§1-§3. When a Euclidean space is stripped of its vector space structure and only its differentiable structure retained, there are many ways of piecing together domains of it in a smooth manner, thereby obtaining a so-called differentiable manifold. Local concepts like a differentiable function and a tangent vector can still be given a meaning whereby the manifold can be viewed "tangentially," that is, through its family of tangent spaces as a curve in the plane is, roughly speaking, determined by its family of tangents. This viewpoint leads to the study of tensor fields, which are important tools in local and global differential geometry. They form an algebra \( \mathcal{D}(M) \), the mixed tensor algebra over the manifold \( M \). The alternate covariant tensor fields (the differential forms) form a submodule \( \mathfrak{A}(M) \) of \( \mathcal{D}(M) \) which inherits a multiplication from \( \mathcal{D}(M) \), the exterior multiplication. The resulting algebra is called the Grassmann algebra of \( M \). Through the work of É. Cartan the Grassmann algebra with the exterior differentiation \( d \) has become an indispensable tool for dealing with submanifolds, these being analytically described by the zeros of differential forms. Moreover, the pair \( (\mathcal{D}(M), d) \) determines the cohomology of \( M \) via de Rham's theorem, which however will not be dealt with here.

§4-§8. The concept of an affine connection was first defined by Levi-Civita for Riemannian manifolds, generalizing significantly the notion of parallelism for Euclidean spaces. On a manifold with a countable basis an affine connection always exists (see the exercises following this chapter). Given an affine connection on a manifold \( M \) there is to each curve \( \gamma(t) \) in \( M \) associated an isomorphism between any two tangent spaces \( M_{\gamma(t)} \) and \( M_{\gamma(t')} \). Thus, an affine connection makes it possible to relate tangent spaces at distant points of the manifold. If the tangent vectors of the curve \( \gamma(t) \) all correspond under these isomorphisms we have the analog of a straight line, the so-called geodesic. The theory of affine connections mainly amounts to a study of the mappings \( \text{Exp}_p : M_p \to M \) under which straight lines (or segments of them) through the origin in the tangent space \( M_p \), correspond to geodesics through \( p \) in \( M \). Each mapping \( \text{Exp}_p \) is a diffeomorphism of a neighborhood of 0 in \( M_p \) into \( M \), giving the so-called normal coordinates at \( p \).
§1. Manifolds

Let $\mathbb{R}^m$ and $\mathbb{R}^n$ denote two Euclidean spaces of $m$ and $n$ dimensions, respectively. Let $O$ and $O'$ be open subsets, $O \subset \mathbb{R}^m$, $O' \subset \mathbb{R}^n$ and suppose $\varphi$ is a mapping of $O$ into $O'$. The mapping $\varphi$ is called differentiable if the coordinates $y_j(\varphi(p))$ of $\varphi(p)$ are differentiable (that is, indefinitely differentiable) functions of the coordinates $x_i(p)$, $p \in O$. The mapping $\varphi$ is called analytic if for each point $p \in O$ there exists a neighborhood $U$ of $p$ and $n$ power series $P_j(x) = \sum a_{jk}x^j$ ($1 \leq j \leq n$) for $q \in U$. A differentiable mapping $\varphi : O \to O'$ is called a diffeomorphism of $O$ onto $O'$ if $\varphi(O) = O'$, $\varphi$ is one-to-one, and the inverse mapping $\varphi^{-1}$ is differentiable. In the case when $n = 1$ it is customary to replace the term "mapping" by the term "function."

An analytic function on $\mathbb{R}^m$ which vanishes on an open set is identically 0. For differentiable functions the situation is completely different. In fact, if $A$ and $B$ are disjoint subsets of $\mathbb{R}^m$, $A$ compact and $B$ closed, then there exists a differentiable function $\varphi$ which is identically 1 on $A$ and identically 0 on $B$. The standard procedure for constructing such a function $\varphi$ is as follows:

Let $0 < a < b$ and consider the function $f$ on $\mathbb{R}$ defined by

$$f(x) = \begin{cases} \exp \left( \frac{1}{x-a} - \frac{1}{x-b} \right) & \text{if } a < x < b, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f$ is differentiable and the same holds for the function

$$F(x) = \int_a^x f(t) \, dt / \int_a^b f(t) \, dt,$$

which has value 1 for $x \leq a$ and 0 for $x \geq b$. The function $\psi$ on $\mathbb{R}^m$ given by

$$\psi(x_1, ..., x_m) = F(x_1^2 + ... + x_m^2)$$

is differentiable and has values 1 for $x_1^2 + ... + x_m^2 \leq a$ and 0 for $x_1^2 + ... + x_m^2 \geq b$. Let $S$ and $S'$ be two concentric spheres in $\mathbb{R}^m$, $S'$ lying inside $S$. Starting from $\psi$ we can by means of a linear transformation of $\mathbb{R}^m$ construct a differentiable function on $\mathbb{R}^m$ with value 1 in the interior of $S'$ and value 0 outside $S$. Turning now to the sets $A$ and $B$ we can, owing to the compactness of $A$, find finitely many spheres $S_i$ ($1 \leq i \leq n$), such that the corresponding open balls $B_i$ ($1 \leq i \leq n$), form a covering of $A$ (that is, $A \subset \bigcup_{i=1}^n B_i$) and such that the closed balls $B_i$ ($1 \leq i \leq n$) do not intersect $B$. Each sphere $S_i$ can be shrunk to a concentric sphere $S'_{i}$ such that the corresponding open balls $B'_{i}$ still form a covering of $A$. Let $\psi_i$ be a differentiable function on $\mathbb{R}^m$ which is identically 1 on $B'_{i}$ and identically 0 in the complement of $B'_{i}$. Then the function

$$\varphi = 1 - (1 - \psi_1)(1 - \psi_2) ... (1 - \psi_n)$$

is a differentiable function on $\mathbb{R}^m$ which is identically 1 on $A$ and identically 0 on $B$.

Let $M$ be a topological space. We assume that $M$ satisfies the Hausdorff separation axiom which states that any two different points in $M$ can be separated by disjoint open sets. An open chart on $M$ is a pair $(U, \varphi)$ where $U$ is an open subset of $M$ and $\varphi$ is a homeomorphism of $U$ onto an open subset of $\mathbb{R}^m$. 

6
Definition. Let $M$ be a Hausdorff space. A \textit{differentiable structure} on $M$ of dimension $m$ is a collection of open charts $(U_a, \varphi_a)_{a \in A}$ on $M$ where $\varphi_a(U_a)$ is an open subset of $\mathbb{R}^m$ such that the following conditions are satisfied:

1. \((M_1)\) $M = \bigcup_{a \in A} U_a$.

2. \((M_2)\) For each pair $a, b \in A$ the mapping $\varphi_b \circ \varphi_a^{-1}$ is a differentiable mapping of $\varphi_a(U_a \cap U_b)$ onto $\varphi_b(U_a \cap U_b)$.

3. \((M_3)\) The collection $(U_a, \varphi_a)_{a \in A}$ is a maximal family of open charts for which $(M_1)$ and $(M_2)$ hold.

A \textit{differentiable manifold} (or \textit{$C^\infty$ manifold} or simply \textit{manifold}) of dimension $m$ is a Hausdorff space with a differentiable structure of dimension $m$. If $M$ is a manifold, a \textit{local chart} on $M$ (or a \textit{local coordinate system} on $M$) is by definition a pair $(U_a, \varphi_a)$ where $a \in A$. If $p \in U_a$ and $\varphi_a(p) = (x_1(p), \ldots, x_m(p))$, the set $U_a$ is called a \textit{coordinate neighborhood} of $p$ and the numbers $x_i(p)$ are called \textit{local coordinates} of $p$. The mapping $\varphi_a : q \mapsto (x_1(q), \ldots, x_m(q))$, $q \in U_a$, is often denoted $(x_1, \ldots, x_m)$.

Remark 1. Condition $(M_3)$ will often be cumbersome to check in specific instances. It is therefore important to note that the condition $(M_3)$ is not essential in the definition of a manifold. In fact, if only $(M_1)$ and $(M_2)$ are satisfied, the family $(U_a, \varphi_a)_{a \in A}$ can be extended in a unique way to a larger family $\mathfrak{M}$ of open charts such that $(M_1)$, $(M_2)$, and $(M_3)$ are all fulfilled. This is easily seen by defining $\mathfrak{M}$ as the set of all open charts $(V, \varphi)$ on $M$ satisfying: (1) $\varphi(V)$ is an open set in $\mathbb{R}^m$; (2) for each $a \in A$, $\varphi_a \circ \varphi^{-1}$ is a diffeomorphism of $\varphi(V \cap U_a)$ onto $\varphi_a(V \cap U_a)$.

Remark 2. If we let $\mathbb{R}^m$ mean a single point for $m = 0$, the preceding definition applies. The manifolds of dimension 0 are then the discrete topological spaces.

Remark 3. A manifold is connected if and only if it is pathwise connected. The proof is left to the reader.

An \textit{analytic structure} of dimension $m$ is defined in a similar fashion. In $(M_3)$ we just replace "differentiable" by "analytic." In this case $M$ is called an \textit{analytic manifold}.

In order to define a \textit{complex manifold} of dimension $m$ we replace $\mathbb{R}^m$ in the definition of differentiable manifold by the $m$-dimensional complex space $\mathbb{C}^m$. The condition $(M_3)$ is replaced by the condition that the $m$ coordinates of $\varphi_b \circ \varphi_a^{-1}(p)$ should be holomorphic functions of the coordinates of $p$. Here a function $f(x_1, \ldots, x_m)$ of $m$ complex variables is called \textit{holomorphic} if at each point $(x_1^0, \ldots, x_m^0)$ there exists a power series

$$\sum a_{n_1, \ldots, n_m} (x_1 - x_1^0)^{n_1} \cdots (x_m - x_m^0)^{n_m},$$

which converges absolutely to $f(x_1, \ldots, x_m)$ in a neighborhood of the point.

The manifolds dealt with in the later chapters of this book (mostly
Lie groups and their coset spaces) are analytic manifolds. From Remark 1 it is clear that we can always regard an analytic manifold as a differentiable manifold. It is often convenient to do so because, as pointed out before for $\mathbb{R}^m$, the class of differentiable functions is much richer than the class of analytic functions.

Let $f$ be a real-valued function on a $C^\infty$ manifold $M$. The function $f$ is called differentiable at a point $p \in M$ if there exists a local chart $(U, \varphi_a)$ with $p \in U_a$ such that the composite function $f \circ \varphi_a^{-1}$ is a differentiable function on $\varphi_a(U_a)$. The function $f$ is called differentiable if it is differentiable at each point $p \in M$. If $M$ is analytic, the function $f$ is said to be analytic at $p \in M$ if there exists a local chart $(U, \varphi_a)$ with $p \in U_a$ such that $f \circ \varphi_a^{-1}$ is an analytic function on the set $\varphi_a(U_a)$.

Let $M$ be a differentiable manifold and let $\mathcal{F}$ denote the set of differential functions on $M$.

We shall often write $C^\infty(M)$ instead of $\mathcal{F}$ and will sometimes denote by $C^\infty(p)$ the set of functions on $M$ which are differentiable at $p$. The set $C^\infty(M)$ is an algebra over $\mathbb{R}$, the operations being

$$(\lambda f)(p) = \lambda f(p),$$

$$(f + g)(p) = f(p) + g(p),$$

$$(fg)(p) = f(p)g(p)$$

for $\lambda \in \mathbb{R}, p \in M, f, g \in C^\infty(M)$.

**Lemma 1.2.** Let $C$ be a compact subset of a manifold $M$ and let $V$ be an open subset of $M$ containing $C$. Then there exists a function $\psi \in C^\infty(M)$ which is identically 1 on $C$, identically 0 outside $V$.

This lemma has already been established in the case $M = \mathbb{R}^m$. We shall now show that the general case presents no additional difficulties.

Let $(U, \varphi_a)$ be a local chart on $M$ and $S$ a compact subset of $U_a$. There exists a differentiable function $f$ on $\varphi_a(U_a)$ such that $f$ is identically 1 on $\varphi_a(S)$ and has compact support contained in $\varphi_a(U_a)$. The function $F$ on $M$ given by

$$F(q) = \begin{cases} f(\varphi_a(q)) & \text{if } q \in U_a, \\ 0 & \text{otherwise} \end{cases}$$

is a differentiable function on $M$ which is identically 1 on $S$ and identically 0 outside $U_a$. Since $C$ is compact and $V$ open, there exist finitely many coordinate neighborhoods $U_1, ..., U_n$ and compact sets $S_1, ..., S_n$ such that

$$C \subset \bigcup_{i=1}^n S_i, \quad S_i \subset U_i$$

$$(U_1 \cap U_i) \subset V.$$

As shown previously, there exists a function $F_i \in C^\infty(M)$ which is identically 1 on $S_i$ and identically 0 outside $U_i$. The function

$$\psi = 1 - (1 - F_1)(1 - F_2) ... (1 - F_n)$$

belongs to $C^\infty(M)$, is identically 1 on $C$ and identically 0 outside $V$.

Let $M$ be a $C^\infty$ manifold and $(U_a, \varphi_a)_{a \in \mathcal{A}}$ a collection satisfying $(M_1)$, $(M_2)$, and $(M_3)$. If $U$ is an open subset of $M$, $U$ can be given a differentiable structure by means of the open charts $(V_a, \psi_a)$, where $V_a = U \cap U_a$ and $\psi_a$ is the restriction of $\varphi_a$ to $V_a$. With this structure, $U$ is called an open submanifold of $M$. In particular, since $M$ is locally connected, each connected component of $M$ is an open submanifold of $M$. 

$\mathcal{8}$
Let $M$ and $N$ be two manifolds of dimension $m$ and $n$, respectively. Let $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ and $(V_\beta, \psi_\beta)_{\beta \in B}$ be collections of open charts on $M$ and $N$, respectively, such that the conditions $(M_1)$, $(M_2)$, and $(M_3)$ are satisfied. For $\alpha \in A$, $\beta \in B$, let $\varphi_\alpha \times \psi_\beta$ denote the mapping $(p, q) \mapsto (\varphi_\alpha(p), \psi_\beta(q))$ of the product set $U_\alpha \times V_\beta$ into $\mathbb{R}^{m+n}$. Then the collection $(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta)_{\alpha \in A, \beta \in B}$ of open charts on the product space $M \times N$ satisfies $(M_1)$ and $(M_2)$ so by Remark 1, $M \times N$ can be turned into a manifold the product of $M$ and $N$.

An immediate consequence of Lemma 1.2 is the following fact which will often be used: Let $V$ be an open submanifold of $M$, $f$ a function in $C^\infty(V)$, and $p$ a point in $V$. Then there exists a function $\tilde{f} \in C^\infty(M)$ and an open neighborhood $N$, $p \in N \subset V$ such that $f$ and $\tilde{f}$ agree on $N$.

1. Vector Fields and 1-Forms

Let $A$ be an algebra over a field $K$. A derivation of $A$ is a mapping $D: A \to A$ such that

(i) $D(\alpha f + \beta g) = \alpha Df + \beta Dg$ for $\alpha, \beta \in K$, $f, g \in A$;

(ii) $D(fg) = f(Dg) + (Df)g$ for $f, g \in A$. 


Definition. A vector field $X$ on a $C^\infty$ manifold is a derivation of the algebra $C^\infty(M)$.

Let $\mathfrak{D}$ (or $\mathfrak{X}(M)$) denote the set of all vector fields on $M$. If $f \in C^\infty(M)$ and $X, Y \in \mathfrak{D}(M)$, then $fX$ and $X + Y$ denote the vector fields

$fX : g \mapsto f(Xg), \quad g \in C^\infty(M),$

$X + Y : g \mapsto Xg + Yg, \quad g \in C^\infty(M).$

This turns $\mathfrak{D}(M)$ into a module over the ring $\mathcal{R} = C^\infty(M)$. If $X, Y \in \mathfrak{D}(M)$, then $XY - YX$ is also a derivation of $C^\infty(M)$ and is denoted by the bracket $[X, Y]$. As is customary we shall often write $\theta(X) Y = [X, Y]$. The operator $\theta(X)$ is called the Lie derivative with respect to $X$.

The bracket satisfies the Jacobi identity $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ or, otherwise written $\theta(X) ([Y, Z]) = [\theta(X) Y, Z] + [Y, \theta(X) Z]$.

It is immediate from (ii) that if $f$ is constant and $X \in \mathfrak{D}$, then $Xf = 0$.

Suppose now that a function $g \in C^\infty(M)$ vanishes on an open subset $V \subset M$. Let $p$ be an arbitrary point in $V$. According to Lemma 1.2 there exists a function $f \in C^\infty(M)$ such that $f(p) = 0$, and $f = 1$ outside $V$. Then $g = fg$ so

$Xg = f(Xg) + g(Xf),$

which shows that $Xg$ vanishes at $p$. Since $p$ was arbitrary, $Xg = 0$ on $V$. We can now define $Xf$ on $V$ for every function $f \in C^\infty(V)$. If $p \in V$, select $f \in C^\infty(M)$ such that $f$ and $f$ coincide in a neighborhood of $p$ and put $(Xf)(p) = (Xf)(p)$. The consideration above shows that this is a valid definition, that is, independent of the choice of $f$. This shows that a vector field on a manifold induces a vector field on any open submanifold.

On the other hand, let $Z$ be a vector field on an open submanifold $V \subset M$ and $p$ a point in $V$. Then there exists a vector field $Z$ on $M$ and an open neighborhood $N$, $p \in N \subset V$ such that $Z$ and $Z$ induce the same vector field on $N$. In fact, let $C$ be any compact neighborhood of $p$ contained in $V$ and let $N$ be the interior of $C$. Choose $\psi \in C^\infty(M)$ of compact support contained in $V$ such that $\psi = 1$ on $C$. For any $g \in C^\infty(M)$, let $g_V$ denote its restriction to $V$ and define $Zg$ by

$Zg(q) = \begin{cases} \psi(q) (Zg_V)(q) & \text{for } q \in V, \\ 0 & \text{if } q \notin V. \end{cases}$

Then $g \mapsto Zg$ is the desired vector field on $M$.

Now, let $(U, \varphi)$ be a local chart on $M$, $X$ a vector field on $U$, and let $p$ be an arbitrary point in $U$. We put $\varphi(q) = (x_1(q), ..., x_m(q))$ ($q \in U$),
and \( f^* = f \circ \varphi^{-1} \) for \( f \in C^\infty(M) \). Let \( V \) be an open subset of \( U \) such that \( \varphi(V) \) is an open ball in \( \mathbb{R}^m \) with center \( \varphi(p) = (a_1, \ldots, a_m) \). If \( (x_1, \ldots, x_m) \in \varphi(V) \), we have
\[
f^*(x_1, \ldots, x_m) = f^*(a_1, \ldots, a_m) + \int_0^1 \frac{\partial}{\partial t} f^*(a_1 + t(x_1 - a_1), \ldots, a_m + t(x_m - a_m)) \, dt
\]
(Here \( f^* \) denotes the partial derivative of \( f^* \) with respect to the \( j \)th argument.) Transferring this relation back to \( M \) we obtain
\[
f(q) = f(p) + \sum_{i=1}^m (x(q) - x_i(p)) g_i(q) \quad (q \in V),
\]
where \( g_i \in C^\infty(V) \) (\( 1 \leq i \leq m \)), and
\[
g_i(p) = \left( \frac{\partial f^*}{\partial x_i} \right)_{\varphi(p)}.
\]
It follows that
\[
(Xf)(p) = \sum_{i=1}^m \left( \frac{\partial f^*}{\partial x_i} \right)_{\varphi(p)} (Xx_i)(p) \quad \text{for } p \in U.
\]
The mapping \( f \to \left( \frac{\partial f^*}{\partial x_i} \right)_{\varphi(p)} \circ \varphi (f \in C^\infty(U)) \) is a vector field on \( U \) and is denoted \( \partial_i \). We write \( \partial_i \) instead of \( \partial/\partial x_i(f) \). Now, by (2)
\[
X = \sum_{i=1}^m (Xx_i) \frac{\partial}{\partial x_i} \quad \text{on } U.
\]
Thus, \( \partial_i \) (\( 1 \leq i \leq m \)) is a basis of the module \( \mathcal{D}^1(U) \).

For \( p \in M \) and \( X \in \mathcal{D}^1 \), let \( X_p \) denote the linear mapping \( X_p: f \to (Xf)(p) \) of \( C^\infty(p) \) into \( \mathbb{R} \). The set \( \{X_p: X \in \mathcal{D}^1(M)\} \) is called the tangent space to \( M \) at \( p \); it will be denoted by \( \mathcal{D}^1(p) \) or \( M_p \) and its elements are called the tangent vectors to \( M \) at \( p \). Relation (2) shows that \( M_p \) is a vector space over \( \mathbb{R} \) spanned by the \( m \) linearly independent vectors
\[
e_i: f \to \left( \frac{\partial f^*}{\partial x_i} \right)_{\varphi(p)}, \quad f \in C^\infty(M).
\]
This tangent vector \( e_i \) will often be denoted by \( (\partial/\partial x_i)_p \). A linear mapping \( L: C^\infty(p) \to \mathbb{R} \) is a tangent vector to \( M \) at \( p \) if and only if the condition
\( L(fg) = f(p) L(g) + g(p) L(f) \) is satisfied for all \( f, g \in C^\infty(p) \). In fact, the necessity of the condition is obvious and the sufficiency is a simple consequence of (1). Thus, a vector field \( X \) on \( M \) can be identified with a collection \( X_p(p \in M) \) of tangent vectors to \( M \) with the property that for each \( f \in C^\infty(M) \) the function \( p \mapsto X_p f \) is differentiable.

Suppose the manifold \( M \) is analytic. The vector field \( X \) on \( M \) is then called \textit{analytic} at \( p \) if \( X_f \) is analytic at \( p \) whenever \( f \) is analytic at \( p \).

Remark. Let \( V \) be a finite-dimensional vector space over \( R \). If \( X_1, \ldots, X_n \) is any basis of \( V \), the mapping \( \Sigma_{i=1}^n x_i X_i \rightarrow (x_1, \ldots, x_n) \) is an open chart valid on the entire \( V \). The resulting differentiable structure is independent of the choice of basis. If \( X \in V \), the tangent space \( V_x \) is identified with \( V \) itself by the formula

\[
(Yf)(X) = \left. \frac{d}{dt} f(X + tY) \right|_{t=0}, \quad f \in C^\infty(V),
\]

which to each \( Y \in V \) assigns a tangent vector to \( V \) at \( X \).

§ 3. Mappings

1. The Interpretation of the Jacobian

Let \( M \) and \( N \) be \( C^\infty \) manifolds and \( \Phi \) a mapping of \( M \) into \( N \). Let \( p \in M \). The mapping \( \Phi \) is called \textit{differentiable at} \( p \) if \( g \circ \Phi \in C^\infty(p) \) for each \( g \in C^\infty(\Phi(p)) \). The mapping \( \Phi \) is called \textit{differentiable} if it is differentiable at each \( p \in M \). Similarly \textit{analytic} mappings are defined. Let \( \psi: q \rightarrow (x_1(q), \ldots, x_n(q)) \) be a system of coordinates on a neighborhood \( U \) of \( p \in M \) and \( \psi: r \rightarrow (y_1(r), \ldots, y_n(r)) \) a system of coordinates on a neighborhood \( U' \) of \( \Phi(p) \) in \( N \). Assume \( \Phi(U) \subset U' \). The mapping \( \psi \circ \Phi \circ \psi^{-1} \) of \( \psi(U) \) into \( \psi'(U') \) is given by a system of \( n \) functions

\[
y_j = \varphi_j(x_1, \ldots, x_n) \quad (1 \leq j \leq n),
\]

which we call the \textit{expression} of \( \Phi \) in coordinates. The mapping \( \Phi \) is differentiable at \( p \) if and only if the functions \( \varphi_j \) have partial derivatives of all orders in some fixed neighborhood of \( (x_1(p), \ldots, x_n(p)) \).

The mapping \( \Phi \) is called a \textit{diffeomorphism of} \( M \) onto \( N \) if \( \Phi \) is a one-to-one differentiable mapping of \( M \) onto \( N \) and \( \Phi^{-1} \) is differentiable. If in addition \( M, N, \Phi, \) and \( \Phi^{-1} \) are analytic, \( \Phi \) is called an \textit{analytic diffeomorphism}.

If \( \Phi \) is differentiable at \( p \in M \) and \( A \in M_p \), then the linear mapping \( B: C^\infty(\Phi(p)) \rightarrow R \) given by \( B(g) = A(g \circ \Phi) \) for \( g \in C^\infty(\Phi(p)) \) is a tangent vector to \( N \) at \( \Phi(p) \). The mapping \( A \rightarrow B \) of \( M_p \) into \( N_{\Phi(p)} \) is denoted \( d\Phi_p \) (or just \( \Phi_p \)) and is called the \textit{differential} of \( \Phi \) at \( p \). We have seen that the vectors

\[
\xi_i: f \rightarrow \left( \frac{\partial f^*}{\partial x_i} \right)_{\psi(p)} \quad (1 \leq i \leq m), \quad f^* = f \circ \psi^{-1},
\]

\[
\xi_j: g \rightarrow \left( \frac{\partial g^*}{\partial y_j} \right)_{\psi'(\Phi(p))} \quad (1 \leq j \leq n), \quad g^* = g \circ (\psi')^{-1}
\]

form a basis of \( M_p \) and \( N_{\Phi(p)} \), respectively.
\[ d\Phi_p(e_i) g = e_i(g \circ \Phi) = \left( \frac{\partial (g \circ \Phi)^*}{\partial x_i} \right)_{\psi(p)} . \]

But
\[ (g \circ \Phi)^*(y_1, ..., y_n) = g^*(y_1, ..., y_n), \]

where \( y_j = \varphi_j(x_1, ..., x_m) \) (1 \( \leq j \leq n \)). Hence
\[ d\Phi_p(e_i) = \sum_{j=1}^{n} \left( \frac{\partial \varphi_j}{\partial x_i} \right)_{\psi(p)} e_j . \]  

This shows that if we use the bases \( e_i (1 \leq i \leq m) \), \( e_j (1 \leq j \leq n) \) to express the linear transformation \( d\Phi_p \) in matrix form, then the matrix we obtain is just the Jacobian of the system (1). From a standard theorem on the Jacobian (the inverse function theorem), we can conclude:

**Proposition 3.1.** If \( d\Phi_p \) is an isomorphism of \( M_p \) onto \( N_{\phi(p)} \), then there exist open submanifolds \( U \subset M \) and \( V \subset N \) such that \( p \in U \) and \( \Phi \) is a diffeomorphism of \( U \) onto \( V \).

**Definition.**

Let \( M \) and \( N \) be differentiable (or analytic) manifolds.

(a) A mapping \( \Phi : M \rightarrow N \) is called *regular* at \( p \in M \) if \( \Phi \) is differentiable (analytic) at \( p \in M \) and \( d\Phi_p \) is a one-to-one mapping of \( M_p \) into \( N_{\phi(p)} \).

(b) \( M \) is called a submanifold of \( N \) if (1) \( M \subset N \) (set theoretically); (2) the identity mapping \( I \) of \( M \) into \( N \) is regular at each point of \( M \).

For example, the sphere \( x_1^2 + x_2^2 + x_3^2 = 1 \) is a submanifold of \( \mathbb{R}^3 \) and a topological subspace as well. However, a submanifold \( M \) of a manifold \( N \) is not necessarily a topological subspace of \( N \). For example, let \( N \) be a torus and let \( M \) be a curve on \( N \) without double points, dense in \( N \) (Chapter II, §2). Proposition 3.1 shows that a submanifold \( M \) of a manifold \( N \) is an open submanifold of \( N \) if and only if \( \dim M = \dim N \).

**Proposition 3.2.** Let \( M \) be a submanifold of a manifold \( N \) and let \( p \in M \). Then there exists a coordinate system \( \{x_1, ..., x_n\} \) valid on an open neighborhood \( V \) of \( p \) in \( N \) such that \( x_1(p) = ... = x_n(p) = 0 \) and such that the set

\[ U = \{ q \in V : x_j(q) = 0 \text{ for } m + 1 \leq j \leq n \} \]

together with the restrictions of \( \{x_1, ..., x_n\} \) to \( U \) form a local chart on \( M \) containing \( p \).

**Proof.** Let \( \{y_1, ..., y_m\} \) and \( \{z_1, ..., z_n\} \) be coordinate systems valid on open neighborhoods of \( p \) in \( M \) and \( N \), respectively, such that \( y_i(p) = z_i(p) = 0 \), (1 \( \leq i \leq m \), 1 \( \leq j \leq n \)). The expression of the
identity mapping $I : M \to N$ is (near $p$) given by a system of functions $z_j = \psi_j(y_1, \ldots, y_m), 1 \leq j \leq n$. The Jacobian matrix $(\partial \psi_j/\partial y_k)$ of this system has rank $m$ at $p$ since $I$ is regular at $p$. Without loss of generality we may assume that the square matrix $(\partial \psi_j/\partial y_k)_{1 \leq i, j \leq m}$ has determinant $\neq 0$ at $p$. In a neighborhood of $(0, \ldots, 0)$ we have therefore $y_i = \psi_i(x_1, \ldots, x_m), 1 \leq i \leq m$, where each $\psi_i$ is a differentiable function. If we now put $x_i = z_i, \quad 1 \leq i \leq m,$

$\quad x_j = z_j - \psi_j(\psi_1(x_1, \ldots, x_m), \ldots, \psi_m(x_1, \ldots, x_m)), \quad m + 1 \leq j \leq n,$

it is clear that

\[
\det \left( \frac{\partial x_i}{\partial y_j} \right)_{1 \leq i, j \leq m} \neq 0, \quad \det \left( \frac{\partial x_i}{\partial z_k} \right)_{1 \leq i, k \leq n} \neq 0.
\]

Therefore $\{x_1, \ldots, x_n\}$ gives the desired coordinate system.

2. Transformation of Vector Fields

Let $M$ and $N$ be $C^\infty$ manifolds and $\Phi$ a differentiable mapping of $M$ into $N$. Let $X$ and $Y$ be vector fields on $M$ and $N$, respectively; $X$ and $Y$ are called $\Phi$-related if

\[
d\Phi_p(X_p) = Y_{\Phi(p)} \quad \text{for all } p \in M.
\]

It is easy to see that (3) is equivalent to

\[
(Yf) \circ \Phi = X(f \circ \Phi) \quad \text{for all } f \in C^\infty(N).
\]

It is convenient to write $d\Phi : X = Y$ or $X^* = Y$ instead of (3).

**Proposition 3.3.**

(i) Suppose $X_i$ and $Y_i$ are $\Phi$-related, $(i = 1, 2)$. Then

$\{X_1, X_2\}$ and $\{Y_1, Y_2\}$ are $\Phi$-related.

(ii) Suppose $\Phi$ is a diffeomorphism of $M$ onto itself and put $f^* = f \circ \Phi^{-1}$ for $f \in C^\infty(M)$. Then if $X \in \mathfrak{X}(M),$

$$(fX)^* = f^*X^*, \quad (Xf)^* = X^*f^*.$$  

**Proof.** From (4) we have $(Y_1(Y_2f)) \circ \Phi = X_1(Y_2f \circ \Phi) = X_1(X_2(f \circ \Phi)), \text{so (i) follows. The last relation in (ii) is also an immediate consequence of (4). As to the first one, we have for } g \in C^\infty(M)$

$$(fX)^*g \circ \Phi = (fX)(g \circ \Phi) = f((X^*g) \circ \Phi),$$

so

$$(fX)^*g = f^*(X^*g).$$
Remark. Since \( Xf^\circ = (Xf^{\circ -1})^\circ \) it is natural to make the following definition. Let \( \Phi \) be a diffeomorphism of \( M \) onto \( M \) and \( A \) a mapping of \( C^\infty(M) \) into itself. The mapping \( A^\circ \) is defined by \( A^\circ f = (Af^{\circ -1})^\circ \) for \( f \in C^\infty(M) \). We also write \([Af](p)\) for the value of the function \( Af \) at \( p \in M \). If \( \Phi \) and \( \Psi \) are two diffeomorphisms of \( M \), then \( f^{\circ \Psi} = (f^{\circ \Phi})^\circ \) and \( A^{\circ \Psi} = (A^{\circ \Phi})^\circ \).

Let \( M \) be a differentiable manifold, \( S \) a submanifold. Let \( m = \dim M \) and \( s = \dim S \). A curve in \( S \) is of course a curve in \( M \), but a curve in \( M \) contained in \( S \) is not necessarily a curve in \( S \), because it may not even be continuous. However, we have:

**Lemma 3.4.** Let \( \varphi \) be a differentiable mapping of a manifold \( V \) into the manifold \( M \) such that \( \varphi(V) \) is contained in the submanifold \( S \). If the mapping \( \varphi: V \to S \) is continuous it is also differentiable.

Let \( p \in V \). In view of Prop. 3.2, there exists a coordinate system \( \{x_1, \ldots, x_m\} \) valid on an open neighborhood \( N \) of \( \varphi(p) \) in \( M \) such that the set

\[
N_S = \{ r \in N : x_j(r) = 0 \text{ for } s < j \leq m \}
\]

together with the restrictions of \( \{x_1, \ldots, x_m\} \) to \( N_S \) form a local chart on \( S \) containing \( \varphi(p) \). By the continuity of \( \varphi \) there exists a local chart \( (W, \psi) \) around \( p \) such that \( \varphi(W) \subset N_S \). The coordinates \( x_j(\varphi(q)) \) \( (1 \leq j \leq m) \) depend differentiably on the coordinates of \( q \in W \). In particular, this holds for the coordinates \( x_j(\varphi(q)) \) \( (1 \leq j \leq s) \) so the mapping \( \varphi: V \to S \) is differentiable.

As an immediate consequence of this lemma we have the following statement: Suppose that \( V \) and \( S \) are submanifolds of \( M \) and \( V \subset S \). If \( S \) has the relative topology of \( M \), then \( V \) is a submanifold of \( S \).

### §4. Affine Connections

**Definition.** An affine connection on a manifold \( M \) is a rule \( \nabla \) which assigns to each \( X \in \mathfrak{X}(M) \) a linear mapping \( \nabla_X \) of the vector space \( \mathfrak{X}(M) \) into itself satisfying the following two conditions:

1. \( \nabla_{fX+gY} = f \nabla_X + g \nabla_Y \)
2. \( \nabla_X(fY) = f \nabla_X(Y) + (XYf) Y \)

for \( f, g \in C^\infty(M) \), \( X, Y \in \mathfrak{X}(M) \). The operator \( \nabla_X \) is called covariant differentiation with respect to \( X \). For a motivation see Exercises.

**Lemma 4.1.** Suppose \( M \) has the affine connection \( X \mapsto \nabla_X \) and let \( U \) be an open submanifold of \( M \). Let \( X, Y \in \mathfrak{X}(M) \). If \( X \) or \( Y \) vanishes identically on \( U \), then so does \( \nabla_X(Y) \).

**Proof.** Suppose \( Y \) vanishes on \( U \). Let \( p \in U \) and \( g \in C^\infty(M) \). To prove that \( \nabla_X(Y)g(p) = 0 \), we select \( f \in C^\infty(M) \) such that \( f(p) = 0 \) and \( f = 1 \) outside \( U \) (Lemma 1.2). Then \( fY = Y \) and

\[
\nabla_X(Y)g = \nabla_X(fY)g = (Xf)(Yg) + f(\nabla_X(Y)g)
\]

which vanishes at \( p \). The statement about \( X \) follows similarly.
An affine connection $\nabla$ on $M$ induces an affine connection $\nabla_U$ on an arbitrary open submanifold $U$ of $M$. In fact, let $X$, $Y$ be two vector fields on $U$. For each $p \in U$ there exist vector fields $X'$, $Y'$ on $M$ which agree with $X$ and $Y$ in an open neighborhood $V$ of $p$. We then put $((\nabla_U)x(Y))_q = (\nabla_{X'}(Y'))_q$ for $q \in V$. By Lemma 4.1, the right-hand side of this equation is independent of the choice of $X'$, $Y'$. It follows immediately that the rule $\nabla_U: X \rightarrow (\nabla_U)x(X \in \mathfrak{X}(U))$ is an affine connection on $U$.

In particular, suppose $U$ is a coordinate neighborhood where a coordinate system $\varphi: q \rightarrow (x_1(q), ..., x_m(q))$ is valid. For simplicity, we write $\nabla_\xi$ instead of $((\nabla_U)x(Y))$ for simplicity of notation we write also $\Gamma_{i\xi}^k$ for the function $\Gamma_{i\xi}^k \circ \varphi^{-1}$. If $\{y_1, ..., y_m\}$ is another coordinate system valid on $U$, we get another set of functions $\Gamma_{a\gamma}^\nu$ by

$$\nabla_a \left( \frac{\partial}{\partial y^\gamma} \right) = \sum_{\nu} \Gamma_{\alpha\nu}^a \frac{\partial}{\partial y^\nu}.$$ 

Using the axioms $\nabla_1$ and $\nabla_2$ we find easily

$$\Gamma_{a\gamma}^\nu = \sum_{i,j,k} \frac{\partial x_i}{\partial y^\alpha} \frac{\partial x_j}{\partial y^\beta} \frac{\partial y^\gamma}{\partial x^\alpha} \Gamma_{i\beta}^k + \sum_{j} \frac{\partial x_i}{\partial y^\alpha} \frac{\partial y^\gamma}{\partial y^\beta} \Gamma_{j\beta}^k.$$ 

On the other hand, suppose there is given a covering of a manifold $M$ by open coordinate neighborhoods $U$ and in each neighborhood a system of functions $\Gamma_{i\xi}^k$ such that (2) holds whenever two of these neighborhoods overlap. Then we can define $\nabla_\xi$ by (1) and thus get an affine connection $\nabla_U$ in each coordinate neighborhood $U$. We finally define an affine connection $\nabla$ on $M$ as follows: Let $X$, $Y \in \mathfrak{X}(M)$ and $p \in M$. If $U$ is a coordinate neighborhood containing $p$, let

$$\nabla_{(X(Y))}_p = ((\nabla_U)x(Y))_p$$

if $X_i$ and $Y_i$ are the vector fields on $U$ induced by $X$ and $Y$, respectively. Then $\nabla$ is an affine connection on $M$ which on each coordinate neighborhood $U$ induces the connection $\nabla_U$.

**Lemma 4.2.** Let $X$, $Y \in \mathfrak{X}(M)$. If $X$ vanishes at a point $p$ in $M$, then so does $\nabla_X(Y)$.

Let $\{x_1, ..., x_m\}$ be a coordinate system valid on an open neighborhood $U$ of $p$. On the set $U$ we have $X = \sum_i f_i(\partial/\partial x_i)$ where $f_i \in C^\infty(U)$ and $f_i(p) = 0$, $1 \leq i \leq m$. Using Lemma 4.1 we find $((\nabla_X(Y))_p = \sum_i f_i(\partial)(\nabla_X(Y))_p = 0$.

Remark. Thus if $v \in M_p$, $\nabla_v(Y)$ is a well-defined vector in $M_p$.

Definition. Suppose $\nabla$ is an affine connection on $M$ and that $\Phi$ is a diffeomorphism of $M$. A new affine connection $\nabla'$ can be defined on $M$ by

$$\nabla'_X(Y) = (\nabla_X(\Phi Y^\Phi))^{\Phi^{-1}}, \quad X, Y \in \mathfrak{X}(M).$$

That $\nabla'$ is indeed an affine connection on $M$ is best seen from Prop. 3.3.
The affine connection $\nabla$ is called \textit{invariant} under $\Phi$ if $\nabla' = \nabla$. In this case $\Phi$ is called an \textit{affine transformation} of $M$. Similarly one can define an affine transformation of one manifold onto another.

\section*{§ 5. Parallellism}

Let $M$ be a $C^\infty$ manifold. A \textit{curve} in $M$ is a regular mapping of an open interval $I \subset \mathbb{R}$ into $M$. The restriction of a curve to a closed subinterval is called a \textit{curve segment}. The curve segment is called finite if the interval is finite.

Let $\gamma : t \to \gamma(t)$ ($t \in I$) be a curve in $M$. Differentiation with respect to the parameter will often be denoted by a dot ($\cdot$). In particular, $\dot{\gamma}(t)$ stands for the tangent vector $dy(d/dt)_t$. Suppose now that to each $t \in I$ is associated a vector $Y(t) \in M_{\gamma(t)}$. Assuming $Y(t)$ to vary differentiably with $t$, we shall now define what it means for the family $Y(t)$ to be parallel with respect to $\gamma$. Let $J$ be a compact subinterval of $I$ such that the finite curve segment $\gamma_J : t \to \gamma(t)$ ($t \in J$) has no double points and such that $\gamma(J)$ is contained in a coordinate neighborhood $U$. Owing to the regularity of $\gamma$ each point of $I$ is contained in such an interval $J$ with nonempty interior. Let $\{x_1, \ldots, x_m\}$ be a \textit{coordinate system} on $U$.

\textbf{Lemma 5.1.} Let $t \to \gamma(t)$ ($t \in I$) be a curve in a manifold $M$. Let $t_0 \in I$ and $g$ a smooth function on a neighborhood of $t_0$ in $I$. Then \exists open interval $I_{t_0}$ around $t_0$ in $I$ and a function $G \in C^\infty(M)$ such that

$$G(\gamma(t)) = g(t) \quad t \in I_{t_0}.$$  

\textbf{Proof:} 

Let $\{x_1, \ldots, x_m\}$ be a coordinate system on a neighborhood of $\gamma(t_0)$ in $M$. There exists an index $i$ such that $t \to x_i(\gamma(t))$ has a nonzero derivative at $t = t_0$. Thus \exists smooth function $\eta_i$ of 1-variable smooth in a neighborhood of $x_i(\gamma(t_0))$ in $\mathbb{R}$ such that $t = \eta_i(x_i(\gamma(t)))$ for all $t$ in a neighborhood of $t_0$. The function

$$q \to g(\eta_i(x_i(q)))$$

is defined and smooth in a neighborhood of $\gamma(t_0)$ in $M$. In a smaller neighborhood it coincides with a function $G \in C^\infty(M)$. But then

$$G(\gamma(t)) = g(\eta_i(x_i(\gamma(t)))) = g(t)$$

for $t$ in an interval around $t_0$. 

17
We put $X(t) = \dot{\gamma}(t)$ ($t \in I$). Using Lemma 5.1 it is easy to see that there exist vector fields $X$, $Y \in \mathfrak{V}(M)$ such that $(Y(t)$ being as before)

$$X_{\gamma(t)} = X(t), \quad Y_{\gamma(t)} = Y(t) \quad (t \in I).$$

Given an affine connection $\nabla$ on $M$, the family $Y(t)$ ($t \in J$) is said to be parallel with respect to $\gamma$, (or parallel along $\gamma$) if

$$\nabla X(Y)_{\gamma(t)} = 0 \quad \text{for all } t \in I.$$  \hfill (1)

To show that this definition is independent of the choice of $X$ and $Y$, we express (1) in the coordinates $\{x_1, ..., x_m\}$. There exist functions $X_i, Y_j$ ($1 \leq i, j \leq m$) on $U$ such that

$$X = \sum_i X^i \frac{\partial}{\partial x_i}, \quad Y = \sum_j Y^j \frac{\partial}{\partial x_j} \quad \text{on } U.$$

For simplicity we put $x_i(t) = x_i(\gamma(t))$, $X^i(t) = X^i(\gamma(t))$, and $Y^j(t) = Y^j(\gamma(t))$ ($t \in I$). Then

$$\nabla X(Y) = \sum_k \left( \sum_i X^i \frac{\partial Y^k}{\partial x_i} + \sum_{i,j} X^i Y^j \Gamma_{ij}^k \right) \frac{\partial}{\partial x_k} \quad \text{on } U.$$

we obtain

$$\frac{dY^k}{dt} + \sum_{i,j} \Gamma_{ij}^k \frac{dx_i}{dt} Y^j = 0 \quad (t \in J).$$  \hfill (2)

This equation involves $X$ and $Y$ only through their values on the curve. Consequently, condition (1) for parallelism is independent of the choice of $X$ and $Y$. It is now obvious how to define parallelism with respect to any finite curve segment $\gamma_j$ and finally with respect to the entire curve $\gamma$.

Definition. Let $\gamma: t \mapsto \gamma(t)$ ($t \in I$) be a curve in $M$. The curve $\gamma$ is called a geodesic if the family of tangent vectors $\dot{\gamma}(t)$ is parallel with respect to $\gamma$. A geodesic $\gamma$ is called maximal if it is not a proper restriction of any geodesic.

Suppose $\gamma_j$ is a finite geodesic segment without double points con-
tained in a coordinate neighborhood $U$ where the coordinates $\{x_1, \ldots, x_m\}$ are valid. Then (2) implies

$$\frac{dx_k}{dt} + \sum_{i,j} R^k_{ij} \frac{dx_i}{dt} \frac{dx_j}{dt} = 0 \quad (t \in J). \quad (3)$$

If we change the parameter on the geodesic and put $t = f(s)$, $(f'(s) \neq 0)$, then we get a new curve $s \rightarrow \gamma(f(s))$. This curve is a geodesic if and only if $f$ is a linear function, as (3) shows.

**Proposition 5.2.** Let $p, q$ be two points in $M$ and $\gamma$ a curve segment from $p$ to $q$. The parallelism $\tau$ with respect to $\gamma$ induces an isomorphism of $M_p$ onto $M_q$.

Proof. Without loss of generality we may assume that $\gamma$ has no double points and lies in a coordinate neighborhood $U$. Let $\{x_1, \ldots, x_m\}$ be a system of coordinates on $U$. Suppose the curve segment $\gamma$ is given by the mapping $t \rightarrow \gamma(t)$ $(a \leq t \leq b)$ such that $\gamma(a) = p, \gamma(b) = q$. As before we put $x_i(t) = x_i(\gamma(t))$ $(a \leq t \leq b)$ $(1 \leq i \leq m)$.

Consider the system (2). From the theory of systems of ordinary, linear differential equations of first order we can conclude:

There exist $m$ functions $\varphi_i(t, y_1, \ldots, y_m)$ $(1 \leq i \leq m)$ defined and differentiable† for $a \leq t \leq b, -\infty < y_i < \infty$ such that

(i) For each $m$-tuple $(y_1, \ldots, y_m)$, the functions $Y_i(t) = \varphi_i(t, y_1, \ldots, y_m)$ satisfy the system (2).

(ii) $\varphi_i(a, y_1, \ldots, y_m) = y_i \quad (1 \leq i \leq m)$.

The functions $\varphi_i$ are uniquely determined by these properties.

The properties (i) and (ii) show that the family of vectors $Y(t) = \sum_i Y_i(t) (\partial / \partial x_i)$ $(a \leq t \leq b)$ is parallel with respect to $\gamma$ and that $Y(a) = \sum_i y_i (\partial / \partial x_i)_p$. The mapping $Y(a) \rightarrow Y(b)$ is a linear mapping of $M_p$ into $M_q$ since the functions $\varphi_i$ are linear in the variables $y_1, \ldots, y_m$. This mapping is one-to-one owing to the uniqueness of the functions $\varphi_i$.

Consequently, it is an isomorphism.

**Proposition 5.3.** Let $M$ be a differentiable manifold with an affine connection. Let $p$ be any point in $M$ and let $X \neq 0$ in $M_p$. Then there exists a unique maximal geodesic $t \rightarrow \gamma(t)$ in $M$ such that

$$\gamma(0) = p, \quad \gamma(0) = X. \quad (4)$$

† A function on a closed interval $I$ is called differentiable on $I$ if it is extendable to a differentiable function on some open interval containing $I$. 19
Proof. Let \( q : q \to (x_1(q), \ldots, x_m(q)) \) be a system of coordinates on a neighborhood \( U \) of \( p \) such that \( \varphi(U) \) is a cube \( \{(x_1, \ldots, x_m) : |x_i| < c\} \) and \( \varphi(p) = 0 \). Then \( X \) can be written \( X = \sum \alpha_i \frac{\partial}{\partial x_i} \) where \( \alpha_i \in \mathbb{R} \).

We consider the system of differential equations

\[
\frac{dx_i}{dt} = z_i \quad (1 \leq i \leq m),
\]

\[
\frac{dx_k}{dt} = -\sum_{i=1}^m \Gamma_{ik}^j (x_1, \ldots, x_m) x_i x_j \quad (1 \leq k \leq m),
\]

with the initial conditions

\[
(x_1, \ldots, x_m, x_1, \ldots, x_m)_{t=0} = (0, \ldots, 0, \alpha_1, \ldots, \alpha_m).
\]

Let \( c_1, K \) satisfy \( 0 < c_1 < c, 0 < K < \infty \). In the interval \( |x_i| < c_1, |x_i| < K (1 \leq i \leq m) \), the right-hand sides of the foregoing equations satisfy a Lipschitz condition.

From the existence and uniqueness theorem (see, e.g., Miller and Murray\(^k\), p. 42) for a system of ordinary differential equations we conclude:

There exists a constant \( b_1 > 0 \) and differentiable functions \( x_i(t), z_i(t) (1 \leq i \leq m) \) in the interval \( |t| \leq b_1 \) such that

(i) \( \frac{dx_i(t)}{dt} = x_i(t) \quad (1 \leq i \leq m), \quad |t| < b_1 \),

(ii) \( \frac{dx_k(t)}{dt} = -\sum_{i=1}^m \Gamma_{ik}^j (x_1(t), \ldots, x_m(t)) x_i(t) x_j(t) \quad (1 \leq k \leq m), \quad |t| < b_1 \);

(iii) \( |x_i(t)| < c_1, |x_i(t)| < K \quad (1 \leq i \leq m), \quad |t| < b_1 \);

(iv) \( x_i(t), z_i(t) \quad (1 \leq i \leq m) \) is the only set of functions satisfying the conditions (i), (ii), and (iii).

This shows that there exists a geodesic \( t \to \gamma(t) \) in \( M \) satisfying (4) and that two such geodesics coincide in some interval around \( t = 0 \). Moreover, we can conclude from (iv) that if two geodesics \( t \to \gamma_1(t) \) \( (t \in I_1) \), \( t \to \gamma_2(t) \) \( (t \in I_2) \) coincide in some open interval, then they coincide for all \( t \in I_1 \cap I_2 \). Proposition 5.3 now follows immediately.

Definition. The geodesic with the properties in Prop. 5.3 will be denoted \( \gamma_x \). If \( X = 0 \), we put \( \gamma_x(t) = p \) for all \( t \in \mathbb{R} \).

§ 6. The Exponential Mapping

Suppose again $M$ is a $C^\infty$ manifold with an affine connection. Let $p \in M$. We use the notation from the proof of Prop. 5.3. We shall now study the solutions of (5) and (5') and their dependence on the initial values. From the existence and uniqueness theorem (see, e.g., Miller and Murray, p. 64) for the system (5), (5'), we can conclude:

There exists a constant $b$ ($0 < b < c$) and differentiable functions $\varphi_i(t, \xi, ..., \xi_m, \zeta, ..., \zeta_m)$ for $|t| \leq 2b$, $|\xi| \leq b$, $|\zeta| \leq b$ ($1 \leq i, j \leq m$) such that:

(i) For each fixed set $(\xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$ the functions

$$x_i(t) = \varphi_i(t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$$

$$z_i(t) = \left[ \frac{\partial \varphi_i}{\partial t} \right] (t, \xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m), \quad 1 \leq i \leq m, \quad |t| \leq 2b,$$

satisfy (5) and (5') and $|x_i(t)| < c_1$, $|z_i(t)| < K$.

(ii) $(x_1(t), ..., x_m(t), z_1(t), ..., z_m(t)))_{t=0} = (\xi_1, ..., \xi_m, \zeta_1, ..., \zeta_m)$.

(iii) The functions $\varphi_i$ are uniquely determined by the above properties.

Theorem 6.1. Let $M$ be a manifold with an affine connection. Let $p$ be any point in $M$. Then there exists an open neighborhood $N_0$ of 0 in $M_p$ and an open neighborhood $N_p$ of $p$ in $M$ such that the mapping $X \rightarrow \gamma_X(1)$ is a diffeomorphism of $N_0$ onto $N_p$.

Proof. Using the notation above, we put

$$\psi_i(t, \xi_1, ..., \zeta) = \varphi_i(t, 0, ..., 0, \xi_1, ..., \zeta_m)$$

for $1 \leq i \leq m$, $|t| \leq 2b$, $|\xi| \leq b$. Then

$$\psi_i(0, \xi_1, ..., \zeta_m) = 0,$$

$$\frac{\partial \psi_i}{\partial t}(0, \xi_1, ..., \zeta_m) = \xi_i.$$

Since $\gamma_X(st) = \gamma_{sx}(t)$, the uniqueness (iii) implies

$$\psi_i(st, \xi_1, ..., \zeta_m) = \psi_i(t, s\xi_1, ..., s\zeta_m) \quad (1)$$
The map $X \rightarrow \gamma X(b)$ has coordinate expression
\[ \Psi : (\zeta_1, \ldots, \zeta_m) \rightarrow (\psi_1(b, \zeta_1, \ldots, \zeta_m), \ldots, \psi_m(b, \zeta_1, \ldots, \zeta_m)) \]
and we calculate its Jacobian at $(0, \ldots, 0)$.
\[ \left( \frac{\partial \psi_i}{\partial \zeta_j} \right)_{(0, \ldots, 0)} = \lim_{h \to 0} \frac{\psi_i(b, 0, \ldots, hb, \ldots, 0) - \psi_i(b, 0, \ldots, 0)}{hb} \]
Using (1) and the relation $0 = \psi_i(b, 0, \ldots, 0) = \psi_i(0, \ldots, b, \ldots, 0)$ this limit is
\[ \lim_{h \to 0} \frac{\psi_i(hb, 0, \ldots, b, \ldots, 0) - \psi_i(0, \ldots, b, \ldots, 0)}{hb} = \left( \frac{\partial \psi_i}{\partial t} \right)(0, \ldots, b, \ldots, 0) = b \delta_{ij} \]
Thus the Jacobian at $(0, \ldots, 0)$ equals $b^m$. Since $\gamma X(b) = \gamma X(1)$ the theorem follows.

**Definition.** The mapping $X \rightarrow \gamma X(1)$ described in Theorem 6.1 is called the **Exponential mapping** at $p$ and will be denoted by Exp (or Exp$_p$).

**Definition.** Let $M$ be a manifold with an affine connection and $p$ a point in $M$. An open neighborhood $N_0$ of the origin in $M_0$ is said to be **normal** if: (1) the mapping Exp is a diffeomorphism of $N_0$ onto an open neighborhood $N_p$ of $p$ in $M$; (2) if $X \in N_0$ and $0 \leq t \leq 1$, then $tX \in N_p$.

The last condition means that $N_0$ is "star-shaped." A neighborhood $N_p$ of $p$ in $M$ is called a **normal neighborhood** of $p$ if $N_p = \text{Exp} N_0$ where $N_0$ is a normal neighborhood of 0 in $M_0$. Assuming this to be the case, and letting $X_1, \ldots, X_m$ denote some basis of $M_p$, the inverse mapping
\[ \text{Exp}_p(a_1X_1 + \ldots + a_nX_n) \rightarrow (a_1, \ldots, a_n) \]
of $N_p$ into $\mathbb{R}^m$ is called a system of normal coordinates at $p$.

**§ 7. Covariant Differentiation**

In § 5, parallelism was defined by means of the covariant differentiation $\nabla_X$. Theorem 7.1 below shows that it is also possible to go the other way and describe the covariant derivative by means of parallel translation. This makes it possible to define the covariant derivative of other objects.
Definition. Let $X$ be a vector field on a manifold $M$. A curve $s \to \varphi(s)$ ($s \in I$) is called an integral curve of $X$ if

$$\varphi(s) = X_{\varphi(s)}, \quad s \in I.$$ (1)

Assuming $0 \in I$, let $p = \varphi(0)$ and let $\{x_1, \ldots, x_m\}$ be a system of coordinates valid in a neighborhood $U$ of $p$. There exist functions $X^i \in C^\infty(U)$ such that $X = \sum_{i} X^i \partial / \partial x_i$ on $U$. For simplicity let $x_i(s) = x_i(\varphi(s))$ and write $X^i$ instead of $(X^i)^*$ (§ 2, No. 1). Then (1) is equivalent to

$$\frac{dx_i}{ds} = X^i(x_1(s), \ldots, x_m(s)) \tag{1} \quad (1 \leq i \leq m).$$ (2)

Therefore if $X_p \neq 0$ there exists an integral curve of $X$ through $p$.

**Theorem 7.1.** Let $M$ be a manifold with an affine connection. Let $p \in M$ and let $X, Y$ be two vector fields on $M$. Assume $X, Y \neq 0$. Let $s \to \varphi(s)$ be an integral curve of $X$ through $p = \varphi(0)$ and $\tau_t$ the parallel translation from $p$ to $\varphi(t)$ with respect to the curve $\varphi$. Then

$$(\nabla_X(Y))_p = \lim_{s \to 0} \frac{1}{s} (\tau^{-1}_s Y_{\varphi(s)} - Y_p).$$

**Proof.** We shall use the notation introduced above. Consider a fixed $s > 0$ and the family $Z_{\varphi(t)}$ ($0 \leq t \leq s$) which is parallel with respect to the curve $\varphi$ such that $Z_{\varphi(0)} = \tau^{-1}_s Y_{\varphi(0)}$. We can write

$$Z_{\varphi(t)} = \sum_i Z^i(t) \left( -\frac{\partial}{\partial x_i} \right)_{\varphi(0)}, \quad Y_{\varphi(t)} = \sum_i Y^i(t) \left( -\frac{\partial}{\partial x_i} \right)_{\varphi(0)},$$

and have the relations

$$Z^k(t) + \sum_{i,j} \Gamma_{ij}^k Z_i(t) Z^j(t) = 0 \quad (0 \leq t \leq s)$$

$$Z^k(s) = Y^k(s) \quad (1 \leq k \leq m).$$

By the mean value theorem

$$Z^k(s) = Z^k(0) + sZ^k(t^*)$$

for a suitable number $t^*$ between 0 and $s$. Hence the $k$th component of $(1/s) (\tau^{-1}_s Y_{\varphi(s)} - Y_p)$ is

$$\frac{1}{s} (Z^k(0) - Y^k(0)) = \frac{1}{s} (Z^k(s) - sZ^k(t^*) - Y^k(0))$$

$$= \sum_{i,j} \Gamma_{ij}^k(\varphi(t^*)) \dot{x}_i(t^*) Z^j(t^*) + \frac{1}{s} (Y^k(s) - Y^k(0)).$$

As $s \to 0$ this expression has the limit

$$\frac{dY^k}{ds} + \sum_{i,j} \Gamma_{ij}^k \frac{ds_i}{ds} Y^j.$$ (3)

Let this last expression be denoted by $A_k$. It was shown earlier that

$$\nabla_X(Y)_p = \sum_k A_k \left( -\frac{\partial}{\partial x_k} \right)_p.$$ (4)

This proves the theorem.