DIFFERENTIAL GEOMETRY

AND

MATHEMATICAL PHYSICS

by

VICTORIA VOROTILKINA

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Masters of science

in the

Mathematics

Program

YOUNGSTOWN STATE UNIVERSITY

MAY, 2003

DIFFERENTIAL GEOMETRY

AND

MATHEMATICAL PHYSICS

Victoria Vorotilkina

Signature: Vorotilkina 04.25.03

Victoria Vorotilkina, Student — Date

Approvals: Steven L. Kent 4/25/03

Steven Kent, Thesis Advisor — Date

Eric J. Wingler 4/25/03

Eric Wingler, Committee Member — Date

Richard A. Goldthwait 4/25/03

Richard Goldthwait, Committee Member — Date

Peter J. Kasvinsky 4/29/03

Peter J. Kasvinsky, Dean of Graduate Studies — Date

ABSTRACT

We will begin with basic definitions in the study of differentiable manifolds, including relevant definitions and properties from point set topology. After developing both the geometric and coordinate dependent approaches to the study of tensors on a manifold, we will investigate some of the applications of the mathematical ideas to the study of electricity and magnetism, and to its mathematical generalization, Yang-Mills field theory.

ACKNOWLEDGMENTS

I wish to express my gratitude to my thesis advisor, Dr. Steven Kent, whose guidance and patience led me to believe in my own abilities.

I wish to thank Dr.Wingler and Dr.Goldthwait for their corrections to a draft of this paper.

Ultimately, I would like to express my appreciation to my family for their love and support.

TABLE OF CONTENTS

INTRODUCTION

The purpose of this thesis is to show a detailed analysis of the difficult concepts based on differential geometry, tensor theory, and some of their applications to mathematical physics. We are going to explain all of the concepts and notation in such a manner that will lead to a readable presentation of inherently difficult material. Some of the material appears together in a manner which is hard to find elsewhere.

First in this thesis we introduce the concept of a differentiable manifold (a knowledge of which has become useful in an increasing number of areas of mathematics and of its applications) and the concept of vectors and tensors, which are the natural geometric objects defined on the manifold. We will treat the manifold as being a space which is locally similar to Euclidean space and will study important concepts defined by the manifold structure which are independent of the choice of a coordinate system.

A discussion of maps of manifolds will lead to the definitions of the induced maps of tensors. We will study the operation of exterior differentiation, which depends only on the manifold structure. And by imposing extra structure, the connection, we will define the covariant derivative and the curvature tensor.

We will also give a brief discussion of fibre bundles since these are used in some applications of mathematical physics.

We will investigate some of the applications of the mathematical ideas to the study of electricity and magnetism, and to its mathematical generalization, Yang-Mills field theory.

Some Topological Preliminaries.

Definition 1 : A subset U of $\mathbb{R}^n$ is defined to be an open subset of $\mathbb{R}^n$ if for each $p \in$U, there is an $\varepsilon > 0$ such that $\mathbb{N}_\varepsilon(p) \subseteq$U ($\mathbb{N}_\varepsilon(p) = \{q \in \mathbb{R}^n : \delta(p,q) < \varepsilon\}$).

Definition 2 : The collection of all open subsets of $\mathbb{R}^n$ is called the **topology** of $\mathbb{R}^n$.

A topological space is a set S equipped with a topology on it.

We refer to the pair ($S, \urcorner$) as a **topological space.**

Definition 3 : Suppose $(S, \urcorner)$ is a topological space and $A \subseteq S$.

Let $\urcorner' = \{A \cap O$ such that $O \in \urcorner\}$.Then $\{A, \urcorner'\}$ is called the topology of A derived from $(S, \urcorner)$ (or **the relative topology**).

Definition 4 : A manifold M of dimension n , or n**-manifold,** is a topological space with the properties:

i) M is **Hausdorff.**

ii) M is **locally Euclidean** of dimension n.

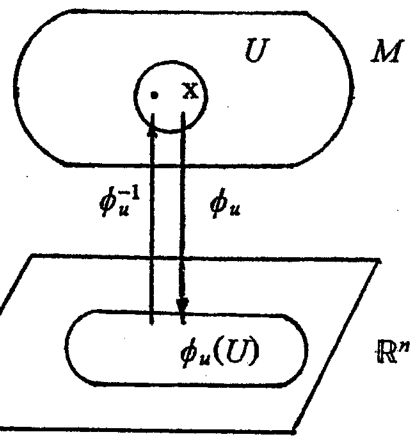
iii) M has a **countable basis** of open sets.

- M is a Hausdorff space if for any distinct points x,y $\in M$ such that x $\neq$y there exist U,V $\in \urcorner$ such that x $\in$ U, y $\in$ V and U $\cap$V $= \varnothing$.

- Each point p has a neighborhood U homeomorphic to an n-ball in $\mathbb{R}^n$. (Example: a manifold of dim 1 is locally homeomorphic to an open interval , a manifold of dim 2 is locally homeomorphic to an open disk , etc.)

So $\forall\, x \in M$ $\exists$ U$_x \in \urcorner$ such that x$\in$ U and U is homeomorphic to a subset of $\mathbb{R}^n$; that is, $\exists\, \phi_u : U \to \phi_u(U) \subseteq \mathbb{R}^n$ such that ϕ_u is one to one and continuous with continuous inverse , ϕ^{-1}.

Figure 1.



Let $(X, \mathcal{T})$ be given, and $Y \subseteq X$.

Define $\mathcal{T}_Y = \{U_Y : U_Y = U \cap Y, \forall U \in \mathcal{T}\}$.

Denote by $(Y, \mathcal{T}_Y)$, the subspace Y with relative topology.

<u>**Lemma 1**</u>**:** If $(X, \mathcal{T})$ is Hausdorff, then $(Y, \mathcal{T}_Y)$ is Hausdorff.

Proof: <u>Show</u> $\forall p, q \in Y \;\; \exists U_Y, V_Y$ such that $p \in U_Y$, $q \in V_Y$, $U_Y \cap V_Y = \varnothing$.

Let $p, q \in Y$. Since $Y \subseteq X$ then $p, q \in X$.

Since $(X, \mathcal{T})$ is Hausdorff then $\exists U, V \in \mathcal{T}$ such that $p \in U, q \in V, U \cap V = \varnothing$.

Then $U \cap Y = U_Y \; (U \in \mathcal{T})$

$V \cap Y = V_Y \; (V \in \mathcal{T})$, and clearly $p \in U_Y, q \in V_Y$.

<u>Show</u> $U_Y \cap V_Y = \varnothing$.

Suppose that $U_Y \cap V_Y \neq \varnothing$. Let $z \in U_Y \cap V_Y$, so $z \in U_Y$ and $z \in V_Y$

then $z \in U \cap Y$ and $z \in V \cap Y$

then $z \in U$ and $z \in Y$, $z \in V$ and $z \in Y$

then $z \in U \cap V$, contradicting $U \cap V = \varnothing$.

So $U_Y \cap V_Y = \varnothing$.

So (Y, τ_Y) is Hausdorff.

<u>Definition 5</u> : Let (S, τ) be a topological space. A collection $\beta \subseteq \tau$ is **a basis** for the topology τ if every open subset in τ is a union of elements of β

$(K \in \tau \Rightarrow \exists \hat{\beta} \subseteq \beta$ such that $\bigcup_{B \in \hat{\beta}} B = K)$.

Let β be a countable basis of (X, τ).

<u>Lemma 2.</u> : If (X, τ) has a countable basis then (Y, τ_Y) has a countable basis.

Define $\beta_Y = \{B \cap Y : B \in \beta\}$

Show β_Y is a basis for (Y, τ_Y).

So we need to show

$\forall U_Y \in \tau_Y \quad \exists \beta'_Y \subseteq \beta_Y$ such that $U_Y = \bigcup_{B_Y \in \beta'_Y} B_Y$

<u>Proof</u>:

Let $U_Y \in \tau_Y$. So $U_Y = U \cap Y$ for some $U \in \tau$. Since $U \in \tau \quad \exists\, \beta' \subseteq \beta$ such that $U = \bigcup_{B \in \beta' \subseteq \beta} B$

Then $U_Y = U \cap Y = \left[\bigcup_{B \in \beta' \subseteq \beta} B \right] \cap Y = \bigcup_{B \in \beta' \subseteq \beta} [B \cap Y]$

For each $B \in \beta'$, $B \cap Y \in \beta_Y$,

so take $\beta'_Y = \{B \cap Y \ : \ B \in \beta'\}$ then $U_Y = \bigcup_{B_Y \in \beta'_Y \subseteq \beta_Y} B_Y$.

So $U_Y = \bigcup B_Y$.

So $\exists\ \beta'_Y \subseteq \beta_Y$ such that $U_Y = \bigcup_{B_Y \in \beta'_Y \subseteq \beta_Y} B_Y$.

So β_Y is a basis for (Y, τ_Y). And since β_Y is a collection of sets which is indexed by β (which is countable), we have that β_Y is countable.

Differentiable functions and Mappings.

Definition 6 : let f be a function on an open set $U \subset \mathbb{R}^n$. We shall say that f is **differentiable** at $a \in U$ if there is a (homogeneous) linear expression $\sum_{i=1}^{n} b_i(x^i - a^i)$ such that the (inhomogeneous) linear function defined by $f(a) + \sum_{i=1}^{n} b_i(x^i - a^i)$ approximates $f(x)$ near a in the following sense :

$$\lim_{x \to a} \frac{f(x) - f(a) - \sum b_i(x^i - a^i)}{\|x - a\|} = 0,$$

or equivalently, if there exist constants $b_1, \ldots, b_n$ and a function $r(x, a)$ defined on a neighborhood V of $a \in U$ which satisfy the following two conditions :

$$f(x) = f(a) + \sum b_i(x^i - a^i) + \|x - a\| r(x, a) \qquad \text{on } V \text{ , and}$$

$$\lim_{x \to a} r(x, a) = 0.$$

If f is differentiable for every $a \in U$, we say it is differentiable on U.

Definition 7 : A mapping $F : U \rightarrow \mathbb{R}^m$, U an open subset of $\mathbb{R}^n$, is **differentiable** at $a \in U$ (or on U) if there exists an $m \times n$ matrix A of constants (respectively, functions on U) and an m-tuple $R(x, a) = (r^1(x, a), \ldots, r^m(x, a))$ of functions defined on U (on $U \times U$) such that $\|R(x, a)\| \rightarrow 0$ as $x \rightarrow a$ and for each $x \in U$ we have

$$F(x) = F(a) + A(x - a) + \|x - a\| R(x, a).$$

A is called the Jacobian matrix.

The Definition of a Differentiable Manifold

Each pair U, φ ,where U is an open subset of M and φ is a homeomorphism of U to an open subset of $\mathbb{R}^n$, is called **a coordinate neighborhood** : to $q \in U$ we assign the n coordinates $x^1(q), \ldots, x^n(q)$ of its image $\varphi(q)$ in $\mathbb{R}^n$ - each $x^i(q)$ is a real-valued function on U, the ith coordinate function.

If q lies also in a second coordinate neighborhood V, ψ, then it has coordinates $y^1(q), \ldots y^n(q)$ in this neighborhood. Since φ and ψ are homeomorphisms, this defines a homeomorphism

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V),$$

the domain and range being the two open subsets of $\mathbb{R}^n$ which correspond to the points of $U \cap V$ by the two coordinate maps φ, ψ, respectively. In coordinates , $\psi \circ \varphi^{-1}$ is given by continuous functions

$$y^i = h^i(x^1, \ldots\ldots, x^n)\ , i = 1, \ldots\ldots n\ ,$$

giving the y-coordinates of each $q \in U \cap V$ in terms of its x –coordinates.

Similarly $\varphi \circ \psi^{-1}$ gives the inverse mapping which expresses the x –coordinates as functions of the y –coordinates

$$x^i = g^i(y^1, \ldots\ldots, y^n)\ , i = 1, \ldots\ldots n\ .$$

The fact that $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are homeomorphisms and are inverse to each other is equivalent to the continuity of $h^i(x)$ and $g^j(y)$, $i,j = 1, \ldots\ldots n$ together with the identities

$$h^i(g^1(y), \ldots\ldots, g^n(y)) \equiv y^i\ , i = 1, \ldots\ldots n\ ,$$

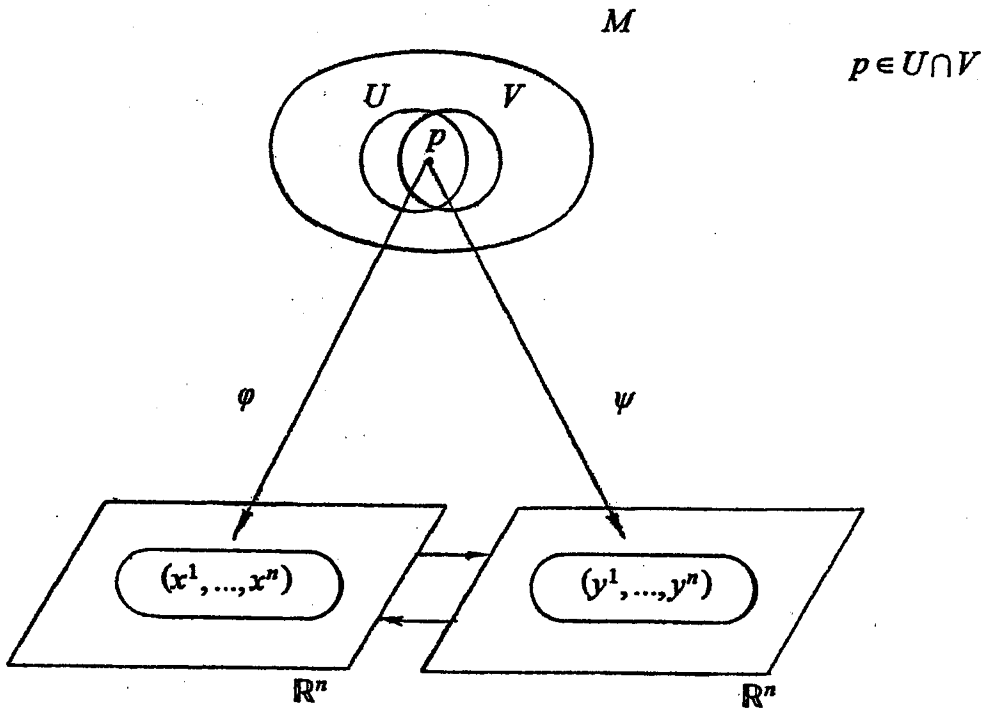
and

$$g^j(h^1(x), \ldots\ldots, h^n(x)) \equiv x^j\ , j = 1, \ldots\ldots n\ .$$

These two mappings $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are called **transition functions**.

Thus every point of a n –manifold M lies in a very large collection of coordinate neighborhoods, but whenever two neighborhoods overlap we have the formulas just given for a change of coordinates. The basic idea that leads to differentiable manifolds is to try to select a family or subcollection of neighborhoods so that the change of coordinates h^i and g^j are always given by differentiable functions.

Figure 2.



$\varphi : U \to \varphi(U) \subseteq \mathbb{R}^n$

$\varphi(p) = (x^1, ...x^n)$

$\psi : V \to \psi(V) \subseteq \mathbb{R}^n$

$\psi(p) = (y^1, ...y^n)$

φ and ψ are homeomorphisms

Definition 8 : U, φ and V, ψ are C^∞ **–compatible** if $U \cap V \neq \varnothing$ implies that the change of coordinates is always given by C^∞ functions; this is equivalent to requiring $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ to be differentiable from $\psi(U \cap V)$ to $\varphi(U \cap V)$ in $\mathbb{R}^n$ and $\varphi(U \cap V)$ to $\psi(U \cap V)$ in $\mathbb{R}^n$, respectively.

Definition 9 : **A differentiable** or C^∞ **structure** on a topological manifold M is a family $\wp = \{U_a, \varphi_a\}$ of coordinates neighborhoods such that:

1) the U_a cover M ($M = \cup U_a$),

2) for any α, β the neighborhoods U_a, φ_a and U_β, φ_β are C^∞-compatible,

3) any coordinate neighborhood V, ψ compatible with every $U_a, \varphi_a \in \wp$ is itself in $\wp$.

A C^∞ manifold is a topological manifold together with a C^∞ –differentiable structure.

Theorem 1. Let M be a Hausdorff space with a countable basis of open sets. If $V = \{V_\beta, \psi_\beta\}$ is a covering of M by C^∞-compatible coordinate neighborhoods, then there is a unique C^∞ structure on M containing these coordinate neighborhoods.

The reason this Theorem is important is that using the Theorem, we only need to produce a specific covering of M (a Hausdorff space with countble basis of an open sets) which consists of C^∞- compatible coordinate neighborhoods. Then all 3 conditions of the definition of C^∞ structure will be satisfied. In particular the Theorem gives an alternate way establishing condition 3, which generally would be too difficult to verify.

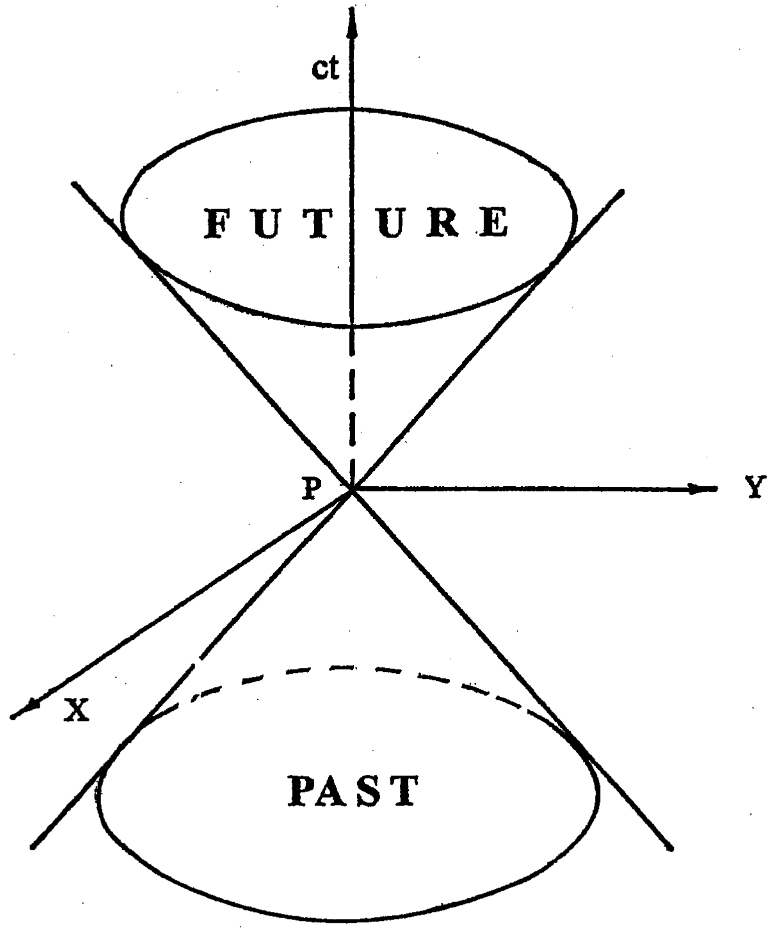
The following is an example of differentiable manifold and we will show that all properties of a definition C^∞ structure are satisfied. Consider a sphere S^2 in $\mathbb{R}^3$.

We will now discuss how we can think of S^2 as a cross-section of what we will call a **light cone** at a point in 4-dimention spacetime by which we will mean a set of points of a form (t, x, y, z), where the concept of distance will be replaced by what we will call an **interval.**

These concepts will be further elucidated as this thesis progresses.

First consider a three-dimensional coordinate system. Consider point P in a spacetime as being on earth.

Figure 3.



Suppose you see a star which is very far away. That means when you see the light from that star you see the light that was emitted several years ago. You are seeing this from the past. The light is coming in from the past. Future is where you are headed, so after you leave that point you imagine yourself moving along the ct-ray. You are still in the same point but you are moving in a sense of advancing time. You are moving to a future, light moves on the cone. That way is called **a light cone**. The cross-section of that light cone is a circle. The entire cone could be generated by taking a single ray and rotating it around the ct-axis, so as to form a circular cross-section. So the way of describing a direction that light can travel in a three-dimensional system is in any one of the directions on that circle.

But in reality we are in a 4-dimensional system, where the light-cone still exists but its cross-section is a sphere. So the direction the light can travel from the past or into a future is a sphere.

It is important to anybody studying the universe to be able to describe the various paths that light might travel. Since light is going to play such a major role in study of the universe, it would be nice to have a coordinate system that somehow incorporates this traveling light as part of a coordinate system. It becomes important to be able to coordinatize a sphere, because a sphere is representative of how light travels. There are a lot of ways to put the coordinates on a sphere. We are going to use a spherical coordinate system. We are going to use the xy-plane as a basis of our coordinate system where instead of thinking of points in the xy-plane as being labeled with a pair (x,y), you think of it as being labeled with a single complex number $a+bi$, where $i^2=-1$. Making a coordinate system based on a complex number allows easier study on that sphere.

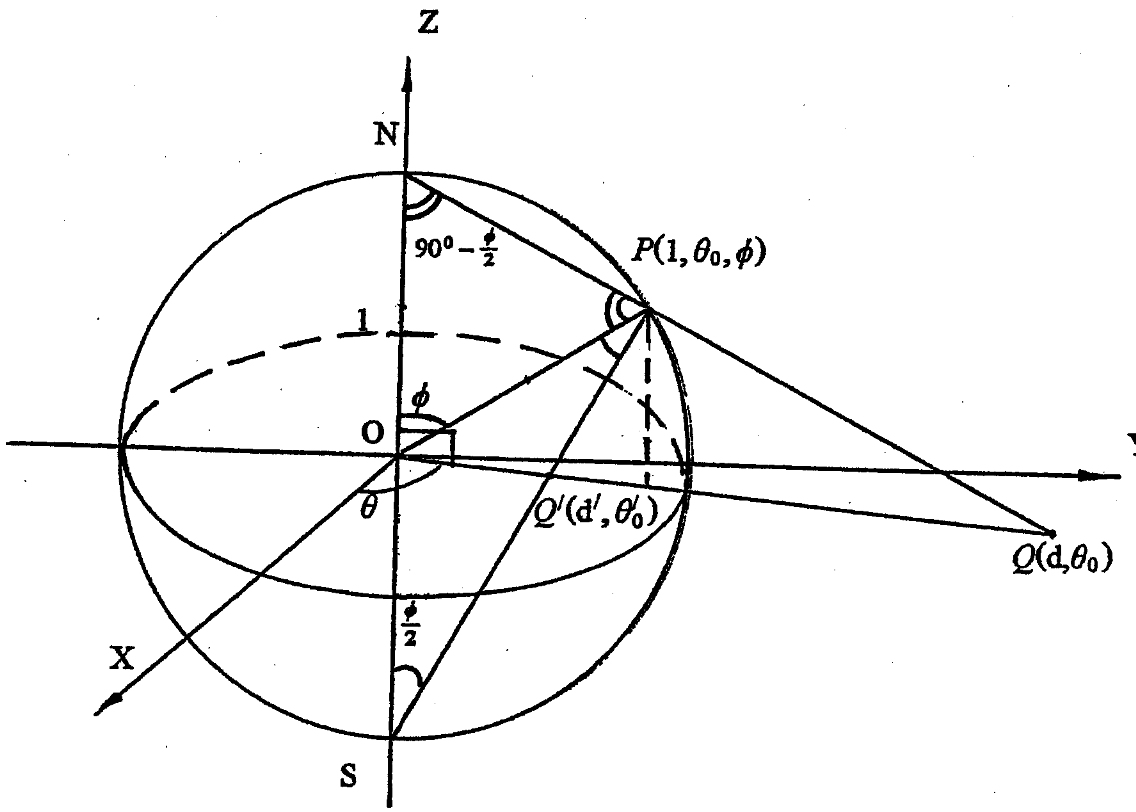
So $P(\theta,\phi) \rightarrow Z \in \mathbb{C}$, where $Z = x + iy$, $i^2 = -1$

$$= d(\cos\theta + i\sin\theta) = d\,e^{i\theta} \text{ (exponential form).}$$

Define N and S to be the north pole and the south pole respectively .

We will show that all properties of the definition C^∞structure ar satisfied.

Figure 4.



1) Consider $U_1 = S^2 \setminus \{N\}$

$$U_2 = S^2 \setminus \{S\}$$

$S^2 = U_1 \cup U_2$.

So the first condition of the definition is satisfied.

2) Using stereographic projection from the north pole N determine a coordinate neighborhood U_1, ψ_1. In the same way determine by projection from the south pole S a neighborhood U_2, ψ_2. We need to show that these two neighborhoods determine a C^∞ structure on S^2.

Note $U_1 \cap U_2 \neq \varnothing$.

Let $p \in U_1 \cap U_2$ with θ–coordinate θ_0.

We will consider the plane $\theta = \theta_0$ and the geometry of ray $\overrightarrow{Np}$ in this plane.

$p = (1, \theta_0, \phi)$.

$\overrightarrow{Np} \cap (xy-plane) = Q(d, \theta_0) = \psi_1(p)$,

$\overrightarrow{Sp} \cap (xy-plane) = Q'(d', \theta'_0) = \psi_2(p)$

$\left(\psi_1 : U_1 \to \mathbb{R}^2 \quad , \quad \psi_2 : U_2 \to \mathbb{R}^2\right)$

We are going to look for some relationships between (d', θ') and (d, θ). So we try to find transition function such that $(d', \theta') = F[(d, \theta)]$.

First we need to find d and d'.

Consider $\triangle NOP$: $\quad ON = OP = 1 \Rightarrow \angle NOP = \angle NPO$, so $2p = 180^0 - \phi \Rightarrow p = 90^0 - \frac{\phi}{2}$

$\triangle NOQ$: $\angle NOQ = 90^0$, $ON = 1$, $\angle ONQ = 90^0 - \frac{\phi}{2}$

$$\tan\left(90^0 - \tfrac{\phi}{2}\right) = \tfrac{OQ}{ON} = \tfrac{d}{1} \quad \Rightarrow \quad d = \tan\left(90^0 - \tfrac{\phi}{2}\right) = \cot\tfrac{\phi}{2}$$

Now consider $\triangle SOP$: $\quad OB = OP = 1$,

$$\angle OSP = \angle OPS = \tfrac{180^0 - (180^0 - \phi)}{2} = \tfrac{\phi}{2}$$

$$\triangle SOQ' : \quad \tan\tfrac{\phi}{2} = \tfrac{OQ'}{1} \quad \Rightarrow OQ' = \tan\tfrac{\phi}{2} \quad \Rightarrow d' = \tan\tfrac{\phi}{2}$$

So we have shown that $d = \cot\frac{\phi}{2}$ and $d' = \tan\frac{\phi}{2}$,

we can see that $d' = \tan\frac{\phi}{2} = \frac{1}{d}$.

d and $d' = \frac{1}{d}$ are functions of two variables and they are differentiable functions.

So F is C^∞.

Thus the coordinate neighborhoods U_1, ψ_1 and U_2, ψ_2 are C^∞ compatible.

3) property number 3 can be checked by using the Theorem : we have a covering by C^∞ compatible neighborhoods U_1 and U_2, and S^2 is Hausdorff and has a countable basis

(by Lemma 1 and Lemma 2), therefore there is a unique C^∞ structure on S^2 .

So we can say that S^2 is a differentiable manifold.

Diffeomorphism.

Let f be a real-valued function defined on an open set W_f of a C^∞ manifold M.

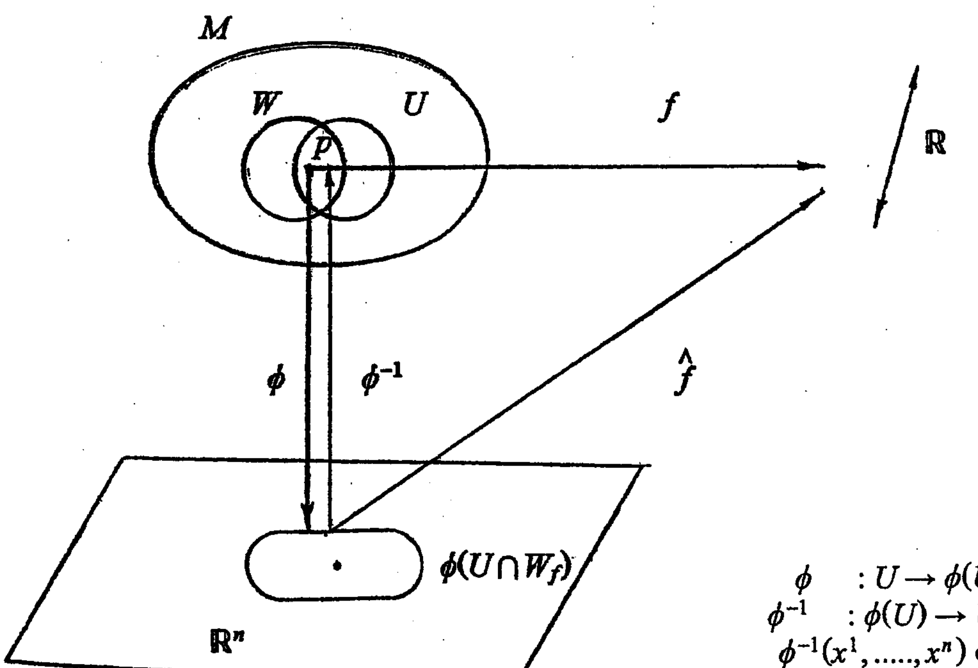
$$f: W_f \to \mathbb{R}$$

U, ϕ is a coordinate neighborhood such that $W_f \cap U \neq \varnothing$, and if $x^1, \ldots\ldots, x^n$ denotes the local coordinates , then f corresponds to a function $\hat{f}(x^1, \ldots\ldots, x^n)$ on $\phi(W_f \cap U)$ defined by $\hat{f} = f \circ \phi^{-1}$, that is , so that $f(p) = \hat{f}(x^1(p), \ldots\ldots, x^n(p)) = \hat{f}(\phi(p))$.

Definition 10 : $f: W_f \to \mathbb{R}$ is a C^∞ function if each $p \in W_f$ lies in a coordinate neighborhood U, ϕ such that $f \circ \phi^{-1}(x^1, \ldots\ldots, x^n) = \hat{f}(x^1, \ldots\ldots, x^n)$ is C^∞ on $\phi(W_f \cap U)$.

[Clearly, a C^∞ function is continuous.]

Figure 5.



$\phi \quad : U \to \phi(U) \subseteq \mathbb{R}^n$

$\phi^{-1} \quad : \phi(U) \to U$

$\phi^{-1}(x^1, \ldots\ldots, x^n) \in U$

$f \quad : W_f \to \mathbb{R}$

$f[\phi^{-1}(x^1, \ldots\ldots, x^n)] \in \mathbb{R}$

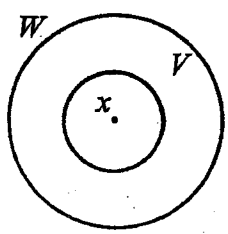
$$\hat{f} = (f \circ \phi^{-1})(x^1, \ldots\ldots, x^n) \in \mathbb{R}$$

$\hat{f}(x^1, \ldots\ldots, x^n) \in \mathbb{R}$ such that $f \circ \phi^{-1} : \phi(U \cap W_f) \to \mathbb{R}$

$$\hat{f} : \phi\Big(\underset{U \subseteq M}{U}\Big) \to \mathbb{R} \text{ is differentiable.}$$

It is a consequence of the definition that if f is C^∞ on W and $V \subset W$ is an open set, then $f \mid V$ is C^∞ on V. Moreover, if W is a union of open sets on each of which a real-valued function f is C^∞, then f is C^∞ on W.

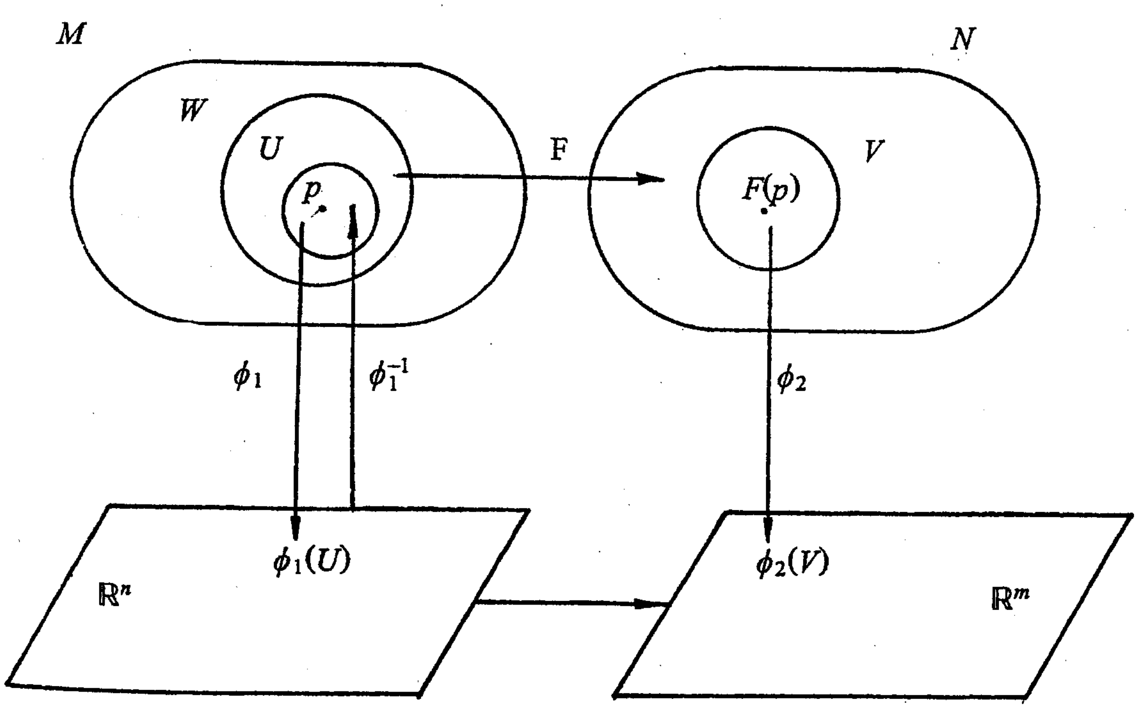
Figure 6.



$$f : W \to \mathbb{R}$$
$$f \mid V : V \to \mathbb{R},\ V \subset W$$
$$(f \mid V)(x) = f(x) \quad \forall x \in V$$

Suppose that M and N are C^∞ manifolds, $W \subset M$ is an open subset and $F : W \to N$ is a mapping, then we have the following definition.

Figure 7.



Definition 11 : F is a C^∞ mapping of W into N if for every $p \in W$ there exist coordinate neighborhoods U, ϕ_1 of p and V, ϕ_2 of $F(p)$ with

$F(U) \subset V$ such that $\phi_2 \circ F \circ \phi_1^{-1} : \phi_1(U) \to \phi_2(V)$ is C^∞.

Definition 12 : A C^∞ mapping $F : M \to N$ between C^∞ manifolds is a **diffeomorphism** if it is a homeomorphism and F^{-1} is C^∞. M and N are diffeomorphic if there exists a diffeomorphism $F : M \to N$.

For example, the transition functions $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ which were discussed in the section, the definition of a differentiable manifold, are diffeomorphisms of open subsets in $\mathbb{R}^n$.

Dual Vector Space.

Definition 13 : Let V be a finite dimensional vector space over F. The **dual vector space** V^* of V is defined to be the vector space of linear transformations from ($L(V,F)$) where F is identified with the vector space overitself. The elements of V^* are simply functions f from V into F such that $f(v_1 + v_2) = f(v_1) + f(v_2)$ $\forall v_1, v_2 \in V$ and $f(av) = af(v)$, $a \in F$, $v \in V$. Elements of V^* are called linear functions on V.

Lemma 3 : Let $\{v_1, \ldots\ldots, v_n\}$ be a basis for V over F. Then there exist linear functions $\{f_1, \ldots\ldots, f_n\}$ such that for each i, $f_i(v_i) = 1$, $f_i(v_j) = 0$, $j \neq i$.

The linear functions $\{f_1, \ldots\ldots, f_n\}$ form a basis for V^* over F, called the dual basis to $\{v_1, \ldots\ldots, v_n\}$.

Proof. See [3].

Tangent space.

We begin with a discussion of the tangent space at a point a of $\mathbb{R}^n$.

Let us denote by $C^\infty(a)$ the collection of all C^∞ functions whose domain includes a, since we are only interested in their derivatives at a. Let $X_a = \sum_{i=1}^{n} a^i E_{ia}$ be an expression for a vector of $T_a(\mathbb{R}^n)$ in the canonical basis; we define the **directional derivative** Δf of f at a in the "direction of X_a" by $\Delta f = \sum_{i=1}^{n} a^i \frac{\partial f}{\partial x^i}$ evaluated at $a = (a^1, \ldots, a^n)$. This is a slight extension of the usual definition in that we do not require X_a to be a unit vector. Since Δf depends on f, a, and X_a we shall write it as $X_a^* f$. Thus

$$X_a^* f = \sum_{i=1}^{n} a^i \left(\frac{\partial f}{\partial x^i}\right)_a .$$

We may take the directional derivative in the "direction of X_a" of any C^∞ function defined in a neighborhood of a. Hence $f \to X_a^* f$ defines a mapping assigning to each $f \in C^\infty(a)$ a real number; $X_a^* : C^\infty(a) \to \mathbb{R}$.

It is reasonable to denote this mapping by $X_a^* = \sum_{i=1}^{n} a^i \left(\frac{\partial}{\partial x^i}\right)$, where we must remember that the derivatives are to be evaluated at a. We remark that $X_a^* \pi^i = a^i$, $i = 1, \ldots, n$.

Indeed, $X_a^* \pi^i = \left[\sum_{j=1}^{n} a^j \frac{\partial}{\partial x^j} \mid_a \right] \pi^i = \sum_{j=1}^{n} a^j \frac{\partial \pi^i}{\partial x^j} \mid_a = \sum_{j=1}^{n} a^j \delta_j^i = a^i$. Since X_a is completely determined by the a^i, we now see that X_a is determined by what it does to each of the coordinate functions π^i, $1 \le i \le n$. In other words, the vectors which comprise $T_a(\mathbb{R}^n)$ are defined by the above discussion.

Now we will define the **tangent space** $T_p(M)$ to a more general manifold, M, at p to be the set of all mappings $X_p : C^\infty(p) \to \mathbb{R}$ satisfying for

$\forall a, \beta \in \mathbb{R}$, $f, g \in C^{\infty}(p)$ the two conditions

i) $X_p(af + \beta g) = a(X_p f) + \beta(X_p g)$ (linearity)

ii) $X_p(fg) = (X_p f)g(p) + f(p)(X_p g)$ (Leibniz rule)

with the vector space operations in $T_p(M)$ defined by

$$(X_p + Y_p)f = X_p f + Y_p f$$

$$(aX_p)f = a(X_p f)$$

A tangent vector to M at p is any $X_p \in T_p(M)$.

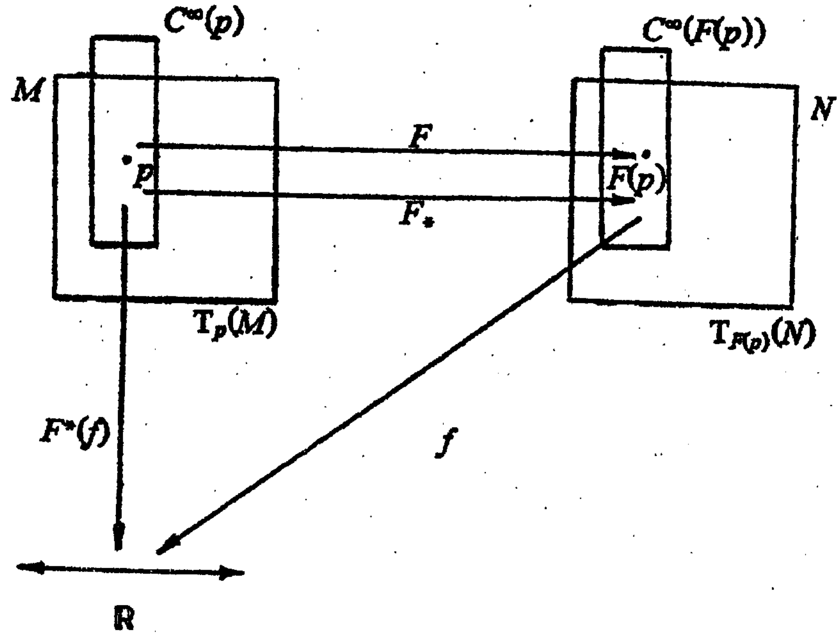
We see that $T_p(M)$ is a vector space over $\mathbb{R}$ for if

$X_{1p}, X_{2p} : C^{\infty}(p) \to \mathbb{R}$ and $a, \beta \in \mathbb{R}$, then we define

$(aX_{1p} + \beta X_{2p})f = a(X_{1p}f) + \beta(X_{2p}f)$, where the operations on the right are in $\mathbb{R}$. This defines in $T_p(M)$ both vector addition and multiplication by real numbers a, β.

Theorem 2. Let $F : M \to N$ be a C^{∞} map of manifolds. Then for $p \in M$ the map $F^* : C^{\infty}(F(p)) \to C^{\infty}(p)$ defined by $F^*(f) = f \circ F$ is a homomorphism (linear transformation) of algebras and induces a dual vector space homomorphism $F_* : T_p(M) \to T_{F(p)}(N)$, defined by $F_*(X_p)f = X_p(F^* f)$, which gives $F_*(X_p)$ as a map of $C^{\infty}(F(p))$ to $\mathbb{R}$.

Figure 8.



Define $F_* : T_p(M) \to T_{F(p)}(N)$ by $F_*(X_p) = X_{F(p)}$.

$$X_p : C^\infty(p) \to \mathbb{R},$$

$$X_{F(p)} : C^\infty(F(p)) \to \mathbb{R}.$$

What does $F_*(X_p)$ do to $f \in C^\infty(F(p))$?

$$[F_*(X_p)](f) = X_{F(p)} f = X_p(f \circ F) = [X_p](F^*(f)).$$

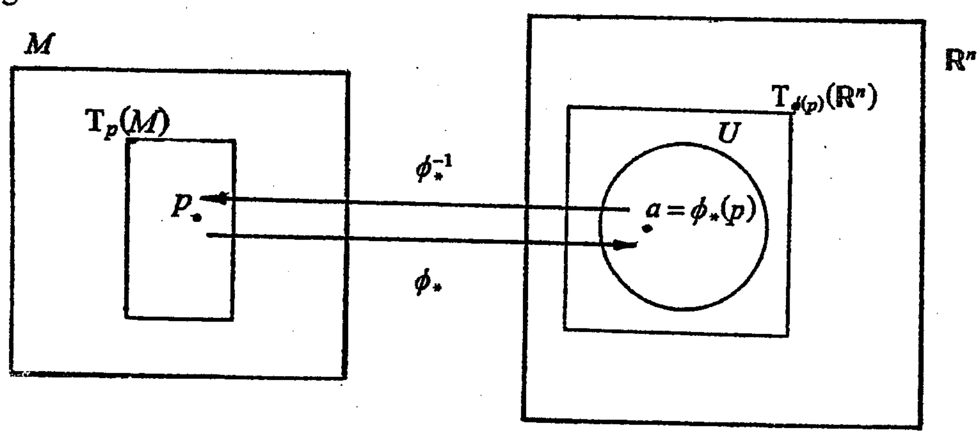
__Corollary 1.__ If $F : M \to N$ is a diffeomorphism of M onto an open set $U \subset N$ and $p \in M$, then $F_* : T_p(M) \to T_{F(p)}(N)$ is an isomorphism.

Remembering that any open subset of a manifold is a (sub)manifold of the same dimension, we see that if U, ϕ is a coordinate neighborhood on M, then the coordinate map ϕ induces an isomorphism $\phi_* : T_p(M) \to T_{\phi(p)}(\mathbb{R}^n)$ of the tangent space at each point $p \in U$ onto $T_a(\mathbb{R}^n)$, $a = \phi(p)$. The map ϕ_*^{-1} on the other hand, maps $T_a(\mathbb{R}^n)$ isomorphically onto $T_p(M)$. ϕ_*^{-1} is a linear transformation which is one-to-one and onto, so it takes a basis for $T_a(\mathbb{R}^n)$ into a basis for $T_p(M)$. A basis for $T_p(M)$ is this $\phi_*^{-1}\left(\frac{\partial}{\partial x^i}\right)$.

The images $E_{ip} = \phi_*^{-1}\left(\frac{\partial}{\partial x^i}\right), i = 1, \ldots, n$ of the basis $\frac{\partial}{\partial x^1}, \ldots\ldots, \frac{\partial}{\partial x^n}$ at each $a \in \phi(U) \subset \mathbb{R}^n$ determine at $p = \phi^{-1}(a) \in M$ a basis $E_{1p}, \ldots\ldots, E_{np}$ of $T_p(M)$; we call these bases the coordinate frames.

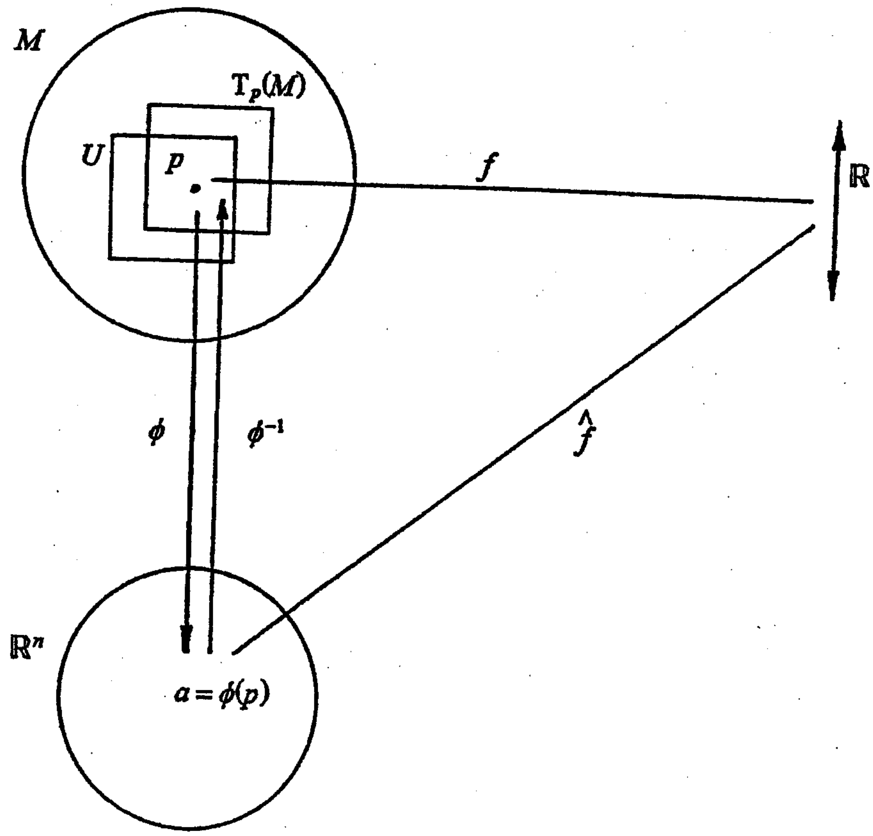
When we do calculus on M we can essentially treat it as though we are doing calculus in $\mathbb{R}^n$ (locally M is $\mathbb{R}^n$). So when we do calculus on a manifold it is often customery to drop the notation of the ϕ_*^{-1} as though $\frac{\partial}{\partial x^i}$ form a basis for the tangent space at a point of a manifold.

Figure 9.



__Corollary 2.__ To each coordinate neighborhood U on M there corresponds a natural basis $E_{1p}, \ldots\ldots, E_{np}$ of $T_p(M)$ for every $p \in U$; in particular, $\dim T_p(M) = \dim M$. Let f be a C^∞ function defined in a neighborhood of p, and $\hat{f} = f \circ \phi^{-1}$ its expression in local coordinates relative to U, ϕ. Then $E_{ip}f = \left(\frac{\partial \hat{f}}{\partial x^i}\right)_{\phi(p)}$.

Figure 10.



Tangent Covectors.

Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and let V^* denote its dual space. Then V^* is the space whose elements are linear functions from V to $\mathbb{R}$, and we call them **covectors.**

If $\sigma \in V^*$, then $\sigma : V \to \mathbb{R}$ and for $\forall v \in V$ we denote the value of σ on v by $\sigma(v)$ or by $\langle v, \sigma \rangle$.

The vector addition and multiplication by scalars in V^* are defined by the equations :

$$(\sigma_1 + \sigma_2)(v) = \sigma_1(v) + \sigma_2(v)$$

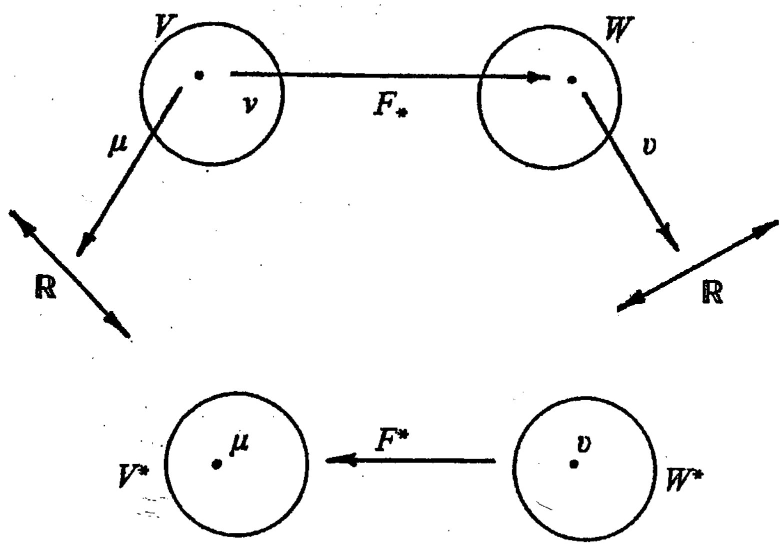
$$(a\sigma)(v) = a(\sigma(v)),$$

giving the values of $\sigma_1 + \sigma_2$ and $a\sigma$, $a \in \mathbb{R}$, on an arbitrary $v \in V$, the right hand operations taking place in $\mathbb{R}$.

1) If $F_* : V \to W$ is a linear map of vector spaces, then it uniquely determines a dual linear map $F^* : W^* \to V^*$ by $(F^*\sigma)(v) = \sigma(F_*(v))$ or $\langle v, F^*(\sigma)\rangle = \langle F_*(v), \sigma\rangle$.

When F_* is injective (surjective), then F^* is surjective (injective).

Figure 11.



$F_*(v) \in W$

V and V^*, W and W^* are dual of each other

$F^* : W^* \to V^*$ if $\underset{\upsilon : W \to \mathbb{R}}{\upsilon} \in W^*$, then $F^*(\upsilon) \in V^*$ (so there exists an element μ of V^* such that $\mu = F^*(\upsilon) : V \to \mathbb{R}$)

$v \in V$

is given by $F^*(\upsilon)(v) = \upsilon(F_*(v))$.

2) If $e_1, \ldots\ldots, e_n$ is a basis of V, then there exists a unique dual basis $w^1, \ldots\ldots, w^n$ of V^*

such that $w^i(e_j) = \delta^i_j$ $\begin{pmatrix} \delta^i_j = 0 & \text{if} & i \neq j \\ \delta^i_j = 1 & \text{if} & i = j \end{pmatrix}$

(each element of the basis is a linear function on V)

If $v \in V$, then $w^1(v), \ldots\ldots, w^n(v)$ are exactly the components of v with respect to the basis

$e_1, \ldots\ldots, e_n$.

In other words $v = \sum\limits_{j=1}^{n} w^j(v)e_j$.

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis for V. Then $v = \sum\limits_{j=1}^{n} a_j e_j$.

Consider $w^{i_0}(v) = w^{i_0}\left(\sum\limits_{j=1}^{n} a_j e_j\right) = \sum\limits_{j=1}^{n} a_j w^{i_0}(e_j) = \sum\limits_{j=1}^{n} a_j \delta_j^{i_0} = a_1\delta_1^{i_0} + \ldots + a_n\delta_n^{i_0} = a_{i_0}$, where

$i_0 = 1, \ldots, n$; and since i_0 is a dummy letter, we can replace it by j.

So $v = \sum\limits_{j=1}^{n} w^j(v)e_j$.

3) There is a natural isomorphism of V onto $(V^*)^*$ given by $v \to \langle v, . \rangle$; that is, v is mapped to the linear function on V^* whose value on any $\sigma \in V^*$ is $\langle v, \sigma \rangle$. Note that $\langle v, \sigma \rangle$ is linear in each variable separately (with the other fixed).

Covectors on Manifolds.

Let M be a C^∞ manifold , $p \in M$. $T_p^*(M)$ is the dual space to $T_p(M)$; thus $\sigma_p \in T_p^*(M)$ is a linear mapping $\sigma_p : T_p(M) \to \mathbb{R}$ and its value on $X_p \in T_p(M)$ is denoted by

$\sigma_p(X_p)$ or $\langle X_p, \sigma_p \rangle$.

Given a basis $E_{1p}, \ldots\ldots E_{np}$ of $T_p(M)$, there is a uniquely determined dual basis $w_p^1, \ldots\ldots, w_p^n$ satisfying by definition , $w_p^i(E_{jp}) = \delta_j^i$. The components of σ_p relative to this basis w_p^i are

equal to the values σ_p on the basis vectors $E_{1p}, \ldots\ldots E_{np}$, that is

$\sigma_p = \sum \sigma_p(E_{ip})w_p^i$, $i = 1, \ldots, n$. And now we are going to prove it.

Proof. Let $\sigma_p \in T_p^*(M)$. Since $\{w^1, \ldots, w^n\}$ is a basis of $T_p^*(M)$, $\sigma_p = \sum a_i w_p^i$, $a_i \in \mathbb{R}$

$\sigma_p(E_{jp}) = \left(\sum a_i w_p^i\right)(E_{jp})$

$$w_p^i(\mathrm{E}_{jp}) = \delta_j^i = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

then $\sum_{j=1}^{n}\left[a_i w_p^i(\mathrm{E}_{jp})\right] = a_j.$

so $\sigma_p(\mathrm{E}_{jp}) = a_j \quad , 1 \leq j \leq n$.

Then $\sigma_p = \sum \sigma_p(\mathrm{E}_{ip}) w_p^i \quad , \; i = 1, ..., n.$

T_p^* is a set of linear mappings from T_p to $\mathbb{R}$, and we can view elements of T_p as linear mappings from T_p^* to $\mathbb{R}$.

From the space T_p of vectors at p and the space T_p^* which consists of elements we call one-forms at p , we can form the **Cartesian product**,

$$\prod{}_r^s = \underbrace{\mathrm{T}_p^* \times ..... \times \mathrm{T}_p^*}_{r \text{ factors}} \times \underbrace{\mathrm{T}_p \times ..... \times \mathrm{T}_p}_{s \text{ factors}} \; ,$$

i.e. the ordered set of one-forms and vectors $(\eta^1, ..., \eta^r, \mathrm{Y}_1, ...\mathrm{Y}_s)$ where the Y's and η's are arbitrary vectors and one-forms respectively.

Example $\prod{}_1^1 = \mathrm{T}_p^* \times \mathrm{T}_p$

A tensor of type (r, s) at p is a function on $\prod{}_r^s$ which is linear in each argument .

If T is a tensor of type (r, s) at p , we write the real number into which T maps the element $(\eta^1, ..., \eta^r, \mathrm{Y}_1, ...\mathrm{Y}_s)$ of $\prod{}_r^s$ as $\mathrm{T}(\eta^1, ..., \eta^r, \mathrm{Y}_1, ...\mathrm{Y}_s)$. So,

$$\mathrm{T}: \prod{}_r^s \rightarrow \mathbb{R}$$

(we write $\mathrm{T}(\eta, \mathrm{Y})$ when $r = 1, s = 1$).

For example , for $\forall a, b \in \mathbb{R}, \quad \mu, \upsilon \in T_p^*$:

$$\mathrm{T}(a\mu + b\upsilon, \mathrm{Y}_1, ..., \mathrm{Y}_s) = a\mathrm{T}(\mu, \mathrm{Y}_1, ..., \mathrm{Y}_s) + b\mathrm{T}(\upsilon, \mathrm{Y}_1, ..., \mathrm{Y}_s)$$

and for $\forall a, b \in \mathbb{R}, \quad \mathrm{X,Y} \in T_p$:

$$T(\eta^1,...\eta^r, aX+bY, Y_2,...Y_s) = aT(\eta^1,...,\eta^r, X, Y_2,...Y_s) + bT(\eta^1,...\eta^r, Y, Y_2,..., Y_s)$$

The space of all such tensors is called the **tensor product**

$$T^r_s(p) = \underbrace{T_p \otimes ..... \otimes T_p}_{r \text{ factors}} \otimes \underbrace{T^*_p \otimes ..... \otimes T^*_p}_{s \text{ factors}}.$$

In particular ,

$$T^1_0(p) = T_p,$$

$$T^0_1(p) = T^*_p.$$

Addition of tensors of type (r,s) is defined by the rule :

$(T+T')$ is the tensor of type (r,s) at p such that for $\forall Y_i \in T_p, \eta^j \in T^*_p$:

$(T+T')(\eta^1,...,\eta^r, Y_1,..., Y_s) = T(\eta^1,...,\eta^r, Y_1,..., Y_s) + T'(\eta^1,...,\eta^r, Y_1,..., Y_s)$

Similarly , multiplication of a tensor by a scalar $a \in \mathbb{R}$ is defined by the rule:

(aT) is the tensor such that $\forall Y_i \in T_p \quad ,\eta^j \in T^*_p$:

$(aT)(\eta^1,...,\eta^r, Y_1,..., Y_s) = aT(\eta^1,...,\eta^r, Y_1,..., Y_s)$

With these rules of addition and scalar multiplication, the tensor product space $T^r_s(p)$ is a vector space of dimension n^{r+s} over $\mathbb{R}$, since each factor of $T_p(M)$ and $T^*_p(M)$ is of dimension n.

Let $X_i \in T_p$ $(i = 1,...,r)$, $\omega^j \in T^*_p$, $(j = 1,...,s)$. We denote by $X_1 \otimes ..... \otimes X_r \otimes \omega^1 \otimes ..... \otimes \omega^s$ that element of $T^r_s(p)$ which maps the element $(\eta^1,...,\eta^r, Y_1,..., Y_s)$ of $\prod^s_r$ into

$$\langle \eta^1, X_1 \rangle \cdots \langle \eta^r, X_r \rangle \langle \omega^1, Y_1 \rangle \cdots \langle \omega^s, Y_s \rangle \in \mathbb{R}.$$

Let $\omega \in T^*_p(M)$, and let $\{dx^1,......, dx^n\}$ be a basis for $T^*_p(M)$

$$dx^i(\tfrac{\partial}{\partial x^j}) \mid_p = \delta^i_j , \quad dx^i : T_p(M) \to \mathbb{R} , \ 1 \le i \le n.$$

So $\omega = \sum_{i=1}^{n} \omega_i \, dx^i$, where $\omega_i = \omega(E_i)$.

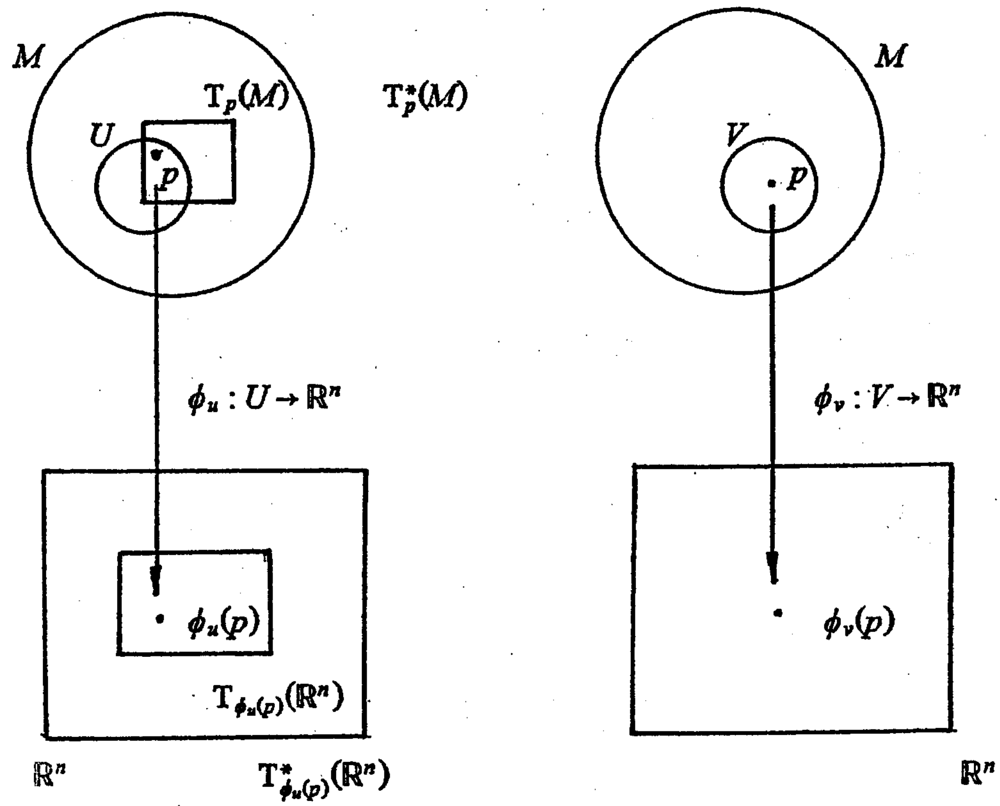
We are going to introduce the Einstein summation convention : $\omega = \omega_i \mathrm{d}x^i \equiv \sum_{i=1}^{n} \omega_i \mathrm{d}x^i$.

Now, suppose we have a different basis for $\mathrm{T}_p^*(M)$: $\{\mathrm{d}x^{1'}, ..., \mathrm{d}x^{n'}\}$; then $\omega = \omega_{i'} \mathrm{d}x^{i'}$

(like two different basis from two different neighborhoods U and V)

What is the relationship between $\mathrm{d}x^i$ and $\mathrm{d}x^{i'}$, ω_i and $\omega_{i'}$?

Figure 12.



We know $\phi_u : U \to \phi(U)$, $\phi_u^{-1} : \phi(U) \to U$

By Theorem3 we have $(\phi_u)_* : \mathrm{T}_p(U) \to \mathrm{T}_{\phi(p)}(\phi(U))$

$$(\phi_u^{-1})_* : \mathrm{T}_{\phi(p)}(\phi(U)) \to \mathrm{T}_p(U)$$

Now since $\frac{\partial}{\partial x^j}$ is a basis for $\mathrm{T}_{\phi_u(p)}(\mathbb{R}^n)$ and $\mathrm{T}_{\phi_{u(p)}}(\mathbb{R}^n)$ is isomorphic to $\mathrm{T}_p(M)$, then basis $\frac{\partial}{\partial x^j}$ is a basis for $\mathrm{T}_p(M)$.

So, a basis for $T_p\left(\underset{U\subseteq M}{U}\right)$ is technically given by

$$\left\{(\phi_u^{-1})_* \frac{\partial}{\partial x^1}\mid_{\phi(p)}\ ,\ (\phi_u^{-1})_* \frac{\partial}{\partial x^2}\mid_{\phi(p)}\ ,\ \ldots\ldots,\ (\phi_u^{-1})_* \frac{\partial}{\partial x^n}\mid_{\phi(p)}\ \right\},$$

and technically a basis for $T_p^*(U)$ is

$$\left\{\phi_u^*\,(\mathrm{d}x^1\ \mid_{\phi_u(p)})\ ,\ \ldots\ldots,\ \phi_u^*(\,\mathrm{d}x^n\ \mid_{\phi_u(p)})\right\}.$$

So we say that $\left\{(\phi_u^{-1})_* \frac{\partial}{\partial x^i}\mid_{\phi_u(p)}\right\}$ is a basis for $T_p(U)$

and $\left\{\phi_u^*\,(\mathrm{d}x^i\ \mid_{\phi_u(p)})\right\}$ is a basis for $T_p^*(U)$. ($\phi_u^* : T^*_{\phi_u(p)}(\phi(U)) \rightarrow T_p^*(M)$).

That is, a basis for $T_p(U)$ is identified with and written as $\left\{\frac{\partial}{\partial x^i}\mid_p\right\}, i = 1, \ldots, n$

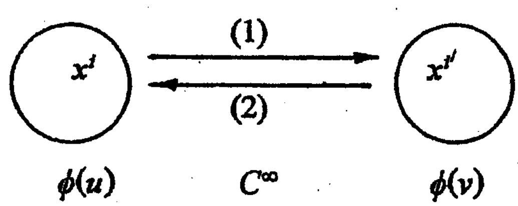
and a basis for $T_p^*(U)$ is identified with and written as $\{\mathrm{d}x^i\mid_p\},\ i = 1, \ldots, n.$

Consider a vector ω which has two different coordinate representations x^i and $x^{i'}$ with respect to old and new coordinates .

Then we have transition functions : $x^{i'} = x^{i'}(x^i)$ (1)

$x^i = x^i(x^{i'})$ (2) , $i, i' = 1, \ldots, n$

Figure 13.



Before we going further let us consider the following: if f is a C^∞ function on M, then we can define $\mathrm{d}f$ by the formula

$$\langle X_p, \mathrm{d}f_p\rangle = X_p f \quad \text{or} \quad \mathrm{d}f_p(X_p) = X_p f.$$

As p varies we obtain df, the differential of p. In the case of an open set $U \subset \mathbb{R}^n$, the coordinates x^i of a point of U are functions on U and, by our definition, $\mathrm{d}x^i$ assigns to each vector X at $p \in U$ a number $X_p x^i$, its ith component in the natural basis of $\mathbb{R}^n$.

In paricular $\langle \frac{\partial}{\partial x^j}, \mathrm{d}x^i \rangle = \frac{\partial x^i}{\partial x^j} = \delta^i_j$ so we see that $\mathrm{d}x^1, ..., \mathrm{d}x^n$ is exactly the field of coframes dual to $\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}$. Now if f is a C^∞ function on U, then we may express $\mathrm{d}f$ as a linear combination of $\mathrm{d}x^1, ..., \mathrm{d}x^n$. We know that the coefficients in this combinations, that is the components of $\mathrm{d}f$, are given by $\mathrm{d}f(\frac{\partial}{\partial x^i}) = \frac{\partial f}{\partial x^i}$. Thus we have

$$\mathrm{d}f = \frac{\partial f}{\partial x^1}\mathrm{d}x^1 + \ldots + \frac{\partial f}{\partial x^n}\mathrm{d}x^n.$$

Now using the above result we have

$$\mathrm{d}x^{i'} = \frac{\partial x^{i'}}{\partial x^1}\mathrm{d}x^1 + \frac{\partial x^{i'}}{\partial x^2}\mathrm{d}x^2 + \ldots.. + \frac{\partial x^{i'}}{\partial x^n}\mathrm{d}x^n$$

So $\mathrm{d}x^{i'} = \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \mathrm{d}x^i$ (recall summation on i) $i, i' = 1, ..., n$

(this $\mathrm{d}x^{i'}$ is expressed in terms of a linear combination of the $\mathrm{d}x^i$).

Now take arbitrary $f \in C^\infty(\phi(p))$ or $C^\infty(\psi(p))$: $f(x^{i'}) = f(x^{i'}(x^i))$. By the Chain rule, we have $\frac{\partial f}{\partial x^i} = \sum_{i=1}^{n} \frac{\partial f}{\partial x^{i'}} \frac{\partial x^{i'}}{\partial x^i}$.

So $\frac{\partial}{\partial x^i} \mid_{\phi(p)} = \sum_{i=1}^{n} \frac{\partial x^{i'}}{\partial x^i} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)} \Rightarrow \frac{\partial}{\partial x^i} \rightarrow \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)}$ (Einstein Summation Convention).

So $F_*\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)}$ (3)

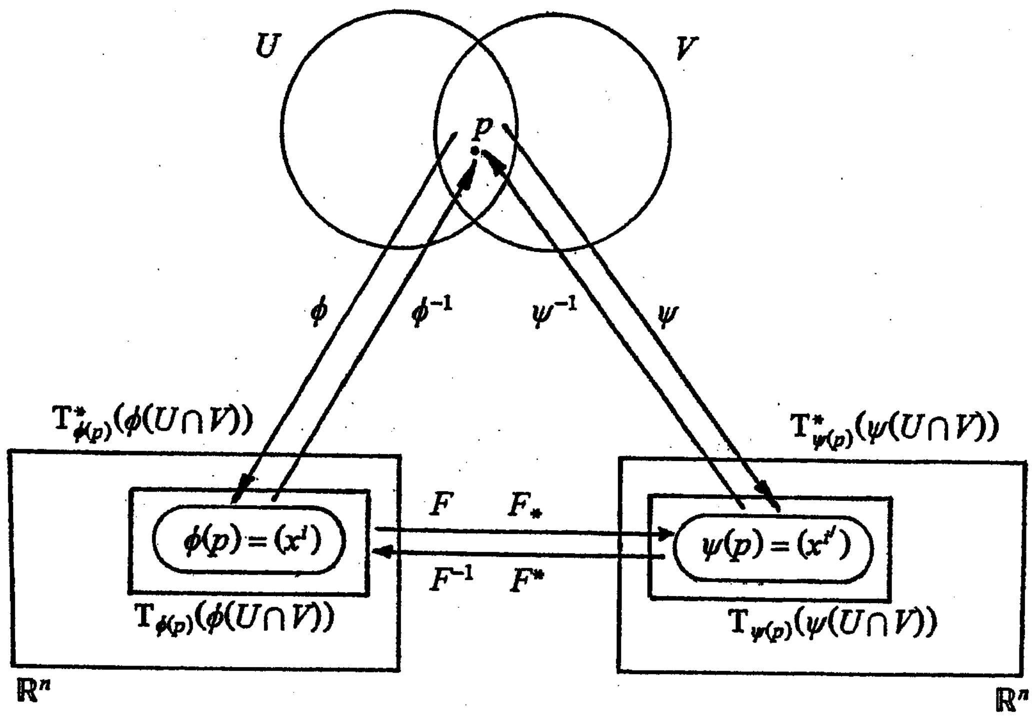
and $F^*(\mathrm{d}x^{i'}) = \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \mathrm{d}x^i$. (4)

Consider a covector $\omega = \omega_i\, \mathrm{d}x^i = \omega'_i\, \mathrm{d}x^{i'}$ (same covector with different basis)

Consider $U \cap V$ and $p \in U \cap V$.

The function ϕ takes p into $\phi(U) \subseteq \mathbb{R}^n$, the function ψ takes p into $\psi(V) \subseteq \mathbb{R}^n$.

Figure 14.



$$F=\psi\circ\phi^{-1} \quad : \quad \phi(U\cap V)\rightarrow\psi(U\cap V),$$

$$F^{-1}=\phi\circ\psi^{-1} : \quad \psi(U\cap V)\rightarrow\phi(U\cap V)$$

By the Theorem2 we can define linear transformations

$$F_* : T_{\phi(p)}(\phi(U\cap V))\rightarrow T_{\psi(p)}(\psi(U\cap V))$$

$$F^* : T^*_{\psi(p)}(\psi(U\cap V))\rightarrow T^*_{\phi(p)}(\phi(U\cap V))$$

We would like to answer several questions :

- what happens to a basis for $T_{\phi(p)}(\phi(U\cap V))$ under F_* ;

- what happens to a basis for $T^*_{\psi(p)}(\psi(U\cap V))$ under F^* ;

- what happens to the coordinates of a general tangent vector under F_* ;

- what happens to the coordinates of a general covector under F^* .

We have a vector in $T_p(M)$ (z and z' are coordinate representations of this vector) and a covector in $T_p^*(M)$ (ω and ω' are coordinate representations of this vector).

Consider a covector $\omega = \omega_i \, dx^i, \quad \omega_i \in \mathbb{R}$ $\qquad \omega \in T^*_{\psi(p)}(\psi(U \cap V))$

a covector $\omega' = \omega_{i'} \, dx^{i'}$ $\qquad \omega = F^*(\omega')$

a vector $z = z^i \frac{\partial}{\partial x^i} \mid_{\phi(p)}$ $\qquad z \in T_{\phi(p)}(\phi(U \cap V))$

a vector $z' = z^{i'} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)}$ $\qquad z' = F_*(z)$

We have $\quad F^*(dx^{i'}) = \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} dx^i$ by (3)

($dx^{i'}$ is a basis for $T^*_{\psi(p)}(\psi(U \cap V))$ and

$$F_*\left(\frac{\partial}{\partial x^i} \mid_{\phi(p)}\right) = \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)} \text{ by (4)}$$

($\frac{\partial}{\partial x^i}$ is a basis for $T_{\phi(p)}(\phi(U \cap V))$)

Given a covector (or vector) at p, we can express it in terms of x^i or $x^{i'}$:

$$\omega_i \, dx^i = \omega = F^*(\omega_{i'} \, dx^{i'}) = \omega_{i'} F^*(dx^{i'}) \overset{\text{by (3)}}{=} \omega_{i'} \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} dx^i$$

Since $\{dx^i\}$ form a basis it follows that

$\omega_i = \omega_{i'} \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)}$ (i' represents the column of a matrix , i represents the row of a matrix).

The (i, i') entry of the matrix corresponding to the linear transformation F^* with respect to the bases $\{dx^{i'}\}$ in the domain and $\{dx^i\}$ in the codomain of F^* is given by $\frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)=(x^i)}$.

Next $z^{i'} \frac{\partial}{\partial x^{i'}} = z' = F_*(z) = F_*\left(z^i \frac{\partial}{\partial x^i} \mid_{\phi(p)}\right) = z^i F_*\left(\frac{\partial}{\partial x^i} \mid_{\phi(p)}\right) = z^i \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)} \frac{\partial}{\partial x^{i'}} \mid_{\psi(p)}$.

Since $\left\{\frac{\partial}{\partial x^{i'}}\right\}$ form a basis, we have that

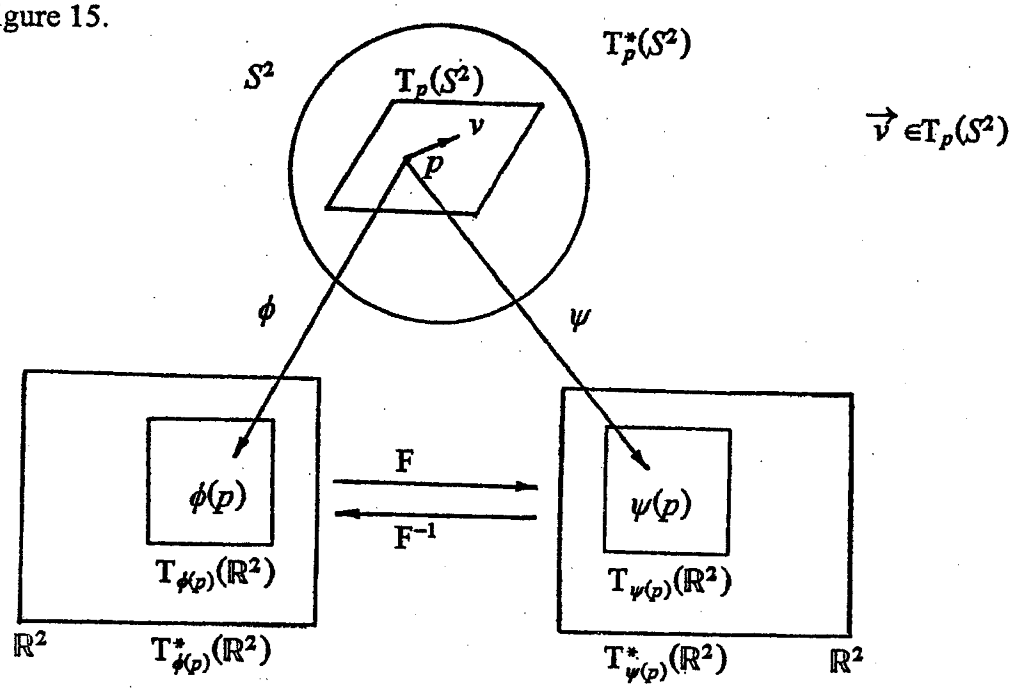
$$z^{i'} = z^i \frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)}, \quad i, i' = 1, \ldots, n$$

(i', i) entry of the matrix corresponding to the linear transformation F_* with respect to the basis $\{\frac{\partial}{\partial x^i}\}$ in the domain and $\left\{\frac{\partial}{\partial x^{i'}}\right\}$ in the codomain of F_* is given by $\frac{\partial x^{i'}}{\partial x^i} \mid_{\phi(p)}$.

Thus the matrices representing F^* and F_* with respect to the bases given are the transposes of each other.

Example: Consider a sphere. $p \in$ sphere but $p \neq N, p \neq S$.

Figure 15.



Determine two neighborhoods

$\phi : p \to \phi(p)$
$\psi : p \to \psi(p)$

We have defined earlier that $\phi(p) = (d, \theta); \psi(p) = (d', \theta')$, and we found a relationship between d and d', θ and θ'; i.e. $d' = \frac{1}{d}$, $\theta' = \theta$.

At p define the tangent space $T_p(S^2)$. At $\phi(p)$ define $T_{\phi(p)}(\mathbb{R}^2)$

And $\phi_* : T_p(S^2) \to T_{\phi(p)}(\mathbb{R}^2)$

Also we have a tangent space $T_{\psi(p)}(\mathbb{R}^2)$ with

$$\psi_* : T_p(S^2) \to T_{\psi(p)}(\mathbb{R}^2)$$

By the Theorem (2) we can define

$$F_* : T_{\phi(p)}(\mathbb{R}^2) \to T_{\psi(p)}(\mathbb{R}^2)$$

the basis for $T_{\phi(p)}(\mathbb{R}^2)$ is $\left\{\frac{\partial}{\partial d}\mid_{\phi\{p\}}, \frac{\partial}{\partial \theta}\mid_{\phi\{p\}}\right\}$, and

the basis for $T_{\psi(p)}(\mathbb{R}^2)$ is $\left\{\frac{\partial}{\partial d'}\mid_{\psi\{p\}}, \frac{\partial}{\partial \theta'}\mid_{\psi\{p\}}\right\}$;

and we can define

$$F^* : T^*_{\psi(p)}(\mathbb{R}^2) \to T^*_{\phi(p)}(\mathbb{R}^2)$$

Find the matrix corresponding to F_* - linear transformation with respect to the basis

$\left\{\frac{\partial}{\partial d}\mid_{\phi\{p\}}, \frac{\partial}{\partial \theta}\mid_{\phi\{p\}}\right\}$ in the domain and $\left\{\frac{\partial}{\partial d'}\mid_{\psi\{p\}}, \frac{\partial}{\partial \theta'}\mid_{\psi\{p\}}\right\}$ in the codomain .

$\left(z^{i'} = z^i \frac{\partial x^{i'}}{\partial x^i}\mid_{\phi(p)} \right)$

This is going to be a 2×2 matrix since we are in $\mathbb{R}^2$.

Since $d' = \frac{1}{d}$, $\theta' = \theta$ we will have

$$\frac{\partial d'}{\partial d} = \left(\frac{1}{d}\right)' = -\frac{1}{d^2}, \qquad \frac{\partial d'}{\theta} = 0$$

$$\frac{\partial \theta'}{\partial d} = 0, \qquad \frac{\partial \theta'}{\partial \theta} = 1$$

So we have $\begin{pmatrix} \frac{-1}{d^2} & 0 \\ 0 & 1 \end{pmatrix}$ for our transformation matrix

So $\begin{pmatrix} z^{d'} \\ z^{\theta'} \end{pmatrix} = \begin{pmatrix} \frac{-1}{d^2} & 0 \\ 0 & 1 \end{pmatrix}_{\phi(p)} \begin{pmatrix} z^{d} \\ z^{\theta} \end{pmatrix}$.

Since the matrix is diagonal then we have the special case when the matrix corresponding to F^* with respect to the basis $\left\{ \mathrm{d}d' \mid_{\psi(p)}, \mathrm{d}\theta' \mid_{\psi(p)} \right\}$ in the domain and $\left\{ \mathrm{d}d \mid_{\phi(p)}, \mathrm{d}\theta \mid_{\phi(p)} \right\}$ in the codomain is the same as before, i.e. $\begin{pmatrix} \frac{-1}{d^2} & 0 \\ 0 & 1 \end{pmatrix}$. So

$$\begin{pmatrix} \omega^d \\ \omega^\theta \end{pmatrix} = \begin{pmatrix} \frac{-1}{d^2} & 0 \\ 0 & 1 \end{pmatrix}_{\phi(p)} \begin{pmatrix} \omega^{d'} \\ \omega^{\theta'} \end{pmatrix}.$$

Now take an element T of T^r_s, the tensor product. We want to write down the form that T takes with respect to a basis for this tensor product :

$$T = T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots\ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots\ldots \otimes dx^{j_s}$$

(implying $r+s$ summation symbols)

That is for each $i_1, \ldots, i_r$ take a basis vector from T_p ;

for each $j_1, \ldots, j_s$ take a basis vector from T^*_p ;

$i_1, \ldots, i_r, j_1, \ldots, j_s$ range from 1 to n where n is the dimension of the our manifold .

T is a mapping from $\prod^s_r$ into $\mathbb{R}$:

$$T: \prod\nolimits^s_r \to \mathbb{R} .$$

What does it do to a general element of set $\prod^s_r$?

Given $X \in \prod^s_r$, define T(X).

We also know that T is a multilinear mapping by definition, which means that it is linear in each factor, which means that just like a linear transformation what T does to an arbitrary element of our vector space is completely determined by what T does to a basis. Thus it is sufficient to find T on a basis for $\prod^s_r$.

$$T\left(\left(dx^{k_1}, \ldots\ldots, dx^{k_r}, \frac{\partial}{\partial x^{e_1}}, \ldots\ldots, \frac{\partial}{\partial x^{e_s}}\right)\right) =$$

($k_1, \ldots, k_r, e_1, \ldots, e_s = 1, \ldots, n$; $\dim(\prod^s_r) = n^{s+r}$)

(take each corresponding operator and operate on the corresponding argument)

$$= T^{i_1 \ldots\ldots i_r}{}_{j_1 \ldots\ldots j_s} \frac{\partial}{\partial x^{i_1}} (dx^{k_1}) \cdot \ldots\ldots \cdot \frac{\partial}{\partial x^{i_r}} (dx^{k_r}) \cdot dx^{j_1} \left(\frac{\partial}{\partial x^{e_1}}\right) \cdot \ldots\ldots \cdot dx^{j_s} \left(\frac{\partial}{\partial x^{e_s}}\right)$$

($\frac{\partial}{\partial x^{i_1}}$ and dx^{k_1} are dual of each other)

$=T^{i_1 \dots i_r}_{j_1 \dots j_s} \delta^{k_1}_{i_1} \cdot \dots \cdot \delta^{k_r}_{i_r} \delta^{j_1}_{e_1} \cdot \dots \cdot \delta^{j_s}_{e_s}$ (where "·" denotes ordinary multiplication of real numbers and $\delta^{k_1}_{i_1} = \begin{cases} 0, k_1 \neq i_1 \\ 1, k_1 = i_1 \end{cases}$ and so on).

So the only terms surviving will be when $i_1 = k_1, \dots, i_r = k_r$, $j_1 = e_1, \dots, j_s = e_s$

$$\Rightarrow \quad T\left(\left(dx^{k_1}, \dots, dx^{k_r}, \frac{\partial}{\partial x^{e_1}}, \dots, \frac{\partial}{\partial x^{e_s}}\right)\right) = T^{k_1 \dots k_r}{}_{e_1 \dots e_s}$$

Example 1). $T\left(dx^1, dx^2, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\right) = T^{12}{}_{12}$

(we are not saying how big the dimension of our manifold is)

Example 2). Suppose $T = \underbrace{\frac{\partial}{\partial x^1} \otimes \dots \otimes \frac{\partial}{\partial x^1}}_{r \text{ times}} \otimes \underbrace{dx^1 \otimes \dots \otimes dx^1}_{s \text{ times}}$. In other words, $T^{1 \dots 1}_{1 \dots 1} = 1$, all other components of T are zero.

Find T(X) $\forall X \in \prod^s_r$ ($\prod^s_r$ is the domain).

1. Find T on a basis for $\prod^s_r$.
2. Write X in terms of a basis for $\prod^s_r$ and apply T.

1. $T\left(dx^{k_1}, \dots, dx^{k_r}, \frac{\partial}{\partial x^{e_1}}, \dots, \frac{\partial}{\partial x^{e_s}}\right) =$

$$= T^{1 \dots 1}{}_{1 \dots 1} \frac{\partial}{\partial x^1}(dx^{k_1}) \cdot \dots \cdot \frac{\partial}{\partial x^1}(dx^{k_r}) \cdot dx^1\left(\frac{\partial}{\partial x^{e_1}}\right) \cdot \dots \cdot dx^1\left(\frac{\partial}{\partial x^{e_s}}\right)$$

$= T^{1 \dots 1}{}_{1 \dots 1} \delta^{k_1}_1 \cdot \dots \cdot \delta^{k_r}_1 \cdot \delta^1_{e_1} \cdot \dots \cdot \delta^1_{e_s}$ (everything will go away unless $k_1 = 1, \dots, k_r = 1, e_1 = 1, \dots, e_s = 1$)

$= T^{1 \dots 1}_{1 \dots 1} = 1$

So $T\left(dx^{k_1}, \dots, dx^{k_r}, \frac{\partial}{\partial x^{e_1}}, \dots, \frac{\partial}{\partial x^{e_s}}\right) = 1$.

2. Consider $X \in \prod^s_r$: $X = (\omega^1, \dots, \omega^r, z_1, \dots, z_s)$

$$T = \underbrace{\frac{\partial}{\partial x^1} \otimes \dots \otimes \frac{\partial}{\partial x^1}}_{r \text{ times}} \otimes \underbrace{dx^1 \otimes \dots \otimes dx^1}_{s \text{ times}}.$$

So $T(X) = \frac{\partial}{\partial x^1}(\omega^1) \cdot \dots \cdot \frac{\partial}{\partial x^1}(\omega^r) \cdot dx^1(z_1) \cdot \dots \cdot dx^1(z_s)$

$$= \frac{\partial}{\partial x^1}\left(\omega^1_{i_1} dx^{i_1}\right) \cdot \dots \cdot \frac{\partial}{\partial x^1}\left(\omega^r_{i_r} dx^{i_r}\right) \cdot dx^1\left(z^{j_1}_1 \frac{\partial}{\partial x^{j_1}}\right) \cdot \dots \cdot dx^1\left(z^{j_s}_s \frac{\partial}{\partial x^{j_s}}\right)$$

$$= \omega^1_{i_1}\left(\tfrac{\partial}{\partial x^1}\, \mathrm{d}\, x^{i_1}\right) \cdot \ldots \cdot \omega^r_{i_r}\left(\tfrac{\partial}{\partial x^1}\, \mathrm{d}x^{i_r}\right) \cdot \mathrm{z}^{j_1}_1\left(\mathrm{d}x^1\, \tfrac{\partial}{\partial x^{j_1}}\right) \cdot \ldots \cdot \mathrm{z}^{j_s}_s\left(\mathrm{d}x^1\, \tfrac{\partial}{\partial x^{j_s}}\right)$$

$$= \omega^1_{i_1}\delta^{i_1}_1 \cdot \ldots \cdot \omega^r_{i_r}\delta^{i_r}_1 \cdot \mathrm{z}^{j_1}_1\delta^1_{j_1} \cdot \ldots \cdot \mathrm{z}^{j_s}_s\delta^1_{j_s}$$ (the only terms surviving will be with $i_k, j_e = 1$)

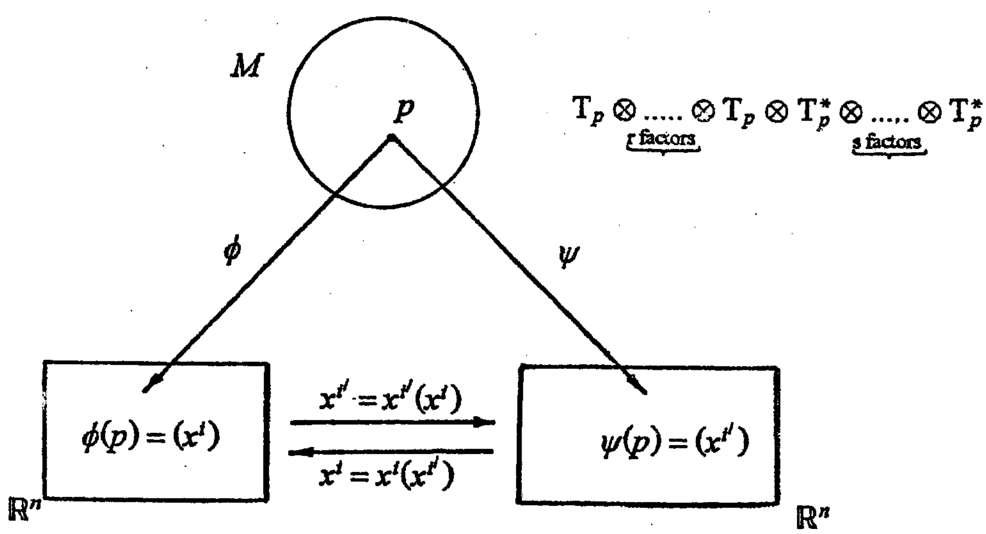
$$= \omega^1_1 \cdot \ldots \cdot \omega^r_1 \cdot \mathrm{z}^1_1 \cdot \ldots \cdot \mathrm{z}^1_s \in \mathbb{R}\,.$$

Now we have a question: what happens to the components of a general tensor when we change the coordinates?

Suppose we have a tensor $\mathrm{T} \in \mathrm{T}^r_s(p)$. Geometrically, this means that we have a point p on the manifold, and we have a space attached to that point p .

This space is $\underbrace{\mathrm{T}_p \otimes \ldots\ldots \otimes \mathrm{T}_p}_{\text{r factors}} \otimes \underbrace{\mathrm{T}^*_p \otimes \ldots\ldots \otimes \mathrm{T}^*_p}_{\text{s factors}}$, which is a vector space with $\dim M = n$; hence the dimension of this vector space is n^{r+s}.

Figure 16.



For a point p in the intersection of two coordinate neighborhoods, (U, ϕ) and (V, ψ), we have two mappings ϕ and ψ : ϕ produces coordinates $\phi(p)$

ψ produces coordinates $\psi(p)$.

And we have transition functions : $x^{i'} = x^{i'}(x^i)$ and

$$x^i = x^i(x^{i'}) .$$

We know how tangent vectors change in their coordinate representations and how covectors change in their coordinate representations. Now we are going to do this for a general tensor.

The tensor T is the same object regardless of what coordinate system we use , but has two different representations; one representation, with respect to the old coordinates:

$$(*) \quad \mathrm{T}=\mathrm{T}^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots.. \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots.. \otimes dx^{j_s} ,$$

and a second representation with respect to the new coordinates:

$$(**) \quad \mathrm{T}=\mathrm{T}^{i'_1 \ldots i'_r}{}_{j'_1 \ldots j'_s} \frac{\partial}{\partial x^{i'_1}} \otimes \ldots.. \otimes \frac{\partial}{\partial x^{i'_r}} \otimes dx^{j'_1} \otimes \ldots.. \otimes dx^{j'_s} .$$

Suppose we know (*). So we can replace each of the factors in (**) individually in terms of factors in (*). So we replace $\frac{\partial}{\partial x^{i'_1}}$ with $\frac{\partial x^{i_1}}{\partial x^{i'_1}} \frac{\partial}{\partial x^{i_1}}$ and so on. Thus we have (with all partial derivatives evaluated at "*p*")

$$\mathrm{T}=\mathrm{T}^{i'_1 \ldots i'_r}{}_{j'_1 \ldots j'_s} \frac{\partial}{\partial x^{i'_1}} \otimes \ldots.. \otimes \frac{\partial}{\partial x^{i'_r}} \otimes dx^{j'_1} \otimes \ldots.. \otimes dx^{j'_s}$$

$$=\mathrm{T}^{i'_1 \ldots i'_r}{}_{j'_1 \ldots j'_s} \frac{\partial x^{i_1}}{\partial x^{i'_1}} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial x^{i_r}}{\partial x^{i'_r}} \frac{\partial}{\partial x^{i_r}} \otimes \frac{\partial x^{j'_1}}{\partial x^{j_1}} dx^{j_1} \otimes \ldots \otimes \frac{\partial x^{j'_s}}{\partial x^{j_s}} dx^{j_s}$$

but $\frac{\partial x^{i_1}}{\partial x^{i'_1}}, \ldots, \frac{\partial x^{j'_s}}{\partial x^{j_s}}$ are just numbers, so they can be pulled out in front.

$$=\mathrm{T}^{i'_1 \ldots i'_r}{}_{j'_1 \ldots j'_s} \frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdot \ldots.. \cdot \frac{\partial x^{i_r}}{\partial x^{i'_r}} \cdot \frac{\partial x^{j'_1}}{\partial x^{j_1}} \cdot \ldots.. \cdot \frac{\partial x^{j'_s}}{\partial x^{j_s}} \cdot \frac{\partial}{\partial x^{i_1}} \otimes \ldots.. \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots.. \otimes dx^{j_s}$$

Now setting this equal to the right-hand side of (*) and using the fact that these two expressions for T are now with respect to the same basis, and coordinates are unique, yields

$$T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = T^{i'_1 \ldots i'_r}{}_{j'_1 \ldots j'_s} \cdot \frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdot \ldots\ldots \cdot \frac{\partial x^{i_r}}{\partial x^{i'_r}} \cdot \frac{\partial x^{j'_1}}{\partial x^{j_1}} \cdot \ldots\ldots \cdot \frac{\partial x^{j'_s}}{\partial x^{j_s}},$$

which is called **the Tensor Transformation Law**.

But the important fact is that the tensor itself as an object does not change. Since we want physical laws of the universe to not depend on which coordinate system we use, we do not want what happens to us in the universe to depend on how we label where we are. This is called the covariance of the laws of physics, which basically means that we want our basic physical principles to remain unchanged when we change our coordinate system.

For example, let us consider an arbitrary element $T \in T_p(M)$ given by

$T = T^i \frac{\partial}{\partial x^i}$ (in unprime coordinates) and $T' = T^{i'} \frac{\partial}{\partial x^{i'}}$ (in prime coordinates).

We need to show that $T = T'$.

We know that $T^{i'} = T^i \frac{\partial x^{i'}}{\partial x^i}$ (by the Transformation Law) and

$$\frac{\partial}{\partial x^{i'}} = \frac{\partial}{\partial x^i}\frac{\partial x^i}{\partial x^{i'}} = \frac{\partial}{\partial x^j}\frac{\partial x^j}{\partial x^{i'}} \text{ (changing } i \text{ to } j).$$

So we have $T' = T^{i'} \frac{\partial}{\partial x^{i'}} = T^i \frac{\partial x^{i'}}{\partial x^i}\left[\frac{\partial}{\partial x^j}\frac{\partial x^j}{\partial x^{i'}}\right]$

$$= T^i \left[\frac{\partial x^{i'}}{\partial x^i}\frac{\partial x^j}{\partial x^{i'}}\right]\frac{\partial}{\partial x^j}$$

$$= T^i \frac{\partial x^j}{\partial x^i}\frac{\partial}{\partial x^j}$$

$$= T^i \delta^j_i \frac{\partial}{\partial x^j}$$

$$= T^j \frac{\partial}{\partial x^j} = T^i \frac{\partial}{\partial x^i} = T \text{ (change } j \text{ to } i)$$

Thus $T' = T$ (tensors are independent of coordinate system).

Now consider a special kind of tensor : $F = F_{ab}\left(\frac{dx^a \otimes dx^b - dx^b \otimes dx^a}{2}\right)$, $a, b = 0, 1, 2, 3.$

$T^*_p(M)$ has a basis $\{dx^a\}$, so $\{dx^a \otimes dx^b\}$ is an element of $T^*_p \otimes T^*_p$, and

$\dim(T_p^* \otimes T_p^*)=16$ (4 elements for a , 4 elements for b).

Also note that $dx^a \otimes dx^b$ is not the same as $dx^b \otimes dx^a$.

<u>Example</u> : for $a = 0, b = 1$ consider $dx^0 \otimes dx^1$ versus $dx^1 \otimes dx^0$.

These are different mappings in $T_p^* \otimes T_p^*$.

Remember that $T_p^* \otimes T_p^*$ acts on $T_p \otimes T_p$ where the first covector acts on the first vector in our pair and the second covector acts on the second vector. But if we switch the mappings then acting on the first vector with dx^0 and acting on the second vector with dx^1 is not necessarily the same thing as acting on the first vector with dx^1 and acting on the second vector with dx^0 .(e.g. $dx^1\left(\frac{\partial}{\partial x^0}\right) = 0$, but $dx^0\left(\frac{\partial}{\partial x^0}\right) = 1$).

Thus $dx^0 \otimes dx^1$ is not the same as $dx^1 \otimes dx^0$. Of course if $a = b$ then they are the same.

In fact if $a = b$, then $dx^a \otimes dx^a - dx^a \otimes dx^a = 0$.

Later in this thesis, we will describe how the F given above can be interpreted as the electromagnetic field tensor in physics.

So let us look at F_{ab} more closely : F_{ab} will be a skew-symmetric 4×4 matrix.

When $a = b$, it is going to be 0 : $F_{aa} = 0, a = 0, 1, 2, 3.$

Next compare F_{ab} and F_{ba} . We have $F_{ab}= -F_{ba}$.

Now $F_{ab} = -F_{ba}$ and $F_{aa} = 0$ are properties which define a **skew-symmetric matrix**:

that is $F^T = -F$,

$$F_{ab} = \begin{matrix} a=0 \\ a=1 \\ a=2 \\ a=3 \end{matrix} \begin{pmatrix} 0 & & & \\ & 0 & * & \\ - & * & 0 & \\ & & & 0 \\ {\scriptstyle b=0} & {\scriptstyle b=1} & {\scriptstyle b=2} & {\scriptstyle b=3} \end{pmatrix}$$

Maxwell was able to write down equations that show how E's (electric) and B's (magnetic) fields are related to each other. And those relationships became known as

Maxwell's equations, which describe the behavior between electric and magnetic fields.

F_{ab} will consist of entries giving the various components of E and B.

Next suppose $g = g_{ab}\left(\frac{dx^a \otimes dx^b + dx^b \otimes dx^a}{2}\right)$, a, b=0,1,2,3.

If we change a and b, we will get the same matrix .

g_{ab} becomes as a matrix

$$\begin{pmatrix} g_{00} & & & * \\ & g_{11} & & \\ & & g_{22} & \\ * & & & g_{33} \end{pmatrix} \text{ and}$$

g is **symmetric**; that is $g^T = g$.

We would like to prove that every element of $T_p^* \otimes T_p^*$ ($\dim M = n$) can be uniquely written as the sum of a symmetric and skew symmetric tensor .

Proof: Take $S \in T_p^* \otimes T_p^*$, then $S = S_{ab}\, dx^a \otimes dx^b$

First define $S_{(ab)} = \frac{1}{2}(S_{ab} + S_{ba})$

and define $S_{[ab]} = \frac{1}{2}(S_{ab} - S_{ba})$.

Show $S_{ab} = S_{(ab)} + S_{[ab]}$.

Show $S_{(ab)}$ is symmetric. (That is, show $S_{(ab)} = S_{(ba)}$)

$$S_{(ba)} = \tfrac{1}{2}(S_{ba} + S_{ab}) = \tfrac{1}{2}(S_{ab} + S_{ba}) = S_{(ab)}.$$

Show $S_{[ab]}$ is skew-symmetric (that is, show $S_{[ab]} = -S_{[ba]}$).

$$S_{[ba]} = \tfrac{1}{2}(S_{ba} - S_{ab}) = -\tfrac{1}{2}(S_{ab} - S_{ba}) = -S_{[ab]}.$$

Show $S_{ab} = S_{(ab)} + S_{[ab]}$

$$S_{(ab)} + S_{[ab]} = \tfrac{1}{2}(S_{ab} + S_{ba}) + \tfrac{1}{2}(S_{ab} - S_{ba})$$

$$= \tfrac{1}{2}S_{ab} + \tfrac{1}{2}S_{ba} + \tfrac{1}{2}S_{ab} - \tfrac{1}{2}S_{ba}$$

$=\frac{1}{2}S_{ab}+\frac{1}{2}S_{ab}$

$=S_{ab}$

So every element of $T_p^* \otimes T_p^*$ can be written as the sum of a symmetric and skew-symmetric tensor .

Finally, show that this decomposition is unique.

Let A be $n \times n$ matrix; A=B+C such that $B^T = B$ and $C^T = -C$.

Then $A^T = (B+C)^T = B^T + C^T = B - C$, A=B+C , and $2C = A - A^T$.

So $C = \frac{1}{2}(A - A^T)$

and $B = \frac{1}{2}(A + A^T)$

and these are unique solutions for B and C, with B symmetric and

C skew-symmetric.

Thus we have shown that every element of $T_p^* \otimes T_p^*$ can be uniquely written as the sum of a symmetric and skew-symmetric tensor. And the components of B and C are as given above.

Now we will look at $\bigwedge^r$ - skew-symmetric elements of $\underbrace{T_p^* \otimes \ldots \otimes T_p^*}_{r \text{ times}}$.

Suppose we have a tensor $T = T_{a_1 \ldots\ldots a_r} dx^{a_1} \otimes \ldots\ldots \otimes dx^{a_r}$.

We want to define what we mean by $T_{[a_1 \ldots\ldots a_r]}$:

$T_{[a_1 \ldots\ldots a_r]} = \frac{1}{r!}$ (alternating sum over all permutations of $a_1, \ldots, a_r$)

Example : $T_{[abcd]} =$

$$\frac{1}{24}\begin{pmatrix} T_{abcd} - T_{bacd} + T_{cabd} - T_{dabc} - T_{acbd} + T_{adbc} - T_{abcd} + T_{acdb} \\ -T_{adcb} + T_{bcad} - T_{bdac} - T_{bcda} + T_{bdca} + T_{badc} - T_{cadb} - T_{cbad} \\ +T_{cbda} + T_{cdab} - T_{cdba} + T_{dacb} + T_{dbac} - T_{dbca} - T_{dcab} + T_{dcba} \end{pmatrix}$$

$\bigwedge^r \subseteq \underbrace{T_p^* \otimes \ldots\ldots \otimes T_p^*}_{r \text{ times}}$

A basis for $\bigwedge^r$ is $dx^{a_1} \wedge dx^{a_2} \wedge \ldots\ldots \wedge dx^{a_r}$ (we call $\wedge$ a wedge product)

Example : Consider $\bigwedge^2 \subseteq T_p^* \otimes T_p^*$;

$$\text{if } T = T_{ab}\, dx^a \otimes dx^b \text{, then } T_{[ab]} dx^a \otimes dx^b = \tfrac{1}{2}[T_{ab} - T_{ba}] dx^a \otimes dx^b$$

$$= \tfrac{1}{2}[T_{ab} dx^a \otimes dx^b - T_{ba} dx^a \otimes dx^b] = \tfrac{1}{2}[T_{ab} dx^a \otimes dx^b - T_{ab} dx^b \otimes dx^a]$$

$$= T_{ab}\left(\frac{dx^a \otimes dx^b - dx^b \otimes dx^a}{2}\right) = T_{ab} dx^a \wedge dx^b \text{ and } dx^a \wedge dx^b \text{ is a basis for } \bigwedge^2.$$

A typical element of $\bigwedge^2$ with n=4 can be represented by

$$\begin{pmatrix} 0 & a & \beta & \gamma \\ -a & 0 & \delta & \in \\ -\beta & -\delta & 0 & \phi \\ -\gamma & -\in & -\phi & 0 \end{pmatrix}.$$

Example. Consider a skew-symmetric matrix $\begin{pmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 4 & 5 \\ -2 & -4 & 0 & 7 \\ -3 & -5 & -7 & 0 \end{pmatrix}$. $dx^a \wedge dx^b = -dx^b \wedge dx^a$,

and the entries above the main diagonal will have the property that the row number is less

than the column number (e.g. $(a,b) = (1,2) : b > a$).

So consider $\{dx^a \wedge dx^b : a, b = 0, 1, 2, 3 \text{ such that } b > a\}$. Then since $dx^a \wedge dx^b = -$

$dx^b \wedge dx^a$, this set will form a basis for $\bigwedge^2$.

-find the dimension of this space :

when $\dim(M) = 4$, the dimension of this space is 6 ,

-write out a basis for $\bigwedge^2(M)$ when $\dim M = 4$:

$$\{\ dx^0 \wedge dx^1 ,\ dx^0 \wedge dx^2 ,\ dx^0 \wedge dx^3 ,\ dx^1 \wedge dx^2 , dx^1 \wedge dx^3 , dx^2 \wedge dx^3 \}.$$

Now find $\dim\left(\bigwedge^2(M)\right)$ if $\dim M = n$.

Then we have $1+2+3+\ldots\ldots+n-1 \equiv S_{n-1}$ independent entries in our $n \times n$ skew-symmetric matrix.

So $S_{n-1} = \frac{n(n-1)}{2}$. We will prove this by induction :

1) for $n = 2$, $1 = \frac{2(2-1)}{2} = 1$

2) suppose $S_{n-1} = \frac{n(n-1)}{2}$ is true for some $n-1$ ($n \geq 3$).

3) show it is true for $(n-1)+1$ (for n) .

$$1+2+3+\ldots+(n-1)+n = \frac{n(n-1)}{2} + n = \frac{n(n+1)}{2} = S_n \, .$$

By the principle of mathematical induction , we conclude that $S_{n-1} = \frac{n(n-1)}{2}$ is true for all positive integers $n-1$.

We want to define a differentiation operation, d, on this set $\bigwedge^r(M)$, which is known as the **exterior derivative.** (d operates on $\bigwedge^r(M)$)

__**Theorem 3.**__ Let M be any C^∞ manifold and let $\bigwedge(M)$ be the algebra of skew-symmetric forms on M . Then there exists a unique R-linear map

$\mathrm{d}_m : \bigwedge(M) \rightarrow \bigwedge(M)$ such that

1) if $f \in \bigwedge^0(M) = C^\infty(M)$, then $\mathrm{d}_m f = \mathrm{d}f$, the differential of f;

2) if $\theta \in \bigwedge^r(M)$ and $\sigma \in \bigwedge^s(M)$, then $\mathrm{d}_m(\theta \wedge \sigma) = \mathrm{d}_m\theta \wedge \sigma + (-1)^r \theta \wedge \mathrm{d}_m \sigma$;

3) $\mathrm{d}_m^2 = 0$.

Explanations :

1) If $r = 0$, then $\bigwedge^0(M) = C^\infty$ functions from M into $\mathbb{R}$.

If $r = 1$ we use $\{\mathrm{d}x^a\}$ as a basis for $\bigwedge^1(M)$.

If $r = 2$ we use $\{\mathrm{d}x^a \wedge \mathrm{d}x^b , b > a \}$ as a basis for $\bigwedge^2(M)$.

Let $f \in C^\infty$. How does d act on it ?

$\mathrm{d}f = \frac{\partial f}{\partial x^i} \mathrm{d}x^i$ (from calculus), $i = 1, \ldots, n$, $n =$ dim of M.

Since dx^i is a basis for T_p^* then $df = \frac{\partial f}{\partial x^i} dx^i$ is an element of T_p^*

2) take $\theta \in \bigwedge^r(M)$: $\theta = a dx^{i_1} \wedge ..... \wedge dx^{i_r}$

$\sigma \in \bigwedge^s(M)$: $\sigma = b dx^{j_1} \wedge ..... \wedge dx^{j_s}$

$$d_M(\theta \wedge \sigma) = d_M[(a dx^{i_1} \wedge ... \wedge dx^{i_r}) \wedge (b dx^{j_1} \wedge ..... \wedge dx^{j_s})]$$

$$= d_M(ab) \wedge (dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge (dx^{j_1} \wedge ..... \wedge dx^{j_s})$$

$$= ((d_M a)b + a(d_M b)) \wedge (dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge (dx^{j_1} \wedge ..... \wedge dx^{j_s})$$

$$= (d_M a \wedge dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge (b dx^{j_1} \wedge ..... \wedge dx^{j_s}) +$$

$$+ (-1)^r (a dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge (d_M b \wedge dx^{j_1} \wedge ..... \wedge dx^{j_s})$$, where we can explain this last step by the following :

we are going use the fact that $db = \frac{\partial b}{\partial x^e} dx^e$.

Now we have $a \frac{\partial b}{\partial x^e} dx^e \wedge (dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge (dx^{j_1} \wedge ..... \wedge dx^{j_s})$.

We are going to interchange the dx^e factor with each of the dx^{i_k} factors (where k=1,...,r) . And we are going to have r interchanges.

$$(-1)\, dx^{i_1} \wedge dx^e \wedge ....$$

$$(-1)^2\, dx^{i_1} \wedge dx^{i_2} \wedge dx^e \wedge ....$$

$$(-1)^3\, dx^{i_1} \wedge dx^{i_2} \wedge dx^{i_3} \wedge dx^e \wedge ....$$

.

.

$$(-1)^r (dx^{i_1} \wedge ..... \wedge dx^{i_r}) \wedge db$$

The exterior differentiation operator d maps r–form fields linearly to $(r+1)$–form fields:

$$d : \quad \bigwedge^r(M) \to \bigwedge^{r+1}(M) .$$

Let $\theta = \theta_{i_1 ... i_r} dx^{i_1} \wedge ... \wedge dx^{i_r}$. What does $d\theta$ belong to ?

$\theta_{i_1 \ldots i_r} \in \bigwedge^0(M) \quad (r=0)$

By 1) $\mathrm{d}\theta_{i_1 \ldots i_r} = \frac{\partial \theta_{i_1 \ldots i_r}}{\partial x^e} \mathrm{d}x^e$

By 2) $\mathrm{d}\theta = (\mathrm{d}\theta_{i_1 \ldots i_r})\, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r} + \theta_{i_1 \ldots i_r} (-1)^0\, \mathrm{d}\,(\mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r})$

$$= \frac{\partial \theta_{i_1 \ldots i_r}}{\partial x^e} \mathrm{d}x^e \wedge \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r} + 0 \quad (\, \mathrm{d}^2 = 0 \,)$$

$$= \frac{\partial \theta_{i_1 \ldots i_r}}{\partial x^e} \mathrm{d}x^e \wedge \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r}$$

we can see that we have extra factor in our wedge product, so $\mathrm{d}\theta \in \bigwedge^{r+1}(M)$.

So $\mathrm{d}: \quad \bigwedge^r(M) \to \bigwedge^{r+1}(M)$.

Now we give an example of the calculation of an exterior derivative in the case that we have a 1-form covector in 3 dimensions.

Let $\gamma \in \bigwedge^1(M) \quad : \gamma = \gamma_a\, \mathrm{d}x^a$, $a = 1,2,3$ or (x,y,z).

We will write out in detail what $\mathrm{d}\gamma$ looks like ($\mathrm{d}\gamma \in \bigwedge^2(M)$) and give a "calculus/vector analysis" interpretation to this.

$$\mathrm{d}\gamma = \mathrm{d}(\gamma_a\, \mathrm{d}x^a) = (\mathrm{d}\gamma_a)\, \mathrm{d}x^a + \gamma_a\, \mathrm{d}\,(\mathrm{d}x^a)$$

$$= \frac{\partial \gamma_a}{\partial x^b} \mathrm{d}x^b \wedge \mathrm{d}x^a + 0$$

$$= \frac{\partial \gamma_a}{\partial x^b} \mathrm{d}x^b \wedge \mathrm{d}x^a$$

$$\equiv \gamma_{a,b}\, \mathrm{d}x^b \wedge \mathrm{d}x^a \,.$$

Now we are going to show that $\gamma_{a,b}\mathrm{d}x^b \wedge \mathrm{d}x^a = \gamma_{[a,b]}\mathrm{d}x^b \wedge \mathrm{d}x^a$.

First consider an example. Suppose $a = 1,2$ and $b = 1,2$ then

$$\gamma_{a,b}\mathrm{d}x^b \wedge \mathrm{d}x^a = (\gamma_{(a,b)} + \gamma_{[b,a]})\mathrm{d}x^b \wedge \mathrm{d}x^a$$

$$= \tfrac{1}{2}(\gamma_{a,b} + \gamma_{b,a} + \gamma_{a,b} - \gamma_{a,b})\mathrm{d}x^b \wedge \mathrm{d}x^a$$

$$= \tfrac{1}{2}(\gamma_{1,2} + \gamma_{2,1} + \gamma_{1,2} - \gamma_{2,1})\mathrm{d}x^2 \wedge \mathrm{d}x^1 + \tfrac{1}{2}(\gamma_{2,1} + \gamma_{1,2} + \gamma_{2,1} - \gamma_{1,2})\mathrm{d}x^1 \wedge \mathrm{d}x^2$$

$=\frac{1}{2}\gamma_{1,2}dx^2 \wedge dx^1 + \frac{1}{2}\gamma_{2,1}dx^2 \wedge dx^1 + \frac{1}{2}\gamma_{1,2}dx^2 \wedge dx^1 + \frac{1}{2}\gamma_{2,1}dx^1 \wedge dx^2$

$-\frac{1}{2}\gamma_{2,1}dx^2 \wedge dx^1 - \frac{1}{2}\gamma_{1,2}dx^2 \wedge dx^1 - \frac{1}{2}\gamma_{2,1}dx^2 \wedge dx^1 - \frac{1}{2}\gamma_{1,2}dx^1 \wedge dx^2$

$=\frac{1}{2}\gamma_{1,2}dx^2 \wedge dx^1 - \frac{1}{2}\gamma_{2,1}dx^2 \wedge dx^1 + \frac{1}{2}\gamma_{2,1}dx^1 \wedge dx^2 - \frac{1}{2}\gamma_{1,2}dx^1 \wedge dx^2$

$=\gamma_{[1,2]}dx^2 \wedge dx^1 + \gamma_{[2,1]}dx^1 \wedge dx^2$

$=\gamma_{[a,b]}dx^b \wedge dx^a$

In general we have

$$\gamma_{a,b}dx^b \wedge dx^a = (\gamma_{(a,b)} + \gamma_{[a,b]})dx^b \wedge dx^a$$

$$=\frac{1}{2}(\gamma_{a,b} + \gamma_{b,a} + \gamma_{a,b} - \gamma_{b,a})dx^b \wedge dx^a$$

$$=\frac{1}{2}\sum_{a<b}\sum (\gamma_{a,b} + \gamma_{b,a} + \gamma_{a,b} - \gamma_{b,a})dx^b \wedge dx^a$$

$$+\frac{1}{2}\sum_{a>b}\sum (\gamma_{a,b} + \gamma_{b,a} + \gamma_{a,b} - \gamma_{b,a})dx^b \wedge dx^a$$

Each $\frac{1}{2}\underset{a<b}{\gamma_{a,b}}dx^b \wedge dx^a$ will be canceled with each $\frac{1}{2}\underset{a>b}{\gamma_{b,a}}dx^b \wedge dx^a$ and

$\frac{1}{2}\underset{a<b}{\gamma_{b,a}}dx^b \wedge dx^a$ with $\frac{1}{2}\underset{a>b}{\gamma_{a,b}}dx^b \wedge dx^a$ (since $dx^b \wedge dx^a = -dx^a \wedge dx^b$).

Each $\frac{1}{2}\underset{a<b}{\gamma_{a,b}}dx^b \wedge dx^a$ will be combined with $\frac{1}{2}\underset{a>b}{\gamma_{a,b}}dx^b \wedge dx^a = -\frac{1}{2}\underset{a>b}{\gamma_{a,b}}dx^a \wedge dx^b$ and

$-\frac{1}{2}\underset{a<b}{\gamma_{b,a}}dx^b \wedge dx^a = \frac{1}{2}\underset{a<b}{\gamma_{b,a}}dx^a \wedge dx^b$ with $-\frac{1}{2}\underset{a>b}{\gamma_{b,a}}dx^b \wedge dx^a$,

so we would have $\gamma_{[a,b]}dx^b \wedge dx^a$.

So $\gamma_{a,b}dx^b \wedge dx^a = \gamma_{[a,b]}dx^b \wedge dx^a$.

But $\gamma_{[a,b]}dx^b \wedge dx^a$

$$=\frac{1}{2}(\gamma_{1,2} - \gamma_{2,1})\, dx^2 \wedge dx^1 + \frac{1}{2}(\gamma_{1,3} - \gamma_{3,1})\, dx^3 \wedge dx^1 + \frac{1}{2}(\gamma_{2,3} - \gamma_{3,2})dx^3 \wedge dx^2$$

$$=\frac{1}{2}(\gamma_{x,y} - \gamma_{y,x})dy \wedge dx + \frac{1}{2}(\gamma_{x,z} - \gamma_{z,x})dz \wedge dx + \frac{1}{2}(\gamma_{y,z} - \gamma_{z,y})dz \wedge dy$$

$$=\frac{1}{2}(\gamma_{y,x} - \gamma_{x,y})dx \wedge dy + \frac{1}{2}(\gamma_{z,x} - \gamma_{x,z})dx \wedge dz + \frac{1}{2}(\gamma_{z,y} - \gamma_{y,z})dy \wedge dz$$

$$=\frac{1}{2}\left(\frac{\partial\gamma_y}{\partial x} - \frac{\partial\gamma_x}{\partial y}\right) dx \wedge dy + \frac{1}{2}\left(\frac{\partial\gamma_z}{\partial x} - \frac{\partial\gamma_x}{\partial z}\right) dx \wedge dz + \frac{1}{2}\left(\frac{\partial\gamma_z}{\partial y} - \frac{\partial\gamma_y}{\partial z}\right) dy \wedge dz$$

In this case $\gamma = (\gamma_x, \gamma_y, \gamma_z)$, and we have from calculus/vector analysis that

$$\begin{bmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \gamma_x & \gamma_y & \gamma_z \end{bmatrix} = \text{curl}\,\gamma = i\left(\frac{\partial \gamma_z}{\partial y} - \frac{\partial \gamma_y}{\partial z}\right) - j\left(\frac{\partial \gamma_z}{\partial x} - \frac{\partial \gamma_x}{\partial z}\right) + k\left(\frac{\partial \gamma_y}{\partial x} - \frac{\partial \gamma_x}{\partial y}\right)$$

where $\frac{1}{2}\mathrm{d}y \wedge \mathrm{d}z$ corresponds to i

$\frac{1}{2}\mathrm{d}x \wedge \mathrm{d}z$ corresponds to $-j$

$\frac{1}{2}\mathrm{d}x \wedge \mathrm{d}y$ corresponds to k.

Now consider $\mathrm{A} = \mathrm{A}_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r}$.

The question is how the components of dA transform under a change of coordinates using the transition functions. We have (using properties 2 and 3 on pg38)

$$\mathrm{dA} = \mathrm{dA}_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r},$$

and next we are going to take dA using the primed coordinates and show that we get the same answer as if we take dA using the unprimed coordinates. That way we can say that the exterior differentiation operator is covariant (the coordinates might change but the basic form of the rule should not change).

So consider another set of coordinates $\{x^{i'}\}$. Then $\mathrm{A} = \mathrm{A}_{i'_1 \ldots i'_r} \mathrm{d}x^{i'_1} \wedge \ldots \wedge \mathrm{d}x^{i'_r}$, where the components $\mathrm{A}_{i'_1 \ldots i'_r}$ are given by

$$\mathrm{A}_{i'_1 \ldots i'_r} = \frac{\partial x^{i_1}}{\partial x^{i'_1}} \frac{\partial x^{i_2}}{\partial x^{i'_2}} \cdots\cdots \frac{\partial x^{i_r}}{\partial x^{i'_r}} \mathrm{A}_{i_1 \ldots i_r} .$$

Thus the $(r+1)$ –form, dA, defined by these coordinates is

$$\mathrm{dA} = \mathrm{d}\left(A_{i'_1 \ldots i'_r} \mathrm{d}x^{i'_1} \wedge \ldots \wedge \mathrm{d}x^{i'_r}\right) = \mathrm{d}\left(\frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_r}}{\partial x^{i'_r}} \mathrm{A}_{i_1 \ldots i_r}\right) \wedge \mathrm{d}x^{i'_1} \wedge \ldots \wedge \mathrm{d}x^{i'_r}$$

(using the fact that $\mathrm{d}^2 = 0$)

$$= \left(\frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_r}}{\partial x^{i'_r}} \mathrm{dA}_{i_1 \ldots i_r}\right) \wedge \mathrm{d}x^{i'_1} \wedge \ldots \wedge \mathrm{d}x^{i'_r}$$

$$+\frac{\partial^2 x^{i_1}}{\partial x^{i'_1}\partial x^{e'}}\frac{\partial x^{i_2}}{\partial x^{i'_2}}\cdots\frac{\partial x^{i_r}}{\partial x^{i'_r}}A_{i_1\ldots i_r}\,dx^{e'}\wedge dx^{i'_1}\wedge\ldots\wedge dx^{i'_r}$$

$$+\frac{\partial x^{i_1}}{\partial x^{i'_1}}\frac{\partial^2 x^{i_2}}{\partial x^{i'_2}\partial x^{e'}}\cdots\frac{\partial x^{i_r}}{\partial x^{i'_r}}A_{i_1\ldots i_r}\,dx^{i'_1}\wedge dx^{e'}\wedge dx^{i'_2}\wedge\ldots\quad\wedge dx^{i'_r}+\ldots\ldots+$$

$$+\frac{\partial x^{i_1}}{\partial x^{i'_1}}\frac{\partial x^{i_2}}{\partial x^{i'_2}}\cdots\frac{\partial^2 x^{i_r}}{\partial x^{i'_r}\partial x^{e'}}A_{i_1\ldots i_r}\,dx^{i'_1}\wedge dx^{i'_2}\wedge\ldots\wedge dx^{e'}\wedge dx^{i'_r}$$

$$=\frac{\partial x^{i_1}}{\partial x^{i'_1}}\cdots\frac{\partial x^{i_r}}{\partial x^{i'_r}}dA_{i_1\ldots i_r}\wedge dx^{i'_1}\wedge\ldots\wedge dx^{i'_r}$$

$$=dA_{i_1\ldots i_r}\wedge\left(\frac{\partial x^{i_1}}{\partial x^{i'_1}}dx^{i'_1}\right)\wedge\left(\frac{\partial x^{i_2}}{\partial x^{i'_2}}dx^{i'_2}\right)\wedge\ldots\wedge\left(\frac{\partial x^{i_r}}{\partial x^{i'_r}}dx^{i'_r}\right)$$

$$=dA_{i_1\ldots i_r}\wedge dx^{i_1}\wedge dx^{i_2}\wedge\ldots\wedge dx^{i_r}$$

(Since $\frac{\partial^2 x^{i_1}}{\partial x^{i'_1}\partial x^{e'}}dx^{e'}\wedge dx^{i'_1}=\sum_{i'=1}^{n}\sum_{e'=1}^{n}\frac{\partial^2 x^{i_1}}{\partial x^{i'_1}\partial x^{e'}}dx^{e'}\wedge dx^{i'_1}=0$, and similarly for all terms involving second order mixed partial derivatives.

For example, suppose $i'_1=1, e'=2$ and then

$i'_1=2, e'=1$, in the double sum given. Then $\frac{\partial^2 x^{i_1}}{\partial x^1\partial x^2}\left(\frac{dx^1\otimes dx^2-dx^2\otimes dx^1}{2}\right)$

$$+\frac{\partial^2 x^{i_1}}{\partial x^2\partial x^1}\left(\frac{dx^2\otimes dx^1-dx^1\otimes dx^2}{2}\right)=$$

$$=\frac{\partial^2 x^{i_1}}{\partial x^1\partial x^2}\left(\frac{dx^1\otimes dx^2-dx^2\otimes dx^1}{2}\right)-\frac{\partial^2 x^{i_1}}{\partial x^2\partial x^1}\left(\frac{dx^1\otimes dx^2-dx^2\otimes dx^1}{2}\right)=0$$

(since $\frac{\partial^2 x^{i_1}}{\partial x^1\partial x^2}=\frac{\partial^2 x^{i_1}}{\partial x^2\partial x^1}$)).

Now let $y\in T_p(M)\quad:\quad y=y^b\frac{\partial}{\partial x^b}$.

Consider $y^b_{,c}\equiv\frac{\partial y^b}{\partial x^c}$ (partial derivative of y^b with respect to x^c)

We will show that the components $y^b_{,c}$ do not transform as a tensor should when we change coordinates.

At each point of M select a tangent vector in such a way that we now have a function on M where the function's codomain is the tangent space associated with each point of M. They connect up in a nice way that we can differentiate them, but the problem is when we do that the components of what we get are not going to transform in a way that a tensor should transform when we change coordinates.

Specifically, $(y^{b'})_{,c'} = \left(\frac{\partial x^{b'}}{\partial x^b} y^b\right)_{,c'} = \left[\frac{\partial}{\partial x^c}\left(\frac{\partial x^{b'}}{\partial x^b} y^b\right)\right]\frac{\partial x^c}{\partial x^{c'}} = \left[\frac{\partial^2 x^{b'}}{\partial x^c \partial x^b} y^b + \frac{\partial y^b}{\partial x^c}\frac{\partial x^{b'}}{\partial x^b}\right]\frac{\partial x^c}{\partial x^{c'}}$

$$= \frac{\partial^2 x^{b'}}{\partial x^c \partial x^b}\frac{\partial x^c}{\partial x^{c'}} y^b + \frac{\partial x^{b'}}{\partial x^b} y^b_{,c}\frac{\partial x^c}{\partial x^{c'}},$$

and we can see that $y^{b'}_{,c'} \neq y^b_{,c}\frac{\partial x^{b'}}{\partial x^b}\frac{\partial x^c}{\partial x^{c'}}$.

Thus the conclusion would be that $y^b_{,c}$ are not the components of the tensor, and hence are not covariant . That is going to motivate us looking at another kind of derivative, the covariant derivative.

Covariant differentiation.

<u>Definition 14</u> : A connection ∇ (del) at a point p of M is a rule which assigns to each vector field X a differential operator ∇_X which maps any C^r (r continuous derivatives) vector field Y into a vector field $\nabla_X Y$ where:

1) $\nabla_X Y$ is a tensor in X , i.e.

   for all functions f, g and vectors X,Y,Z – C^1 vector fields,

   $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.

2) $\nabla_X Y$ is linear in Y (usual derivative rules)

   $\nabla_X(\alpha Y + \beta Z) = \alpha\nabla_X Y + \beta\nabla_X Z$

3) for any C^1 function f and C^1 vector field Y,

   $\nabla_X(fY) = X(f)Y + f\nabla_X Y$.

Then we call $\nabla_X Y$ the covariant derivative of Y with respect to ∇ in the direction of the vector field X. $\nabla_X Y$ is a vector field, which is a tensor field of type (1,0).

We can define ∇Y, the covariant derivative of Y, as a tensor of type (1,1), which means it is a tensor of a vector field and covector field. When the covector field part acts on the vector field X, we call this contraction of ∇Y with X.

Now define ∇Y, the covariant derivative of Y, as a tensor of type (1,1) which, when contracted with X, produces the vector $\nabla_X Y$. Then we have that

(3) holds if and only if $\nabla(fY) = df \otimes Y + f\nabla Y$.

(Since $f\nabla Y$ contracted with X gives us $f\nabla_X Y$,

$\nabla(fY)$ contracted with X gives us $\nabla_X (fY)$.

Now consider $df \otimes Y$, where df is a covector field, Y is a vector field.

$$df(X) = \frac{\partial f}{\partial x^i} dx^i(X) = \frac{\partial f}{\partial x^i} dx^i\left(a^j \frac{\partial}{\partial x^j}\right)$$

$$= \frac{\partial f}{\partial x^i} a^j \, dx^i\left(\frac{\partial}{\partial x^j}\right)$$

$$= \frac{\partial f}{\partial x^i} a^j \, \delta^i_j$$

$= a^i \frac{\partial f}{\partial x^i}$ (which is a real number when evaluated at some point p).

So when $df \otimes Y$ is contracted with X, we get $a^i \frac{\partial f}{\partial x^i} Y$, which is the same as $X(f)Y$, since $X(f) = a^i \frac{\partial f}{\partial x^i}$, where $X = a^i \frac{\partial}{\partial x^i}$).

Given any C^{r+1} vector basis $\frac{\partial}{\partial x^a}$ and dual one-form basis dx^a on a neighborhood U, we shall write the components of ∇Y as $Y^a_{;b}$, so $\nabla Y = Y^a_{;b} \, dx^b \otimes \frac{\partial}{\partial x^a}$.

The connection is determined on U by n^3 C^r functions, Γ^a_{bc}, defined by

$$\nabla \frac{\partial}{\partial x^c} = \Gamma^a_{bc} \, dx^b \otimes \frac{\partial}{\partial x^a}$$

For any C^1 vector field Y,

$$\nabla Y = \nabla\left(y^c \frac{\partial}{\partial x^c}\right) = dy^c \otimes \frac{\partial}{\partial x^c} + y^c \nabla\left(\frac{\partial}{\partial x^c}\right) = dy^c \otimes \frac{\partial}{\partial x^c} + y^c \Gamma^a_{bc}\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$$

$$= \frac{\partial y^c}{\partial x^b}\, dx^b \otimes \frac{\partial}{\partial x^c} + y^c \Gamma^a_{bc}\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$$

$$= Y^c_{,b}\, dx^b \otimes \frac{\partial}{\partial x^c} + y^c \Gamma^a_{bc}\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$$

$$= Y^a_{,b}\, dx^b \otimes \frac{\partial}{\partial x^a} + y^c \Gamma^a_{bc}\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$$ (since we are summing on c which is just a dummy index, we can replace c by a)

$$= \left[Y^a_{,b} + y^c \Gamma^a_{bc}\right]\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$$

Now let us compare $\nabla Y = Y^a_{;b}\, dx^b \otimes \frac{\partial}{\partial x^a}$ and $\nabla Y = \left[Y^a_{,b} + y^c \Gamma^a_{bc}\right]\left(dx^b \otimes \frac{\partial}{\partial x^a}\right)$.

The components of ∇Y with respect to coordinate basis $\frac{\partial}{\partial x^a}$ and dx^b are

$$Y^a_{;b} = Y^a_{,b} + Y^c \Gamma^a_{bc},$$

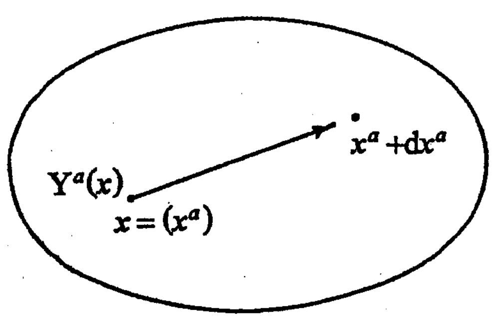
where $a, b, c = 1, \ldots, n$.

And $Y^a_{;b}$ transforms as the components of a tensor should transform, under a change in coordinates.

Suppose we have a manifold and a vector field. So for each point on a M we have a vector.

In other words, $Y^a(x_0) \frac{\partial}{\partial x^a} |_{x_0} \in T_{x_0}(M)$

Figure 17.



Then Y^a is a vector field on M. At each point we select a vector from the tangent space at that point, and we put them together in such a way that if we move from point to point we get a function.

Let us suppose Y^a are the differentiable components of a vector field, and x on M has coordinates x^a in some coordinate neighborhood . Let us take a point $x^a + dx^a$, where dx^a is a small change in each of the coordinates. dx^a can be thought of as a vector going from one point to another. If a vector field is differentiable, we can express the value of $Y^a(x^a + dx^a)$ in terms of $Y^a(x)$, using a Taylor expansion.

We want to write $Y^a(x^a + dx^a) \approx Y^a(x^a) + Y^a_{,b} dx^b$ (approximation to first order; two terms of Taylor expansion).

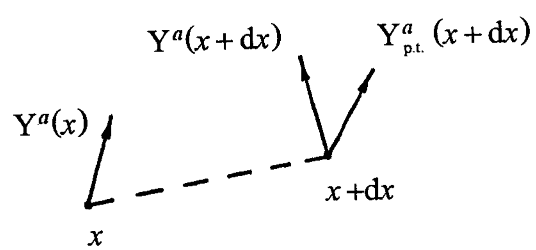
As we change from point to point we are going to have different values of a vector field. Now we are going to introduce a very important concept: **parallel transfer** (transport). We are going to define a new vector field $Y^a_{\text{p.t.}}$ at the new point $x^a + dx^a$ as follows:

__**Definition 15**__ : $Y^a_{\text{p.t.}}(x^a + dx^a) = Y^a(x^a) - \Gamma^a_{bc}(x^a)\, Y^c(x^a)\, dx^b$, $a, b, c = 1, \ldots, n$

(we have n^3 given functions, Γ^a_{bc}, or $Y^a_{\text{p.t.}}(x + dx) = Y^a(x) - \Gamma^a_{bc}(x)\, Y^c(x^a)\, dx^b$)

In this interpretation the Γ^a_{bc} define the parallel transport.

Figure 18.



We wrote Y^a as the components of a vector in a tangent space. We can think of Y^a as either the components of a vector in a tangent space or the components of a vector in the manifold, since the tangent space at a point has the same dimension as the manifold and a manifold is locally homeomorphic to $\mathbb{R}^n$. So we can think of Y^a as being in a manifold instead of being in a tangent space. Hence if we draw a tangent vector in a tangent space it will have "n" components and dx will give us a direction.

In Euclidean space the connection terms are zero (because we have the same vector in terms of components (vectors are parallel and have the same length)). It is true if you are in $\mathbb{R}^n$ in Euclidean space. Euclidean space is flat (zero curvature).

Now what is the difference between $Y^a(x+dx)$ and $Y^a_{\text{p.t.}}(x+dx)$?

This will give us a formula for of $Y^a_{;b}$.

So we need to find

$$Y^a(x+dx) - Y^a_{\text{p.t.}}(x+dx) = Y^a(x) + Y^a_{,b}dx^b - Y^a(x) + \Gamma^a_{bc}(x)Y^c(x)dx^b$$

$$= Y^a_{,b}dx^b + \Gamma^a_{bc}(x)Y^c(x)dx^b$$

$$= \left[Y^a_{,b} + \Gamma^a_{bc}(x)Y^c\right] dx^b.$$

Put $\left[Y^a_{,b} + \Gamma^a_{bc}(x)Y^c\right] dx^b = DY^a$. Then $Y^a_{,b} + \Gamma^a_{bc}(x)Y^c = \frac{DY^a}{dx^b}$.

$Y^a_{;b} = \frac{DY^a}{dx^b}$ is called **the total or absolute derivative.**

Now let $\omega = \omega_a dx^a$, and find an expression for $\omega_{a\,;\,c}$.

We start with any Y^a components of an arbitrary vector and consider

$(Y^a\omega_a)_{;b} = (Y^a\omega_a)_{,b}$.

[Y^a - components of a vector ; ω_a - components of a covector ; Y^a and ω_a are real valued functions, a=1,...,n

note : " ; " on a real valued function is the ordinary partial derivative .

The definition of a covariant derivative can be extended to any C^r tensor field if $r \geq 1$.

One of the rules is $\nabla f = df$.(see definition of "d" on pg 40)

So $(Y^a\omega_a)_{;b} = (Y^a\omega_a)_{,b}$] .

Consider $(Y^a\omega_a)_{;b} = Y^a_{;b}\,\omega_a + Y^a\,\omega_{a\,;\,b} = Y^a_{,b}\,\omega_a + Y^c\,\Gamma^a_{bc}\,\omega_a + Y^a\,\omega_{a\,;\,b}$

and $(Y^a\omega_a)_{,b} = Y^a_{,b}\omega_a + Y^a\omega_{a,b}$

$\Rightarrow \quad Y^a_{,b}\,\omega_a + Y^c\,\Gamma^a_{bc}\,\omega_a + Y^a\,\omega_{a\,;\,b} = Y^a_{,b}\omega_a + Y^a\omega_{a,b}$

$\Rightarrow \quad Y^a\omega_{a\,;\,b} = Y^a\omega_{a,b} - Y^c\Gamma^a_{bc}\omega_a$ or $Y^c\omega_{c\,;\,b} = Y^c\omega_{c,b} - Y^c\Gamma^a_{bc}\omega_a$.

So $Y^c\omega_{c\,;\,b} - Y^c\omega_{c,b} + Y^c\Gamma^a_{bc}\omega_a = 0$

$Y^c[\omega_{c\,;\,b} - \omega_{c,b} + \Gamma^a_{bc}\omega_a] = 0$ for any Y^c.

$\Rightarrow [\omega_{c\,;\,b} - \omega_{c,b} + \Gamma^a_{bc}\omega_a] = 0$

$\Rightarrow \omega_{c\,;\,b} = \omega_{c,b} - \Gamma^a_{bc}\omega_a$

Now we will write down the formula for the covariant derivative of the components of a general tensor.

Suppose the tensor is called T and suppose the components (with respect to a basis) are given by

$T^{b....e..}{}_{c....l..;\,a}$ ($b....e....$- "upstairs" indices , $c....l....$ - "downstairs" indices).

$T^{b...e..}{}_{c...l..;\,a} = T^{b...e..}{}_{c...l..,\,a} + \Gamma^e_{af}\,T^{b..f..}{}_{c...l..} + ... +$ similarly for each upstairs index

$- \Gamma^f_{al}\,T^{b...e..}{}_{c..f..} - ... -$ similarly for each downstairs index.

Suppose we start with the components of a vector. We would like to take $\left(\lambda^a_{;b}\right)_{;c}$ or $\lambda^a_{;bc}$ (take covariant derivative with respect to x^b and then take covariant derivative of that answer with respect to x^c) and take $\left(\lambda^a_{;c}\right)_{;b}$ .Then find $\left(\lambda^a_{;b}\right)_{;c} - \left(\lambda^a_{;c}\right)_{;b}$ (it is not necessarily 0).

The expression we get leads us to the **components of the Riemann curvature tensor**

$$\left(\lambda^a_{;b}\right)_{;c} = \left(\lambda^a_{,b} + \lambda^k\Gamma^a_{bk}\right)_{;c} = \left(\lambda^a_{,b} + \lambda^k\Gamma^a_{bk}\right)_{,c} + \Gamma^a_{cf}\left(\lambda^f_{,b} + \lambda^k\Gamma^f_{bk}\right) - \Gamma^f_{cb}\left(\lambda^a_{,f} + \lambda^k\Gamma^a_{fk}\right)$$

$= \lambda^a_{,bc} + \lambda^k_{,c}\,\Gamma^a_{bk} + \lambda^k\,\Gamma^a_{bk,c} + \Gamma^a_{cf}\,\lambda^f_{,b} + \Gamma^a_{cf}\,\lambda^k\,\Gamma^f_{bk} - \Gamma^f_{cb}\,\lambda^a_{,f} - \Gamma^f_{cb}\,\lambda^k\,\Gamma^a_{fk}$

Next find $\left(\lambda^a_{;c}\right)_{;b} = \lambda^a_{,cb} + \lambda^k_{,b}\,\Gamma^a_{ck} + \lambda^k\,\Gamma^a_{ck,b} + \Gamma^a_{bf}\,\lambda^f_{,c} + \Gamma^a_{bf}\,\lambda^k\,\Gamma^f_{ck} - \Gamma^f_{bc}\,\lambda^a_{,f} - \Gamma^f_{bc}\,\lambda^k\,\Gamma^a_{fk}$.

(This second expression can be easily found from the first by simply interchanging the b and c).

So $\lambda^a_{;bc} - \lambda^a_{;cb} = \lambda^a_{,bc} + \lambda^k_{,c}\,\Gamma^a_{bk} + \lambda^k\,\Gamma^a_{bk,c} + \Gamma^a_{cf}\,\lambda^f_{,b} + \Gamma^a_{cf}\,\lambda^k\,\Gamma^f_{bk} - \Gamma^f_{cb}\,\lambda^a_{f} - \Gamma^f_{cb}\,\lambda^k\,\Gamma^a_{fk}$

$$-\lambda^a_{,cb} - \lambda^k_{,b}\,\Gamma^a_{ck} - \lambda^k\,\Gamma^a_{ck,b} - \Gamma^a_{bf}\,\lambda^f_{,c} - \Gamma^a_{bf}\,\lambda^k\,\Gamma^f_{ck} + \Gamma^f_{bc}\,\lambda^a_{f} + \Gamma^f_{bc}\,\lambda^k\,\Gamma^a_{fk}$$

(in $\Gamma^a_{cf}\,\lambda^f_{,b}$ and $\lambda^k_{,b}\,\Gamma^a_{ck}$, $\lambda^k_{,c}\,\Gamma^a_{bk}$ and $\Gamma^a_{bf}\,\lambda^f_{,c}$ the indices k and f are dummy , so replace k by f)

$$= \lambda^k\left[\Gamma^a_{bk,c} + \Gamma^a_{cf}\Gamma^f_{bk} - \Gamma^a_{ck,b} - \Gamma^a_{bf}\Gamma^f_{ck}\right] + \left(\lambda^a_{f} + \lambda^k\,\Gamma^a_{fk}\right)(\Gamma^f_{bc} - \Gamma^f_{cb})$$

We will deal only with torsion-free connections , i.e. we will assume $T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj} = 0$, where this is the torsion tensor. In this case , the coordinate components of the connection obey $\Gamma^i_{jk} = \Gamma^i_{kj}$, so such a connection is often called a **symmetric connection.**

(In physics we use that assumption.)

So $(\Gamma^f_{bc} - \Gamma^f_{cb}) = 0$

Now replace k by d and get :

$$2\lambda^a_{;[bc]} = \lambda^a_{;bc} - \lambda^a_{;cb} = \lambda^d\left[\Gamma^a_{bd,c} + \Gamma^a_{cf}\Gamma^f_{bd} - \Gamma^a_{cd,b} - \Gamma^a_{bf}\Gamma^f_{cd}\right] = \lambda^d\,R^a_{dcb}$$, where R^a_{dcb} is called the **Riemann** (curvature) **tensor.**

So $\lambda^a_{;bc} - \lambda^a_{;cb} = \lambda^d\,R^a_{dcb}$.

R_{dcb} can be represented in terms of the coordinate components of the connection.

We can define $z^a_{;dc} - z^a_{;cd} = R^a_{bcd}\,z^b = \left(\Gamma^a_{bd,c} - \Gamma^a_{bc,d} - (\Gamma^e_{bc}\,\Gamma^a_{de} - \Gamma^e_{bd}\,\Gamma^a_{ce})\right) z^b$.

Note: this is skew-symmetric in c and d.

So we can define $R^a_b = R^a_{bcd}\, dx^c \wedge dx^d$ to be the curvature two-form and define

$\Gamma^a_b = \Gamma^a_{bd}\, dx^d$ as a 1-form representing n^3 functions, when dim M = n.

We would like to find the exterior derivative $d\Gamma^a_b$ and wedge-product of the two

one-forms $\Gamma^e_b \wedge \Gamma^a_e$ and then consider $2[d\Gamma^a_b - \Gamma^e_b \wedge \Gamma^a_e]$ and compare what we get with

the expression for $R^a_b = R^a_{bcd}\, dx^c \wedge dx^d$.

So find $d\Gamma^a_b = d(\Gamma^a_{bd}\, dx^d) = \Gamma^a_{bd,c} dx^c \wedge dx^d = \Gamma^a_{b[d,c]} dx^c \wedge dx^d$ (skew-symmetric on c and d)

Now find $\Gamma^e_b \wedge \Gamma^a_e = \Gamma^e_{bc} dx^c \wedge \Gamma^a_{de} dx^d = \Gamma^e_{bc} \Gamma^a_{de} dx^c \wedge dx^d = \Gamma^e_{b[c}\, \Gamma^a_{d]e} dx^c \wedge dx^d$.

Find $2[d\Gamma^a_b - \Gamma^e_b \wedge \Gamma^a_e] = 2\left[\Gamma^a_{b[d,c]}\, dx^c \wedge dx^d \;-\; \Gamma^e_{b[c}\;\Gamma^a_{d]e}\, dx^c \wedge dx^d\right]$

$$= 2\left[\Gamma^a_{b[d,c]} \;-\; \Gamma^e_{b[c}\;\Gamma^a_{d]e}\right] dx^c \wedge dx^d$$

$$= 2\left[\tfrac{1}{2}\left(\Gamma^a_{bd,c} - \Gamma^a_{bc,d}\right) \;-\; \tfrac{1}{2}\left(\Gamma^e_{bc}\,\Gamma^a_{de} \;-\; \Gamma^e_{bd}\,\Gamma^a_{ce}\right)\right] dx^c \wedge dx^d$$

$$= \left[\left(\Gamma^a_{bd,c} - \Gamma^a_{bc,d}\right) \;-\; \left(\Gamma^e_{bc}\,\Gamma^a_{de} \;-\; \Gamma^e_{bd}\,\Gamma^a_{ce}\right)\right] dx^c \wedge dx^d$$

$$= R^a_{bcd}\, dx^c \wedge dx^d = R^a_b$$

So we can conclude that

$$R^a_b = 2\,[d\Gamma^a_b \;-\; \Gamma^e_b \wedge \Gamma^a_e] \qquad (5)$$

Fibre bundles.

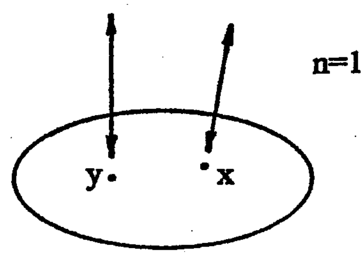
We will find it useful to examine a concept of fibre bundles since these are used in some applications of mathematical physics. We can construct a manifold M called a **fibre bundle** which is a direct product of M and a suitable space. We start with a manifold M and take its Cartesian product with $\mathbb{R}^n$: $\quad M \times \mathbb{R}^n = E$. $\;$ dim $M = 4$, dim $\mathbb{R}^n = n \geq 1$.

In a special case when $n = 1$ we sometimes call this a "line bundle".

We have a manifold and at each point of the manifold there is a line attached to it, because the point in E is described by specifying a point of the manifold together with a point in $\mathbb{R}^n$. But if $n = 1$, then points are in $\mathbb{R}^1$ which are specified by giving a real number.

For example, $M \times \mathbb{R}^1 = \{(x,a) : x \in M, a \in \mathbb{R}^1\}$.

Figure 19.



So, $E = \{(x, a_1, ..., a_n) : x \in M, a_i \in \mathbb{R}\}$.

$\Pi : E \rightarrow M$, is a projection that takes us from one of the points of E and maps us down to the point of M that it is attached to . That is why M is called the base space. This mapping is not one-to-one because all points on a line get mapped to the same point x.

Given $p \in M$, $\Pi : E \rightarrow M$, we define $\Pi^{-1}(p) = \{z \in E : \Pi\{z\} = p\}$

A C^k bundle over a C^s $(s \geq k)$ manifold M is a C^k manifold E and a C^k surjective map $\Pi : E \rightarrow M$ .The manifold E is called the total space , M is called the base space, and Π, the projection.

The simplest example of a bundle is a product bundle $(M \times A, M, \Pi)$, where A is some manifold and the projection Π is defined by $\Pi(p,v) = p$ for all $p \in M$, $v \in A$.

For example, if one chooses M as the circle S^1 and A is the real line R^1 , one constructs the cylinder C^2 as a product bundle over S^1 .

In this thesis we are mostly concerned with the mathematical-physics application where the base is 4-dimensional space-time and A is $\mathbb{R}^n$.

The metric.

From geometry, if we are in 3-dimensional space and want to find a distance, we are going to have

$$\Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \qquad (\Delta s^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)$$

It turns out that we can recognize Δs^2 as being like a product of a matrix with 2 vectors :

$$\begin{pmatrix} \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = \Delta s^2$$

Now consider $\mathbb{R}^4$, $(t,x,y,z) = (x^0, x^1, x^2, x^3)$;

t is going to be treated a little bit differently than x, y, z . It turns out that the distance is replaced with the concept of the interval, and is going to be

$(\Delta t)^2 - ((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)$.

(If we have two people - one is at a certain space and time, another is at a certain space and time - then we have the interval between them.)

We have the matrix $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ and

$$(\Delta t, \Delta x, \Delta y, \Delta z) A \begin{pmatrix} \Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = (\Delta t)^2 - ((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)$$

The matrix A is called a **metric** in flat 4-dimensional space-time. That 4-dimensional space-time has a special name, **Minkowski space**, and the metric is not positive definite; that is, it is possible for two different points in our space-time to have a zero interval between them . That can happen when

$$(\Delta t)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2.$$

Now consider (x_0, x_1, x_2, x_3), where $x_0 = ct$ and c (speed of light) is constant.

The speed (velocity) of light is independent of the motion of the source . For example, suppose I shine a flashlight at you and we are standing still relative to each other, and suppose you measure how fast the light is coming at you from that flashlight. Now suppose I am running toward you with flashlight and you again measure the speed of the light coming at you. Then the two speeds will be the same. But this is not true (for example) about sound.

Now we going to consider how to define the components of a tensor,

$$\partial_a = g_{a\beta}\, \partial^\beta = \sum_{\beta=0}^{3} g_{a\beta}\, \partial^\beta \quad .$$

We can think of $g_{a\beta}\, \partial^\beta$ as $g_{a\beta}$ operating on the components of the vector ∂^β. But there is a free index , a , and so we can think of $g_{a\beta}\, \partial^\beta$ as components of a covector in the following way :

first write $g_{a\beta}$ as the components of a tensor

$$g = g_{a\beta}\, \mathrm{d}x^a \otimes \mathrm{d}x^\beta = g_{a\beta}\, (\, \tfrac{1}{2}(\mathrm{d}x^a \otimes \mathrm{d}x^\beta + \mathrm{d}x^\beta \otimes \mathrm{d}x^a))$$

(i.e.,we know that $g_{a\beta} = g_{\beta a}$ (symmetric) is given),

and think of $g_{a\beta}\, \partial^\beta$ as operating on vectors, to produce a real number answer.

To see this , let x have components $x = (x^a)$ and

let y have components $y = (y^\beta)$.

Then $g(x, y)$ is a real number.

This is defined by taking a pair of vectors from the tangent space and giving us a number

($g : T_p \times T_p \to \mathbb{R}$).

Now suppose we have $g(\_,y) : T_p \to \mathbb{R}$. This mapping now has only one argument.

So it is a mapping from $T_p \to \mathbb{R}$ which is linear in the argument. Such mappings are covectors.

So this mapping can be identified with a covector (an element of the dual space of T_p, which is T_p^*).

The question is what should we call this element of T_p^* ? Every element can be expressed in terms of a basis : $\omega_a \, dx^a$. And we are going to identify ω_a with the symbol ∂_a . Thus we are defining the components of a tensor by the formula

$\partial_a = g_{a\beta} \, \partial^\beta$.

Now in our case , $g_{a\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, $\partial^\beta = \begin{pmatrix} \partial^0 \\ \partial^1 \\ \partial^2 \\ \partial^3 \end{pmatrix}$.

So $g_{a\beta} \, \partial^\beta = \begin{pmatrix} \partial^0 \\ -\partial^1 \\ -\partial^2 \\ -\partial^3 \end{pmatrix}$.

But $\partial^1 = -\frac{\partial}{\partial x_1}$ (by $\partial^a = \left(\frac{\partial}{\partial x_0}, -\nabla\right)$), where ∇ is the usual 3-dimensional gradient operator.

$\partial^2 = -\frac{\partial}{\partial x_2}$

$\partial^3 = -\frac{\partial}{\partial x_3}$.

So $\partial_0 = \partial^0$

$\partial_1 = -\partial^1 = \frac{\partial}{\partial x}$

$\partial_2 = -\partial^2 = \frac{\partial}{\partial y}$

$\partial_3 = -\partial^3 = \frac{\partial}{\partial z}$.

Application to Electrodynamics.

Using this notation, we will now illustrate the covariance of electrodynamics by casting Maxwell's equations in tensor form.

First, the electromagnetic fields E and B are expressed in terms of the potentials as

$$E = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \nabla\phi \qquad \text{- electric field}$$

$$B = \nabla \times \vec{A} \qquad \text{- magnetic field.}$$

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \nabla\phi = -\frac{1}{c}\,c\,\frac{\partial \vec{A}}{\partial x_0} - \nabla\phi = -\frac{\partial \vec{A}}{\partial x_0} - \nabla\phi$$

because $\frac{\partial \vec{A}}{\partial t} = \frac{\partial \vec{A}}{\partial x_0}\frac{dx^0}{dt} = c\frac{\partial \vec{A}}{\partial x_0}$

The potentials ϕ (a scalar function) and $\vec{A}$ (a 3-vector) form a 4-vector potential

$$A^a = (\phi, A) = \left(\underbrace{A_0}_{\phi}, \underbrace{A_1, A_2, A_3}_{A}\right), \text{ where } A_0 = \phi \text{ is the time part.}$$

1) Define $\partial^a = \left(\frac{\partial}{\partial x_0}, -\nabla\right) = \left(\frac{\partial}{\partial x_0}, -\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}, -\frac{\partial}{\partial x_3}\right)$ and write down the x, y, z components of E and B :

$$E_x = -\frac{\partial A_1}{\partial x_0} - (\nabla\phi)_x = -\frac{\partial A_1}{\partial x_0} - \frac{\partial A_0}{\partial x_1} = -\partial^0 A^1 + \partial^1 A^0 = -(\partial^0 A^1 - \partial^1 A^0)$$

$$E_y = -\frac{\partial A_2}{\partial x_0} - (\nabla\phi)_y = -\frac{\partial A_2}{\partial x_0} - \frac{\partial A_0}{\partial x_2} = -\partial^0 A^2 + \partial^2 A^0 = -(\partial^0 A^2 - \partial^2 A^0)$$

$$E_z = -\frac{\partial A_3}{\partial x_0} - (\nabla\phi)_z = -\frac{\partial A_3}{\partial x_0} - \frac{\partial A_0}{\partial x_3} = -\partial^0 A^3 + \partial^3 A^0 = -(\partial^0 A^3 - \partial^3 A^0)$$

So $E_x = -(\partial^0 A^1 - \partial^1 A^0)$

$E_y = -(\partial^0 A^2 - \partial^2 A^0)$

$E_z = -(\partial^0 A^3 - \partial^3 A^0)$

$$B = \nabla \times \vec{A} = \det\begin{bmatrix} i & j & k \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ A_1 & A_2 & A_3 \end{bmatrix} = i\left(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}\right) - j\left(\frac{\partial A_3}{\partial x_1} - \frac{\partial A_1}{\partial x_3}\right) + k\left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right)$$

$$B_x = \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} = -\partial^2 A^3 + \partial^3 A^2 = -(\partial^2 A^3 - \partial^3 A^2)$$

$$B_y = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} = -\partial^3 A^1 + \partial^1 A^3 = -(\partial^3 A^1 - \partial^1 A^3)$$

$$B_z = \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} = -\partial^1 A^2 + \partial^2 A^1 = -(\partial^1 A^2 - \partial^2 A^1)$$

So $B_x = -(\partial^2 A^3 - \partial^3 A^2)$

$B_y = -(\partial^3 A^1 - \partial^1 A^3)$

$B_z = -(\partial^1 A^2 - \partial^2 A^1)$

These equations imply that the electric and magnetic fields, six components in all, are the elements of a second-rank, anti symmetric field-strength tensor:

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha$$

Explicitly, the field-strength tensor is, in matrix form,

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \quad , \quad \alpha, \beta = 0, 1, 2, 3.$$

Indeed , $F^{00} = 0$

$F^{01} = \partial^0 A^1 - \partial^1 A^0 = -E_x$

$F^{02} = \partial^0 A^2 - \partial^2 A^0 = -E_y$

$F^{03} = \partial^0 A^3 - \partial^3 A^0 = -E_z$

$F^{13} = \partial^1 A^3 - \partial^3 A^1 = B_y$

$F^{20} = \partial^2 A^0 - \partial^0 A^2 = E_y$

$F^{21} = \partial^2 A^1 - \partial^1 A^2 = B_z$

$F^{10} = \partial^1 A^0 - \partial^0 A^1 = E_x$

$F^{11} = \partial^1 A^1 - \partial^1 A^1 = 0$

$F^{12} = \partial^1 A^2 - \partial^2 A^1 = -B_z$

$F^{30} = \partial^3 A^0 - \partial^0 A^3 = E_z$

$F^{31} = \partial^3 A^1 - \partial^1 A^3 = -B_y$

$F^{32} = \partial^3 A^2 - \partial^2 A^3 = B_x$

$F^{33} = 0$

$$F^{22} = 0$$

$$F^{23} = \partial^2 A^3 - \partial^3 A^2 = -B_x$$

For reference, we record the field-strength tensor with two covariant indices,

$$F_{\alpha\beta} = g_{\alpha\gamma}\, F^{\gamma\delta}\, g_{\delta\beta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \qquad (6)$$

and in fact $g_{\alpha\gamma}\, F^{\gamma\delta}\, g_{\delta\beta}$ is an ordinary matrix product.

In particular, the expression, $g_{\alpha\gamma}\, F^{\gamma\delta}\, g_{\delta\beta}$, represents the matrix product

$\sum_{s=0}^{3}\left(\sum_{r=0}^{3} g_{ir}F_{rs}\right)g_{sj}$ (the inner summation is the (i,s) entry, the outer summation is the

(i,j) entry).

Explicitly, let $g_{\alpha\gamma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. Then

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

We note that the elements of $F_{\alpha\beta}$ are obtained from $F^{\alpha\beta}$ by putting $E \rightarrow -E$.

Another useful quantity is the dual field-strength tensor $F^{\alpha\beta}$. We first define the totally anti-symmetric fourth rank tensor $\in^{\alpha\beta\gamma\delta} = \begin{cases} +1 \text{ , for } \alpha = 0, \beta = 1, \gamma = 2, \delta = 3, \\ \quad \text{and any even permutation} \\ -1 \text{ , for any odd permutation} \\ 0 \text{ , if any two indices are equal} \end{cases}$

Note : $\in_{\alpha\beta\gamma\delta} = -\in^{\alpha\beta\gamma\delta}$.

The dual field-strength tensor is defined by

$$F^{\star\alpha\beta} = \frac{1}{2}\in^{\alpha\beta\gamma\delta} F_{\gamma\delta} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix}$$

For example, $F^{\star 00} = \frac{1}{2}\in^{00\gamma\delta} F_{\gamma\delta} = 0$

$$F^{\star 01} = \frac{1}{2}\in^{01\gamma\delta} F_{\gamma\delta} = \frac{1}{2}\sum_{\gamma=2}^{3}\sum_{\delta=2}^{3} \in^{01\gamma\delta} F_{\gamma\delta} = \frac{1}{2}\in^{0123}F_{23} + \frac{1}{2}\in^{0132}F_{32} = \frac{1}{2}F_{23} - \frac{1}{2}F_{32} =$$

$$= -\frac{1}{2}B_x - \frac{1}{2}B_x = -B_x$$

$$F^{\star 02} = \frac{1}{2}\in^{02\gamma\delta} F_{\gamma\delta} = \frac{1}{2}\sum_{\gamma\neq 0,2}\sum_{\delta\neq 0,2} \in^{02\gamma\delta}F_{\gamma\delta} = \frac{1}{2}\in^{0213}F_{13} + \frac{1}{2}\in^{0231}F_{31} = -\frac{1}{2}F_{13} + \frac{1}{2}F_{31} =$$

$$= -\frac{1}{2}B_y - \frac{1}{2}B_y = -B_y$$

Analogously we can do the rest.

But what does it mean physically? The elements of the dual tensor $F^{\star\alpha\beta}$are obtained from

$F^{\alpha\beta}$ by putting E→B and B→ −E in $F^{\alpha\beta}$ (physically we changed fields).

Indeed , put E→B and B→ −E in $F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}$, and we would have

$$\begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix}.$$

Every one of the components of the electric and magnetic fields is included in $F^{\star\alpha\beta}$.

It is important because we want to put all information about our fields into a single object called the field-strength or dual field-strength tensor.

We are trying to come up with a more compact version of the equations, and we want to show that we can write them in terms of tensors, so that we know it is covariant in the sense that if we change the coordinates, the laws of electricity and magnetism are the same no matter what our reference frame is.

So we must write the Maxwell equations themselves in an explicitly covariant form.

The inhomogeneous equations are

$$\nabla \cdot \vec{E} = 4\pi\rho\ , \qquad (7)$$

$$\nabla \times \vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} = \frac{4\pi}{c}\vec{J}\ . \qquad (8)$$

These two equations lead us to an equation for $F^{\alpha\beta}$ and we can write them in terms of this field-strength tensor . So in terms of $F^{\alpha\beta}$ and the 4-current $J^{\alpha} = (c\rho, \vec{J})$ these take on the covariant form,

$$\partial_{\alpha}F^{\alpha\beta} = \frac{4\pi}{c}J^{\beta}\ . \qquad (9)$$

Indeed , for $\beta = 0$, $\partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = \frac{\partial}{\partial x}E_x + \frac{\partial}{\partial y}E_y + \frac{\partial}{\partial z}E_z = \nabla \cdot E$,

for $\beta = 1$, $\partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31} = \frac{1}{c}\frac{\partial}{\partial t}(-E_x) + \frac{\partial}{\partial y}B_z + \frac{\partial}{\partial z}(-B_y) =$

$-\frac{1}{c}\frac{\partial E_x}{\partial t} + \frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}$ where $\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}$ is an x -component of a curl,

for $\beta = 2$, $\partial_0 F^{02} + \partial_1 F^{12} + \partial_2 F^{22} + \partial_3 F^{32} = \frac{1}{c}\frac{\partial}{\partial t}(-E_y) + \frac{\partial}{\partial y}(-B_z) + \frac{\partial}{\partial z}(B_x) =$

$-\frac{1}{c}\frac{\partial E_y}{\partial t} - \frac{\partial B_z}{\partial x} + \frac{\partial B_x}{\partial z} = -\frac{1}{c}\frac{\partial E_y}{\partial t} + \frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}$

where $\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}$ is a y- component of a curl,

and for $\beta = 3$, $\partial_0 F^{03} + \partial_1 F^{13} + \partial_2 F^{23} + \partial_3 F^{33} = \frac{1}{c}\frac{\partial}{\partial t}(-E_z) + \frac{\partial}{\partial x}B_y + \frac{\partial}{\partial z}(-B_x) =$

$-\frac{1}{c}\frac{\partial E_z}{\partial t}+\frac{\partial B_y}{\partial x}-\frac{\partial B_x}{\partial y}$ where $\frac{\partial B_y}{\partial x}-\frac{\partial B_x}{\partial v}$ is a z -component of a curl.

Similarly, the homogeneous Maxwell equations

$$\nabla\cdot\vec{B}=0 \qquad (10)$$

$$\nabla\times\vec{E}+\frac{1}{c}\frac{\partial\vec{B}}{\partial t}=0\,, \qquad (11)$$

can be written in terms of the dual field-strength tensor as

$$\partial_\alpha F^{\star\alpha\beta}=0\,.$$

So these four equations (7, 8, 10, 11) can be replaced by a pair of equations which are written in a tensor form.

In terms of $F^{\alpha\beta}$, rather than $F^{\star\alpha\beta}$, these homogeneous equations are the four equations,

$$\partial^\alpha F^{\beta\gamma}+\partial^\beta F^{\gamma\alpha}+\partial^\gamma F^{\alpha\beta}=0 \qquad (12)$$

where α,β,γ are any three of the integers 0, 1, 2, 3.

Let us consider

$$\partial_\alpha F^{\star\alpha\beta}=0. \qquad (13)$$

Indeed, for $\beta=0$ we have $-\partial^1B_x-\partial^2B_y-\partial^3B_z=0;$

for $\beta=1$, $\quad -\partial^0B_x+\partial^2E_z-\partial^3E_y=0;$

for $\beta=2$, $\quad -\partial^0B_y-\partial^1E_z+\partial^3E_x=0;$

for $\beta=3$, $\quad -\partial^0B_z+\partial^1E_y-\partial^2E_x=0.$

Now consider (12) : for $\alpha,\beta,\gamma=0,1,2,3$ we are going to have 64 equations. These equations can be reduced to 12 equations:

$$-\partial^0B_z+\partial^1E_y-\partial^2E_x=0,$$

$$-\partial^0B_y+\partial^3E_x-\partial^1E_z=0,$$

$$-\partial^0B_x+\partial^2E_z-\partial^3E_y=0,$$

$$-\partial^1 E_y + \partial^0 B_z + \partial^2 E_x = 0,$$

$$-\partial^1 E_z - \partial^0 B_y + \partial^3 E_x = 0,$$

$$-\partial^1 B_x - \partial^2 B_y - \partial^3 B_z = 0,$$

$$-\partial^2 E_x - \partial^0 B_z + \partial^1 E_y = 0,$$

$$-\partial^2 E_z + \partial^0 B_x + \partial^3 E_y = 0,$$

$$\partial^2 B_y + \partial^1 B_x + \partial^3 B_z = 0,$$

$$-\partial^3 E_x + \partial^0 B_y + \partial^1 E_z = 0,$$

$$-\partial^3 E_y - \partial^0 B_x + \partial^2 E_z = 0,$$

$$-\partial^3 B_z - \partial^1 B_x - \partial^2 B_y = 0.$$

When we look closer, we see that these 12 equations can be reduced to the 4 equations in (13) we are looking for.

Thus (12) and (13) are equivalent.

These four equations are the Bianchi identities for $F^{\alpha\beta}$.

Applications to Yang-Mills Field Theory.

On Minkowski space M we will consider the vector bundle B (each fiber being an n-complex-dimensional vector space), i.e. B=M×C^n.

The global vector fields e_A (vector-valued functions of $x^a, a = 0, 1, 2, 3, A = 1, ..., n$) form a basis set as does

$e'_A = G^B_A \, e_B$, where G^B_A is a non-singular matrix-valued function on M.

The connection or parallel transfer of vectors is introduced by defining ∇_a by

$$\nabla_a e_A = \gamma^B_{Aa} \, e_B , \qquad (14)$$

with $\gamma^B_A = \gamma^B_{Aa} \, dx^a$ being the connection (matrix-valued) one-form

(γ^B_{Aa} is a matrix and is a component of a one-form with respect to a basis dx^a).

By the definition of covariant derivative, the covariant derivative of a vector field is a vector field, so $\nabla_a e_A$ is a vector field. Thus we can express $\nabla_a e_A$ as a linear combination of a basis e_B, and for each e_A we will get different linear combinations of e_B since $a = 0, 1, 2, 3,$ thus leading to (14).

Now suppose we have an arbitrary vector $V = V^A e_A$. We can define the covariant derivative of an arbitrary vector V by

$$\nabla_a V = (V^A{}_{,a} + V^B \gamma^A_{Ba}) e_A , \qquad (15)$$

with a comma denoting the partial derivatives with respect to the Minkowski coordinates x^a.

Indeed, $\nabla_a V = \nabla_a (V^A e_A) = V^A{}_{,a}\, e_A + V^A \nabla_a e_A$

$$= V^A{}_{,a}\, e_A + V^A \gamma^B_{Aa}\, e_B$$

$$= V^A{}_{,a}\, e_A + V^B \gamma^A_{Ba}\, e_A$$

$$= (V^A{}_{,a} + V^B \gamma^A_{Ba})\, e_A ,$$

which establishes (15).

We will be interested in examining how the connection and other related quantities transform if we choose a different basis labeled by e'_A , $A = 1, \ldots, n$. In other words we can rewrite (14) as

$$\nabla_a e'_A = \gamma'^B_{Aa}\, e'_B , \qquad (16)$$

where γ'^B_{Aa} are the new connection components when we change to the new basis.

For given $e'_A = G^B_A\, e_B$ we can find $e_B = G^{-1A}_B\, e'_A$, then $\nabla_a e'_A = \nabla_a (G_A^B e_B) = G_{A,a}^B e_B + G_A^B \nabla_a e_B$

$$= G_{A,a}^B G_B^{-1C} e'_C + G_A^B \gamma_{Ba}^C e_C$$

$$= G_{A,a}^B G_B^{-1C} e'_C + G_A^C \gamma_{Ca}^B e_B$$

$$= G_{A,a}^B G_B^{-1C} e'_C + G_A^C \gamma_{Ca}^B G_B^{-1D} e'_D$$

$$= G_{A,a}^C G_C^{-1B} e'_B + G_A^C \gamma_{Ca}^D G_D^{-1B} e'_B$$

$$= \left(G_{A,a}^C G_C^{-1B} + G_A^C \gamma_{Ca}^D G_D^{-1B}\right) e'_B.$$

If we compare this with (16) , we can conclude that

$$\gamma'^B_{Aa} = G_{A,a}^C G_C^{-1B} + G_A^C \gamma_{Ca}^D G_D^{-1B} ,$$

or in matrix notation,

$$\gamma'_a = G_{,a} G^{-1} + G \gamma_a G^{-1} \qquad (17)$$

This is also refered to as a gauge transformation.

We say γ'_a and γ_a are gauge- equivalent.

The curvature tensor or gauge field of this connection is defined by

$$\mathrm{F}_{ab} = \gamma_{b,a} - \gamma_{a,b} - [\gamma_a, \gamma_b] = \gamma_{b,a} - \gamma_{a,b} - [\gamma_a\gamma_b - \gamma_b\gamma_a], \qquad (18)$$

which is skew-symmetric in a and b. We are going to examine how this expression is consistent with how we define curvature in (5).

We start with $2[\mathrm{d}\Gamma_b^a - \Gamma_b^e \wedge \Gamma_e^a]$ and establish a new notation.

Think of the 1-form $\Gamma_b^a = \Gamma_{bk}^a \,\mathrm{d}x^k$ as being $\gamma_B^A = \gamma_{Ak}^B \,\mathrm{d}x^k$, the connection (matrix-valued) one form, where $A, B = 1, \ldots, n$; k= 0, 1, 2, 3.

Then $2[\mathrm{d}\Gamma_b^a - \Gamma_b^e \wedge \Gamma_e^a] = 2[\mathrm{d}\gamma_B^A - \gamma_B^E \wedge \gamma_E^A]$

$$= 2[\mathrm{d}(\gamma_{Bk}^A \,\mathrm{d}x^k) - [\gamma_{Bc}^E \,\mathrm{d}x^c \wedge \gamma_{Ek}^A \,\mathrm{d}x^k]]$$

$$= 2\left[\gamma_{Bk,c}^A \,\mathrm{d}x^c \wedge \mathrm{d}x^k - [\gamma_{Bc}^E \gamma_{kE}^A] \,\mathrm{d}x^c \wedge \mathrm{d}x^k\right]$$

$$= 2\left[\gamma^{A}_{B[k,c]} - \gamma^{E}_{B[c}\, \gamma^{A}_{k]E}\right] \mathrm{d}x^{c} \wedge \mathrm{d}x^{k}$$

$$= 2\left[\tfrac{1}{2}\left(\gamma^{A}_{Bk,c} - \gamma^{A}_{Bc,k}\right) - \tfrac{1}{2}\left(\gamma^{E}_{Bc}\gamma^{A}_{kE} - \gamma^{E}_{Bk}\gamma^{A}_{cE}\right)\right] \mathrm{d}x^{c} \wedge \mathrm{d}x^{k}$$

$$= \left[\gamma^{A}_{Bk,c} - \gamma^{A}_{Bc,k} - \left(\gamma^{E}_{Bc}\gamma^{A}_{kE} - \gamma^{E}_{Bk}\gamma^{A}_{cE}\right)\right] \mathrm{d}x^{c} \wedge \mathrm{d}x^{k}$$

Now let us examine the components of this expression.

We see that F_{ab} becomes F_{ck}, which is skew-symmetric in c and k, i.e.

$$\mathrm{F}^{A}_{Bck} = \gamma^{A}_{Bk,c} - \gamma^{A}_{Bc,k} - (\gamma^{E}_{Bc}\gamma^{A}_{kE} - \gamma^{E}_{Bk}\gamma^{A}_{cE}),\ \text{or}$$

$\mathrm{F}_{ck} = \gamma_{k,c} - \gamma_{c,k} - (\gamma_c\gamma_k - \gamma_k\gamma_c)$, which is the curvature tensor or gauge field.

So $\mathrm{F}_{ck} = \gamma_{k,c} - \gamma_{c,k} - (\gamma_c\gamma_k - \gamma_k\gamma_c) = 2\gamma_{[k,c]} - 2\gamma_{[c}\,\gamma_{k]}$.

Now we will examine how the Yang-Mills field F_{ab} transforms under the gauge transformation given by (17).

Consider $\mathrm{F}_{ab} = \gamma_{b,a} - \gamma_{a,b} - [\gamma_a, \gamma_b]$. Now let us write down F'_{ab} using the fact that

$\gamma'_a = G_{,a}G^{-1} + G\,\gamma_a\,G^{-1}$.

Note that $G^{-1}_{,a} = -G^{-1}G_{,a}G^{-1}$. Since (GG^{-1}) =I,

$$(GG^{-1})_{,a} = 0\,,$$

$$(GG^{-1})_{,a} = G_{,a}G^{-1} + GG^{-1}_{,a} = 0.$$

$$\text{So } -G_{,a}G^{-1} = GG^{-1}_{,a}.$$

$$\text{So } G^{-1}_{,a} = -G^{-1}G_{,a}G^{-1}\,.$$

So $\mathrm{F}'_{ab} = \gamma'_{b,a} - \gamma'_{a,b} - [\gamma'_a, \gamma'_b]$

$= [G_{,b}G^{-1} + G\gamma_b G^{-1}]_{,a} - [G_{,a}G^{-1} + G\gamma_a G^{-1}]_{,b} -$

$-[(G_{,a}G^{-1} + G\gamma_a G^{-1})(G_{,b}G^{-1} + G\gamma_b G^{-1})] +$

$+[(G_{,b}G^{-1} + G\gamma_b G^{-1})(G_{,a}G^{-1} + G\gamma_a G^{-1})]$

$= G_{;ba}G^{-1} + G_{,b}G^{-1}_{,a} + G_{,a}\gamma_b G^{-1} + G\gamma_{b,a}G^{-1} + G\gamma_b G^{-1}_{,a} - G_{;ab}G^{-1} - G_{,a}G^{-1}_{,b}$

$$-G_{,b}\gamma_a G^{-1} - G\gamma_{a,b}G^{-1} - G\gamma_a G^{-1}_{,b}$$

$$-[G_{,a}G^{-1}G_{,b}G^{-1} + G_{,a}G^{-1}G\gamma_b G^{-1} + G\gamma_a G^{-1}G_{,b}G^{-1} + G\gamma_a G^{-1}G\gamma_b G^{-1}]$$

$$+[G_{,b}G^{-1}G_{,a}G^{-1} + G_{,b}G^{-1}G\gamma_a G^{-1} + G\gamma_b G^{-1}G_{,a}G^{-1} + G\gamma_b G^{-1}G\gamma_a G^{-1}]$$

$$= G_{,b}G^{-1}_{,a} + G_{,a}\gamma_b G^{-1} + G\gamma_{b,a}G^{-1} - G_{,a}G^{-1}_{,b} - G_{,b}\gamma_a G^{-1} - G\gamma_{a,b}G^{-1}$$

$$-G_{,a}G^{-1}G_{,b}G^{-1} - G_{,a}\gamma_b G^{-1} - G\gamma_a\gamma_b G^{-1} + G_{,b}G^{-1}G_{,a}G^{-1} + G_{,b}\gamma_a G^{-1} + G\gamma_b\gamma_a G^{-1}$$

$$= -G_{,b}G^{-1}G_{,a}G^{-1} + G_{,a}G^{-1}G_{,b}G^{-1} - G_{,a}G^{-1}G_{,b}G^{-1} + G_{,b}G^{-1}G_{,a}G^{-1}$$

$$+G[\gamma_{b,a} - \gamma_{a,b} - \gamma_a\gamma_b + \gamma_b\gamma_a]G^{-1}$$

$$= G[\gamma_{b,a} - \gamma_{a,b} - \gamma_a\gamma_b + \gamma_b\gamma_a]G^{-1}$$

$$= G\mathrm{F}_{ab}G^{-1}.$$

In a similar manner to what we did with the Maxwell electromagnetic field we have the following Bianchi identities, which are satisfied by the curvature tensor F_{ab}, where the partial derivative for the Maxwell case is replaced by the covariant derivative

$$\nabla_{[c}\ \mathrm{F}_{ab]} = 0. \qquad (19)$$

Now we will define the dual field by

$$\mathrm{F}^{\star}_{ab} = \tfrac{1}{2}\zeta_{abcd}\mathrm{F}^{cd}, \zeta_{abcd} = (-g)^{\frac{1}{2}} \in_{abcd}, \quad (20)$$

with $\in_{abcd}$ the alternating symbol with $\in_{0123} = -1$ ($\in_{\alpha\beta\gamma\delta} = -\in^{\alpha\beta\gamma\delta}$)

We now write the Yang-Mills equations are given by the following two sets of equations.

One of them is $g^{bc}\nabla_c \mathrm{F}_{ab} = J_a$, (21)

where g^{bc} is Minkowski metric and J_a is the current, and the other one

$$g^{bc}\nabla_c \mathrm{F}^{\star}_{ab} = 0 \qquad (22)$$

which is equivalent to the Bianchi identities given by $\nabla_{[c}\ \mathrm{F}_{ab]} = 0$ which is always satisfied because F_{ab} is given by

$$F_{ab} = \gamma_{b,a} - \gamma_{a,b} - [\gamma_a, \gamma_b].$$

Now consider (20). If we take $J_a = 0$, then the Yang-Mills equations become

$$g^{bc} \nabla_c F_{ab} = 0, \qquad (23)$$

which is equivalent to

$$\nabla_{[c} F^{\star}_{ab]} = 0. \qquad (24)$$

In the case that $F^{\star}_{ab} = \pm i F_{ab}$ (i.e. F_{ab} is self-dual or anti-self dual) then (24) implies

$$\nabla_{[c}\ iF_{ab]} = 0 \text{ or } \nabla_{[c} F_{ab]} = 0,$$

which is identically satisfied by (19).

Therefore saying that the F_{ab} is self-dual or anti-self dual is equivalent to saying that F_{ab}

satisfies the Yang-Mills equations with $J_a = 0$.

Now consider the special case when n=1. We are going to show that all parts of the above discussion reduce to the Maxwell case, where (21), which is $g^{bc} \nabla_c F_{ab} = J_a$ is a generalization of $\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} J^\beta$.

On Minkowski space M we will consider the vector bundle B, i.e. $B = M \times C^1$ (n=1), which now becomes a line bundle. The global vector fields e_A ($A = 1$) form a basis set as does $e'_A = G^B_A(x^a)\, e_B$, with the matrix-valued function $G^B_A(x^a)$ becoming a scalar function on M, call it $g(x^a)$, $a = 0, 1, 2, 3$.

$\nabla_a\ e_1 = \gamma_a\, e_1$, where γ_a is the connection (electromagnetic potential) and is simply a one-form on Minkowski space.

Under a change in basis (gauge transformation), the new potential $\gamma'_a = g_{,a}\, g^{-1} + \gamma_a$ (since $gg^{-1} = 1$).

Note that $(\log g)_{,a} = \frac{1}{g} g_{,a} = g_{,a} g^{-1}$.

So we can rewrite this as $\gamma'_a = \varphi_{,a} + \gamma_a$, where $\varphi = \log g$.

And (18) becomes $F_{ab} = \gamma_{b,a} - \gamma_{a,b}$ $(\gamma_a \gamma_b = \gamma_b \gamma_a)$.

And under a change in basis $F'_{ab} = F_{ab}$. In other words when we change the potential by adding a gradient of some scalar function φ, the Maxwell field remains unchanged. And this is a well-established fact of electricity and magnetism.

We also note that the Yang-Mills field equation (21) reduced to the Maxwell equations.

CONCLUSION

In this thesis we started with a discussion of the concept of a differentiable manifold, which is a topological manifold with a C^{∞} differentiable structure, and the concepts of vectors and tensors which are defined on the manifold.

We were able to show the important fact that the tensor itself does not change as an object, independent of the choice of a coordinate system. This is important because it is used to illustrate the covariance of the laws of physics as well as therefore in the study of these laws. Covariance of these laws means that we want our basic physical principles to remain unchanged when we change our coordinate system.

In our study of an exterior differentiation, a differential operation which depends only on the manifold structure, we were able to show that the exterior differentiation operator is covariant.

An extra structure, the connection, defined the covariant derivative and the Riemann curvature tensor, which gives us an indication of the curvature of the manifold.

We were mostly concerned with the mathematical-physics application where the 4-dimensional space-time has a special name, Minkowski space.

We were able to show the covariance of electrodynamics by casting Maxwell's equations, which describe the behavior of electromagnetic fields, in tensor form.

We examined the definition of a Yang-Mills field and how it could be thought of as a generalization of a Maxwell field as well as illustrating many of the mathematical concepts earlier discussed as becoming part of the definition of the Yang-Mills field equations.

This thesis could be used as a starting point for somebody interested in studying an area of mathematical physics which makes use of differential geometry, for example, general relativity or fluid mechanics. This work can lead us to the subject of Bäcklund transformations, where the basic idea is to generate new solutions of the self-dual or anti-self-dual Yang-Mills equations from a seed solution.

APPENDIX

In this Appendix we include a derivation of some basic relationships from elementary vector analysis which would be useful in obtaining the left hand sides of some of Maxwell's equations in terms of the electric and magnetic potentials, ϕ and A, respectively, instead of the electric and magnetic fields, E and B, respectively.

Let $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$ and ϕ at a scalar function of x, y, z, and t.

1) Show $\nabla \cdot \left(\vec{a} - \vec{b}\right) = \nabla \cdot \vec{a} - \nabla \cdot \vec{b}$.

Let $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$

$$\nabla \cdot \left(\vec{a} - \vec{b}\right) = \tfrac{\partial}{\partial x}(a_1 - b_1) + \tfrac{\partial}{\partial y}(a_2 - b_2) + \tfrac{\partial}{\partial z}(a_3 - b_3) =$$

$$= \frac{\partial a_1}{\partial x} - \frac{\partial b_1}{\partial x} + \frac{\partial a_2}{\partial y} - \frac{\partial b_2}{\partial y} + \frac{\partial a_3}{\partial z} - \frac{\partial b_3}{\partial z} =$$

$$= \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} - \left(\frac{\partial b_1}{\partial x} + \frac{\partial b_2}{\partial y} + \frac{\partial b_3}{\partial z}\right) = \nabla \cdot \vec{a} - \nabla \cdot \vec{b} .$$

2) Show $\nabla \cdot \frac{\partial \vec{A}}{\partial t} = \frac{\partial}{\partial t}\left(\nabla \cdot \vec{A}\right)$.

Take $\nabla \cdot \vec{A} = \frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}$; now take $\frac{\partial}{\partial t}\left(\nabla \cdot \vec{A}\right)$:

$$\frac{\partial}{\partial t}\left(\frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}\right) = \frac{\partial^2 A_1}{\partial t \partial x} + \frac{\partial^2 A_2}{\partial t \partial y} + \frac{\partial^2 A_3}{\partial t \partial z} = \frac{\partial^2 A_1}{\partial x \partial t} + \frac{\partial^2 A_2}{\partial y \partial t} + \frac{\partial^2 A_3}{\partial z \partial t} =$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial A_1}{\partial t}\right) + \frac{\partial}{\partial y}\left(\frac{\partial A_2}{\partial t}\right) + \frac{\partial}{\partial z}\left(\frac{\partial A_3}{\partial t}\right) = \nabla \cdot \frac{\partial \vec{A}}{\partial t}$$

$$= \nabla \cdot \frac{\partial \vec{A}}{\partial t} .$$

3) Show $\nabla \cdot (\nabla \phi) = \nabla^2 \phi$

$$\nabla \cdot (\nabla \phi) = \nabla \cdot (\phi_{,x}, \phi_{,y}, \phi_{,z}) = \phi_{,xx} + \phi_{,yy} + \phi_{,zz} = \nabla^2 \phi$$

4) Show $\frac{\partial}{\partial t}(\nabla\phi)=\nabla\frac{\partial\phi}{\partial t}$.

$$\nabla\phi=\left(\frac{\partial\phi}{\partial x},\frac{\partial\phi}{\partial y},\frac{\partial\phi}{\partial z}\right)$$

$$\frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial x},\frac{\partial\phi}{\partial y},\frac{\partial\phi}{\partial z}\right)=\left(\frac{\partial^2\phi}{\partial t\partial x},\frac{\partial^2\phi}{\partial t\partial y},\frac{\partial^2\phi}{\partial t\partial z}\right)=\left(\frac{\partial^2\phi}{\partial x\partial t},\frac{\partial^2\phi}{\partial y\partial t},\frac{\partial^2\phi}{\partial z\partial t}\right)=\left(\frac{\partial}{\partial x}\left(\frac{\partial\phi}{\partial t}\right),\frac{\partial}{\partial y}\left(\frac{\partial\phi}{\partial t}\right),\frac{\partial}{\partial z}\left(\frac{\partial\phi}{\partial t}\right)\right)=\nabla\frac{\partial\phi}{\partial t}$$

5) Show $\nabla\times\left(\nabla\times\vec{a}\right)=\nabla\left(\nabla\cdot\vec{a}\right)-\nabla^2\vec{a}$.

$$\nabla\times\vec{a}=\begin{bmatrix} i & j & k\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ a_1 & a_2 & a_3\end{bmatrix}=i\left(\frac{\partial a_3}{\partial y}-\frac{\partial a_2}{\partial z}\right)+j\left(\frac{\partial a_1}{\partial z}-\frac{\partial a_3}{\partial x}\right)+k\left(\frac{\partial a_2}{\partial x}-\frac{\partial a_1}{\partial y}\right),$$

$$\nabla\times\left(\nabla\times\vec{a}\right)=\begin{bmatrix} i & j & k\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ \frac{\partial a_3}{\partial y}-\frac{\partial a_2}{\partial z} & \frac{\partial a_1}{\partial z}-\frac{\partial a_3}{\partial x} & \frac{\partial a_2}{\partial x}-\frac{\partial a_1}{\partial y}\end{bmatrix}=$$

$$=i\left[\frac{\partial}{\partial y}\left(\frac{\partial a_2}{\partial x}-\frac{\partial a_1}{\partial y}\right)-\frac{\partial}{\partial z}\left(\frac{\partial a_1}{\partial z}-\frac{\partial a_3}{\partial x}\right)\right]$$

$$=j\left[\frac{\partial}{\partial z}\left(\frac{\partial a_3}{\partial y}-\frac{\partial a_2}{\partial z}\right)-\frac{\partial}{\partial x}\left(\frac{\partial a_2}{\partial x}-\frac{\partial a_1}{\partial y}\right)\right]+k\left[\frac{\partial}{\partial x}\left(\frac{\partial a_1}{\partial z}-\frac{\partial a_3}{\partial x}\right)-\frac{\partial}{\partial y}\left(\frac{\partial a_3}{\partial y}-\frac{\partial a_2}{\partial z}\right)\right].$$

$$\nabla^2\vec{a}=\frac{\partial^2\vec{a}}{\partial x^2}+\frac{\partial^2\vec{a}}{\partial y^2}+\frac{\partial^2\vec{a}}{\partial z^2}\ :\qquad \frac{\partial^2\vec{a}}{\partial x^2}=\left(\frac{\partial^2\vec{a_1}}{\partial x^2},\frac{\partial^2\vec{a_2}}{\partial x^2},\frac{\partial^2\vec{a_3}}{\partial x^2}\right)$$

$$\frac{\partial^2\vec{a}}{\partial y^2}=\left(\frac{\partial^2\vec{a_1}}{\partial y^2},\frac{\partial^2\vec{a_2}}{\partial y^2},\frac{\partial^2\vec{a_3}}{\partial y^2}\right)$$

$$\frac{\partial^2\vec{a}}{\partial z^2}=\left(\frac{\partial^2\vec{a_1}}{\partial z^2},\frac{\partial^2\vec{a_2}}{\partial z^2},\frac{\partial^2\vec{a_3}}{\partial z^2}\right)$$

$$\nabla\left(\nabla\cdot\vec{a}\right)=\nabla\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)=i\frac{\partial}{\partial x}\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)+j\frac{\partial}{\partial y}\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)+$$

$$+k\frac{\partial}{\partial z}\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)$$

Consider x components from both sides :

x component for $\nabla\times\left(\nabla\times\vec{a}\right)$ is $\frac{\partial}{\partial y}\left(\frac{\partial a_2}{\partial x}-\frac{\partial a_1}{\partial y}\right)-\frac{\partial}{\partial z}\left(\frac{\partial a_1}{\partial z}-\frac{\partial a_3}{\partial x}\right)=\frac{\partial^2 a_2}{\partial y\partial x}-\frac{\partial^2 a_1}{\partial y^2}-\frac{\partial^2 a_1}{\partial z^2}+\frac{\partial^2 a_3}{\partial z\partial x}$,

x component for $\nabla\left(\nabla\cdot\vec{a}\right)-\nabla^2\vec{a}$ is $\frac{\partial}{\partial x}\left(\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial y}+\frac{\partial a_3}{\partial z}\right)-\left(\frac{\partial^2 a_1}{\partial x^2}+\frac{\partial^2 a_1}{\partial y^2}+\frac{\partial^2 a_1}{\partial z^2}\right)=$

$$= \frac{\partial^2 a_1}{\partial x^2} + \frac{\partial^2 a_2}{\partial x \partial y} + \frac{\partial^2 a_3}{\partial x \partial z} - \frac{\partial^2 a_1}{\partial x^2} - \frac{\partial^2 a_1}{\partial y^2} - \frac{\partial^2 a_1}{\partial z^2} = \frac{\partial^2 a_2}{\partial x \partial y} - \frac{\partial^2 a_1}{\partial y^2} - \frac{\partial^2 a_1}{\partial z^2} + \frac{\partial^2 a_3}{\partial x \partial z} .$$

and the X component for $\nabla \times (\nabla \times \vec{a})$ is the same as the X component for $\nabla(\nabla \cdot \vec{a}) - \nabla^2 \vec{a}$. (using the fact that the mixed partial derivatives in the two expressions are equal)

Similarly we can show for the y and z components .

So $\nabla \times (\nabla \times \vec{a}) = \nabla(\nabla \cdot \vec{a}) - \nabla^2 \vec{a}$.

7) Show $\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = \nabla \times \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t}$.

$\nabla \times \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} = \nabla \times (\nabla \times \vec{A}) - \frac{1}{c} \frac{\partial}{\partial t}\left(-\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \phi\right)$ (say from where these equations)

$$= \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} + \frac{1}{c} \frac{\partial}{\partial t}(\nabla \phi)$$

$$= \nabla\left(-\frac{1}{c} \frac{\partial \phi}{\partial t}\right) - \nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} + \frac{1}{c} \frac{\partial}{\partial t}(\nabla \phi)$$

$$= -\frac{1}{c} \nabla \frac{\partial \phi}{\partial t} - \nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} + \frac{1}{c} \frac{\partial}{\partial t}(\nabla \phi) = \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} .$$

8) Show $\frac{1}{c} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \nabla \cdot \vec{E}$.

$$\nabla \cdot \vec{E} = \nabla \cdot \left(-\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \phi\right) = \nabla \cdot \left(-\frac{1}{c} \frac{\partial \vec{A}}{\partial t}\right) - \nabla \cdot (\nabla \phi) = -\frac{1}{c} \frac{\partial}{\partial t}(\nabla \cdot \vec{A}) - \nabla^2 \phi =$$

$$= -\frac{1}{c} \frac{\partial}{\partial t}\left(-\frac{1}{c} \frac{\partial \phi}{\partial t}\right) - \nabla^2 \phi = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi .$$

References

[1] Boothby W.M. 1975 *An Introduction to differentiable manifolds and Riemannian geometry* Academic Press, New York, San francisco, London

[2] Chakravarty S., Halburd R.G., Kent S.L 2003 *Self-similar solutions of certain coupled integrable systems* J.Phys A: Mathematical and General **36** 1371-1384

[3] Curtis C.W.1991 *Linear Algebra, An Introductory Approach,* Spinger, New York

[4] Fitzpatrick P.M. 1996 *Advanced calculus,* PWS Publishing company

[5] Hawking S.W. and Ellis G.F.R. 1973 *The large scale structure of spase-time* Cambridge University press

[6] Jacson J.D. 1975 *Classical Electrodynamics, second edition,* John Wiley and Sons Inc.

[7] Kent S.L., Newman E.T. April 1983 *Yang-Mills theory in null-path space* J.Math.Phys. **24**

[8] Mason L., Chakravarty S., Newman E.T. April 1988 *Bäcklund Transformation for the anti-self-dual Yang-Mills equetions* J.Math.Phys. **29**

[9] Roseman D. 1999 *Elementry Topology,* Upper Saddle River, New Jersey

[10] Spinger C.E. 1962 *Tensor and vector analysis* The Ronald Press Company, New York

[11] Stewart J. *Calculus, third edition* Brooks/Cole Publishing Company

[12] Taylor E.F., Wheeler J.A. 1966 *Spacetime physics* W.H.Freeman and Company

[13] Valenza R.J. 1993 *Linear Algebra, An Introduction to Abstract Mathematics* Springer-Verlag New York, Inc.