Chapter 7

An Introduction to Manifolds

This chapter contains a brief introduction to the classical theory of differential geometry. The fundamental notions presented here deal with differentiable manifolds, tangent space, vector fields, differentiable maps, 1-forms, tensors, linear connections, Riemannian manifolds, and the Levi–Civita connection. The material of this chapter forms the basis for next chapters.

7.1 The Concept of Manifold

A manifold is a multidimensional geometric object that can be considered as a space which is locally similar to the Euclidean space. Since differentiation is a locally defined property, then the differentiation can be defined on a manifold in a similar way as it is defined on the Euclidean space. A point on a manifold can be described by several sets of parameters, which are regarded as local coordinate systems.

The advantage of working on a manifold is that one can consider and study those geometric concepts (functions, invariants, vector fields, tensor fields, connections, etc.) that make sense globally and can also be described quantitatively in local coordinate systems. This property initially made sense in Physics and Relativity Theory, where each coordinate system corresponds to a system of reference.

O. Calin and C. Udrişte, *Geometric Modeling in Probability and Statistics*,
DOI 10.1007/978-3-319-07779-6_7,

Therefore, the main objects of study in that case are velocity, acceleration, force, matter fields, momenta, etc., i.e., objects that remain invariant under a change of the system of reference. This means that while these objects make sense globally, they can be described quantitatively in terms of local coordinates.

The earth's surface is one of the most suggestive examples of manifolds. One is aware of this manifold only locally, where it resembles a piece of plane. A local observer situated on earth's surface can measure coordinates at any point by choosing an origin and a unit of measure, the result of this work being a local map of the region. Even if drawn at different scales, any two maps of overlapping regions are correlated, in the sense that one can "understand" their relationship. If these maps constitute an entire cartography[1] of the planet, then they form an *atlas*. Nowadays people are more familiar with the googlemaps system. The maps can be transformed by translation, contraction, or dilation, which move from one map to another, the transformation being smooth and assuring the correlation between maps. The local knowledge of the earth surface contained in an atlas forms the notion of manifold.

Consider now the system of artificial satellites rotating around the earth. Each satellite can cover a certain region of the earth surface. All satellites are supposed to cover the entire earth surface, with some overlap. The information retrieved from the satellites forms an atlas and the manifold notion emerges again.

Suppose now that a certain country is monitored by a grid of cellular phone towers, each tower servicing a specific region. This is an example that can be considered as a manifold again, each tower region being considered as a local chart. In general, the manifold notion emerges when we can describe only locally an entire global object. The word "local" in this case describes a region one can encompass with the eye, or an area which can be covered by the local cellular phone tower. "Global" describes either an entire country or continent or even the whole earth surface.

There are two distinct points of view when studying a manifold. One is the *intrinsic* point of view of a local observer, situated on the manifold, whose knowledge is bound to a local chart, or system of reference. For instance, digging a ditch or measuring the distance between a house and a nearby tree is an intrinsic activity. The other

[1] Cartography is the study and practice of making maps.

point of view is called *extrinsic*. The extrinsic knowledge is acquired by an observer while elevating himself above the manifold and looking at it from outside. The information about the earth surface obtained by a monkey climbing an eucalypt tree that grows on the earth surface, is extrinsic, while the point of view of a microorganism living on the ground is intrinsic. The notions of intrinsic and extrinsic can be described geometrically by considering either only the metric of the manifold, or taking into account also the normal vector to the manifold. There are some geometrical notions that can be described exclusively in an extrinsic way. Since the round shape of the earth was recently fully mapped by satellites, understanding the shape of the earth is an extrinsic feature. However, this should not be mistaken with curvature, which can be described intrinsically in terms of the local metric (Gauss' Egregium Theorem).

A useful tool used in describing some geometric objects on a manifold is the concept of *tensor*. Many physical quantities, such as force, velocity, acceleration, work, etc., can be described successfully as tensors. Their main feature of a tensor is that it can be described quantitatively in a local chart, and its coordinates transform by a matrix multiplication when changing charts. Therefore, if a tensor vanishes in one chart, then it vanishes in all charts. Since it turns out that the difference of two tensors is also a tensor, the last two features allow for a very powerful method of proving relations between tensors by checking them in a suitable local chart. The work in local coordinates used to prove global relations has been proved extremely useful and has been developed into the so-called tensorial formalism. For instance, if one needs to show that the tensors T and P are equal, it suffices to only show that their components are equal in a chart, $T_{ij} = P_{ij}$.

Many geometrical objects studied in differential geometry are tensors; however, they are called by distinct names, such as metric, vector field, 1-form, volume form, curvature, etc. All these are objects independent of the system of coordinates and can be defined globally but may be written locally in a local system of coordinates using local components. For example, a *vector field* is an object that may be written in local coordinates as $V = \sum V^i \frac{\partial}{\partial x^i}$, where $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1,\dots,n}$ is a basis of the local system of coordinates chosen. This means that its components measured in this system of reference are given by $V^1, \dots, V^n$. Similarly, a *1-form* is an object that can be written in local coordinates as $\omega = \sum \omega_i dx^i$, where $\{dx^i\}_{i=1,\dots,n}$ is a basis of the 1-forms of the local system of coordinates chosen. A *metric* is a tensor written as $g = \sum g_{ij} dx^i \otimes dx^j$, where $\otimes$ is an operation called *tensorial product.*

7.2 Manifold Definition

This section presents the precise definition of manifolds. All manifolds considered in this book are *real*, i.e., the local model is the Euclidean space $\mathbb{R}^n$.

The construction of a manifold starts with a *metric space* (the underlying structure of the manifold), i.e., a space on which is defined a distance function.

Definition 7.2.1 *Let M be a set of points. A distance function is a mapping $d : M \times M \to [0, \infty)$ with the following properties:*

(*i*) *non-degenerate: $d(x, y) = 0$ if and only if $x = y$;*

(*ii*) *symmetric: $d(x, y) = d(y, x)$, for all $x, y \in M$;*

(*iii*) *satisfies the triangle inequality: $d(x, z) \le d(x, y) + d(y, z)$, for all $x, y, z \in M$.*

The pair (M, d) is called a metric space.

Example 7.2.2 Let $M = \mathbb{R}^n$ and consider $x, y \in M$, with $x = (x^1, \dots, x^n)$, $y = (y^1, \dots, y^n)$. Then the Euclidean distance is given by $d_E(x, y) = \Big[\sum_{k=1}^{n} (x^k - y^k)^2\Big]^{1/2}$. The metric space (M, d_E) is called the Euclidean space.

Example 7.2.3 The mapping $d_T : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ given by $d_T(x, y) = \sum_{k=1}^{n} |x^k - y^k|$ is called the taxi-cab distance. It bears its name after the distance followed by a cab in a city with perpendicular and equidistant streets such as New York city.

Definition 7.2.4 *Let (M, d) be a metric space. Consider $x \in M$ and $r > 0$. The ball of radius r and centered at x is the set $B_r(x) = \{y \in M; d(x, y) < r\}$. A subset U of M is called open if for any $x \in U$, there is a $r > 0$ such that $B_r(x) \subset U$.*

The equivalence of the definitions of functions continuity in the framework of metric spaces is stated as in the following.

Proposition 7.2.5 *Let $f : (M, d_M) \to (N, d_N)$ be a mapping between two metric spaces. The following are equivalent:*

(*a*) *For any open set V in N, the pullback $f^{-1}(V) = \{x \in M; f(x) \in V\}$ is an open set in M.*

(*b*) *For any convergent sequence $x_n \to x$ in M, (i.e., $d_M(x_n, x) \to 0, n \to \infty$) we have $f(x_n) \to f(x)$ in N, (i.e., $d_N\big(f(x_n), f(x)\big) \to 0,\ n \to \infty$).*

A function $f : M \to N$ is called *continuous* if any of the foregoing parts (a) or (b) holds true. If f is invertible and both f and f^{-1} are continuous, then f is called a *homeomorphism* between M and N.

Definition 7.2.6 *Let $U \subset M$ be an open set. Then the pair (U, ϕ) is called a chart (coordinate system) on M, if $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a homeomorphism of the open set U in M onto an open set $\phi(U)$ of $\mathbb{R}^n$. The coordinate functions on U are defined as $x^j : U \to \mathbb{R}$, and $\phi(p) = (x^1(p), \dots, x^n(p))$, namely $x^j = u^j \circ \phi$, where $u^j : \mathbb{R}^n \to \mathbb{R}$, $u^j(a_1, \dots, a_n) = a_j$ is the jth projection.*

The integer n is the dimension of the coordinate system. Roughly speaking, the dimension is the number of coordinates needed to describe the position of a point in M.

Definition 7.2.7 *An atlas $\mathcal{A}$ of dimension n associated with the metric space M is a collection of charts $\{(U_\alpha, \phi_\alpha)\}_\alpha$ such that*

1) $U_\alpha \subset M, \forall\alpha,\ \ \bigcup_\alpha U_\alpha = M$ *(i.e., U_α covers M),*

2) if $U_\alpha \cap U_\beta \neq \emptyset$, the restriction to $\phi_\alpha(U_\alpha \cap U_\beta)$ of the map

$$F_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

is differentiable from $\mathbb{R}^n$ to $\mathbb{R}^n$ (i.e., the systems of coordinates overlap smoothly), see Fig. 7.1.

There might be several atlases on a given metric space M. Two atlases $\mathcal{A}$ and $\mathcal{A}'$ are called *compatible* if their union is an atlas on M. The set of compatible atlases with a given atlas $\mathcal{A}$ can be partially ordered by inclusion. Its maximal element is called the *complete atlas* $\overline{\mathcal{A}}$. This

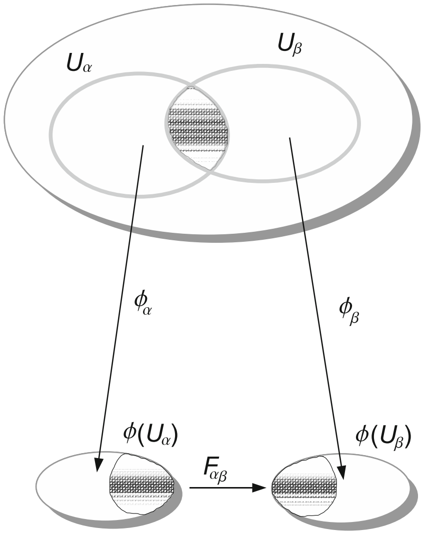

Figure 7.1: Correlated charts on a differential manifold

atlas contains all the charts that overlap smoothly with the charts of the given atlas $\mathcal{A}$. The dimension n of the space $\mathbb{R}^n$, which models the manifold structure, is called the dimension of the atlas $\mathcal{A}$.

Definition 7.2.8 *A differentiable manifold M is a metric space endowed with a complete atlas. The dimension n of the atlas is called the dimension of the manifold.*

We owe a remark about the completeness of an atlas. The completeness feature is required to assure for maximum chartographic information, in the sense that any considered chart is already filed in the atlas; equivalently, no new charts can be considered besides the ones that are already part of the atlas.

However, in practice it suffices to supply an arbitrary atlas (usually not the maximal one), the maximal atlas resulting from the combination of all atlases.

7.3 Examples of Manifolds

In this section we supply a few examples of useful manifolds.

1) The simplest differentiable manifold is the Euclidean space itself, $\mathbb{R}^n$. In this case the atlas has only one chart, the identity map, $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathrm{Id}(x) = x$.

2) Any open set U of $\mathbb{R}^n$ is a differential manifold, with only one chart, (U, Id).

3) Any non-intersecting curve $c : (a, b) \to \mathbb{R}^n$, with $\dot{c}(t) \neq 0$, is a one-dimensional manifold. In this case $M = c\big((a, b)\big)$ and the atlas consists of only one chart (U, ϕ), with $U = c\big((a, b)\big)$, and $\phi : U \to (a, b)$, $\phi = c^{-1}_{|U}$.

4) The sphere $\mathbb{S}^2 = \{x = (x^1, x^2, x^3) \in \mathbb{R}^3 \; ; \; (x^1)^2 + (x^2)^2 + (x^3)^2 = 1\}$ is a differentiable manifold of dimension 2. We shall supply in the following two atlases. The first atlas contains six charts, being given by $\mathcal{A} = \{U_i, \phi_i\}_{i=\overline{1,3}} \cup \{V_i, \psi_i\}_{i=\overline{1,3}}$, where

$$U_1 = \{x \; ; \; x^1 > 0\}\,, \quad \phi_1 : U_1 \to \mathbb{R}^2\,, \quad \phi_1(x) = (x^2, x^3),$$

$$V_1 = \{x \; ; \; x^1 < 0\}, \quad \psi_1 : V_1 \to \mathbb{R}^2, \quad \psi_1(x) = (x^2, x^3),$$

$$U_2 = \{x \; ; \; x^2 > 0\}\,, \quad \phi_2 : U_2 \to \mathbb{R}^2\,, \quad \phi_2(x) = (x^1, x^3),$$

$$V_2 = \{x \; ; \; x^2 < 0\}, \quad \psi_2 : V_2 \to \mathbb{R}^2, \quad \psi_2(x) = (x^1, x^3),$$

$$U_3 = \{x \; ; \; x^3 > 0\}\,, \quad \phi_3 : U_3 \to \mathbb{R}^2\,, \quad \phi_3(x) = (x^1, x^2),$$

$$V_3 = \{x \; ; \; x^3 < 0\}, \quad \psi_3 : V_3 \to \mathbb{R}^2, \quad \psi_3(x) = (x^1, x^2).$$

The second atlas is $\mathcal{A}' = \{(U, \phi_N), (V, \phi_S)\}$, where $U = \mathbb{S}^2 \backslash \{(0, 0, 1)\}$, $V = \mathbb{S}^2 \backslash \{(0, 0, -1)\}$, and the stereographic projections $\phi_N : U \to \mathbb{R}^2$, $\phi_S : V \to \mathbb{R}^2$, see Fig. 7.2, are given by

$$\phi_N(x^1, x^2, x^3) = \Big(\frac{2x^1}{1 - x^3}, \frac{2x^2}{1 - x^3}\Big),$$

$$\phi_S(x^1, x^2, x^3) = \Big(\frac{2x^1}{1 + x^3}, \frac{2x^2}{1 + x^3}\Big).$$

It can be shown as an exercise that the atlases $\mathcal{A}$ and $\mathcal{A}'$ are compatible, so they can be extended to the same complete atlas, i.e., the differential manifold structures induced by $\mathcal{A}$ and $\mathcal{A}'$ are the same.

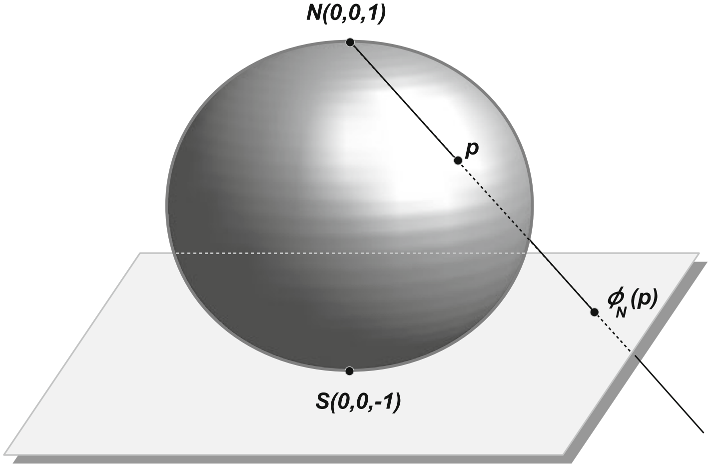

Figure 7.2: The stereographic projection from the north pole

5) Let $M = GL(n, \mathbb{R})$ be the set al all nonsingular $n \times n$ matrices. M is a metric space with the metric

$$d(A, B) = \Big[\sum_{i,j}^{n} (a_{ij} - b_{ij})^2\Big]^{1/2}, \qquad \forall A, B \in M,$$

where $A = (a_{ij})$ and $B = (b_{ij})$. Then M becomes a differential manifold with an atlas consisting of one chart, namely $\phi : M \to \mathbb{R}^{n^2}$,

$$\phi(A) = (a_{11}, a_{12}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn}).$$

We note that $\phi(M)$ is open in $\mathbb{R}^{n^2}$. This follows from considering the continuous mapping $\rho : \mathbb{R}^{n^2} \to \mathbb{R}$ given by $\rho(a_{11}, \dots, a_{nn}) = \det A$. Write $\phi(M) = \rho^{-1}(\mathbb{R}\backslash\{0\})$ for the pre-image of ρ for all nonzero real numbers. Using Proposition 7.2.5, part (a), implies $\phi(M)$ open in $\mathbb{R}^{n^2}$.

6) If M, N are differentiable manifolds of dimensions m and n, respectively, then $M \times N$ can be endowed with a structure of differentiable manifold, called the *product manifold*. If $\mathcal{A}_M$ and $\mathcal{A}_N$ are atlases on M and N, respectively, then an atlas $\mathcal{A}_{M\times N}$ on $M \times N$ can be constructed by considering the charts $\big(U \times V, \Psi\big)$, with $\Psi : U \times V \to \mathbb{R}^{n+m}$, $\Psi(x, y) = \big(\phi(x), \psi(y)\big)$, where $(U, \phi) \in \mathcal{A}_M$ and $(V, \psi) \in \mathcal{A}_N$.

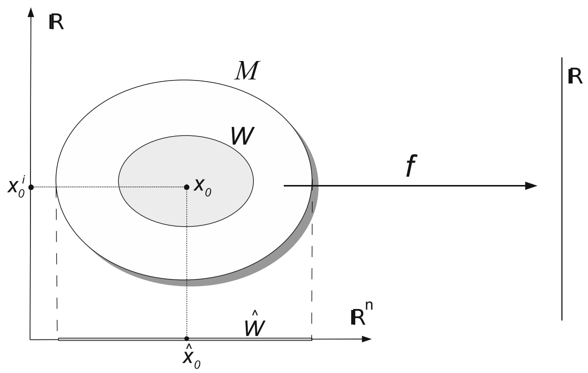

Figure 7.3: By the Implicit Functions Theorem $x^i = g(\hat{x})$ for any $x \in W$

The torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ and the cylinder $\mathbb{S}^1 \times (0,1)$ are two usual examples of product manifolds.

7) Consider the set $M = f^{-1}(0) = \{x \in \mathbb{R}^{n+1}; f(x) = 0\}$, where $f : \mathbb{R}^{n+1} \to \mathbb{R}$ is a C^∞-differentiable function (i.e., a function for which the partial derivatives exist for any order), such that

$$(grad\ f)(x) = \Big(\frac{\partial f}{\partial x^1}(x), \dots, \frac{\partial f}{\partial x^{n+1}}(x)\Big) \neq 0, \quad \forall x \in M.$$

Then M is a differentiable manifold of dimension n, called the *hypersurface defined by* f.

The charts in this manifold are constructed as in the following. Consider a point $x^0 \in M$. Since $(grad\ f)(x^0) \neq 0$, there is an $i \in \{1, \dots, n+1\}$ such that $\dfrac{\partial f}{\partial x^i}(x^0) \neq 0$. By the Implicit Function Theorem, there is an open set V around x^0 such that the equation $f(x^1, \dots, x^{n+1}) = 0$ can be solved uniquely for x^i as $x^i = g(\hat{x})$, where $\hat{x} = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1})$ and $g : \hat{V} \to \mathbb{R}$ is a differentiable function, see Fig. 7.3. Let $U = V \cap M$, and consider $\phi : I \to \mathbb{R}^n$ given by $\phi(x) = \hat{x}$. Then (U, ϕ) is a chart about the point x^0. The set of all charts of this type produces an atlas on M. The compatibility between these charts is left as an exercise to the reader.

This is an effective and practical way of constructing differentiable manifolds. For instance, if consider $f(x^1, \dots, x^{n+1}) = \sum_{k=1}^n (x^k)^2 - 1$, then $S^n = f^{-1}(0)$ is the n-dimensional sphere of radius 1.

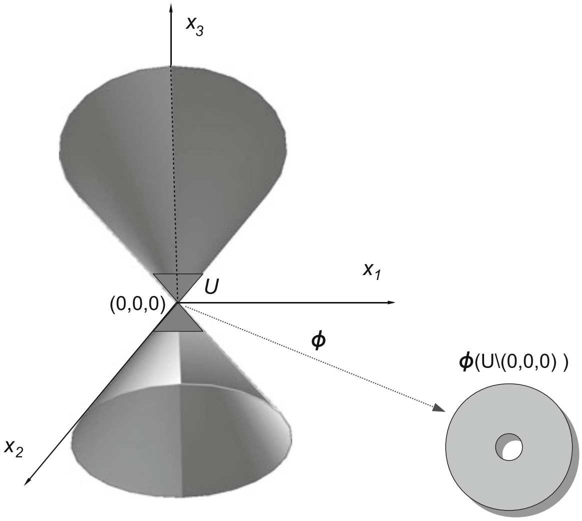

Figure 7.4: The cone is not a differentiable manifold

It is worth noting that the regularity condition $(grad\ f)(x) \neq 0$, for all $x \in M$, in general cannot be waived. For instance, if $f(x^1, x^2, x^3) = (x^1)^2 + (x^2)^2 - (x^3)^2$, then $\mathcal{C} = f^{-1}(0)$ is a cone in $\mathbb{R}^3$. We have that $(grad\ f)(x) = (2x^1, 2x^2, -2x^3)$ vanishes for $x = 0$. As a consequence, the cone $\mathcal{C} = \{(x^1)^2 + (x^2)^2 = (x^3)^2\}$ is not necessarily differentiable manifold. We investigate this by considering a chart (U, ϕ) around the origin $(0, 0, 0)$. Then $V = U\backslash(0, 0, 0)$ has two connected components, while $\phi(V) = \phi(U)\backslash\phi(0, 0, 0)$ has only one component, fact that leads to a contradiction, see Fig. 7.4. Hence the cone $\mathcal{C}$ is not a differentiable manifold.

7.4 Tangent Space

Before defining the concept of tangent vector, we need to introduce the notion of differentiable function on a manifold. We assume well-known from Calculus the concept of a differentiable function on $\mathbb{R}^n$. Since the differentiability has a local character, in the case of differentiable manifolds the function is required to be differentiable in a local chart.

Definition 7.4.1 *A function $f : M \to \mathbb{R}$ is said to be differentiable if for any chart (U, ϕ) on M the function $f \circ \phi^{-1} : \phi(U) \to \mathbb{R}$ is* *differentiable. The set of all differentiable functions on the manifold M will be denoted by $\mathcal{F}(M)$.*

The notion of "differentiable" is not made too precise on the degree of smoothness. It can mean C^∞ or just C^k-differentiable, for some $k \geq 1$, which depends on the nature of the problem.

Since a vector on $\mathbb{R}^n$ at a point can serve as a directional derivative of functions in $\mathcal{F}(\mathbb{R}^n)$, a similar idea can be used when defining the tangent vector on a manifold.

Definition 7.4.2 *A tangent vector of M at a point $p \in M$ is a function $X_p : \mathcal{F}(M) \to \mathbb{R}$ such that*

i) X_p is $\mathbb{R}$-linear

$$X_p(af + bg) = aX_p(f) + bX_p(g), \forall a, b \in \mathbb{R}, \forall f, g \in \mathcal{F}(M);$$

ii) the Leibniz rule is satisfied

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g), \quad \forall f, g \in \mathcal{F}(M). \quad (7.4.1)$$

Definition 7.4.3 *Consider a differentiable curve $\gamma : (-\epsilon, \epsilon) \to M$ on the manifold M, with $\gamma(0) = p$. The tangent vector*

$$X_p(f) = \frac{d(f \circ \gamma)}{dt}(0), \qquad \forall f \in \mathcal{F}(M) \quad (7.4.2)$$

is called the tangent vector to $\gamma(-\epsilon, \epsilon)$ at $p = \gamma(0)$ and is denoted by $\dot{\gamma}(0)$.

We note that the derivative in formula (7.4.2) is the usual derivative of the real-valued function $f \circ \gamma : (-\epsilon, \epsilon) \to \mathbb{R}$. Also, X_p satisfies the conditions from the definition of the tangent vector. Condition $i)$ follows from the linearity of the derivative d/dt, while condition $ii)$ is an application of the product rule. Sometimes, the vector $\dot{\gamma}(0)$ is called the *velocity vector* of γ at p.

Now consider the particular case of the ith *coordinate curve* γ. This means there is a chart (U, ϕ) around $p = \gamma(0)$ in which $\phi\big(\gamma(t)\big) = (x_0^1, \dots, x^i, \dots, x_0^n)$, where $\phi(p) = (x_0^1, \dots, x_0^i, \dots, x_0^n)$. Then the tangent vector to γ

$$\dot{\gamma}(0) = \frac{\partial}{\partial x^i}\Big|_p$$

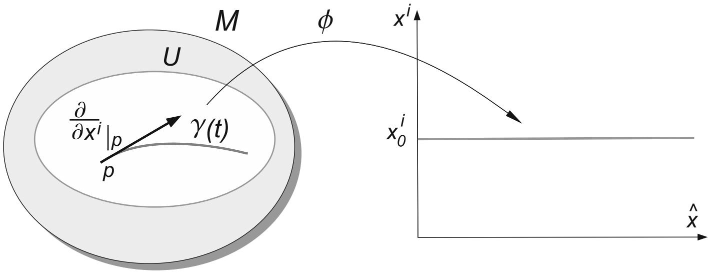

Figure 7.5: The geometric interpretation of the coordinate vector field $\dfrac{\partial}{\partial x^i}{}_{|p}$

is called a *coordinate tangent vector* at p, see Fig. 7.5. This can be defined equivalently as a derivation

$$\frac{\partial}{\partial x^i}\Big|_p(f) = \frac{\partial(f\circ\phi^{-1})}{\partial u^i}(\phi(p)), \qquad \forall f \in \mathcal{F}(M), \tag{7.4.3}$$

where $\phi = (x^1, \ldots, x^n)$ is a system of coordinates around p and $u^1, \ldots, u^n$ are the coordinate functions on $\mathbb{R}^n$.

Definition 7.4.4 *The set of all tangent vectors at p to M is called the tangent space of M at p, and is denoted by T_pM.*

T_pM is a vectorial space of dimension n with a basis given by the coordinate tangent vectors $\left\{\frac{\partial}{\partial x^1}\Big|_p, \ldots, \frac{\partial}{\partial x^n}\Big|_p\right\}$. For a detailed proof of this fact the reader can consult, for instance, Millman and Parker [58]. The tangent space T_pM can be also visualized geometrically as the set of velocities at p along all curves passing through this point.

Using the aforementioned basis any vector $V \in T_pM$ can be written locally as $V = \sum_i V^i \frac{\partial}{\partial x^i}\big|_p$, where $V^i = V(x^i) \in \mathbb{R}$ are called the components of V with respect to the system of coordinates $(x^1, \ldots, x^n)$.

It is worth noting that if the vector V is written with respect to a new system of coordinates $(\bar{x}^1, \ldots, \bar{x}^n)$ as $V = \sum_i \bar{V}^i \frac{\partial}{\partial \bar{x}^i}\big|_p$, then the components in the two coordinates systems are related by

$$\bar{V}^k = \sum_{i=1}^n \frac{\partial \bar{x}^k}{\partial x^i} V^i. \tag{7.4.4}$$

It is also worthy to note that the change of coordinates matrix $\left(\frac{\partial \bar{x}^k}{\partial x^i}\right)_{i,k}$ is nonsingular, fact implied by the nonvanishing Jacobian[2] of a diffeomorphism, as stated by the Inverse Function Theorem.

The tangent vector X_p acts on differentiable functions f on M as

$$X_p f = \sum_{i=1}^{n} X^i(p) \frac{\partial f}{\partial x^i}_{|p}.$$

Definition 7.4.5 *A vector field X on M is a smooth map X that assigns to each point $p \in M$ a vector X_p in T_pM. For any function $f \in \mathcal{F}(M)$ we define the real-valued function $(Xf)_p = X_p f$. By "smooth" we mean the following: for each $f \in \mathcal{F}(M)$ then $Xf \in \mathcal{F}(M)$.*

Vector fields can be visualized as fields of forces on velocities for ocean currents, air currents, or convection currents, or river flows. They are important geometric objects used to model the dynamics on a manifold.
The set of all vector fields on M will be denoted by $\mathcal{X}(M)$. In a local system of coordinates a vector field is given by $X = \sum X^i \frac{\partial}{\partial x^i}$, where the components $X^i \in \mathcal{F}(M)$ because they are given by $X^i = X(x^i)$, $1 \leq i \leq n$, where x^i is the ith coordinate function of the chart.

We show next that to each vector field we can associate a family of non-intersecting curves. Given a vector field X, consider the ordinary differential equations system

$$\frac{dc^k}{dt}(t) = X^k_{c(t)}, \qquad 1 \leq k \leq n. \tag{7.4.5}$$

Standard theorems of existence and uniqueness of ODEs imply that the system (7.4.5) can be solved locally around any point $x_0 = c(0)$.

Theorem 7.4.6 *Given $x_0 \in M$ and a nonzero vector field X on an open set $U \subset M$, then there is an $\epsilon > 0$ such that the system (7.4.5) has a unique solution $c : [0, \epsilon) \to U$ satisfying $c(0) = x_0$.*

The solution $t \to c(t)$ is called the *integral curve* associated with the vector field X through the point x_0. The integral curves play

[2] If $\phi(x) = (\phi^1(x), \dots, \phi^n(x))$ is a function of n variables $x^1, \dots, x^n$, the Jacobian is the determinant of the matrix $\left(\frac{\partial \phi^j}{\partial x^k}\right)_{jk}$.

an important role in describing the evolution of a dynamical system modeled on the manifold. An effective description of the evolution of a dynamical system is usually done using conservation laws, i.e., relations whose value remains invariant along the integral curves of a vector field.

Definition 7.4.7 *A function $f \in \mathcal{F}(M)$ is called a first integral of motion for the vector field X if it remains constant along the integral curves of X, i.e,*

$$f\big(c(t)\big) = constant, \quad 0 \leq t \leq \epsilon,$$

where $c(t)$ verifies (7.4.5).

Proposition 7.4.8 *Let $f \in \mathcal{F}(M)$, with M differentiable manifold. Then f is a first integral of motion for the vector field X if and only if $X_{c(t)}(f) = 0$.*

Proof: Consider a local system of coordinates $(x^1, \dots, x^n)$ in which the vector field writes as $X = \sum_k X^k \frac{\partial}{\partial x^k}$. Then

$$\begin{aligned} X_{c(t)}(f) &= \sum_k X^k_{c(t)} \frac{\partial f}{\partial x^k} = \sum_k \frac{dc^k}{dt}(t) \frac{\partial f}{\partial x^k} \\ &= \frac{d}{dt} f\big(c(t)\big). \end{aligned}$$

Then $X_{c(t)}(f) = 0$ if and only if $\frac{d}{dt} f\big(c(t)\big) = 0$, which is equivalent to $f\big(c(t)\big) =$ constant, $0 \leq t \leq \epsilon$, with ϵ small enough such that $c((0, \epsilon))$ is included in the initially considered chart. ∎

7.5 Lie Bracket

This section deals with an important operation on vector fields, called the *Lie bracket*, which is given by $[\ ,\ ] : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$,

$$[X, Y]_p f = X_p(Yf) - Y_p(Xf), \quad \forall f \in \mathcal{F}(M),\ p \in M. \tag{7.5.6}$$

The Lie bracket will be used in later sections of the chapter to define the concepts of torsion and curvature of a linear connection, as well as the differential of a 1-form.

The vector fields X and Y *commute* if $[X, Y] = 0$. The Lie bracket $[X, Y]$, which at first sight looks to be a differential operator of second degree, turns out to be a vector field (a first order differential operator), which measures the noncommutativity between vector fields. In local coordinates, the Lie bracket takes the form (see Problem 7.3.)

$$[X,Y] = \sum_{i,j=1}^{n} \Big(\frac{\partial Y^i}{\partial x^j} X^j - \frac{\partial X^i}{\partial x^j} Y^j\Big) \frac{\partial}{\partial x^i}. \tag{7.5.7}$$

The bracket satisfies the following properties

1) $\mathbb{R}$-bilinearity:

$$\begin{aligned} [aX + bY, Z] &= a[X,Z] + b[Y,Z], \\ [Z, aX + bY] &= a[Z,X] + b[Z,Y], \qquad \forall a, b \in \mathbb{R}; \end{aligned}$$

2) Skew-symmetry:

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

3) The cyclic sum is zero (Jacobi identity):

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0;$$

4) The Lie bracket is not $\mathcal{F}(M)$-linear, because $[fX, gY] \neq fg[X,Y]$. We have instead

$$[fX, gY] = fg[X,Y] + f(Xg)Y - g(Yf)X\ , \quad \forall f, g \in \mathcal{F}(M).$$

Example 7.5.1 *Consider on $\mathbb{R}^2$ the vector fields $X = \partial_{x^1}$, $Y = x^1\partial_{x^2}$, called the Grushin vector fields. Then $[X,Y] = \partial_{x^2} \neq 0$, and hence X and Y do not commute.*

7.6 Differentiable Maps

The concept of differentiability on a manifold is defined locally with respect to charts.

Definition 7.6.1 *A map $F : M \to N$ between two manifolds M and N is differentiable about $p \in M$ if for any charts (U, ϕ) on M about p and $(V, \psi) \in N$ about $F(p)$, the map $\psi \circ F \circ \phi^{-1}$ is differentiable from $\phi(U) \subset \mathbb{R}^m$ to $\psi(V) \subset \mathbb{R}^n$, see Fig. 7.6.*

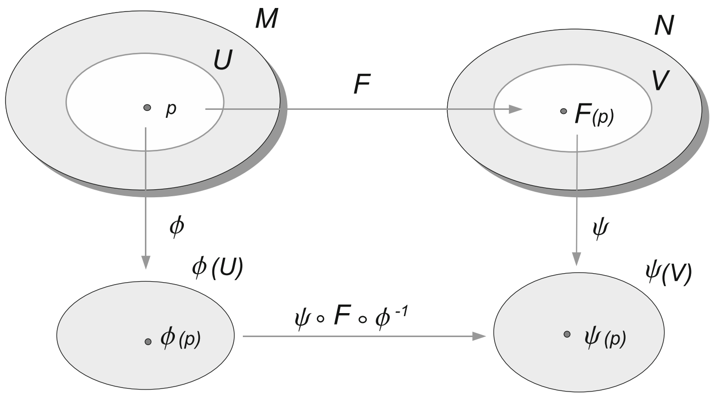

Figure 7.6: The diagram of a differentiable function

Definition 7.6.2 *Let $F : M \to N$ be a differentiable map. For every $p \in M$, the differential map dF at p is defined by*

$$dF_p : T_pM \to T_{F(p)}N$$

$$(dF_p)(v)(f) = v(f \circ F)\ , \quad \forall v \in T_pM\ , \ \forall f \in \mathcal{F}(N). \tag{7.6.8}$$

The picture can be seen in Fig. 7.7. A few important properties of the differential of a map at a point, dF_p, are given in the following:

1) dF_p is an $\mathbb{R}$-linear application between the tangent spaces T_pM and $T_{F(p)}N$:

$$\begin{aligned} dF_p(v+w) &= dF_p(v) + dF_p(w), \quad \forall v, w \in T_pM; \\ dF_p(\lambda v) &= \lambda dF_p(v), \quad \forall v \in T_pM, \ \forall \lambda \in \mathbb{R}. \end{aligned}$$

2) Let $\{\frac{\partial}{\partial x^j}\big|_p\}$ and $\{\frac{\partial}{\partial y^j}\big|_{F(p)}\}$ be bases associated with the tangent spaces T_pM and $T_{F(p)}N$. Consider the function $F{=}(F^1, \ldots, F^n)$ and denote by $J_{kj} = \frac{\partial F^k}{\partial x^j}$ the Jacobian matrix of F with respect to the charts $(x^1, \ldots x^m)$ and $(y^1, \ldots, y^n)$ on M and N, respectively. Then dF_p can be represented locally by

$$dF_p\Big(\frac{\partial}{\partial x^j}_{|p}\Big) = \sum_{k=1}^{n} J_{kj}(p) \frac{\partial}{\partial y^k}_{|F(p)}. \tag{7.6.9}$$

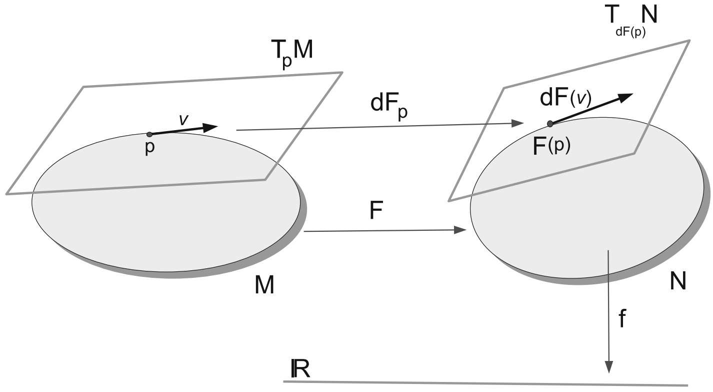

Figure 7.7: The differential of a map

3) Assume $\dim M = \dim N = n$. Then the following conditions are equivalent:

   (i) $dF_p : T_pM \to T_{F(p)}N$ is an isomorphism of vectorial spaces;

   (ii) F is a local diffeomorphism in a neighborhood of p;

   (iii) There are two charts $(x^1, \dots, x^n)$ and $(y^1, \dots, y^n)$ on M around p and on N around $F(p)$, respectively, such that the associated Jacobian is non-degenerate, i.e. $\det J_{kj}(p) \neq 0$.

   The foregoing assertion is usually called the Inverse Function Theorem on manifolds. For a proof the reader can consult the comprehensive book of Spivak [77].

4) Let $F : M \to N$ be a differentiable map. Then the differential dF commutes with the Lie bracket

$$dF_p[v, w] = [dF_p(v), dF_p(w)], \qquad \forall v, w \in T_pM.$$

7.7 1-Forms

The differential of a function $f \in \mathcal{F}(M)$ is defined at any point p by $(df)_p : T_pM \to \mathbb{R}$,

$$(df)_p(v) = v(f) \qquad \forall v \in T_pM. \tag{7.7.10}$$

In local coordinates $(x^1, \dots, x^n)$ this takes the form $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$, where $\{dx^i\}$ is the dual basis of $\{\frac{\partial}{\partial x^i}\}$ of T_pM, i.e.

$$dx^i\Big(\frac{\partial}{\partial x^j}\Big) = \delta^i_j,$$

where δ^i_j denotes the Kronecker symbol. The space spanned by $\{dx^1, \dots, dx^n\}$ is called the *cotangent space* of M at p, and is denoted by T_p^*M. The elements of T_p^*M are called *covectors*. The differential df is an example of 1-form.

In general, a one form ω on the manifold M is a mapping which assigns to each point $p \in M$ an element $\omega_p \in T_p^*M$. A 1-form can be written in local coordinates as

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^i, \tag{7.7.11}$$

where $\omega_i = \omega(\frac{\partial}{\partial x^i})$ is the ith coordinate of the form with respect to the basis $\{dx^i\}$. The set of all 1-forms on the manifold M will be denoted by $\mathcal{X}^*(M)$.

The interested reader can find more details about differential forms in DoCarmo [36].

7.8 Tensors

Let T_pM and T_p^*M be the tangent and the cotangent spaces of M at p. We adopt the following useful notations

$$(T_p^*M)^r = \underbrace{T_p^*M \times \cdots \times T_p^*M}_{r \text{ times}}, \quad (T_pM)^s = \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}}.$$

Definition 7.8.1 *A tensor of type (r, s) at $p \in M$ is an $\mathcal{F}(M)$-multilinear function $T : (T_p^*M)^r \times (T_pM)^s \to \mathbb{R}$.*
A tensor field $\mathcal{T}$ of type (r, s) is a differential map, which assigns to each point $p \in M$ an $(r, s)-$tensor $\mathcal{T}_p$ on M at the point p.

Since $\{dx^{j_1} \otimes \cdots \otimes dx^{j_r}\}_{j_1 < \cdots < j_r}$ and $\{\frac{\partial}{\partial x_{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_s}}\}_{i_1 < \cdots < i_s}$ are bases in the vectorial spaces $(T_p^*M)^r$ and $(T_pM)^s$, respectively, the tensor field $\mathcal{T}$ can be written using local coordinates as (with summation over repeated indices)

$$\mathcal{T} = \mathcal{T}^{i_1 i_2 \dots i_r}_{j_1 j_2 \dots j_s} \, dx^{j_1} \otimes \cdots \otimes dx^{j_s} \otimes \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}}, \tag{7.8.12}$$

where "$\otimes$" stands for the usual tensorial product. This means that $\mathcal{T}$ acts on r 1-forms and s vector fields as

$$\begin{aligned} & \mathcal{T}(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \\ = \; & \mathcal{T}^{i_1 i_2 \dots i_r}_{j_1 j_2 \dots j_s} \, dx^{j_1}(X_1) \otimes \cdots \otimes dx^{j_s}(X_s) \otimes \frac{\partial}{\partial x^{i_1}}(\omega^1) \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}(\omega^r) \\ = \; & \mathcal{T}^{i_1 \dots i_r}_{j_1 \dots j_s} X_1^{j_1} \dots X_s^{j_s} \omega^1_{i_1} \dots \omega^r_{i_r}. \end{aligned}$$

We say the tensor $\mathcal{T}$ is s covariant and r contravariant. It is worth noting the following particular examples of tensors:

1. Any 1-form ω is a tensor of type $(0, 1)$. For any vector field X

$$\omega(X) = \omega_i dx^i(X) = \omega_i dx^i(X^j \frac{\partial}{\partial x^j}) = \omega_i X^i,$$

   with summation in the repeated index. In particular, the differential of a function, df, is a $(0, 1)$-tensor.

2. Any vector field X is a $(1, 0)$-tensor on M, with

$$X(\omega) = \omega(X) = \omega_i X^i, \qquad \forall \omega.$$

3. An s-differentiable form is a skew-symmetric tensor of type $(0, s)$. In particular, a 2-form is a 2-covariant tensor Ω whose coordinates satisfy $\Omega_{ij} = -\Omega_{ji}$

4. A volume form on an n-dimensional manifold is an n-form, i.e., a skew-symmetric tensor of type $(0, n)$.

In order to show that $\mathcal{T}$ is a tensor, in practice we check the $\mathcal{F}(M)$-linearity in each argument. For instance, if $\mathcal{T}$ is 2-covariant, then we need to show that for any $f_1, f_2 \in \mathcal{F}(M)$ and vector fields X_1, X_2, Y_1, Y_2 we have

$$\mathcal{T}(f_1 X_1, f_2 X_2) = f_1 f_2 \mathcal{T}(X_1, X_2)$$

$$\mathcal{T}(X_1 + Y_1, X_2) = \mathcal{T}(X_1, X_2) + \mathcal{T}(Y_1, X_2)$$

$$\mathcal{T}(X_1, X_2 + Y_2) = \mathcal{T}(X_1, X_2) + \mathcal{T}(X_1, Y_2).$$

In the case of a symmetric tensor, $\mathcal{T}(X, Y) = \mathcal{T}(Y, X)$, it suffices to show the previous relations only in the first argument.

If we like to show that a tensor, or a tensorial expression vanishes, then in the virtue of the previous properties it suffices to show that it vanishes in just one system of coordinates.

7.9 Riemannian Manifolds

A Riemannian manifold is a manifold on which one is able to measure distances between points, angles between vectors, length of curves and volumes. Roughly speaking, it is a manifold endowed with a *metric structure*. The precise definitions are stated in the following.

Definition 7.9.1 *A Riemannian metric g on a differentiable manifold M is a symmetric, positive definite 2-covariant tensor field.*
A Riemannian manifold is a differentiable manifold M endowed with a Riemannian metric g.

A Riemannian manifold will be denoted from now on by the pair (M, g). The Riemannian metric g can be considered as a positive definite scalar product $g_p : T_pM \times T_pM \to \mathbb{R}$ that depends differentially on the point $p \in M$. In local coordinates we write

$$g = g_{ij}\, dx^i dx^j, \qquad (7.9.13)$$

with $g_{ij} = g_{ji} = g(\partial_i, \partial_j)$. The Riemannian metric g acts on a pair of vector fields as $g(X, Y) = g_{ij}X^iY^j$, where we assume the summation convention over the repeated indices.

The most obvious example of Riemannian manifold is the n-dimensional Euclidean space $\mathbb{E}^n = (\mathbb{R}^n, \delta_{ij})$, which induces the scalar product $\langle X, Y\rangle = \sum_i X^iY^i$.

It can be proved that any differentiable manifold has a Riemannian metric structure. The idea of this construction is that a Riemannian manifold can be seen as a collection of local charts that resemble the Euclidean space $\mathbb{E}^n$. Using methods of global analysis, one can unify this local metrics into a global defined metric tensor, see, for instance, Auslander and MacKenzie [9].

A metric g induces a natural bijective correspondence between 1-forms and vector fields on a Riemannian manifold M. If X is a vector field, then one may associate with it the 1-form ω such that

$$\omega(Y) = g(Y, X), \qquad \forall Y \in \mathcal{X}(M). \qquad (7.9.14)$$

In local coordinates this becomes $\omega_k = g_{jk}X^j$, where $\omega = \omega_i dx^i$ and $X = X^j \frac{\partial}{\partial x^j}$.

7.10 Linear Connections

A linear connection allows differentiation of a function, a vector field, or, in general, a tensor with respect to a given vector field. It can be seen as an extension of the directional derivative from the Euclidean case. The precise definition follows. Recall that $\mathcal{X}(M)$ denotes the set of vector fields on M.

Definition 7.10.1 *A linear connection ∇ on a differentiable manifold M is a map $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ with the following properties:*

1) *$\nabla_X Y$ is $\mathcal{F}(M)$-linear in X;*
2) *$\nabla_X Y$ is $\mathbb{R}$-linear in Y;*
3) *it satisfies the Leibniz rule:*

$$\nabla_X(fY) = (Xf)Y + f\,\nabla_X Y, \quad \forall f \in \mathcal{F}(M).$$

For fixed vector fields X and Y, the object $\nabla_X Y$ is also a vector field on M, which measures the vector rate change of Y in the direction of X. In a local coordinates system $(x^1, \ldots, x^n)$ we can write

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k,$$

where Γ_{ij}^k are the coordinates of the connection with respect to the local base $\{\partial_i\}$, where $\partial_i = \dfrac{\partial}{\partial x^i}$. If $X = X^i\partial_i$ and $Y = Y^j\partial_j$, then a straightforward computation provides the formula

$$\nabla_X Y = (\nabla_X Y)^k \partial_k,$$

where $(\nabla_X Y)^k = X^i\Big(\partial_i Y^k + Y^j\Gamma_{ij}^k\Big)$, with summation over i and j.

An example of a linear connection on the Euclidean space $\mathbb{R}^n$ is given by $\overline{\nabla}_X Y = X(Y^j)e_j$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the jth basis vector on $\mathbb{R}^n$ and $Y = (Y^1, \ldots, Y^n) = Y^j e_j$. The coordinates of this connection are zero, $\overline{\Gamma}_{ij}^k = 0$.

A connection can be also used to differentiate tensors. If T is an r-covariant tensor field, we may differentiate it along a vector field X with respect to the linear connection ∇ as

$$(\nabla_X T)(Y_1, \ldots, Y_r) = X\,T(Y_1, \ldots, Y_r) - \sum_{i=1}^{n} T(Y_1, \ldots, \nabla_X Y_i, \ldots, Y_r). \tag{7.10.15}$$

In particular, we have the following concept:

Definition 7.10.2 *Let g be the Riemannian metric tensor. A linear connection ∇ is called metric connection if g is parallel with respect to ∇, i.e.,*

$$\nabla_Z g = 0, \quad \forall\, Z \in \mathcal{X}(M). \tag{7.10.16}$$

This can be stated equivalently as

$$Z\, g(X,Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y), \quad \forall\, X,Y,Z \in \mathcal{X}(M). \tag{7.10.17}$$

Let $X = \partial_i$, $Y = Y^j\partial_j$ and $Z = Z^k\partial_k$. Choosing $X = \dfrac{\partial}{\partial x^i}$, $Y = \dfrac{\partial}{\partial x^j}$, and $Z = \dfrac{\partial}{\partial x^k}$, a straightforward computation transforms (7.10.17) into

$$\partial_k g_{ij} = \Gamma^p_{ki} g_{pj} + \Gamma^r_{kj} g_{ir}. \tag{7.10.18}$$

It is worth noting that given the metric coefficients g_{ij}, there are $\frac{n^2(n+1)}{2}$ linear equations in Γ^p_{ki} of type (7.10.18). The total number of unknowns Γ^p_{ki} is n^3, where n is the dimension of the manifold. The excess $\epsilon(n) = n^3 - \frac{n^2(n+1)}{2} = \frac{n^2(n-1)}{2}$ represents the number of arbitrary functions the family of linear connections depends on. For instance, on a curve there is only one linear connection, because $\epsilon(1) = 0$, but on a surface, the family of linear connections depends on $\epsilon(2) = 2$ arbitrary functions.

A linear connection is described by two other tensors, the *torsion* and *curvature*, which are defined shortly.

Definition 7.10.3 *Let ∇ be a linear connection. The torsion is defined as*

$$T : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{7.10.19}$$

The torsion measures the noncommutativity of the derivation with respect to two vector fields. The last term, $[X,Y]$, is necessary because it confers tensorial properties to $T(\cdot,\cdot)$:

$$\begin{aligned} T(fX, hY) &= fhT(X,Y), \quad \forall X,Y,Z \in \mathcal{X}(M), \forall f,h \in \mathcal{F}(M) \\ T(X, Y+Z) &= T(X,Y) + T(X,Z). \end{aligned}$$

Since $T(X,Y) = -T(Y,X)$, then T is a 2-covariant skew-symmetric tensor. Since in local coordinates we have

$$\begin{aligned} T_{ij} &= T(\partial_i, \partial_j) = \nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i - \underbrace{[\partial_i, \partial_j]}_{=0} \\ &= \left(\Gamma_{ij}^k - \Gamma_{ji}^k\right)\partial_k, \end{aligned}$$

it follows that the torsion coordinates are given by $T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k$. A connection ∇ is called *torsion-free* if $T = 0$. This can be described equivalently as $\Gamma_{ij}^k = \Gamma_{ji}^k$, which is a symmetry relation for the connection coefficients. This is the reason why these type of connections are also called *symmetric*. There are exactly $\frac{n^2(n-1)}{2}$ equations of type $T_{ij}^k = T_{ji}^k$, which is exactly the excess $\epsilon(n)$. If these are considered as constraints applied to the linear system of equations (7.10.18), it follows that there is only one solution to this system. This leads to a unique linear connection, which is both symmetric and metric. We shall get in more detail regarding this issue later, when discussing the Levi–Civita connection.

Definition 7.10.4 *The curvature of the linear connection ∇ is given by*

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$
$$R(X,Y,Z) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z. \qquad (7.10.20)$$

If we write the curvature as

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

it follows that R is a measure of the noncommutativity of the connections with respect to X and Y. It can be shown that R satisfies the following properties

$$\begin{aligned} R(f_1X, f_2Y, f_3Z) &= f_1f_2f_3R(X,Y,Z) \\ R(X_1 + X_2, Y, Z) &= R(X_1,Y,Z) + R(X_2,Y,Z) \\ R(X, Y_1 + Y_2, Z) &= R(X,Y_1,Z) + R(X,Y_2,Z) \\ R(X,Y,Z_1 + Z_2) &= R(X,Y,Z_1) + R(X,Y,Z_2), \end{aligned}$$

for all $f_i \in \mathcal{F}(M)$ and $X_i, Y_j, Z_k \in \mathcal{X}(M)$, so that R becomes a 3-covariant tensor field. The tensor R is skew-symmetric in the first pair of arguments, i.e., $R(X,Y,Z) = -R(Y,X,Z)$. Since the first pair is more special, the curvature tensor is sometimes denoted by $R(X,Y)Z$. In a local system of coordinates we write

$$R\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big)\frac{\partial}{\partial x^k} = R^p_{ijk}\frac{\partial}{\partial x^p}.$$

Definition 7.10.5 *The Ricci curvature associated with the linear connection* ∇ *is given by*

$$Ric : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{F}(M),$$

$$Ric(Y,Z) = Trace\Big(X \to R(X,Y)Z\Big).$$

This means that if $\{E_1, \dots E_n\}$ is an orthonormal set of tangent vectors at p, then $Ric(X,Y)_p = \sum_{j=1}^n g\big(R(E_j,X,Y),E_j\big)$. In local coordinates we write $R_{ij} = Ric(\partial_i, \partial_j)$. We can show that $R_{ij} = R^k_{ikj}$, with summation over k, see Problem 7.14. It is worth noting that *Ric* is a 2-covariant tensor. It will play an important role in the study of equiaffine connections in Chap. 9.

7.11 Levi–Civita Connection

One of the most remarkable facts of Riemannian geometry is the existence and uniqueness of a metric connection that has zero torsion. This is called the *Levi–Civita connection* of the Riemannian manifold (M,g), see, for instance, O'Neill [66]. Sometimes this is also called the *Riemannian connection* and will be denoted throughout the book by $\nabla^{(0)}$. For the purpose of this section we shall keep the notation ∇.

The next theorem, also known as the fundamental lemma of Riemannian geometry, provides the Levi–Civita connection as an explicit expression in terms of the Riemannian metric g. This is an useful result that allows to eliminate the connection from a formula and write it in terms of the Riemannian metric only.

Theorem 7.11.1 *On a Riemannian manifold there is a unique torsion-free, metric connection* ∇*. Furthermore,* ∇ *is given by the following Koszul formula*

$$\begin{aligned} 2g(\nabla_X Y, Z) &= X\,g(Y,Z) + Y\,g(X,Z) - Z\,g(X,Y) \\ &\quad + g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X). \end{aligned} \qquad (7.11.21)$$

Proof: The proof has two parts, the existence and uniqueness.

Existence: We shall show that connection ∇ defined by formula (7.11.21) is a metric and torsion-free connection.

First we need to show that ∇ is a linear connection. Using the properties of vector fields and Lie brackets we can show by a direct computation that

$$2g(\nabla_{fX}Y, Z) = 2fg(\nabla_X Y, Z), \quad \forall Z \in \mathcal{X}(M),$$

so $\nabla_{fX}Y = f\nabla_X Y$, $\forall X, Y \in \mathcal{X}(M)$, i.e., ∇ is $\mathcal{F}(M)$-linear in the first argument. Next we check the second property of connections:

$$\begin{aligned} 2g(\nabla_X(fY), Z) &= X\,g(fY, Z) + fY\,g(X, Z) - Z\,g(X, fY) \\ &\quad + g([X, fY], Z) - g([X, Z], fY) - g([fY, Z], X) \\ &= X(f)g(Y, Z) + fXg(Y, Z) + fYg(X, Z) \\ &\quad - Z(f)g(X, Y) - fZg(X, Y) \\ &\quad + fg([X, Y], Z) + X(f)g(Y, Z) - fg([X, Z], Y) \\ &\quad - fg([Y, Z], X) + Z(f)g(Y, X) \\ &= 2f\,g(\nabla_X Y, Z) + 2X(f)g(Y, Z) \\ &= 2g(f\nabla_X Y + X(f)Y, Z). \end{aligned}$$

Dropping the Z-argument yields Leibniz formula. Therefore, ∇ is a linear connection.

The next computation verifies that the connection is torsion-free. Using (7.11.21) yields

$$\begin{aligned} 2g(T(X, Y), Z)) &= 2g(\nabla_X Y, Z) - 2g(\nabla_Y X, Z) - 2g([X, Y], Z) \\ &= Xg(Y, Z) + Yg(X, Z) - Zg(X, Y) \\ &\quad + g([X, Y], Z) - g([X, Z], Y) - g([Y, Z], X) \\ &\quad - Yg(X, Z) - Xg(Y, Z) + Zg(Y, X) \\ &\quad - g([Y, X], Z) + g([Y, Z], X) + g([X, Z], Y) \\ &\quad - 2g([X, Y], Z) \\ &= g(2[X, Y] - 2[X, Y], Z) = 0, \qquad \forall Z \in \mathcal{X}(M). \end{aligned}$$

Dropping the vector field Z and using that $g(\cdot, \cdot)$ is non-degenerate yields $T(X, Y) = 0$, for all $X, Y \in \mathcal{X}(M)$.

Applying formula (7.11.21) twice and then cancelling in pairs, we have

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

Therefore $g(\nabla_Z X, Y) + g(X, \nabla_Z Y) = Zg(X,Y)$, i.e., ∇ is a metric connection.

Uniqueness: We need to prove that any metric and symmetric connection ∇ is given by formula (7.11.21). It suffices to do the verification in a local system of coordinates $(x^1, \dots, x^n)$. Let $X = \partial_i$, $Y = \partial_j$, $Z = \partial_k$. Using $\Gamma_{ij}^k = g(\nabla_{\partial_i}\partial_j, \partial_k)$ and $g_{ij} = g(\partial_i, \partial_j)$, then formula (7.11.21) becomes

$$2\Gamma_{ij}^p g_{pk} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}. \qquad (7.11.22)$$

Writing that ∇ is a metric connection in three different ways, using cyclic permutation of indices, see formula (7.10.18), we have

$$\begin{aligned}
\partial_i g_{jk} &= \Gamma_{ij}^p g_{pk} + \Gamma_{ik}^r g_{jr} \\
\partial_j g_{ki} &= \Gamma_{jk}^p g_{pi} + \Gamma_{ji}^r g_{kr} \\
\partial_k g_{ij} &= \Gamma_{ki}^p g_{pj} + \Gamma_{kj}^r g_{ir}.
\end{aligned}$$

Adding the first two equations and subtracting the last, using the symmetry $\Gamma_{ij}^k = \Gamma_{ji}^k$, yields exactly the Eq. (7.11.22). This ends the proof of uniqueness. ∎

Solving for the connection coefficient in (7.11.22) we obtain

$$\Gamma_{ij}^p = \frac{1}{2} g^{pk} \big(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij} \big), \qquad (7.11.23)$$

where (g^{pk}) denotes the inverse matrix of (g_{ij}). The coordinates Γ_{ij}^p of the Levi–Civita connection, see (7.11.23), are called the *Christoffel symbols of second kind*. The *Christoffel symbols of first kind* are obtained lowering the indices

$$\Gamma_{ij,k} = \Gamma_{ij}^p g_{pk}.$$

Conversely, if the coordinates of a linear connection on a Riemannian manifold (M, g) are given by formula (7.11.23), then the connection has to be the Levi–Civita connection.

The curvature tensor of type $(1, 3)$ associated with the Levi–Civita connection by formula (7.10.20) is called the *Riemann curvature tensor* of type $(1, 3)$. If in local coordinates we have $R(\partial_i, \partial_j)\partial_k = R^p_{ijk}\partial_p$, then the coordinate R^P_{ijk} can be expressed in terms of Christoffel symbols as

$$R^r_{ijk} = \partial_i \Gamma^r_{jk} - \partial_j \Gamma^r_{ik} + \Gamma^r_{ih}\Gamma^h_{jk} - \Gamma^r_{jh}\Gamma^h_{ik}.$$

In Riemannian geometry the following $(0, 4)$-type curvature tensor is also useful

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{F}(M),$$

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

If in local coordinates we write $R(\partial_i, \partial_j, \partial_k, \partial_l) = R_{ijkl}$, then we have $R_{ijkl} = R^p_{ijk} g_{pl}$. The coordinates R_{ijkl} satisfy several relations, the most useful being provided in the following:

1. Skew symmetry in the first and second pair:
$$R_{ijkl} = -R_{jikl} = -R_{ijlk}.$$
2. Interchange symmetry between pairs: $R_{ijkl} = R_{klij}$.
3. First Bianchi identity: $R_{ijkl} + R_{iklj} + R_{iljk} = 0$.

Another important 2-covariant tensor is the *Ricci tensor*, which is defined by the contraction

$$Ric(X, Y) = Trace\Big(V \to R(X, V, Y)\Big) = Trace\Big(V \to R(V, X, Y)\Big).$$

It can be shown that the Ricci tensor associated with the Levi–Civita connection is symmetric, $R(X, Y) = R(Y, X)$.

For more details about Calculus and Differential Geometry on differentiable manifolds the reader may consult Spivak [77, 78], and doCarmo [34, 35].

7.12 Problems

7.1. Let $p \in M$ be a point on the differentiable manifold M, and let $\mathcal{V}_p$ be a neighborhood of p. Show that there is a differentiable function $f \in \mathcal{F}(M)$ such that $f(p) = 1$ and $f(x) = 0$ if $x \notin \mathcal{V}_p$.

7.2. Let p be a point on the differentiable manifold M. If $f \in \mathcal{F}(M)$ has a local extremum at p, then $X_p(f) = 0$, for any tangent vector X_p at p.

7.3. (a) Let $X, Y \in \mathcal{X}(M)$ be two vector fields on the differentiable manifold M. Prove that the Lie bracket $[X, Y]$ is a vector field on M, which in local coordinates can be written as

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

where $X = \sum_i X^i \frac{\partial}{\partial x^i}$ and $Y = \sum_i Y^i \frac{\partial}{\partial x^i}$.

(b) Let $M = \mathbb{R}^2$ and consider the vector fields $X = x^1 x^2 \frac{\partial}{\partial x^1}$ and $Y = x^2 \frac{\partial}{\partial x^2}$. Show that $[X, Y] = -x^1 x^2 \frac{\partial}{\partial x^1}$.

7.4. Let (M, g) be a Riemannian manifold. If $\omega = \omega_i dx^i$ is a 1-form, define the vector filed $\omega^\# = \omega^k \partial_{x^k}$, where $\omega^k g_{kr} = \omega_r$. Show that $g(\omega^\#, X) = \omega(X)$, $\forall X \in \mathcal{X}(M)$.

7.5. Show that the following properties of tangent vectors, $X_p \in T_pM$, hold:

(i) $X_p(c) = 0$, for any constant c;

(ii) $X_p(f^2) = 2f X_p(f)$, $\forall f \in \mathcal{F}(M)$;

(iii) If $f, g \in \mathcal{F}(M)$ such that $f(p) = g(p) = 0$, then $X_p(fg)$=0;

7.6. Let $M \simeq \mathbb{R}^{n^2}$ be the manifold of square $n \times n$-matrices, and $X_a(x) = a \cdot x$, $Y_a(x) = a \cdot x - x \cdot a$ be two vector fields on it, where $a, x \in M$.

(a) Compute the flow of the field V_a. Find first integrals for this flow.

(b) Compute the commutator $[X_a, X_b]$ for two vector fields X_a and X_b, defined by two matrices $a, b \in M$.

(*c*) The same questions relative to the vector field $V_a(x)$.

7.7. A vector field X on a manifold M is called complete, if any of its trajectory can be infinitely continued forward and backward.

(*a*) Prove that on a compact manifold any vector field is complete.

(*b*) Show that on any manifold M and any vector field X on it, there exists a positive function $f \in C^1(M)$ such that the vector field fX is complete.

7.8. Let (M, g) be a Riemannian manifold and the corresponding volume form $\omega \in \Lambda^n(M)$. Prove that for any $2n$ vector fields $X_1, \cdots, X_n$, $Y_1, \cdots, Y_n$, we have

$$\omega(X_1, \cdots, X_n) \cdot \omega(Y_1, \cdots, Y_n) = \det[g(X_i, Y_j)].$$

7.9. (Hessian of Rosenbrok's banana function) Let us consider the Riemannian manifold $(\mathbb{R}^2, g)$, with the metric

$$g(x^1, x^2) = \begin{pmatrix} 1 + 4(x^1)^2 & -2x^1 \\ -2x^1 & 1 \end{pmatrix}.$$

Show that the Hessian of the Rosenbrok's banana function

$$f : \mathbb{R}^2 \to R,\ f(x^1, x^2) = 100(x^2 - (x^1)^2) + (1 - x^1)^2$$

is a Riemannian metric.

7.10. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^3-function such that its Hessian $Hess(f)$ is positive definite. From the Euclidean space $(\mathbb{R}^n, \delta_{ij})$ we pass to the Riemannian manifold $(\mathbb{R}^2, Hess(f))$. Show that the equations of geodesics in this new manifold are

$$2\frac{\partial^2 f}{\partial x^i \partial x^k}(x(t))\, \ddot{x}^i(t) + \frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}(x(t))\, \dot{x}^i(t)\, \dot{x}^j(t) = 0.$$

7.11. (*a*) Find the Christoffel coefficients on the Riemannian manifold $(\mathbb{R}^2_+, g)$, where $g = \mathrm{diag}\left(\frac{1}{x^2}, \frac{1}{y^2}\right)$.

(*b*) Compute the Hessian of the function

$$f : \mathbb{R}^2_+ \to \mathbb{R},\ f(x,y) = \frac{1}{x} + \sqrt{x} + \frac{1}{y} + \sqrt{y},$$

with respect to g.

(*c*) Find the geodesics of the Riemannian manifold $(\mathbb{R}^2_+, Hess_g(f))$.

7.12. Find the geodesics of the Riemannian manifold

$$(\mathbb{R}^2_+, g(x,y)),$$

when g is a posinomial metric

$$g(x,y) = \begin{pmatrix} a_{11}x^{\alpha_{11}}y^{\beta_{11}} & a_{12}x^{\alpha_{12}}y^{\beta_{12}} \\ a_{12}x^{\alpha_{12}}y^{\beta_{12}} & a_{22}x^{\alpha_{22}}y^{\beta_{22}} \end{pmatrix}.$$

7.13. Let (M, g) be a Riemannian manifold.

(*a*) Show that $R^r_{ijk} = -R^r_{jik}$.

(*b*) Assume $\dim M = 1$. Show that M is flat, i.e., $R = 0$.

7.14. Let $R_{ij} = Ric(\partial_i, \partial_j)$ be the components of the Ricci tensor in local coordinates. Show that $R_{ij} = R^k_{ikj}$, with summation over k.

7.13 Historical Remarks

Differential Geometry started with the study of curves since around 1700s. Among the first mathematicians who had investigated the theory of curves were Euler, Monge, Venant, Serret, and Darboux. In 1827 Gauss published his celebrated work *Disquisitiones generales circa superficies curvas*, where he introduced the first and the second fundamental forms on surfaces in $\mathbb{R}^3$ and had shown that they characterize the surface up to a rigid motion. Gauss proved that the curvature is an intrinsic invariant of the surface, result that is called *Theorema Egregium.* The name emphasizes its profound philosophical implications, since the curvature is usually perceived as an extrinsic object.

Gauss' ideas of intrinsic geometry of a surface influenced his pupil, Riemann, who at only 28, presents his Ph.D. dissertation *Ueber die*

Hypothesen welche der Geometrie zu Grunde liegen at Göttingen in 1954. Riemann associated a metric with each hypersurface, fact that led to the concept of Riemannian manifold later. These results flourished into an elegant theory, which generalized Gauss' results on manifolds.

However, this theory requires laborious computations, fact that needed the construction of the tensorial formalism. Ricci developed the tensorial calculus on manifolds and Levi–Civita introduced the linear connection with the same name in 1900s.

Differential geometry has important consequences and applications. First, it closed the celebrated problem of the 5th postulate of Euclid. This was accomplished by finding examples of non-Euclidean spaces among Riemannian manifolds.

Another application is the use of differential geometry to General Theory of Relativity. Einstein's theory published in 1917 used tensorial calculus to write the equations of space-time invariantly. This way, the concept of inertial system from Newtonian mechanics is generalized and the new theory was able to explain the Mercury's perihelion advance and the light deflection about sun.

If Lorentz geometry, which is the geometry of a manifold endowed with a space-time type metric, is a good environment for relativity theory, then Riemannian geometry was proved to be suited for the Classical Mechanics, see Abraham and Marsden [1] or Calin and Chang [22]. The conservation laws of Newtonian Physics can be written in an elegant way in terms of the Riemannian Geometry language.

Another direction where Differential Geometry has recently been applied is the geometric theory of differential equations. Each differential operator is associated with a principal symbol, which can be considered as a Hamiltonian. This defines a metric on an associated manifold. The study of heat kernels and fundamental solutions can be geometrically based on the study of geodesics on the associated Riemannian manifold. The interested reader can consult this topic in Calin et al. [23] and [24]. For convex functions and optimization methods on Riemannian manifolds the reader is referred to Udriste [81].

Another related branch of Riemannian Geometry has been developed over the last several decades. It is known under the names of SubRiemannian Geometry, Non-holonomic geometry, or Carnot-Carathéodory geometry. It is related with Quantum Mechanics behavior of particles and Thermodynamics, see Calin and Chang [23].

The goal of the present book is to deal with one of the branches of Differential Geometry which applies to Information Theory, Probability and Statistics. This is known under the name of Information Geometry. Its main object of study is the statistical manifold, which is a Riemannian manifold that holds a dualistic structure and studies the relationship between dual geometric objects. All the next chapters deal with notions which culminate with the study of statistical manifolds.