Chapter 11

Riemannian Metrics, Riemannian Manifolds

11.1 Frames

Fortunately, the rich theory of vector spaces endowed with a Euclidean inner product can, to a great extent, be lifted to the tangent bundle of a manifold.

The idea is to equip the tangent space T_pM at p to the manifold M with an inner product $\langle -, - \rangle_p$, in such a way that these inner products vary smoothly as p varies on M.

It is then possible to define the length of a curve segment on a M and to define the distance between two points on M.

The notion of local (and global) frame plays an important technical role.

Definition 11.1. Let M be an n-dimensional smooth manifold. For any open subset, $U \subseteq M$, an n-tuple of vector fields, $(X_1, \ldots, X_n)$, over U is called a *frame over U* iff $(X_1(p), \ldots, X_n(p))$ is a basis of the tangent space, T_pM, for every $p \in U$. If $U = M$, then the X_i are global sections and $(X_1, \ldots, X_n)$ is called a *frame* (of M).

The notion of a frame is due to Élie Cartan who (after Darboux) made extensive use of them under the name of *moving frame* (and the *moving frame method*).

Cartan's terminology is intuitively clear: As a point, p, moves in U, the frame, $(X_1(p), \ldots, X_n(p))$, moves from fibre to fibre. Physicists refer to a frame as a choice of *local gauge*.

If $\dim(M) = n$, then for every chart, (U, φ), since $d\varphi^{-1}_{\varphi(p)} \colon \mathbb{R}^n \to T_pM$ is a bijection for every $p \in U$, the n-tuple of vector fields, $(X_1, \ldots, X_n)$, with $X_i(p) = d\varphi^{-1}_{\varphi(p)}(e_i)$, is a frame of TM over U, where $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^n$. See Figure 11.1.

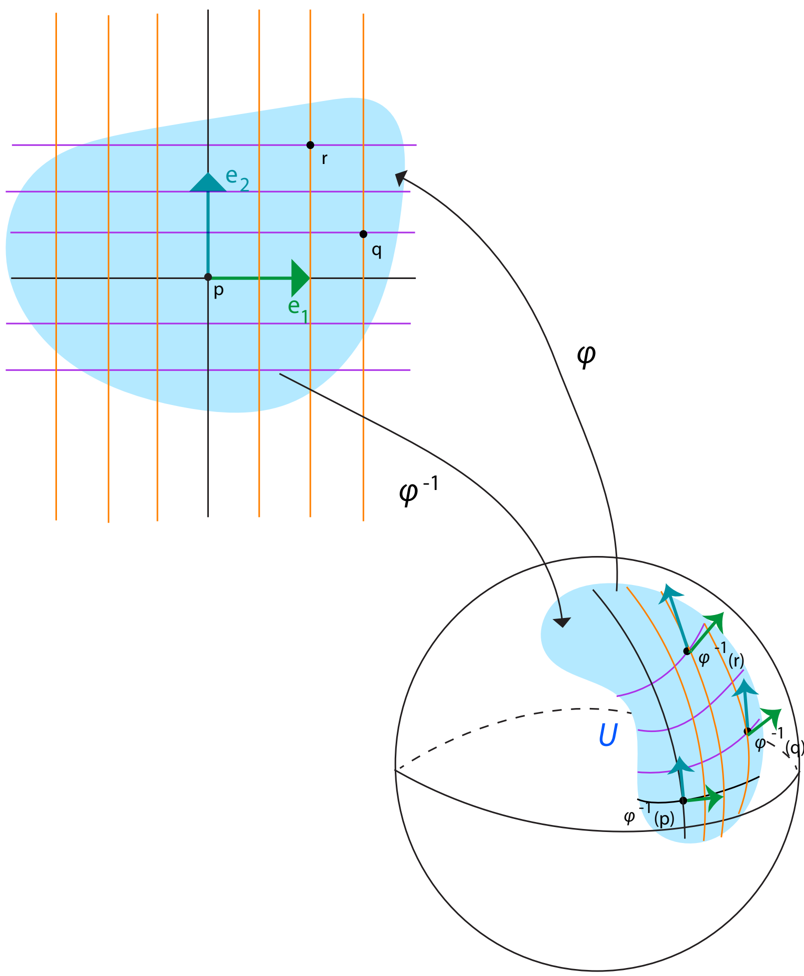

Figure 11.1: A frame on S^2.

The following proposition tells us when the tangent bundle is trivial (that is, isomorphic to the product, $M \times \mathbb{R}^n$):

Proposition 11.1. *The tangent bundle, TM, of a smooth n-dimensional manifold, M, is trivial iff it possesses a frame of global sections (vector fields defined on M).*

As an illustration of Proposition 11.1 we can prove that the tangent bundle, TS^1, of the circle, is trivial.

Indeed, we can find a section that is everywhere nonzero, *i.e.* a non-vanishing vector field, namely

$$X(\cos\theta, \sin\theta) = (-\sin\theta, \cos\theta).$$

The reader should try proving that TS^3 is also trivial (use the quaternions).

However, TS^2 is nontrivial, although this not so easy to prove.

More generally, it can be shown that TS^n is nontrivial for all even $n \geq 2$. It can even be shown that S^1, S^3 and S^7 are the only spheres whose tangent bundle is trivial. This is a rather deep theorem and its proof is hard.

Remark: A manifold, M, such that its tangent bundle, TM, is trivial is called *parallelizable*.

We now define Riemannian metrics and Riemannian manifolds.

11.2 Riemannian Metrics

Definition 11.2. Given a smooth n-dimensional manifold, M, a *Riemannian metric on M (or TM)* is a family, $(\langle -, - \rangle_p)_{p \in M}$, of inner products on each tangent space, T_pM, such that $\langle -, - \rangle_p$ depends smoothly on p, which means that for every chart, $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^n$, for every frame, $(X_1, \ldots, X_n)$, on U_α, the maps

$$p \mapsto \langle X_i(p), X_j(p) \rangle_p, \qquad p \in U_\alpha,\ 1 \leq i, j \leq n$$

are smooth. A smooth manifold, M, with a Riemannian metric is called a *Riemannian manifold*.

If $\dim(M) = n$, then for every chart, (U, φ), we have the frame, $(X_1, \ldots, X_n)$, over U, with $X_i(p) = d\varphi^{-1}_{\varphi(p)}(e_i)$, where $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^n$. Since every vector field over U is a linear combination, $\sum_{i=1}^n f_i X_i$, for some smooth functions, $f_i \colon U \to \mathbb{R}$, the condition of Definition 11.2 is equivalent to the fact that the maps,

$$p \mapsto \langle d\varphi^{-1}_{\varphi(p)}(e_i), d\varphi^{-1}_{\varphi(p)}(e_j) \rangle_p, \qquad p \in U,\ 1 \leq i, j \leq n,$$

are smooth.

If we let $x = \varphi(p)$, the above condition says that the maps,

$$x \mapsto \langle d\varphi_x^{-1}(e_i), d\varphi_x^{-1}(e_j)\rangle_{\varphi^{-1}(x)}, \quad x \in \varphi(U), 1 \leq i, j \leq n,$$

are smooth.

If M is a Riemannian manifold, the metric on TM is often denoted $g = (g_p)_{p \in M}$. In a chart, using local coordinates, we often use the notation $g = \sum_{ij} g_{ij} dx_i \otimes dx_j$ or simply $g = \sum_{ij} g_{ij} dx_i dx_j$, where

$$g_{ij}(p) = \left\langle \left(\frac{\partial}{\partial x_i}\right)_p, \left(\frac{\partial}{\partial x_j}\right)_p \right\rangle_p.$$

For every $p \in U$, the matrix, $(g_{ij}(p))$, is symmetric, positive definite.

The standard Euclidean metric on $\mathbb{R}^n$, namely,

$$g = dx_1^2 + \cdots + dx_n^2,$$

makes $\mathbb{R}^n$ into a Riemannian manifold.

Then, every submanifold, M, of $\mathbb{R}^n$ inherits a metric by restricting the Euclidean metric to M.

For example, the sphere, S^{n-1}, inherits a metric that makes S^{n-1} into a Riemannian manifold.

It is instructive to find the local expression of this metric for S^2 in spherical coordinates.

We can parametrize the sphere S^2 in terms of two angles θ (the *colatitude*) and φ (the *longitude*) as follows:

$$\begin{aligned} x &= \sin\theta\cos\varphi \\ y &= \sin\theta\sin\varphi \\ z &= \cos\theta. \end{aligned}$$

See Figure 11.2.

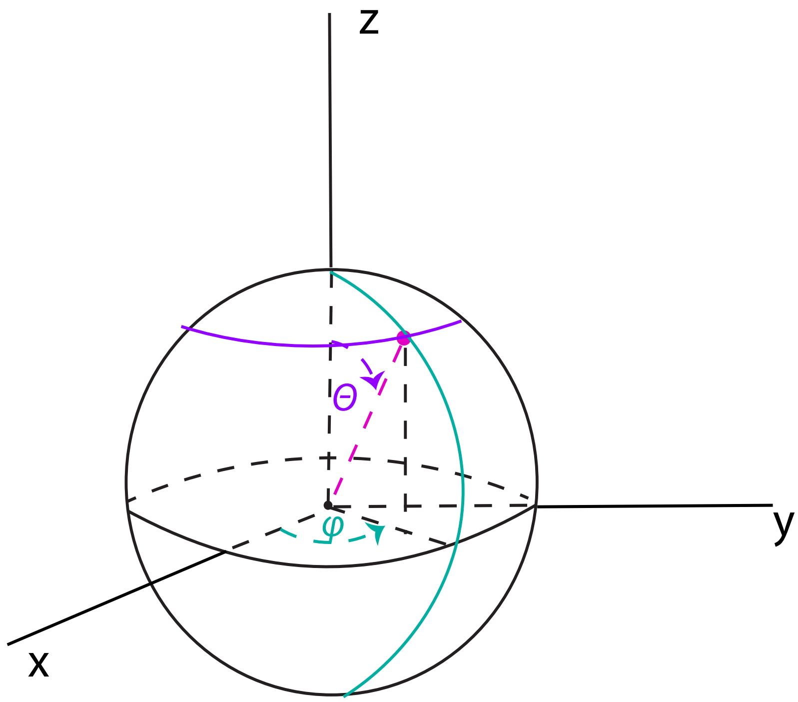

Figure 11.2: The spherical coordinates of S^2.

In order for the above to be a parametrization, we need to restrict its domain to $V = \{(\theta, \varphi) \mid 0 < \theta < \pi, 0 < \varphi < 2\pi\}$.

Then the semicircle from the north pole to the south pole lying in the xz-plane is omitted from the sphere.

In order to cover the whole sphere, we need another parametrization obained by choosing the axes in a suitable fashion; for example, to omit the semicircle in the xy-plane from $(0, 1, 0)$ to $(0, -1, 0)$ and with $x \leq 0$.

To compute the matrix giving the Riemannian metric in this chart, we need to compute a basis $(u(\theta, \varphi), v(\theta, \varphi))$ of the the tangent plane T_pS^2 at
$p = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$.

We can use

$$\begin{aligned} u(\theta, \varphi) &= \frac{\partial p}{\partial \theta} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta) \\ v(\theta, \varphi) &= \frac{\partial p}{\partial \varphi} = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0), \end{aligned}$$

and we find that

$$\begin{aligned} \langle u(\theta, \varphi), u(\theta, \varphi)\rangle &= 1 \\ \langle u(\theta, \varphi), v(\theta, \varphi)\rangle &= 0 \\ \langle v(\theta, \varphi), v(\theta, \varphi)\rangle &= \sin^2\theta, \end{aligned}$$

so the metric on T_pS^2 w.r.t. the basis $(u(\theta, \varphi), v(\theta, \varphi))$ is given by the matrix

$$g_p = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}.$$

Thus, for any tangent vector

$$w = au(\theta, \varphi) + bv(\theta, \varphi), \quad a, b \in \mathbb{R},$$

we have

$$g_p(w, w) = a^2 + \sin^2\theta\, b^2.$$

A nontrivial example of a Riemannian manifold is the *Poincaré upper half-space*, namely, the set
$H = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ equipped with the metric

$$g = \frac{dx^2 + dy^2}{y^2}.$$

Consider the Lie group $\mathbf{SO}(n)$.

We know from Section 7.2 that its tangent space at the identity $T_I\mathbf{SO}(n)$ is the vector space $\mathfrak{so}(n)$ of $n \times n$ skew symmetric matrices, and that the tangent space $T_Q\mathbf{SO}(n)$ to $\mathbf{SO}(n)$ at Q is isomorphic to

$$Q\mathfrak{so}(n) = \{QB \mid B \in \mathfrak{so}(n)\}.$$

If we give $\mathfrak{so}(n)$ the inner product

$$\langle B_1, B_2\rangle = \mathrm{tr}(B_1^\top B_2) = -\mathrm{tr}(B_1 B_2),$$

the inner product on $T_Q\mathbf{SO}(n)$ is given by

$$\langle QB_1, QB_2\rangle = \mathrm{tr}((QB_1)^\top QB_2) = \mathrm{tr}(B_1^\top B_2).$$

We will see in Chapter 13 that the length $L(\gamma)$ of the curve segment γ from I to e^B given by $t \mapsto e^{tB}$ (with $B \in \mathfrak{so}(n)$) is given by

$$L(\gamma) = \left(\operatorname{tr}(-B^2)\right)^{\frac{1}{2}}.$$

More generally, given any Lie group G, any inner product $\langle -, - \rangle$ on its Lie algebra $\mathfrak{g}$ induces by left translation an inner product $\langle -, - \rangle_g$ on T_gG for every $g \in G$, and this yields a Riemannian metric on G (which happens to be left-invariant; see Chapter 18).

Going back to the second example of Section 7.5, where we computed the differential df_R of the function $f\colon \mathbf{SO}(3) \to \mathbb{R}$ given by

$$f(R) = (u^\top R v)^2,$$

we found that

$$df_R(X) = 2u^\top X v u^\top R v, \quad X \in R\mathfrak{so}(3).$$

Since each tangent space $T_R\mathbf{SO}(3)$ is a Euclidean space under the inner product defined above, by duality (see Proposition **??** applied to the pairing $\langle -, -\rangle$), there is a unique vector $Y \in T_R\mathbf{SO}(3)$ defining the linear form df_R; that is,

$$\langle Y, X\rangle = df_R(X), \quad \text{for all } X \in T_R\mathbf{SO}(3).$$

By definition, the vector Y is the *gradient of f at R*, denoted $(\mathrm{grad}(f))_R$.

We leave it as an exercise to prove that the gradient of f at R is given by

$$(\mathrm{grad}(f))_R = u^\top RvR(R^\top uv^\top - vu^\top R).$$

More generally, the notion of gradient is defined as follows.

Definition 11.3. If $(M, \langle -, - \rangle)$ is a smooth manifold with a Riemannian metric and if $f\colon M \to \mathbb{R}$ is a smooth function on M, then the unique smooth vector field $\mathrm{grad}(f)$ defined such that

$$\langle (\mathrm{grad}(f))_p, u \rangle_p = df_p(u), \quad \text{for all } p \in M \text{ and all } u \in T_pM$$

is called the *gradient of f*.

It is usually complicated to find the gradient of a function.

If (U, φ) is a chart of M, with $p \in M$, and if

$$\left(\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p\right)$$

denotes the basis of T_pM induced by φ, the local expression of the metric g at p is given by the $n \times n$ matrix $(g_{ij})_p$, with

$$(g_{ij})_p = g_p\left(\left(\frac{\partial}{\partial x_i}\right)_p, \left(\frac{\partial}{\partial x_j}\right)_p\right).$$

The inverse is denoted by $(g^{ij})_p$.

We often omit the subscript p and observe that for every function $f \in C^\infty(M)$,

$$\operatorname{grad} f = \sum_{ij} g^{ij} \frac{\partial f}{\partial x_j} \frac{\partial}{\partial x_i}.$$

A way to obtain a metric on a manifold, N, is to pull-back the metric, g, on another manifold, M, along a local diffeomorphism, $\varphi\colon N \to M$.

Recall that φ is a local diffeomorphism iff

$$d\varphi_p\colon T_pN \to T_{\varphi(p)}M$$

is a bijective linear map for every $p \in N$.

Given any metric g on M, if φ is a local diffeomorphism, we define the *pull-back metric*, φ^*g, on N induced by g as follows: For all $p \in N$, for all $u, v \in T_pN$,

$$(\varphi^*g)_p(u, v) = g_{\varphi(p)}(d\varphi_p(u), d\varphi_p(v)).$$

We need to check that $(\varphi^*g)_p$ is an inner product, which is very easy since $d\varphi_p$ is a linear isomorphism.

Our map, φ, between the two Riemannian manifolds (N, φ^*g) and (M, g) is a local isometry, as defined below.

Definition 11.4. Given two Riemannian manifolds, (M_1, g_1) and (M_2, g_2), a *local isometry* is a smooth map, $\varphi\colon M_1 \to M_2$, such that $d\varphi_p\colon T_pM_1 \to T_{\varphi(p)}M_2$ is an isometry between the Euclidean spaces $(T_pM_1, (g_1)_p)$ and $(T_{\varphi(p)}M_2, (g_2)_{\varphi(p)})$, for every $p \in M_1$, that is,

$$(g_1)_p(u, v) = (g_2)_{\varphi(p)}(d\varphi_p(u), d\varphi_p(v)),$$

for all $u, v \in T_pM_1$ or, equivalently, $\varphi^* g_2 = g_1$. Moreover, φ is an *isometry* iff it is a local isometry and a diffeomorphism.

The isometries of a Riemannian manifold, (M, g), form a group, $\mathrm{Isom}(M, g)$, called the *isometry group of* (M, g).

An important theorem of Myers and Steenrod asserts that the isometry group, $\mathrm{Isom}(M, g)$, is a Lie group.

An interesting example of the notion of isometry arises in machine learning, namely with respect to the *multinomial manifold*.

Example 11.1. Let Δ_+^n be the standard open simplex

$$\Delta_+^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1 + \cdots + x_{n+1} = 1, x_i > 0\}.$$

This is an open submanifold of the hyperplane of equation $x_1 + \cdots + x_{n+1} = 1$, which is itself a submanifold of $\mathbb{R}^{n+1}$.

The manifold Δ_+^n is diffeomorphic to the positive quadrant of the unit sphere in $\mathbb{R}^{n+1}$ given by

$$S_+^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1, x_i > 0\}.$$

See Figure 11.3.

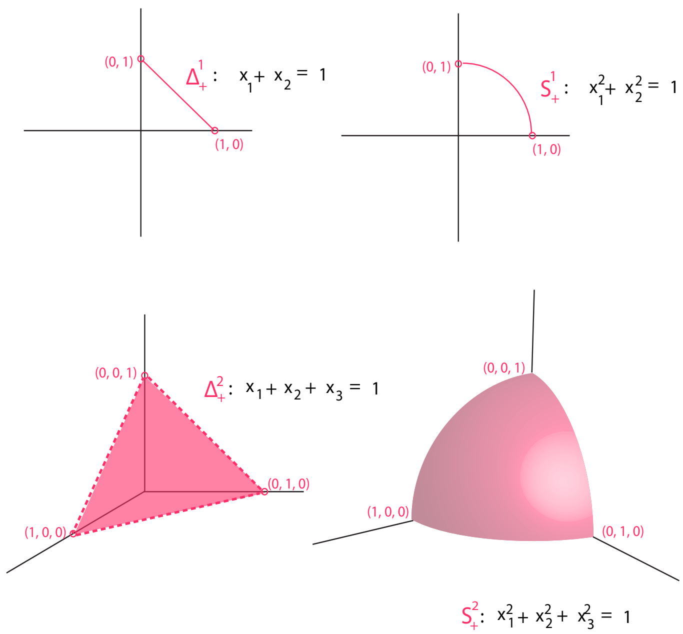

Figure 11.3: The open simplexes Δ^1_+ and Δ^2_+ along with the diffeomorphic S^1_+ and S^2_+.

The maps $\varphi\colon S^n_+ \to \Delta^n_+$ and $\psi\colon \Delta^n_+ \to S^n_+$ given by

$$\begin{aligned}\varphi(x_1\dots,x_{n+1}) &= (x_1^2,\dots,x_{n+1}^2)\\ \psi(x_1\dots,x_{n+1}) &= (\sqrt{x_1},\dots,\sqrt{x_{n+1}})\end{aligned}$$

are clearly inverse diffeomorphisms. The map $\varphi\colon S^n_+ \to \Delta^n_+$ is often called the *real moment map*.

For any $x \in S^n_+$, the tangent space $T_xS^n_+$ is given by

$$\begin{aligned}T_xS^n_+ &= \{u \in \mathbb{R}^{n+1} \mid \langle x, u\rangle = 0\}\\ &= \{u \in \mathbb{R}^{n+1} \mid x_1u_1 + \cdots + x_{n+1}u_{n+1} = 0\},\end{aligned}$$

where $\langle -, -\rangle$ is the standard Euclidean inner product in $\mathbb{R}^{n+1}$, and for any $x \in \Delta^n_+$, the tangent space $T_x\Delta^n_+$ is given by

$$T_x\Delta^n_+ = \{u \in \mathbb{R}^{n+1} \mid u_1 + \cdots + u_{n+1} = 0\}.$$

It is easily verified that the derivative $d\varphi_x$ of φ at $x \in S^n_+$ is given by

$$d\varphi_x(u_1, \ldots, u_{n+1}) = 2(x_1u_1, \ldots, x_{n+1}u_{n+1}).$$

As a consequence, if we give Δ^n_+ the Riemannian metric defined by

$$\langle u, v\rangle^F_x = \frac{1}{4}\sum_{i=1}^{n+1}\frac{u_iv_i}{x_i}, \quad x \in \Delta^n_+,$$

then we have

$$\begin{aligned}\langle d\varphi_x(u), d\varphi_x(v)\rangle^F_{\varphi(x)} &= \langle 2(x_1u_1, \ldots, x_{n+1}u_{n+1}), \\ &\qquad 2(x_1v_1, \ldots, x_{n+1}v_{n+1})\rangle^F_{(x_1^2, \ldots, x_{n+1}^2)} \\ &= \frac{1}{4}\sum_{i=1}^{n+1}\frac{2x_iu_i2x_iv_i}{x_i^2} \\ &= \sum_{i=1}^{n+1}u_iv_i = \langle u, v\rangle.\end{aligned}$$

Therefore, φ is an isometry between the Riemannian manifold $(S^n_+, \langle -, - \rangle)$ (equipped with the restriction of the Euclidean metric of $\mathbb{R}^{n+1}$) to the manifold $(\Delta^n_+, \langle -, - \rangle^F)$ equipped with the metric

$$\begin{aligned} \langle u, v \rangle^F_x &= \frac{1}{4} \sum_{i=1}^{n+1} \frac{u_i v_i}{x_i} = \frac{1}{4} \sum_{i=1}^{n+1} x_i \frac{u_i}{x_i} \frac{v_i}{x_i} \\ &= \frac{1}{4} \sum_{i=1}^{n+1} x_i \frac{d(\log x_i)}{dx_i} \frac{d(\log x_i)}{dx_i} u_i v_i, \quad x \in \Delta^n_+, \end{aligned}$$

known as the *Fisher information metric* (actually, one fourth of the Fisher information metric).

The above shows that the Fisher information metric is the pullback of the Euclidean metric on S^n_+ along the inverse ψ of the real moment map φ.

In machine learning the manifold $(\Delta^n_+, \langle -, - \rangle^F)$ is called the *multinomial manifold*. Unfortunately, it is often denoted by $\mathbb{P}^n$, which clashes with the standard notation for projective space.

Given a map, $\varphi\colon M_1 \to M_2$, and a metric g_1 on M_1, in general, φ does not induce any metric on M_2.

However, if φ has some extra properties, it does induce a metric on M_2. This is the case when M_2 arises from M_1 as a quotient induced by some group of isometries of M_1. For more on this, see Gallot, Hulin and Lafontaine [19], Chapter 2, Section 2.A.

Now, because a manifold is *paracompact* (see Section 9.1), a Riemannian metric always exists on M. This is a consequence of the existence of partitions of unity (see Theorem 9.4).

Theorem 11.2. *Every smooth manifold admits a Riemannian metric.*