A Course in Riemannian Geometry

David R. Wilkins

Contents

1 Smooth Manifolds

1.1 Smooth Manifolds

A topological space M is said a *topological manifold* of dimension n if it is metrizable (i.e., there exists a distance function d on M which generates the topology of M) and every point of M has an open neighbourhood homeomorphic to an open set in n-dimensional Euclidean space $\mathbb{R}^n$.

Let M be a topological manifold. A *continuous coordinate system* defined over an open set U in M is defined to be an n-tuple $(x^1, x^2, \ldots, x^n)$ of continuous real-valued functions on U such that the map $\varphi\colon U \to \mathbb{R}^n$ defined by

$$\varphi(u) = \bigl(x^1(u), x^2(u), \ldots, x^n(u)\bigr)$$

maps U homeomorphically onto some open set in $\mathbb{R}^n$. The domain U of the coordinate system $(x^1, x^2, \ldots, x^n)$ is referred to as a *coordinate patch* on M.

Two continuous coordinate systems $(x^1, x^2, \ldots, x^n)$ and $(y^1, y^2, \ldots, y^n)$ defined over coordinate patches U and V are said to be *smoothly compatible* if the coordinates $(x^1, x^2, \ldots, x^n)$ depend smoothly on $(y^1, y^2, \ldots, y^n)$ and vica versa on the overlap $U \cap V$ of the coordinate patches. Note in particular that two coordinate systems are smoothly compatible if the corresponding coordinate patches are disjoint.

A *smooth atlas* on M is a collection of continuous coordinate systems on M such that the following two conditions hold:—

(i) every point of M belongs to the coordinate patch of at least one of these coordinate systems,

(ii) the coordinate systems in the atlas are smoothly compatible with one another.

Let $\mathcal{A}$ be a smooth atlas on a topological manifold M of dimension n. Let $(u^1, u^2, \ldots, u^n)$ and $(v^1, v^2, \ldots, v^n)$ be continuous coordinate systems, defined over coordinate patches U and V respectively. If the coordinate systems (u^i) and (v^i) are smoothly compatible with all the coordinate systems in the atlas $\mathcal{A}$ then they are smoothly compatible with each other. Indeed suppose that $U \cap V \neq 0$, and let m be a point of $U \cap V$. Then (by condition (i) above) there exists a coordinate system (x^i) belonging to the atlas $\mathcal{A}$ whose coordinate patch includes that point m. But the coordinates (v^i) depend smoothly on the coordinates (x^i), and the coordinates (x^i) depend smoothly on the coordinates (u^i) around m (since the coordinate systems (u^i) and (v^i) are smoothly compatible with all coordinate systems in the atlas $\mathcal{A}$). It follows from the Chain Rule that the coordinates (v^i) depend smoothly on the coordinates (u^i) around m, and similarly the coordinates (u^i) depend smoothly on the coordinates (v^i). Therefore the continuous coordinate systems (u^i) and (v^i) are smoothly compatible with each other.

We deduce that, given a smooth atlas $\mathcal{A}$ on a topological manifold M, we can enlarge $\mathcal{A}$ by adding to to $\mathcal{A}$ all continuous coordinate systems on M that are smoothly compatible with each of the coordinate systems of $\mathcal{A}$. In this way we obtain a smooth atlas on M which is *maximal* in the sense that any coordinate system smoothly compatible with all the coordinate systems in the atlas already belongs to the atlas.

Definition A *smooth manifold* $(M, \mathcal{A})$ consists of a topological manifold M together with a maximal smooth atlas $\mathcal{A}$ of coordinate systems on M. A *smooth coordinate system* $(x^1, x^2, \ldots, x^n)$ on M is a coordinate system belonging to the maximal smooth atlas $\mathcal{A}$.

Note that $\mathbb{R}^n$ is a smooth manifold of dimension n. The maximal smooth atlas on $\mathbb{R}^n$ consists of all (curvilinear) coordinate systems that are smoothly compatible with the standard Cartesian coordinate system on $\mathbb{R}^n$.

1.2 Submanifolds

Let M be a subset of a k-dimensional smooth manifold N. We say that M is a smooth *embedded submanifold* of N of dimension n if, given any point m of M, there exists a smooth coordinate system $(u^1, u^2, \ldots, u^k)$ defined over some open set U in N, where $m \in U$, with the property that

$$M \cap U = \{p \in U : u^i(p) = 0 \text{ for } i = n+1, \ldots, k\}.$$

Given such a coordinate system $(u^1, u^2, \ldots, u^k)$, the restrictions of the coordinate functions $u^1, u^2, \ldots, u^n$ to $U \cap M$ provide a coordinate system on M around the point m. The collection of all such coordinate systems constitutes a smooth atlas on M. Thus any smooth embedded submanifold M of a smooth manifold N is itself a smooth manifold (with respect to the unique maximal smooth atlas containing the smooth atlas on M just described).

Example Consider the unit sphere S^n in $\mathbb{R}^{n+1}$ consisting of those vectors $\mathbf{x}$ in $\mathbb{R}^{n+1}$ satisfying $|\mathbf{x}| = 1$. Given any integer i between 1 and $n+1$, let

$$u^j(\mathbf{x}) = \begin{cases} x^j & \text{if } j < i; \\ x^{j+1} & \text{if } i \le j \le n; \\ (x^1)^2 + (x^2)^2 + \cdots + (x^{n+1})^2 - 1 & \text{if } j = n+1. \end{cases}$$

Then $(u^1, u^2, \ldots, u^{n+1})$ is a smooth coordinate system on U_i^+ and on U_i^-, where

$$U_i^+ = \{\mathbf{x} \in \mathbb{R}^{n+1} : x^i > 0\}, \qquad U_i^- = \{\mathbf{x} \in \mathbb{R}^{n+1} : x^i < 0\},$$

and

$$S^n \cap U_i^{\pm} = \{\mathbf{x} \in U_i^{\pm} : u^{n+1}(\mathbf{x}) = 0\}.$$

Moreover S^n is covered by $U_1^{\pm}, U_2^{\pm}, \ldots, U_{n+1}^{\pm}$. This shows that S^n is a smooth embedded submanifold of $\mathbb{R}^{n+1}$.

1.3 Smooth Mappings between Smooth Manifolds

Let M and N be smooth manifolds of dimension n and k respectively. A mapping $\varphi\colon M \to N$ from M to N is said to be *smooth* around a point m of M if, given smooth coordinate systems $(x^1, x^2, \ldots, x^n)$ and $(y^1, y^2, \ldots, y^k)$ around m and $\varphi(m)$, the coordinates $(y^1(\varphi(u)), y^2(\varphi(u)), \ldots, y^k(\varphi(u))$ of $\varphi(u)$ depend smoothly on the coordinates $(x^1(u), x^2(u), \ldots, x^n(u))$ of u for all points u belonging to some sufficiently small neighbourhood of m. (Note that if there exist smooth coordinate systems $(x^1, x^2, \ldots, x^n)$ and $(y^1, y^2, \ldots, y^k)$ around m and $\varphi(m)$ for which this condition is satisfied, then the condition is satisfied for all such smooth coordinate systems around m and $\varphi(m)$; this follows easily from the fact that a composition of smooth functions is smooth.) The mapping $\varphi\colon M \to N$ is said to be *smooth* if it is smooth around every point of M.

1.4 Bump Functions and Partitions of Unity

Let $f\colon X \to \mathbb{R}$ be a real-valued function defined defined over a topological space X. The *support* $\operatorname{supp} f$ of f is defined to be the closure of the set $\{x \in X : f(x) \neq 0\}$. Thus $\operatorname{supp} f$ is the smallest closed set in X with the property that the function f vanishes on the complement of that set.

Lemma 1.1 *Let U be an open set in a smooth manifold M of dimension n, and let m be a point of M. Then there exists an open subset V of U containing the point m and a smooth non-negative function $f\colon M \to \mathbb{R}$ such that* $\operatorname{supp} f \subset U$ *and $f(v) = 1$ for all $v \in V$.*

Proof We may assume, without loss of generality, that U is contained in the coordinate patch of some smooth coordinate system $(x^1, x^2, \ldots, x^n)$ and that $x^i(m) = 0$ for $i = 1, 2, \ldots, n$. Thus it suffices to show that, given any $r > 0$, there exists a smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$, taking values in the interval $[0,1]$, such that $f(\mathbf{x}) = 1$ if $|\mathbf{x}| \leq r$ and $f(\mathbf{x}) = 0$ if $|\mathbf{x}| \geq 2r$. Let

$$g(t) = \begin{cases} \exp\Bigl(\dfrac{-1}{(t-1)(4-t)}\Bigr) & \text{if } 1 < t < 4, \\ 0 & \text{if } t \leq 1 \text{ or } t \geq 4. \end{cases}$$

Then $g\colon \mathbb{R} \to \mathbb{R}$ is a smooth non-negative function. Indeed $g(t) = h(3t - 3)h(12 - 3t)$ for all t, where $h\colon \mathbb{R} \to \mathbb{R}$ is the smooth function defined by

$$h(t) = \begin{cases} \exp(-1/t) & \text{if } t > 0; \\ 0 & \text{if } t \leq 0. \end{cases}$$

Thus if we define

$$f(\mathbf{x}) = C \int_{|\mathbf{x}|^2/r^2}^{+\infty} g(t)\,dt,$$

where $1/C = \int_0^{+\infty} g(t)\,dt$, then the function $f\colon \mathbb{R}^n \to \mathbb{R}$ has the required properties. ∎

We now prove an existence theorem for finite 'partitions of unity' on a compact smooth manifold.

Theorem 1.2 *Let M be a compact smooth manifold, and let $\mathcal{V}$ be an open cover of M. Then there exist smooth non-negative functions $f_1, f_2, \ldots, f_k$ on M with the following properties:—*

(i) $f_1 + f_2 + \cdots + f_k = 1$,

(ii) *for each function f_i there exists an open set V belonging to $\mathcal{V}$ such that* $\operatorname{supp} f_i \subset V$.

Proof For each point m of M there exists a smooth non-negative function $g_m\colon M \to \mathbb{R}$ such that $g_m(m) = 1$ and $\operatorname{supp} g_m \subset V$ for at least one open set V belonging to $\mathcal{V}$ (Lemma 1.1). For each $m \in M$, let

$$W_m = \{x \in M : g_m(x) > \frac{1}{2}\}.$$

Then $\{W_m : m \in M\}$ is an open cover of M. It follows from the compactness of M that there exists a finite collection $m_1, m_2, \ldots, m_k$ of points of M such that

$$M = W_{m_1} \cup W_{m_2} \cup \cdots \cup W_{m_k}.$$

Set $f_i(x) = g_{m_i}(x)/G(x)$ for all $x \in M$, where

$$G(x) = g_{m_1}(x) + g_{m_2}(x) + \cdots g_{m_k}(x).$$

Then $f_1, f_2, \ldots, f_k$ is a collection of smooth functions on M with the required properties. ∎

A collection $f_1, f_2, \ldots, f_k$ of functions with the properties stated in Theorem 1.2 is referred to as a finite *partition of unity* subordinate to the open cover $\mathcal{V}$ of the manifold M.

Suppose that M is a (not necessarily compact) smooth manifold. A *locally finite partition of unity* on M is a collection of smooth non-negative functions such that

- each point of M has an open neighbourhood on which at most finitely many of the functions are non-zero,
- the sum of the values of the functions at any point of M equals 1.

Using the metrizability of M, it can be shown that there exists a locally finite partition of unity subordinate to any open cover of M.

2 Tangent Spaces

Let M be a smooth manifold of dimension n, and let m be a point of M. We say that a real-valued function f is defined *around* m if f is defined throughout some open neighbourhood of m. A *tangent vector* X_m at the point m can be regarded as an operator, associating a real number $X_m[f]$ to any smooth real-valued function f defined around m, where

(i) $X_m[\alpha f + \beta g] = \alpha X_m[f] + \beta X_m[g]$ for all real numbers α and β and smooth functions f and g defined around m,

(ii) $X_m[f.g] = X_m[f]\, g(m) + f(m)\, X_m[g]$ for all smooth functions f and g defined around m,

(iii) if f and g are smooth real-valued functions defined around m and if $f = g$ on some open set V containing the point m then $X_m[f] = X_m[g]$.

(Here $f.g$ denotes the product of the functions f and g, defined by $(f.g)(m) = f(m)g(m)$ for all $m \in M$.) The quantity $X_m[f]$ is referred to as the *directional derivative* of the function f along the vector X_m.

If X_m and Y_m are tangent vectors at the point m then, for any real numbers α and β, $\alpha X_m + \beta Y_m$ is also a tangent vector at the point m, where $(\alpha X_m + \beta Y_m)[f] = \alpha X_m[f] + \beta Y_m[f]$ for all smooth real-valued functions f defined around m. It follows that the collection of all tangent vectors at the point m is a vector space $T_m M$, referred to as the *tangent space* to M at the point m.

Lemma 2.1 *Let M be a smooth manifold, and let X_m be a tangent vector at some point m of M. let c_λ denote the constant function on M with value λ. Then $X_m[c_\lambda] = 0$ for all $\lambda \in \mathbb{R}$.*

Proof It follows from condition (ii) above that

$$X_m[c_1] = X_m[c_1.c_1] = 2c_1(m)X_m[c_1] = 2X_m[c_1],$$

where c_1 is the constant function with value 1. Therefore $X_m[c_1] = 0$, and hence $X_m[c_\lambda] = \lambda X_m[c_1] = 0$ for all $\lambda \in \mathbb{R}$. ∎

Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system defined over an open set U in M. Then the mapping $\varphi: U \to \mathbb{R}^n$ which sends a point u of U to $(x^1(u), x^2(u), \ldots, x^n(u))$ maps the coordinate patch U homeomorphically onto an open set $\varphi(U)$ in $\mathbb{R}^n$. The inverse φ^{-1} of φ is thus a well-defined smooth map on the image $\varphi(U)$ of U.

Let m be a point of the coordinate patch U. Given any smooth real-valued function f defined around m, we denote by

$$\left.\frac{\partial f}{\partial x^i}\right|_m$$

the ith partial derivative of the function f with respect to the coordinate system $(x^1, x^2, \ldots, x^n)$, defined by

$$\left.\frac{\partial f}{\partial x^i}\right|_m = \left.\frac{\partial (f \circ \varphi^{-1})}{\partial t^i}\right|_{(t^1,\ldots,t^n)=\varphi(m)},$$

where $(t^1, t^2, \ldots, t^n)$ is the standard Cartesian coordinate system on $\mathbb{R}^n$.

Given any real numbers $a^1, a^2, \ldots, a^n$, the operator sending any smooth real-valued function f defined around the point m to

$$a^1 \left.\frac{\partial f}{\partial x^1}\right|_m + a^2 \left.\frac{\partial f}{\partial x^2}\right|_m + \cdots + a^n \left.\frac{\partial f}{\partial x^n}\right|_m$$

satisfies conditions (i)–(iii) and therefore represents a tangent vector at m which we denote by

$$a^1 \left.\frac{\partial}{\partial x^1}\right|_m + a^2 \left.\frac{\partial}{\partial x^2}\right|_m + \cdots + a^n \left.\frac{\partial}{\partial x^n}\right|_m.$$

Conversely, we shall show that any tangent vector at m is of this form for suitable real numbers $a^1, \ldots, a^n$. The following lemma is is the basic result needed to prove this fact.

Lemma 2.2 *Let M be a smooth manifold of dimension n and let m be a point of M. Let f be a smooth function defined over some neighbourhood of the point m. Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system defined around the point m. Then there exist smooth functions $g_1, g_2, \ldots g_n$, defined over some suitable open set U containing the point m, such that*

$$f(u) = f(m) + \sum_{i=1}^{n} \left(x^i(u) - x^i(m)\right) g_i(u)$$

for all $u \in U$. Moreover

$$g_i(m) = \left.\frac{\partial f}{\partial x^i}\right|_m \qquad \text{for } i = 1, 2, \ldots, n.$$

Proof Without loss of generality, we may assume that f is a real-valued function defined over some open ball B about the origin in $\mathbb{R}^n$. We must show that there exist smooth real-valued functions $g_1, g_2, \ldots, g_n$ on B such that

$$f(\mathbf{x}) = f(\mathbf{0}) + x^1 g_1(\mathbf{x}) + x^2 g_2(\mathbf{x}) + \cdots + x^n g_n(\mathbf{x})$$

for all $\mathbf{x} \in B$. Now

$$f(\mathbf{x}) - f(\mathbf{0}) = \int_0^1 \frac{d}{dt}\left(f(t\mathbf{x})\right)\, dt = \sum_{i=1}^{n} x^i \int_0^1 (\partial_i f)(t\mathbf{x})\, dt,$$

Let

$$g_i(\mathbf{x}) = \int_0^1 (\partial_i f)(t\mathbf{x})\, dt$$

for $i = 1, 2, \ldots, n$. Then $g_1, g_2, \ldots, g_n$ satisfy the required conditions. ∎

Proposition 2.3 *Let M be a smooth manifold of dimension n, and let X_m be a tangent vector at some point m of M. Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system around the point m. Then*

$$X_m = a^1 \left.\frac{\partial}{\partial x^1}\right|_m + a^2 \left.\frac{\partial}{\partial x^2}\right|_m + \cdots + a^n \left.\frac{\partial}{\partial x^n}\right|_m.$$

where $a^i = X_m[x^i]$. If $(y^1, y^2, \ldots, y^n)$ is another smooth coordinate system around m then

$$X_m = b^1 \left.\frac{\partial}{\partial y^1}\right|_m + b^2 \left.\frac{\partial}{\partial y^2}\right|_m + \cdots + b^n \left.\frac{\partial}{\partial y^n}\right|_m,$$

where

$$b^j = \sum_{i=1}^{n} a^i \left.\frac{\partial y^j}{\partial x^i}\right|_m \qquad (j = 1, 2, \ldots, n).$$

Proof Let f be a smooth real-valued function defined around m. It follows from Lemma 2.2 that there exist smooth functions $g_1, g_2, \ldots, g_n$ defined around m such that

$$f(u) = f(m) + \sum_{i=1}^{n} \left(x^i(u) - x^i(m)\right) g_i(u)$$

for all points u belonging to some sufficiently small open set containing m. Moreover

$$g_i(m) = \left.\frac{\partial f}{\partial x^i}\right|_m \qquad \text{for } i = 1, 2, \ldots, n.$$

Let $p^i(u) = x^i(u) - x^i(m)$. Now the operator X_m annihilates constant functions, by Lemma 2.1. Therefore $X_m[p^i] = X_m[x^i] = a^i$ for all i, and hence

$$X_m[f] = \sum_{i=1}^{n} \left(X_m[p^i]\, g_i(m) + p^i(m)\, X_m[g^i]\right) = \sum_{i=1}^{n} a^i \left.\frac{\partial f}{\partial x^i}\right|_m .$$

If $(y^1, y^2, \ldots, y^n)$ is another smooth coordinate system around m then

$$\frac{\partial f}{\partial x^i} = \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i} \frac{\partial f}{\partial y^j},$$

by the Chain Rule, and hence

$$X_m = b^1 \left.\frac{\partial}{\partial y^1}\right|_m + b^2 \left.\frac{\partial}{\partial y^2}\right|_m + \cdots + b^n \left.\frac{\partial}{\partial y^n}\right|_m,$$

where

$$b^j = \sum_{i=1}^{n} a^i \left.\frac{\partial y^j}{\partial x^i}\right|_m \qquad (j = 1, 2, \ldots, n),$$

as required. ∎

Corollary 2.4 *Let M be a smooth manifold of dimension n. Then the tangent space $T_m M$ to M at any point m of M has dimension n. Moreover, given any smooth coordinate system $(x^1, x^2, \ldots, x^n)$ around m, the tangent vectors*

$$\left.\frac{\partial}{\partial x^1}\right|_m, \left.\frac{\partial}{\partial x^2}\right|_m, \ldots, \left.\frac{\partial}{\partial x^n}\right|_m$$

constitute a basis for the tangent space $T_m M$.

Proof It follows immediately from Proposition 2.3 that these tangent vectors span the tangent space $T_m M$. It thus suffices to show that they are linearly independent. Suppose that

$$\sum_{i=1}^{n} a^i \left.\frac{\partial}{\partial x^i}\right|_m = 0$$

for some real numbers $a^1, a^2, \ldots, a^n$. Then

$$a^j = \left(\sum_{i=1}^{n} a^i \left.\frac{\partial}{\partial x^i}\right|_m \right) (x^j) = 0$$

for $j = 1, 2, \ldots, n$, showing that the tangent vectors $\partial/\partial x^i$ $(i = 1, 2, \ldots, n)$ are linearly independent at m, as required. ∎

Let $\gamma: I \to M$ be a smooth curve in the smooth manifold M, where I is some interval in $\mathbb{R}$. Then γ determines, for each $t \in I$, a tangent vector $\gamma'(t)$ at the point $\gamma(t)$, defined by

$$\gamma'(t)[f] = \frac{df(\gamma(t))}{dt}.$$

We refer to the tangent vector $\gamma'(t)$ as the *velocity vector* of the curve γ at $\gamma(t)$.

Every tangent vector at a point m of the smooth manifold M is the velocity vector of some smooth curve passing through the point m. Indeed let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system around the point m chosen such that $x^i(m) = 0$ for $i = 1, 2, \ldots, n$. Let X_m be a tangent vector at the point m. Then

$$X_m = a^1 \left.\frac{\partial}{\partial x^1}\right|_m + a^2 \left.\frac{\partial}{\partial x^2}\right|_m + \cdots + a^n \left.\frac{\partial}{\partial x^n}\right|_m$$

bu Proposition 2.3, where $a^i = X_m[x^i]$ for $i = 1, 2, \ldots, n$. Let $\gamma: (-\varepsilon, \varepsilon) \to M$ be the smooth curve in M given by $x^i(\gamma(t)) = a^i t$ for $i = 1, 2, \ldots, n$ (where ε is some suitably small positive real number). It follows from the Chain Rule that

$$\gamma'(0)[f] = \left.\frac{df(\gamma(t))}{dt}\right|_{t=0} = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x^i}\right|_m \left.\frac{d(x^i(\gamma(t)))}{dt}\right|_{t=0} = \sum_{i=1}^{n} a^i \left.\frac{\partial f}{\partial x^i}\right|_m = X_m[f]$$

for all smooth real-valued functions f defined around m, so that $\gamma'(0) = X_m$.

2.1 Derivatives of Smooth Maps

Let $\varphi\colon M \to N$ be a smooth map between smooth manifolds of dimensions n and k respectively. Let X_m be a tangent vector at some point m of M. Given any smooth function f, defined around $\varphi(m)$ in N, we define $(\varphi_* X_m)[f] = X_m[f \circ \varphi]$.

Lemma 2.5 *The operator $\varphi_* X_m$ is a tangent vector at $\varphi(m)$.*

Proof Let f and g be smooth real-valued functions defined around $\varphi(m)$, and let α and β be real numbers. Then

$$\begin{aligned}
(\varphi_* X_m)[\alpha f + \beta g] &= X_m[\alpha(f \circ \varphi) + \beta(g \circ \varphi)] = \alpha X_m[f \circ \varphi] + \beta X_m[g \circ \varphi] \\
&= \alpha(\varphi_* X_m)[f] + \beta(\varphi_* X_m)[g], \\
(\varphi_* X_m)[f.g] &= X_m[f \circ \varphi] g(\varphi(m)) + f(\varphi(m)) X_m[g \circ \varphi] \\
&= (\varphi_* X_m[f] g(\varphi(m)) + f(\varphi(m))(\varphi_* X_m)[g]
\end{aligned}$$

Moreover if the functions f and g agree on some open set containing $\varphi(m)$ then the functions $f \circ \varphi$ and $g \circ \varphi$ agree on some open set containing m (since $\varphi\colon M \to N$ is continuous), and therefore $(\varphi_* X_m)[f] = (\varphi_* X_m)[g]$. Thus the operator $\varphi_* X_m$ is a tangent vector at $\varphi(m)$. ∎

Given a smooth map $\varphi\colon M \to N$ between smooth manifolds M and N, the linear transformation $\varphi_*\colon T_m M \to T_{\varphi(m)} N$ between the tangent spaces at m and $\varphi(m)$ is referred to as the *derivative* of the smooth map φ at the point m.

Let $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^k)$ be smooth coordinate systems on M and N respectively around the points m and $\varphi(m)$. Then the map φ is determined around m by smooth functions $F^1, F^2, \ldots, F^k$, where

$$F^j(x^1, x^2, \ldots, x^n) = y^j \circ \varphi \qquad (j = 1, 2, \ldots, k).$$

A straightforward application of the Chain Rule for functions of real variables shows that

$$\varphi_* \left(\sum_{i=1}^n a^i \left. \frac{\partial}{\partial x^i} \right|_m \right) = \sum_{j=1}^k \left(\sum_{i=1}^n a^i \left. \frac{\partial F^j}{\partial x^i} \right|_m \right) \left. \frac{\partial}{\partial y^j} \right|_{\varphi(m)}.$$

2.2 Vector Fields on Smooth Manifolds

Given any open set U in a smooth manifold M we denote by $C^\infty(U)$ the algebra of all smooth real-valued functions on U.

Definition Let U be an open set in a smooth manifold M. A smooth *vector field* X on U is an operator $X\colon C^\infty(U) \to C^\infty(U)$, sending $f \in C^\infty(U)$ to $X[f]$, where

$$X[f + h] = X[f] + X[h], \quad X[\lambda f] = \lambda X[f] \text{ and } X[f.h] = X[f].h + f.X[h]$$

for all $f, h \in C^\infty(U)$ and $\lambda \in \mathbb{R}$.

If X is a smooth vector field on U then $X[c] = 0$ for all constant functions c on U.

Lemma 2.6 *Let X be a smooth vector field on an open set U in a smooth manifold M, and let V be an open subset of U. Let f and h be smooth real-valued functions on U. Suppose that $f = h$ on V. Then $X[f] = X[h]$ on V.*

Proof Let m be a point of V. Then there exists a smooth bump function $\beta\colon U \to \mathbb{R}$ such that $\beta(m) = 1$ and $\beta(u) = 0$ for all $u \in U \setminus V$ (see Lemma 1.1). Then $\beta.(f - h) = 0$, hence

$$0 = X[\beta.(f - h)] = X[\beta].(f - h) + \beta.\left(X[f] - X[h]\right)$$

But $f(m) - h(m) = 0$ and $\beta(m) = 1$. Therefore $X[f](m) = X[h](m)$, as required. ∎

Let f be a smooth real-valued function defined around the point m. Then there exists a smooth function $\tilde{f}\colon M \to \mathbb{R}$ on M such that $f = \tilde{f}$ throughout some open set V containing the point m. Indeed we can take

$$\tilde{f}(p) = \begin{cases} \beta(p) f(p) & \text{if } p \in \operatorname{supp} \beta, \\ 0 & \text{if } p \notin \operatorname{supp} \beta, \end{cases}$$

where $\beta\colon M \to \mathbb{R}$ is a smooth bump function with the following properties:

- $\beta(v) = 1$ for all points v of some open set V containing the point m,
- the support $\operatorname{supp} \beta$ of β is contained in the domain of the function f.

Given a smooth vector field X defined over some open set U in M, and let m be a point of U. Given any smooth real-valued function f defined around m, we define $X_m[f] = X[\tilde{f}](m)$, where $\tilde{f}\colon M \to \mathbb{R}$ is some smooth function on M with the property that $\tilde{f} = f$ over some open set V containing the point m. It follows from Lemma 2.6 that the value of $X_m[f]$ does not depend on the choice of $\tilde{f}$. Moreover the operator X_m sending any smooth real-valued function f defined around m to the real number $X_m[f]$ is a tangent vector at the point m. Moreover $X_m[f] = X[f](m)$ for all $f \in C^\infty(U)$. We refer to X_m as the *value* of the vector field X at the point m.

Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system defined over some open set U in M. Then the operator $\partial/\partial x^i$ sending any smooth function $f\colon U \to \mathbb{R}$ to $\partial f/\partial x^i$ is a smooth vector field on U for $i = 1, 2, \ldots, n$. Thus, given any smooth real-valued functions $a^1, a^2, \ldots, a^m$ on U, the operator

$$a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + \cdots + a^n \frac{\partial}{\partial x^n}$$

is a smooth vector field on U. The following result now follows directly using Proposition 2.3.

Proposition 2.7 *Let M be a smooth manifold of dimension n, and let X be a smooth vector field defined over the domain U of some smooth coordinate system $(x^1, x^2, \ldots, x^n)$ on M. Then*

$$X = a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + \cdots + a^n \frac{\partial}{\partial x^n}$$

where $a^1, a^2, \ldots, a^n$ are the smooth functions on U given by $a^i = X[x^i]$, for $i = 1, 2, \ldots, n$.

2.3 The Lie Bracket

Let M be a smooth manifold and let X and Y be smooth vector fields on some open set U in M. We denote by $[X, Y]$ the operator on $C^\infty(U)$ defined by

$$[X, Y][f] = X[Y[f]] - Y[X[f]]$$

for all smooth real-valued functions f on U. Now

$$[X, Y][f + g] = [X, Y][f] + [X, Y][g], \qquad [X, Y][\lambda f] = \lambda[X, Y][f]$$

for all $f, g \in C^\infty(U)$ and $\lambda \in \mathbb{R}$. Also

$$\begin{aligned} X[Y[f.g]] &= X[(Y[f].g + f.Y[g])] \\ &= X[Y[f]].g + Y[f].X[g] + X[f].Y[g] + f.X[Y[g]], \end{aligned}$$

and hence

$$\begin{aligned} [X, Y][f.g] &= X[Y[f.g]] - Y[X[f.g]] \\ &= X[Y[f]].g - Y[X[f].g + f.X[Y[g]] - f.Y[X[g]] \\ &= [X, Y][f].g + f.[X, Y][g]. \end{aligned}$$

We deduce that the operator $[X,Y]$ on $C^\infty(U)$ is a smooth vector field on U. This vector field is referred to as the *Lie bracket* of the vector fields X and Y on U.

Note that if X is a smooth vector field on U, then so is fX for any $f \in C^\infty(U)$, where $fX[g] = f.X[g]$ for all $g \in C^\infty(U)$.

Lemma 2.8 *Let M be a smooth manifold. Let X and Y be smooth vector fields on M and let f and g be smooth real-valued functions on M. Then*

$$[fX, gY] = (f.g)[X,Y] + (f.X[g])Y - (g.Y[f])X.$$

Proof Let h be a smooth real-valued function defined over some open set in M. Then

$$\begin{aligned}[fX,gY][h] &= f.X[g.Y[h]] - g.Y[f.X[h]] \\ &= (f.g).X[Y[h]] + f.X[g].Y[h] - (f.g).Y[X[h]] - g.Y[f].X[h] \\ &= ((f.g)[X,Y] + (f.X[g])Y - (g.Y[f])X)\,[h],\end{aligned}$$

as required. ∎

Lemma 2.9 *Let M be a smooth manifold of dimension n and let X and Y be smooth vector fields on some open set U in M. Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system on U. Suppose that*

$$X = \sum_{i=1}^n u^i \frac{\partial}{\partial x^i}, \qquad Y = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}.$$

Then

$$[X,Y] = \sum_{i=1}^n \sum_{j=1}^n \left(u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j} \right) \frac{\partial}{\partial x^i}$$

Proof Let $f \in C^\infty(U)$. Then

$$\begin{aligned}[X,Y][f] &= X\left[\sum_{i=1}^n v^i \frac{\partial f}{\partial x^i}\right] - Y\left[\sum_{i=1}^n u^i \frac{\partial f}{\partial x^i}\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \left(u^j \frac{\partial}{\partial x^j}\left(v^i \frac{\partial f}{\partial x^i}\right) - v^j \frac{\partial}{\partial x^j}\left(u^i \frac{\partial f}{\partial x^i}\right)\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \left(u^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial u^i}{\partial x^j} \right) \frac{\partial f}{\partial x^i},\end{aligned}$$

since

$$\sum_{i=1}^n \sum_{j=1}^n (u^j v^i - v^j u^i) \frac{\partial^2 f}{\partial x^j \partial x^i} = \sum_{i=1}^n \sum_{j=1}^n u^j v^i \left(\frac{\partial^2 f}{\partial x^j \partial x^i} - \frac{\partial^2 f}{\partial x^i \partial x^j} \right) = 0.$$ ∎

Note in particular that, for any smooth coordinate system $(x^1, x^2, \ldots, x^n)$ on M,

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0 \qquad (i, j = 1, 2, \ldots, n).$$

Let M and N be smooth manifolds and let $\varphi\colon M \to N$ be a smooth map from M to N. Let X and $\tilde{X}$ be smooth vector fields on M and N respectively. We say that X and $\tilde{X}$ are *φ-related* if $\tilde{X}[g] \circ \varphi = X[g \circ \varphi]$ for all smooth real-valued functions g on N.

Lemma 2.10 *Let M and N be smooth manifolds and let $\varphi\colon M \to N$ be a smooth map from M to N. Let X and Y be smooth vector fields on M, and let $\tilde{X}$ and $\tilde{Y}$ be smooth vector fields on N. Suppose that the vector fields X and $\tilde{X}$ are φ-related and that the vector fields Y and $\tilde{Y}$ are φ-related. Then the vector fields $[X, Y]$ and $[\tilde{X}, \tilde{Y}]$ are also φ-related.*

Proof Let $g \in C^\infty(N)$. Then

$$\begin{aligned} [\tilde{X}, \tilde{Y}][g] \circ \varphi &= \tilde{X}[\tilde{Y}[g]] \circ \varphi - \tilde{Y}[\tilde{X}[g]] \circ \varphi = X[\tilde{Y}[g] \circ \varphi] - Y[\tilde{X}[g] \circ \varphi] \\ &= X[Y[g \circ \varphi]] - Y[X[g \circ \varphi]] = [X, Y][g \circ \varphi], \end{aligned}$$

as required. ∎

3 Affine Connections on Smooth Manifolds

Definition Let M be a smooth manifold. An *affine connection* on M is a differential operator, sending smooth vector fields X and Y to a smooth vector field $\nabla_X Y$, which satisfies the following conditions:

$$\begin{array}{ll} \nabla_{X+Y} Z = \nabla_X Z + \nabla_Y Z, & \nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z, \\ \nabla_{fX} Y = f\, \nabla_X Y, & \nabla_X (fY) = X[f]\, Y + f\, \nabla_X Y \end{array}$$

for all smooth vector fields X, Y and Z and real-valued functions f on M. The vector field $\nabla_X Y$ is known as the *covariant derivative* of the vector field Y along X (with respect to the affine connection ∇). The *torsion tensor* T and the *curvature tensor* R of an affine connection ∇ are the operators sending smooth vector fields X, Y and Z on M to the smooth vector fields $T(X, Y)$ and $R(X, Y)Z$ given by

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

An affine connection ∇ on M is said to be *torsion-free* if its torsion tensor is everywhere zero (so that $\nabla_X Y - \nabla_Y X = [X, Y]$ for all smooth vector fields X and Y on M).

Example Let U be an open set in $\mathbb{R}^n$, and let X and Y be smooth vector fields on U. Then

$$X = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}, \qquad Y = \sum_{i=1}^n b^i \frac{\partial}{\partial x^i},$$

where $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ are the components of the vector fields X and Y with respect to the Cartesian coordinate system $(x^1, x^2, \ldots, x^n)$ on $\mathbb{R}^n$. The *directional derivative* $\partial_X Y$ of the vector field Y along the vector field X is then given by the formula

$$\partial_X Y = \sum_{i=1}^n X[b^i] \frac{\partial}{\partial x^i} = \sum_{i,j=1}^n a^j \frac{\partial b^i}{\partial x^j} \frac{\partial}{\partial x^i}$$

(where $X[b^i]$ denotes the directional derivative of the function b^i along the vector field X). Then the differential operator sending smooth vector fields X and Y to $\partial_X Y$ is an affine connection on U. We refer to this affine connection as the *canonical* (or *usual*) *flat connection* on the open set U. Now

$$\partial_X Y - \partial_Y X = \sum_{i=1}^n \sum_{j=1}^n \left(a^j \frac{\partial b^i}{\partial x^j} - b^j \frac{\partial a^i}{\partial x^j} \right) \frac{\partial}{\partial x^i} = [X, Y]$$

(see Lemma 2.9). Thus the canonical flat connection ∂ on U is torsion-free. Moreover, given any smooth vector field Z on U with Cartesian compoents $c^1, c^2, \ldots, c^n$, we see that

$$\begin{aligned} \partial_X \partial_Y Z - \partial_Y \partial_X Z &= (\partial_X \partial_Y - \partial_Y \partial_X) \left(\sum_{i=1}^n c^i \frac{\partial}{\partial x^i} \right) \\ &= \sum_{i=1}^n \left(X[Y[c^i]] - Y[X[c^i]] \right) \frac{\partial}{\partial x^i} = \sum_{i=1}^n [X, Y][c^i] \frac{\partial}{\partial x^i} \\ &= \partial_{[X,Y]} Z. \end{aligned}$$

We deduce that the curvature tensor of the canonical flat connection ∂ on U is zero everywhere on U.

Example Let M be a smooth n-dimensional submanifold of $\mathbb{R}^k$. This means that M is a subset of $\mathbb{R}^k$ with the property that, for all $m \in M$, there exists a smooth (curvilinear) coordinate system $(u^1, u^2, \ldots, u^k)$ defined over a neighbourhood U of m which has the property that

$$M \cap U = \{\mathbf{p} \in U : u^j(\mathbf{p}) = 0 \text{ for } j = n+1, \ldots, k\}.$$

Let X and Y be smooth vector fields on M that are everywhere tangent to M, and let $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ be the components of X and Y with respect to the curvilinear coordinate system $(u^1, u^2, \ldots, u^n)$ on M. Thus

$$X = \sum_{j=1}^{n} a^j \frac{\partial}{\partial u^j}, \qquad Y = \sum_{j=1}^{n} b^j \frac{\partial}{\partial u^j} = \sum_{i=1}^{k} c^i \frac{\partial}{\partial x^i},$$

where $(x^1, x^2, \ldots, x^k)$ is the standard Cartesian coordinate system of $\mathbb{R}^k$, and

$$c^i = \sum_{j=1}^{n} b^j \frac{\partial x^i}{\partial u^j} \qquad (i = 1, 2, \ldots, k).$$

The directional derivative $\partial_X Y$ of the vector field Y along X at each point of M is defined to be the vector in $\mathbb{R}^k$ whose Cartesian components are the directional derivatives $X[c^1], X[c^2], \ldots, X[c^k]$ of the Cartesian components $c^1, c^2, \ldots, c^k$ of Y along the vector field X. Thus

$$\begin{aligned} \partial_X Y &= \sum_{i=1}^{k} X[c^i] \frac{\partial}{\partial x^i} = \sum_{i=1}^{k} \sum_{j,l=1}^{n} a^j \frac{\partial}{\partial u^j} \left(b^l \frac{\partial x^i}{\partial u^l} \right) \frac{\partial}{\partial x^i} \\ &= \sum_{i=1}^{k} \sum_{j,l=1}^{n} \left(a^j \frac{\partial b^l}{\partial u^j} \frac{\partial x^i}{\partial u^l} + a^j b^l \frac{\partial^2 x^i}{\partial u^j \partial u^l} \right) \frac{\partial}{\partial x^i}. \end{aligned}$$

Now $\partial_X Y$ is not necessarily tangential to to the submanifold M at each point of M. However $\partial_X Y - \partial_Y X$ is always tangential to M, and $\partial_X Y - \partial_Y X = [X, Y]$. Indeed

$$\begin{aligned} \partial_X Y - \partial_Y X &= \sum_{i=1}^{k} \sum_{j,l=1}^{n} \left(a^j \frac{\partial b^l}{\partial u^j} - b^j \frac{\partial a^l}{\partial u^j} \right) \frac{\partial x^i}{\partial u^l} \frac{\partial}{\partial x^i} \\ &= \sum_{j,l=1}^{n} \left(a^j \frac{\partial b^l}{\partial u^j} - b^j \frac{\partial a^l}{\partial u^j} \right) \frac{\partial}{\partial u^l} \\ &= [X, Y]. \end{aligned}$$

We split $\partial_X Y$ into its tangential and normal components, writing

$$\partial_X Y = \nabla_X Y + S(X, Y),$$

where $\nabla_X Y$ is everywhere tangential to M and $S(X, Y)$ is everywhere normal to M (i.e., at each point m of M, the vector $S(X, Y)$ is orthogonal to the tangent space $T_m M$ to M at the point m). Now

$$\partial_{X_1 + X_2} Y = \partial_{X_1} Y + \partial_{X_2} Y, \qquad \partial_X (Y_1 + Y_2) = \partial_X Y_1 + \partial_X Y_2$$

$$\partial_{fX}Y = f\,\partial_X Y, \qquad \partial_X(fY) = f\,\partial_X Y + X[f]\,Y$$

for all smooth real-valued functions f and (tangential) vector fields X, X_1, X_2, Y, Y_1, Y_2 on M. On splitting these identities into their tangential and normal components, we deduce that the differential operator ∇ is an affine connection on M, and

$$\begin{aligned} S(X_1 + X_2, Y) &= S(X_1, Y) + S(X_2, Y), \\ S(X, Y_1 + Y_2) &= S(X, Y_1) + S(X, Y_2), \end{aligned}$$

$$S(fX, Y) = fS(X, Y) = S(X, fY).$$

Moreover

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

since $\partial_X Y - \partial_Y X = [X, Y]$. Thus the affine connection ∇ is torsion-free.

Lemma 3.1 *Let ∇ be an affine connection on a smooth manifold M. Then the value of the covariant derivative $\nabla_X Y$ at a point m of M depends only on the vector field Y and on the value X_m of the vector field X at the point m.*

Proof Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system defined around m. Then

$$X = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}, \qquad Y = \sum_{i=1}^n b^i \frac{\partial}{\partial x^i},$$

where $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ are smooth real-valued functions defined around m. Now

$$\nabla_{\frac{\partial}{\partial x^j}} \left(\frac{\partial}{\partial x^k} \right) = \sum_{i=1}^n \Gamma^i_{jk} \frac{\partial}{\partial x^i},$$

where Γ^i_{jk} are smooth functions defined over the domain of the coordinate system $(x^1, x^2, \ldots, x^n)$. It follows from the definition of an affine connection that

$$\begin{aligned} \nabla_X Y &= \sum_{k=1}^n \left(X[b^k] \frac{\partial}{\partial x^k} + b^k \nabla_X \left(\frac{\partial}{\partial x^k} \right) \right) \\ &= \sum_{i,j=1}^n \left(a^j \frac{\partial b^i}{\partial x^j} + \sum_{k=1}^n a^j b^k \Gamma^i_{jk} \right) \frac{\partial}{\partial x^i}. \end{aligned}$$

Therefore the value of $\nabla_X Y$ at the point m depends only on the functions $b^1, \ldots, b^n$ and their partial derivatives at m and on the values of the functions $a^1, a^2, \ldots, a^n$ at m. We deduce that the value of $\nabla_X Y$ at m depends only on the vector field Y and on the value X_m of X at m. ∎

Let ∇ be an affine connection on a smooth manifold M, let Y be a smooth vector field on M, and let X_m be a tangent vector at some point m of M. We define $\nabla_{X_m} Y$ to be the value of $\nabla_X Y$ at m, where X is any smooth vector field on M whose value at m is X_m. Lemma 3.1 shows that $\nabla_{X_m} Y$ is well-defined. We refer to $\nabla_{X_m} Y$ as the *covariant derivative* of the vector field Y along the tangent vector X_m at the point m.

Lemma 3.2 *Let ∇ be an affine connection on a smooth manifold M and let T be the torsion tensor of the affine connection ∇. Then*

$$T(X, Y) = -T(Y, X)$$

for all vector fields X and Y on M. Also

$$\begin{array}{ll} T(X + Y, Z) = T(X, Z) + T(Y, Z), & T(X, Y + Z) = T(X, Y) + T(X, Z), \\ T(fX, Y) = f\,T(X, Y), & T(X, fY) = f\,T(X, Y) \end{array}$$

for all smooth vector fields X, Y and Z on M and for all smooth real-valued functions f on M. Moreover the value of $T(X, Y)$ at any point m of M depends only on the values X_m and Y_m of X and Y at m.

Proof The identities

$$T(X, Y) = -T(Y, X), \qquad T(X + Y, Z) = T(X, Z) + T(Y, Z),$$

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

follow immediately from the definition of T. Now if f is a smooth real-valued function on M then $[X, fY] = f\,[X, Y] + X[f]\,Y$ by Lemma 2.8. Thus

$$\begin{aligned} T(X, fY) &= \nabla_X (fY) - \nabla_{fY} X - [X, fY] \\ &= f\,\nabla_X Y + X[f]\,Y - f\,\nabla_Y X - f\,[X, Y] - X[f]\,Y \\ &= fT(X, Y). \end{aligned}$$

Also

$$T(fX, Y) = -T(Y, fX) = -f\,T(Y, X) = f\,T(X, Y).$$

Let m be a point of M and let $(x^1, x^2, \ldots, x^n)$ be a coordinate system defined around m. Then

$$X = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}, \qquad Y = \sum_{i=1}^n b^i \frac{\partial}{\partial x^i},$$

where $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ are smooth real-valued functions defined around m. Let $T^i{}_{jk}$ be the real-valued functions on the domain of the coordinate system characterized by the property that

$$T\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = \sum_{i=1}^{n} T^i{}_{jk} \frac{\partial}{\partial x^i}.$$

It follows from the identities derived above that

$$T(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(T^i{}_{jk} a^j b^k \frac{\partial}{\partial x^i} \right).$$

This shows that the value of $T(X, Y)$ at m depends only on the values of the functions $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ at m, and thus depends only on the values X_m and Y_m of the vector fields X and Y at m. ∎

Lemma 3.3 *Let ∇ be an affine connection on a smooth manifold M with curvature tensor R. Then*

$$\begin{aligned}
R(X,Y)Z &= -R(Y,X)Z, \\
R(X_1 + X_2, Y)Z &= R(X_1, Y)Z + R(X_2, Y)Z, \\
R(X, Y_1 + Y_2)Z &= R(X, Y_1)Z + R(X, Y_2)Z, \\
R(X,Y)(Z_1 + Z_2) &= R(X,Y)Z_1 + R(X,Y)Z_2, \\
R(fX, Y)Z &= R(X, fY)Z = R(X,Y)(fZ) = f\, R(X,Y)Z,
\end{aligned}$$

for all vector fields X, X_1, X_2, Y, Y_1, Y_2, Z, Z_1 and Z_2 and smooth functions f on M. Moreover the value of $R(X,Y)Z$ at any point m of M depends only on the values X_m, Y_m and Z_m of X, Y and Z at m.

Proof The first four identities follow directly from the definition of the curvature tensor. Let X, Y and Z be smooth vector fields on M and let f be a smooth real-valued function on M. Then

$$[X, fY] = f\,[X,Y] + X[f]\,Y,$$

hence

$$\begin{aligned}
R(X, fY)Z &= \nabla_X \nabla_{fY} Z - \nabla_{fY} \nabla_X Z - \nabla_{(f\,[X,Y] + X[f]\,Y)} Z \\
&= \nabla_X (f\, \nabla_Y Z) - f\, \nabla_Y \nabla_X Z - f\, \nabla_{[X,Y]} Z - X[f]\, \nabla_Y Z \\
&= f\, \nabla_X \nabla_Y Z + X[f] \nabla_Y Z - f\, \nabla_Y \nabla_X Z - f\, \nabla_{[X,Y]} Z \\
&\quad - X[f]\, \nabla_Y Z \\
&= f\, R(X,Y)Z.
\end{aligned}$$

Also

$$R(fX,Y)Z = -R(Y,fX)Z = -f\,R(Y,X)Z = f\,R(X,Y)Z,$$

and

$$\begin{aligned} R(X,Y)(fZ) &= \nabla_X\,(f\,\nabla_Y Z + Y[f]\,Z) - \nabla_Y\,(f\,\nabla_X Z + X[f]\,Z) \\ &\quad - f\,\nabla_{[X,Y]}Z - [X,Y][f]\,Z \\ &= (f\,\nabla_X\nabla_Y Z + X[f]\,\nabla_Y Z + Y[f]\,\nabla_X Z + X[Y[f]]\,Z) \\ &\quad - (f\,\nabla_Y\nabla_X Z + Y[f]\,\nabla_X Z + X[f]\,\nabla_Y Z + Y[X[f]]\,Z) \\ &\quad - f\,\nabla_{[X,Y]}Z - X[Y[f]]\,Z + Y[X[f]]\,Z \\ &= f\,R(X,Y)Z. \end{aligned}$$

Thus

$$R(fX,Y)Z = R(X,fY)Z = R(X,Y)(fZ) = f\,R(X,Y)Z.$$

We now show that the value of $R(X,Y)Z$ depends only on the values of the vector fields X, Y and Z at m. Choose a smooth coordinate system $(x^1, x^2, \ldots, x^n)$ around m. Let the functions $(R^i{}_{jkl})$ be defined such that

$$R\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right)\frac{\partial}{\partial x^j} = \sum_{i=1}^n R^i{}_{jkl}\frac{\partial}{\partial x^i}.$$

Let

$$X = \sum_{k=1}^n u^k\frac{\partial}{\partial x^k}, \qquad Y = \sum_{l=1}^n v^l\frac{\partial}{\partial x^l}, \qquad Z = \sum_{j=1}^n a^j\frac{\partial}{\partial x^j}.$$

Then

$$R(X,Y)Z = \sum_{i,j,k,l}\left(R^i{}_{jkl}a^j u^k v^l\frac{\partial}{\partial x^i}\right).$$

We deduce that the value of $R(X,Y)Z$ depends only on the values of X, Y and Z at m, as required. ∎

Let ∇ be an affine connection on a smooth manifold M, let Y be a smooth vector field on M, and let X_m, Y_m and Z_m be tangent vectors at some point m of M. We define $T(X_m, Y_m)$ and $R(X_m, Y_m)Z_m$ to be the values of the vector fields $T(X,Y)$ and $R(X,Y)Z$ at m, where X, Y and Z are any smooth vector fields on M whose values at m are X_m, Y_m and Z_m respectively. Lemma 3.2 and Lemma 3.3 show that $T(X_m, Y_m)$ and $R(X_m, Y_m)Z_m$ are well-defined, and do not depend on the choice of X, Y and Z.

Lemma 3.4 (The First Bianchi Identity) *Let ∇ be a torsion-free affine connection on a smooth manifold M. Let R denote the curvature operator of ∇. Then*

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

Proof The connection ∇ is torsion-free, hence $\nabla_X Y - \nabla_Y X = [X,Y]$ for all vector fields X and Y on M. Therefore

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

3.1 Vector Fields along Smooth Maps

Let P and M be smooth manifolds, and let $\alpha\colon P \to M$ be a smooth map. A vector field V along the map α is defined to be a map which associates to each point p of P a tangent vector $V(p)$ to M at $\alpha(p)$. If $(x^1, x^2, \ldots, x^n)$ is a smooth coordinate system defined over some open set U in M then we can write

$$V(p) = \sum_{i=1}^{n} c^i(p) \left.\frac{\partial}{\partial x^i}\right|_{\alpha(p)}$$

for all $p \in \alpha^{-1}(U)$. We say that a vector field V along the map α is smooth if, given any smooth coordinate system $(x^1, x^2, \ldots, x^n)$ defined over an open subset U of M, the components $c^1, c^2, \ldots, c^n$ of V with respect to this coordinate system are smooth functions on $\alpha^{-1}(U)$. In particular, one can define in this way smooth vector fields along curves and surfaces in the smooth manifold M.

Let $\gamma\colon I \to M$ be a smooth curve in the smooth manifold M, where I is some open interval in $\mathbb{R}$. Then a vector field V along the curve γ is a function which associates to $t \in I$ a tangent vector $V(t)$ to M at $\gamma(t)$. In particular, the mapping sending $t \in I$ to the velocity vector $\gamma'(t)$ of the curve γ at $\gamma(t)$ provides an example of a smooth vector field along the curve γ.

3.2 Covariant Differentiation of Vector Fields along Curves

Let M be a smooth manifold, and let ∇ be an affine connection on M. Let $\gamma: I \to M$ be a smooth curve in the smooth manifold M, where I is some open interval in $\mathbb{R}$. Given any $t \in I$, we can find smooth vector fields $X_1, X_2, \ldots, X_n$, defined over some open set U containing $\gamma(t)$, with the property that the values $(X_1)_m, (X_2)_m, \ldots, (X_n)_m$ of these vector fields at any point m of U constitute a basis for the tangent space $T_m M$ of M at m. (For example, we can take $X_i = \partial/\partial x^i$, where $(x^1, x^2, \ldots, x^n)$ is any smooth coordinate system defined over U.) If V is any smooth vector field along the curve γ, we can write

$$V(t) = \sum_{i=1}^{n} v^i(t)(X_i)_{\gamma(t)},$$

where $v^1, v^2, \ldots, v^n$ are smooth functions on $\gamma^{-1}(U)$. We define the *covariant derivative* of V along the curve γ at $\gamma(t)$ by the formula

$$\frac{DV(t)}{dt} = \sum_{i=1}^{n} \left(\frac{dv^i(t)}{dt}(X_i)_{\gamma(t)} + v^i(t)\nabla_{\gamma'(t)} X_i \right),$$

where $\gamma'(t)$ denotes the velocity vector of the curve γ at $\gamma(t)$. We claim that this expression defining the covariant derivative does not depend on the choice of the vector fields $X_1, X_2, \ldots, X_n$ around $\gamma(t)$. Indeed suppose that $Y_1, Y_2, \ldots Y_n$ are vector fields on U which form a basis of the tangent space at each point of U. Then

$$X_i = \sum_{j=1}^{n} a_i^j Y_j$$

on U, where (a_i^j) is an $n \times n$ matrix of smooth functions on U which is non-singular at each point of U. Thus

$$V(t) = \sum_{j=1}^{n} w^j(t)(Y_j)_{\gamma(t)}, \text{ where } w^j(t) = \sum_{i=1}^{n} v^i(t) a_i^j(\gamma(t))$$

for all $t \in \gamma^{-1}(U)$, and

$$\begin{aligned}
&\sum_{j=1}^{n} \left(\frac{dw^j(t)}{dt}(Y_j)_{\gamma(t)} + w^j(t)\nabla_{\gamma'(t)} Y_j \right) \\
&\qquad = \sum_{i,j=1}^{n} \Big(\frac{dv^i(t)}{dt} a_i^j(\gamma(t))(Y_j)_{\gamma(t)} + v^i(t)\frac{d(a_i^j \circ \gamma)(t)}{dt}(Y_j)_{\gamma(t)}
\end{aligned}$$

$$
\begin{aligned}
&\quad + v^i(t)a_i^j(\gamma(t))\nabla_{\gamma'(t)}Y_j\Big) \\
&= \sum_{i=1}^n \Big(\frac{dv^i(t)}{dt}(X_i)_{\gamma(t)} + v^i(t)\nabla_{\gamma'(t)}\Big(\sum_{j=1}^n a_i^j Y_j\Big)\Big) \\
&= \sum_{i=1}^n \left(\frac{dv^i(t)}{dt}(X_i)_{\gamma(t)} + v^i(t)\nabla_{\gamma'(t)}X_i\right),
\end{aligned}
$$

showing that the covariant derivative of the vector field V along γ is indeed well-defined, and does not depend on the choice of the vector fields $X_1, X_2, \ldots, X_n$. The properties of the covariant derivative operator D/dt stated in the following lemma follow easily from the definition of the covariant derivative.

Lemma 3.5 *Let M be a smooth manifold, let ∇ be an affine connection on M, and let $\gamma: I \to M$ be a smooth curve in M. Let V and W be smooth vector fields along γ and let $f: I \to \mathbb{R}$ be a smooth real-valued function. Then*

(i) $\dfrac{D(V(t)+W(t))}{dt} = \dfrac{DV(t)}{dt} + \dfrac{DW(t)}{dt}$,

(ii) $\dfrac{D(f(t)V(t))}{dt} = \dfrac{df(t)}{dt}V(t) + f(t)\dfrac{DV(t)}{dt}$,

(iii) *if $V(t) = X_{\gamma(t)}$ for all t, where X is some smooth vector field defined over an open set in M, then* $\dfrac{DV(t)}{dt} = \nabla_{\gamma'(t)}X$.

Moreover the differential operator D/dt is the unique operator on the space of smooth vector fields along the curve γ satisfying (i), (ii) and (iii).

A smooth vector field V along a smooth curve γ is said to be *parallel* if $\dfrac{DV(t)}{dt} = 0$ for all t.

3.3 Vector Fields along Parameterized Surfaces

Let M be a smooth manifold, let U be a connected open set in $\mathbb{R}^m$, and let $\alpha: U \to M$ be a smooth map from U to M. Given $(t^1, t^2, \ldots, t^m) \in U$, we define

$$\frac{\partial\alpha(t^1, t^2, \ldots, t^m)}{\partial t^i}$$

to be the velocity vector of the curve $t \mapsto \alpha(t^1, \ldots, t^{i-1}, t, t^{i+1}, \ldots, t^m)$ at $t = t^i$. Then $\partial\alpha/\partial t^i$ is a smooth vector field along the map α for $i = 1, 2, \ldots, m$.

Let ∇ be an affine connection on M. Given any smooth vector field V along the map α, and given $(t^1, t^2, \ldots, t^m) \in U$, we define

$$\frac{DV(t^1, t^2, \ldots, t^m)}{\partial t^i}$$

to be the covariant derivative of the vector field

$$t \mapsto V(t^1, \ldots, t^{i-1}, t, t^{i+1}, \ldots, t^m)$$

along the curve $t \mapsto \alpha(t^1, \ldots, t^{i-1}, t, t^{i+1}, \ldots, t^m)$ at $t = t^i$. Then the partial covariant derivative $DV/\partial t^i$ is a smooth vector field along the map α of $i = 1, 2, \ldots, m$. We now consider the case $m = 2$.

Let M be a smooth manifold of dimension n. A smooth *parameterized surface* in M is a smooth map $\alpha\colon U \to M$ defined on a connected open subset U on $\mathbb{R}^2$.

Lemma 3.6 *Let M be a smooth manifold and let ∇ be an affine connection on M. Let V be a smooth vector field along a smooth parameterized surface $\alpha\colon U \to M$ in M. Then*

$$\begin{aligned}
\frac{D}{\partial s}\frac{\partial \alpha(s,t)}{\partial t} - \frac{D}{\partial t}\frac{\partial \alpha(s,t)}{\partial s} &= T\Big(\frac{\partial \alpha(s,t)}{\partial s}, \frac{\partial \alpha(s,t)}{\partial t}\Big), \\
\frac{D}{\partial s}\frac{DV(s,t)}{\partial t} - \frac{D}{\partial t}\frac{DV(s,t)}{\partial s} &= R\left(\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\right) V(s,t),
\end{aligned}$$

where T and R are the torsion and curvature tensors of the affine connection ∇.

Proof Without loss of generality, we may suppose that the image of the map $\alpha\colon U \to M$ is contained in the domain of some smooth coordinate system $(x^1, x^2, \ldots, x^n)$. Let $B_1, B_2, \ldots, B_n$ be the smooth vector fields over this coordinate patch defined by $B_i = \partial/\partial x^i$ for $i = 1, 2, \ldots, n$. The vectors $B_1, B_2, \ldots, B_n$ constitute a basis of the tangent space at each point of this coordinate patch. Moreover $[B_j, B_k] = 0$ for all j and k, so that

$$\nabla_{B_j} B_k - \nabla_{B_k} B_j = T(B_j, B_k), \qquad \nabla_{B_j} \nabla_{B_k} B_i - \nabla_{B_k} \nabla_{B_j} B_i = R(B_j, B_k) B_i,$$

for all i, j and k.

The map $\alpha\colon U \to M$ is specified, with respect to the coordinate system $(x^1, x^2, \ldots, x^n)$, by smooth real-valued functions $\alpha^1, \alpha^2, \ldots, \alpha^n$ on U. (Thus the coordinates of $\alpha(s,t)$ are given by $x^i = \alpha^i(s,t)$ for all $i = 1, 2, \ldots, n$.) It follows that

$$\frac{\partial \alpha}{\partial s} = \sum_{j=1}^n \frac{\partial \alpha^j}{\partial s} B_j, \qquad \frac{\partial \alpha}{\partial t} = \sum_{k=1}^n \frac{\partial \alpha^k}{\partial t} B_k.$$

Thus

$$\frac{DX}{\partial s} = \sum_{j=1}^{n} \frac{\partial \alpha^j}{\partial s} \nabla_{B_j} X, \qquad \frac{DX}{\partial t} = \sum_{k=1}^{n} \frac{\partial \alpha^k}{\partial t} \nabla_{B_k} X.$$

for all smooth vector fields X on the coordinate patch. Now

$$\begin{aligned} \frac{D}{\partial s}\frac{\partial \alpha}{\partial t} &= \sum_{k=1}^{n} \frac{D}{\partial s}\Bigl(\frac{\partial \alpha^k}{\partial t} B_k\Bigr) \\ &= \sum_{k=1}^{n} \frac{\partial^2 \alpha^k}{\partial s \partial t} B_k + \sum_{j,k=1}^{n} \frac{\partial \alpha^j}{\partial s}\frac{\partial \alpha^k}{\partial t} \nabla_{B_j} B_k. \end{aligned}$$

Thus

$$\begin{aligned} \frac{D}{\partial s}\frac{\partial \alpha}{\partial t} - \frac{D}{\partial t}\frac{\partial \alpha}{\partial s} &= \sum_{j,k=1}^{n} \frac{\partial \alpha^j}{\partial s}\frac{\partial \alpha^k}{\partial t} \left(\nabla_{B_j} B_k - \nabla_{B_k} B_j\right) \\ &= \sum_{j,k=1}^{n} \frac{\partial \alpha^j}{\partial s}\frac{\partial \alpha^k}{\partial t} T(B_j, B_k) = T\Bigl(\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\Bigr). \end{aligned}$$

Let $f\colon U \to \mathbb{R}$ be a smooth real-valued function on U, and let V be a smooth vector field along the map α. Then

$$\begin{aligned} \frac{D}{\partial s}\frac{D(fV)}{\partial t} &= \frac{D}{\partial s}\Bigl(\frac{\partial f}{\partial t} V + f \frac{DV}{\partial t}\Bigr) \\ &= \frac{\partial^2 f}{\partial s \partial t} V + \frac{\partial f}{\partial t}\frac{DV}{\partial s} + \frac{\partial f}{\partial s}\frac{DV}{\partial t} + f \frac{D}{\partial s}\frac{DV}{\partial t}, \end{aligned}$$

and thus

$$\Bigl(\frac{D}{\partial s}\frac{D}{\partial t} - \frac{D}{\partial t}\frac{D}{\partial s}\Bigr)(fV) = f\Bigl(\frac{D}{\partial s}\frac{DV}{\partial t} - \frac{D}{\partial t}\frac{DV}{\partial s}\Bigr)$$

Now any smooth vector field V along the map α can be expressed in the form

$$V(s,t) = \sum_{i=1}^{n} v^i(s,t) (B_i)_{\alpha(s,t)}$$

for some smooth real-valued functions $v^1, v^2, \ldots, v^n$ on U. It follows that

$$\frac{D}{\partial s}\frac{DV}{\partial t} - \frac{D}{\partial t}\frac{DV}{\partial s} = \sum_{i=1}^{n} v^i \Bigl(\frac{D}{\partial s}\frac{D}{\partial t} - \frac{D}{\partial t}\frac{D}{\partial s}\Bigr) B_i.$$

But

$$\begin{aligned} \frac{D}{\partial s}\frac{D}{\partial t} B_i &= \sum_{k=1}^{n} \frac{D}{\partial s}\Bigl(\frac{\partial \alpha^k}{\partial t} \nabla_{B_k} B_i\Bigr) \\ &= \sum_{k=1}^{n} \frac{\partial^2 \alpha^k}{\partial s \partial t} \nabla_{B_k} B_i + \sum_{j,k=1}^{n} \frac{\partial \alpha^j}{\partial s}\frac{\partial \alpha^k}{\partial t} \nabla_{B_j} \nabla_{B_k} B_i, \end{aligned}$$

and hence

$$\begin{aligned}
\Big(\frac{D}{\partial s}\frac{D}{\partial t} - \frac{D}{\partial t}\frac{D}{\partial s}\Big)B_i &= \sum_{j,k=1}^{n} \frac{\partial\alpha^j}{\partial s}\frac{\partial\alpha^k}{\partial t}\left(\nabla_{B_j}\nabla_{B_k}B_i - \nabla_{B_k}\nabla_{B_j}B_i\right) \\
&= \sum_{j,k=1}^{n} \frac{\partial\alpha^j}{\partial s}\frac{\partial\alpha^k}{\partial t} R(B_j, B_k)B_i \\
&= R\Big(\frac{\partial\alpha}{\partial s}, \frac{\partial\alpha}{\partial t}\Big)B_i.
\end{aligned}$$

We deduce that

$$\frac{D}{\partial s}\frac{DV}{\partial t} - \frac{D}{\partial t}\frac{DV}{\partial s} = \sum_{i=1}^{n} v^i R\Big(\frac{\partial\alpha}{\partial s}, \frac{\partial\alpha}{\partial t}\Big)B_i = R\Big(\frac{\partial\alpha}{\partial s}, \frac{\partial\alpha}{\partial t}\Big)V,$$

as required. ∎

4 Riemannian Manifolds

Definition Let M be a smooth manifold. A *Riemannian metric* g on M assigns to any smooth vector fields X and Y on M a smooth function $g(X, Y)$, where

$$g(X_1+X_2, Y) = g(X_1, Y)+g(X_2, Y), \qquad g(X, Y_1+Y_2) = g(X, Y_1)+g(X, Y_2),$$

$$g(fX, Y) = f\,g(X, Y) = g(X, fY), \qquad g(X, Y) = g(Y, X)$$

for all smooth real-valued functions f and vector fields X, X_1, X_2, Y, Y_1, Y_2, and

$$g(X, X) > 0 \text{ wherever } X \neq 0.$$

Let $(x^1, x^2, \ldots, x^n)$ be a smooth coordinate system defined over some open set U in M. Define smooth real-valued functions g_{ij} on U by the formula

$$g_{ij} = g\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big)$$

The $n\times n$ matrix (g_{ij}) is positive-definite at each point of U, since $g(X, X) > 0$ for all non-zero vector fields X on U. If X and Y are smooth vector fields on M then

$$X = \sum_{i=1}^{n} a^i \frac{\partial}{\partial x^i}, \qquad Y = \sum_{i=1}^{n} b^i \frac{\partial}{\partial x^i},$$

where $a^1, a^2, \ldots, a^n$ and $b^1, b^2, \ldots, b^n$ are smooth functions on U, and therefore

$$g(X, Y) = \sum_{i,j=1}^{n} g_{ij} a^i b^j.$$

We deduce that the value of $g(X, Y)$ at any point m of M depends only on the values X_m and Y_m of X and Y at m.

Given tangent vectors X_m and Y_m at a point m of M, we denote by $g(X_m, Y_m)$ the value of $g(X, Y)$ at m, where X and Y are any smooth vector fields on M with values X_m and Y_m at m. Then $(X_m, Y_m) \mapsto g(X_m, Y_m)$ is a well-defined inner product on the tangent space $T_m M$ to M at m. Thus one can regard a Riemannian metric on a smooth manifold M as a smooth assignment of an inner product to each tangent space of M.

Definition A *Riemannian manifold* (M, g) consists of a smooth manifold M together with a (smooth) Riemannian metric g on M.

Remark It should be noted that the metric tensor of the theory of general relativity is not a Riemannian metric, since it is not positive-definite at each point of the space-time manifold. It is in fact a 'semi-Riemannian' metric.

Lemma 4.1 *Let M be a Riemannian manifold with Riemannian metric g. Let θ be a transformation mapping smooth vector fields on M to smooth functions on M. Suppose that*

$$\theta(X + Y) = \theta(X) + \theta(Y), \qquad \theta(fX) = f\,\theta(X)$$

for all smooth real-valued functions f and vector fields X and Y on M. Then there exists a unique smooth vector field V on M with the property that $\theta(X) = g(V, X)$ for all smooth vector fields X on M.

Proof First we verify the uniqueness of the vector field V. Let U be an open set in M and let V and W be vector fields on U with the property that $g(V, X) = \theta(X) = g(W, X)$ for all smooth vector fields X on U. Then $g(V - W, X) = 0$ on U for all vector fields X. In particular, $g(V - W, V - W) = 0$ on U. It follows from the definition of a Riemannian metric that $V - W = 0$, so that $V = W$ on U. This proves the uniqueness of the vector field V.

Now suppose that the open set U is the domain of some smooth coordinate system $(x^1, x^2, \ldots, x^n)$. Let

$$g_{ij} = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \qquad (i, j = 1, 2, \ldots, n)$$

Then (g_{ij}) is a matrix of smooth functions on U which is positive definite, and hence invertible, at each point of U. Let (g^{ij}) be the smooth functions on U characterized by the property that, at each point of U, the matrix (g^{ij}) is the inverse of (g_{ij}). Thus

$$\sum_{j=1}^{n} g^{ij} g_{jk} = \begin{cases} 1 & \text{if } i = k; \\ 0 & \text{if } i \neq k. \end{cases}$$

Define a smooth vector field V on U by

$$V = \sum_{i,j=1}^{n} \theta_i g^{ij} \frac{\partial}{\partial x^j},$$

where

$$\theta_i = \theta\Big(\frac{\partial}{\partial x^i}\Big) \qquad (i = 1, 2, \dots, n).$$

Let X be a smooth vector field on U, given by

$$X = \sum_{k=1}^{n} a^k \frac{\partial}{\partial x^k}.$$

Then

$$g(V, X) = \sum_{i,j,k=1}^{n} \theta_i g^{ij} g_{jk} a^k = \sum_{i=1}^{n} \theta_i a^i = \theta(X)$$

on U. We thus obtain a smooth vector field V over any coordinate patch U with the property that $g(V, X) = \theta(X)$ for all vector fields X on U. If we are given two overlapping coordinate systems on M then the uniqueness result already proved shows that the vector fields over the coordinate patches obtained in the manner just described must agree on the overlap of the coordinate patches. Thus we obtain a smooth vector field V defined over the whole of M such that $\theta(X) = g(V, X)$ for all smooth vector fields X on M, as required. ∎

4.1 The Levi-Civita Connection

Let (M, g) be a Riemannian manifold and let ∇ be an affine connection on M. We say that ∇ is *compatible with the Riemannian metric* g if

$$Z[g(X, Y)] = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

for all smooth vector fields X, Y and Z on M. We shall show that on every Riemannian manifold there exists a unique torsion-free connection that is compatible with the Riemannian metric.

Lemma 4.2 *Let ∇ be an affine connection on a Riemannian manifold M that is compatible with the Riemannian metric g. Let V and W be smooth vector fields along some smooth curve $\gamma\colon I \to M$ in M (where I denotes some open interval in $\mathbb{R}$). Then*

$$\frac{d}{dt} g(V(t), W(t)) = g\Bigl(\frac{DV(t)}{dt}, W(t)\Bigr) + g\Bigl(V(t), \frac{DW(t)}{dt}\Bigr)$$

(where DV/dt and DW/dt are the covariant derivatives of the vector fields V and W along the curve γ).

Proof Suppose that $\gamma(t) \in U$, where U is the domain of some smooth coordinate system $(x^1, x^2, \ldots, x^n)$ on M. Then

$$V(t) = \sum_{i=1}^n v^i(t)(B_i)_{\gamma(t)}, \qquad W(t) = \sum_{j=1}^n w^j(t)(B_j)_{\gamma(t)},$$

where $v^1, v^2, \ldots, v^n$ and $w^1, w^2, \ldots, w^n$ are smooth functions on $\gamma^{-1}(U)$ and $B_i = \partial/\partial x^i$ for $i = 1, 2, \ldots, n$. Thus

$$\begin{aligned}
\frac{d}{dt} g(V(t), W(t)) &= \sum_{i,j=1}^n \frac{d}{dt}\bigl(v^i(t) w^j(t) g(B_i, B_j)\bigr) \\
&= \sum_{i,j=1}^n \Bigl(\frac{dv^i(t)}{dt} w^j(t) g(B_i, B_j) + v^i(t) \frac{dw^j(t)}{dt} g(B_i, B_j) \\
&\qquad + v^i(t) w^j(t)(g(\nabla_{\gamma'(t)} B_i, B_j) + g(B_i, \nabla_{\gamma'(t)} B_j))\Bigr) \\
&= g\Bigl(\sum_{i=1}^n \Bigl(\frac{dv^i(t)}{dt} B_i + v^i(t) \nabla_{\gamma'(t)} B_i\Bigr), W(t)\Bigr) \\
&\qquad + g\Bigl(V(t), \sum_{j=1}^n \Bigl(\frac{dw^j(t)}{dt} B_j + w^j(t) \nabla_{\gamma'(t)} B_j\Bigr)\Bigr) \\
&= g\Bigl(\frac{DV(t)}{dt}, W(t)\Bigr) + g\Bigl(V(t), \frac{DW(t)}{dt}\Bigr),
\end{aligned}$$

as required. ∎

Theorem 4.3 *Let (M, g) be a Riemannian manifold. Then there exists a unique torsion-free affine connection ∇ on M compatible with the Riemannian metric g. This connection is characterized by the identity*

$$\begin{aligned}
2g(\nabla_X Y, Z) &= X[g(Y, Z)] + Y[g(X, Z)] - Z[g(X, Y)] \\
&\qquad + g([X, Y], Z) - g([X, Z], Y) - g([Y, Z], X)
\end{aligned}$$

for all smooth vector fields X, Y and Z on M.

Proof Given smooth vector fields X, Y and Z on M, let $A(X,Y,Z)$ be the smooth function on M defined by

$$\begin{aligned} A(X,Y,Z) &= \tfrac{1}{2}(X[g(Y,Z)] + Y[g(X,Z)] - Z[g(X,Y)] \\ &\quad + g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X)). \end{aligned}$$

Then $A(X,Y,Z_1+Z_2) = A(X,Y,Z_1) + A(X,Y,Z_2)$ for all smooth vector fields X, Y, Z_1 and Z_2 on M. Using the identities

$$[X,fZ] = f\,[X,Z] + X[f]\,Z, \qquad [Y,fZ] = f\,[Y,Z] + Y[f]\,Z$$

one can readily verify that $A(X,Y,fZ) = f\,A(X,Y,Z)$ for all smooth real-valued functions f and vector fields X, Y and Z on M. On applying Lemma 4.1 to the transformation $Z \mapsto A(X,Y,Z)$, we see that there is a unique vector field $\nabla_X Y$ on M with the property that $A(X,Y,Z) = g(\nabla_X Y, Z)$ for all smooth vector fields X, Y and Z on M. Moreover

$$\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y, \qquad \nabla_X(Y_1+Y_2) = \nabla_X Y_1 + \nabla_X Y_2.$$

Straightforward calculations show that

$$\begin{aligned} g(\nabla_{fX} Y, Z) &= A(fX,Y,Z) = f\,A(X,Y,Z) = g(f\,\nabla_X Y, Z), \\ g(\nabla_X (fY), Z) &= A(X,fY,Z) = f\,A(X,Y,Z) + X[f]\,g(Y,Z) \\ &= g(f\,\nabla_X Y + X[f]\,Y, Z), \end{aligned}$$

for all smooth real-valued functions f on M, so that

$$\nabla_{fX} Y = f\,\nabla_X Y, \qquad \nabla_X(fY) = f\,\nabla_X Y + X[f]\,Y.$$

These properties show that ∇ is indeed an affine connection on M. Moreover

$$A(X,Y,Z) - A(Y,X,Z) = g([X,Y],Z),$$

so that $\nabla_X Y - \nabla_Y X = [X,Y]$. Thus the affine connection ∇ is torsion-free. Also

$$g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = A(X,Y,Z) + A(X,Z,Y) = X[g(Y,Z)],$$

showing that the affine connection ∇ preserves the Riemannian metric.

Finally suppose that ∇' is any torsion-free affine connection on M which preserves the Riemannian metric. Then

$$\begin{aligned} X[g(Y,Z)] &= g(\nabla'_X Y, Z) + g(Y, \nabla'_X Z), \\ Y[g(X,Z)] &= g(\nabla'_Y X, Z) + g(X, \nabla'_Y Z), \\ Z[g(X,Y)] &= g(\nabla'_Z X, Y) + g(X, \nabla'_Z Y). \end{aligned}$$

A straightforward calculation (using the fact that ∇' is torsion-free) shows that $A(X,Y,Z) = g(\nabla'_X Y, Z)$. Therefore $\nabla'_X Y = \nabla_X Y$ for all smooth vector fields X and Y on M, as required. ∎

Let (M, g) be a Riemannian manifold. The unique torsion-free affine connection on M which preserves the Riemannian metric is known as the *Levi-Civita connection* on M.

Example Let M be a smooth n-dimensional submanifold of k-dimensional Euclidean space $\mathbb{R}^k$. Given (tangential) vector fields X and Y on M, we decompose the directional derivative $\partial_X Y$ of Y along X as $\partial_X Y = \nabla_X Y + S(X, Y)$, where $\nabla_X Y$ is tangential to M and $S(X, Y)$ is orthogonal to M. Then ∇ is a torsion-free affine connection on M (as was shown in an earlier example). Now the restriction to the tangent spaces of M of the standard inner product $\langle ., . \rangle$ on $\mathbb{R}^k$ gives a Riemannian metric g on M. Moreover

$$g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = \langle \partial_X Y, Z \rangle + \langle Y, \partial_X Z \rangle = X[\langle Y, Z \rangle] = X[g(Y, Z)]$$

for all vector fields X, Y and Z on M that are everywhere tangential to M. We conclude that the affine connection ∇ on M coincides with the Levi-Civita connection of the Riemannian manifold (M, g).

The *Riemann curvature tensor* R of a Riemannian manifold (M, g) is given by the formula

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

where

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

for all smooth vector fields X, Y and Z on M. It follows immediately from Lemma 3.3 that the value of $R(W, Z, X, Y)$ at a point m of M depends only on the values W_m, Z_m, X_m and Y_m of the vector fields W, Z, X and Y at the point m.

Proposition 4.4 *Let (M, g) be a Riemannian manifold. The Riemann curvature tensor on M satisfies the following identities:—*

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

(ii) $R(W, X, Y, Z) + R(W, Y, Z, X) + R(W, Z, X, Y) = 0$,

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

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

for all smooth vector fields X, Y, Z and W on M.

Proof (i) follows directly from the definition of the Riemann curvature tensor, and (ii) corresponds to the First Bianchi Identity

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

(see Lemma 3.4). Now

$$\begin{aligned} X[Y[g(W,Z)]] &= X\left[g(\nabla_Y W, Z) + g(W, \nabla_Y Z)\right] \\ &= g(\nabla_X \nabla_Y W, Z) + g(\nabla_Y W, \nabla_X Z) \\ &\quad + g(\nabla_X W, \nabla_Y Z) + g(W, \nabla_X \nabla_Y Z), \end{aligned}$$

and hence

$$\begin{aligned} [X,Y][g(W,Z)] &= X[Y[g(W,Z)]] - Y[X[g(W,Z)]] \\ &= g(\nabla_X \nabla_Y W - \nabla_Y \nabla_X W, Z) \\ &\quad + g(W, \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z). \end{aligned}$$

Therefore

$$\begin{aligned} &R(W,Z,X,Y) + R(Z,W,X,Y) \\ &\quad = g(W, R(X,Y)Z) + g(R(X,Y)W, Z) \\ &\quad = [X,Y][g(W,Z)] - g(\nabla_{[X,Y]} W, Z) - g(W, \nabla_{[X,Y]} Z) \\ &\quad = 0. \end{aligned}$$

This proves (iii). Using (i), (ii) and (iii), we see that

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

This proves (iv). ∎

Let m be point in M and let P be a two-dimensional vector subspace (plane) in the tangent space $T_m M$ to M at m. Let (E_1, E_2) be an orthonormal basis of P. We define the *sectional curvature* $K(P)$ of M in the plane P by the formula

$$K(P) = R(E_1, E_2, E_1, E_2).$$

Note that if X and Y are tangent vectors in P then

$$X = a_{11}E_1 + a_{12}E_2, \qquad Y = a_{21}E_1 + a_{22}E_2,$$

for some real numbers a_{11}, a_{12}, a_{21} and a_{22}, and hence

$$\begin{aligned} R(X,Y,X,Y) &= R(X,Y,a_{11}E_1 + a_{12}E_2, a_{21}E_1 + a_{22}E_2) \\ &= (a_{11}a_{22} - a_{12}a_{21})R(X,Y,E_1,E_2) \\ &= (\det A)R(X,Y,E_1,E_2) = (\det A)^2 R(E_1,E_2,E_1,E_2) \\ &= (\det A)^2 K(P), \end{aligned}$$

where A is the matrix given by

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

In particular, if (X,Y) is any orthonormal basis of P then the matrix A is an orthogonal matrix, and thus $\det A = \pm 1$. It follows that the value of the sectional curvature $K(P)$ does not depend on the choice of the orthonormal basis (E_1, E_2) of P.

Lemma 4.5 *Let (M,g) be a Riemannian manifold, and let m be a point of M. Then the values of the sectional curvatures $K(P)$ for all planes P in the tangent space T_mM to M at m determine the Riemann curvature tensor at m.*

Proof The calculation given above shows that the sectional curvatures determine the values of $R(X,Y,X,Y)$ for all $X,Y \in T_mM$.

Now suppose that we are given $X,Y,Z \in T_mM$. Using the symmetries of the Riemann curvature tensor listed in Proposition 4.4, we see that

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

Thus the sectional curvatures $K(P)$ determine the values of $R(X,Y,X,Z)$ for all tangent vectors X, Y and Z at m. It follows from this that the sectional curvatures determine $R(X,Y,Z,X)$, $R(Y,X,X,Z)$ and $R(Y,X,Z,X)$. But

$$\begin{aligned} 3R(W,X,Y,Z) &= 2R(W,X,Y,Z) - R(W,Y,Z,X) - R(W,Z,X,Y) \\ &= (R(W,X,Y,Z) + R(W,Y,X,Z)) \\ &\quad + (R(W,X,Y,Z) + R(W,Z,Y,X)) \\ &= R(W,X+Y,X+Y,Z) - R(W,X,X,Z) \\ &\quad - R(W,Y,Y,Z) + R(W,X+Z,Y,X+Z) \\ &\quad - R(W,X,Y,X) - R(W,Z,Y,Z). \end{aligned}$$

We conclude that $R(W, X, Y, Z)$ is determined by the sectional curvatures of M, as required. ∎

5 Geometry of Surfaces in $\mathbb{R}^3$

A *smooth surface* in $\mathbb{R}^3$ is defined to be a smooth 2-dimensional submanifold of $\mathbb{R}^3$. Let M be a smooth surface in $\mathbb{R}^3$. If this surface is orientable, then we can choose a unit normal vector ν_p at each point p of M (where ν_p is orthogonal to the tangent space T_pM to M at p) such that ν_p varies smoothly with p. We shall assume that M is orientable and that such a smooth unit normal vector field ν has been chosen on M.

Recall that if X and Y are vector fields on M then we denote by $\partial_{X_p} Y$ the derivative of the vector field Y along X. Thus if X_p is the tangent vector to the smooth curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ at zero then

$$\partial_{X_p} Y \equiv \left. \frac{dY(\gamma(t))}{dt} \right|_{t=0}.$$

Note that this definition makes sense for any vector field Y along the surface M which takes values in $\mathbb{R}^3$, whether or not that vector field is everywhere tangential to M. In particular, we can define $\partial_X \nu$, where ν is the chosen unit normal vector field on M by

$$\partial_{X_p} \nu \equiv \left. \frac{d\nu(\gamma(t))}{dt} \right|_{t=0},$$

where γ is a smooth curve such that $\gamma(0) = p$ and $\gamma'(0) = X_p$.

Now let X and Y be two vector fields on M which are everywhere tangential to the surface M. We split $\partial_X Y$ into components tangential to the surface and normal to the surface. We denote the tangential component by $\nabla_X Y$ Thus

$$\partial_X Y = \nabla_X Y + \sigma(X, Y)\nu$$

where $\sigma(X, Y)$ is a smooth function on M determined by the vector fields X and Y. Now

$$\sigma(X + Y, Z) = \sigma(X, Z) + \sigma(Y, Z), \qquad \sigma(X, Y + Z) = \sigma(X, Y) + \sigma(X, Z).$$

Next we note that

$$\sigma(fX, Y) = f\sigma(X, Y) = \sigma(X, fY)$$

since

$$\partial_{fX} Y = f\partial_X Y, \qquad \partial_X(fY) = f\partial_X Y + X[f]Y,$$

and $X[f]Y$ is clearly tangential to M. Also $\sigma(X,Y) = \sigma(Y,X)$ for all tangential vector fields X and Y on M, since

$$\partial_X Y - \partial_Y X = [X,Y],$$

and the Lie bracket $[X,Y]$ is tangential to M.

We can regard the surface M as a 2-dimensional Riemannian manifold, where the Riemannian metric on M is the restriction to each tangent space of the standard inner product $\langle .,. \rangle$ on $\mathbb{R}^3$. Then the connection ∇ on M defined as above is torsion-free and preserves the Riemannian metric (i.e.,

$$Z[\langle X,Y\rangle] = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle$$

for all tangential vector fields X, Y and Z on M. Thus this connection is the Levi-Civita connection on the surface M.

Observe that $\partial_X \nu$ is tangential to X for all tangental vectors X on M (where $\partial_X \nu$ denotes the directional derivative of the unit normal vector field ν along X. This follows from the fact that

$$2\,\langle \partial_X \nu, \nu\rangle = X[|\nu|^2] = 0.$$

Let Y be a vector field on M that is everywhere tangential to M. Then $\langle \nu, Y\rangle = 0$, hence

$$\begin{aligned}\langle \partial_X \nu, Y\rangle &= \partial_X \langle \nu, Y\rangle - \langle \nu, \partial_X Y\rangle \\ &= -\sigma(X,Y).\end{aligned}$$

Observe that if X and Y are vector fields on M everywhere tangential to M and if V is a vector field on M taking values in $\mathbb{R}^3$ (i.e., V is a 'vector field along the embedding map $i\colon M \hookrightarrow \mathbb{R}^3$') then

$$\partial_X \partial_Y V - \partial_Y \partial_X V = \partial_{[X,Y]} V.$$

Indeed if $V = (v^1, v^2, v^3)$ then

$$\partial_X V \equiv (X[v^1], X[v^2], X[v^3]), \qquad \partial_Y V \equiv (Y[v^1], Y[v^2], Y[v^3]),$$

so that

$$\begin{aligned}&\partial_X \partial_Y V - \partial_Y \partial_X V \\ &\quad= (X[Y[v^1]] - Y[X[v^1]], X[Y[v^2]] - Y[X[v^2]], X[Y[v^3]] - Y[X[v^3]]) \\ &\quad= ([X,Y][v^1], [X,Y][v^2], [X,Y][v^3]) = \partial_{[X,Y]} V.\end{aligned}$$

Let W, Z, X and Y be smooth vector fields on M, everywhere tangential to M. Then

$$\begin{aligned}\langle W, \partial_X \partial_Y Z\rangle &= \langle W, \partial_X(\nabla_Y Z + \sigma(Y,Z)\nu)\rangle \\ &= \langle W, \partial_X \nabla_Y Z\rangle + \sigma(Y,Z)\,\langle W, \partial_X \nu\rangle \\ &= \langle W, \partial_X \nabla_Y Z\rangle + \sigma(Y,Z)\,\langle W, \partial_X \nu\rangle \\ &= \langle W, \nabla_X \nabla_Y Z\rangle - \sigma(Y,Z)\sigma(X,W).\end{aligned}$$

Using the fact that

$$\partial_X \partial_Y Z - \partial_Y \partial_X Z = \partial_{[X,Y]} Z.$$

we see that the Riemann curvature tensor R of M is given by

$$\begin{aligned}R(W,Z,X,Y) &\equiv \left\langle W, \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z\right\rangle \\ &= \langle W, \partial_X \partial_Y Z - \partial_Y \partial_X Z\rangle \\ &\quad + \sigma(Y,Z)\sigma(X,W) - \sigma(X,Z)\sigma(Y,W) \\ &\quad - \left\langle W, \partial_{[X,Y]} Z\right\rangle \\ &= \sigma(Y,Z)\sigma(X,W) - \sigma(X,Z)\sigma(Y,W).\end{aligned}$$

This identity expresses the Riemann curvature tensor of the surface M in terms of the second fundamental form of the surface. This result is essentially due to Gauss (though Gauss employed an alternative formalism).

We now derive an identity satisfied by the second fundamental form σ on M. Note that

$$\begin{aligned}\langle \nu, \partial_X \partial_Y Z\rangle &= \langle \nu, \partial_X(\nabla_Y Z + \sigma(Y,Z)\nu)\rangle \\ &= \sigma(X, \nabla_Y Z) + \langle \nu, X[\sigma(Y,Z)]\nu + \sigma(Y,Z)\partial_X \nu\rangle .\end{aligned}$$

But $\langle \nu, \partial_X \nu\rangle = 0$, hence

$$\langle \nu, \partial_X \partial_Y Z\rangle = \sigma(X, \nabla_Y Z) + X[\sigma(Y,Z)].$$

Also

$$\left\langle \nu, \partial_{[X,Y]} Z\right\rangle = \sigma([X,Y], Z) = \sigma(\nabla_X Y, Z) - \sigma(\nabla_Y X, Z).$$

But

$$\langle \nu, \partial_X \partial_Y Z\rangle - \langle \nu, \partial_X \partial_Y Z\rangle = \left\langle \nu, \partial_{[X,Y]} Z\right\rangle .$$

Therefore

$$\begin{aligned}X[\sigma(Y,Z)] - Y[\sigma(X,Z)] + \sigma(X, \nabla_Y Z) - \sigma(Y, \nabla_X Z)& \\ = \sigma(\nabla_X Y, Z) - \sigma(\nabla_Y X, Z),&\end{aligned}$$

so that

$$(\nabla_X \sigma)(Y, Z) = (\nabla_Y \sigma)(X, Z),$$

where $\nabla\sigma$ denotese the covariant derivative of the tensor σ, defined by

$$(\nabla_X \sigma)(Y, Z) \equiv X[\sigma(Y, Z)] - \sigma(\nabla_X Y, Z) - \sigma(Y, \nabla_X Z).$$

Gathering together the identities proved above, we see that we have verified the *Gauss-Codazzi* identities, stated in the following lemma.

Lemma 5.1 (The Gauss-Codazzi Identities) *Let M be a smooth surface embedded in $\mathbb{R}^3$. Then the Riemann curvature tensor R on M is expressed in terms of the second fundamental form σ of the embedding of M in $\mathbb{R}^3$ by the identity*

$$R(W, Z, X, Y) = \sigma(W, X)\sigma(Z, Y) - \sigma(W, Y)\sigma(Z, X). \qquad \textit{(Gauss' Identity)}$$

Also the covariant derivative $\nabla\sigma$ of the second fundamental form σ satisfies the identity

$$(\nabla_X \sigma)(Y, Z) = (\nabla_Y \sigma)(X, Z). \qquad \textit{(Codazzi's Identity)}$$

Let (E_1, E_2) be an orthonormal basis of the tangent space T_pM to M at the point p. The *Gaussian curvature* of M at p is defined to be the quantity $K(p)$ defined by

$$K(p) = R(E_1, E_2, E_1, E_2).$$

The Gaussian curvature $K(p)$ of M at p is well-defined independently of the choice of orthonormal basis (E_1, E_2) of the tangent space T_pM to M at p. It follows from the Gauss-Codazzi equations that

$$K(p) = \sigma(E_1, E_1)\sigma(E_2, E_2) - \sigma(E_1, E_2)\sigma(E_2, E_1).$$

Thus if (s_{ij}) is the 2×2 matrix representing σ at p with respect to the orthonormal basis (E_1, E_2), then $K(p)$ is the determinant of this matrix.

Lemma 5.2 *Let M be a smooth surface in $\mathbb{R}^3$ and let (X, Y) be a smooth moving frame on M. Then the Gaussian curvature K is given by*

$$K = \frac{\sigma(X, X)\sigma(Y, Y) - \sigma(X, Y)^2}{\langle X, X\rangle \langle Y, Y\rangle - \langle X, Y\rangle^2}$$

Proof Let (E_1, E_2) be an orthonormal moving frame defined around some point of M. We can write

$$\begin{aligned} X &= a_{11}E_1 + a_{12}E_2 \\ Y &= a_{21}E_1 + a_{22}E_2. \end{aligned}$$

Let us define matrices A, B, and C by

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \\ B &= \begin{pmatrix} \sigma(E_1, E_1) & \sigma(E_1, E_2) \\ \sigma(E_2, E_1) & \sigma(E_2, E_2) \end{pmatrix}. \\ C &= \begin{pmatrix} \sigma(X, X) & \sigma(X, Y) \\ \sigma(Y, X) & \sigma(Y, Y) \end{pmatrix}. \end{aligned}$$

The matrices B and C are symmetric and $C = A^T BA$, where A^T is the transpose of A. Moreover the Gaussian curvature K of M is the determinant $\det B$ of the matrix B. Thus $\det C = K(\det A)^2$. But

$$A^T A = \begin{pmatrix} \langle X, X\rangle & \langle X, Y\rangle \\ \langle Y, X\rangle & \langle Y, Y\rangle \end{pmatrix},$$

so that

$$(\det A)^2 = \langle X, X\rangle \langle Y, Y\rangle - \langle X, Y\rangle^2 .$$

The required result follows immediately from this. ∎

Let S^2 denote the standard unit sphere in $\mathbb{R}^3$. We can regard the unit normal vector field ν on the smooth surface M as defining a smooth map from M to S^2 (since $\nu(p)$ is a vector of unit length in $\mathbb{R}^2$). This smooth map from M to S^2 is customarily referred to as the *Gauss map*.

Let us consider the derivative $\nu_*: T_pM \to T_{\nu(p)}S^2$ of the Gauss map $\nu: M \to S^2$ at p. If $\gamma: (-\varepsilon, \varepsilon) \to M$ is a smooth curve in M with $\gamma(0) = p$ and $\gamma'(0) = X_p$ then

$$\nu_* X_p = \left. \frac{d\nu(\gamma(t))}{dt} \right|_{t=0} = \partial_{X_p}\nu.$$

Note that the tangent space $T_{\nu(p)}S^2$ of S^2 at $\nu(p)$ is parallel to the tangent space T_pM of M at p (since both are orthogonal to the vector $\nu(p)$). Using the fact that

$$\left\langle \partial_{X_p}\nu, Y_p \right\rangle = -\sigma(X_p, Y_p)$$

we see that if (E_1, E_2) is an orthonormal basis of the tangent space to M at p and if $(\tilde{E}_1, \tilde{E}_2)$ is the corresponding orthonormal basis of $T_{\nu(p)}S^2$ (where the components of E_1 and E_2 with respect to the standard basis of $\mathbb{R}^3$ are equal to the corresponding components of $\tilde{E}_1$ and $\tilde{E}_2$ respectively) then

$$\begin{aligned} \nu_* E_1 &= -\sigma(E_1, E_1)\tilde{E}_1 - \sigma(E_2, E_1)\tilde{E}_2 \\ \nu_* E_2 &= -\sigma(E_1, E_2)\tilde{E}_1 - \sigma(E_2, E_2)\tilde{E}_2 \end{aligned}$$

Thus the matrix representing the derivative of the Gauss map $\nu\colon M \to S^2$ at p with respect to these orthonormal bases is the symmetric matrix

$$\begin{pmatrix} -s_{11} & -s_{12} \\ -s_{21} & -s_{22} \end{pmatrix},$$

where $s_{ij} = \sigma(E_1, E_2)$. Now the Gaussian curvature $K(p)$ at p is the determinant of this matrix (see above). Therefore the derivative $\nu_*\colon T_pM \to T_{\nu(p)}S^2$ of the Gauss map $\nu\colon M \to S^2$ at p multiplies areas by a factor of $|K(p)|$. Using this fact we deduce the following result.

Lemma 5.3 *Let M be a smooth oriented surface in $\mathbb{R}^3$ and let $\nu\colon M \to S^2$ be the Gauss map of this surface. Suppose that ν is injective on a region D of M. Then*

$$\textit{Area of } \nu(D) = \int_D |K| dA,$$

where the integral is taken with respect to the usual area measure on A.

Informally we can regard the magnitude $|K|$ of the Gaussian curvature at p as the limit of the ratio

$$\frac{\text{area of } \nu(D)}{\text{area of } D}$$

as the region D in M shrinks down inside ever smaller neighbourhoods of the point p. If $K(p) > 0$ then the Gauss map $\nu\colon M \to S^2$ is orientation-preserving around p. If $K(p) < 0$ then the Gauss map is orientation-reversing around p.

However although this description of the Gaussian curvature depends on the manner in which the surface M is embedded in $\mathbb{R}^3$, yet the expression for the Gaussian curvature in terms of the Riemann curvature tensor of M (with respect to the metric on M induced from the standard inner product on $\mathbb{R}^3$) shows that the Gaussian curvature of M is completely determined by the induced Riemannian metric on M. Thus isometric surfaces in $\mathbb{R}^3$ have the same Gaussian curvature at corresponding points of the surfaces.

(Recall that an *isometry* $\varphi\colon M_1 \to M_2$ is a diffeomorphism which preserves the lengths of all vectors.) The well-known mathematician Gauss discovered this fact, expressing it in his *Theorema Egregium* (which means 'remarkable theorem' in Latin).

Example Let us consider a surface M in $\mathbb{R}^3$ of the form $z = f(x, y)$, where f is a smooth function of x and y. Let us write

$$f_x \equiv \frac{\partial f}{\partial x}, \qquad f_y \equiv \frac{\partial f}{\partial y}$$

$$f_{xx} \equiv \frac{\partial^2 f}{\partial x^2}, \qquad f_{yy} \equiv \frac{\partial^2 f}{\partial y^2}, \qquad f_{xy} = f_{yx} \equiv \frac{\partial^2 f}{\partial x \partial y}.$$

We let X and Y be the smooth vector fields tangent to M given by

$$X = (1, 0, f_x(x, y)), \qquad Y = (0, 1, f_y(x, y)).$$

Note that (X, Y) is a smooth moving frame on M. We let ν be the unit normal vector field on M given by

$$\nu = \frac{1}{\sqrt{1 + (f_x)^2 + (f_y)^2}} \left(-f_x, -f_y, 1\right).$$

Now

$$\begin{aligned} \partial_X X &= (0, 0, f_{xx}), \\ \partial_X Y &= \partial_Y X \\ &= (0, 0, f_{xy}), \\ \partial_Y Y &= (0, 0, f_{yy}), \end{aligned}$$

therefore the second fundamental form σ is given by

$$\begin{aligned} \sigma(X, X) &= \langle \nu, \partial_X X \rangle \\ &= \frac{f_{xx}}{\sqrt{1 + (f_x)^2 + (f_y)^2}}, \\ \sigma(X, Y) &= \langle \nu, \partial_X Y \rangle \\ &= \frac{f_{xy}}{\sqrt{1 + (f_x)^2 + (f_y)^2}}, \\ \sigma(Y, X) &= \sigma(X, Y), \\ \sigma(Y, Y) &= \langle \nu, \partial_X Y \rangle \\ &= \frac{f_{yy}}{\sqrt{1 + (f_x)^2 + (f_y)^2}}. \end{aligned}$$

Now

$$\begin{aligned} \langle X, X\rangle \langle Y, Y\rangle - \langle X, Y\rangle^2 &= \left(1 + (f_x)^2\right)\left(1 + (f_y)^2\right) - (f_x)^2(f_y)^2 \\ &= 1 + (f_x)^2 + (f_y)^2. \end{aligned}$$

The Gaussian curvature of the surface M is therefore given by

$$K = \frac{f_{xx}f_{yy} - (f_{xy})^2}{\left(1 + (f_x)^2 + (f_y)^2\right)^2}.$$

Now let M be an oriented smooth surface in $\mathbb{R}^3$ and let p be a point of M. By performing a translation we can ensure that the point p is at the origin $(0, 0, 0)$ of $\mathbb{R}^3$. Moreover, by performing a rotation of the surface, we can ensure that the tangent space to M at p is given by the plane $z = 0$.

An application of the Inverse Function Theorem shows that the surface M can be represented around p by an equation of the form $z = f(x, y)$. Indeed if $\pi\colon M \to \mathbb{R}^2$ is the map sending $(x, y, z) \in M$ to (x, y) then the derivative of π at p is an isomorphism. Therefore π has a continuous inverse $\pi^{-1}\colon U \to M$ which maps some open neighbourhood U of $(0, 0)$ in $\mathbb{R}^2$ onto an open neighbourhood of p in M. If $\varphi\colon M \to \mathbb{R}$ is the restriction to M of the function that sends (x, y, z) to z then M is given by the equation $z = f(x, y)$ around p, where $f = \varphi \circ \pi^{-1}$.

Using the fact that the tangent space to M at p is the plane $z = 0$ (where p is the origin of $\mathbb{R}^3$), we see that $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$. It follows from the formula for the Gaussian curvature calculated in the example above that

$$K(p) = f_{xx}f_{yy} - (f_{xy})^2,$$

where

$$f_{xx} \equiv \frac{\partial^2 f}{\partial x^2}, \qquad f_{yy} \equiv \frac{\partial^2 f}{\partial y^2}, \qquad f_{xy} = f_{yx} \equiv \frac{\partial^2 f}{\partial x \partial y}.$$

Now Taylor's theorem implies that

$$f(x, y) = \tfrac{1}{2}\left(f_{xx}x^2 + 2f_{xy}xy + f_{yy}y^2\right) + O\left((x^2 + y^2)^{\frac{3}{2}}\right)$$

(where $O\left((x^2 + y^2)^{\frac{3}{2}}\right)$ is a quantity with the property that

$$(x^2 + y^2)^{-\frac{3}{2}} O\left((x^2 + y^2)^{\frac{3}{2}}\right)$$

remains bounded as $(x, y) \to (0, 0)$). Thus the paraboloid represented by the equation

$$z = \tfrac{1}{2}\left(f_{xx}x^2 + 2f_{xy}xy + f_{yy}y^2\right)$$

is the paraboloid which best approximates to the surface M around the point p. This paraboloid is known as the *osculating paraboloid* at the point p of M. We can find such an osculating paraboloid approximating the surface M around each point of M. The osculating paraboloid at any point of M is completely determined by the second fundamental form of M at that point. Indeed if p is a point of M then the osculating paraboloid at p is the paraboloid consisting of all points of the form

$$\{q \in \mathbb{R}^3 : q = p + X_p + \tfrac{1}{2}\sigma(X_p, X_p)\nu_p \text{ for some } X_p \in T_pM\}.$$

It follows from the theorem concerning diagonalization of symmetric matrices that, for each point p of M there exists an orthonormal basis (E_1, E_2) of the tangent space T_pM to M at p such that $\sigma(E_1, E_2) = 0$. Let us denote $\sigma(E_1, E_1)$ and $\sigma(E_2, E_2)$ by κ_1 and κ_2 respectively. The quantities κ_1 and κ_2 are known as the *principal curvatures* of M at p. The lines represented by the vectors E_1 and E_2 are referred to as the *principal directions* at p.

Now the Gaussian curvature K of M at p is equal to the determinant of the matrix

$$\begin{pmatrix} \sigma(E_1, E_1) & \sigma(E_1, E_2) \\ \sigma(E_2, E_1) & \sigma(E_2, E_2) \end{pmatrix}$$

and is therefore equal to the product $\kappa_1\kappa_2$ of the principal curvatures of M at p. The *mean curvature* H at p is defined by

$$H(p) \equiv \tfrac{1}{2}\left(\sigma(E_1, E_1) + \sigma(E_2, E_2)\right)$$

(where (E_1, E_2) is an orthonormal basis of the tangent space T_pM to M at p). If κ_1 and κ_2 are the principal curvatures at p then

$$H = \tfrac{1}{2}(\kappa_1 + \kappa_2).$$

Unlike the Gaussian curvature K, the mean curvature H does not depend simply on the induced Riemannian metric on M but depends on the manner in which the surface M is embedded in $\mathbb{R}^3$. It is not invariant under isometries (unlike the Gaussian curvature). The principal curvatures κ_1 and κ_2 at the point p are given by the formula $H \pm \sqrt{H^2 - K}$ (where K is the Gaussian curvature and H is the mean curvature at p).

The shape of the surface M around the point p approximates to the shape of the osculating paraboloid at p, which is in turn determined by the principal curvatures κ_1 and κ_2.

Consider the case when κ_1 and κ_2 are both non-zero and have the same sign, so that the second fundamental form is either positive-definite or negative-definite. This is the case if and only if $K > 0$. In this case the osculating paraboloid to the surface M at p is bowl-shaped, and it is easy to prove that all points of M sufficiently close to p all lie on one side of the tangent plane to M at p. Such a point p of M at which the Gaussian curvature K satisfies $K > 0$ is referred to as an *elliptic point* of M.

Next consider the case when κ_1 and κ_2 are non-zero and have opposite signs, so that the second fundamental form is non-degenerate and indefinite. This is the case if and only if $K < 0$. In this case the osculating paraboloid is saddle-shaped, and the surface M around p lies on both sides of the tangent plane to M at p. Such a point at which the Gaussian curvature K satisfies $K < 0$ is referred to as a *hyperbolic point* of M.

Next consider the case when exactly one of the principal curvatures κ_1 and κ_2 is non-zero, so that the second fundamental form is degenerate but is non-zero. This is the case if and only if $K = 0$ and $H \neq 0$. In this case the osculating paraboloid is trough-shaped. Such a point at which $K = 0$ and $H \neq 0$ is referred to to as a *parabolic point* of M.

The remaining case occurs when both κ_1 and κ_2 are zero, so that the second fundamental form vanishes. This is the case if and only if $K = 0$ and $H = 0$. The osculating paraboloid at p is then the tangent plane at p. Such a point is referred to as a *planar point* of M.

A point p of M is said to be an *umbilic point* if and only if $\kappa_1 = \kappa_2$. This is the case if and only if $H^2 = K$. An umbilic point is either an ellipic point or a planar point.

Example Let M be a sphere of radius R about the origin in $\mathbb{R}^3$. The Gauss map sends a point (x, y, z) of M to the point

$$\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R}\right)$$

on the unit sphere S^2. Therefore the derivative of the Gauss map $\nu\colon M \to S^2$ sends a vector (a^1, a^2, a^3) tangent to M at a point (x, y, z) of M to the vector

$$\left(\frac{a^1}{R}, \frac{a^2}{R}, \frac{a^3}{R}\right)$$

We conclude that

$$\sigma(X, Y) = -\frac{1}{R}\langle X, Y\rangle$$

(with respect to the outward normal direction) for all tangent vectors X and Y on M. The principal curvatures are equal to $-R^{-1}$ everywhere on M. The mean curvature is everywhere equal to $-R^{-1}$ and the Gaussian curvature is everywhere equal to R^{-2}. Every point on the sphere M of radius R is an elliptic point and an umbilic point.

Let M be an oriented smooth surface in $\mathbb{R}^3$ and let ν be the normal vector field on M. Let $\gamma\colon I_\gamma \to M$ be a curve in M parameterized by arclength such that $\gamma(0) = p$. The *normal curvature* of γ at $\gamma(s)$ is defined to be the quantity

$$\left\langle \nu_{\gamma(s)}, \frac{d^2\gamma(s)}{ds^2} \right\rangle$$

(where γ is parameterized by arclength). We claim that the normal curvature of γ at $\gamma(s)$ is equal to $\sigma(\gamma'(s), \gamma'(s))$, where σ is the second fundamental form. Indeed let us choose an orthonormal moving frame (E_1, E_2) around $\gamma(s)$. Then we can write

$$\frac{d\gamma(s)}{dt} = a^1(s)E_1 + a^2(s)E_2.$$

Therefore

$$\frac{d^2\gamma(s)}{dt^2} = \frac{da^1(s)}{dt}E_1 + \frac{da^2(s)}{dt}E_2 + a^1(s)\partial_{\gamma'(s)}E_1 + a^2(s)\partial_{\gamma'(s)}E_2,$$

so that

$$\begin{aligned} \left\langle \nu_{\gamma(s)}, \frac{d^2\gamma(s)}{ds^2} \right\rangle &= a^1(s)\sigma(\gamma'(s), E_1) + a^2(s)\sigma(\gamma'(s), E_2) \\ &= \sigma(\gamma'(s), \gamma'(s)). \end{aligned}$$

as required. Now $\gamma'(t)$ is a vector of unit length (since γ is parameterized by arclength). Therefore if κ_1 and κ_2 are the principal curvatures of M at $\gamma(s)$, where $\kappa_1 \leq \kappa_2$, then the value of the normal curvature of γ at $\gamma(s)$ lies between κ_1 and κ_2. More precisely there is an orthonormal basis (E_1, E_2) of the tangent space at $\gamma(s)$ such that $\kappa_1 = \sigma(E_1, E_1)$, $\kappa_2 = \sigma(E_2, E_2)$ and $\sigma(E_1, E_2) = 0$ (i.e., the vectors E_1 and E_2 point along the principal directions at $\gamma(s)$). Let us write

$$\gamma'(s) = \cos\varphi\, E_1 + \sin\varphi\, E_2$$

for some angle φ. Then the normal curvature of γ at $\gamma(s)$ is

$$\cos^2\varphi\,\kappa_1 + \sin^2\varphi\,\kappa_2.$$

6 Geodesics in Riemannian Manifolds

Let (M, g) be a Riemannian manifold of dimension n, and let $\gamma\colon I \to M$ be a smooth curve in M, defined over some interval I in $\mathbb{R}$. We say that γ is a *geodesic* if and only if

$$\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right) = 0.$$

Thus γ is a geodesic if and only if the velocity vector field $t \mapsto \gamma'(t)$ is parallel along γ (with respect to the Levi-Civita connection on M). The geodesic $\gamma: I \to M$ is said to be *maximal* if it cannot be extended to a geodesic defined over some interval J, where $I \subset J$ and $I \neq J$.

Let $\gamma: I \to M$ be a geodesic in the Riemannian manifold (M, g). Then the length $|\gamma'(t)|$ of the velocity vector $\gamma'(t)$ is constant along the curve, since

$$\frac{d}{dt}|\gamma'(t)|^2 = \frac{d}{dt} g\left(\frac{d\gamma(t)}{dt}, \frac{d\gamma(t)}{dt}\right) = g\left(\frac{D}{dt}\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\right) + g\left(\frac{d\gamma}{dt}, \frac{D}{dt}\frac{d\gamma}{dt}\right) = 0.$$

Let us choose a smooth coordinate system $(x^1, x^2, \ldots, x^n)$ over some open set U in the smooth manifold M. Let the smooth functions Γ^i_{jk} on U be characterized by the property that

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = \sum_{i=1}^n \Gamma^i_{jk} \frac{\partial}{\partial x^i}.$$

Then $\Gamma^i_{jk} = \Gamma^i_{kj}$ for all j and k, since

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} - \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} = \left[\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right] = 0.$$

Let $\gamma: I \to U$ be a smooth curve in U, and let $\gamma^i(t) = x^i \circ \gamma(t)$ for all $t \in \gamma^{-1}(U)$. Then

$$\frac{d\gamma(t)}{dt} = \sum_{k=1}^n \frac{d\gamma^k(t)}{dt} \frac{\partial}{\partial x^k},$$

so that

$$\begin{aligned} \frac{D}{dt}\frac{d\gamma(t)}{dt} &= \sum_{k=1}^n \left(\frac{d^2\gamma^k(t)}{dt^2} \frac{\partial}{\partial x^k} + \frac{d\gamma^k(t)}{dt} \sum_{j=1}^n \frac{d\gamma^j(t)}{dt} \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} \right) \\ &= \sum_{i=1}^n \left(\frac{d^2\gamma^i(t)}{dt^2} + \sum_{j=1}^n \sum_{k=1}^n \Gamma^i_{jk}(\gamma(t)) \frac{d\gamma^j(t)}{dt} \frac{d\gamma^k(t)}{dt} \right) \frac{\partial}{\partial x^i}. \end{aligned}$$

Thus $\gamma: I \to U$ is a geodesic if and only if

$$\frac{d^2\gamma^i(t)}{dt^2} + \sum_{j=1}^n \sum_{k=1}^n \Gamma^i_{jk}(\gamma(t)) \frac{d\gamma^j(t)}{dt} \frac{d\gamma^k(t)}{dt} = 0 \qquad (i = 1, 2, \ldots, n).$$

Standard existence and uniqueness theorems for solutions of ordinary differential systems of equations ensure that, given a tangent vector V at any point m of M, and given any real number t_0, there exists a unique maximal geodesic $\gamma: I \to M$, defined on some open interval I containing t_0, such that $\gamma(t_0) = m$ and $\gamma'(t_0) = V$. Moreover, for each point m of M, there exists some $\delta_m > 0$ such that every tangent vector V at m whose length is less than δ satisfies $V = \gamma_V'(0)$ for some geodesic γ_V, defined over some interval containing $[0, 1]$. The *exponential map*

$$\exp_m: \{V \in T_m M : |V| < \delta_m\} \to M$$

at m is defined by $\exp_m(V) = \gamma_V(1)$, where $\gamma_V: [0, 1] \to M$ is the unique geodesic in M satisfying $\gamma_V(0) = m$ and $\gamma_V'(0) = V$. Standard smoothness results for solutions of ordinary differential systems ensure that $\exp_m(V)$ depends smoothly on both V and m.

A *diffeomorphism* is a homeomorphism between smooth manifolds which is smooth and has a smooth inverse.

Lemma 6.1 *Let (M, g) be a Riemannian manifold, let m be a point of M and let $\exp_m: D \to M$ be the exponential map at p, defined over a neighbourhood D of zero in the tangent space $T_m M$ at m. If D is chosen sufficiently small then $\exp_m$ maps D diffeomorphically onto some open set in m. The inverse of $\exp_m |D$ therefore provides a smooth chart around the point m.*

Proof The tangent space $T_{\mathbf{0}}(T_m M)$ to $T_m M$ at the zero vector is naturally isomorphic to $T_m M$ itself. Moreover, given any $V \in T_m M$, the element of $T_{\mathbf{0}}(T_m M)$ that is mapped to V under this natural isomorphism is the velocity vector of the curve $t \mapsto tV$ at $t = 0$. But the exponential map $\exp_m$ maps the curve $t \mapsto tV$ in $T_m M$ onto the geodesic γ_V in M with initial condition $\gamma_V'(0) = V$. We deduce that the derivative of $\exp_m: D \to M$ at the zero vector sends the vector in $T_{\mathbf{0}}(T_m M)$ corresponding to V onto the vector V itself. We deduce that the derivative of the exponential map $\exp_m$ at the zero vector is an isomorphism of tangent spaces. It follows from the Inverse Function Theorem that the exponential map $\exp_m$ maps some sufficiently small open neighbourhood D of zero in $T_m M$ diffeomorphically onto an open set in M, as required. ∎

Lemma 6.2 *Let $\gamma: [a, b] \to M$ be a smooth curve in a Riemannian manifold (M, g), and let X be a smooth vector field along the curve γ. Then there exists a smooth map $\alpha: [a, b] \times (-\varepsilon, \varepsilon) \to M$ such that $\alpha(t, 0) = \gamma(t)$ and*

$$\left. \frac{\partial \alpha(t, u)}{\partial u} \right|_{u=0} = X(t)$$

for all $t \in [a, b]$. Moreover if $X(a) = 0$ and $X(b) = 0$ then the map α may be chosen such that $\alpha(a, u) = \gamma(a)$ and $\alpha(b, u) = \gamma(b)$ for all $u \in (-\varepsilon, \varepsilon)$.

Proof A map $\alpha: I \times (-\varepsilon, \varepsilon) \to M$ with the required property is given by

$$\alpha(t, u) = \exp_{\sigma(t)}(uX(t))$$

(so that, for any $t \in I$, $u \mapsto \alpha(t, u)$ is the unique geodesic in M with velocity vector $X(t)$ at $u = 0$). Standard existence and smoothness theorems for solutions of systems of ordinary differential equations can be used to show that the map α is well-defined and smooth for some sufficiently small value of ε. ∎

6.1 Length-Minimizing Curves in Riemannian Manifolds

Let (M, g) be a Riemannian manifold, and let $\gamma: [a, b] \to M$ be a smooth curve in M. The *length* $L(\gamma)$ of γ is defined by

$$L(\gamma) = \int_a^b |\gamma'(t)|\, dt,$$

where $|\gamma'(t)|^2 = g(\gamma'(t), \gamma'(t))$.

A continuous curve $\gamma: [a, b] \to M$ is said to be *piecewise smooth* if there exists a partition $\{t_0, t_1, \ldots, t_k\}$ of the interval $[a, b]$, where $a = t_0 < t_1 < \cdots < t_k = b$, such that the restriction $\gamma|[t_{i-1}, t_i]$ of γ to the interval $[t_{i-1}, t_i]$ is smooth for $i = 1, 2, \ldots, k$. The length $L(\gamma)$ of γ is the sum of the lengths of the smooth curves $\gamma|[t_{i-1}, t_i]$ for $i = 1, 2, \ldots, k$.

Lemma 6.3 *Let (M, g) be a Riemannian manifold, let $\alpha: [a, b] \times (-\varepsilon, \varepsilon) \to M$ be a smooth map, and let $\gamma: [a, b] \to M$ and $\alpha_u: [a, b] \to M$ be the smooth curves in M defined by $\gamma(t) = \alpha(t, 0)$ and $\alpha_u(t) = \alpha(t, u)$ for all $t \in [a, b]$ and $u \in (-\varepsilon, \varepsilon)$ (so that $\gamma = \alpha_0$), and let $L(\alpha_u)$ denote the length of α_u. Then*

$$\begin{aligned} \left.\frac{dL(\alpha_u)}{du}\right|_{u=0} &= \frac{1}{|\gamma'(b)|} g(\gamma'(b), X(b)) - \frac{1}{|\gamma'(a)|} g(\gamma'(a), X(a)) \\ &\quad - \int_a^b g\left(\frac{D}{dt}\left(\frac{1}{|\gamma'(t)|}\frac{d\gamma(t)}{dt}\right), X(t)\right) dt, \end{aligned}$$

where

$$X(t) = \left.\frac{\partial \alpha(t, u)}{\partial u}\right|_{u=0}.$$

In particular, if $\gamma: [a, b] \to M$ is parameterized by arclength, then

$$\left.\frac{dL(\alpha_u)}{du}\right|_{u=0} = g(\gamma'(b), X(b)) - g(\gamma'(a), X(a)) - \int_a^b g\left(\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right), X(t)\right) dt.$$

Proof Using Lemma 4.2, we see that

$$\begin{aligned}
\frac{\partial}{\partial u}\left\|\frac{\partial\alpha(t,u)}{\partial t}\right\| &= \frac{\partial}{\partial u} g\left(\frac{\partial\alpha(t,u)}{\partial t}, \frac{\partial\alpha(t,u)}{\partial t}\right)^{\frac{1}{2}} \\
&= \frac{1}{2} g\left(\frac{\partial\alpha(t,u)}{\partial t}, \frac{\partial\alpha(t,u)}{\partial t}\right)^{-\frac{1}{2}} \frac{\partial}{\partial u} g\left(\frac{\partial\alpha(t,u)}{\partial t}, \frac{\partial\alpha(t,u)}{\partial t}\right) \\
&= \left\|\frac{\partial\alpha(t,u)}{\partial t}\right\|^{-1} g\left(\frac{\partial\alpha(t,u)}{\partial t}, \frac{D}{du}\frac{\partial\alpha(t,u)}{\partial t}\right) \\
&= \left\|\frac{\partial\alpha(t,u)}{\partial t}\right\|^{-1} g\left(\frac{\partial\alpha(t,u)}{\partial t}, \frac{D}{dt}\frac{\partial\alpha(t,u)}{\partial u}\right)
\end{aligned}$$

Thus

$$\begin{aligned}
\left.\frac{dL(\alpha_u)}{du}\right|_{u=0} &= \int_a^b \frac{1}{|\gamma'(t)|} g\left(\gamma'(t), \frac{DX(t)}{dt}\right) dt \\
&= \int_a^b \frac{d}{dt}\left(\frac{1}{|\gamma'(t)|} g(\gamma'(t), X(t))\right) dt \\
&\quad - \int_a^b g\left(\frac{D}{dt}\left(\frac{1}{|\gamma'(t)|}\frac{d\gamma(t)}{dt}\right), X(t)\right) dt \\
&= \frac{1}{|\gamma'(b)|} g(\gamma'(b), X(b)) - \frac{1}{|\gamma'(a)|} g(\gamma'(a), X(a)) \\
&\quad - \int_a^b g\left(\frac{D}{dt}\left(\frac{1}{|\gamma'(t)|}\frac{d\gamma(t)}{dt}\right), X(t)\right) dt,
\end{aligned}$$

as required. ∎

Theorem 6.4 *Let p and q be distinct points in a Riemannian manifold (M, g), and let $\gamma\colon [a,b] \to M$ be a piecewise smooth curve in M from p to q, parameterized by arclength. Suppose that the length of γ is less than or equal to the length of every other piecewise smooth curve from p to q. Then γ is a smooth (unbroken) geodesic in M.*

Proof First consider the case when $\gamma\colon [a,b] \to M$ is smooth. Choose a smooth function $f\colon [a,b] \to \mathbb{R}$ for which $f(a) = 0$, $f(b) = 0$ and $f(t) > 0$ for all $t \in (a,b)$, and let X be the smooth vector field along γ given by

$$X(t) = f(t)\frac{D}{dt}\frac{d\gamma(t)}{dt}.$$

Then there exists a smooth map $\alpha\colon [a,b] \times (-\varepsilon, \varepsilon) \to M$ such that $\alpha(a,u) = p$ and $\alpha(b,u) = q$ for all u, and

$$\left.\frac{\partial\alpha(t,u)}{\partial u}\right|_{u=0} = X(t)$$

for all $t \in [a, b]$ (see Lemma 6.2). Let $\alpha_u\colon [a, b] \to M$ be given by $\alpha_u(t) = \alpha(t, u)$. It follows from Lemma 6.3 that

$$\begin{aligned} 0 &= \left.\frac{dL(\alpha_u)}{du}\right|_{u=0} = -\int_a^b g\left(\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right), X(t)\right) dt \\ &= -\int_a^b f(t)\left\|\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right)\right\|^2 dt. \end{aligned}$$

But the integrand is everywhere non-negative. We deduce that

$$f(t)\left\|\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right)\right\|^2 dt = 0,$$

for all $t \in [a, b]$, and thus

$$\frac{D}{dt}\left(\frac{d\gamma(t)}{dt}\right) = 0$$

for all $t \in [a, b]$, showing that $\gamma\colon [a, b] \to M$ is a geodesic.

Now suppose that γ is piecewise smooth. Then there exists a partition $\{t_0, t_1, \ldots, t_k\}$ of $[a, b]$, where

$$a = t_0 < t_1 < \cdots < t_k = b$$

such that $\gamma_i\colon [t_{i-1}, t_i] \to M$ is smooth for $i = 1, 2, \ldots, k$, where $\gamma_i = \gamma|[t_{i-1}, t_i]$. Now γ_i must minimize length amongst all smooth curves from $\gamma(t_{i-1})$ to $\gamma(t_i)$. It follows from the result already proved that γ_i is a geodesic for all i.

Choose $V_i \in T_{\gamma(t_i)}M$ for $i = 1, 2, \ldots, k-1$. One can then construct a continuous vector field X along γ with the properties that $X(a) = 0$, $X(b) = 0$, $X(t_i) = V_i$ for $i = 1, 2, \ldots, k-1$ and $X|[t_{i-1}, t_i]$ is smooth along γ_i. It follows, as in the proof of Lemma 6.2, that there exists a continuous map $\alpha\colon [a, b] \times (-\varepsilon, \varepsilon) \to M$ for some $\varepsilon > 0$ satisfying

$$\left.\frac{\partial\alpha(t, u)}{\partial u}\right|_{u=0} = X(t)$$

for all $t \in [a, b]$, where $\alpha|[t_{i-1}, t_i] \times (-\varepsilon, \varepsilon)$ is smooth for all i. For each $u \in (-\varepsilon, \varepsilon)$ let $\alpha_u\colon [a, b] \to M$ be the piecewise smooth curve given by $\alpha_u(t) = \alpha(t, u)$. Now γ_i is a geodesic for $i = 1, 2, \ldots, k$. It follows from Lemma 6.3 that

$$\begin{aligned} \left.\frac{dL(\alpha_u)}{du}\right|_{u=0} &= \sum_{i=1}^{k} \left(g(\gamma_i'(t_i), V_i) - g(\gamma_i'(t_{i-1}), V_{i-1})\right) \\ &= \sum_{i=1}^{k-1} g(\gamma_i'(t_i) - \gamma_{i+1}'(t_i), V_i), \end{aligned}$$

where $V_0 \in T_pM$ and $V_k \in T_qM$ are given by $V_0 = 0$ and $V_k = 0$. But $\gamma = \alpha_0$, and γ minimizes length amongst all piecewise smooth curves from p to q. Therefore

$$\sum_{i=1}^{k-1} g(\gamma_i'(t_i) - \gamma_{i+1}'(t_i), V_i) = 0$$

for all possible choices of $V_1, V_2, \ldots, V_{k-1}$, and thus $\gamma_i'(t_i) = \gamma_{i+1}'(t_i)$ for all i. We deduce that γ_i and γ_{i+1} are portions of the unique maximal geodesic in M which passes through $\gamma(t_i)$ at time $t = t_i$ with velocity vector $\gamma_i'(t_i)$. We deduce from this that the curve γ is smooth around t_i for all i, and is thus a smooth (unbroken) geodesic from p to q, as required. ∎

6.2 Geodesic Spheres and Gauss' Lemma

Let (M, g) be a Riemannian manifold, and let m be a point of M. Now Lemma 6.1 ensures that some open neighbourhood of zero in the tangent space T_mM is mapped diffeomorphically onto an open set in M under the exponential map. We define the *injectivity radius* δ_m of M at the point m as follows: if the exponential map can be defined over the whole of the tangent space T_mM at m then we set $\delta_m = +\infty$; otherwise we define δ_m to be the supremum of all positive real numbers r with the property that the exponential map $\exp_m$ maps the open ball of radius r about the origin diffeomorphically onto some open set in M. It is not difficult to verify that the open ball of radius δ_m about zero in the tangent space T_mM is itself mapped diffeomorphically onto an open set in M, where δ_m is the injectivity radius of M at m.

Let (M, g) be a Riemannian manifold, and let m be a point of M. Given any real number r satisfying $0 < r < \delta_m$, where δ_m is the injectivity radius of M at m, the *geodesic sphere* $S(m, r)$ of radius r about the point m is defined by

$$S(m, r) = \exp_m (\{V \in T_mM : \|V\| = r\}),$$

where $\|V\|^2 = g(V, V)$. (Thus $S(m, r)$ is the set of all points in M which lie at a distance r from m along some geodesic radiating from the point m.) Note that the geodesic sphere $S(m, r)$ is the image under the exponential map $\exp_m$ of the sphere of radius r about the origin in the Euclidean space T_mM. It follows from this that $S(m, r)$ is a smooth submanifold of M diffeomorphic to the $(n-1)$-sphere S^{n-1} for all sufficiently small r, where n is the dimension of M.

The following result is known as *Gauss' Lemma*. (It was proved by Gauss in his treatise *Disquisitiones generales circa superficies curvas*, published in 1827, in the case when M is a smooth surface in $\mathbb{R}^3$.)

Lemma 6.5 (Gauss' Lemma) *Let (M, g) be a Riemannian manifold, and let m be a point of M. Then, for all sufficiently small r, every geodesic passing through m intersects the geodesic sphere $S(m, r)$ orthogonally (i.e., the velocity vector of the geodesic at the point of intersection is orthogonal to the tangent space to the geodesic sphere $S(m, r)$).*

Proof Choose r such that r is less than the injectivity radius of M at the point m. Let V be a vector of length r in $T_m M$, for some sufficiently small positive number r, let $\gamma\colon [0, 1] \to M$ be the geodesic with initial conditions $\gamma(0) = m$ and $\gamma'(0) = V$, and let $q = \gamma(1)$. Then $q \in S(m, r)$.

Let W be any vector tangent to the geodesic sphere $S(m, r)$ at the point q. Now the sphere of radius r about the origin in $T_m M$ is mapped diffeomorphically onto $S(m, r)$ by the exponential map $\exp_m$. Therefore Then $W = (\exp_m \circ \sigma)'(0)$ for some smooth curve $\sigma\colon (-\varepsilon, \varepsilon) \to T_m M$ with the property that $\sigma(0) = V$ and $\|\sigma(u)\| = r$ for all $u \in (-\varepsilon, \varepsilon)$. Define $\alpha\colon [0, 1] \times (-\varepsilon, \varepsilon) \to M$ by $\alpha(t, u) = \exp_m(t\sigma(u))$ for all $t \in [0, 1]$ and $u \in (-\varepsilon, \varepsilon)$. Then $\alpha(t, 0) = \gamma(t)$ (where γ is the geodesic joining m and q), and, for each $u \in (-\varepsilon, \varepsilon)$, the curve $t \mapsto \alpha(t, u)$ is a geodesic in M joining m to $\exp_m(\sigma(u))$ whose velocity vector at $t = 0$ is $\sigma(u)$. Thus

$$\frac{D}{\partial t}\frac{\partial \alpha}{\partial t} = 0,$$

and

$$g\left(\frac{\partial \alpha(t, u)}{\partial t}, \frac{\partial \alpha(t, u)}{\partial t}\right) = g(\sigma(u), \sigma(u)) = r^2$$

for all $t \in [0, 1]$ and $u \in (-\varepsilon, \varepsilon)$, since the length of the velocity vector is constant along any geodesic. But

$$\frac{D}{\partial u}\frac{\partial \alpha}{\partial t} = \frac{D}{\partial t}\frac{\partial \alpha}{\partial u},$$

by Lemma 3.6 (since the Levi-Civita connection is torsion-free by definition). Therefore

$$\begin{aligned}
\frac{\partial}{\partial t} g\left(\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial u}\right) &= g\left(\frac{D}{\partial t}\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial u}\right) + g\left(\frac{\partial \alpha}{\partial t}, \frac{D}{\partial t}\frac{\partial \alpha}{\partial u}\right) \\
&= g\left(\frac{\partial \alpha}{\partial t}, \frac{D}{\partial t}\frac{\partial \alpha}{\partial u}\right) = g\left(\frac{\partial \alpha}{\partial t}, \frac{D}{\partial u}\frac{\partial \alpha}{\partial t}\right) \\
&= \frac{1}{2}\frac{\partial}{\partial u} g\left(\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial t}\right) = 0.
\end{aligned}$$

But $\partial\alpha/\partial u = 0$ when $t = 0$ (since $\alpha(0, u) = m$ for all u). Therefore

$$g\left(\frac{\partial\alpha}{\partial t}, \frac{\partial\alpha}{\partial u}\right) = 0$$

for all $t \in [0, 1]$ and $u \in (-\varepsilon, \varepsilon)$. In particular, if we set $t = 1$ and $u = 0$ then we deduce that $W \in T_q S(m, r)$ is orthogonal to $\gamma'(1)$, as required. ∎

Lemma 6.6 *Let (M, g) be a Riemannian manifold, let m be a point of M, and let $S(m, r)$ be the geodesic sphere of radius r about the point m. If r is chosen sufficiently small then the length $L(\sigma)$ of every piecewise smooth curve σ in M from m to a point on $S(m, r)$ satisfies $L(\sigma) \geq r$.*

Proof Let r be chosen such that r is less than the injectivity radius of M at the point m. It suffices to prove that $L(\sigma) \geq r$ for every piecewise smooth curve σ in the closed ball about m bounded by the geodesic sphere $S(m, r)$ (since it then follows that any piecewise smooth path leaving this ball must cross $S(m, r)$ and thus must have length greater than r). Let $\sigma: [a, b] \to M$ be a piecewise smooth path contained in this closed ball. Without loss of generality we may assume that $\sigma(t) \neq m$ for all $t \in (a, b]$. We can then write $\sigma(t) = \exp_m(u(t)\theta(t))$ for all $t \in [a, b]$, where $u: [a, b] \to [0, +\infty)$ is a piecewise smooth non-negative function on $[a, b]$ and $\theta: [a, b] \to T_m M$ is a piecewise smooth curve contained in the unit sphere of $T_m M$ (so that $\|\theta(t)\| = 1$ for all $t \in [a, b]$). Now

$$\frac{d}{dt}(u(t)\theta(t)) = u'(t)\theta(t) + u(t)\theta'(t),$$

But the vectors $\theta(t)$ and $u(t)\theta'(t)$ in the tangent space to $T_m M$ at $u(t)\theta(t)$ are tangent to the curves $s \mapsto s\theta(t)$ and $s \mapsto u(t)\theta(s)$ at at $s = u(t)$ and $s = t$ respectively. It follows that

$$\sigma'(t) = u'(t)N(t) + V(t),$$

where $N(t)$ is the unit vector tangent to the geodesic $s \mapsto \exp_m(s\theta(t))$ at $s = u(t)$ and $V(t)$ is tangent to the curve $s \mapsto \exp_m(u(t)\theta(s))$ at $s = t$. But $\exp_m(u(t)\theta(s))$ belongs to the geodesic sphere $S(m, u(t))$ of radius $u(t)$ about m for all s, and hence the vector $V(t)$ is tangent to the geodesic sphere $S(m, u(t))$ at $\sigma(t)$. It follows from Gauss' Lemma (Lemma 6.5) that the vectors $N(t)$ and $V(t)$ are orthogonal, so that

$$\|\sigma'(t)\|^2 = g(\sigma'(t), \sigma'(t)) = u'(t)^2 g(N(t), N(t)) + g(V(t), V(t)) \geq= |u'(t)|^2.$$

Therefore

$$L(\sigma) = \int_a^b \|\sigma'(t)\|\, dt \geq \int_a^b |u'(t)|\, dt \geq |u(b) - u(a)| = r.$$

as required. ∎

7 Complete Riemannian Manifolds

Let (M, g) be a Riemannian manifold. We say that this Riemannian manifold is *geodesically complete* if and only if every geodesic in M can be extended to a geodesic $\gamma: \mathbb{R} \to M$ defined over the whole of $\mathbb{R}$.

Note that if (M, g) is a geodesically complete Riemannian manifold then, for each point m of M, the exponential map $\exp_m: T_m M \to M$ is defined over the whole of the tangent space $T_m M$ to M at m.

Given points p and q in a connected Riemannian manifold (M, g) we define the *Riemannian distance* $d(p, q)$ from p to q to be the infimum (i.e., the greatest lower bound) of the lengths of all piecewise smooth curves starting at p and ending at q. Note that $d(p, q) \geq 0$ and $d(p, q) = d(q, p)$. The Triangle Inequality

$$d(m, q) \leq d(m, p) + d(p, q)$$

is clearly satisfied for all $m, p, q \in M$ (since every piecewise smooth curve from m to p combines with every piecewise smooth curve from p to q to give a piecewise smooth curve from m to q). If $p \neq q$ then $d(p, q) > 0$. Indeed there exists some positive real number r such that the closed ball about p bounded by the geodesic sphere $S(p, r)$ of radius r about p does not contain the point q. It follows from Lemma 6.6 that the length of every piecewise smooth curve from p to q must exceed r, since the curve must cross $S(p, r)$. We deduce that every connected Riemannian manifold (M, g) is a metric space with respect to the Riemannian distance function. Moreover it follows easily from Lemma 6.1 and Lemma 6.6 that the topology generated by this distance function is the given topology on M.

The following results are due to Hopf and Rinow.

Theorem 7.1 *Let (M, g) be a connected Riemannian manifold and let m be a point of M. Suppose that the exponential map $\exp_m: T_m M \to M$ at m is defined over the whole of the tangent space $T_m M$ to M at m (i.e., every geodesic passing through the point m can be extended to a geodesic $t \mapsto \gamma(t)$ defined for all t). Then, given any $q \in M$, there exists a geodesic from m to q which minimizes length amongst all piecewise smooth curves from m to q.*

Proof Let q be a point of M which is distinct from m, and let r denote the Riemannian distance $d(m, q)$ between the points m and q. Choose $\delta > 0$ such that δ is less than the injectivity radius of M at m and q does not belong to the closed ball about m in M bounded by the geodesic sphere $S(m, \delta)$ of radius δ about m. Then every piecewise smooth curve from m to q must cross the geodesic sphere $S(m, \delta)$. It follows from the compactness of $S(m, \delta)$

that there exists some tangent vector $V \in T_m M$ of unit length at m with the property that

$$d(\exp_m(\delta V), q) = \inf\{d(s, q) : s \in S(m, \delta)\}.$$

Let $\gamma: \mathbb{R} \to M$ be the geodesic, parameterized by arclength, defined by $\gamma(t) = \exp_m(tV)$. We claim that $\gamma(r) = q$, where $r = d(m, q)$.

Now every path from m to q must cross the geodesic sphere $S(m, \delta)$ of radius δ about m. But $d(m, q)$ is defined to be the infimum of the lengths of all smooth curves from m to q. Therefore

$$d(m, q) = \delta + \inf\{d(p, q) : p \in S(m, \delta)\} = \delta + d(\gamma(\delta), q)$$

(since every piecewise smooth curve from m to q must cross $S(m, \delta)$, and $d(m, p) = \delta$ for all $p \in S(m, \delta)$). Thus $d(\gamma(\delta), q) = r - \delta$, where $r = d(m, q)$. Let

$$u = \sup\{t \in [\delta, r] : d(\gamma(t), q) = r - t\},$$

where $r = d(m, q)$. We shall show that $u = r$.

Now follows from the continuity of the Riemannian distance function that $d(\gamma(u), q) = r - u$. Suppose that it were the case that $u < r$. We show that this would lead to a contradiction. For if $u < r$ then we could find some sufficiently small $\eta > 0$ with the property that every path from $\gamma(u)$ to q must cross the geodesic sphere $S(\gamma(u), \eta)$ of radius η about $\gamma(u)$. But then there would exist $p \in S(\gamma(u), \eta)$ such that

$$d(p, q) = \inf\{d(p', q) : p' \in S(\gamma(u), \eta)\},$$

since $S(\gamma(u), \eta)$ is compact. Moreover

$$d(\gamma(u), q) = \eta + \inf\{d(p', q) : p' \in S(\gamma(u), \eta)\} = \eta + d(p, q),$$

so that $d(p, q) = r - u - \eta$. Thus

$$r = d(m, q) \leq d(m, p) + d(p, q) \leq d(m, p) + r - u - \eta,$$

and hence $d(m, p) \geq u + \eta$. But if $\alpha: [0, u + \eta] \to M$ is the piecewise smooth curve from m to p, parameterized by arclength, obtained by concatenating the geodesic γ from m to $\gamma(u)$ with the geodesic of length η from $\gamma(u)$ to p, then the length of the curve α is $u + \eta$, and therefore α minimized length amongst all piecewise smooth curves from m to p. It follows from Lemma 6.4 that α is a smooth (unbroken) geodesic, so that $\alpha(t) = \gamma(t)$ for all $t \in [0, u + \eta]$. But then $p = \gamma(u + \eta)$ so that $d(\gamma(u + \eta), q) = r - u - \eta$, contradicting the definition of u. Thus excludes the possibility that $u < r$. Therefore $u = r$, and hence $d(\gamma(r), q) = 0$. Thus $q = \gamma(r)$, as required. ∎

Corollary 7.2 *Let (M, g) be a connected Riemannian manifold. Then the following three conditions are equivalent:*

(i) *the Riemannian distance function on M is complete (i.e., every Cauchy sequence in M converges),*

(ii) *the Riemannian manifold (M, g) is geodesically complete,*

(iii) *there exists a point m of M with the property that the exponential map $\exp_m$ is defined over the whole of the tangent space $T_m M$ to M at m (i.e., every geodesic passing through the point m can be extended to a geodesic from $\mathbb{R}$ into M).*

Proof First we show that (i) implies (ii). Thus suppose that the Riemannian distance function of M is complete. Let $\gamma: I \to M$ be a maximal geodesic in M, defined on an open interval I in $\mathbb{R}$. We must show that $I = \mathbb{R}$. Suppose that it were the case the the interval I is bounded above. Let $u = \sup I$. Then we could find an ascending sequence $(t_j : j \in \mathbb{N})$ of elements of I which converges to u. Now the length $\|\gamma'(t)\|$ of the velocity vector $\gamma'(t)$ of γ is constant along the geodesic. Let $\|\gamma'(t)\| = C$. Then $d(\gamma(t_j), \gamma(t_k)) \leq C(t_k - t_j)$ for all natural numbers j and k (where $d(\gamma(t_j), \gamma(t_k))$ is the Riemannian distance from t_j to t_k). But the sequence $(t_j : j \in \mathbb{N})$ is a Cauchy sequence in $\mathbb{R}$. It follows that the sequence $(\gamma(t_j) : j \in \mathbb{N})$ is a Cauchy sequence in M and therefore converges to some point m of M, since M is complete with respect to the Riemannian distance function.

Now it follows from standard existence theorems for solutions of differential equations, applied to the second order system characterizing geodesics in M with respect to some smooth coordinate system, that there exists some open neighbourhood U of m and some $\delta > 0$ such that $\exp_q(V)$ is well-defined for all $q \in U$ and $V \in T_q M$ satisfying $\|V\| < \delta$. Let j be chosen large enough to ensure that $t_j \in U$ and $u - t_j < \varepsilon$, where $\varepsilon = \delta/(2C)$. Then $\|(t-t_j)\gamma'(t_j)\| < \delta$ for all $t \in (t_j - \varepsilon, t_j + \varepsilon)$. Thus if $\sigma(t) = \exp_q((t-t_j)\gamma'(t_j))$ for all $t \in (t_j - \varepsilon, t_j + \varepsilon)$ then $\sigma: (t_j - \varepsilon, t_j + \varepsilon)$ is a geodesic in M, and $\sigma'(t_j) = \gamma'(t_j)$. We deduce that $\sigma(t) = \gamma(t)$ for all t sufficiently close to t_j. It follows from the maximality of γ that $(t_j - \varepsilon, t_j + \varepsilon) \subset I$ (since otherwise we could extend γ be defining $\gamma(t) = \sigma(t)$ for all $t \in (t_j - \varepsilon, t_j + \varepsilon) \setminus I$). But this is impossible since $t_j + \varepsilon > u$, where $u = \sup I$. We deduce that the interval I is not bounded above. Similarly the interval I is not bounded below. Therefore $I = \mathbb{R}$. This shows that (i) implies (ii).

A Riemannian manifold (M, g) is geodesically complete if and only if, for every $m \in M$, the exponential map $\exp_m: T_m M \to M$ is defined over the whole of the tangent space $T_m M$ at m. Thus (ii) implies (iii).

Finally we show that (iii) implies (i). Suppose that (iii) holds. Let m be a point of M with the property that the exponential map $\exp_m$ is defined over the whole of the tangent space $T_m M$ to M at m. Let $(m_j : j \in \mathbb{N})$ be a Cauchy sequence in M. Then $d(m, m_j)$ remains bounded as $j \to +\infty$. Therefore there exists some $R > 0$ with the property that $d(m, m_j) \leq R$. Now it follows from Theorem 7.1 that each point m_j can be joined to m by a geodesic of length $d(m, m_j)$. Therefore $m_j \in B(m, R)$ for all j, where

$$B(m, R) = \exp_m \left(\{ V \in T_m M : |V| \leq R \} \right).$$

Now $B(m, R)$ is compact, since it is the image of a compact subset of $T_m M$ under the continuous map $\exp_m \colon T_m M \to M$. Moreover any compact metric space is complete. Therefore the Cauchy sequence m_1, m_2, m_3 converges (in M) to some point of $B(m, R)$. This shows that M is complete. Thus (iii) implies (i). Thus conditions (i), (ii) and (iii) are equivalent, as required. ∎

We see that a connected Riemannian manifold is geodesically complete if and only if it is a complete metric space with respect to the Riemannian distance function. Such a Riemannian manifold is said to be *complete*.

7.1 Local Isometries and Covering Maps

Definition Let (M_1, g_1) and (M_2, g_2) be Riemannian manifolds of the same dimension, and let $\varphi \colon M_1 \to M_2$ be a smooth map. We say that φ is a *local isometry* if and only if

$$g_2(\varphi_* X, \varphi_* Y) = g_1(X, Y)$$

for all $m \in M_1$ and $X, Y \in T_m M_1$, where $\varphi_* \colon T_m M_1 \to T_{\varphi(m)} M_2$ is the derivative of the map φ at m. A local isometry $\varphi \colon M_1 \to M_2$ is said to be an *isometry* if it is also a diffeomorphism from M_1 to M_2.

Observe that if $\varphi \colon M_1 \to M_2$ is a local isometry then $\|\varphi_* X\| = \|X\|$ for all tangent vectors X. It follows that the derivative $\varphi_* \colon T_m M_1 \to T_{\varphi(m)} M_2$ of φ at each point m of M_1 is an injective homomorphism between vector spaces of the same dimension and is thus an isomorphism of vector spaces. It follows from the Inverse Function Theorem that φ maps some open neighbourhood of each point m of M_1 diffeomorphically onto an open neighbourhood of $\varphi(m)$ in M_2. Moreover the Levi-Civita connection and the Riemann curvature tensor around a point in a Riemannian manifold is completely determined by the Riemannian metric around that point. It follows that the Levi-Civita connections and Riemann curvature tensors of the Riemannian manifolds

(M_1, g_1) and (M_2, g_2) correspond under any local isomorphism $\varphi\colon M_1 \to M_2$. Moreover a smooth curve $\gamma\colon I \to M_1$ is a geodesic in M_1 if and only if $\varphi\colon \gamma\colon I \to M_2$ is a geodesic in M_2.

We now recall the definition of covering maps. Let $\tilde{X}$ and X be topological spaces and let $\varphi\colon \tilde{X} \to X$ be a continuous map. An open set U in X is said to be *evenly covered* by the map $\varphi\colon \tilde{X} \to X$ if the preimage $\varphi^{-1}(U)$ of U is a disjoint union of open sets in $\tilde{X}$, each of which is mapped homeomorphically onto U by the map φ. The map $\varphi\colon \tilde{X} \to X$ is said to be a *covering map* if it is surjective and each point of X has an open neighbourhood evenly covered by the map φ. If the topological space X is simply-connected then any covering map $\varphi\colon \tilde{X} \to X$ is a homeomorphism.

Theorem 7.3 *Let (M, g) and $(\tilde{M}, \tilde{g})$ be connected (non-empty) Riemannian manifolds of the same dimension and let $\varphi\colon \tilde{M} \to M$ be a local isometry. Suppose that $\tilde{M}$ is complete. Then M is complete and the map $\varphi\colon \tilde{M} \to M$ is a covering map. In particular, $\varphi\colon \tilde{M} \to M$ is surjective.*

Proof We first prove the following lifting property for geodesics in M: given any geodesic $\gamma\colon I \to M$, defined on some interval I in $\mathbb{R}$, and given $t_0 \in I$ and $q \in \tilde{M}$ satisfying $\varphi(q) = \gamma(t_0)$, there exists a unique geodesic $\tilde{\gamma}\colon I \to \tilde{M}$ such that $\tilde{\gamma}(t_0) = q$ and $\varphi \circ \tilde{\gamma} = \gamma$. Indeed the derivative φ_* of φ at q is an isomorphism of vector spaces, and hence there exists a unique tangent vector $V \in T_q\tilde{M}$ at q satisfying $\varphi_* V = \gamma'(t_0)$. Moreover there exists a geodesic $\eta\colon \mathbb{R} \to \tilde{M}$ in $\tilde{M}$ satisfying $\eta(t_0) = q$ and $\eta'(t_0) = V$, since $\tilde{M}$ is complete. Then $\varphi \circ \eta\colon \mathbb{R} \to M$ is a geodesic in M, and $(\varphi \circ \eta)'(t_0) = \gamma'(t_0)$. But the geodesic γ is uniquely determined by $\gamma'(t_0)$. We deduce that $\gamma = (\varphi \circ \eta)|I$. Thus if we define $\tilde{\gamma}\colon I \to \tilde{M}$ to be the restriction $\eta|I$ of $\eta\colon \mathbb{R} \to M$ to the interval I then $\tilde{\gamma}$ is the required lift of the geodesic γ to $\tilde{M}$. We deduce also that any geodesic passing through a point of the image $\varphi(\tilde{M})$ of the map $\varphi\colon \tilde{M} \to M$ can be extended to a geodesic defined over the whole of $\mathbb{R}$. We deduce from Theorem 7.1 that the exponential map $\exp_m\colon T_m M \to M$ of M at a point m of $\varphi(M)$ is surjective, and, since every geodesic passing through m is of the form $\varphi \circ \tilde{\gamma}$ for some geodesic $\tilde{\gamma}$ in $\tilde{M}$, it follows that $\varphi(\tilde{M}) = M$. We also deduce that every geodesic in M can be extended to a geodesic defined over the whole of $\mathbb{R}$, so that M is complete.

Let m be a point in M. It follows from Lemma 6.1 that there exists some $\delta_m > 0$ such that the exponential map $\exp_m$ at m maps the open ball of radius δ about the zero vector in the tangent space $T_m M$ to M at m diffeomorphically onto an open neighbourhood U of m. Then every point of U can be joined to m by a unique geodesic in U of length less than δ_m. We show that U is evenly covered by the map φ.

Let p be an element of $\tilde{M}$ satisfying $\varphi(p) = m$ and let U_p be the subset of $\varphi^{-1}(U)$ consisting of all points that can be joined to p by a geodesic contained wholly in $\varphi^{-1}(U)$. Now if $\tilde{\gamma}$ is any geodesic in $\varphi^{-1}(U)$ joining p to a point q in U_p then $\varphi \circ \tilde{\gamma}$ is a geodesic in U joining m to $\varphi(q)$. Conversely, given any geodesic γ in U joining m to a point u of U, there exists a unique geodesic $\tilde{\gamma}$ in U_p joining p to some point q of U_p satisfying $\varphi(q) = u$. Moreover there is exactly one point q of U_p satisfying $\varphi(q) = u$, since the geodesic γ is the unique geodesic in U joining m to u, and therefore $\tilde{\gamma}$ is the unique geodesic in U_p joining p to a point of $\varphi^{-1}(\{u\})$. We deduce that U_p is mapped bijectively on U under the map φ. But $\varphi\colon \tilde{M} \to M$ is a local diffeomorphism. It follows easily that U_p is mapped diffeomorphically onto U under the map φ.

Now let q be any element of $\varphi^{-1}(U)$. Then $\varphi(q)$ can be joined to m by a unique geodesic γ contained wholly in U. The geodesic γ lifts to a geodesic $\tilde{\gamma}$ in $\varphi^{-1}(U)$ joining q to some element p of $\tilde{M}$ satisfying $\varphi(p) = m$, where $\varphi \circ \tilde{\gamma} = \gamma$. Moreover p is uniquely determined by q, since any geodesic in $\varphi^{-1}(U)$ joining q to an element of $\varphi^{-1}(\{m\})$ is mapped under φ to the unique geodesic in U joining $\varphi(q)$ to m. We deduce that $\varphi^{-1}(U)$ is the disjoint union of the sets U_p, as p ranges over the elements of $\varphi^{-1}(\{m\})$, and thus U is evenly covered by the map φ. This shows that $\varphi\colon \tilde{M} \to M$ is a covering map, as required. ∎

We recall that if X is a simply-connected topological space then any covering map $\varphi\colon \tilde{X} \to X$ is a homeomorphism. Also any local isometry is a local diffeomorphism, and therefore any local isometry which also a homeomorphism is a diffeomorphism, and is thus an isometry of Riemannian manifolds. The follows result therefore follows directly from Theorem 7.3.

Corollary 7.4 *Let (M, g) and $(\tilde{M}, \tilde{g})$ be connected (non-empty) Riemannian manifolds of the same dimension and let $\varphi\colon \tilde{M} \to M$ be a local isometry. Suppose that $\tilde{M}$ is complete and that M is simply-connected. Then the map $\varphi\colon \tilde{M} \to M$ is an isometry of Riemannian manifolds.*

8 Jacobi Fields

Let (M, g) be a Riemannian manifold, and let $\gamma\colon [a, b] \to M$ be a geodesic in M. A *Jacobi field* along γ is a vector field V along γ which satisfies the *Jacobi equation*

$$\frac{D^2V(t)}{dt^2} = R(\gamma'(t), V(t))\gamma'(t),$$

where R denotes the curvature tensor of the Levi-Civita connection on M. First we show that Jacobi fields arise naturally from variations of the geodesic γ through neighbouring geodesics.

Lemma 8.1 *Let $\gamma: I \to M$ be a geodesic in a Riemannian manifold (M, g) and let*

$$\alpha: I \times (-\varepsilon, \varepsilon) \to M$$

be a smooth map satisfying $\alpha(t, 0) = \gamma(t)$ for all $t \in I$. Let V be the vector field along the geodesic γ defined by

$$V(t) = \left.\frac{\partial \alpha(t, u)}{\partial u}\right|_{u=0}.$$

Suppose that, for each $u \in (-\varepsilon, \varepsilon)$, the curve $t \mapsto \alpha(t, u)$ is a geodesic in M. Then the vector field V satisfies the Jacobi equation

$$\frac{D^2 V(t)}{dt^2} = R(\gamma'(t), V(t))\gamma'(t).$$

Proof First we note that

$$\frac{D}{\partial t}\frac{\partial \alpha}{\partial t} = 0,$$

since each curve $t \mapsto \alpha(t, u)$ is a geodesic Also

$$\frac{D}{\partial t}\frac{D}{\partial u}\frac{\partial \alpha}{\partial t} - \frac{D}{\partial u}\frac{D}{\partial t}\frac{\partial \alpha}{\partial t} = R\Big(\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial u}\Big)\frac{\partial \alpha}{\partial t}$$

and

$$\frac{D}{\partial t}\frac{\partial \alpha}{\partial u} = \frac{D}{\partial u}\frac{\partial \alpha}{\partial t}$$

by Lemma 3.6, using the fact that the Levi-Civita connection is torsion-free. Therefore

$$\begin{aligned}\frac{D^2}{\partial t^2}\frac{\partial \alpha}{\partial u} &= \frac{D}{\partial t}\frac{D}{\partial u}\frac{\partial \alpha}{\partial t} \\ &= R\Big(\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial u}\Big)\frac{\partial \alpha}{\partial t} + \frac{D}{\partial u}\frac{D}{\partial t}\frac{\partial \alpha}{\partial t} \\ &= R\Big(\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial u}\Big)\frac{\partial \alpha}{\partial t}.\end{aligned}$$

Now

$$\left.\frac{\partial \alpha(t, u)}{\partial t}\right|_{u=0} = \gamma'(t), \qquad \left.\frac{\partial \alpha(t, u)}{\partial u}\right|_{u=0} = V(t).$$

Thus, on setting $u = 0$, we deduce that

$$\frac{D^2 V(t)}{dt^2} = R(\gamma'(t), V(t))\gamma'(t),$$

as required. ∎

The next result shows how Jacobi fields can be used to study the derivative of the exponential map.

Lemma 8.2 *Let (M, g) be a Riemannian manifold, let $\gamma\colon [0,1] \to M$ be a geodesic in M defined on the interval $[0,1]$, let $m = \gamma(0)$, and let $X = \gamma'(0)$. Then $\exp_{m*}(Y_X) = V(1)$ for any $Y \in T_m M$, where*

$$\exp_{m*}\colon T_X(T_m M) \to T_{\gamma(1)} M$$

is the derivative of the exponential map at X, $Y_X \in T_X(T_m M)$ is the velocity vector of the curve $u \mapsto X + uY$ at $u = 0$, and V is the Jacobi field along γ satisfying the initial conditions

$$V(0) = 0, \qquad \left.\frac{DV(t)}{dt}\right|_{t=0} = Y.$$

Proof Let $\alpha\colon [0,1] \times (-\varepsilon, \varepsilon) \to M$ be the map defined by $\alpha(t, u) = \exp_m(tX + tuY)$ for all $t \in [0,1]$ and $u \in (-\varepsilon, \varepsilon)$, where $\varepsilon > 0$ is chosen small enough to ensure that the map α is well-defined. For fixed u, the curve $t \mapsto \alpha(t,u)$ is a geodesic in M. It follows from Lemma 8.1 that the vector field V along γ defined by

$$V(t) = \left.\frac{\partial \alpha(t,u)}{\partial u}\right|_{u=0}$$

is a Jacobi field along γ. Moreover

$$\left.\frac{DV(t)}{dt}\right|_{t=0} = \left.\frac{D}{\partial t}\frac{\partial \alpha}{\partial u}\right|_{(t,u)=(0,0)} = \left.\frac{D}{\partial u}\frac{\partial \alpha}{\partial t}\right|_{(t,u)=(0,0)} = \left.\frac{D}{\partial u}(X + uY)\right|_{u=0} = Y,$$

(where we have used Lemma 3.6), and

$$\begin{aligned} V(1) &= \left.\frac{d\alpha(1,u)}{du}\right|_{u=0} = \left.\frac{d}{du}\exp_m(X + uY)\right|_{u=0} \\ &= \exp_{m*}\left(\left.\frac{d}{du}(X + uY)\right|_{u=0}\right) = \exp_{m*}(Y_X), \end{aligned}$$

as required. ∎

8.1 Flat Riemannian manifolds

Now let us consider the case when the Riemann curvature tensor of a Riemannian manifold (M, g) vanishes everywhere on M. Such a Riemannian manifold is said to be *flat*.

The tangent space $T_m M$ to a Riemannian manifold (M, g) has a natural flat Riemannian metric corresponding to the inner product on $T_m M$ given by the Riemannian metric.

Theorem 8.3 *Let (M, g) be a flat Riemannian manifold, and let m be a point of M. Then the exponential map $\exp_m: U \to M$ is a local isometry, where $U \subset T_m M$ is the domain of the exponential map $\exp_m$ at the point m.*

Proof Let X be an element of U, and let Y and Z be tangent vectors at m. Let $\gamma: \mathbb{R} \to M$ be the geodesic defined by $\gamma(t) = \exp_m(tX)$ for all $t \in \mathbb{R}$, and let fields $t \mapsto Y(t)$ and $t \mapsto Z(t)$ be the parallel vector fields along γ satisfying $Y(0) = Y$, $Z(0) = Z$,

$$\frac{DY(t)}{dt} = 0 \text{ and } \frac{DZ(t)}{dt} = 0.$$

Then

$$\frac{D^2}{dt^2}(tY(t)) = 0, \qquad \frac{D^2}{dt^2}(tZ(t)) = 0.$$

But the curvature tensor of M is zero everywhere on M, since M is flat, and therefore the vector fields $t \mapsto tY(t)$ and $t \mapsto tZ(t)$ satisfy the Jacobi equation along γ. Thus $\exp_{m*} Y_X = Y(1)$ and $\exp_{m*} Z_X = Z(1)$ by Lemma 8.2, where $Y_X \in T_X(T_m M)$ and $Z_X \in (T_m M)$ are tangent to the curves $u \mapsto X + uY$ and $u \mapsto X + uZ$ at $u = 0$. But

$$\frac{d}{dt} g(Y(t), Z(t)) = g\Big(\frac{D}{dt} Y(t), Z(t)\Big) + g\Big(Y(t), \frac{D}{dt} Z(t)\Big) = 0,$$

and hence

$$g(\exp_* Y_X, \exp_* Z_X) = g(Y(1), Z(1)) = g(Y, Z) = \langle Y_X, Z_X \rangle,$$

where $\langle ., . \rangle$ denotes the inner product on the tangent spaces of $T_m M$ defined by the Riemannian metric g at the point m. Thus the derivative of the exponential map at $X \in T_m M$ is an isometry from $T_X(T_m M)$ to $T_{\gamma(1)} M$. Thus $\exp_m: U \to M$ is a local isometry. ∎

Corollary 8.4 *Let (M, g) be a complete simply-connected Riemannian manifold of dimension n. Suppose that M is flat. Then M is isometric to the Euclidean space $\mathbb{R}^n$.*

Proof If M is complete and simply-connected then the exponential map $\exp_m$ is defined over the whole of $T_m M$ and is a local isometry. Moreover $T_m M$ is complete (since any Euclidean space is a complete Riemannian manifold). It follows from Corollary 7.4 that $\exp_m: T_m M \to M$ is an isometry, so that M is isometric to n-dimensional Euclidean space. ∎

8.2 The Cartan-Hadamard Theorem

Let (M, g) be a Riemannian manifold, let m be a point of M, and let P be a 2-dimensional subspace of the tangent space $T_m M$ to M at m. We recall that the *sectional curvature* $K(P)$ in the plane P is given terms of the curvature tensor of M by the formula

$$K(P) = R(E_1, E_2, E_1, E_2) = g(E_1, R(E_1, E_2)E_2),$$

where (E_1, E_2) is any orthonormal basis for P. We conclude that the sectional curvatures of the Riemannian manifold (M, g) are all non-positive at a point m if and only if $g(X, R(X, Y)Y) \leq 0$ for all tangent vectors X and Y at m.

Theorem 8.5 (Cartan-Hadamard) *Let (M, g) be a complete simply-connected Riemannian manifold. Suppose that the sectional curvatures $K(P)$ of M all satisfy $K(P) \leq 0$. Then M is diffeomorphic to $\mathbb{R}^n$.*

Proof Choose a point m of M. Then the exponential map $\exp_m: T_m M \to M$ at m is defined over the whole of the tangent space of M at m, since (M, g) is complete. Given $X \in T_m M$, let $\gamma: \mathbb{R} \to M$ be the geodesic in M given by $\gamma(t) = \exp_m(tX)$ for all $t \in \mathbb{R}$, and let $t \mapsto V(t)$ be a Jacobi field along γ. satisfying

$$V(0) = 0, \qquad DV(t)/dt|_{t=0} \neq 0.$$

We show that $t \mapsto \|V(t)\|$ is an increasing function of t for $t = 0$, so that $V(t) \neq 0$ for all $t > 0$. Now $t \mapsto V(t)$ satisfies the Jacobi equation

$$\frac{D^2 V(t)}{dt^2} = R(\gamma'(t), V(t))\gamma'(t) = -R(V(t), \gamma'(t))\gamma'(t),$$

and hence

$$\begin{aligned} \frac{d^2}{dt^2}\|V(t)\|^2 &= 2\frac{d}{dt} g\left(V(t), \frac{DV(t)}{dt}\right) \\ &= 2g\left(\frac{DV(t)}{dt}, \frac{DV(t)}{dt}\right) + 2g\left(V(t), \frac{D^2V(t)}{dt^2}\right) \\ &= 2\left\|\frac{DV(t)}{dt}\right\|^2 - 2g(V(t), R(V(t), \gamma'(t))\gamma'(t)). \end{aligned}$$

But $g(V(t), R(V(t), \gamma'(t))\gamma'(t) \leq 0$ for all t, since the sectional curvatures of M are all non-positive. Thus

$$\frac{d^2}{dt^2}\|V(t)\|^2 \geq 2\left\|\frac{DV(t)}{dt}\right\|^2 \geq 0$$

for all t. But there exist values of t arbitrarily close to zero for which $d/dt\|V(t)\|^2 > 0$, since $V(0) = 0$ and $V(t) \neq 0$ for all sufficiently small non-zero values of t. We deduce that $t \mapsto \|V(t)\|$ is an increasing function for $t > 0$, and thus $\|V(t)\| > 0$ for all $t > 0$. Now it follows from Lemma 8.2 that the derivative $\exp_{m*}$ of $\exp_m: T_m M \to M$ at X sends Y_X to $V(1)$, where

$$Y = \left.\frac{DV(t)}{dt}\right|_{t=0},$$

and $Y_X \in T_X(T_m M)$ is the velocity to $u \mapsto X + uY$ at $u = 0$. We deduce that $\exp_{m*}: T_X(T_m M) \to T_{\exp_m(X)} M$ is injective and is thus an isomorphism of vector spaces.

Let $\tilde{g}$ be the Riemannian metric on $T_m M$ defined at $X \in T_m M$ by $\tilde{g}(Y, Z) = g(\exp_{m*} Y, g(\exp_{m*} Z)$ for all $Y, Z \in T_X(T_m M)$. The metric tensor $\tilde{g}$ is positive-definite at X since the derivative $\exp_{m*}$ of $\exp_m$ at X is an isomorphism of vector spaces. Then $\exp_m: T_m M \to M$ is a local isometry between the Riemannian manifolds $(T_m, \tilde{g})$ and (M, g).

We claim that the Riemannian manifold $(T_m M, \tilde{g})$ is complete. Now a curve $\tilde{\gamma}$ is a geodesic in $T_m M$ if and only if $\exp_m \circ \gamma$ is a geodesic in M. But the exponential map $\exp_m$ sends straight lines through the origin in $T_m M$ to geodesics in M. We conclude that every geodesic through the origin in the Riemannian manifold $(T_m M, \tilde{g})$ can be extended to a geodesic defined over the whole of the real line $\mathbb{R}$. It follows from Corollary 7.2 that $(T_m M, \tilde{g})$ is complete. But M is simply-connected, by hypothesis, and any local isometry from a complete Riemannian manifold to a simply-connected Riemannian manifold is an isometry (see Corollary 7.4), and is thus a diffeomorphism. We conclude that M is diffeomorphic to the Euclidean space $T_m M$, as required. ∎

8.3 The Second Variation of Energy

Let (M, g) be a Riemannian manifold and let $\gamma: [a, b] \to M$ be a geodesic in M. Let $\alpha: [a, b] \times (-\varepsilon, \varepsilon) \to M$ be a smooth map with the properties that

$$\begin{aligned} \alpha(t, 0) &= \gamma(t) \text{ for all } t \in [a, b], \\ \alpha(a, u) &= \gamma(a) \text{ for all } u \in (-\varepsilon, \varepsilon), \\ \alpha(b, u) &= \gamma(b) \text{ for all } u \in (-\varepsilon, \varepsilon). \end{aligned}$$

Thus if $\alpha_u: [a, b] \to M$ is the smooth curve defined by $\alpha_u(t) = \alpha(t, u)$ then each α_u starts at $\gamma(a)$ and ends at $\gamma(b)$. We calculate

$$\left.\frac{d^2 E(\gamma(\alpha_u)}{du^2}\right|_{u=0},$$

where $E(\alpha_u)$ is the energy of α_u. Let X and Y be the smooth vector fields along the map α defined by

$$X(t,u) = \frac{\partial \alpha(t,u)}{\partial t}, Y(t,u) = \frac{\partial \alpha(t,u)}{\partial u}.$$

Note that $Y(a,u) = 0$ and $Y(b,u) = 0$ for all $u \in (-\varepsilon, \varepsilon)$, on account of the fact that $\alpha(a,u) = \gamma(a)$ and $\alpha(b,u) = \gamma(b)$. The energy of the curve α_u is given by

$$E(\alpha_u) = \tfrac{1}{2}\int_a^b g(X(t,u), X(t,u))\, dt.$$

Now

$$\frac{DX}{\partial u} = \frac{D}{\partial u}\frac{\partial \alpha}{\partial t} = \frac{D}{\partial t}\frac{\partial \alpha}{\partial u} = \frac{DY}{\partial t}$$

by Lemma 3.6. Thus

$$\frac{dE(\alpha_u)}{du} = \int_a^b g\left(X, \frac{DX}{\partial u}\right) dt = \int_a^b g\left(X, \frac{DY}{\partial t}\right) dt,$$

hence

$$\begin{aligned} \frac{d^2E(\alpha_u)}{du^2} &= \int_a^b \left(g\left(\frac{DX}{\partial u}, \frac{DY}{\partial t}\right) + g\left(X, \frac{D}{\partial u}\frac{DY}{\partial t}\right)\right) dt \\ &= \int_a^b \left(g\left(\frac{DY}{\partial t}, \frac{DY}{\partial t}\right) + g\left(X, \frac{D}{\partial u}\frac{DY}{\partial t}\right)\right) dt. \end{aligned}$$

But

$$\frac{D}{\partial u}\frac{DY}{\partial t} = \frac{D}{\partial t}\frac{DY}{\partial u} + \mathcal{R}(Y,X)Y$$

by Lemma 3.6. Therefore

$$\begin{aligned} \frac{d^2E(\alpha_u)}{du^2} &= \int_a^b g\left(\frac{DY}{\partial t}, \frac{DY}{\partial t}\right) dt \\ &\quad + \int_a^b g\left(X, \frac{D}{\partial t}\frac{DY}{\partial u} + \mathcal{R}(Y,X)Y\right) dt. \end{aligned}$$

But

$$\begin{aligned} \int_a^b g\left(X, \frac{D}{\partial t}\frac{DY}{\partial u}\right) dt &= \int_a^b \frac{\partial}{\partial t}\left(g\left(X, \frac{DY}{\partial u}\right)\right) dt - \int_a^b g\left(\frac{DX}{\partial t}, \frac{DY}{\partial u}\right) dt \\ &= g\left(X(b,u), \frac{DY(b,u)}{\partial u}\right) - g\left(X(a,u), \frac{DY(a,u)}{\partial u}\right) \\ &\quad - \int_a^b g\left(\frac{DX}{\partial t}, \frac{DY}{\partial u}\right) dt \\ &= -\int_a^b g\left(\frac{DX}{\partial t}, \frac{DY}{\partial u}\right) dt, \end{aligned}$$

because $Y(a,u)=0$ and $Y(b,u)=0$ for all $u\in(-\varepsilon,\varepsilon)$. Thus

$$\begin{aligned}\frac{d^2E(\alpha_u)}{du^2} &= \int_a^b \left(g\left(\frac{DY}{\partial t},\frac{DY}{\partial t}\right)+g(X,\mathcal{R}(Y,X)Y)\right)\,dt \\ &\quad -\int_a^b g\left(\frac{DX}{\partial t},\frac{DY}{\partial u}\right)\,dt.\end{aligned}$$

Now let us set $u=0$. We define the vector field V along γ by

$$V(t)=Y(t,0)=\left.\frac{\partial\alpha(t,u)}{\partial u}\right|_{u=0}.$$

Note that $X(t,0)=\gamma'(t)$ and

$$\frac{DX(t,0)}{dt}=\frac{D\gamma'(t)}{dt}=0$$

(since γ is a geodesic. Therefore

$$\begin{aligned}&\left.\frac{d^2E(\alpha_u)}{du^2}\right|_{u=0}\\ &= \int_a^b \left(g\left(\frac{DV(t)}{\partial t},\frac{DV(t)}{\partial t}\right)+g(\gamma'(t),\mathcal{R}(V(t),\gamma'(t))V(t))\right)\,dt\\ &= \int_a^b \left(g\left(\frac{DV(t)}{\partial t},\frac{DV(t)}{\partial t}\right)+R(\gamma'(t),V(t),V(t),\gamma'(t))\right)\,dt.\end{aligned}$$

We can integrate the first term in this formula by parts. Using the fact that $V(a)=0$ and $V(b)=0$ we see that

$$\int_a^b g\left(\frac{DV(t)}{\partial t},\frac{DV(t)}{\partial t}\right)\,dt=-\int_a^b g\left(V(t),\frac{D^2V(t)}{\partial t^2}\right)\,dt$$

Also

$$\begin{aligned}R(\gamma'(t),V(t),V(t),\gamma'(t)) &= -R(V(t),\gamma'(t),V(t),\gamma'(t))\\ &= R(V(t),\gamma'(t),\gamma'(t),V(t))\\ &= g(V(t),\mathcal{R}(\gamma'(t),V(t))\gamma'(t)),\end{aligned}$$

by Lemma 4.4. We conclude that

$$\left.\frac{d^2E(\alpha_u)}{du^2}\right|_{u=0}=\int_a^b g\left(V(t),\mathcal{R}(\gamma'(t),V(t))\gamma'(t)-\frac{D^2V(t)}{dt^2}\right)\,dt.$$

Thus if V is a Jacobi field along γ then

$$\left.\frac{d^2E(\alpha_u)}{du^2}\right|_{u=0}=0.$$

Lemma 8.6 *Let (M, g) be a Riemannian manifold and let p and q be points of M. Let $\gamma\colon [a, b] \to M$ be a geodesic with $\gamma(a) = p$ and $\gamma(b) = q$ which minimizes length among all curves from p to q that approximate sufficiently closely to γ. Then*

$$\int_a^b \left(g\left(V(t), \frac{D^2 V(t)}{\partial t^2} \right) + R(V(t), \gamma'(t), V(t), \gamma'(t)) \right) dt \leq 0$$

for all smooth vector fields V along γ which satisfy $V(a) = 0$ and $V(b) = 0$.

Proof Let V be a smooth vector field along $\gamma\colon [a, b] \to \mathbb{R}$ which satisfies $V(a) = 0$ and $V(b) = 0$. It follows from Lemma 6.2 that there exists a smooth map $\alpha\colon [a, b] \times (-\varepsilon, \varepsilon) \to M$ such that $\alpha(t, 0) = \gamma(t)$ for all t and

$$V(t) = \left. \frac{\partial \alpha(t, u)}{\partial u} \right|_{u=0}.$$

Moreover the map α can be chosen such that $\alpha(a, u) = p$ and $\alpha(b, u) = q$ for all $u \in (-\varepsilon, \varepsilon)$. Now γ minimizes length among all smooth curves from p to q that approximate sufficiently closely to γ. Therefore the length $L(\alpha_u)$ of the curve $\alpha_u\colon [a, b] \to M$ satisfies $L(\alpha_u) \geq L(\gamma)$ for all u sufficiently close to zero (where $\alpha_u(t) = \alpha(t, u)$ for all $t \in [a, b]$). We show that this implies that $E(\alpha(u) \geq E(\gamma)$ for all u sufficiently close to zero. To do this we use a form of Schwarz' Inequality which states that if f and g are functions on the closed interval $[a, b]$ whose squares are integrable on $[a, b]$ then

$$\left(\int_a^b f(t) g(t)\, dt \right)^2 \leq \int_a^b f(t)^2\, dt \int_a^b g(t)^2\, dt,$$

where equality holds if and only if the functions f and g are proportional to one another. If we apply this result where

$$f(t) = |\alpha_u'(t)| = \sqrt{g(\alpha_u'(t), \alpha_u'(t))}$$

and where g is the constant function equal to 1 everywhere on $[a, b]$ we see that

$$L(\alpha_u)^2 = \left(\int_a^b |\alpha_u'(t)|\, dt \right)^2 \leq (b - a) \int_a^b |\alpha_u'(t)|^2\, dt = 2(b - a) E(\alpha_u)$$

with equality if and only if $|\alpha'(t)|$ is constant along the curve. Now $|\gamma'(t)|$ is constant along the geodesic γ, hence $L(\gamma)^2 = 2(b - a)E(\gamma)$. Therefore

$$E(\alpha_u) \geq \frac{L(\alpha_u)^2}{2(b - a)} \geq \frac{L(\gamma)^2}{2(b - a)} = E(\gamma)$$

for all u sufficiently close to zero. We conclude that

$$\left.\frac{d^2E(\alpha_u)}{du^2}\right|_{u=0} \geq 0.$$

But it follows from the remarks above that

$$\left.\frac{d^2E(\alpha_u)}{du^2}\right|_{u=0} = -\int_a^b \left(g\left(V(t), \frac{D^2V(t)}{\partial t^2} \right) + R(V(t), \gamma'(t), V(t), \gamma'(t)) \right) dt.$$

The required result follows directly from this. ∎

We can use this inequality to derive the following result.

Lemma 8.7 *Let (M, g) be a Riemannian manifold all of whose sectional curvatures $K(P)$ satisfy $K(P) > \pi^2/L^2$ for some $L > 0$ and let $\gamma\colon [0, L] \to M$ be a geodesic in M of length L, parameterized by arclength. Then γ does not minimize length amongst all smooth curves from $\gamma(0)$ to $\gamma(L)$.*

Proof Let E be a parallel vector field along the geodesic γ which is of unit length and which is everywhere perpendicular to γ and let V be the vector field along γ defined by

$$V(t) = \sin\left(\frac{\pi t}{L}\right) E(t)$$

Then

$$\frac{D^2V(t)}{dt^2} = -\frac{\pi^2}{L^2} \sin\left(\frac{\pi t}{L}\right) E(t),$$

so that

$$g\left(V(t), \frac{D^2V(t)}{dt^2} \right) = -\frac{\pi^2}{L^2} \sin^2\left(\frac{\pi t}{L}\right).$$

Also

$$\begin{aligned} R(V(t), \gamma'(t), V(t), \gamma'(t)) &= \sin^2\left(\frac{\pi t}{L}\right) R(E(t), \gamma'(t), E(t), \gamma'(t)) \\ &> \frac{\pi^2}{L^2} \sin^2\left(\frac{\pi t}{L}\right) \end{aligned}$$

for all $t \in (0, L)$ (since $R(E(t), \gamma'(t), E(t), \gamma'(t))$ is the sectional curvature of the plane spanned by the orthonormal vectors $\gamma'(t)$ and $E(t)$ and this sectional curvature exceeds π^2/L^2). Therefore

$$\int_a^b \left(g\left(V(t), \frac{D^2V(t)}{\partial t^2} \right) + R(V(t), \gamma'(t), V(t), \gamma'(t)) \right) dt > 0.$$

But $V(0) = 0$ and $V(L) = 0$. It follows from Lemma 8.6 that the geodesic $\gamma\colon [0, L] \to M$ is not a curve of shortest length from $\gamma(0)$ to $\gamma(L)$. ∎

One can strengthen the above result. To do this we introduce the *Ricci tensor* ρ of the Riemannian manifold (M, g). The Ricci tensor on M is the tensor of type $(0, 2)$ defined at the point p of M by

$$\rho(U, V) = \sum_{i=1}^{n} R(E_i, U, E_i, V)$$

for all $U, V \in T_pM$, where $(E_1, E_2, \ldots, E_n)$ is an orthonormal basis of T_pM. One can easily verify by simple linear algebra that the Ricci tensor is well-defined independently of the choice of orthonormal basis $(E_1, E_2, \ldots, E_n)$. Now

$$R(E_i, U, E_i, V) = R(E_i, V, E_i, U)$$

for all U and V, by Lemma 4.4. Thus $\rho(U, V) = \rho(V, U)$ for all $U, V \in T_pM$. It can be shown that if V is a tangent vector of unit length at some point p of M then the avarage of the sectional curvatures in all planes in T_pM that contain the vector V is equal to

$$\frac{1}{n-1}\rho(V, V),$$

where n is the dimension of n.

Theorem 8.8 (Myers) *Let (M, g) be a Riemannian manifold of dimension n whose Ricci tensor ρ satisfies the condition $\rho(V, V) \geq C|V|^2$ for all tangent vectors V at all points of M. Let $\gamma\colon [0, L] \to M$ be a geodesic of length L in M parameterized by arclength. If $L^2 > (n-1)\pi^2/C$ then the geodesic γ is not of minimal length. Thus if (M, g) is complete and connected then M is compact and the diameter of M does not exceed*

$$\sqrt{\frac{(n-1)\pi^2}{C}}.$$

Proof Let $\gamma\colon [0, L] \to M$ be a geodesic in M parameterized by arclength whose length L satisfies $L^2 > (n-1)\pi^2/C$. Choose parallel vector fields $E_1, E_2, \ldots, E_{n-1}$ along γ with the property that

$$(E_1(0), E_2(0), \ldots, E_{n-1}(0), \gamma'(0))$$

is an orthonormal basis of $T_{\gamma(0)}M$. Then

$$(E_1(t), E_2(t), \ldots, E_{n-1}(t), \gamma'(t))$$

is an orthonormal basis of $T_{\gamma(t)}M$ for all $t \in [0, L]$. (This follows from Lemma 4.2 on account of the fact that the vector fields $E_1, E_1, \ldots, E_{n-1}, \gamma'$ are covariantly constant along the geodesic γ.) Thus

$$\rho(\gamma'(t), \gamma'(t)) = \sum_{i=1}^{n-1} R(E_i(t), \gamma'(t), E_i(t), \gamma'(t))$$

(where we have used the fact that $R(\gamma'(t), \gamma'(t), \gamma'(t), \gamma'(t)) = 0$.) Let us define smooth vector fields $V_1, V_2, \ldots, V_{n-1}$ along γ by

$$V_i(t) = \sin\left(\frac{\pi t}{L}\right) E_i(t).$$

Then $V_i(0) = 0$ and $V_i(L) = 0$ for all i. Then

$$\sum_{i=1}^{n-1} R(V_i(t), \gamma'(t), V_i(t), \gamma'(t)) = \sin^2\left(\frac{\pi t}{L}\right) \rho(\gamma'(t), \gamma'(t)) \geq C \sin^2\left(\frac{\pi t}{L}\right).$$

But

$$g\left(V_i(t), \frac{D^2 V_i(t)}{dt^2}\right) = -\frac{\pi^2}{L^2} \sin^2\left(\frac{\pi t}{L}\right).$$

Thus if $L^2 > (n-1)\pi^2/C$ then

$$\sum_{i=1}^{n-1} \left(g\left(V_i(t), \frac{D^2 V_i(t)}{dt^2}\right) + R(V_i(t), \gamma'(t), V_i(t), \gamma'(t)) \right) > 0$$

for all $t \in (0, L)$. But if γ were a geodesic of minimal length then we would have

$$\sum_{i=1}^{n-1} \int_a^b \left(g\left(V_i(t), \frac{D^2 V_i(t)}{\partial t^2}\right) + R(V_i(t), \gamma'(t), V_i(t), \gamma'(t)) \right) dt \leq 0,$$

by Lemma 8.6. This shows that if the length L of γ satisfies $L^2 > (n-1)\pi^2/C$ then γ is not a geodesic of minimal length.

Now suppose that (M, g) is complete and connected. Then any two points of M can be joined by a geodesic of minimal length, by Theorem 7.1. We conclude that the length L of this geodesic must satisfy $L^2 > (n-1)\pi^2/C$, so that the diameter of M does not exceed

$$\sqrt{\frac{(n-1)\pi^2}{C}}.$$

Moreover if we choose some $p \in M$ then M is contained in the image under the exponential map $\exp_p$ of the closed ball in T_pM of radius R about the zero vector, where $R^2 = (n-1)\pi^2/C$. Thus M is compact (since M is the image of a compact set under a continuous map). ∎

Remark Let S^n denote the unit sphere in $\mathbb{R}^{n+1}$. We define the standard Riemannian metric g on S^n to be the restriction to the tangent spaces of S^n of the standard inner product on $\mathbb{R}^{n+1}$. When one calculates the Riemann curvature tensor of S^n one finds that all the sectional curvatures of S^n are equal to $+1$, and the Ricci tensor ρ aof S^n is given by $\rho(U,V) = (n-1)g(U,V)$ for all tangent vectors U and V at any given point of S^n. Any meridian of S^n joining the points $(1, 0, \ldots, 0)$ and $(-1, 0, \ldots, 0)$ is a geodesic of length π. It follows that the inequality $L^2 \leq (n-1)\pi^2/C$ on the length L of minimal geodesics provided by Myers' Theorem is an equality for this geodesic on S^n (where C is given by $C = n - 1$). Thus the upper bound on the length of geodesics of minimal length provided by Myers' Theorem is attained when the Riemannian manifold in question is the standard n-dimensional sphere S^n.

Remark It can be shown that if M is a connected smooth manifold then there exists a covering map $\pi\colon \tilde{M} \to M$ over M, where $\tilde{M}$ is a simpy-connected smooth manifold. The smooth manifold $\tilde{M}$ is known as the *universal cover* of the manifold M. Moreover the covering map $\pi \tilde{M} \to M$ is a local diffeomorphism. The inverse image $\pi^{-1}(m)$ of a point m of M is in bijective correspondence with the fundamental group $\pi_1(M, m)$ of M based at M. Any Riemannian metric g on M induces a corresponding Riemannian metric $\tilde{g}$ on $\tilde{M}$ characterized by the property that the covering map $\pi\colon \tilde{M} \to M$ is a local isometry. The Riemannian manifold $(\tilde{M}, \tilde{g})$ is complete if and only if (M, g) is complete.

Now suppose that M is connected and complete and that there exists a positive constant $C > 0$ with the property that $\rho(V,V) \geq C|V|^2$ for all tangent vectors V on M, where ρ is the Ricci tensor of M. Then the same condition will hold for the universal cover $(\tilde{M}, \tilde{g})$ of M. We conclude that $\tilde{M}$ is compact. Now given $m \in M$, there exists an open neighbourhood U of M such that $\pi^{-1}(U)$ is a disjoint union of copies of U. It follows from the compactness of $\tilde{M}$ that the number of such copies must be finite. Therefore the fundamental group of M is finite.

With substantially more effort, one can prove the following theorem.

> **Sphere Theorem** Let (M, g) be a complete simply-connected Riemannian manifold of dimension n. Suppose that the sectional curvatures $K(P)$ of M satisfy $\frac{1}{4}K_0 < K(P) \leq K_0$ for some positive constant K_0. Then M is homeomorphic to the n-dimensional sphere S^n.