CHAPTER 2

Curvature

With the comforting feeling that there are indeed a variety of Riemannian manifolds out there, we shall now immerse ourselves in the theory. In this chapter we confine ourselves to infinitesimal considerations. The most important and often also least understood object of Riemannian geometry is the connection and its function as covariant differentiation. We shall give a motivation of this concept that depends on exterior and Lie derivatives (The basic definitions and properties of Lie derivatives are recaptured in the appendix). It is hoped that this makes the concept a little less of a *deus ex machina.* Covariant differentiation, in turn, gives us nice formulae for exterior derivatives, Lie derivatives, divergence and much more (see also the appendix). It is also important in the development of curvature which is the central theme of Riemannian geometry. The idea of a Riemannian metric having curvature, while intuitively appealing and natural, is for most people the stumbling block for further progress into the realm of geometry.

In the third section of the chapter we shall study what we call the fundamental equations of Riemannian geometry. These equations relate curvature to the Hessian of certain geometrically defined functions (Riemannian submersions onto intervals). These formulae hold all the information that we shall need when computing curvatures in new examples and also for studying Riemannian geometry in the abstract.

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 the Taylor expansion of a Riemannian metric at a point, where coordinates are chosen such that the zeroth-order term is the Euclidean metric and the first-order term is zero. Lipschitz, Killing, and Christoffel introduced the connection in various ways as an intermediate step in computing the curvature. Also, they found it was a natural invariant for what is called the equivalence problem in Riemannian geometry. This problem, which seems rather odd nowadays (although it really is important), comes out of the problem one faces when writing the same metric in two different coordinates. Namely, how is one to know that they are the same or equivalent. The idea is to find invariants of the metric that can be computed in coordinates and then try to show that two metrics are equivalent if their invariant expressions are equal. 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 with the hope of using it 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 semi-Riemannian setting. The connection and curvature tensor are generalized without changes. But the formulas for divergence and Ricci curvature do require some modifications. The thing to watch for is that the trace of an operator has a slightly different formula in this setting (see exercises to chapter 1).

1. Connections

1.1. Directional Differentiation. First we shall introduce some important notation. There are many ways of denoting the *directional derivative* of a function on a manifold. Given a function $f : M \to \mathbb{R}$ 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 $f : M \to \mathbb{R}$ on a manifold, then the differential $df : TM \to \mathbb{R}$ 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 $\mathrm{grad} f = \nabla f$, 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 the next section). 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 had used polar coordinates on $\mathbb{R}^2$, say, then we mostly have that

$$\begin{aligned} \nabla f &= \partial_x (f) \partial_x + \partial_y (f) \partial_y \\ &\neq \partial_r (f) \partial_r + \partial_\theta (f) \partial_\theta, \end{aligned}$$

One rule of thumb for items that are invariantly defined is that they should satisfy the Einstein summation convention, where one sums over identical super- and subscripts. 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.

1.2. Covariant Differentiation. We now come to the question of attaching a meaning to the change of a vector field. In $\mathbb{R}^n$ we can use the standard Cartesian coordinate vector fields to write $X = a^i\partial_i$. If we think of the coordinate vector fields as being constant, then it is natural to define the *covariant derivative* of X in the direction of Y as

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

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

$$\begin{aligned}\partial_r &= \frac{1}{r}\left(x\partial_x + y\partial_y\right)\\ \partial_\theta &= -y\partial_x + x\partial_y\end{aligned}$$

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

In order to better understand what is happening we need to find a coordinate independent definition of this change. 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., $\left(i_X g\right)(Y) = g\left(X, Y\right)$, 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 $d\theta_X$ therefore measures the extend 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 chapter 7). If this happens then we say that X is a *Killing field*. Lie derivatives will be used heavily below. The results we use are standard from manifold theory and are all explained in the appendix.

In case $X = \nabla f$ is a gradient field the expression $L_{\nabla f} g$ is essentially the Hessian of f. We can prove this in $\mathbb{R}^n$ were we already know what the Hessian should be. Let

$$\begin{aligned}X &= \nabla f = a^i\partial_i,\\ a^i &= \partial_i f,\end{aligned}$$

then

$$\begin{aligned}L_X\left(\delta_{ij}dx^i dx^j\right) &= \left(L_X\delta_{ij}\right) + \delta_{ij}L_X\left(dx^i\right)dx^j + \delta_{ij}dx^i L_X\left(dx^j\right)\\ &= 0 + \delta_{ij}\left(dL_X\left(x^i\right)\right)dx^j + \delta_{ij}dx^i\left(dL_X\left(x^j\right)\right)\\ &= \delta_{ij}\left(da^i\right)dx^j + \delta_{ij}dx^i da^j\\ &= \delta_{ij}\left(\partial_k a^i\right)dx^k dx^j + \delta_{ij}dx^i\left(\partial_k a^j\right)dx^k\\ &= \partial_k a^i dx^k dx^i + \partial_k a^i dx^i dx^k\\ &= \left(\partial_k a^i + \partial_i a^k\right)dx^i dx^k\\ &= \left(\partial_k\partial_i f + \partial_i\partial_k f\right)dx^i dx^k\\ &= 2\left(\partial_i\partial_k f\right)dx^i dx^k\\ &= 2\mathrm{Hess}f.\end{aligned}$$

From this calculation we can also quickly see what the Killing fields on $\mathbb{R}^n$ should be. If $X = a^i\partial_i$, then X is a Killing field iff $\partial_k a^i + \partial_i a^k = 0$. This shows that

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

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

$$a^i = \alpha^i_j x^j + \beta^i$$

with the extra conditions that

$$\alpha^i_j = \partial_j a^i = -\partial_i a^j = -\alpha^j_i.$$

The angular field ∂_θ is therefore 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

$$\begin{aligned}F^t(x) &= \exp(At)\,x + t\beta, \\ A &= \left[\alpha^i_j\right], \\ \beta &= \left[\beta^i\right].\end{aligned}$$

In this way we see that a vector field on $\mathbb{R}^n$ is constant iff it is a Killing field that is also a gradient field.

The important observation we can make on $\mathbb{R}^n$ is that

Proposition 3. *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

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

□

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

(1) $Y \to \nabla_Y X$ is a $(1,1)$-tensor:

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

(2) $X \to \nabla_Y X$ 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 $f : \mathbb{R}^n \to \mathbb{R}$.

(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

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

Thus

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

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

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

We then see that (3) follows from

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

And (4) from

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

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

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

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 which 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

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

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 $X\to\nabla_Y X$ is a derivation, it is definitely not tensorial in X. Therefore, we can not expect that $\left(\nabla_Y X\right)|_p$ depends only on $X|_p$ and $Y|_p$. The next two lemmas explore how $\left(\nabla_Y X\right)|_p$ depends on X.

Lemma 1. *Let M be a manifold and ∇ an affine connection on M. If $p\in M$, $v\in T_pM$, 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 $\lambda:M\to\mathbb{R}$ 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

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

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

$$\begin{aligned} \nabla_v X &= \nabla_v \lambda X \\ &= \nabla_v \lambda Y \\ &= \nabla_v Y \end{aligned}$$

□

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

PROOF. Choose a framing $\{Z_1, \ldots, Z_n\}$ in a neighborhood of p and write $Y = \sum \alpha^i \cdot Z_i$, $X = \sum \beta^i Z_i$ on this neighborhood. From the assumption that $X \circ \gamma = Y \circ \gamma$ we get that $\alpha^i \circ \gamma = \beta^i \circ \gamma$. Thus,

$$\begin{aligned} \nabla_v Y &= \nabla_v \alpha^i Z_i \\ &= \alpha^i(p) \nabla_v Z_i + Z_i(p) d\alpha^i(v) \\ &= \beta^i(p) \nabla_v Z_i + Z_i(p) d\beta^i(v) \\ &= \nabla_v X. \end{aligned}$$

□

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 $i = 1, \ldots, n$ and $v \in T_pM$. It is an easy exercise to show that such coordinates and frames always exist.

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

If S is a $(0, r)$- or $(1, r)$-tensor field, then we can define a *covariant derivative* ∇S that we interpret as a $(0, r+1)$- or $(1, r+1)$-tensor field. (Remember 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, define

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

Here ∇_X is interpreted as the directional derivative when applied to a function, while we use it as covariant differentiation on vector fields.

A tensor is said to be *parallel* if $\nabla S \equiv 0$. In $(\mathbb{R}^n, \mathrm{can})$ one can easily see that if a tensor is written in Cartesian coordinates, then it is parallel iff it has constant coefficients. Thus $\nabla X \equiv 0$ for constant vector fields. On a Riemannian manifold (M, g) we always have that $\nabla g \equiv 0$ since

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

from property (4) of the connection.

If $f : M \to \mathbb{R}$ is smooth, then we already have ∇f defined as the vector field satisfying

$$g(\nabla f, v) = D_v f = df(v).$$

There is some confusion here, with ∇f now also being defined as df. In any given context it will generally be clear what we mean. The *Hessian* $\mathrm{Hess}f$ is defined as the symmetric $(0,2)$-tensor $\frac{1}{2} L_{\nabla f} g$. We know that this conforms with our definition on $\mathbb{R}^n$. It can also be defined as a self-adjoint $(1,1)$-tensor by $S(X) = \nabla_X \nabla f$. These two tensors are naturally related by

$$\mathrm{Hess}f\,(X, Y) = g(S(X), Y).$$

To see this we observe that $d\left(\theta_{\nabla f}\right) = 0$ so

$$\begin{aligned}2g(S(X), Y) &= 2g\left(\nabla_X \nabla f, Y\right) \\ &= \left(L_{\nabla f} g\right)(Y, Z) + d\left(\theta_{\nabla f}\right)(Y, Z) \\ &= 2\mathrm{Hess}f\,(X, Y).\end{aligned}$$

The trace of S is the *Laplacian*, and we will use the notation $\Delta f = \mathrm{tr}(S)$. On $\mathbb{R}^n$ this is also written as $\Delta f = \mathrm{div}\nabla f$. The *divergence* of a vector field, $\mathrm{div}X$, on (M, g) is defined as

$$\mathrm{div}X = \mathrm{tr}(\nabla X).$$

In coordinates this is

$$\mathrm{tr}(\nabla X) = dx^i\left(\nabla_{\partial_i} X\right),$$

and with respect to an orthonormal basis

$$\mathrm{tr}(\nabla X) = \sum_{i=1}^{n} g\left(\nabla_{e_i} X, e_i\right).$$

Thus, also

$$\Delta f = \mathrm{tr}(\nabla(\nabla f)) = \mathrm{div}(\nabla f).$$

In analogy with our definition of $\mathrm{div}X$ we can also define the divergence of a $(1, r)$-tensor S to be the $(0, r)$-tensor

$$(\mathrm{div}S)(v_1, \ldots, v_r) = \mathrm{tr}\left(w \to (\nabla_w S)(v_1, \ldots, v_r)\right).$$

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

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

With this we get the $(0,2)$ version of the Hessian of a function defined as

$$\begin{aligned}\nabla^2_{X,Y} f &= \nabla_X \nabla_Y f - \nabla_{\nabla_X Y} f \\ &= \nabla_X g(Y, \nabla f) - g(\nabla_X Y, \nabla f) \\ &= g(Y, \nabla_X \nabla f) \\ &= g(S(X), Y).\end{aligned}$$

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

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

2. The Connection in Local Coordinates

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

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

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

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

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

$$\delta^i_j = g^{ik} g_{kj}.$$

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

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

In other words we have

$$\theta_X = g_{ij} a^i dx^j = \alpha_j dx^j = \omega.$$

This shows that

$$g_{ij} a^i = \alpha_j.$$

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

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

Therefore

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

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

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

We now go on to find a formula for $\nabla_Y X$ in local coordinates

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

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

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

and on the right hand side

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

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

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

Thus we have the formula

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

The symbols

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

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

$$\begin{aligned}\begin{Bmatrix} k \\ ij \end{Bmatrix} &= \Gamma_{ij}^{k},\\ [ij,k] &= \Gamma_{ij,k}\end{aligned}$$

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 are zero in Cartesian coordinates, but not in polar coordinates:

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

In fact, 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,

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

The covariant derivative is then computed exactly as in Euclidean space

$$\begin{aligned}\nabla_Y X|_p &= \left(\nabla_{b^i\partial_i}a^j\partial_j\right)|_p\\ &= b^i(p)\,\partial_i\left(a^j\right)|_p\partial_j|_p.\end{aligned}$$

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

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

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

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

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

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

Thus we have

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

3. Curvature

Having now developed the idea of covariant derivatives and explained their relation to the classical concepts of gradient, Hessian, and Laplacian, one might hope that somehow these concepts carry over to tensors. As we have seen, this is true with one important exception, namely, the most important tensor for us, the Riemannian metric g. This tensor is parallel and therefore has no gradient, etc.

Instead, we think of the connection itself as a sort of gradient of the metric. The next question then is, what should the Laplacian and Hessian be? The answer is, curvature.

Any connection on a manifold gives rise to a *curvature tensor*. This operator measures in some sense how far away the connection is from being our standard connection on $\mathbb{R}^n$, which we assume is our canonical curvature-free, or flat, space. If we are on a Riemannian manifold, then it is possible to take traces of this curvature operator to obtain various averaged curvatures.

3.1. The Curvature Tensor. We shall work exclusively in the Riemannian setting. So let (M, g) be a Riemannian manifold and ∇ the Riemannian connection. The curvature tensor is a $(1,3)$-tensor defined by

$$\begin{aligned} R(X,Y)Z &= \nabla^2_{X,Y}Z - \nabla^2_{Y,X}Z \\ &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X - \nabla_{[X,Y]}Z \\ &= [\nabla_X, \nabla_Y]\, Z - \nabla_{[X,Y]}Z. \end{aligned}$$

on vector fields X, Y, Z. Of course, it needs to be proved that this is indeed a tensor. Since both of the second covariant derivatives are tensorial in X and Y, we need only check that R is tensorial in Z. This is easily done:

$$\begin{aligned} R(X,Y)\,fZ &= \nabla^2_{X,Y}(fZ) - \nabla^2_{Y,X}(fZ) \\ &= f\nabla^2_{X,Y}(Z) - f\nabla^2_{Y,X}(Z) \\ &\quad + \left(\nabla^2_{X,Y}f\right)Z - \left(\nabla^2_{Y,X}f\right)Z \\ &\quad + (\nabla_Y f)\,\nabla_X Z + (\nabla_X f)\,\nabla_Y Z \\ &\quad - (\nabla_X f)\,\nabla_Y Z - (\nabla_Y f)\,\nabla_X Z \\ &= f\left(\nabla^2_{X,Y}(Z) - \nabla^2_{Y,X}(Z)\right) \\ &= fR(X,Y)\,Z. \end{aligned}$$

Notice that X, Y appear skew-symmetrically in $R(X,Y)Z$, while Z plays its own role on top of the line, hence the unusual notation. One could also write $R_{X,Y}Z$. Using the metric g we can change this to a $(0,4)$-tensor as follows:

$$R(X,Y,Z,W) = g(R(X,Y)Z, W).$$

The variables are now treated on a more equal footing, which is also justified by the next proposition.

Proposition 4. *The Riemannian curvature tensor $R(X,Y,Z,W)$ satisfies the following properties:*

(1) R is skew-symmetric in the first two and last two entries:

$$R(X,Y,Z,W) = -R(Y,X,Z,W) = R(Y,X,W,Z).$$

(2) R is symmetric between the first two and last two entries:

$$R(X,Y,Z,W) = R(Z,W,X,Y).$$

(3) R satisfies a cyclic permutation property called Bianchi's first identity*:*

$$R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0.$$

(4) ∇R satisfies a cyclic permutation property called Bianchi's second identity*:*

$$(\nabla_Z R)(X,Y)\,W + (\nabla_X R)(Y,Z)\,W + (\nabla_Y R)(Z,X)\,W = 0.$$

PROOF. The first part of (1) has already been established. For part two of (1) observe that $[X,Y]$ is the unique vector field defined by

$$D_X D_Y f - D_Y D_X f - D_{[X,Y]} f = 0.$$

In other words, $R(X,Y)f = 0$. This is the idea behind the calculations that follow:

$$\begin{aligned} g(R(X,Y)Z,Z) &= g(\nabla_X\nabla_Y Z,Z) - g(\nabla_Y\nabla_X Z,Z) - g(\nabla_{[X,Y]}Z,Z) \\ &= D_X g(\nabla_Y Z,Z) - g(\nabla_Y Z,\nabla_X Z) \\ &\quad -D_Y g(\nabla_X Z,Z) + g(\nabla_X Z,\nabla_Y Z) - \frac{1}{2}D_{[X,Y]}g(Z,Z) \\ &= \frac{1}{2}D_X D_Y g(Z,Z) - \frac{1}{2}D_Y D_X g(Z,Z) - \frac{1}{2}D_{[X,Y]}g(Z,Z) \\ &= 0. \end{aligned}$$

Now (1) follows by polarizing the identity $R(X,Y,Z,Z) = 0$ in Z.

Part (3) is proved using the torsion free property of the connection. We introduce some special notation. Let T be any mapping with 3 vector field variables and values that can be added. Summing over cyclic permutations of the variables gives us a new map

$$\mathfrak{S}T(X,Y,Z) = T(X,Y,Z) + T(Z,X,Y) + T(Y,Z,X)$$

that is invariant under cyclic permutations. Note that T doesn't have to be a tensor. As an example we can use $T(X,Y,Z) = [X,[Y,Z]]$ and observe that the Jacobi identity for vector fields says:

$$\mathfrak{S}[X,[Y,Z]] = 0.$$

For the proof of (3) we have

$$\begin{aligned} \mathfrak{S}R(X,Y)Z &= \mathfrak{S}\nabla_X\nabla_Y Z - \mathfrak{S}\nabla_Y\nabla_X Z - \mathfrak{S}\nabla_{[X,Y]}Z \\ &= \mathfrak{S}\nabla_Z\nabla_X Y - \mathfrak{S}\nabla_Z\nabla_Y X - \mathfrak{S}\nabla_{[X,Y]}Z \\ &= \mathfrak{S}\nabla_Z(\nabla_X Y - \nabla_Y X) - \mathfrak{S}\nabla_{[X,Y]}Z \\ &= \mathfrak{S}[X,[Y,Z]] \\ &= 0. \end{aligned}$$

Part (2) is a combinatorial consequence of (1) and (3):

$$\begin{aligned} R(X,Y,Z,W) &= -R(Z,X,Y,W) - R(Y,Z,X,W) \\ &= R(Z,X,W,Y) + R(Y,Z,W,X) \\ &= -R(W,Z,X,Y) - R(X,W,Z,Y) \\ &\quad -R(W,Y,Z,X) - R(Z,W,Y,X) \\ &= 2R(Z,W,X,Y) + R(X,W,Y,Z) + R(W,Y,X,Z) \\ &= 2R(Z,W,X,Y) - R(Y,X,W,Z) \\ &= 2R(Z,W,X,Y) - R(X,Y,Z,W), \end{aligned}$$

which implies $2R(X,Y,Z,W) = 2R(Z,W,X,Y)$.

Now for part (4). We use again the cyclic sum notation and in addition that

$$R(X,Y)Z = [\nabla_X,\nabla_Y]Z - \nabla_{[X,Y]}Z,$$

$$
\begin{aligned}
(\nabla_Z R)(X,Y)W &= \nabla_Z(R(X,Y)W) - R(\nabla_Z X,Y)W \\
&\quad - R(X,\nabla_Z Y)W - R(X,Y)\nabla_Z W \\
&= [\nabla_Z, R(X,Y)]W - R(\nabla_Z X,Y)W - R(X,\nabla_Z Y)W.
\end{aligned}
$$

Keeping in mind that we only do cyclic sums over X, Y, Z and that we have the Jacobi identity for operators:

$$\mathfrak{S}[\nabla_X,[\nabla_Y,\nabla_Z]] = 0$$

we obtain

$$
\begin{aligned}
\mathfrak{S}(\nabla_X R)(Y,Z)W &= \mathfrak{S}[\nabla_X, R(Y,Z)]W - \mathfrak{S}R(\nabla_X Y,Z)W - \mathfrak{S}R(Y,\nabla_X Z)W \\
&= \mathfrak{S}[\nabla_X,[\nabla_Y,\nabla_Z]]W - \mathfrak{S}\left[\nabla_X,\nabla_{[Y,Z]}\right]W \\
&\quad - \mathfrak{S}R(\nabla_X Y,Z)W - \mathfrak{S}R(Y,\nabla_X Z)W \\
&= -\mathfrak{S}\left[\nabla_X,\nabla_{[Y,Z]}\right]W - \mathfrak{S}R(\nabla_X Y,Z)W + \mathfrak{S}R(\nabla_Y X,Z)W \\
&= -\mathfrak{S}\left[\nabla_X,\nabla_{[Y,Z]}\right]W - \mathfrak{S}R([X,Y],Z)W \\
&= -\mathfrak{S}\left[\nabla_X,\nabla_{[Y,Z]}\right]W - \mathfrak{S}\left[\nabla_{[X,Y]},\nabla_Z\right]W + \mathfrak{S}\nabla_{[[X,Y],Z]}W \\
&= \mathfrak{S}\left[\nabla_{[X,Y]},\nabla_Z\right]W - \mathfrak{S}\left[\nabla_{[X,Y]},\nabla_Z\right]W \\
&= 0.
\end{aligned}
$$

□

Notice that part (1) is related to the fact that ∇ is metric, i.e.,

$$d(g(X,Y)) = g(\nabla X,Y) + g(X,\nabla Y),$$

while part (3) follows from ∇ being torsion free, i.e.,

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

Example 20. $(\mathbb{R}^n, \mathrm{can})$ *has* $R \equiv 0$ *since* $\nabla_{\partial_i}\partial_j = 0$ *for the standard Cartesian coordinates.*

More generally for any tensor field S of type $(\cdot, r)$ we can define the curvature as the new $(\cdot, r)$ tensor field

$$R(X,Y)S = \nabla^2_{X,Y}S - \nabla^2_{Y,X}S.$$

Again one needs to check that this is indeed a tensor. This is done in the same way we checked that $R(X,Y)Z$ was tensorial in Z. Clearly, $R(X,Y)S$ is also tensorial and skew symmetric in X and Y.

From the curvature tensor R we can derive several different curvature concepts.

3.2. The Curvature Operator. First recall that we have the space $\Lambda^2 M$ of bivectors. If e_i is an orthonormal basis for T_pM, then the inner product on $\Lambda^2_p M$ is such that the bivectors $e_i \wedge e_j$, $i < j$ will form an orthonormal basis. The inner product that $\Lambda^2 M$ inherits in this way is also denoted by g. Alternatively, we can define the inner product g on $\Lambda^2_p M$ using

$$
\begin{aligned}
g(x\wedge y, v\wedge w) &= g(x,v)\,g(y,w) - g(x,w)\,g(y,v) \\
&= \det\begin{pmatrix} g(x,v) & g(x,w) \\ g(y,v) & g(y,w) \end{pmatrix}
\end{aligned}
$$

and then extend it by linearity to all of $\Lambda_p^2 M$. It is also useful to interpret bivectors as skew symmetric maps. This is done by the formula:

$$(v \wedge w)(x) = g(w, x)\, v - g(v, x)\, w.$$

With this definition we have a Bianchi or Jacobi type identity:

$$(x \wedge y)(z) + (y \wedge z)(x) + (z \wedge x)(y) = 0.$$

From the symmetry properties of the curvature tensor we see that R actually defines a symmetric bilinear map

$$\begin{aligned} R &: \Lambda^2 M \times \Lambda^2 M \to \mathbb{R} \\ R(X \wedge Y, V \wedge W) &= R(X, Y, W, V). \end{aligned}$$

Note the reversal of V and W! The relation

$$g(\mathfrak{R}(X \wedge Y), V \wedge W) = R(X \wedge Y, V \wedge W)$$

therefore defines a self-adjoint operator $\mathfrak{R} : \Lambda^2 M \to \Lambda^2 M$. This operator is called the *curvature operator*. It is clearly just a different manifestation of the curvature tensor. The switch between Z and W is related to our definition of the next curvature concept.

3.3. Sectional Curvature. For any $v \in T_p M$ let

$$R_v(w) = R(w, v)v : T_p M \to T_p M$$

be the *directional curvature operator* . This operator is also known as the *tidal force operator*. The latter name accurately describes in physical terms the meaning of the tensor. The above conditions imply that this operator is self-adjoint and that v is always a zero eigenvector. The normalized quadratic form

$$\begin{aligned} \sec(v, w) &= \frac{g(R_v(w), w)}{g(v,v)g(w,w) - g(v,w)^2} \\ &= \frac{g(R(w,v)v, w)}{g(v \wedge w, v \wedge w)} \\ &= \frac{g(\mathfrak{R}(v \wedge w), v \wedge w)}{(\text{area}\square(v,w))^2} \end{aligned}$$

is called the *sectional curvature* of (v, w). Since the denominator is the square of the area of the parallelogram $\{tv + sw : 0 \leq t,\ s \leq 1\}$, we can easily check that $\sec(v, w)$ depends only on the plane $\pi = \text{span}\{v, w\}$. One of the important relationships between directional and sectional curvature is the following observation by Riemann.

Proposition 5. (Riemann, 1854) *The following properties are equivalent:*
(1) $\sec(\pi) = k$ *for all 2-planes in* $T_p M$.
(2) $R(v_1, v_2)v_3 = k(v_1 \wedge v_2)(v_3)$ *for all* $v_1, v_2, v_3 \in T_p M$.
(3) $R_v(w) = k \cdot (w - g(w,v)v) = k \cdot pr_{v^\perp}(w)$ *for all* $w \in T_p M$ *and* $|v| = 1$.
(4) $\mathfrak{R}(\omega) = k \cdot \omega$ *for all* $\omega \in \Lambda_p^2 M$.

Proof. (2) $\Rightarrow$ (3) $\Rightarrow$ (1) are easy. For (1) $\Rightarrow$ (2) we introduce the multilinear maps:

$$\begin{aligned} R_k(v_1, v_2)v_3 &= k(v_1 \wedge v_2)(v_3), \\ R_k(v_1, v_2, v_3, v_4) &= kg((v_1 \wedge v_2)(v_3), v_4). \end{aligned}$$

The first observation is that these maps behave exactly like the curvature tensor in that they satisfy properties 1, 2, and 3 of the above proposition. Now consider the map

$$T(v_1, v_2, v_3, v_4) = R(v_1, v_2, v_3, v_4) - R_k(v_1, v_2, v_3, v_4)$$

which also satisfies the same symmetry properties. Moreover, the assumption that $\sec = k$ implies

$$T(v, w, w, v) = 0$$

for all $v, w \in T_pM$. Using polarization $w = w_1 + w_2$ we get

$$\begin{aligned} 0 &= T(v, w_1 + w_2, w_1 + w_2, v) \\ &= T(v, w_1, w_2, v) + T(v, w_2, w_1, v) \\ &= 2T(v, w_1, w_2, v) \\ &= -2T(v, w_1, v, w_2). \end{aligned}$$

Using properties 1 and 2 of the curvature tensor we now see that T is alternating in all four variables. That, however, is in violation of Bianchi's first identity unless $T = 0$, which is exactly what we wish to prove.

To see why (2) $\Rightarrow$ (4), choose an orthonormal basis e_i for T_pM; then $e_i \wedge e_j$, $i < j$, is a basis for $\Lambda^2_p M$. Using (2) we see that

$$\begin{aligned} g(\mathfrak{R}(e_i \wedge e_j), e_t \wedge e_s) &= R(e_i, e_j, e_s, e_t) \\ &= k \cdot (g(e_j, e_s)g(e_i, e_t) - g(e_i, e_s)g(e_j, e_t)) \\ &= k \cdot g(e_i \wedge e_j, e_t \wedge e_s). \end{aligned}$$

But this implies that

$$\mathfrak{R}(e_i \wedge e_j) = k \cdot (e_i \wedge e_j).$$

For (4) $\Rightarrow$ (1) just observe that if $\{v, w\}$ are orthogonal unit vectors, then

$$k = g(\mathfrak{R}(v \wedge w), v \wedge w) = \sec(v, w).$$

□

A Riemannian manifold (M, g) that satisfies either of these four conditions for all $p \in M$ and the same $k \in \mathbb{R}$ for all $p \in M$ is said to have *constant curvature k*. So far we only know that $(\mathbb{R}^n, \mathrm{can})$ has curvature zero. In chapter 3 we shall prove that $dr^2 + \mathrm{sn}_k^2(r)ds_{n-1}^2$ has constant curvature k. When $k > 0$, recall that these represent $\left(S^n\left(\frac{1}{\sqrt{k}}\right), \mathrm{can}\right)$, while when $k < 0$ we still don't have a good picture yet. A whole section in chapter 3 is devoted to these constant negative curvature metrics.

3.4. Ricci Curvature. Our next curvature is the Ricci curvature, which should be thought of as the Laplacian of g.

The *Ricci curvature* Ric is a trace of R. If $e_1, \ldots, e_n \in T_pM$ is an orthonormal basis, then

$$\begin{aligned} \mathrm{Ric}(v,w) &= \mathrm{tr}\,(x \to R\,(x,v)\,w) \\ &= \sum_{i=1}^{n} g\,(R\,(e_i,v)\,w, e_i) \\ &= \sum_{i=1}^{n} g\,(R\,(v,e_i)\,e_i, w) \\ &= \sum_{i=1}^{n} g\,(R\,(e_i,w)\,v, e_i)\,. \end{aligned}$$

Thus Ric is a symmetric bilinear form. It could also be defined as the symmetric $(1,1)$-tensor

$$\mathrm{Ric}(v) = \sum_{i=1}^{n} R\,(v,e_i)\,e_i.$$

We adopt the language that $\mathrm{Ric} \geq k$ if all eigenvalues of $\mathrm{Ric}(v)$ are $\geq k$. In $(0,2)$ language this means more precisely that $\mathrm{Ric}\,(v,v) \geq kg\,(v,v)$ for all v. If (M,g) satisfies $\mathrm{Ric}(v) = k \cdot v$, or equivalently $\mathrm{Ric}(v,w) = k \cdot g(v,w)$, then (M,g) is said to be an *Einstein manifold* with *Einstein constant* k. If (M,g) has constant curvature k, then (M,g) is also Einstein with Einstein constant $(n-1)k$.

In chapter 3 we shall exhibit several interesting Einstein metrics that do not have constant curvature. Three basic types are

(1) $(S^n(1) \times S^n(1),\ ds_n^2 + ds_n^2)$ with Einstein constant $n-1$.

(2) The Fubini-Study metric on $\mathbb{C}P^n$ with Einstein constant $2n+2$.

(3) The Schwarzschild metric on $\mathbb{R}^2 \times S^2$, which is a doubly warped product metric: $dr^2 + \varphi^2(r)d\theta^2 + \psi^2(r)ds_2^2$ with Einstein constant 0.

If $v \in T_pM$ is a unit vector and we complete it to an orthonormal basis $\{v, e_2, \ldots, e_n\}$ for T_pM, then

$$\mathrm{Ric}\,(v,v) = g\,(R\,(v,v)\,v,v) + \sum_{i=2}^{n} g\,(R\,(e_i,v)\,v,e_i) = \sum_{i=2}^{n} \sec\,(v,e_i)\,.$$

Thus, when $n=2$, there is no difference from an informational point of view in knowing R or Ric. This is actually also true in dimension $n=3$, because if $\{e_1,e_2,e_3\}$ is an orthonormal basis for T_pM, then

$$\begin{aligned} \sec\,(e_1,e_2) + \sec\,(e_1,e_3) &= \mathrm{Ric}\,(e_1,e_1)\,, \\ \sec\,(e_1,e_2) + \sec\,(e_2,e_3) &= \mathrm{Ric}\,(e_2,e_2)\,, \\ \sec\,(e_1,e_3) + \sec\,(e_2,e_3) &= \mathrm{Ric}\,(e_3,e_3)\,. \end{aligned}$$

In other words:

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sec\,(e_1,e_2) \\ \sec\,(e_2,e_3) \\ \sec\,(e_1,e_3) \end{bmatrix} = \begin{bmatrix} \mathrm{Ric}\,(e_1,e_1) \\ \mathrm{Ric}\,(e_2,e_2) \\ \mathrm{Ric}\,(e_3,e_3) \end{bmatrix}.$$

Here, the matrix has det $= 2$, therefore any sectional curvature can be computed from Ric. In particular, we see that $\left(M^3,g\right)$ is Einstein iff $\left(M^3,g\right)$ has constant sectional curvature. The search for Einstein metrics should therefore begin in dimension 4.

3.5. Scalar Curvature. The last curvature quantity we wish to mention is the *scalar curvature*:

$$\text{scal} = \text{tr}\,(\text{Ric}) = 2 \cdot \text{tr}\mathfrak{R}.$$

Notice that scal depends only on $p \in M$ and is therefore a function, scal : $M \to \mathbb{R}$. In an orthonormal basis $e_1, \ldots, e_n$ for T_pM we have

$$\begin{aligned}
\text{scal} &= \text{tr}\,(\text{Ric}) \\
&= \sum_{j=1}^{n} g\left(\text{Ric}\,(e_j)\,, e_j\right) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} g\left(R\,(e_i, e_j)\, e_j, e_i\right) \\
&= \sum_{i,j=1}^{n} g\left(\mathfrak{R}\,(e_i \wedge e_j)\,, e_i \wedge e_j\right) \\
&= 2\sum_{i<j} g\left(\mathfrak{R}\,(e_i \wedge e_j)\,, e_i \wedge e_j\right) \\
&= 2\text{tr}\mathfrak{R} \\
&= 2\sum_{i<j} \sec\,(e_i, e_j)\,.
\end{aligned}$$

When $n = 2$ we see that $\text{scal}(p) = 2 \cdot \sec(T_pM)$. In chapter 3 we shall show that when $n = 3$ there are metrics with constant scalar curvature that are not Einstein. When $n \geq 3$ there is also another interesting phenomenon occurring related to scalar curvature.

LEMMA 3. (Schur, 1886) *Suppose that a Riemannian manifold* (M, g) *of dimension* $n \geq 3$ *satisfies one of the following two conditions:*

a) $\sec(\pi) = f(p)$ *for all 2-planes* $\pi \subset T_pM,\ p \in M$.

b) $\text{Ric}(v) = (n-1) \cdot f(p) \cdot v$ *for all* $v \in T_pM,\ p \in M$.

Then in either case f *must be constant. In other words, the metric has constant curvature or is Einstein, respectively.*

PROOF. It clearly suffices to show (b), as the conditions for (a) imply that (b) holds. To show (b) we need the important identity:

$$d\text{scal} = 2\text{div}\,(\text{Ric})\,.$$

Let us see how this implies (b). First we have

$$\begin{aligned}
d\text{scal} &= d\text{tr}\,(\text{Ric}) \\
&= d\,(n \cdot (n-1) \cdot f) \\
&= n \cdot (n-1) \cdot df.
\end{aligned}$$

On the other hand

$$\begin{aligned}
2\mathrm{div}\,(\mathrm{Ric})\,(v) &= 2\sum g\left(\left(\nabla_{e_i}\mathrm{Ric}\right)(v), e_i\right) \\
&= 2\sum g\left(\left(\nabla_{e_i}\left((n-1)\,f\cdot I\right)\right)(v), e_i\right) \\
&= 2\sum g\left((n-1)\left(\nabla_{e_i}f\right)v, e_i\right) + 2\sum g\left((n-1)\,f\left(\nabla_{e_i}I\right)(v), e_i\right) \\
&= 2\,(n-1)\,g\left(v, \sum\left(\nabla_{e_i}f\right)e_i\right) \\
&= 2\,(n-1)\,g\left(v, \nabla f\right) \\
&= 2\,(n-1)\,df\,(v)\,.
\end{aligned}$$

Thus, we have shown that $n\cdot df = 2\cdot df$, but this is impossible unless $n = 2$ or $df \equiv 0$ (i.e., f is constant). □

PROPOSITION 6.

$$d\mathrm{tr}\,(\mathrm{Ric}) = 2\mathrm{div}\,(\mathrm{Ric})\,.$$

PROOF. The identity is proved by a long and uninspired calculation that uses the second Bianchi identity. Choose a normal orthonormal frame E_i at $p\in M$, i.e., $\nabla E_i|_p = 0$, and let W be a vector field such that $\nabla W|_p = 0$. Using the second Bianchi identity

$$\begin{aligned}
\left(d\mathrm{tr}\,(\mathrm{Ric})\right)(W)\,(p) &= D_W\sum g\left(\mathrm{Ric}\,(E_i), E_i\right) \\
&= D_W\sum g\left(R\,(E_i,E_j)\,E_j, E_i\right) \\
&= \sum g\left(\nabla_W\left(R\,(E_i,E_j)\,E_j\right), E_i\right) \\
&= \sum g\left(\left(\nabla_W R\right)(E_i,E_j)\,E_j, E_i\right) \\
&= -\sum g\left(\left(\nabla_{E_j}R\right)(W,E_i)\,E_j, E_i\right) \\
&\quad -\sum g\left(\left(\nabla_{E_i}R\right)(E_j,W)\,E_j, E_i\right) \\
&= -\sum\left(\nabla_{E_j}R\right)(W,E_i,E_j,E_i) - \sum\left(\nabla_{E_i}R\right)(E_j,W,E_j,E_i) \\
&= \sum\left(\nabla_{E_j}R\right)(E_j,E_i,E_i,W) + \sum\left(\nabla_{E_i}R\right)(E_i,E_j,E_j,W) \\
&= 2\sum\left(\nabla_{E_j}R\right)(E_j,E_i,E_i,W) \\
&= 2\sum\nabla_{E_j}\left(R\,(E_j,E_i,E_i,W)\right) \\
&= 2\sum\nabla_{E_j}g\left(\mathrm{Ric}\,(E_j), W\right) \\
&= 2\sum\nabla_{E_j}g\left(\mathrm{Ric}\,(W), E_j\right) \\
&= 2\sum g\left(\nabla_{E_j}\left(\mathrm{Ric}\,(W)\right), E_j\right) \\
&= 2\sum g\left(\left(\nabla_{E_j}\mathrm{Ric}\right)(W), E_j\right) \\
&= 2\mathrm{div}\,(\mathrm{Ric})\,(W)\,(p)\,.
\end{aligned}$$

□

COROLLARY 1. *An $n\,(>2)$-dimensional Riemannian manifold (M,g) is Einstein iff*

$$\mathrm{Ric} = \frac{\mathrm{scal}}{n} g.$$

3.6. Curvature in Local Coordinates. As with the connection it is sometimes convenient to know what the curvature tensor looks like in local coordinates. We first observe that if $X = \alpha^i \partial_i$, $Y = \beta^j \partial_j$, $Z = \gamma^k \partial_k$, then we can write

$$\begin{aligned} R(X,Y)Z &= \alpha^i \beta^j \gamma^k R^l_{ijk} \partial_l, \\ R^l_{ijk}\partial_l &= R(\partial_i, \partial_j)\partial_k. \end{aligned}$$

Using the definition of R we see that

$$\begin{aligned} R^l_{ijk}\partial_l &= R(\partial_i,\partial_j)\partial_k \\ &= \nabla_{\partial_i}\nabla_{\partial_j}\partial_k - \nabla_{\partial_j}\nabla_{\partial_i}\partial_k \\ &= \nabla_{\partial_i}\left(\Gamma^s_{jk}\partial_s\right) - \nabla_{\partial_j}\left(\Gamma^t_{ik}\partial_t\right) \\ &= \partial_i\left(\Gamma^s_{jk}\right)\partial_s + \Gamma^s_{jk}\nabla_{\partial_i}\partial_s \\ &\quad -\partial_j\left(\Gamma^t_{ik}\right)\partial_t - \Gamma^t_{ik}\nabla_{\partial_j}\partial_t \\ &= \partial_i\left(\Gamma^l_{jk}\right)\partial_l - \partial_j\left(\Gamma^l_{ik}\right)\partial_l \\ &\quad +\Gamma^s_{jk}\Gamma^l_{is}\partial_l - \Gamma^t_{ik}\Gamma^l_{jt}\partial_l \\ &= \left(\partial_i\Gamma^l_{jk} - \partial_j\Gamma^l_{ik} + \Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js}\right)\partial_l. \end{aligned}$$

So

$$R^l_{ijk} = \partial_i\Gamma^l_{jk} - \partial_j\Gamma^l_{ik} + \Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js}.$$

This coordinate expression can also be used, in conjunction with the properties of the Christoffel symbols, to prove all of the symmetry properties of the curvature tensor. The formula clearly simplifies if we are at a point p where $\Gamma^k_{ij}|_p = 0$

$$R^l_{ijk}|_p = \partial_i\Gamma^l_{jk}|_p - \partial_j\Gamma^l_{ik}|_p.$$

If we use the formulas for the Christoffel symbols we can evidently get an expression for R^l_{ijk} that depends on the metric g_{ij} and its first two derivatives.

4. The Fundamental Curvature Equations

In this section we are going to study how curvature comes up naturally in the investigation of certain types of functions. This will lead us to various formulae that make it possible to calculate the curvature tensor on all of the rotationally symmetric and doubly warped product metrics from chapter 1. With this information we can then exhibit the above mentioned examples. This will be accomplished in the next chapter.

4.1. Distance Functions. The functions we wish to look into are *distance functions*. As we don't have a concept of distance yet, we will say that $r : U \to \mathbb{R}$, where $U \subset (M,g)$ is open, is a *distance function* if $|\nabla r| \equiv 1$ on U. Distance functions are therefore simply solutions to the *Hamilton-Jacobi equation*

$$|\nabla r|^2 = 1.$$

This is a nonlinear first-order PDE and can be solved by the method of characteristics see e.g. [**5**]. For now we shall assume that solutions exist and investigate their

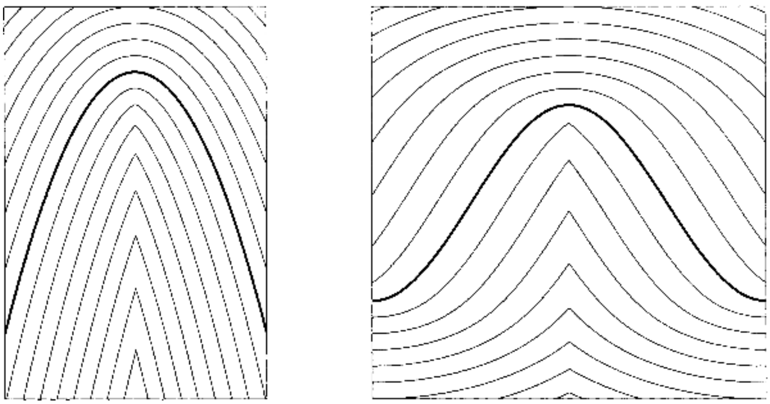

Figure 2.1

properties. Later, when we have developed the theory of geodesics, we shall show the existence of such functions and also show that their name is appropriate.

EXAMPLE 21. *On* $(\mathbb{R}^n, \mathrm{can})$ *define* $r(x) = |x - y|$. *Then* r *is smooth on* $\mathbb{R}^n - \{y\}$ *and has* $|\nabla r| \equiv 1$. *If we have two different points* $\{y, z\}$, *then*

$$r(x) = d\left(x, \{y, z\}\right) = \min\{d(x, y),\ d(x, z)\}$$

is smooth away from $\{y, z\}$ *and the hyperplane* $\{x \in \mathbb{R}^n : |x - y| = |x - z|\}$ *equidistant from* y *and* z.

EXAMPLE 22. *More generally if* $M \subset \mathbb{R}^n$ *is a submanifold, then it can be shown that*

$$r(x) = d(x, M) = \inf\{d(x, y) : y \in M\}$$

is a distance function on some open set $U \subset \mathbb{R}^n$. *If* M *is an orientable hypersurface, then we can see this as follows. Since* M *is orientable, we can choose a unit normal vector field* N *on* M. *Now "coordinatize"* $\mathbb{R}^n$ *as* $x = tN + y$, *where* $t \in \mathbb{R},\ y \in M$. *In some neighborhood* U *of* M *these "coordinates" are actually well-defined. In other words, there is some function* $\varepsilon(y) : M \to (0, \infty)$ *such that any point in*

$$U = \{tN + y : y \in M,\ |t| < \varepsilon(y)\}$$

has unique coordinates (t, y). *We can now define* $r(x) = t$ *on* U *or* $r(x) = d(x, M) = |t|$ *on* $U - M$. *Both functions will then define distance functions on their respective domains. Here* r *is usually referred to as the* signed distance *to* M, *while* f *is just the regular distance. Figure 2.1 shows some pictures of the level sets of a distance function together with the orthogonal trajectories that form the integral curves for the gradient of the distance function.*

EXAMPLE 23. *On* $I \times M$, *where* $I \subset \mathbb{R}$, *is an interval we have metrics of the form* $dr^2 + g_r$, *where* dr^2 *is the standard metric on* I *and* g_r *is a metric on* $\{r\} \times M$ *that depends on* r. *In this case the projection* $I \times M \to I$ *is a distance function. Special cases of this situation are rotationally symmetric metrics, doubly warped products, and our submersion metrics on* $I \times S^{2n-1}$.

LEMMA 4. *Given* $r : U \to I \subset \mathbb{R}$, *then* r *is a distance function iff* r *is a Riemannian submersion.*

PROOF. In general, we have $dr(v) = g(\nabla r, v)$, so $Dr(v) = dr(v)\partial_t = 0$ iff $v \perp \nabla r$. Thus, v is perpendicular to the kernel of Dr iff it is proportional to ∇r. For such $v = \alpha \nabla r$ we have that

$$Dr(v) = \alpha Dr(\nabla r) = \alpha g(\nabla r, \nabla r)\partial_t.$$

Now ∂_t has length 1 in I, so

$$\begin{aligned} |v| &= |\alpha| |\nabla r|, \\ |Dr(v)| &= |\alpha| |\nabla r|^2. \end{aligned}$$

Thus, r is a Riemannian submersion iff $|\nabla r| = 1$ □

Before continuing we need some simplifying notation. A distance function $r : U \to \mathbb{R}$ is fixed and $U \subset (M, g)$ is an open subset of a Riemannian manifold. The gradient ∇r will usually be denoted by $\partial_r = \nabla r$. The ∂_r notation comes from our warped product metrics $dr^2 + g_r$. The level sets for r are denoted $U_r = \{x \in U : r(x) = r\}$, and the induced metric on U_r is g_r. In this spirit ∇^r, R^r are the Riemannian connection and curvature on (U_r, g_r). The $(1,1)$ version of the Hessian of r is denoted by $S(\cdot) = \nabla_\cdot \partial_r$, i.e., $\text{Hess}r(X,Y) = g(S(X), Y)$. S stands for second derivative or *shape operator* or *second fundamental form*, depending on the point of view of the observer. The last two terms are more or less synonymous and refer to the shape of (U_r, g_r) in $(U, g) \subset (M, g)$. The idea is that $S = \nabla \partial_r$ measures how the induced metric on U_r changes by computing how the unit normal to U_r changes.

EXAMPLE 24. *Let $M \subset \mathbb{R}^n$ be an orientable hypersurface, N the unit normal, and S the shape operator defined by $S(v) = \nabla_v N$ for $v \in TM$. If $S \equiv 0$ on M then N must be a constant vector field on M, and hence M is an open subset of the hyperplane*

$$H = \{x + p \in \mathbb{R}^n : x \cdot N_p = 0\},$$

where $p \in M$ is fixed. As an explicit example of this, recall our isometric immersion or embedding $(\mathbb{R}^{n-1}, \text{can}) \to (\mathbb{R}^n, \text{can})$ from chapter 1 defined by

$$(x^1, \ldots, x^{n-1}) \to (\gamma(x^1), x^2, \ldots, x^{n-1}),$$

where γ is a unit speed curve $\gamma : \mathbb{R} \to \mathbb{R}^2$. In this case,

$$N = (N(x^1), 0, \ldots, 0)$$

is a unit normal, where $N(x^1)$ is the unit normal to γ in $\mathbb{R}^2$. We can write this as

$$N = (-\dot\gamma^2(x^1), \dot\gamma^1(x^1), 0, \ldots, 0)$$

in Cartesian coordinates. So

$$\begin{aligned} \nabla N &= -d(\dot\gamma^2)\partial_1 + d(\dot\gamma^1)\partial_2 \\ &= -\ddot\gamma^2 dx^1 \partial_1 + \ddot\gamma^1 dx^1 \partial_2 \\ &= (-\ddot\gamma^2 \partial_1 + \ddot\gamma^1 \partial_2) dx^1. \end{aligned}$$

Thus, $S \equiv 0$ iff $\ddot\gamma^1 = \ddot\gamma^2 = 0$ iff γ is a straight line iff M is an open subset of a hyperplane. The shape operator therefore really captures the idea that the hypersurface bends in $\mathbb{R}^n$, even though $\mathbb{R}^{n-1}$ cannot be seen to bend inside itself.

We have seen here the difference between *extrinsic* and *intrinsic* geometry. Intrinsic geometry is everything we can do on a Riemannian manifold (M,g) that does not depend on how (M,g) might be isometrically immersed in some other Riemannian manifold. Extrinsic geometry is the study of how an isometric immersion $(M,g)\to(N,g_N)$ bends (M,g) inside (N,g_N). Thus, the curvature tensor on (M,g) measures how the space bends intrinsically, while the shape operator measures extrinsic bending.

4.2. Curvature Equations. We are now ready to establish our first fundamental equation.

THEOREM 2. (The Radial Curvature Equation) *If* $U\subset(M,g)$ *is an open set and* $r:U\to\mathbb{R}$ *a distance function, then*

$$\nabla_{\partial_r}S+S^2=-R_{\partial_r}.$$

PROOF. We proceed by straightforward computation. If X is a vector field on U, then

$$\begin{aligned}(\nabla_{\partial_r}S)(X)+S^2(X)&=\nabla_{\partial_r}(S(X))-S(\nabla_{\partial_r}X)+S(S(X))\\&=\nabla_{\partial_r}\nabla_X\partial_r-\nabla_{\nabla_{\partial_r}X}\partial_r+\nabla_{\nabla_X\partial_r}\partial_r\\&=\nabla_{\partial_r}\nabla_X\partial_r-\nabla_{\nabla_{\partial_r}X-\nabla_X\partial_r}\partial_r\\&=\nabla_{\partial_r}\nabla_X\partial_r-\nabla_{[\partial_r,X]}\partial_r.\end{aligned}$$

In order for this to equal $-R(X,\partial_r)\partial_r$ we only need to check what happened to $-\nabla_X\nabla_{\partial_r}\partial_r$. However, as $\partial_r=\nabla r$ is unit, we see that for any vector field Y on U:

$$\begin{aligned}g(\nabla_{\partial_r}\partial_r,Y)&=\mathrm{Hess}r\,(\partial_r,Y)\\&=\mathrm{Hess}r\,(Y,\partial_r)\\&=g(\nabla_Y\partial_r,\partial_r)\\&=\frac{1}{2}D_Yg(\partial_r,\partial_r)\\&=\frac{1}{2}D_Y1=0.\end{aligned}$$

In particular, $\nabla_{\partial_r}\partial_r=S(\partial_r)=0$ on all of U. □

This result tells us two things: First, that ∂_r is always a zero eigenvector for S and secondly how certain "radial curvatures" relate to the Hessian of r. The Hessian of a generic function cannot, of course, exhibit such predictable behavior (namely, being a solution to a PDE). It is only geometrically relevant functions that behave so nicely.

The second and third fundamental equations are also known as the *Gauss equations* and *Codazzi-Mainardi equations,* respectively. They will be proved simultaneously but stated separately. For a vector we use the notation for decomposing it into normal and tangential components to U_r:

$$\begin{aligned}v&=\tan v+\mathrm{nor}v\\&=v-g(v,\partial_r)\,\partial_r+g(v,\partial_r)\,\partial_r.\end{aligned}$$

THEOREM 3. (The Tangential Curvature Equation)

$$\begin{aligned}\tan R(X,Y)Z&=R^r(X,Y)Z-(S(X)\wedge S(Y))(Z),\\g(R(X,Y)Z,W)&=g_r(R^r(X,Y)Z,W)-\mathrm{II}(Y,Z)\mathrm{II}(X,W)+\mathrm{II}(X,Z)\mathrm{II}(Y,W).\end{aligned}$$

Here X, Y, Z, W are tangent to the level sets U_r and

$$\mathrm{II}\,(U,V) = \mathrm{Hess}r\,(U,V) = g\,(S\,(U)\,,V)$$

is the classical second fundamental form.

THEOREM 4. (The Normal or Mixed Curvature Equation)

$$\begin{aligned} g\,(R(X,Y)Z,\partial_r) &= g(-(\nabla_X S)(Y) + (\nabla_Y S)(X), Z) \\ &= -\,(\nabla_X \mathrm{II})\,(Y,Z) + (\nabla_Y \mathrm{II})\,(X,Z)\,. \end{aligned}$$

where X, Y, Z are tangent to the level sets U_r.

PROOF. The proofs hinge on the important fact that if X, Y are vector fields that are tangent to the level sets U_r, then:

$$\begin{aligned} \nabla^r_X Y &= \tan(\nabla_X Y) \\ &= \nabla_X Y - g\,(\nabla_X Y, \partial_r)\,\partial_r \\ &= \nabla_X Y + g(S(X), Y)\partial_r \\ &= \nabla_X Y + \mathrm{II}(X,Y)\partial_r \end{aligned}$$

Here the first equality is a consequence of the uniqueness of the Riemannian connection on (U_r, g_r). One can check either that $\tan(\nabla_X Y)$ satisfies properties 1-4 of a Riemannian connection or alternatively that it satisfies the Koszul formula. The latter task is almost immediate. The second and fourth equality are obvious. The third follows as $Y \perp \partial_r$ implies

$$\begin{aligned} 0 &= \nabla_X g(Y, \partial_r) \\ &= g(\nabla_X Y, \partial_r) + g(Y, S(X)), \end{aligned}$$

whence

$$g(S(X), Y) = -g(\nabla_X Y, \partial_r).$$

Both of the curvature equations are now verified by calculating $R(X,Y)Z$ using

$$\nabla_X Y = \nabla^r_X Y - g(S(X), Y) \cdot \partial_r.$$

$$\begin{aligned} R(X,Y)Z &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ &= \nabla_X (\nabla^r_Y Z - g(S(Y), Z) \cdot \partial_r) - \nabla_Y (\nabla^r_X Z - g(S(X), Z) \cdot \partial_r) \\ &\quad -\nabla^r_{[X,Y]} Z + g(S([X,Y]), Z) \cdot \partial_r \\ &= \nabla_X \nabla^r_Y Z - \nabla_Y \nabla^r_X Z - \nabla^r_{[X,Y]} Z \\ &\quad -\nabla_X \left(g(S(Y), Z) \cdot \partial_r\right) + \nabla_Y \left(g(S(X), Z) \cdot \partial_r\right) + g(S([X,Y]), Z) \cdot \partial_r \\ &= R^r(X,Y)Z - g\,(S\,(X)\,, \nabla_Y Z) \cdot \partial_r + g(S(Y), \nabla_X Z) \cdot \partial_r \\ &\quad -g(S(Y), \nabla_X Z) \cdot \partial_r + g(S(X), \nabla_Y Z) \cdot \partial_r \\ &\quad -g(\nabla_X S(Y), Z) \cdot \partial_r + g(\nabla_Y S(X), Z) \cdot \partial_r + g(S([X,Y]), Z) \cdot \partial_r \\ &\quad -g(S(Y), Z)S(X) + g(S(X), Z)S(Y) \\ &= R^r(X,Y)Z - (S(X) \wedge S(Y))\,(Z) \\ &\quad +g\,(-(\nabla_X S)(Y) + (\nabla_Y S)(X), Z) \cdot \partial_r \end{aligned}$$

This establishes the first part of each formula. The second parts follow from using the definitions of the involved concepts. □

The three fundamental equations give us a way of computing curvature tensors by induction on dimension. More precisely, if we know how to do computations on U_r and also how to compute S, then we can compute anything on U. We shall clarify and exploit this philosophy in subsequent chapters.

Here we confine ourselves to some low dimensional observations. Recall that the three curvature quantities sec, Ric, and scal obeyed some special relationships in dimensions 2 and 3. Curiously enough this also manifests itself in our three fundamental equations.

If M has dimension 1, then there aren't too many distance functions. Our equations don't even seem to apply here since the level sets are points. This is related to the fact that $R \equiv 0$ on all 1 dimensional spaces.

If M has dimension 2, then any distance function $r : U \subset M \to \mathbb{R}$ has 1-dimensional level sets. Thus $R^r \equiv 0$ and the three vectors X, Y and Z are proportional. Our equations therefore reduce to the single equation:

$$\nabla_{\partial_r} S + S^2 = -R_{\partial_r}.$$

Actually, since $S(\partial_r) = 0$, we know that S depends only on its value on a unit vector $v \in TU_r$ thus $S(v) = \alpha v$, where $\alpha = \mathrm{tr} S = \Delta r$. The radial curvature equation can therefore be reduced to:

$$\partial_r(\Delta r) + (\Delta r)^2 = -\sec(T_p M).$$

To be even more concrete, we have that g_r on U_r can be written: $g_r = \varphi^2(r,\theta)d\theta^2$; so

$$g = dr^2 + \varphi^2(r,\theta)d\theta^2,$$

and since

$$\begin{aligned} \varphi \partial_r \varphi &= \frac{1}{2}\partial_r g(\partial_\theta, \partial_\theta) \\ &= g(\nabla_{\partial_r}\partial_\theta, \partial_\theta) \\ &= g(S(\partial_\theta), \partial_\theta) \\ &= \alpha |\partial_\theta|^2 \\ &= \alpha \varphi^2, \end{aligned}$$

we have

$$\mathrm{tr} S = \frac{\partial_r \varphi}{\varphi},$$

implying

$$-\sec(T_p M) = \frac{\partial_r^2 \varphi}{\varphi}.$$

When M has dimension 3, the level sets of r are 2-dimensional. The radial curvature equation therefore doesn't reduce, but in the other two equations we have that one of the three vectors X, Y, Z is a linear combination of the other two. We might as well assume that $X \perp Y$ and $Z = X$ or Y. So, if $\{X, Y, \partial_r\}$ represents an orthonormal framing, then the complete curvature tensor depends on the quantities: $g(R(X,\partial_r)\partial_r, Y)$, $g(R(X,\partial_r)\partial_r, X)$, $g(R(Y,\partial_r)\partial_r, Y)$, $g(R(X,Y)Y,X)$, $g(R(X,Y)Y,\partial_r)$, $g(R(Y,X)X,\partial_r)$. The first three quantities can be computed from the radial curvature equation, the fourth from the tangential curvature equation, and the last two from the mixed curvature equation.

In the special case where $M^3 = \mathbb{R}^3$, $R = 0$, the tangential curvature equation is particularly interesting:

$$\begin{aligned} \sec(T_pU_r) &= R^r(X,Y,Y,X) \\ &= g(S(X),X)g(S(Y),Y) - g(S(X),Y)g(S(X),Y) \\ &= \det S \end{aligned}$$

This was *Gauss's wonderful observation*! namely, that the extrinsic quantity $\det S$ for U_r is actually the intrinsic quantity, $\sec(T_pU_r)$.

Finally, in dimension 4 everything reaches its most general level. We can start with an orthonormal framing $\{X, Y, Z, \partial_r\}$, and there will be twenty curvature quantities to compute.

5. The Equations of Riemannian Geometry

In this section we shall investigate the connection between the metric tensor and curvature. This is done by using the radial curvature equation together with some new formulae. Having established these fundamental equations, we shall introduce some useful vector fields that make it possible to see how the curvature influences the metric in some unexpected ways.

Recall from the end of the last section that we arrived at a very nice formula for the relationship between the metric and curvature on a surface, namely, if $g = dr^2 + \varphi^2(r,\theta)d\theta^2$, then $\partial_r^2\varphi = -\sec\cdot\varphi$. This formula can be used not only to compute curvatures from knowledge of the metric, but also in reverse to conclude things about the metric from the curvature. This relationship, which is classical for surfaces, will be generalized in this section to manifolds of any dimension and then extensively used throughout the entire text as a universal tool for understanding the relationship between the metric and curvature.

5.1. The Coordinate-Free Equations. We need to introduce an ad hoc concept for Hessians and symmetric bilinear forms on Riemannian manifolds. If $B(X,Y)$ is a symmetric $(0,2)$-tensor and $L(X)$ the corresponding self-adjoint $(1,1)$-tensor defined via

$$g(L(X),Y) = B(X,Y),$$

then the *square* of B is the symmetric bilinear form corresponding to L^2

$$B^2(X,Y) = g\left(L^2(X),Y\right) = g(L(X),L(Y)).$$

Note that this symmetric bilinear form is always nonnegative, i.e., $B^2(X,X) \geq 0$ for all X.

PROPOSITION 7. *If we have a smooth distance function $r : (U,g) \to \mathbb{R}$ and denote $\nabla r = \partial_r$, then*

(1) $L_{\partial_r} g = 2\mathrm{Hess}r$,
(2) $(\nabla_{\partial_r}\mathrm{Hess}r)(X,Y) + \mathrm{Hess}^2 r(X,Y) = -R(X,\partial_r,\partial_r,Y)$,
(3) $(L_{\partial_r}\mathrm{Hess}r)(X,Y) - \mathrm{Hess}^2 r(X,Y) = -R(X,\partial_r,\partial_r,Y)$.

PROOF. (1) is simply the definition of the Hessian.

To prove (2) and (3) we use that $\nabla_{\partial_r}\partial_r = 0$ and perform virtually the same calculations that were used for the radial curvature equation. Keep in mind that

$\nabla_X \partial_r = S(X)$ is the self-adjoint operator corresponding to Hess$r$.

$$\begin{aligned}
(\nabla_{\partial_r}\mathrm{Hess}r)(X,Y) &= \partial_r \mathrm{Hess}r(X,Y) - \mathrm{Hess}r(\nabla_{\partial_r}X,Y) - \mathrm{Hess}r(X,\nabla_{\partial_r}Y)\\
&= \partial_r g(\nabla_X\partial_r,Y) - g\left(\nabla_{\nabla_{\partial_r}X}\partial_r,Y\right) - g(\nabla_X\partial_r,\nabla_{\partial_r}Y)\\
&= g(\nabla_{\partial_r}\nabla_X\partial_r,Y) - g\left(\nabla_{\nabla_{\partial_r}X}\partial_r,Y\right)\\
&\quad +g(\nabla_X\partial_r,\nabla_{\partial_r}Y) - g(\nabla_X\partial_r,\nabla_{\partial_r}Y)\\
&= g(R(\partial_r,X)\partial_r,Y) - g(\nabla_{\nabla_X\partial_r}\partial_r,Y)\\
&= -R(X,\partial_r,\partial_r,Y) - g(\nabla_Y\partial_r,\nabla_X\partial_r)\\
&= -R(X,\partial_r,\partial_r,Y) - \mathrm{Hess}^2 r(X,Y).
\end{aligned}$$

$$\begin{aligned}
(L_{\partial_r}\mathrm{Hess}r)(X,Y) &= \partial_r \mathrm{Hess}r(X,Y) - \mathrm{Hess}r([\partial_r,X],Y) - \mathrm{Hess}r(X,[\partial_r,Y])\\
&= \partial_r g(\nabla_X\partial_r,Y) - g\left(\nabla_{[\partial_r,X]}\partial_r,Y\right) - g(\nabla_X\partial_r,[\partial_r,Y])\\
&= g(\nabla_{\partial_r}\nabla_X\partial_r,Y) - g\left(\nabla_{[\partial_r,X]}\partial_r,Y\right)\\
&\quad +g(\nabla_X\partial_r,\nabla_{\partial_r}Y) - g(\nabla_X\partial_r,\nabla_{\partial_r}Y - \nabla_Y\partial_r)\\
&= g(R(\partial_r,X)\partial_r,Y) + g(\nabla_X\partial_r,\nabla_Y\partial_r)\\
&= -R(X,\partial_r,\partial_r,Y) + \mathrm{Hess}^2 r(X,Y).
\end{aligned}$$

□

The first equation shows how the Hessian controls the metric. The second and third equations give us control over the Hessian if we have information about the curvature. These two equations are different in a very subtle way. The third equation is at the moment the easiest to work with as it only uses Lie derivatives and hence can be put in a nice form in an appropriate coordinate system. The second equation is ultimately more useful, but requires that we find a way of making it easier to interpret.

In the next two sections we shall see how appropriate choices for vector fields can give us a better understanding of these fundamental equations.

5.2. Jacobi Fields. A *Jacobi field* for a smooth distance function r is a smooth vector field J that does not depend on r, i.e., it satisfies the *Jacobi equation*

$$L_{\partial_r}J = 0.$$

This is a first order linear PDE, which can be solved by the *method of characteristics*. To see how this is done we locally select a coordinate system $\left(r, x^2, ..., x^n\right)$ where r is the first coordinate. Then $J = a^r\partial_r + a^i\partial_i$ and the Jacobi equation becomes:

$$\begin{aligned}
0 &= L_{\partial_r}J\\
&= L_{\partial_r}\left(a^r\partial_r + a^i\partial_i\right)\\
&= \partial_r(a^r)\partial_r + \partial_r\left(a^i\right)\partial_i.
\end{aligned}$$

Thus the coefficients a^r, a^i have to be independent of r as already indicated. What is more, we can construct such Jacobi fields knowing the values on a hypersurface $H \subset M$ where $\left(x^2, ..., x^n\right)|_H$ is a coordinate system. In this case ∂_r is transverse to H and so we can solve the equations by declaring that a^r, a^i are constant along the integral curves for ∂_r. Note that the coordinate vector fields are themselves Jacobi fields. Jacobi fields satisfy a more general second order equation, also known as the *Jacobi Equation*:

$$\nabla_{\partial_r}\nabla_{\partial_r}J = -R(J,\partial_r)\partial_r,$$

since

$$\begin{aligned}
-R(J,\partial_r)\partial_r &= R(\partial_r,J)\partial_r \\
&= \nabla_{\partial_r}\nabla_J\partial_r - \nabla_J\nabla_{\partial_r}\partial_r - \nabla_{[\partial_r,J]}\partial_r \\
&= \nabla_{\partial_r}\nabla_J\partial_r \\
&= \nabla_{\partial_r}\nabla_{\partial_r}J.
\end{aligned}$$

This is a second order equation and must therefore have more solutions than the above first order equation. This equation will be studied further in chapter 3 for rotationally symmetric metrics and for general Riemannian manifolds in chapter 6.

If we evaluate equations (1) and (3) on Jacobi fields we obtain

$$\begin{aligned}
&(1) \quad \partial_r\left(g\left(J_1,J_2\right)\right) = 2\mathrm{Hess}r\left(J_1,J_2\right), \\
&(3) \quad \partial_r\left(\mathrm{Hess}r\left(J_1,J_2\right)\right) - \mathrm{Hess}^2 r\left(J_1,J_2\right) = -R\left(J_1,\partial_r,\partial_r,J_2\right).
\end{aligned}$$

As we now only have directional derivatives we have a much simpler version of the fundamental equations. Therefore, there is a much better chance of predicting how g and Hess$r$ change depending on our knowledge of Hess$r$ and R respectively.

This can be reduced a bit further if we take a product neighborhood $\Omega = (a,b)\times H \subset M$ such that $r(t,z) = t$. On this product the metric has the form

$$g = dr^2 + g_r$$

where g_r is a one parameter family of metrics on H. If J is a vector field on H, then there is a unique extension to a Jacobi field on $\Omega = (a,b)\times H$. First observe that

$$\begin{aligned}
\mathrm{Hess}r\left(\partial_r,J\right) &= g\left(\nabla_{\partial_r}\partial_r,J\right) = 0, \\
g_r\left(\partial_r,J\right) &= 0.
\end{aligned}$$

Thus we only need to consider the restrictions of g and Hess$r$ to H. By doing this we obtain

$$\partial_r g = \partial_r g_r = 2\mathrm{Hess}r$$

The fundamental equations can therefore be written as

$$\begin{aligned}
&(1) \quad \partial_r g_r = 2\mathrm{Hess}r, \\
&(3) \quad \partial_r\mathrm{Hess}r - \mathrm{Hess}^2 r = -R\left(\cdot,\partial_r,\partial_r,\cdot\right).
\end{aligned}$$

There is a sticky point that is hidden in (3). Namely, how to extract information from R and pass it on to the Hessian. As we usually make assumptions about the sectional curvature we should try to rewrite this term. This can be done as follows:

$$\begin{aligned}
R\left(X,\partial_r,\partial_r,X\right) &= \sec\left(X,\partial_r\right)\left(g\left(X,X\right)g\left(\partial_r,\partial_r\right) - \left(g\left(X,\partial_r\right)\right)^2\right) \\
&= \sec\left(X,\partial_r\right)g\left(X - g\left(X,\partial_r\right)\partial_r, X - g\left(X,\partial_r\right)\partial_r\right) \\
&= \sec\left(X,\partial_r\right)g_r\left(X,X\right).
\end{aligned}$$

So if we evaluate (3) on a Jacobi field J we obtain

$$\partial_r\left(\mathrm{Hess}r\left(J,J\right)\right) - \mathrm{Hess}^2 r\left(J,J\right) = -\sec\left(J,\partial_r\right)g_r\left(J,J\right).$$

This means that (1) and (3) are coupled as we have not eliminated the metric from (3). The next subsection shows how we can deal with this by evaluating on different vector fields.

Nevertheless, we have reduced (1) and (3) to a set of ODEs where r is the independent variable along the integral curve for ∂_r through p.

5.3. Parallel Fields. A *parallel field* for a smooth distance function is a vector field X such that:

$$\nabla_{\partial_r} X = 0.$$

This is, like the Jacobi equation, a first order linear PDE and can be solved in a similar manner. There is, however, one crucial difference: Parallel fields are almost never Jacobi fields.

If we evaluate g on a pair of parallel fields we see that

$$\partial_r g(X,Y) = g(\nabla_{\partial_r} X, Y) + g(X, \nabla_{\partial_r} Y) = 0.$$

This means that (1) is not simplified by using parallel fields. The second equation, on the other hand, now looks like

$$\partial_r (\text{Hess} r(X,Y)) + \text{Hess}^2 r(X,Y) = -R(X,\partial_r,\partial_r,Y).$$

If we rewrite this in terms of sectional curvature we obtain as above

$$\partial_r (\text{Hess} r(X,X)) + \text{Hess}^2 r(X,X) = -\sec(X,\partial_r)\, g_r(X,X).$$

But this time we know that $g_r(X,X)$ is constant in r as X is parallel. We can even assume that $g(X,\partial_r) = 0$ and $g(X,X) = 1$ by first projecting X onto H and then scaling it. Therefore, (2) takes the form

$$\partial_r (\text{Hess} r(X,X)) + \text{Hess}^2 r(X,X) = -\sec(X,\partial_r)$$

on unit parallel fields that are orthogonal to ∂_r. In this way we really have decoupled the equation for the Hessian from the metric. This allows us to glean information about the Hessian from information about sectional curvature. Equation (1), when rewritten using Jacobi fields, then gives us information about the metric from the information we just obtained about the Hessian using parallel fields.

5.4. Conjugate Points. In general, we might think of the curvatures R_{∂_r} as being given. They could be constant or merely satisfy some inequality. We then wish to investigate how the curvature influences the metric. Equation (1) is linear. Thus the metric can't blow up in finite time unless the Hessian also blows up. However, if we assume that the curvature is bounded, then equation (2) tells us that, if the Hessian blows up, then it must be decreasing in r, hence it can only go to $-\infty$. Going back to (1), we then conclude that the only degeneration which can occur along an integral curve for ∂_r, is that the metric stops being positive definite. We say that the distance function r develops a *conjugate, or focal, point* along this integral curve if this occurs. Below we have some pictures of how conjugate points can develop. Note that as the metric itself is Euclidean, these singularities exist only in the coordinates, not in the metric.

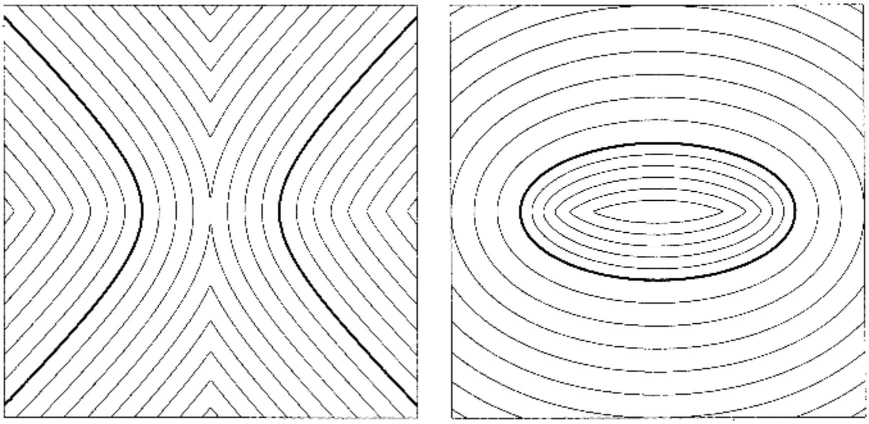

Figure 2.2

It is worthwhile investigating equations (2) and (3) a little further. If we rewrite them as

$$\begin{aligned}(2)\quad & (\nabla_{\partial_r}\mathrm{Hess}r)(X,X) = -R(X,\partial_r,\partial_r,X) - \mathrm{Hess}^2 r(X,X),\\ (3)\quad & (L_{\partial_r}\mathrm{Hess}r)(X,X) = -R(X,\partial_r,\partial_r,X) + \mathrm{Hess}^2 r(X,X),\end{aligned}$$

then we can think of the curvatures as representing fixed *external forces,* while $\mathrm{Hess}^2 r$ describes an *internal reaction (or interaction).* The reaction term is always of a fixed sign and, it will try to force $\mathrm{Hess}r$ blow up in finite time. If, for instance $\sec \leq 0$, then $L_{\partial_r}\mathrm{Hess}r$ is positive. Therefore, if $\mathrm{Hess}r$ is positive at some point, then it will stay positive. On the other hand, if $\sec \geq 0$, then $\nabla_{\partial_r}\mathrm{Hess}r$ is negative, forcing $\mathrm{Hess}r$ to stay nonpositive if it is nonpositive at a point.

In chapters 6, 7, 9, and 11 we shall study and exploit this in much greater detail.

6. Some Tensor Concepts

In this section we shall collect together some notational baggage that is needed from time to time.

6.1. Type Change. The inner product structure on the tangent spaces to a Riemannian manifold makes it possible to view tensors in different ways. We saw this with the Hessian and the Ricci tensor. This is nothing but the elementary observation that a bilinear map can be interpreted as a linear map when one has an inner product present.

If, in general, we have an (s,t)-tensor T, we view it as a section in the bundle

$$\underbrace{TM\otimes\cdots\otimes TM}_{s \text{ times}}\otimes\underbrace{T^*M\otimes\cdots\otimes T^*M}_{t \text{ times}}$$

Then given a Riemannian metric g on M, we can make it into an $(s-k,t+k)$-tensor for any $k\in\mathbb{Z}$ such that both $s-k$ and $t+k$ are nonnegative. Abstractly, this is done as follows: On a Riemannian manifold TM is naturally isomorphic to T^*M; the isomorphism is given by sending $v\in TM$ to the linear map $(w\to g(v,w))\in T^*M$. Using this isomorphism we can therefore replace TM by T^*M or vice versa and thus change the type of the tensor.

At a more concrete level what happens is this: We select a frame $E_1, \dots, E_n$ and construct the coframe $\sigma^1, \dots, \sigma^n$. The vectors and covectors (in T^*M) can be written as

$$\begin{aligned} v &= v^i E_i = \sigma^i (v) E_i, \\ \omega &= \alpha_j \sigma^j = \omega (E_j) \sigma^j. \end{aligned}$$

The tensor T can now be written as

$$T = T^{i_1 \cdots i_s}_{j_1 \cdots j_t} E_{i_1} \otimes \cdots \otimes E_{i_s} \otimes \sigma^{j_1} \otimes \cdots \otimes \sigma^{j_t}.$$

Now we need to know how we can change E_i into a covector and σ^j into a vector. As before, the dual to E_i is the covector $w \to g(E_i, w)$, which can be written as

$$g(E_i, w) = g(E_i, E_j) \sigma^j (w) = g_{ij} \sigma^j (w).$$

Conversely, we have to find the vector v corresponding to the covector σ^j. The defining property is

$$g(v, w) = \sigma^j (w).$$

Thus, we have

$$g(v, E_i) = \delta^j_i.$$

If we write $v = v^k E_k$, this gives

$$g_{ki} v^k = \delta^j_i.$$

Letting g^{ij} denote the ijth entry in the inverse of (g_{ij}), we therefore have

$$v = v^i E_i = g^{ij} E_i.$$

Thus,

$$\begin{aligned} E_i &\to g_{ij} \sigma^j, \\ \sigma^j &\to g^{ij} E_i. \end{aligned}$$

Note that using Einstein notation properly will help keep track of the correct way of doing things as long as the inverse of g is given with superscript indices. With this formula one can easily change types of tensors by replacing Es with σs and vice versa. Note that if we used coordinate vector fields in our frame, then one really needs to invert the metric, but if we had chosen an orthonormal frame, then one simply moves indices up and down as the metric coefficients satisfy $g_{ij} = \delta_{ij}$.

Let us list some examples:

The Ricci tensor: We write the Ricci tensor as a $(1,1)$-tensor: $\mathrm{Ric}(E_i) = \mathrm{Ric}^j_i E_j$; thus

$$\mathrm{Ric} = \mathrm{Ric}^i_j \cdot E_i \otimes \sigma^j.$$

As a $(0,2)$-tensor it will look like

$$\mathrm{Ric} = \mathrm{Ric}_{jk} \cdot \sigma^j \otimes \sigma^k = g^i_{ji} \mathrm{Ric}_k \cdot \sigma^j \otimes \sigma^k,$$

while as a $(2,0)$-tensor acting on covectors it will be

$$\mathrm{Ric} = \mathrm{Ric}^{ik} \cdot E_i \otimes E_k = g^{ij} \mathrm{Ric}^k_j \cdot E_i \otimes E_k.$$

The curvature tensor: We start with the $(1,3)$-curvature tensor $R(X,Y)Z$, which we write as

$$R = R^l_{ijk} \cdot E_l \otimes \sigma^i \otimes \sigma^j \otimes \sigma^k.$$

As a $(0,4)$-tensor we get

$$\begin{aligned} R &= R_{ijkl} \cdot \sigma^i \otimes \sigma^j \otimes \sigma^k \otimes \sigma^l \\ &= R_{ijk}^s g_{sl} \cdot \sigma^i \otimes \sigma^j \otimes \sigma^k \otimes \sigma^l, \end{aligned}$$

while as a $(2,2)$-tensor we have:

$$\begin{aligned} R &= R_{ij}^{kl} \cdot E_k \otimes E_l \otimes \sigma^i \otimes \sigma^j \\ &= R_{ijs}^l g^{sk} \cdot E_k \otimes E_l \otimes \sigma^i \otimes \sigma^j. \end{aligned}$$

Here, however, we must watch out, because there are several different ways of doing this. We choose to raise the last index, but we could also have chosen any other index, thus yielding different $(2,2)$-tensors. The way we did it gives essentially the curvature operator.

6.2. Contractions. Contractions are simply traces of tensors. Thus, the contraction of a $(1,1)$-tensor $T = T_j^i \cdot E_i \otimes \sigma^j$ is simply its trace:

$$C(T) = \mathrm{tr}T = T_i^i.$$

If instead we had a $(0,2)$-tensor T, then we could, using the Riemannian structure, first change it to a $(1,1)$-tensor and then take the trace

$$\begin{aligned} C(T) &= C\left(T_{ij} \cdot \sigma^i \otimes \sigma^j\right) \\ &= C\left(T_{ik} g^{kj} \cdot E_k \otimes \sigma^j\right) \\ &= T_{ik} g^{ki}. \end{aligned}$$

In this way the Ricci tensor becomes a contraction:

$$\begin{aligned} \mathrm{Ric} &= \mathrm{Ric}_j^i \cdot E_i \otimes \sigma^j \\ &= R_{ik}^{kj} \cdot E_i \otimes \sigma^j \\ &= R_{iks}^j g^{sk} \cdot E_i \otimes \sigma^j, \end{aligned}$$

or

$$\begin{aligned} \mathrm{Ric} &= \mathrm{Ric}_{ij} \cdot \sigma^i \otimes \sigma^j \\ &= g^{kl} R_{iklj} \cdot \sigma^i \otimes \sigma^j, \end{aligned}$$

which after type change can be seen to give the same expressions. The scalar curvature can be expressed as:

$$\begin{aligned} \mathrm{scal} &= \mathrm{tr}(\mathrm{Ric}) \\ &= \mathrm{Ric}_i^i \\ &= R_{iks}^i g^{sk} \\ &= \mathrm{Ric}_{ik} g^{ki} \\ &= R_{ijkl} g^{jk} g^{il}. \end{aligned}$$

Again, it is necessary to be careful to specify over which indices one contracts in order to get the right answer.

Note that the divergence of a $(1,k)$-tensor S is nothing but a contraction of the covariant derivative ∇S of the tensor. Here one contracts against the new variable introduced by the covariant differentiation.

6.3. Norms of Tensors. There are several conventions in Riemannian geometry for how one should measure the norm of a linear map. Essentially, there are two different norms in use, the *operator norm* and the *Euclidean norm*. The former is defined for a linear map $L : V \to W$ between inner product spaces as

$$|L| = \sup_{|v|=1} |Lv|$$

The Euclidean norm, in contrast, is given by

$$|L| = \sqrt{\mathrm{tr}\,(L^* \circ L)} = \sqrt{\mathrm{tr}\,(L \circ L^*)},$$

where $L^* : W \to V$ is the adjoint. Despite the fact that we use the same notation for these norms, they are almost never equal. If, for instance, $L : V \to V$ is self adjoint and $\lambda_1 \leq \cdots \leq \lambda_n$ the eigenvalues of L counted with multiplicities, then the operator norm is: $\max\{|\lambda_1|, |\lambda_n|\}$, while the Euclidean norm is $\sqrt{\lambda_1^2 + \cdots + \lambda_n^2}$. The Euclidean norm also has the advantage of actually coming from an inner product:

$$\langle L_1, L_2 \rangle = \mathrm{tr} L_1 \circ L_2^* = \mathrm{tr} L_2 \circ L_1^*.$$

As a general rule we shall always use the Euclidean norm.

It is worthwhile to see how the Euclidean norm of some simple tensors can be computed on a Riemannian manifold. Note that this computation uses type changes to compute adjoints and contractions to take traces.

Let us start with a $(1,1)$-tensor $T = T_j^i \cdot E_i \otimes \sigma^j$. We think of this as a linear map $TM \to TM$. Then the adjoint is first of all the dual map $T^* : T^*M \to T^*M$, which we then change to $T^* : TM \to TM$. This means that

$$T^* = T_i^j \cdot \sigma^i \otimes E_j,$$

which after type change becomes

$$T^* = T_l^k g^{lj} g_{ki} \cdot E_j \otimes \sigma^i.$$

Finally,

$$|T|^2 = T_j^i T_l^k g^{lj} g_{ki}.$$

If the frame is orthonormal, this takes the simple form of

$$|T|^2 = T_j^i T_i^j.$$

For a $(0,2)$-tensor $T = T_{ij} \cdot \sigma^i \otimes \sigma^j$ we first have to change type and then proceed as above. In the end one gets the nice formula

$$|T|^2 = T_{ij} T^{ij}.$$

6.4. Positional Notation. A final remark is in order. Many of the above notations could be streamlined even further so as to rid ourselves of some of the notational problems we have introduced by the way in which we write tensors in frames. Namely, tensors $TM \to TM$ (section of $TM \otimes T^*M$) and $T^*M \to T^*M$ (section of $T^*M \otimes TM$) seem to be written in the same way, and this causes some confusion when computing their Euclidean norms. That is, the only difference between the two objects $\sigma \otimes E$ and $E \otimes \sigma$ is in the ordering, not in what they actually do. We simply interpret the first as a map $TM \to TM$ and then the second as $T^*M \to T^*M$, but the roles could have been reversed, and both could be interpreted as maps $TM \to TM$. This can indeed cause great confusion.

One way to at least keep the ordering straight when writing tensors out in coordinates is to be even more careful with our indices and how they are written down. Thus, a tensor T that is a section of $T^*M \otimes TM \otimes T^*M$ should really be written as

$$T = T_i{}^j{}_k \cdot \sigma^i \otimes E_j \otimes \sigma^k.$$

Our standard $(1,1)$-tensor (section of $TM \otimes T^*M$) could therefore be written

$$T = T^i{}_j \cdot E_i \otimes \sigma^j,$$

while the adjoint (section of $T^*M \otimes TM$) before type change is

$$\begin{aligned} T^* &= T_k{}^l \cdot \sigma^k \otimes E_l \\ &= T^i{}_j g_{ki} g^{lj} \cdot \sigma^k \otimes E_l. \end{aligned}$$

Thus, we have the nice formula

$$|T|^2 = T^i{}_j T_i{}^j.$$

In the case of the curvature tensor one would normally write

$$R = R^l{}_{ijk} \cdot E_l \otimes \sigma^i \otimes \sigma^j \otimes \sigma^k,$$

and when changing to the $(2,2)$ version we have

$$\begin{aligned} R &= R^{kl}{}_{ij} \cdot E_k \otimes E_l \otimes \sigma^i \otimes \sigma^j \\ &= R^l{}_{ijs} g^{sk} \cdot E_k \otimes E_l \otimes \sigma^i \otimes \sigma^j. \end{aligned}$$

It is then clear how to keep track of the other $(2,2)$ versions by writing

$$R_i{}^{jk}{}_l = R_{ist}{}^u g^{js} g^{kt} g_{lu}.$$

Nice as this notation is, it is not used consistently in the literature, probably due to typesetting problems. It would be convenient to use it, but in most cases one can usually keep track of things anyway. Most of this notation can of course also be avoided by using invariant (coordinate-free) notation, but often it is necessary to do coordinate or frame computations both in abstract and concrete situations.

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

$$S = S^i_{j_1 \cdots j_k} \cdot E_i \otimes \sigma^{j_1} \otimes \cdots \otimes \sigma^{j_k},$$

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

$$\nabla S = S^i_{j_1 \cdots j_k, j_{k+1}} \cdot E_i \otimes \sigma^{j_1} \otimes \cdots \otimes \sigma^{j_k} \otimes \sigma^{j_{k+1}}.$$

The coefficient $S^i_{j_1 \cdots j_k, j_{k+1}}$ can be computed via the formula

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

where one must find the expression for

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

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

7. Further Study

It is still too early to give useful references. In the upcoming chapters we shall mention several other books on geometry that the reader might wish to consult. At this stage we shall only list the authoritative guide [**60**]. Every differential geometer must have a copy of these tomes, but their effective usefulness has probably passed away. In a way, it is the Bourbaki of differential geometry and should be treated as such.

8. Exercises

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

(2) If $F : M \to M$ is a diffeomorphism, then the push-forward of a vector field is defined as

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

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

(a) Show that $F_*(\nabla_X Y) = \nabla_{F_* X} F_* Y$ for all vector fields.

(b) If $(M, g) = (\mathbb{R}^n, \mathrm{can})$, then isometries 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.

(3) 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 iff the Lie algebra is Abelian.

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

(5) For any $p \in (M, g)$ and orthonormal basis $e_1, \ldots, e_n$ for $T_p M$, show that there is an orthonormal frame $E_1, \ldots, E_n$ 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 $\left[\alpha_i^j(x)\right] \in SO(n)$ and $\alpha_i^j(p) = \delta_i^j$, then this will yield the desired frame provided that the $D_{e_k} \alpha_i^j$ are appropriately prescribed.

(6) (Riemann) As in the previous problem, but now show that there are coordinates $x^1, \ldots, x^n$ 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. Show that in normal coordinates g viewed as a matrix function of x has the expansion

$$\begin{aligned} g &= \sum_{i,j=1}^{n} g_{ij} dx^i dx^j \\ &= \sum_{i=1}^{n} dx^i dx^i \\ &\quad + \sum_{i<j,k<l} R_{ijkl} \left(x^i dx^j - x^j dx^i\right)\left(x^k dx^l - x^l dx^k\right) + o\left(|x|^2\right), \end{aligned}$$

where $R_{ijkl} = g\left(R\left(\partial_i, \partial_j\right)\partial_k, \partial_l\right)(p)$. In dimension 2 this formula reduces to

$$\begin{aligned} g &= dx^2 + dy^2 + R_{1212}\left(xdy - ydx\right)^2 + o\left(x^2 + y^2\right) \\ &= dx^2 + dy^2 - \sec\left(p\right)\left(xdy - ydx\right)^2 + o\left(x^2 + y^2\right). \end{aligned}$$

(7) Let M be an n-dimensional submanifold of $\mathbb{R}^{n+m}$ with the induced metric and assume that we have a local coordinate system given by a parametrization $x^s\left(u^1, ..., u^n\right)$, $s = 1, ..., n+m$. Show that in these coordinates we have:
(a)
$$g_{ij} = \sum_{s=1}^{n+m} \frac{\partial x^s}{\partial u^i}\frac{\partial x^s}{\partial u^j}.$$
(b)
$$\Gamma_{ij,k} = \sum_{s=1}^{n+m} \frac{\partial x^s}{\partial u^k}\frac{\partial^2 x^s}{\partial u^i \partial u^j}.$$
(c) R_{ijkl} depends only on the first and second partials of x^s.

(8) Show that $\mathrm{Hess} f = \nabla df$.

(9) Let r be a distance function and $S\left(X\right) = \nabla_X \partial_r$ the $(1,1)$ version of the Hessian. Show that

$$\begin{aligned} L_{\partial_r} S &= \nabla_{\partial_r} S, \\ L_{\partial_r} S + S^2 &= -R_{\partial_r}. \end{aligned}$$

How do you reconcile this with what happens for the fundamental equations for the $(0,2)$-version of the Hessian?

(10) Let (M, g) be oriented and define the Riemannian volume form $d\mathrm{vol}$ as follows:

$$d\mathrm{vol}\left(v_1, \ldots, v_n\right) = \det\left(g\left(v_i, e_j\right)\right),$$

where $e_1, \ldots, e_n$ is a positively oriented orthonormal basis for T_pM.
(a) Show that if $v_1, \ldots, v_n$ is positively oriented, then
$$d\mathrm{vol}\left(v_1, \ldots, v_n\right) = \sqrt{\det\left(g\left(v_i, v_j\right)\right)}.$$
(b) Show that the volume form is parallel.
(c) Show that in positively oriented coordinates,
$$d\mathrm{vol} = \sqrt{\det\left(g_{ij}\right)} dx^1 \wedge \cdots \wedge dx^n.$$
(d) If X is a vector field, show that
$$L_X d\mathrm{vol} = \mathrm{div}\left(X\right) d\mathrm{vol}.$$
(e) Conclude that the Laplacian has the formula
$$\Delta u = \frac{1}{\sqrt{\det\left(g_{ij}\right)}}\partial_k\left(\sqrt{\det\left(g_{ij}\right)} g^{kl}\partial_l u\right).$$
Given that the coordinates are normal at p we get as in Euclidean space that
$$\Delta f\left(p\right) = \sum_{i=1}^{n} \partial_i \partial_i f.$$

(11) Let (M,g) be a oriented Riemannian manifold with volume form $d\text{vol}$ as above.

(a) If f has compact support, then

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

(b) Show that

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

(c) Show that

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

(d) Establish the integration by parts formula for functions with compact support:

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

(e) Conclude that if f is sub- 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.*

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

(a) Show that X is incompressible iff the local flows it generates are volume preserving (i.e., leave the Riemannian volume form invariant).

(b) Let X be a unit vector field X on $\mathbb{R}^2$. Show that $\nabla X = 0$ if X is incompressible.

(c) Find a unit vector field X on $\mathbb{R}^3$ that is incompressible but where $\nabla X \neq 0$.

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

(a) Show that X is locally the gradient of a distance function iff the orthogonal distribution is integrable.

(b) Show that X is the gradient of a distance function in a neighborhood of $p \in M$ iff the orthogonal distribution has an integral submanifold through p. Hint: It might help to show that $L_X \theta_X = 0$.

(c) Find X with the given conditions so that it is not a gradient field. Hint: Consider S^3.

(14) Given an orthonormal frame $E_1, \ldots, E_n$ on (M,g), define the *structure constants* c_{ij}^k by $[E_i, E_j] = c_{ij}^k E_k$. Then define the Γs and Rs by

$$\begin{aligned} \nabla_{E_i} E_j &= \Gamma_{ij}^k E_k, \\ R\,(E_i, E_j)\,E_k &= R_{ijk}^l E_l \end{aligned}$$

and compute them in terms of the cs. Notice that on Lie groups with left-invariant metrics the structure constants can be assumed to be constant. In this case, computations simplify considerably.

(15) There is yet another effective method for computing the connection and curvatures, namely, the *Cartan formalism*. Let (M, g) be a Riemannian manifold. Given a frame $E_1, \ldots, E_n$, the connection can be written

$$\nabla E_i = \omega_i^j E_j,$$

where ω_i^j are 1-forms. Thus,

$$\nabla_v E_i = \omega_i^j (v) E_j.$$

Suppose now that the frame is orthonormal and let ω^i be the dual coframe, i.e., $\omega^i (E_j) = \delta_j^i$. Show that the *connection forms* satisfy

$$\begin{aligned} \omega_i^j &= -\omega_j^i, \\ d\omega^i &= \omega^j \wedge \omega_j^i. \end{aligned}$$

These two equations can, conversely, be used to compute the connection forms given the orthonormal frame. Therefore, if the metric is given by declaring a certain frame to be orthonormal, then this method can be very effective in computing the connection.

If we think of $\left[\omega_i^j\right]$ as a matrix, then it represents a 1-form with values in the skew-symmetric $n \times n$ matrices, or in other words, with values in the Lie algebra $\mathfrak{so}(n)$ for $O(n)$.

The *curvature forms* Ω_i^j are 2-forms with values in $\mathfrak{so}(n)$. They are defined as

$$R(\cdot, \cdot) E_i = \Omega_i^j E_j.$$

Show that they satisfy

$$d\omega_i^j = \omega_i^k \wedge \omega_k^j + \Omega_i^j.$$

When reducing to Riemannian metrics on surfaces we obtain for an orthonormal frame E_1, E_2 with coframe ω^1, ω^2

$$\begin{aligned} d\omega^1 &= \omega^2 \wedge \omega_2^1, \\ d\omega^2 &= -\omega^1 \wedge \omega_2^1, \\ d\omega_2^1 &= \Omega_2^1, \\ \Omega_2^1 &= \sec \cdot d\mathrm{vol}. \end{aligned}$$

(16) Show that a Riemannian manifold with parallel Ricci tensor has constant scalar curvature. In chapter 3 it will be shown that the converse is not true, and also that a metric with parallel curvature tensor doesn't have to be Einstein.

(17) Show that if R is the $(1,3)$-curvature tensor and Ric the $(0,2)$-Ricci tensor, then

$$(\mathrm{div} R)(X, Y, Z) = (\nabla_X \mathrm{Ric})(Y, Z) - (\nabla_Y \mathrm{Ric})(X, Z).$$

Conclude that $\mathrm{div} R = 0$ if $\nabla \mathrm{Ric} = 0$. Then show that $\mathrm{div} R = 0$ iff the $(1,1)$ Ricci tensor satisfies:

$$(\nabla_X \mathrm{Ric})(Y) = (\nabla_Y \mathrm{Ric})(X) \text{ for all } X, Y.$$

(18) Let G be a Lie group with a bi-invariant metric. Using left-invariant fields establish the following formulas. Hint: First go back to the exercises to chapter 1 and take a peek at chapter 3 where some of these things are proved.

(a) $\nabla_X Y = \frac{1}{2}[X,Y]$.

(b) $R(X,Y)Z = \frac{1}{4}[Z,[X,Y]]$.

(c) $g(R(X,Y)Z,W) = -\frac{1}{4}(g([X,Y],[Z,W]))$. Conclude that the sectional curvatures are nonnegative.

(d) Show that the curvature operator is also nonnegative by showing that:

$$g\left(\mathfrak{R}\left(\sum_{i=1}^{k} X_i \wedge Y_i\right), \left(\sum_{i=1}^{k} X_i \wedge Y_i\right)\right) = \frac{1}{4}\left|\sum_{i=1}^{k}[X_i,Y_i]\right|^2.$$

(e) Show that $\operatorname{Ric}(X,X) = 0$ iff X commutes with all other left-invariant vector fields. Thus G has positive Ricci curvature if the center of G is discrete.

(f) Consider the linear map $\Lambda^2\mathfrak{g} \to [\mathfrak{g},\mathfrak{g}]$ that sends $X \wedge Y$ to $[X,Y]$. Show that the sectional curvature is positive iff this map is an isomorphism. Conclude that this can only happen if $n = 3$ and $\mathfrak{g} = \mathfrak{su}(2)$.

(19) It is illustrative to use the Cartan formalism in the above problem and compute all quantities in terms of the structure constants for the Lie algebra. Given that the metric is bi-invariant, it follows that with respect to an orthonormal basis they satisfy

$$c_{ij}^k = -c_{ji}^k = c_{jk}^i.$$

The first equality is skew-symmetry of the Lie bracket, and the second is bi-invariance of the metric.

(20) Suppose we have two Riemannian manifolds (M, g_M) and (N, g_N). Then the product has a natural product metric $(M \times N, g_M + g_N)$. Let X be a vector field on M and Y one on N, show that if we regard these as vector fields on $M \times N$, then $\nabla_X Y = 0$. Conclude that $\sec(X,Y) = 0$. This means that product metrics always have many curvatures that are zero.

(21) 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.

(a) Show that the distributions are integrable.

(b) Show that around any point in M there is a product neighborhood $U = V_E \times V_F$ such that $(U,g) = (V_E \times V_F, g|_E + g|_F)$, where $g|_E$ and $g|_F$ are the restrictions of g to the two distributions. In other words, M is locally a product metric.

(22) 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 $(U,g) = \left(V \times I, g|_{TV} + dt^2\right)$.

(23) For 3-dimensional manifolds, show that if the curvature operator in diagonal form looks like

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix},$$

then the Ricci curvature has a diagonal form like

$$\begin{pmatrix} \alpha+\beta & 0 & 0 \\ 0 & \beta+\gamma & 0 \\ 0 & 0 & \alpha+\gamma \end{pmatrix}.$$

Moreover, the numbers α, β, γ must be sectional curvatures.

(24) The *Einstein tensor* on a Riemannian manifold is defined as

$$G = \mathrm{Ric} - \frac{\mathrm{scal}}{2} \cdot I.$$

Show that $G = 0$ in dimension 2 and that $\mathrm{div} G = 0$ in higher dimensions. This tensor is supposed to measure the mass/energy distribution. The fact that it is divergence free tells us that energy and momentum are conserved. In a vacuum, one therefore imagines that $G = 0$. Show that this happens in dimensions > 2 iff the metric is Ricci flat.

(25) This exercise will give you a way of finding the curvature tensor from the sectional curvatures. Using the Bianchi identity show that

$$\begin{aligned} -6R(X, Y, Z, W) &= \left.\frac{\partial^2}{\partial s \partial t}\right|_{s=t=0} \{R(X + sZ, Y + tW, Y + tW, X + sZ) \\ &\quad -R(X + sW, Y + tZ, Y + tZ, X + sW)\}. \end{aligned}$$

(26) Using polarization show that the norm of the curvature operator on $\Lambda^2 T_p M$ is bounded by

$$|\mathfrak{R}|_p| \leq c(n) |\mathrm{sec}|_p$$

for some constant $c(n)$ depending on dimension, and where $|\mathrm{sec}|_p$ denotes the largest absolute value for any sectional curvature of a plane in $T_p M$.

(27) We can artificially complexify the tangent bundle to a manifold: $T_{\mathbb{C}} M = TM \otimes \mathbb{C}$. If we have a Riemannian structure, we can extend all the accompanying tensors to this realm. The metric tensor, in particular, gets extended as follows:

$$g_{\mathbb{C}}(v_1 + iv_2, w_1 + iw_2) = g(v_1, w_1) - g(v_2, w_2) + i(g(v_1, w_2) + g(v_2, w_1)).$$

This means that a vector can have complex length zero without being trivial. Such vectors are called *isotropic*. Clearly, they must have the form $v_1 + iv_2$, where $|v_1| = |v_2|$ and $g(v_1, v_2) = 0$. More generally, we can have isotropic subspaces, i.e., those subspace on which $g_{\mathbb{C}}$ vanishes. If, for instance, a plane is generated by two isotropic vectors $v_1 + iv_2$ and $w_1 + iw_2$, where v_1, v_2, w_1, w_2 are orthogonal, then the plane is isotropic. Note that one must be in dimension ≥ 4 to have isotropic planes. We now say that the isotropic curvatures are positive, if "sectional" curvatures on isotropic planes are positive. This means that if $v_1 + iv_2$ and $w_1 + iw_2$ span the plane and v_1, v_2, w_1, w_2 are orthogonal, then

$$0 < R(v_1 + iv_2, w_1 + iw_2, w_1 - iw_2, v_1 - iv_2).$$

(a) Show that the expression $R(v_1 + iv_2, w_1 + iw_2, w_1 - iw_2, v_1 - iv_2)$ is always a real number.
(b) Show that if the original metric is strictly quarter pinched, i.e., all sectional curvatures lie in an open interval of the form $\left(\frac{1}{4}k, k\right)$, then the isotropic curvatures are positive.
(c) Show that if the sum of the two smallest eigenvalues of the original curvature operator is positive, then the isotropic curvatures are positive.

(28) Consider a Riemannian metric (M, g). Now *scale* the metric by multiplying it by a number λ^2. Then we get a new Riemannian manifold $\left(M, \lambda^2 g\right)$. Show that the new connection and $(1,3)$-curvature tensor remain the same, but that sec, scal, and $\mathfrak{R}$ all get multiplied by λ^{-2}.

(29) For a $(1,1)$-tensor T on a Riemannian manifold, show that if E_i is an orthonormal basis, then

$$|T|^2 = \sum |T(E_i)|^2 .$$

(30) If we have two tensors S, T of the same type (r, s), $r = 0, 1$, define the inner product

$$g(S, T)$$

and show that

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

If S is symmetric and T skew-symmetric show that $g(S, T) = 0$.

(31) Recall that complex manifolds have complex tangent spaces. Thus we can multiply vectors by $\sqrt{-1}$. As a generalization of this we can define an *almost complex* structure. This is a $(1,1)$-tensor J such that $J^2 = -I$. Show that the *Nijenhuis tensor:*

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

is indeed a tensor. If J comes from a complex structure then $N = 0$, conversely Newlander&Nirenberg have shown that J comes from a complex structure if $N = 0$.

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

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

Show that ω is a 2-form. Show that $d\omega = 0$ iff $\nabla J = 0$. If the Kähler form is closed, then we call the metric a Kähler metric.