CHAPTER 4

CURVATURE

1. Introduction

The notion of curvature in a Riemannian manifold was introduced by Riemann (See Riemann [Ri]) in a rather geometric manner, which we are now going to describe. Let p be a point of a Riemannian manifold M and let $\sigma \subset T_pM$ be a two dimensional subspace of the tangent space T_pM of M at p. Consider the set of geodesics that start at p and are tangent to σ. The segments of such geodesics in a normal neighborhood $U \subset M$ of p determine a submanifold of dimension two $S \subset M$ (with our present notation, S is the image of $\exp_p$ restricted to $\sigma \cap \exp_p^{-1}(U)$). S has a metric induced from the inclusion. Since Gauss had proved that the curvature of a surface can be expressed in terms of its metric, so Riemann could speak of the curvature of S at p, and indicate it by $K(p,\sigma)$, (nowadays, $K(p,\sigma)$ is called the sectional curvature of M at p with respect to σ). This was the curvature considered by Riemann in [Ri]. It is a natural generalization of the Gaussian curvature of surfaces and it is clear that if $M = R^n$, $K(p,\sigma) = 0$ for all p and all σ.

Riemann did not indicate a way to calculate the sectional curvature starting with the metric of M; that was done a few years later by Christoffel (see Christoffel [Cf]; Cf. also Eq. (2) of this chapter). Indeed, all the work of Riemann contains just one formula, namely, an expression for the metric for which $K(p,\sigma)$ is constant, for all p and σ, and even this formula was presented without proof. (The formula of Riemann will be presented in Exercise 1(c) of Chap. 8.)

As frequently happens in mathematics, a "workable" formulation of the concept of curvature required a long time for its development. When such a formulation finally appeared it had the advantage of being easy to use to prove theorems but it had the disadvantage of being so far removed from the initial intuitive concept that it looked as if it were some kind of arbitrary creation.

This chapter presents a definition of curvature that, intuitively, measures the amount that a Riemannian manifold deviates from being Euclidean (Cf. Def. 2.1). In Chapter 6, we are going to show that the notion of sectional curvature (Cf. Def. 3.2) obtained by starting with this definition of curvature generalizes the notion of Gaussian curvature for surfaces, and coincides with the concept introduced by Riemann.

2. Curvature

2.1 DEFINITION. The *curvature* R of a Riemannian manifold M is a correspondence that associates to every pair $X, Y \in \mathcal{X}(M)$ a mapping $R(X,Y)\colon \mathcal{X}(M) \to \mathcal{X}(M)$ given by

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

where ∇ is the Riemannian connection of M.

Observe that if $M = R^n$, then $R(X,Y)Z = 0$ for all $X, Y, Z \in \mathcal{X}(R^n)$. In fact, if the vector field Z is given by $Z = (z_1, \dots, z_n)$, with the components of Z coming from the natural coordinates of R^n, we obtain

$$\nabla_X Z = (Xz_1, \dots, Xz_n),$$

hence

$$\nabla_Y \nabla_X Z = (YXz_1, \dots, YXz_n),$$

which implies that

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z = 0,$$

as was stated. We are able, therefore, to think of R as a way of measuring how much M deviates from being Euclidean.

Another way of viewing definition 2.1 is to consider a system of coordinates $\{x_i\}$ around $p \in M$. Since $\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = 0$, we obtain

$$R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})\frac{\partial}{\partial x_k} = (\nabla_{\partial/\partial x_j} \nabla_{\partial/\partial x_i} - \nabla_{\partial/\partial x_i} \nabla_{\partial/\partial x_j})\frac{\partial}{\partial x_k},$$

that is, the curvature measures the non-commutativity of the covariant derivative.

These interpretations, are, however, more or less formal. In this chapter we advise the reader to get used to the formal properties of curvature, postponing until Chapter 6 the proof of a more geometric interpretation of curvature. Let us remark also that a frequently encountered definition of curvature in the literature differs from definition 2.1 by a sign.

2.2 PROPOSITION. *The curvature R of a Riemannian manifold has the following properties:*

(i) R is bilinear in $\mathcal{X}(M) \times \mathcal{X}(M)$, that is,

$$R(fX_1 + gX_2, Y_1) = fR(X_1, Y_1) + gR(X_2, Y_1),$$
$$R(X_1, fY_1 + gY_2) = fR(X_1, Y_1) + gR(X_1, Y_2),$$

$f, g \in \mathcal{D}(M), \quad X_1, X_2, Y_1, Y_2 \in \mathcal{X}(M).$

(ii) For any $X, Y \in \mathcal{X}(M)$, the curvature operator $R(X,Y)$: $\mathcal{X}(M) \to \mathcal{X}(M)$ is linear, that is,

$$R(X,Y)(Z+W) = R(X,Y)Z + R(X,Y)W,$$
$$R(X,Y)fZ = fR(X,Y)Z,$$

$f \in \mathcal{D}(M), \quad Z, W \in \mathcal{X}(M).$

Proof. Let us verify (ii) only, leaving (i) as an exercise for the reader. The first part of (ii) is obvious. As for the second, we have

$$\nabla_Y \nabla_X (fZ) = \nabla_Y (f\nabla_X Z + (Xf)Z) = f\nabla_Y \nabla_X Z + (Yf)(\nabla_X Z) + (Xf)(\nabla_Y Z) + (Y(Xf))Z.$$

Therefore,

$$\nabla_Y \nabla_X (fZ) - \nabla_X \nabla_Y (fZ) = f(\nabla_Y \nabla_X - \nabla_X \nabla_Y)Z + ((YX - XY)f)Z,$$

hence

$$R(X,Y)fZ = f\nabla_Y \nabla_X Z - f\nabla_X \nabla_Y Z + ([Y,X]f)Z + f\nabla_{[X,Y]}Z + ([X,Y]f)Z = fR(X,Y)Z. \quad \square$$

2.3 REMARK. An analysis of the proof above shows that the necessity of the appearance of the term $\nabla_{[X,Y]}Z$ in the definition of the curvature is connected to the fact that we want the mapping $R(X,Y)\colon \mathcal{X}(M) \to \mathcal{X}(M)$ to be linear (see the next Rem. 2.6).

2.4 PROPOSITION. *(Bianchi Identity).*

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

Proof. From the symmetry of the Riemannian connection, we have,

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = \nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z + \nabla_{[X,Y]}Z$$
$$+\nabla_Z\nabla_Y X - \nabla_Y\nabla_Z X + \nabla_{[Y,Z]}X + \nabla_X\nabla_Z Y - \nabla_Z\nabla_X Y + \nabla_{[Z,X]}Y$$
$$= \nabla_Y[X,Z] + \nabla_Z[Y,X] + \nabla_X[Z,Y] - \nabla_{[X,Z]}Y - \nabla_{[Y,X]}Z - \nabla_{[Z,Y]}X$$
$$= [Y,[X,Z]] + [Z,[Y,X]] + [X,[Z,Y]] = 0,$$

where the last equality follows from the Jacobi identity for vector fields. □

From now on, we shall write $\langle R(X,Y)Z,T\rangle = (X,Y,Z,T)$.

2.5 PROPOSITION. (a) $(X,Y,Z,T)+(Y,Z,X,T)+(Z,X,Y,T) = 0$
(b) $(X,Y,Z,T) = -(Y,X,Z,T)$
(c) $(X,Y,Z,T) = -(X,Y,T,Z)$
(d) $(X,Y,Z,T) = (Z,T,X,Y)$.

Proof.
(a) is just the Bianchi identity again;
(b) follows directly from Definition 2.1;
(c) is equivalent to $(X,Y,Z,Z) = 0$, whose proof follows:

$$(X,Y,Z,Z) = \langle \nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z + \nabla_{[X,Y]}Z, Z\rangle.$$

But

$$\langle \nabla_Y\nabla_X Z, Z\rangle = Y\langle \nabla_X Z, Z\rangle - \langle \nabla_X Z, \nabla_Y Z\rangle,$$

and

$$\langle \nabla_{[X,Y]}Z, Z\rangle = \frac{1}{2}[X,Y]\langle Z,Z\rangle.$$

Hence

$$(X,Y,Z,Z) = Y\langle \nabla_X Z, Z\rangle - X\langle \nabla_Y Z, Z\rangle + \frac{1}{2}[X,Y]\langle Z,Z\rangle$$
$$= \frac{1}{2}Y(X\langle Z,Z\rangle) - \frac{1}{2}X(Y\langle Z,Z\rangle)$$
$$+ \frac{1}{2}[X,Y]\langle Z,Z\rangle = -\frac{1}{2}[X,Y]\langle Z,Z\rangle$$
$$+ \frac{1}{2}[X,Y]\langle Z,Z\rangle = 0,$$

which proves (c).

In order to prove (d), we use (a), and write:

$$(X,Y,Z,T)+(Y,Z,X,T)+(Z,X,Y,T)=0,$$

$$(Y,Z,T,X)+(Z,T,Y,X)+(T,Y,Z,X)=0,$$

$$(Z,T,X,Y)+(T,X,Z,Y)+(X,Z,T,Y)=0,$$

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

Summing the equations above, we obtain

$$2(Z,X,Y,T)+2(T,Y,Z,X)=0$$

and, therefore,

$$(Z,X,Y,T)=(Y,T,Z,X). \quad \square$$

It is convenient to express what was seen above in a coordinate system $(U,\mathbf{x})$ based at the point $p \in M$. Let us indicate, as usual, $\frac{\partial}{\partial x_i} = X_i$. We put

$$R(X_i, X_j)X_k = \sum_{\ell} R_{ijk}^{\ell} X_{\ell}.$$

Thus R_{ijk}^{ℓ} are the components of the curvature R in $(U,\mathbf{x})$. If

$$X = \sum_i u^i X_i,\ Y = \sum_j v^j X_j,\ Z = \sum_k w^k X_k,$$

we obtain, from the linearity of R,

(1) $$R(X,Y)Z = \sum_{i,j,k,\ell} R_{ijk}^{\ell} u^i v^j w^k X_{\ell}.$$

To express R_{ijk}^{ℓ} in terms of the coefficients Γ_{ij}^{k} of the Riemannian connection, we write,

$$\begin{aligned} R(X_i, X_j)X_k &= \nabla_{X_j}\nabla_{X_i}X_k - \nabla_{X_i}\nabla_{X_j}X_k \\ &= \nabla_{X_j}\Big(\sum_{\ell}\Gamma_{ik}^{\ell}X_{\ell}\Big) - \nabla_{X_i}\Big(\sum_{\ell}\Gamma_{jk}^{\ell}X_{\ell}\Big), \end{aligned}$$

which by a direct calculation yields

$$(2)\qquad R^s_{ijk} = \sum_\ell \Gamma^\ell_{ik}\Gamma^s_{j\ell} - \sum_\ell \Gamma^\ell_{jk}\Gamma^s_{i\ell} + \frac{\partial}{\partial x_j}\Gamma^s_{ik} - \frac{\partial}{\partial x_i}\Gamma^s_{jk}.$$

Putting

$$\langle R(X_i, X_j)X_k, X_s\rangle = \sum_\ell R^\ell_{ijk} g_{\ell s} = R_{ijks},$$

we can write the identities of Proposition 2.5 as:

$$R_{ijks} + R_{jkis} + R_{kijs} = 0$$

$$R_{ijks} = -R_{jiks}$$

$$R_{ijks} = -R_{ijsk}$$

$$R_{ijks} = R_{ksij}.$$

2.6 Remark. The equation (1), which depends on the linearity of the operator R, shows that the value of $R(X,Y)Z$ at the point p depends uniquely on the values of X, Y, Z at p and the values of the functions R^ℓ_{ijk} at p. Observe that this contrasts with the behavior of the covariant derivative (See Rem. 2.3, Chap. 2), the reason being that the covariant derivative is not linear in all of its arguments. In general, entities, such as the curvature, that are linear, are called tensors on M (more details will be given in Section 5).

3. Sectional curvature

Closely related to the curvature operator is the sectional (or Riemannian) curvature that we are now going to define.

In what follows it is convenient to use the following notation. Given a vector space V, we denote by $|x \wedge y|$ the expression

$$\sqrt{|x|^2\,|y|^2 - \langle x, y\rangle^2},$$

which represents the area of a two-dimensional parallelogram determined by the pair of vectors $x, y \in V$.

3.1 PROPOSITION. *Let $\sigma \subset T_pM$ be a two-dimensional subspace of the tangent space T_pM and let $x, y \in \sigma$ be two linearly independent vectors. Then*

$$K(x,y) = \frac{(x,y,x,y)}{|x \wedge y|^2}$$

does not depend on the choice of the vectors $x, y \in \sigma$.

Proof. To avoid calculating, we observe that we can pass from the basis $\{x,y\}$ of σ to any other basis $\{x', y'\}$ by iterating the following elementary transformations:

(a) $\{x,y\} \to \{y,x\}$,

(b) $\{x,y\} \to \{\lambda x, y\}$,

(c) $\{x,y\} \to \{x + \lambda y, y\}$.

It is easy to see that $K(x,y)$ is invariant by such transformations and that completes the proof. □

3.2 DEFINITION. Given a point $p \in M$ and a two-dimensional subspace $\sigma \subset T_pM$, the real number $K(x,y) = K(\sigma)$, where $\{x,y\}$ is any basis of σ, is called the *sectional curvature* of σ at p.

Besides the fact that the sectional curvature has interesting geometrical interpretations, its importance comes from the fact that knowledge of $K(\sigma)$, for all σ, determines the curvature R completely. This is a purely algebraic fact:

3.3 LEMMA. *Let V be a vector space of dimension ≥ 2, provided with an inner product $\langle\ ,\ \rangle$. Let $R: V \times V \times V \to V$ and $R': V \times V \times V \to V$ be tri-linear mappings such that conditions (a), (b), (c) and (d) of Proposition 2.5 are satisfied by*

$$(x,y,z,t) = \langle R(x,y)z, t\rangle, \quad (x,y,z,t)' = \langle R'(x,y)z, t\rangle.$$

If x, y are two linearly independent vectors, we may write,

$$K(\sigma) = \frac{(x,y,x,y)}{|x \wedge y|^2}, \qquad K'(\sigma) = \frac{(x,y,x,y)'}{|x \wedge y|^2},$$

where σ is the bi-dimensional subspace generated by x and y. If for all $\sigma \subset V$, $K(\sigma) = K'(\sigma)$, then $R = R'$.

Proof. It suffices to prove that $(x,y,z,t) = (x,y,z,t)'$ for any $x,y,z,t \in V$. Observe first that, by hypothesis, we have $(x,y,x,y) = (x,y,x,y)'$, for all $x,y \in V$. Then

$$(x+z,y,x+z,y) = (x+z,y,x+z,y)',$$

hence

$$(x,y,x,y) + 2(x,y,z,y) + (z,y,z,y) = (x,y,x,y)' + 2(x,y,z,y)' + (z,y,z,y)'$$

and, therefore

$$(x,y,z,y) = (x,y,z,y)',$$

for all $x,y,z \in V$.

Using what we have just proved, we obtain

$$(x,y+t,z,y+t) = (x,y+t,z,y+t)',$$

hence

$$(x,y,z,t) + (x,t,z,y) = (x,y,z,t)' + (x,t,z,y)',$$

which can be written further as

$$(x,y,z,t) - (x,y,z,t)' = (y,z,x,t) - (y,z,x,t)'.$$

It follows that, the expression $(x,y,z,t) - (x,y,z,t)'$ is invariant by cyclic permutations of the first three elements. Therefore, by (a) of Proposition 2.5, we have

$$3[(x,y,z,t) - (x,y,z,t)'] = 0,$$

hence

$$(x,y,z,t) = (x,y,z,t)'$$

for all $x,y,z,t \in V$. □

The Riemannian manifolds that have constant sectional curvature played a fundamental role in the development of Riemannian Geometry. We shall treat these manifolds in more detail in Chapter 8 of this book. At the moment, we wish only to show how the lemma above allows us to obtain a characterization of such manifolds by means of the components $R_{ijk\ell}$ of the curvature in an orthonormal basis. This follows from the lemma below.

3.4 LEMMA. *Let M be a Riemannian manifold and p a point of M. Define a tri-linear mapping $R'\colon T_pM \times T_pM \times T_pM \to T_pM$ by*

$$\langle R'(X,Y,W),Z\rangle = \langle X,W\rangle\langle Y,Z\rangle - \langle Y,W\rangle\langle X,Z\rangle,$$

for all $X,Y,W,Z \in T_pM$. Then M has constant sectional curvature equal to K_o if and only if $R = K_oR'$, where R is the curvature of M.

Proof. Assume that $K(p,\sigma) = K_o$ for all $\sigma \subset T_pM$, and set $\langle R'(X,Y,W),Z\rangle = (X,Y,W,Z)'$. Observe that R' satisfies the properties (a), (b), (c) and (d) of Proposition 2.5. Since

$$(X,Y,X,Y)' = \langle X,X\rangle\langle Y,Y\rangle - \langle X,Y\rangle^2,$$

we have that, for all pairs of vectors $X,Y \in T_pM$,

$$R(X,Y,X,Y) = K_o(|X|^2\,|Y|^2 - \langle X,Y\rangle^2) = K_oR'(X,Y,X,Y).$$

Lemma 3.3 implies that, for all X,Y,W,Z,

$$R(X,Y,W,Z) = K_oR'(X,Y,W,Z),$$

hence $R = K_oR'$. The converse is immediate.

3.5 COROLLARY. *Let M be a Riemannian manifold, p a point of M and $\{e_1,\dots,e_n\}$, $n = \dim M$, an orthonormal basis of T_pM. Define $R_{ijk\ell} = \langle R(e_i,e_j)e_k,e_\ell\rangle$, $i,j,k,\ell = 1,\dots,n$. Then $K(p,\sigma) = K_o$ for all $\sigma \subset T_pM$, if and only if*

$$R_{ijk\ell} = K_o(\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}),$$

where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & i \neq j. \end{cases}$$

In other words, $K(p,\sigma) = K_o$ for all $\sigma \subset T_pM$ if and only if $R_{ijij} = -R_{ijji} = K_o$ for all $i \neq j$, and $R_{ijk\ell} = 0$ in the other cases.