a natural and completely algebraic fashion by using that all $(1, 1)$ -tensors naturally act as derivations on tensors.

For clarity we define a derivation on tensors as map $T \mapsto DT$ that preserves the type of the tensor T ; is linear; commutes with contractions; and satisfies the product rule

$$D(T_1 \otimes T_2) = (DT_1) \otimes T_2 + T_1 \otimes DT_2.$$

2.3.1 Endomorphisms as Derivations

The goal is to show that $(1, 1)$ -tensors naturally act as derivations on the space of all tensors.

We use the natural homomorphism

$$\text{GL}(V) \rightarrow \text{GL}(T(V)),$$

where $T(V)$ is the space of all tensors over the vector space V . This respects the natural grading of tensors: The subspace of (s, t) -tensors is spanned by

$$v_1 \otimes \cdots \otimes v_s \otimes \phi_1 \otimes \cdots \otimes \phi_t$$

where $v_1, \dots, v_s \in V$ and $\phi_1, \dots, \phi_t : V \rightarrow \mathbb{R}$ are linear functions. The natural homomorphism acts as follows: for $\alpha \in \mathbb{R}$ we have $g \cdot \alpha = 0$; for $v \in V$ we have $g \cdot v = g(v)$; for $\phi \in V^*$ we have $g \cdot \phi = \phi \circ g^{-1}$; and on general tensors

$$\begin{aligned} & g \cdot (v_1 \otimes \dots \otimes v_s \otimes \phi_1 \otimes \dots \otimes \phi_t) \\ &= g(v_1) \otimes \dots \otimes g(v_s) \otimes (\phi_1 \circ g^{-1}) \otimes \dots \otimes (\phi_t \circ g^{-1}). \end{aligned}$$

The derivative of this action yields a linear map

$$\text{End}(V) \rightarrow \text{End}(T(V)),$$

which for each $L \in \text{End}(V)$ induces a derivation on $T(V)$. Specifically, if $L \in \text{End}(V)$, then $Lv = L(v)$ on vectors; on 1-forms $L\phi = -\phi \circ L$; and on general tensors

$$\begin{aligned} & L(v_1 \otimes \dots \otimes v_s \otimes \phi_1 \otimes \dots \otimes \phi_t) \\ &= L(v_1) \otimes \dots \otimes v_s \otimes \phi_1 \otimes \dots \otimes \phi_t \\ &+ \dots \\ &+ v_1 \otimes \dots \otimes L(v_s) \otimes \phi_1 \otimes \dots \otimes \phi_t \\ &- v_1 \otimes \dots \otimes v_s \otimes (\phi_1 \circ L) \otimes \dots \otimes \phi_t \\ &- \dots \\ &- v_1 \otimes \dots \otimes v_s \otimes \phi_1 \otimes \dots \otimes (\phi_t \circ L). \end{aligned}$$

As the natural derivation comes from an action that preserves symmetries of tensors we immediately obtain.

Proposition 2.3.1. *The linear map*

$$\text{End}(V) \rightarrow \text{End}(T(V))$$

$$L \mapsto LT$$

is a Lie algebra homomorphism that preserves symmetries of tensors.

We also need to show that it is a derivation.

Proposition 2.3.2. *Any $(1, 1)$ -tensor L defines a derivation on tensors.*

Proof. It is easy to see from the definition that it is linear and satisfies the product rule. So it remains to show that it commutes with contractions. Consider a $(1, 1)$ -tensor T and in a local frame X_i with associated coframe σ^i write it as $T = T_j^i X_i \otimes \sigma^j$. The contraction of T is scalar valued and simply the trace of T so we know that $L(\text{tr } T) = 0$. On the other hand we have

$$\begin{aligned}
L(T) &= T_j^i L(X_i) \otimes \sigma^j - T_j^i X_i \otimes \sigma^j \circ L \\
&= T_j^i L_i^k X_k \otimes \sigma^j - T_j^i L_l^j X_i \otimes \sigma^l \\
&= T_l^i L_i^k X_k \otimes \sigma^l - T_j^k L_l^j X_k \otimes \sigma^l \\
&= (T_l^i L_i^k - T_j^k L_l^j) X_k \otimes \sigma^l
\end{aligned}$$

so

$$\text{tr}(L(T)) = T_k^i L_i^k - T_j^k L_k^j = 0.$$

A similar strategy can be used for general tensors $T_{j_1 \dots j_l}^{i_1 \dots i_k}$ where we trace or contract over a fixed superscript and subscript. \square

We also need to know how this derivation interacts with an inner product. The inner product on $T(V)$ is given by declaring

$$e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}$$

an orthonormal basis when e_1, \dots, e_n is an orthonormal basis for V and e^1, \dots, e^n the dual basis for V^* .

Proposition 2.3.3. *Assume V has an inner product:*

- (1) *The adjoint of $L : V \rightarrow V$ extends to become the adjoint for $L : T(V) \rightarrow T(V)$.*
- (2) *If $L \in \mathfrak{so}(V)$, i.e., L is skew-adjoint, then L commutes with type change of tensors.*

2.3.2 Derivatives

One can easily show that both the Lie derivative L_U and the covariant derivative ∇_U act as derivations on tensors (see exercises 2.5.9 and 2.5.10). However, these operations are nontrivial on functions. Therefore, they are not of the type we just introduced above.

Proposition 2.3.4. *If we think of ∇U as the $(1, 1)$ -tensor $X \mapsto \nabla_X U$, then*

$$L_U = \nabla_U - (\nabla U).$$

Proof. It suffices to check that this identity holds on vector fields and functions. On functions it reduces to the definition of directional derivatives, on vectors from the definition of Lie brackets and the torsion free property of the connection. \square

This proposition indicates that one can make sense of the expression $\nabla_T U$ where T is a tensor and U a vector field. It has in other places been named $A_X T$, but as that now generally has been accepted as the A -tensor for a Riemannian submersion we have not adopted this notation.

2.4 The Connection in Tensor Notation

In a local coordinate system the metric is written as $g = g_{ij}dx^i dx^j$. So if $X = X^i \partial_i$ and $Y = Y^j \partial_j$ are vector fields, then

$$g(X, Y) = g_{ij}X^i Y^j.$$

We can also compute the dual 1-form θ_X to X by:

$$\begin{aligned}\theta_X &= g(X, \cdot) \\ &= g_{ij}dx^i(X) dx^j(\cdot) \\ &= g_{ij}X^i dx^j.\end{aligned}$$

The inverse of the matrix $[g_{ij}]$ is denoted $[g^{ij}]$. Thus we have

$$\delta_j^i = g^{ik}g_{kj}.$$

The vector field X dual to a 1-form $\omega = \omega_i dx^i$ is defined implicitly by

$$g(X, Y) = \omega(Y).$$

In other words we have

$$\theta_X = g_{ij}X^i dx^j = \omega_j dx^j = \omega.$$

This shows that

$$g_{ij}X^i = \omega_j.$$

In order to isolate X^i we have to multiply by g^{kj} on both sides and also use the symmetry of g_{ij}

$$\begin{aligned}g^{kj}\omega_j &= g^{kj}g_{ij}X^i \\ &= g^{kj}g_{ji}X^i \\ &= \delta_i^k X^i \\ &= X^k.\end{aligned}$$

Therefore,

$$\begin{aligned}X &= X^i \partial_i \\ &= g^{ij}\omega_j \partial_i.\end{aligned}$$

The gradient field of a function is a particularly important example of this construction

$$\begin{aligned}\nabla f &= g^{ij} \partial_i f \partial_j, \\ df &= \partial_i f dx^i.\end{aligned}$$

We proceed to find a formula for $\nabla_Y X$ in local coordinates

$$\begin{aligned}\nabla_Y X &= \nabla_{Y^i \partial_i} X^j \partial_j \\ &= Y^i \nabla_{\partial_i} X^j \partial_j \\ &= Y^i (\partial_i X^j) \partial_j + Y^i X^j \nabla_{\partial_i} \partial_j \\ &= Y^i (\partial_i X^j) \partial_j + Y^i X^j \Gamma_{ij}^k \partial_k,\end{aligned}$$

where we simply expanded the term $\nabla_{\partial_i} \partial_j$ in local coordinates. The first part of this formula is what we expect to get when using Cartesian coordinates in \mathbb{R}^n . The second part is the correction term coming from having a more general coordinate system and also a non-Euclidean metric. Our next goal is to find a formula for Γ_{ij}^k in terms of the metric. To this end we can simply use our defining implicit formula for the connection keeping in mind that there are no Lie bracket terms. On the left-hand side we have

$$\begin{aligned}2g(\nabla_{\partial_i} \partial_j, \partial_l) &= 2g(\Gamma_{ij}^k \partial_k, \partial_l) \\ &= 2\Gamma_{ij}^k g_{kl},\end{aligned}$$

and on the right-hand side

$$\begin{aligned}(L_{\partial_j} g)(\partial_i, \partial_l) + d\theta_{\partial_j}(\partial_i, \partial_l) &= \partial_j g_{il} + \partial_i (\theta_{\partial_j}(\partial_l)) - \partial_l (\theta_{\partial_j}(\partial_i)) \\ &= \partial_j g_{il} + \partial_i g_{jl} - \partial_l g_{ji}.\end{aligned}$$

Multiplying by g^{lm} on both sides then yields

$$\begin{aligned}2\Gamma_{ij}^m &= 2\Gamma_{ij}^k \delta_k^m \\ &= 2\Gamma_{ij}^k g_{kl} g^{lm} \\ &= (\partial_j g_{il} + \partial_i g_{jl} - \partial_l g_{ji}) g^{lm}.\end{aligned}$$

Thus we have the formula

$$\begin{aligned}\Gamma_{ij}^k &= \frac{1}{2} g^{lk} (\partial_j g_{il} + \partial_i g_{jl} - \partial_l g_{ji}) \\ &= \frac{1}{2} g^{kl} (\partial_j g_{il} + \partial_i g_{jl} - \partial_l g_{ji}) \\ &= \frac{1}{2} g^{kl} \Gamma_{ij,l}.\end{aligned}$$

The symbols

$$\begin{aligned}\Gamma_{ij,k} &= \frac{1}{2} (\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ji}) \\ &= g(\nabla_{\partial_i} \partial_j, \partial_k)\end{aligned}$$

are called the *Christoffel symbols of the first kind*, while Γ_{ij}^k are the *Christoffel symbols of the second kind*. Classically the following notation has also been used

$$\begin{aligned}\left\{ \begin{matrix} k \\ i, j \end{matrix} \right\} &= \Gamma_{ij}^k, \\ [ij, k] &= \Gamma_{ij,k}\end{aligned}$$

so as not to think that these things define a tensor. The reason why they are not tensorial comes from the fact that they may be zero in one coordinate system but not zero in another. A good example of this comes from the plane where the Christoffel symbols vanish in Cartesian coordinates, but not in polar coordinates:

$$\begin{aligned}\Gamma_{\theta\theta,r} &= \frac{1}{2} (\partial_\theta g_{\theta r} + \partial_\theta g_{\theta r} - \partial_r g_{\theta\theta}) \\ &= -\frac{1}{2} \partial_r (r^2) \\ &= -r.\end{aligned}$$

In fact, as is shown in exercise 2.5.20 it is always possible to find coordinates around a point $p \in M$ such that

$$\begin{aligned}g_{ij}|_p &= \delta_{ij}, \\ \partial_k g_{ij}|_p &= 0.\end{aligned}$$

In particular,

$$\begin{aligned}g_{ij}|_p &= \delta_{ij}, \\ \Gamma_{ij}^k|_p &= 0.\end{aligned}$$

In such coordinates the covariant derivative is computed exactly as in Euclidean space

$$\begin{aligned}\nabla_Y X|_p &= (\nabla_{Y^i \partial_i} X^j \partial_j)|_p \\ &= Y^i(p) (\partial_i X^j)|_p \partial_j|_p.\end{aligned}$$

The torsion free property of the connection is equivalent to saying that the Christoffel symbols are symmetric in ij as

$$\begin{aligned}\Gamma_{ij}^k \partial_k &= \nabla_{\partial_i} \partial_j \\ &= \nabla_{\partial_j} \partial_i \\ &= \Gamma_{ji}^k \partial_k.\end{aligned}$$

The metric property of the connection becomes

$$\begin{aligned}\partial_k g_{ij} &= g(\nabla_{\partial_k} \partial_i, \partial_j) + g(\partial_i, \nabla_{\partial_k} \partial_j) \\ &= \Gamma_{ki,j} + \Gamma_{kj,i}.\end{aligned}$$

This shows that the Christoffel symbols completely determine the derivatives of the metric.

Just as the metric could be used to give a formula for the gradient in local coordinates we can use the Christoffel symbols to get a local coordinate formula for the Hessian of a function. This is done as follows

$$\begin{aligned}2 \text{Hess} f(\partial_i, \partial_j) &= (L_{\nabla f} g)(\partial_i, \partial_j) \\ &= D_{\nabla f} g_{ij} - g(L_{\nabla f} \partial_i, \partial_j) - g(\partial_i, L_{\nabla f} \partial_j) \\ &= g^{kl} (\partial_k f) (\partial_l g_{ij}) \\ &\quad + g(L_{\partial_i} (g^{kl} (\partial_k f) \partial_l), \partial_j) \\ &\quad + g(\partial_i, L_{\partial_j} (g^{kl} (\partial_k f) \partial_l)) \\ &= (\partial_k f) g^{kl} (\partial_l g_{ij}) \\ &\quad + \partial_i (g^{kl} (\partial_k f)) g_{lj} \\ &\quad + \partial_j (g^{kl} (\partial_k f)) g_{il} \\ &= (\partial_k f) g^{kl} (\partial_l g_{ij}) \\ &\quad + (\partial_i \partial_k f) g^{kl} g_{lj} + (\partial_j \partial_k f) g^{kl} g_{il} \\ &\quad + (\partial_i g^{kl}) (\partial_k f) g_{lj} + (\partial_j g^{kl}) (\partial_k f) g_{il} \\ &= 2 \partial_i \partial_j f \\ &\quad + (\partial_k f) ((\partial_i g^{kl}) g_{lj} + (\partial_j g^{kl}) g_{il} + g^{kl} (\partial_l g_{ij})).\end{aligned}$$

To compute $\partial_i g^{jk}$ we note that

$$\begin{aligned} 0 &= \partial_i \delta_l^j \\ &= \partial_i (g^{jk} g_{kl}) \\ &= (\partial_i g^{jk}) g_{kl} + g^{jk} (\partial_i g_{kl}). \end{aligned}$$

Thus we have

$$\begin{aligned} 2 \text{Hess} f (\partial_i, \partial_j) &= 2 \partial_i \partial_j f \\ &\quad + (\partial_k f) ((\partial_i g^{kl}) g_{lj} + (\partial_j g^{kl}) g_{il} + g^{kl} (\partial_l g_{ij})) \\ &= 2 \partial_i \partial_j f \\ &\quad + (\partial_k f) (-g^{kl} \partial_i g_{lj} - g^{kl} \partial_j g_{li} + g^{kl} (\partial_l g_{ij})) \\ &= 2 \partial_i \partial_j f - g^{kl} (\partial_i g_{lj} + \partial_j g_{li} - \partial_l g_{ij}) \partial_k f \\ &= 2 (\partial_i \partial_j f - \Gamma_{ij}^k \partial_k f). \end{aligned}$$

Finally we mention yet another piece of notation that is often seen. Namely, if S is a $(1, k)$ -tensor written in a frame as:

$$S = S_{j_1 \dots j_k}^i \cdot E_i \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k},$$

then the covariant derivative is a $(1, k+1)$ -tensor that can be written as

$$\nabla S = S_{j_1 \dots j_k j_{k+1}}^i \cdot E_i \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k} \otimes \sigma^{j_{k+1}}.$$

The coefficient $S_{j_1 \dots j_k j_{k+1}}^i$ can be computed via the formula

$$\begin{aligned} \nabla_{E_{j_{k+1}}} S &= D_{E_{j_{k+1}}} (S_{j_1 \dots j_k}^i) \cdot E_i \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k} \\ &\quad + S_{j_1 \dots j_k}^i \cdot \nabla_{E_{j_{k+1}}} (E_i \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k}), \end{aligned}$$

where one must find the expression for

$$\begin{aligned} \nabla_{E_{j_{k+1}}} (E_i \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k}) &= (\nabla_{E_{j_{k+1}}} E_i) \otimes \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k} \\ &\quad + E_i \otimes (\nabla_{E_{j_{k+1}}} \sigma^{j_1}) \otimes \dots \otimes \sigma^{j_k} \\ &\quad \dots \\ &\quad + E_i \otimes \sigma^{j_1} \otimes \dots \otimes (\nabla_{E_{j_{k+1}}} \sigma^{j_k}) \end{aligned}$$

by writing each of the terms $(\nabla_{E_{j_k+1}} E_i), (\nabla_{E_{j_k+1}} \sigma^{j_1}), \dots, (\nabla_{E_{j_k+1}} \sigma^{j_k})$ in terms of the frame and coframe and substitute back into the formula.

This notation, however, is at odds with the idea that the covariant derivative variable should come first as the notation forces its index to be last. A better index notation, often used in physics, is to write

$$\nabla_{j_0} S = \nabla_{E_{j_0}} S$$

and let

$$\nabla_{j_0} S_{j_1 \dots j_k}^i = (\nabla S)_{j_0 \dots j_k}^i.$$

This notation is also explored in exercise 2.5.34. This will also be our convention when using indices for the curvature tensor.

2.5 Exercises

EXERCISE 2.5.1. Show that the connection on Euclidean space is the only affine connection such that $\nabla X = 0$ for all constant vector fields X .

EXERCISE 2.5.2. Show that the skew-symmetry property $[X, Y] = -[Y, X]$ does not necessarily hold for C^1 vector fields. Show that the Jacobi identity holds for C^2 vector fields.

EXERCISE 2.5.3. Let ∇ be an affine connection on a manifold. Show that the torsion tensor

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

defines a $(2, 1)$ -tensor.

EXERCISE 2.5.4. Show that if $c : I \rightarrow M$ has nonzero speed at $t_0 \in I$, then there is a vector X such that $X|_{c(t)} = \dot{c}(t)$ for t near t_0 .

EXERCISE 2.5.5. Let (M, g) be a Riemannian manifold, f, h functions on M , and X a vector field on M . Show that

$$\operatorname{div}(fX) = D_X f + f \operatorname{div} X,$$

$$\Delta(fh) = h\Delta f + f\Delta h + 2g(\nabla f, \nabla h),$$

$$\operatorname{Hess}(fh) = h\operatorname{Hess} f + f\operatorname{Hess} h + df dh + dh df.$$

EXERCISE 2.5.6. Let (M, g) be a Riemannian manifold, f a function on M , and ϕ a function on \mathbb{R} . Show that

$$\begin{aligned}\Delta(\phi(f)) &= \dot{\phi}(f) \Delta f + \ddot{\phi}(f) |df|^2, \\ \text{Hess}(\phi(f)) &= \dot{\phi}(f) \text{Hess} f + \ddot{\phi}(f) df^2.\end{aligned}$$

EXERCISE 2.5.7. Let (M, g) be a Riemannian manifold, X a vector field on M , and θ_X the dual 1-form. Show that $d\theta_X(Y, Z) = g(\nabla_Y X, Z) - g(Y, \nabla_Z X)$.

EXERCISE 2.5.8. The metric in coordinates satisfies:

- (1) $\partial_s g^{ij} = g^{ik} \partial_s g_{kl} g^{lj}$.
- (2) $\partial_s g^{ij} = -g^{il} \Gamma_{sl}^j - g^{jl} \Gamma_{sl}^i$.

EXERCISE 2.5.9. Let X be a vector field.

- (1) Show that for any $(1, 1)$ -tensor S

$$\text{tr}(\nabla_X S) = \nabla_X \text{tr} S.$$

- (2) Let $T(Y, Z) = g(S(Y), Z)$. Show that

$$(\nabla_X T)(Y, Z) = g((\nabla_X S)(Y), Z).$$

- (3) Show more generally that contraction and covariant differentiation commute.
- (4) Finally show that type change and covariant differentiation commute.

EXERCISE 2.5.10. Let X be a vector field.

- (1) Show that for any $(1, 1)$ -tensor S

$$\text{tr}(L_X S) = L_X \text{tr} S.$$

- (2) Let $T(Y, Z) = g(S(Y), Z)$. Show that

$$(L_X T)(Y, Z) = (L_X g)(S(Y), Z) + g((L_X S)(Y), Z).$$

- (3) Show that contraction and Lie differentiation commute.

EXERCISE 2.5.11. Show that a vector field X on a Riemannian manifold is locally a gradient field if and only if $Z \mapsto \nabla_Z X$ is self-adjoint.

EXERCISE 2.5.12. If $F: M \rightarrow M$ is a diffeomorphism, then the push-forward of a vector field is defined as

$$(F_* X)|_p = DF(X|_{F^{-1}(p)}).$$

Let F be an isometry on (M, g) .

- (1) Show that $F_*(\nabla_X Y) = \nabla_{F_*X} F_*Y$ for all vector fields.
- (2) Use this to show that isometries on $(\mathbb{R}^n, g_{\mathbb{R}^n})$ are of the form $F(x) = Ox + b$, where $O \in O(n)$ and $b \in \mathbb{R}^n$. Hint: Show that F maps constant vector fields to constant vector fields.

EXERCISE 2.5.13. A vector field X is said to be *affine* if $L_X \nabla = 0$.

- (1) Show that Killing fields are affine. Hint: The flow of X preserves the metric.
- (2) Give an example of an affine field on \mathbb{R}^n which is not a Killing field.

EXERCISE 2.5.14. Let G be a Lie group. Show that there is a unique affine connection such that $\nabla X = 0$ for all left-invariant vector fields. Show that this connection is torsion free if and only if the Lie algebra is Abelian.

EXERCISE 2.5.15. Show that the Hessian of a composition $\phi(f)$ is given by

$$\text{Hess } \phi(f) = \phi'' df^2 + \phi' \text{Hess } f.$$

EXERCISE 2.5.16. Consider a vector field X and a $(1, 1)$ -tensor L .

- (1) Show that $L_X + L$ defines a derivation on tensors.
- (2) Show that all derivations are of this form and that X is unique.
- (3) Show that derivations are uniquely determined by how they act on functions and vector fields.
- (4) Show that $L_{fX} = fL_X - X \otimes df$, where $X \otimes df$ is the rank 1 $(1, 1)$ -tensor $Y \mapsto Xdf(Y)$.

EXERCISE 2.5.17. Show that if X is a vector field of constant length on a Riemannian manifold, then $\nabla_v X$ is always perpendicular to X .

EXERCISE 2.5.18. Show that if we have a tensor field T on a Riemannian manifold (M, g) that vanishes at $p \in M$, then for any vector field X we have $L_X T = \nabla_X T$ at p . Conclude that the $(1, 1)$ version of the Hessian of a function is independent of the metric at a critical point. Can you find an interpretation of $L_X T$ at p ?

EXERCISE 2.5.19. For any $p \in (M, g)$ and orthonormal basis e_1, \dots, e_n for $T_p M$, show that there is an orthonormal frame E_1, \dots, E_n in a neighborhood of p such that $E_i = e_i$ and $(\nabla E_i)|_p = 0$. Hint: Fix an orthonormal frame \bar{E}_i near $p \in M$ with $\bar{E}_i(p) = e_i$. If we define $E_i = \alpha_i^j \bar{E}_j$, where $[\alpha_i^j(x)] \in SO(n)$ and $\alpha_i^j(p) = \delta_i^j$, then this will yield the desired frame provided that the directional derivatives $D_{e_k} \alpha_i^j$ are appropriately prescribed at p .

EXERCISE 2.5.20. Show that there are coordinates x^1, \dots, x^n such that $\partial_i = e_i$ and $\nabla \partial_i = 0$ at p . These conditions imply that the metric coefficients satisfy $g_{ij} = \delta_{ij}$ and $\partial_k g_{ij} = 0$ at p . Such coordinates are called normal coordinates at p . Hint: Given a general set of coordinates y^i around p with $y^i(p) = 0$, let $x^i = \alpha_i^j(y) y^j$, adjust

$\alpha_j^i(0)$ to make the fields orthonormal at p , and adjust $\frac{\partial \alpha_j^i}{\partial y^k}(0)$ to make the covariant derivatives vanish at p .

EXERCISE 2.5.21. Consider coordinates x^i and \bar{x}^s around $p \in M$. Show that the Christoffel symbols of a metric g in these two charts are related by

$$\bar{\Gamma}_{ij}^k = \frac{\partial^2 x^s}{\partial \bar{x}^i \partial \bar{x}^j} \frac{\partial \bar{x}^k}{\partial x^s} + \frac{\partial x^s}{\partial \bar{x}^i} \frac{\partial x^t}{\partial \bar{x}^j} \frac{\partial \bar{x}^k}{\partial x^t} \Gamma_{st}^l,$$

$$\frac{\partial^2 x^r}{\partial \bar{x}^i \partial \bar{x}^j} = \bar{\Gamma}_{ij}^k \frac{\partial x^r}{\partial \bar{x}^k} - \frac{\partial x^s}{\partial \bar{x}^i} \frac{\partial x^t}{\partial \bar{x}^j} \Gamma_{st}^r,$$

and

$$\bar{\Gamma}_{ij,k} = \frac{\partial^2 x^s}{\partial \bar{x}^i \partial \bar{x}^j} \frac{\partial x^t}{\partial \bar{x}^k} g_{st} + \frac{\partial x^s}{\partial \bar{x}^i} \frac{\partial x^t}{\partial \bar{x}^j} \frac{\partial x^l}{\partial \bar{x}^k} \Gamma_{st,l}.$$

EXERCISE 2.5.22. Let M be an n -dimensional submanifold of \mathbb{R}^{n+m} with the induced metric. Further assume that we have a local coordinate system given by a parametrization $u^s(x^1, \dots, x^n)$, $s = 1, \dots, n+m$. Show that in these coordinates:

(1)

$$g_{ij} = \sum_{s=1}^{n+m} \frac{\partial u^s}{\partial x^i} \frac{\partial u^s}{\partial x^j}.$$

(2)

$$\Gamma_{ij,k} = \sum_{s=1}^{n+m} \frac{\partial u^s}{\partial x^k} \frac{\partial^2 u^s}{\partial x^i \partial x^j}.$$

EXERCISE 2.5.23. Let (M, g) be an oriented manifold.

(1) Show that if v_1, \dots, v_n is positively oriented, then

$$\text{vol}(v_1, \dots, v_n) = \sqrt{\det(g(v_i, v_j))}.$$

(2) Show that in positively oriented coordinates,

$$\text{vol} = \sqrt{\det(g_{ij})} dx^1 \wedge \dots \wedge dx^n.$$

(3) Conclude that the Laplacian has the formula

$$\Delta u = \frac{1}{\sqrt{\det(g_{ij})}} \partial_k \left(\sqrt{\det(g_{ij})} g^{kl} \partial_l u \right).$$

Given that the coordinates are normal at p we get as in Euclidean space that

$$\Delta f(p) = \sum_{i=1}^n \partial_i^2 f.$$

EXERCISE 2.5.24. Show that if a $(0, 2)$ -tensor T is given by T_{kl} , then ∇T is given by

$$(\nabla T)_{jkl} = \frac{\partial T_{kl}}{\partial x^j} - \Gamma_{jk}^i T_{il} - \Gamma_{jl}^i T_{ki}.$$

Similarly, when a $(1, 1)$ -tensor T is given by T_l^k , then ∇T is given by

$$(\nabla T)_{jl}^k = \frac{\partial T_l^k}{\partial x^j} - \Gamma_{jl}^i T_i^k + \Gamma_{ji}^k T_l^i.$$

EXERCISE 2.5.25. Let $F : (M, g_M) \looparrowright (\bar{M}, g_{\bar{M}})$ be an isometric immersion. For two vector fields X, Y tangent to M we can compute both $\nabla_X^M Y$ and $\nabla_X^{\bar{M}} Y$. Show that the component of $\nabla_X^{\bar{M}} Y$ that is tangent to M is $\nabla_X^M Y$. Show that the normal component

$$\nabla_X^{\bar{M}} Y - \nabla_X^M Y = T_X Y$$

is symmetric in X, Y and use that to show that it is tensorial.

EXERCISE 2.5.26. Let $F : (M, g_M) \looparrowright (\bar{M}, g_{\bar{M}})$ be an isometric immersion and

$$T^\perp M = \{v \in T_p \bar{M} \mid p \in M \text{ and } v \perp T_p M\}$$

the normal bundle. A vector field $V : M \rightarrow T^\perp \bar{M}$ such that $V_p \in T_p^\perp M$ is called a normal field along M . For a vector field X and normal field V show that

- (1) The covariant derivative $\nabla_X^{\bar{M}} V$ can be defined.
- (2) Decompose $\nabla_X^{\bar{M}} V$ into normal $\nabla_X^\perp V$ and tangential $T_X V$ components:

$$\nabla_X^{\bar{M}} V = \nabla_X^\perp V + T_X V.$$

$\nabla_X^\perp V$ is called the normal derivative of V along M . Show that

$$g_{\bar{M}}(T_X Y, V) = -g_M(Y, T_X V).$$

- (3) Show that $\nabla_X^\perp V$ is linear and a derivation in the V variable and tensorial in the X variable.

EXERCISE 2.5.27. Let (M, g) be an oriented Riemannian manifold.

- (1) If f has compact support, then

$$\int_M \Delta f \cdot \text{vol} = 0.$$

- (2) Show that

$$\text{div}(f \cdot X) = g(\nabla f, X) + f \cdot \text{div} X.$$

- (3) Show that

$$\Delta(f_1 \cdot f_2) = (\Delta f_1) \cdot f_2 + 2g(\nabla f_1, \nabla f_2) + f_1 \cdot (\Delta f_2).$$

- (4) Establish Green's formula for functions with compact support:

$$\int_M f_1 \cdot \Delta f_2 \cdot d \text{vol} = - \int_M g(\nabla f_1, \nabla f_2) \text{vol}.$$

- (5) Conclude that if f is subharmonic or superharmonic (i.e., $\Delta f \geq 0$ or $\Delta f \leq 0$), then f is constant. (Hint: first show $\Delta f = 0$; then use integration by parts on $f \cdot \Delta f$.) This result is known as the *weak maximum principle*. More generally, one can show that any subharmonic (respectively superharmonic) function that has a global maximum (respectively minimum) must be constant. For this one does not need f to have compact support. This result is usually referred to as the *strong maximum principle*.

EXERCISE 2.5.28. A vector field and its corresponding flow is said to be *incompressible* if $\text{div} X = 0$.

- (1) Show that X is incompressible if and only if the local flows it generates are volume preserving (i.e., leave the Riemannian volume form invariant).
- (2) Let X be a unit vector field on \mathbb{R}^2 . Show that $\nabla X = 0$ if X is incompressible.
- (3) Find a unit vector field X on \mathbb{R}^3 that is incompressible but where $\nabla X \neq 0$.

EXERCISE 2.5.29. Let X be a unit vector field on (M, g) such that $\nabla_X X = 0$.

- (1) Show that X is locally the gradient of a function if and only if the orthogonal distribution is integrable.
- (2) Show that the orthogonal distribution is integrable in a neighborhood of $p \in M$ if it has an integral submanifold through p . Hint: It might help to show that $L_X \theta_X = 0$.
- (3) Find X with the given conditions so that it is not a gradient field. Hint: Consider S^3 .

EXERCISE 2.5.30. Suppose we have two distributions E and F on (M, g) , that are orthogonal complements of each other in TM . In addition, assume that the distributions are parallel i.e., if two vector fields X and Y are tangent to, say, E , then $\nabla_X Y$ is also tangent to E .

- (1) Show that the distributions are integrable.
- (2) Show that around any point $p \in M$ there is a product neighborhood $U = V_E \times V_F$ such that $(U, g) = (V_E \times V_F, g|_{V_E} + g|_{V_F})$, where V_E and V_F are the integral submanifolds through p .

EXERCISE 2.5.31. Let X be a parallel vector field on (M, g) . Show that X has constant length. Show that X generates parallel distributions, one that contains X and the other that is the orthogonal complement to X . Conclude that locally the metric is a product with an interval $(U, g) = (V \times I, g|_V + dt^2)$, where V is a submanifold perpendicular to X .

EXERCISE 2.5.32. If we have two tensors S, T of the same type show that

$$D_X g(S, T) = g(\nabla_X S, T) + g(S, \nabla_X T).$$

EXERCISE 2.5.33. Recall that complex manifolds have complex tangent spaces. Thus we can multiply vectors by i . As a generalization of this we can define an *almost complex* structure. This is a $(1, 1)$ -tensor J such that $J^2 = -I$. A *Hermitian structure* on a Riemannian manifold (M, g) is an almost complex structure J such that $g(J(X), J(Y)) = g(X, Y)$. The *Kähler form* of a Hermitian structure is $\omega(X, Y) = g(J(X), Y)$.

- (1) Show that the *Nijenhuis tensor*:

$$N(X, Y) = [J(X), J(Y)] - J([J(X), Y]) - J([X, J(Y)]) - [X, Y]$$

is a tensor.

- (2) Show that if J comes from a complex structure, then $N = 0$. The converse is the famous theorem of Newlander and Nirenberg.
- (3) Show that ω is a 2-form.
- (4) Show that $d\omega = 0$ if $\nabla J = 0$.
- (5) Conversely show that if $d\omega = 0$ and J is a complex structure, then $\nabla J = 0$. In this case we call the metric a Kähler metric.

EXERCISE 2.5.34. Define $\nabla_i T$ as the covariant derivative in the direction of the i^{th} coordinate vector field and $\nabla^i T = g^{ij} \nabla_j T$ as the corresponding type changed tensor.

- (1) For a function f show that $df = \nabla_i f dx^i$ and $\nabla f = \nabla^i f \partial_i$.
- (2) For a vector field X show that $(\nabla_i X)^i = \text{div } X$.
- (3) For a $(0, 2)$ -tensor T show that $(\nabla^i T)_{ij} = -(\nabla^* T)_j$.

Riemannian Geometry

Petersen, P.

2016, XVIII, 499 p. 50 illus., 1 illus. in color., Hardcover

ISBN: 978-3-319-26652-7