The Classical Theory of Affine Connections

by

J.W. Auer

THE CLASSICAL THEORY OF AFFINE CONNECTIONS

by

J.W. Auer

Submitted to the Faculty of Graduate Studies and Research, McGill University, in partial fulfilment of the requirements for the degree of Master of Science.

Department of Mathematics
McGill University

July 1966

ACKNOWLEDGEMENT

I would like to express my thanks to Professor B.A. Rattray for his help and advice.

CONTENTS

1. Introduction

The theory of affine connections is, roughly speaking, a generalization of certain concepts of parallelism and differentiation defined in plane differential geometry, to the differential geometry of surfaces, and, more generally, to the geometry of differentiable manifolds. It is the purpose of this essay to relate the various stages of this generalization, and to present the essentials of the classical theory of affine connections on a differentiable manifold.

2. Affine Connections on Surfaces in Three-Dimensional Euclidean Space

2.1 Surfaces

In classical differential geometry, a <u>surface</u> (strictly, a portion of a surface) is defined as a C^k homeomorphism $\bar{x}$ of an open connected subset of the Cartesian plane into Euclidean 3-space, denoted by E_3 .

By a <u>C^k mapping</u> is meant, as usual, a mapping whose first k derivatives (partial, or ordinary) are continuous; we shall assume when speaking of C^k mappings, that k is large enough to make the discussion meaningful, say, $k \geq 3$. In terms of the homeomorphism $\bar{x}$, where $\bar{x}$ represents the position vector of any point of the surface, the equation of the surface may be written as

$$\bar{x} = \bar{x}(u^1,u^2) = (x_1(u^1,u^2),\ x_2(u^1,u^2),\ x_3(u^1,u^2)),$$

where u^1,u^2 are coordinates in the plane. If to the surface we now adjoin the class of "allowable" coordinate transformations, the surface becomes a two-dimensional differentiable manifold.

A <u>curve</u> C on the surface is a C^k homeomorphism of an open subset (a,b) of the real line into the surface, so that as a function of a parameter t, $t \in (a,b)$, the curve may be represented by $\bar{x}(t) = \bar{x}(u^1(t),u^2(t))$.

The <u>tangent</u> at any point $u^i_o = u^i(t_o)$, i=1,2, of the curve has the direction of the vector

$$\bar{x}' = \left.\frac{d\bar{x}}{dt}\right|_{t=t_o} = \bar{x}_i u^{i\prime} \quad \text{(summation)},$$

where $\bar{x}_i = \left.\frac{\partial \bar{x}}{\partial u^i}\right|_{u^i=u^i_o}$. This vector is visualized as pointing into the space E_3 in which the surface is imbedded; furthermore, we may think of attaching to each point u^i_o of the surface a two-dimensional vector space called the <u>tangent plane</u>, and spanned by the vectors $\bar{x}_1,\bar{x}_2$;

geometrically, this is the plane tangent to the surface. A rule which assigns to each point u^i of the surface a tangent vector $v^i\bar{x}_i$ is called a vector field ; if a tangent vector is defined at each point on a curve C, it is called a vector field along the curve C.

2.2 Parallelism in the Plane and in Surfaces

In particular, we now consider a vector field $\bar{v}(t) = \bar{x}_i v^i(t)$ defined along the curve C: $u^i = u^i(t)$ in the plane, which can be considered as the surface

$$\bar{x}(u^1,u^2) = (u^1,u^2,0) .$$

The vector field $\bar{v}(t)$ is called a field of (geodesically*) parallel vectors (also, a parallel field of vectors) if each vector of the field can be obtained from every other vector of the field by a displacement; that is , for all t, $v^i(t) = v^i(t_o)$, where t_o is some fixed value of the parameter t . Equivalently, the field is parallel iff $\bar{v}' = 0$, where as usual, ' denotes the derivative with respect to t . This definition is of course motivated by the well known properties of parallelism in the plane. For example, for any two values t_1,t_2 of t, the vectors $v^i(t_1)$ and $v^i(t_2)$ make equal angles with the geodesic (that is, with the straight line) joining $u^i(t_1)$ with $u^i(t_2)$.

It is obvious that the above definition of parallelism is inadequate for vector fields on an arbitrary

* Laugwitz, page 48.

surface, since the coincidence of the tangent planes at all points of a surface is a unique property of the plane. Accordingly, the following definition of parallelism of a vector field on a surface was proposed by Levi-Cività:

<u>Definition 1</u>

Let C: $\bar{x}(u^i(t))$ be a curve on the surface $\bar{x} = \bar{x}(u^i)$, and $\bar{v}(t)$ a vector field along C. The field $\bar{v}(t)$ is called (geodesically) <u>parallel</u> (with respect to the curve C) iff $\bar{x}_i . \bar{v}' = 0$, $i = 1, 2$; that is, iff the component of $\bar{v}'$ in the tangent plane vanishes. Here "." denotes the inner product.

This definition will be shown in what follows to have the essential properties of parallelism present in the corresponding definition for the plane.

<u>Theorem 1</u>

Geodesic parallelism of vector fields is an intrinsic property of surfaces.

<u>Proof</u>: By definition, the vector field $\bar{v}(t) = \bar{x}_j v^j(t)$ is parallel along C iff

$0 = \bar{x}_i . (\bar{x}_j v^j(t))' = \bar{x}_i . \bar{x}_{jk} u^{k\prime} v^j + \bar{x}_i . \bar{x}_j v^{j\prime}$, $i = 1, 2$;

that is, iff

$$g_{ij} v^{j\prime} + [jki] \, u^{k\prime} v^j = 0 \qquad (1)$$

where $g_{ij} = \bar{x}_i . \bar{x}_j$ denotes the metric tensor of the surface, and $[jki]$ is the Christoffel symbol of the first kind, defined by $[jki] = \bar{x}_{jk} . \bar{x}_i$ ($\bar{x}_{jk} = \frac{\partial \bar{x}_j}{\partial u^k}$)

The result then follows by virtue of the fact that*

* Kreyszig, page 148.

$$[jki] = \frac{1}{2}\left[\frac{\partial g_{ji}}{\partial u^k} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^i}\right] \quad ,$$

and by the definition of "intrinsic" property.

By taking the inner product of equation (1) with g^{ih}, we obtain the more convenient condition for parallelism:

$$v^{h\prime} + \left\{ {h \atop j\ k} \right\} u^{k\prime} v^j = 0 \qquad (2)$$

with

$\left\{ {h \atop j\ k} \right\} = g^{ih}\ [jki]$, the Christoffel symbols of the second kind. As will be seen in the following, the Christoffel symbols of the second kind constitute the components of an (unique in a certain sense) affine connection in the Riemannian space determined by the metric tensor of the surface. We first complete the analogy with parallelism in the plane by means of the following theorem:

<u>Theorem 2</u>

If C: $u^i(t)$ is a curve on the surface $\bar{x} = \bar{x}(u^i)$ and $a^i(t)$, $b^i(t)$ are fields of parallel vectors along C, then 1) the lengths of the $a^i(t)$ are constant along C,

2) the angle between $a^i(t)$ and $b^i(t)$ is constant along C.

<u>Proof</u>: Consider $\frac{d}{dt}(g_{ij}a^i b^j)$

$$= \frac{\partial g_{ij}}{\partial u^k} u^{k\prime} a^i b^j + g_{ij} a^{i\prime} b^j + g_{ij} a^i b^{j\prime} \ .$$

Using (2) with $a^{i\prime}, b^{j\prime}$ for $v^{h\prime}$, we have, as a consequence of the parallelism of a^i and b^j,

$$a^{i\prime} = -\left\{ {i \atop k\ n} \right\} u^{k\prime} a^n \quad \text{and} \quad b^{j\prime} = -\left\{ {j \atop k\ n} \right\} u^{k\prime} b^n$$

Substituting these expressions, we get

$$\frac{d}{dt}(g_{ij}a^ib^j) = \frac{\partial g_{ij}}{\partial u^k} u^{k\prime} a^i b^j - \left\{ {i \atop k\ n} \right\} u^{k\prime}\, a^n g_{ij} b^j - \left\{ {j \atop k\ n} \right\} u^{k\prime} b^n g_{ij} a^i$$

$$= \left[\frac{\partial g_{ij}}{\partial u^k} - \left\{ {m \atop k\ i} \right\} g_{mj} - \left\{ {m \atop k\ j} \right\} g_{im} \right] u^{k\prime} a^i b^j ,$$

after renaming summation indices. Thus,

$$\frac{d}{dt}(g_{ij}a^ib^j) = \left[\frac{\partial g_{ij}}{\partial u^k} - [kij] - [kji] \right] u^{k\prime} a^i b^j = 0,$$

since differentiation of the identity $g_{ij} = \bar{x}_i \cdot \bar{x}_j$ shows that $\frac{\partial g_{ij}}{\partial u^k} = [kij] + [kji]$ (cf. the proof of theorem 1).

As a result, the theorem is proved, since

1) the length of the vector $a^i(t)$ is defined to be $(\bar{x}_i a^i) . (\bar{x}_j a^j) = g_{ij} a^i a^j$, while

2) the angle θ between a^i and b^j can be shown to be given by the expression*

$$\cos \theta = \frac{g_{ij} a^i b^j}{\sqrt{g_{mn} a^m a^n \, g_{rs} b^r b^s}}$$

Corollary

If $a^i(t)$ is a parallel vector field defined

* Kreyszig, page 110.

along a geodesic C: $u^i(t)$, the angle between the tangent vector to C at any point $u^i(t)$, and $a^i(t)$ is constant along C.
Proof: This follows at once from the fact that the tangent vectors to a geodesic C form a parallel vector field . For, the geodesics satisfy the equation $u^{i''} + \left\{ {i \atop j\ k} \right\} u^{j'} u^{k'} = 0$, and the tangent to the curve $\bar{x}(u^i(t))$ is $\bar{x}' = \bar{x}_i u^{i'}(t)$; ie, it is the vector field with components $u^{i'}$.

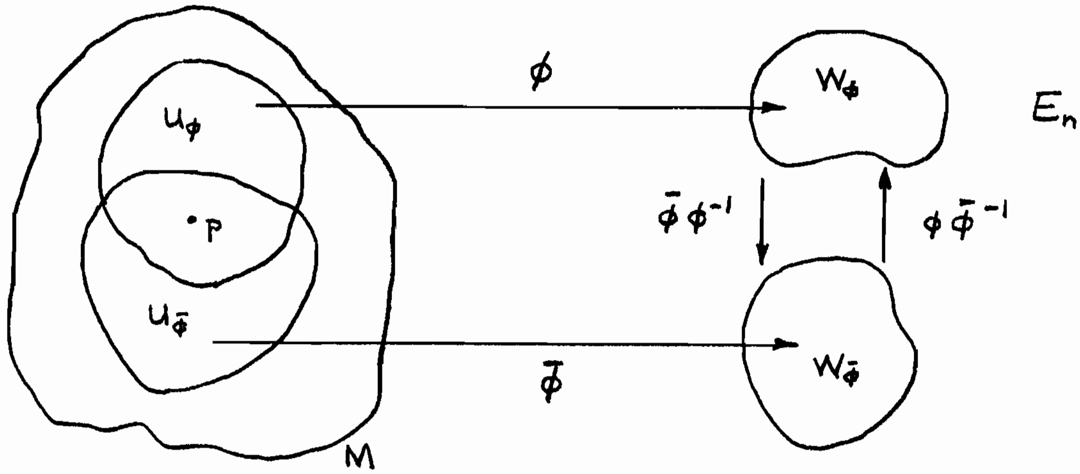
2.3 Summary

The preceding section has shown that the definition of parallelism for surfaces given by Levi-Civitá has the essential properties of parallelism in the plane. Furthermore, the corollary to theorem 2 has strengthened the analogy between geodesics on surfaces and straight lines in the plane: the tangent vectors of both types of curve are geodesically parallel. In the next section we define affine connections on a differentiable manifold, and thereafter show how the Christoffel symbols furnish such a connection on a surface seen as a differentiable manifold.

3. Affine Connections on a Differentiable Manifold

3.1 Notation

In the following, M is a differentiable manifold of class C^k ; this means that the charts at each $p \in M$ are related by coordinate transformations of class C^k.



If ϕ , $\bar\phi$ are charts at a point $p \in M$,

$$\phi : U_\phi \to W_\phi \subset E_n$$
$$\bar\phi : U_{\bar\phi} \to W_{\bar\phi} \subset E_n ,$$

then, as any point of W_ϕ can be represented by an n-tuple $(x^1,..,x^n)$, if $q \in U_\phi$ and $(x^1,...,x^n) = \phi(q)$, $(x^1,...,x^n)$ are called the <u>coordinates</u> of q in the chart ϕ. By definition (of differentiable manifold), the mappings given by

$$(x^1,...,x^n) = \phi\,\bar\phi^{-1}(\bar x^1,...,\bar x^n)$$
$$(\bar x^1,...,\bar x^n) = \bar\phi\,\phi^{-1}(x^1,...,x^n)$$

are differentiable of class C^k; they will be denoted henceforth in the abbreviated form $x^i = x^i(\bar x^j)$, $\bar x^i = \bar x^i(x^j)$. In this notation, by definition,

$$\det \left(\left. \frac{\partial x^i}{\partial \bar x^j} \right|_p \right) \neq 0$$

3.2 The Tangent Space, and Tensor Fields

For completeness, we briefly review some of the definitions regarding tangent spaces and tensor fields on M. (i) A C^m curve $(m \leq k)$ C in M is a mapping $\Theta : (a,b) \to M$ such that for all $t \in (a,b)$ and for all charts ϕ at $\Theta(t)$,

$\emptyset\,\Theta$ is C^m. C is a curve <u>through</u> the point $\Theta(t)$ for all $t \in (a,b)$. When we identify the point $\Theta(t)$ with its coordinates, $(x^1,\dots,x^n) = \emptyset\,\Theta(t)$, we abbreviate the notation for C to C: $x^i(t)$.

(ii) A function f: $M \to$ Reals is <u>differentiable</u> (of class C^k) at a point $p \in M$ iff for all charts $\emptyset$ at p, $\emptyset\, f$ is C^k at p. Again, if p is identified with its coordinates, we write $f(x^i)$ instead of $f(p)$.

(iii) Let D_p be the class of differentiable functions at p, and $\emptyset, x^i$ a chart at p. If C: $x^i(t)$ is a C^m curve through p, with $\emptyset(p) = x^i(t_o)$, the <u>tangent</u> <u>vector</u> to C at p is the operator* L: $D_p \to$ Reals defined by $L = L^i \dfrac{\partial}{\partial x^i}$, with

$$L(f) = L^i \left.\frac{\partial f}{\partial x^i}\right|_p = \left.\frac{d}{dt} f(x^i(t))\right|_{t=t_o}$$

and $L^i = \left.\dfrac{dx^i}{dt}\right|_{t=t_o}$. It can be shown that the set of operators so defined form a vector space of dimension n, called the <u>tangent</u> <u>space</u> T_p at p*. Furthermore, the tangent vectors $\dfrac{\partial}{\partial x^k}$ to the n curves defined by

$$C_k : x^i(t) = x^i(t_o) + (t-t_o)\,\delta^i_k \ , \ k=1,\dots,n,$$

form a basis for this vector space; hence, any tangent vector L at p (ie, any $L \in T_p$) can be represented (in the chart $\emptyset, x^i$) as the sum $L = L(x^i)\dfrac{\partial}{\partial x^i}$. Note that each

* Cohn, page 11.

x^i is a differentiable function on the coordinate neighborhood U_ϕ .

(iv) An important consequence of the above definitions is that a coordinate transformation $x^i \to \bar{x}^i$ at p induces the linear transformation $T_p \to T_p$ defined by

$$\frac{\partial}{\partial \bar{x}^k} = \left.\frac{\partial x^i}{\partial \bar{x}^k}\right|_p \cdot \frac{\partial}{\partial x^i} \quad \text{, as follows}$$

from (iii) . This is important, because if T is an (r,s) tensor*, represented by components $t^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ in the basis $\left\{\frac{\partial}{\partial x^i}\right\}$ of T_p corresponding to the chart x^i, then the law of transformation of the components of T induced by the coordinate transformation $x^i \to \bar{x}^i$ is

$$\bar{t}^{k_1 \ldots k_r}_{h_1 \ldots h_s} = t^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial x^{k_1}}{\partial \bar{x}^{i_1}} \cdots \frac{\partial x^{k_r}}{\partial \bar{x}^{i_r}} \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_s}}{\partial x^{h_s}} \qquad (3)$$

3.3 The Covariant Derivative

In order to generalize the ideas of parallelism to differentiable manifolds (which lack the structure afforded by the metric tensor g_{ij} of surfaces), we first try to find an invariant definition of the derivative of a contravariant vector field. For, if we consider the contravariant vector field with components v^i and $\bar{v}^i$ in the charts x^i and $\bar{x}^i$, respectively, then on account of

---

* that is, $T \in T_p \otimes T_p \otimes \overset{(r)}{\cdots} \otimes T_p \otimes T_p^* \otimes T_p^* \otimes \overset{(s)}{\cdots} \otimes T_p^*$;
cf. Auslander and MacKenzie, page 195.

the law of transformation (3), $\bar{v}^i = \frac{\partial \bar{x}^i}{\partial x^k} v^k$, we have that

$$\bar{v}^i_{,j} = \frac{\partial v^k}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \bar{x}^i}{\partial x^k} + v^k \frac{\partial^2 \bar{x}^i}{\partial x^k \partial x^h} \frac{\partial x^h}{\partial \bar{x}^j} \qquad (4)$$

where $\bar{v}^i_{,j} = \frac{\partial \bar{v}^i}{\partial \bar{x}^j}$, by definition. Because of the presence of the right-hand term in (4), $v^i_{,j}$ is not a tensor, and consequently does not in general have a meaning invariant under coordinate transformations. Accordingly, the following definition is made:

<u>Definition 2</u> *

The <u>covariant</u> <u>derivative</u> of a contravariant vector field v^i (with respect to x^j), denoted by $v^i_{;j}$, is defined by the equation

$$v^i_{;j} = v^i_{,j} + A^i_{jk}(x^r)\, v^k \qquad (5)$$

The functions A^i_{jk} are arbitrary, except that their law of transformation must be such that $v^i_{;j}$ is a (1,1) tensor; that is,

$$\bar{v}^i_{;j} = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial x^h}{\partial \bar{x}^j} v^k_{;h} \qquad (6)$$

To calculate this law of transformation, let $\bar{A}^i_{jk}$ denote the values of the functions A^i_{jk} in the chart $\bar{x}^i$; then,

$$\bar{v}^i_{;j} = \bar{v}^i_{,j} + \bar{A}^i_{jk}\, \bar{v}^k \qquad (7)$$

Substituting (5) and (7) into (6) we get the following:

* Laugwitz, page 96.

$$\bar{v}^i_{,j} + \bar{A}^i_{jk}\,\bar{v}^k = \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^h}{\partial \bar{x}^j}(v^k_{,h} + A^k_{hm}\,v^m)$$

Using (4) for $\bar{v}^i_{,j}$ it follows that

$$v^k_{,h}\frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial \bar{x}^i}{\partial x^k} + v^k \frac{\partial^2 \bar{x}^i}{\partial x^k \partial x^h}\frac{\partial x^h}{\partial \bar{x}^j} + \bar{A}^i_{jk}\,\bar{v}^k$$

$$= \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^h}{\partial \bar{x}^j}(v^k_{,h} + A^k_{hm}\,v^m) \quad ,$$

or, since $\bar{v}^k = \dfrac{\partial \bar{x}^k}{\partial x^m}v^m$,

$$\bar{A}^i_{jk}\frac{\partial \bar{x}^k}{\partial x^m}v^m = A^k_{hm}\,v^m\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^h}{\partial \bar{x}^j} - v^m\frac{\partial^2 \bar{x}^i}{\partial x^m \partial x^h}\frac{\partial x^h}{\partial \bar{x}^j} \qquad (8)$$

Since (8) must hold for all vector fields v^i , we find that the A^i_{jk} must transform according to

$$\bar{A}^i_{jk}\frac{\partial \bar{x}^k}{\partial x^m} = A^k_{hm}\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^h}{\partial \bar{x}^j} - \frac{\partial^2 \bar{x}^i}{\partial x^m \partial x^h}\frac{\partial x^h}{\partial \bar{x}^j} \qquad (9)$$

3.4 Parallel Displacement

As for the existence of such functions A^i_{jk} , it is easily verified that for surfaces the Christoffel symbols of the second kind , $\left\{ {i \atop j\ k} \right\}$, transform according to (9). Furthermore, by referring to equation (2) it can be seen that a vector field $v^i(t)$ defined along a curve C: $u^i(t)$ on a surface $\bar{x} = \bar{x}(u^i)$ is parallel iff

$\frac{Dv^i}{dt} = 0$, where $\frac{Dv^i}{dt} = v^i_{;j}\, u^{j\prime} = \frac{dv^i}{dt} + \left\{ {i \atop j\ k} \right\} v^k u^{j\prime}$,

by definition. $\frac{Dv^i}{dt}$ is clearly a tensor of the same kind as v^i , and is called the absolute derivative of v^i .

Motivated by the theory of surfaces, given a differentiable manifold with functions A^i_{jk} as above, we make the following

Definition 3

Let $v^i(t)$ be a vector field along the curve C: $x^i(t)$ on a differentiable manifold M. Then $v^i(t)$ is called parallel with respect to C iff

$$v^i_{;j}\, x^{j\prime} = v^{i\prime} + A^i_{jk}\, v^k\, x^{j\prime} = 0 \qquad (10)$$

Furthermore, a solution $v^i(t)$ of the differential equations (10) (which are easily seen to be the components of a contravariant vector, cf. Eisenhart, page 13) satisfying $v^i(t_o) = w^i$, $w^i \in T_p$, and where $p \in M$ is the point on C with coordinates $x^i(t_o)$, is said to have been obtained by parallel displacement along C from w^i.

Definition 4

The functions A^i_{jk} are called the components (in the chart x^i) of a linear connection A on M.

This terminology has arisen, because the functions A^i_{jk} connect the tangent spaces at different points of the manifold in the following sense:*

* Laugwitz, page 99.

let $v^i(t)$ be a solution of equations (10) . Then, as these equations are linear in v^i, $v^i(t) = B^i_k(t)v^k(t_o)$, so that the parallel displacement induces a linear transformation B with matrix B^i_k , $B: T_{t_o} \longrightarrow T_t$, where T_t is the tangent space at the point of M with coordinates $x^i(t)$. This linear transformation serves to "connect" the (distinct) tangent spaces at distinct points of M.

3.5 Affine Connections

<u>Definition 5</u>

A linear connection A is called an <u>affine</u> connection iff $A^i_{jk} = A^i_{kj}$; that is, iff the components are symmetric in their subscripts.

From equation (9) it is clear that if the components are symmetric in one coordinate system, then they are symmetric in all coordinate systems.

The terminology "affine connection" arises from the following important theorem:

<u>Theorem 3</u>

A linear connection A on M is symmetric iff for all $p \in M$ there is a chart ϕ at p such that $A^i_{jk}(\phi(p)) = 0$.

<u>Proof</u>: Let $\phi(p) = (x^1_o,\ldots,x^n_o)$, and let $\bar{\phi}$ be a second chart at p with $\bar{\phi}(p) = (\bar{x}^1_o,\ldots,\bar{x}^n_o)$. If $A^i_{jk}(x^r_o) = 0$, then by (9) ,

$$\bar{A}^i_{jk}\frac{\partial\bar{x}^k}{\partial x^m} = \frac{-\partial^2\bar{x}^i}{\partial x^m\partial x^h}\frac{\partial x^h}{\partial\bar{x}^j}\text{ , all evaluated at p.}$$

After multiplication of both sides by $\dfrac{\partial x^m}{\partial\bar{x}^q}$, we get

$$\bar{A}^i_{jk}\,\delta^k_q = \bar{A}^i_{jq} = -\frac{\partial^2 \bar{x}^i}{\partial x^m \partial x^h}\,\frac{\partial x^h}{\partial \bar{x}^j}\,\frac{\partial x^m}{\partial \bar{x}^q} \text{ , or ,}$$

$$\bar{A}^i_{jq} = \left(-\frac{\partial^2 \bar{x}^i}{\partial x^k \partial x^m}\,\frac{\partial x^m}{\partial \bar{x}^q}\right)\frac{\partial x^k}{\partial \bar{x}^j}$$

$$= \left(\bar{A}^i_{qn}\,\frac{\partial \bar{x}^n}{\partial x^k}\right)\left(\frac{\partial x^k}{\partial \bar{x}^j}\right) \text{ (by (9) again)}$$

$$= \bar{A}^i_{qn}\,\delta^n_j = \bar{A}^i_{qj} \text{ ,}$$

so that the components are symmetric.

Conversely, let the components of A be symmetric, and let $\bar{\phi}$ be a chart at p with $\bar{\phi}(p) = \bar{x}^i_o$. Then,

$$\bar{A}^i_{jk}(\bar{x}^r_o) = \bar{A}^i_{kj}(\bar{x}^r_o) \text{ .}$$

Consider the (analytic) coordinate transformation defined by

$$x^i = (\bar{x}^i - \bar{x}^i_o) + \frac{1}{2}\,\bar{A}^i_{jk}\,(\bar{x}^r_o)(\bar{x}^j - \bar{x}^j_o)(\bar{x}^k - \bar{x}^k_o) \text{ .}$$

Then, $$\frac{\partial x^i}{\partial \bar{x}^h} = \delta^i_h + \frac{1}{2}\bar{A}^i_{jk}(\bar{x}^j - \bar{x}^j_o)\,\delta^k_h + \frac{1}{2}\,\bar{A}^i_{jk}(\bar{x}^k - \bar{x}^k_o)\,\delta^j_h$$

$$= \delta^i_h + \frac{1}{2}\,\bar{A}^i_{hj}(\bar{x}^j - \bar{x}^j_o) + \frac{1}{2}\,\bar{A}^i_{hj}(\bar{x}^j - \bar{x}^j_o)$$

$$= \delta^i_h + \bar{A}^i_{hj}(\bar{x}^j - \bar{x}^j_o) \text{ .} \qquad (11)$$

Since $\det\left(\left.\frac{\partial x^i}{\partial \bar{x}^h}\right|_p\right) = 1 \neq 0$, the transformation is admissible.

Further, $$\frac{\partial^2 x^i}{\partial \bar{x}^m \partial \bar{x}^h} = \bar{A}^i_{hj}\,\delta^j_m = \bar{A}^i_{hm} \qquad (12)$$

Rewriting (9) in the appropriate form gives

$$A^i_{jk}\frac{\partial x^k}{\partial \bar{x}^m} = \bar{A}^k_{hm}\frac{\partial x^i}{\partial \bar{x}^k}\frac{\partial \bar{x}^h}{\partial x^j} - \frac{\partial^2 x^i}{\partial \bar{x}^m \partial \bar{x}^h}\frac{\partial \bar{x}^h}{\partial x^j} \quad , \qquad (13)$$

and substitution of (12) into (13) results in

$$A^i_{jk}\frac{\partial x^k}{\partial \bar{x}^m} = \frac{\partial^2 x^k}{\partial \bar{x}^m \partial \bar{x}^h}\frac{\partial x^i}{\partial \bar{x}^k}\frac{\partial \bar{x}^h}{\partial x^j} - \frac{\partial^2 x^i}{\partial \bar{x}^m \partial \bar{x}^h}\frac{\partial \bar{x}^h}{\partial x^j} \ .$$

Evaluation at p then gives, using (11),

$$A^i_{jk}(x^i_o)\,\delta^k_m = \frac{\partial^2 x^k}{\partial \bar{x}^m \partial \bar{x}^h}\frac{\partial \bar{x}^h}{\partial x^j}\,\delta^i_k - \frac{\partial^2 x^i}{\partial \bar{x}^m \partial \bar{x}^h} \quad ; \text{ that is,}$$

$$A^i_{jk}(x^i_o) = 0, \text{ as required.}$$

Note that the chart x^i has the property that $x^i_o = 0$, so that we have in fact that $A^i_{jk}(0) = 0$.

The coordinates x^i are called (by Weyl) <u>geodesic</u> at p. Furthermore, the theorem shows the significance of the term <u>affine</u> <u>connection</u>, since in the plane with rectangular Cartesian coordinates, it is true that $\left\{ {i \atop j\ k} \right\} = 0$ for all points; this follows because when the plane is so parametrized,

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

We note that the previous theorem may be extended to the following more general result by Levi-Cività; a proof is given in Non-Riemannian Geometry, by Eisenhart.

Theorem 5

If $p \in M$ and C is an arbitrary curve through p, then there exist an interval (a,b) and a coordinate system x^i such that $A^i_{jk}(x^r(t)) = 0$ for $t \in (a,b)$.

3.6 Covariant Differentiation of Arbitrary Tensors

By making the following two requirements, the covariant differentiation defined above for contravariant vector fields may be extended to arbitrary tensor fields:

1) covariant differentiation of the sum, difference, inner and outer products should obey the same rules as ordinary differentiation;

2) for a scalar field $\phi(x^i)$, $\phi(x^i)_{;j} = \dfrac{\partial \phi}{\partial x^j}$.

Consider, for example, the covariant vector field w_i, and let v^j be an arbitrary (1,0) tensor. Since $w_i v^i$ is a scalar field, we have from 2),

$$(w_i v^i)_{;j} = (w_i v^i)_{,j} = w_{i,j} v^i + w_i v^i_{,j} \tag{14}$$

and by 1), $(w_i v^i)_{;j} = w_{i;j} v^i + w_i v^i_{;j}$ (15)

Also, $v^i_{;j} = v^i_{,j} + A^i_{jk} v^k$, (16)

so that combining the last three equations we get

$$w_{i;j} v^i + v^i_{,j} w_i + A^i_{jk} v^k w_i = w_{i,j} v^i + w_i v^i_{,j} .$$

Thus, $w_{i;j}v^i = w_{i,j}v^i - A^i_{jk}\, v^k\, w_i$, and as v^i was arbitrary, we must have

$$w_{i;j} = w_{i,j} - A^k_{ji}\, w_k \ .$$

In general, if $t^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ are the components of an arbitrary (r,s) tensor, it must follow that

$$t^{i_1 \ldots i_r}_{j_1 \ldots j_s ; j} = t^{i_1 \ldots i_r}_{j_1 \ldots j_s , j} + \sum_{u=1}^{r} A^{i_u}_{jm}\, t^{i_1 \ldots \overset{(u)}{m} \ldots i_r}_{j_1 \ldots \ldots j_s} - \sum_{u=1}^{s} A^{m}_{j j_u}\, t^{i_1 \ldots \ldots i_r}_{j_1 \ldots \underset{(u)}{m} \ldots j_s} \ .$$

(equation (17))

3.7 The Curvature Tensor

Although covariant differentiation has many of the properties of ordinary differentiation, it does not share them all. In particular, it does not generally follow, for a vector field v^i, that $v^i{}_{;j;k} = v^i{}_{;k;j}$.

For,

$$v^i{}_{;j;k} = v^i{}_{;j,k} + A^i_{kr}\, v^r{}_{;j} - A^r_{kj}\, v^i{}_{;r}$$

$$= v^i{}_{,j,k} + A^i_{jr,k}\, v^r + A^i_{jr}\, v^r{}_{,k}$$

$$+ A^i_{kr}\, v^r{}_{,j} + A^i_{kr}A^r_{js}\, v^s - A^r_{kj}\, v^i{}_{,r} - A^r_{kj}A^i_{rs}v^s \quad (18)$$

and similarly, $v^i{}_{;k;j} = v^i{}_{,k,j} + A^i_{kr,j}v^r + A^i_{kr}v^r{}_{,j}$

$$+ A^i_{jr}v^r{}_{,k} + A^i_{jr}A^r_{ks}v^s - A^r_{jk}v^i{}_{,r} - A^r_{jk}A^i_{rs}\, v^s \quad (19)$$

Subtracting (19) from (18) and using the symmetry of A^i_{jk} we get

$$v^i_{;j;k} - v^i_{;k;j} = (A^i_{jr,k} - A^i_{kr,j})v^r + (A^i_{kr}A^r_{js} - A^i_{jr}A^r_{ks})v^s,$$

or, $v^i_{;j;k} - v^i_{;k;j}$

$$= \left(A^i_{js,k} - A^i_{ks,j} + A^i_{kr}A^r_{js} - A^i_{jr}A^r_{ks}\right)v^s \qquad (20)$$

Putting $R^i_{sjk} = A^i_{js,k} - A^i_{ks,j} + A^i_{kr}A^r_{js} - A^i_{jr}A^r_{ks}$, (21)

then, as the left-hand side of (20) is a tensor, so is R^i_{sjk}, and,

$$v^i_{;j;k} - v^i_{;k;j} = R^i_{sjk}\, v^s. \qquad (22)$$

R^i_{sjk} is the <u>curvature</u> <u>tensor</u> of the affine connection. The terminology arises from the theory of surfaces, in particular, from the Theorema Egregium of Gauss, which states that $K.g = g_{2j}\, R^j_{112}$, where, (23)
$g = \det g_{ij}$, and K is the Gaussian curvature. Equation (23) shows that the curvature tensor is not in general zero, since $K = 0$ iff the surface is developable.

By applying the method used above for v^i to an arbitrary (r,s) tensor, in conjunction with equation (17), one gets the generalized identities of Ricci:

$$t^{i_1\ldots i_r}_{j_1\ldots j_s;j;k} - t^{i_1\ldots i_r}_{j_1\ldots j_s;k;j}$$

$$= \sum_{u=1}^{r} t^{i_1..h..i_r}_{j_1\ldots\ldots j_s} R^{i_u}_{hjk} - \sum_{u=1}^{s} t^{i_1\ldots i_r}_{j_1..h..j_s} R^h_{j_u jk},$$

Finally there is

Theorem 6

The curvature tensor R^i_{sjk} of an affine connection satisfies the following identities:

1) $R^i_{sjk} + R^i_{skj} = 0$

2) $R^i_{sjk} + R^i_{jks} + R^i_{ksj} = 0$

3) $R^i_{sjk;h} + R^i_{skh;j} + R^i_{shj;k} = 0$ (Bianchi)

4) $R^i_{sjk;h} + R^i_{ksh;j} + R^i_{hkj;s} + R^i_{jhs;k} = 0$

(Veblen)

Proof: 1) follows immediately from the definition of R^i_{sjk} . To prove 2), by theorem 3, we may choose at every point of the manifold a geodesic coordinate system x^i, so that $A^i_{jk}(x^i) = 0$. Adding equation (21) to itself three times with the appropriate indices, the result follows for the geodesic coordinate system; as 2) is a tensor equation, it holds in all coordinate systems. To prove 3), we again choose geodesic coordinates x^i, so that $R^i_{sjk;h} = R^i_{sjk,h}$; this follows since $A^i_{jk}(x^r) = 0$. However, from equation (21) we then get

$$R^i_{sjk;h} = A^i_{js,k,h} - A^i_{ks,j,h} \tag{24}$$

Now adding (24) to itself three times with the appropriate indices, the result follows. The proof of 4) is similar.

3.8 Geometric Significance of R^i_{sjk}

Following Laugwitz, we prove the following theorem, which yields a geometric interpretation for the curvature tensor in an affinely connected manifold.

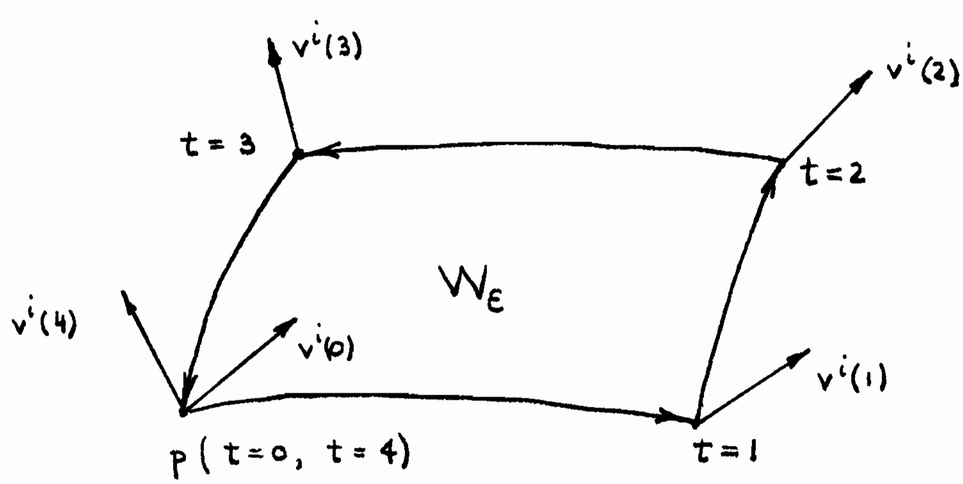
Theorem 7

The curvature tensor is a measure of the change in the components of a vector under parallel displacement around infinitesimal parallelograms.

Proof: Let $\emptyset$ be a chart at $p \in M$ with coordinates x^i such that $\emptyset(p) = 0$, and $A^i_{jk}(0) = 0$; this chart exists by theorem 3. Let n^i , z^i be the components of fixed, linearly independent contravariant vectors, and for each $\varepsilon > 0$, define the curve W_ε : $x^i(t)$ by

$$x^i(t) = \begin{cases} t\,\varepsilon\, n^i , & 0 \leq t < 1 \\ \varepsilon n^i + (t-1)\,\varepsilon\, z^i , & 1 \leq t < 2 \\ \varepsilon (n^i + z^i) - (t-2)\,\varepsilon\, n^i , & 2 \leq t < 3 \\ \varepsilon z^i (1-(t-3)) , & 3 \leq t \leq 4 \end{cases}$$

(equation (25) ; see figure below)



Let $v^i(0) \in T_p$, and let $v^i(t)$ be the vector obtained from $v^i(0)$ by parallel displacement around W_ε .

By Taylor's theorem,

$$v^i(t+1) = v^i(t) + \frac{dv^i}{dt} + \frac{d^2v^i}{dt^2} + \ldots \qquad (26)$$

However, since $v^i(t)$ is a parallel vector field along W_ε ,

$$\frac{dv^i}{dt} = - A^i_{jk}\, v^k x^{j\prime} \qquad (27)$$

and,

$$\frac{d^2v^i}{dt^2} = - A^i_{jk,h}\, v^k x^{j\prime} x^{h\prime} - A^i_{jk} x^{j\prime} \frac{dv^k}{dt}$$

$$= - A^i_{jk,h} v^k x^{j\prime} x^{h\prime} + A^i_{jk} A^k_{rs}\, v^s x^{j\prime} x^{r\prime} \ , \qquad (28)$$

where all functions are evaluated at t.

Again by Taylor's theorem,

$$A^i_{jk}(t) = A^i_{jk}(0) + A^i_{jk,h}(0)\, \Delta x^h(t)$$

$$= A^i_{jk,h}(0)\, \Delta x^h(t) \quad . \qquad (29)$$

Now the change in the component $v^i(t)$ as a result of displacement around W_ε is

$$v^i(4) - v^i(0) = \sum_{j=0}^{3} (v^i(j+1) - v^i(j)) \qquad (30)$$

To calculate this change, we find $\frac{dv^i}{dt}(t)$ and $\frac{d^2v^i}{dt^2}(t)$

at the appropriate values of t, and use (30) and (26) :

from (27),

a) $$\frac{dv^i}{dt}(0) = - A^i_{jk}(0)\ v^k(0)x^{j\prime}(0) = 0 \qquad (31a)$$

b) $$\frac{dv^i}{dt}(1) = - A^i_{jk}(1)v^k(1)x^{j\prime}(1)$$

$$= - A^i_{jk,h}(0)\Delta x^h(1)v^k(1)x^{j\prime}(1) \quad \text{by (29)}$$

$$= - A^i_{jk,h}(0)\,\varepsilon\, n^h v^k(1)\,\varepsilon\, z^j \text{ , since}$$

$$\Delta x^h(1) = \Delta x^h(0) + \frac{dx^h}{dt}(0)\ \Delta t + \ldots$$

$$= \varepsilon n^h \qquad (\text{ as } \Delta t = 1).$$

Therefore,

$$\frac{dv^i}{dt}(1) = - \varepsilon^2 A^i_{jk,h}(0)\ v^k(1)n^h\ z^j \qquad (31b)$$

c) $$\frac{dv^i}{dt}(2) = - A^i_{jk}(2)\ v^k(2)x^{j\prime}(2)$$

$$= - A^i_{jk,h}(0)\ \Delta x^h(2)\ v^k(2)\ (-\varepsilon n^j).$$

Now $\Delta x^h(2) = \Delta x^h(1) + \frac{dx^h}{dt}(1)\ \Delta t = \varepsilon n^h + \varepsilon z^h$, so

$$\frac{dv^i}{dt}(2) = + \varepsilon^2 A^i_{jk,h}(0)\ v^k(2)\ n^j(n^h + z^h) \qquad (31c)$$

d) $$\frac{dv^i}{dt}(3) = - A^i_{jk}(3)\ v^k(3)\ x^{j\prime}(3)$$

$$= + \varepsilon^2 A^i_{jk,h}(0)\ v^k(3)\ z^j z^h\ , \tag{31d}$$

since $\Delta x^h(3) = \Delta x^h(2) + \frac{dx^h}{dt}(2)\,\Delta t$

$$= \varepsilon\,(n^h + z^h) - \varepsilon\, n^h = \varepsilon\, z^h.$$

Similarly, from (28),

e) $\frac{d^2v^i}{dt^2}(0) = - A^i_{jk,h}(0)\ v^k(0)\ x^{j\prime}(0)\ x^{h\prime}(0) + 0$

$$= - A^i_{jk,h}(0) v^k(0) n^j n^h \tag{31e}$$

f) $\frac{d^2v^i}{dt^2}(1) = - A^i_{jk,h}(1) v^k(1) x^{j\prime}(1) x^{h\prime}(1)$

$$+ A^i_{jk}(1)\ A^k_{rs}(1) v^s(1) x^{j\prime}(1) x^{r\prime}(1)$$

$$= - \varepsilon^2 A^i_{jk,h}(1) v^k(1) z^j z^h + \varepsilon^4(\ldots) \tag{31f}$$

Similarly, $\frac{d^2v^i}{dt^2}(2) = \frac{d^2v^i}{dt^2}(0)$ and $\frac{d^2v^i}{dt^2}(3) = \frac{d^2v^i}{dt^2}(1)$

(31g)

Now all functions occurring in equations (31) can in fact be replaced by the same functions evaluated at 0 if we use their Taylor expansions about 0 and neglect terms with factors ε^3 and higher. Calling the resulting equations (31'), substituting (31') into (26), and the resulting equations into (30), we get

$$v^i(4) - v^i(0) = \varepsilon^2 A^i_{jk,h}(0)(n^j z^h - n^h z^j)\ v^k(0) + \varepsilon^3(\ldots)$$

$$= \mathcal{E}^2 (A^i_{jk,h}(0) - A^i_{hk,j}(0)) v^k(0) n^j z^h + \mathcal{E}^3(\ldots)$$

$$= \mathcal{E}^2 R^i_{jkh}(0) v^k(0) n^j z^h + \mathcal{E}^3(\ldots), \qquad (32)$$

as follows from (21) with $A^i_{jk} = 0$.

From (32) it is clear that the change in v^i becomes proportional to $R^i_{jkl}(0)$ when $\mathcal{E}$ is chosen sufficiently small, and this completes the proof.

3.9 Paths

It was shown in section 2 that the geodesics on a surface have many of the properties of parallelism which straight lines have in the plane. In particular, the tangent vectors to a geodesic are related by parallel displacement along the curve. The curves on a differentiable manifold which have this property are called paths* (auto-parallels). Noting that two vectors v^i and w^i defined at a point $p \in M$ have the same direction at p iff $v^i = K w^i$, $K \neq 0$, we have, analogous to the theorem for geodesics, the following result:

Theorem 8

For every $p \in M$ and for every direction v^i at p (ie, specified by a $v^i \in T_p$) , there exists a path through p having the direction v^i. Furthermore, this path is uniquely determined throughout some neighborhood

* Eisenhart, page 14.

of p.

<u>Proof</u>: Let ϕ be a chart at p with coordinates x^i and $\phi(p) = x^i_o$; let C: $x^i(t)$ be a path on M. Then the points of C satisfy the equation

$$x^{i''} + A^i_{jk}\, x^{j'} x^{k'} = 0 \qquad (33a)$$

These differential equations have, by a well known theorem, for sufficiently small t , exactly one solution $x^i(t)$ such that

$$\begin{aligned} x^i(t_o) &= x^i_o \\ x^{i'}(t_o) &= v^i \end{aligned} \qquad (33b)$$

To show that the path $x^i(t)$ is uniquely determined by p and the direction at p, we show that the initial conditions $x^i(t_1) = x^i_o$, $x^{i'}(t_1) = Kv^i$, $K \neq 0$, yield the same solution. However, this follows because a solution corresponding to the latter initial conditions is given by

$$\bar{x}^i(t) = x^i(\ K(t-t_1) + t_o) \text{ ; by the}$$

uniqueness of solutions of (33) , it is the only solution.

4. Riemannian Spaces

It was mentioned in section 2 that the Christoffel symbols $\left\{ {i \atop j\ k} \right\}$ constitute the components of an (unique, in a sense) affine connection on a surface with metric tensor g_{ij} .

We now show that in a general Riemannian space (that is, a differentiable manifold on which a symmetric (0,2) tensor g_{ij} is defined), there is only one affine connection which preserves the lengths of vectors under parallel displacement.

<u>Theorem 9</u>

In a Riemannian space with metric tensor g_{ij} there is an unique affine connection, given by

$$A^i_{jk} = \frac{1}{2} g^{ir} (g_{jr,k} + g_{rk,j} - g_{jk,r}) ,$$

which preserves the lengths of parallel vector fields.

<u>Proof</u>: The proof of the fact that the connection A^i_{jk} defined above preserves the lengths of vectors under parallel displacement is identical to the proof given for $\left\{ {i \atop j\ k} \right\}$ in theorem 2. Thus, it is only necessary to show that if B^i_{jk} is any other affine connection on M, then

$$B^i_{jk} = A^i_{jk} .$$

Let $v^i(t)$ be a parallel vector field along the curve C: $x^i(t)$ on M. If the lengths of parallel vectors are to be preserved, we must have

$$0 = \frac{d}{dt} (g_{ik} v^i v^k) = g_{ik,h} x^{h\prime} v^i v^k + g_{ik} v^{i\prime} v^k + g_{ik} v^i v^{k\prime}$$

$$= g_{ik,h} x^{h\prime} v^i v^k + g_{ik} v^k (-B^i_{mn} v^n x^{m\prime}) + g_{ik} v^i (-B^k_{mn} v^n x^{m\prime})$$

$$= (g_{ik;h} - g_{rk} B^r_{hi} - g_{ir} B^r_{hk}) x^{h\prime} v^i v^k$$

As this must hold for all x^i, v^i, we must have that

$$g_{ik,h} = g_{rk} B^r_{hi} + g_{ir} B^r_{hk} .$$

Using the symmetry of g_{ij} and B^i_{jk} , it follows that

$$g_{ih,k} + g_{hk,i} - g_{ik,h} = 2\, g_{rh} B^r_{ik} .$$

After inner multiplication by $\frac{1}{2} g^{hj}$, we get

$$B^i_{jk} = A^i_{jk} \text{ , as required.}$$

....

BIBLIOGRAPHY

Auslander, L.,
MacKenzie, R.E., Introduction to Differentiable Manifolds, McGraw-Hill, 1963.

Eisenhart, L.P., Non-Riemannian Geometry, A.M.S. Coll. Publications, Vol. 8, 1927.

Kobayashi, S.,
Nomizu, K., Foundations of Differential Geometry, Interscience, 1963.

Kreyszig, E., Differential Geometry, University of Toronto Press, 1959.

Laugwitz, D., Differential and Riemannian Geometry, Academic Press, 1965.

Veblen, O., Invariants of Quadratic Differential Forms, Cambridge University Press, 1927.

Veblen, O.,
Whitehead, J.H,C. The Foundations of Differential Geometry, Cambridge University Press, 1932.