An Introduction to Differential Geometry and General Relativity

A collection of notes for PHYM411

Thomas Haworth,
School of Physics,
Stocker Road,
University of Exeter,
Exeter,
EX4 4QL
tjh202@ex.ac.uk

March 29, 2010

Contents

Now pay attention. I’ll try to be clear. I know these things
are difficult and obscure, but I am full of the hope
of fame: it’s as if Dionysus had struck
my mind , even as the love of the Muses urged me onward
to attempt untrodden paths on the heights of their sacred
mountain.
I love to discover fresh springs that nobody else has drunk
from,
to pluck new flowers and weave a chaplet for my brow
from fields where no one has ever ventured before and the
Muses
have never recognised with this token of novel
achievement.

De Rerum Natura
Lucretius
(D. R. Slavitt)

1 Preamble: Qualitative Picture Of Manifolds

The bulk of the planned first half of this independent study involves building mathematical descriptions of manifolds and developing and manipulating objects on them. For example, looking at the ideas of fields (scalar and vector) and calculus on manifolds. Before diving into the mathematics, it is helpful to discuss relatively qualitatively (and briefly) what a manifold is, as well as some of the ideas early on in the notes which are important in the subsequent considerations.

1.1 Manifolds

Manifolds are generalizations of some ideas that are already very familiar. A curve through 3D Euclidean space can be paramaterised by a single variable "t" via x(t), y(t), z(t). Similarly a plane can be paramaterised by two variables "u, v" via x(u,v), y(u,v) and z(u,v).

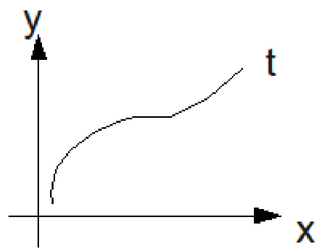

Figure 1: A line can be paramaterised by a single variable "t"'

These objects are considered to be locally very similar to 1D and 2D Euclidean spaces ($\mathbb{R}$ and $\mathbb{R}^2$) respectively. More specifically, we say that there is a continuous, one to one mapping between the local curve and flat space. This is known as being homeomorphic. For example, we consider a curve to be straight (1D Euclidean) at infinitesimal scales (fig. 2).

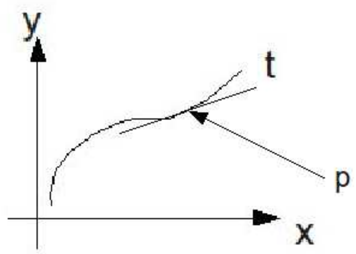

Figure 2: We consider a tangent at some point p to match the curve perfectly at infinitesimal scales

So, a manifold is a space which is homeomorphic to some $\mathbb{R}^m$ locally, but may be different from $\mathbb{R}^m$ globally (i.e. ambient space). We are able to describe the manifold using a set of "m" local coordinates, however, if the manifold is not homeomorphic to $\mathbb{R}^m$ globally then we have to use several of these local ecoordinates between which, transitions must be smooth (see example below).

1.1.1 Example: The Unit Sphere

To illustrate the above, amongst other points, consider a unit sphere (that with unit radius) embedded in 3D Euclidean space ($\mathbb{R}^3$). This is our manifold. We paramaterize our unit sphere by (to take one example) use of polar coordinates θ and ϕ via:

$$x = \sin(\theta)\cos(\phi)$$
$$y = \sin(\theta)\sin(\phi)$$
$$z = \cos(\theta)$$

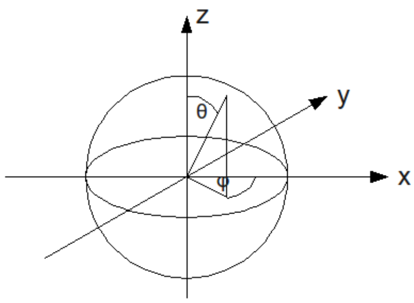

Figure 3: Demonstrating the unit sphere including the polar angles

We can invert these expressions to yield some local (paramaterised) coordinates:

$$\theta = \tan^{-1}(\frac{\sqrt{x^2+y^2}}{z}), \phi = \tan^{-1}(\frac{y}{x})$$

Note, that we do not have to use polar coordinates (cf steregraphic projection later). In the theory of manifolds all coordinate systems are equally good, which is appropriate since in physics we expect the behaviour of a physical system to be independent of the coordinates used.
Further note, that on this manifold no coordinate system may be used everywhere at once...

- At $\theta = \frac{\pi}{2}$ consider ϕ. At $\phi = 2\pi$ we have a discontinuity in ϕ from 2π to 0. In addition, nearby (infinitesimally close) points have very different values. It is not continuous. (since $0 \leq \phi \leq 2\pi$).
- If we were to let $\phi \geq 0$ then every point would correspond to an infinite number of values $(2n\pi, n = 1, 2, ...\infty)$.
- At the poles ϕ is undefined!

We have spotted some flaws in our use of spherical polar coordinates to describe the whole of our manifold. If we are to consider the local behaviour of the manifold we require that the following conditions be met:

1. Nearby points on the manifold should have nearby coordinates on the manifold.
2. Each point on the manifold should have unique coordinates.

To resolve this problem (the study of the first sections of these notes) we introduce 2 or more coordinate systems which overlap, each covers part of the sphere and *does* obey the conditions:

1. Nearby points have nearby coordinates in *at least one* coordinate system.

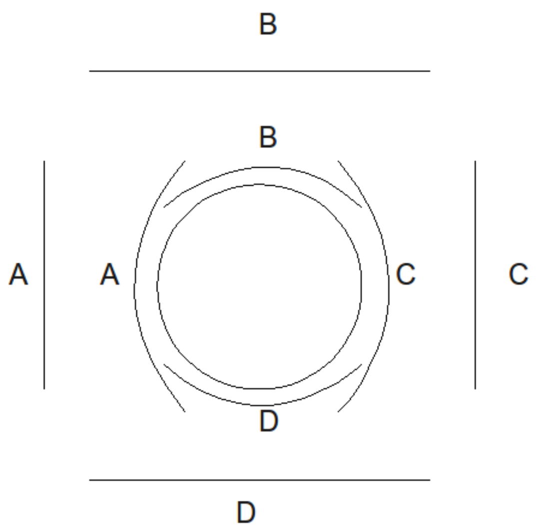

Figure 4: Each chart; A, B, C, D maps part of the circle, however, note that they overlap...

2. Each point has unique coordinates in each coordinate system that contains it.
   For example, consider the unit circle (easier to draw than sphere):
   Due to this overlap, we also require that:

3. If two coordinate systems overlap, they must be related to each other in a sufficiently smooth way.

This final condition is important for retaining continuity over the manifold and subsequently upholding the first two requirements (see later).
So, now that the idea of a manifold and some of its associated concepts that we will have to investigate in these notes have been outlined we can move on to the mathematics with a clearer picture of what will be going on.

2 Distances, Open Sets, Curves and Surfaces

2.1 Defining Space And Distances

2.1.1 Setting Up Space

We can formally declare the space that we are working in. An n-dimensional Euclidean space would be written as:

$$E_n = \{(y_1, y_2, ...y_n) \mid y_i \in \mathbb{R}\}$$

So, here we first declare the number of dimensions, the "set" (...). The symbol $|$ then means "restricted by...". We are finally told with the set membership symbol $\in$ that each dimension of this Euclidean space is an element of real space. So, to take some examples:

$$E_1 = \{y_1 \mid y_1 \in \mathbb{R}\}$$

This represents a set of real numbers, also known as a "real line". Similarly:

$$E_2 = \{(y_1, y_2) \mid (y_1, y_2) \in \mathbb{R}\}$$

This is the familiar cartesian plane of coordinates. E_3 would be 3D Euclidean space and so on for n = 4, 5...

2.1.2 Magnitude (or norm)

The magnitude of some vector $\vec{y} = (y_1, y_2, ...y_n)$ from the origin is defined to be:

$$\|\vec{y}\| = \sqrt{{y_1}^2 + {y_2}^2 + ...{y_n}^2}$$

This can be thought of as the distance from the origin. In a similar manner, the distance between two points $\vec{y} = (y_1, y_2, ...y_n)$ $\vec{z} = (z_1, z_2, ...z_n)$ is then:

$$\|\vec{z} - \vec{y}\| = \sqrt{(z_1 - y_1)^2 + (z_1 - y_2)^2 + ...(z_n - y_n)^2}$$

This is all fairly straightforward.

2.2 Open Sets

A subset "U" of E_n (written as $U \subset E_n$) is called "open" if, for every point $\vec{y}$ in U, all points of E_n within some positive distance "r" of $\vec{y}$ are also in U (the size of r can vary).
Intuitively, an open set is some region minus its boundary. If we include the boundary we get a closed set. A closed set goes right up to the edge of the boundary, so no circle of radius r can be

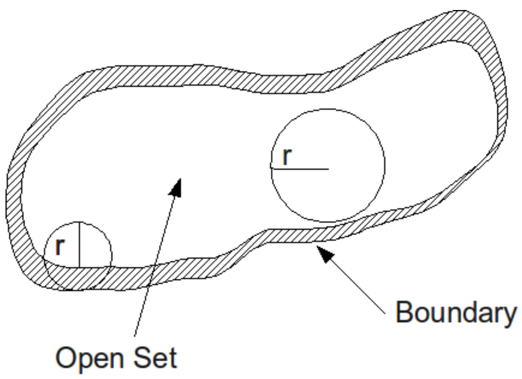

Figure 5: At the edge of the open set we can find some circle of radius r which is still totally contained in U.

drawn at every point and still be in U.

2.2.1 Open Ball

If $\vec{a} \in E_n$, then the open ball with center $\vec{a}$ and radius r is the set of all points in E_n whose distance from a is less than r. An open ball is everything except the ball boundary (as above).
An open ball is written as:

$$B(\vec{a}, r) = \{\vec{x} \in E_n \mid \|\vec{x} - \vec{a}\| < r\}$$

To reiterate, it is open if for all $\vec{x}$ we can travel a distance r in any direction and still be in $B(\vec{a}, r)$

2.2.2 Examples of Open Sets:

- E_n is open (i.e. Euclidean real space has no boundaries)
- $\phi = \{\}$, the empty set, is open
- Unions of sets are open. A union, denoted $\cup$ is the set of all distinct elements of the collection.
- As extension of the above, open sets are actually unions of open balls, $U = \cup B_y$. Put simply, an open set can be considered to have an open ball at each point.

2.2.3 Relatively Open Sets

Consider some manifold M which is a subset of some s-dimensional Euclidean space E_s. A relatively open set applies the same definition of an open set to itself with respect to the manifold M. It is called relatively open because it is not necessarily open in the ambient Euclidean space. For example, if we consider s = 3 and M as some hemisphere with U as a relatively open set in M:

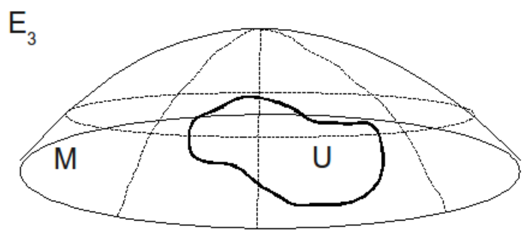

Figure 6: A relatively open set U on M

Then, U *is not open* in E_3, since there are points arbitrarily close to U in E_3 that lie outside of U. However, U *is open* in M, since given any point $\vec{y}$ in U, all points of M within a small enough distance from $\vec{y}$ are still in U.
For example, some relatively open set might contain every point on the Earth within 100 miles of London whilst containing no points above the surface of the Earth.

2.3 Parametric Paths and Surfaces in E3

These concepts are introduced in three dimensions first to aid conceptual understanding.
First, a quick example of paramaterization. We can paramaterize the straight line through a point $\{p_i\}$ in E_3 and parallel to the vector $\vec{a}$ by:

$$y_i = p_i + a_i t$$

2.3.1 Smooth Paths

A *smooth path* in E_3 is a set of three smooth (i.e. infinitely differentiable) real-valued functions of a single variable "t":

$$y_i = y_i(t) \mid i = (1, 2, 3)$$

Where the curve has been paramaterised by t. By infinitely differentiable, we mean that we can differentiate the functions an infinite number of times and still get functions, e.g. a trigonometric

or exponential function is smooth.
If the vector of derivatives $\frac{dy_i}{dt}$ is nowhere zero, we have what we call a non-singular path. Such a path never terminates. Our example of a straight line is not smooth, since after taking the first derivative we are left with a constant.

2.3.2 Smooth Surfaces

Continuing our restricted 3D analysis, a smooth surface immersed in E_3 is a collection of 3 smooth real-valued functions of two paramaterizing variables x^1 and x^2 (these upper indices refer to local coordinates, *not* the square of some value).

$$y_i = y_i(x^1, x^2)$$

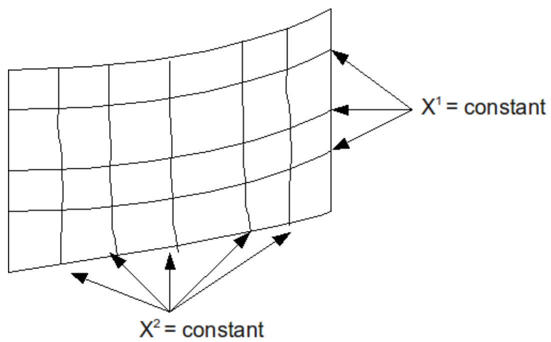

Figure 7: A 2D surface in 3D space with local coordinates marked on

Note that if we hold x^1 contstant we get a smooth path (see figure 7) where different constants yield different paths. Similarly, holding x^2 constant gives another batch of paths which intersect the first ones.
x^1 and x^2 are known as *local coordinates*. This smooth surface requires, by definition, that the 3×2 matrix whose ij^{th} entry is $\frac{\partial y_i}{\partial x^j}$ (the Jacobian matrix) has rank 2. i.e. the 2 rows are linearly independent. This is the requirement that the surface is never singular.

2.3.3 Further Examples of Paramaterization

1. Paramaterization of a plane through $\{p_i\}$, parallel to the independent vectors $\vec{a}, \vec{b}$ is given by:

   $$y_i = p_i + a_i x^1 + b_i x^2 \mid (i = 1, 2, 3)$$

   This is an extension of our consideration of a paramaterized line.

2. Paramaterization of the unit sphere, ${y_1}^2 + {y_2}^2 + {y_3}^2 = 1$, using spherical polar coordinates:

   $$y_1 = \sin(x^1)\cos(x^2), y_2 = \sin(x^1)\sin(x^2), y_3 = \cos(x^1)$$

3. **The Jacobian Matrix**: The matrix of 1st order partial derivatives of ambient coordinates with respect to local coordinates.

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x^1} & \frac{\partial y_2}{\partial x^1} & \frac{\partial y_3}{\partial x^1} \\ \frac{\partial y_1}{\partial x^2} & \frac{\partial y_2}{\partial x^2} & \frac{\partial y_3}{\partial x^2} \end{bmatrix}$$

For the unit sphere, by differentiating the equations in the previous example we get:

$$J = \begin{bmatrix} \cos(x^1)\cos(x^2) & \cos(x^1)\sin(x^2) & -\sin(x^1) \\ -\sin(x^1)\sin(x^2) & \sin(x^1)\cos(x^2) & 0 \end{bmatrix}$$

This Jacobian matrix is rank 2 everwhere except at $x^1 = n\pi$ for which $\sin(n\pi) = 0, cos(n\pi) = 1$ and:

$$J = \begin{bmatrix} \cos(x^2) & \sin(x^2) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where the rank is now 1.

2.3.4 Paramaters As Local Coordinates

The parametric equations of a surface show us how to obtain a point on the surface once we know two local coordinates (x^1, x^2). In other words, we have specified some function which maps E_2 to E_3 (we write this $E_2 \mapsto E_3$). We can use our parametric equations to get x^1, x^2 as functions of $y_i, (i = 1, 2, 3)$. Continuing our example in spherical polars for a unit sphere, we get:

$$x^1 = \cos^{-1}(y_3)$$

$$x^2 = \begin{cases} \cos^- 1(\frac{y_1}{\sqrt{{y_1}^2+{y_2}^2}}) & y_2 \geq 0 \\ 2\pi - \cos^- 1(\frac{y_1}{\sqrt{{y_1}^2+{y_2}^2}}) & y_2 < 0 \end{cases}$$

This allows us to give each point on *most* of the sphere two local coordinates. We do, however, have some issues with continuity in this example, since at $y_2 = 0, x^2$ switches from 0 to 2π. Also at the poles, $x^1 = 0$ and x^2 is not even defined.
We therefore add the restrictions on the portion of the sphere consiered via:

$$0 < x^1 < 2\pi \; ; \; 0 < x^2 < \pi$$

These define an open subset "U" of the spere.
In conclusion our local (x^i) coordinates allow us to map some subset of the manifold to Euclidean (y_i) space. We write this **mapping** as:

$$\mathbf{x} : U \to E_1$$

where:

$$\mathbf{x}(y_1, y_2, y_3) = (x^1(y_1, y_2, y_3), x^2(y_1, y_2, y_3))$$

is known as the *chart*.

2.4 Charts

A chart of a surface is a series of functions which specify each of the local coordinates (x^i) as smooth (infinitely differentiable) functions of some point on the surface in terms of global (ambient) coordinates. For an n dimensional surface it provides an n-1 dimensional map. This is just like a geographic chart gives a 2D representation of our 3 dimensional planet.
In general, we chart an entire manifold "M" by covering it with open sets which become the domains of coordinate charts. (Domain is set of input values for which a function is defined).

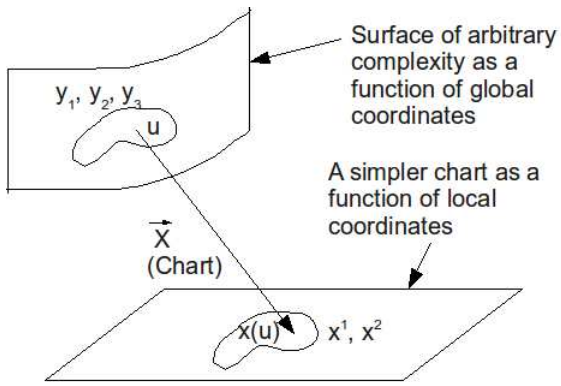

Figure 8: A graphical representation of a chart describing a smooth surface.

3 Smooth Manifolds And Scalar Fields

3.1 Open Cover

An open cover is a generalization of the single subset of a unit sphere that we looked at in the previous section. Formally, an open cover of a manifold $M \subset E_s$ is a collection of open sets $\{U_a\}$ in M, such that M is the union of these subsets.

$$M = \cup_a U_a$$

The open sets comprising the open cover will overlap.

3.2 n-Dimensional Smooth Manifolds and Change of Coordinate Transformations

$M \subset E_s$ is called an n-dimensional smooth manifold if we are given a collection:

$$\{U_a; x^1{}_a, x^2{}_a, ...x^n{}_a\}$$

(That is, a collection of the a^{th} open subsets that are described in terms of local coordinates) where:

- The U_a form an open cover of M.
- Each of these $x^r{}_a$ is a smooth real valued function defined on the subset U in such a manner that the chart:

  $$x(u) = (x^1{}_a, x^2{}_a, ...x^n{}_a)$$

  maps U_a to E_n in a one to one corresspondence. Formally this is written as, $x : U_a \to E_n$ is one to one. In other words, each point on the subset corresponds to one point on the n dimensional chart. Note that these $x^i{}_r$ are independent, thus their being smooth allows for partial derivatives of all orders.

As before we have local charts on this manifold M. The collection of all charts is called a *smooth atlas* of M. Note, that an open subset may also be referred to as a coordinate neighborhood.
Consider a situation where we have two local charts for our manifold M, which we denote (U, x^i) and $(V, \bar{x}^j)$. If neither of these sets are empty at some point, we can write some function:

$$x^i = x^i(\bar{x}^j)$$

With inverse:

$$\bar{x}^k = \bar{x}^k(x^l)$$

These functions are called change of coordinate transformations. Generally, if we are describing a manifold using a number of charts, we can use these functions to move between descriptions of the same points in different charts. This allows us to begin consideration of the overlapping regions between charts. Note that the specified condition that neither of our charts be null ensures that we only consider regions which both charts describe. These concepts are represented graphically in figure 9.

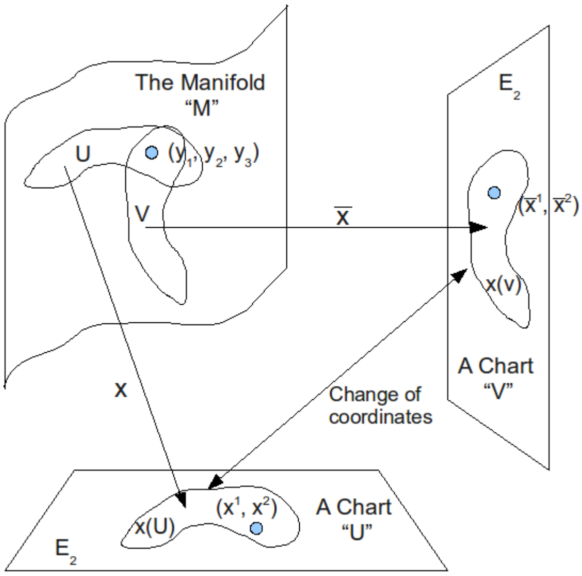

Figure 9: A graphical representation of the use of change of coordinate transformations in relation to our charts and manifold. Note that it only applies in the regions where the charts overlap

One important inference from the introduction of these functions is that our ambient space is not essential. We can define a smooth manifold using a number of charts without any reference to global coordinates.

3.2.1 Some Examples of Manifold Paramaterisation

- Euclidean space E_n is an n-dimensional manifold with a single identity chart defined by:

  $$y_i = x^i(y_1, y_2, ...y_n)$$

  Here the left hand side refers to the global coordinate and the right hand side gives the value of the local coordinate (which is represented in terms of global coordinates). In words we could say that this equation tells us that a global space coordinate of a point is equal to the local coordinate at that point.

- **Generalised unit polar coordinates:** Consider the manifold $M = S^n$ which is the unit n-dimensional sphere:

$$M = S^n = \{(y_1, y_2, ..., y_n, y_{n+1}) \in E_{n+1} \mid \sum {y_i}^2 = 1$$

With manifold local coordinates $(x^1, x^2, ..., x^n)$. Where, to resolve the continuity issues covered in section 2.3.4 we impose the conditions:

$$0 < (x^1, x^2, ...x^{n<1}) < \pi \text{ and } 0 < x^n < 2\pi$$

Generalised polar coordinates for a unit sphere are then given by:

$$\begin{aligned} &y_1 = \cos(x^1) \\ &y_2 = \sin(x^1)\cos(x^2) \\ &y_{n-1} = \sin(x^1)\sin(x^2)...\cos(x^{n-1}) \\ &y_n = \sin(x^1)\sin(x^2)...\sin(x^{n-1})\cos(x^n) \end{aligned}$$

We can apply this definition to non-unit spheres by multiplying each y_i by r.

3.3 The Example of Stereographic Projection

For a 3D unit sphere, stereographic projection involves tracing a line from the north pole through some point on the sphere and mapping the position on a 2D plane.

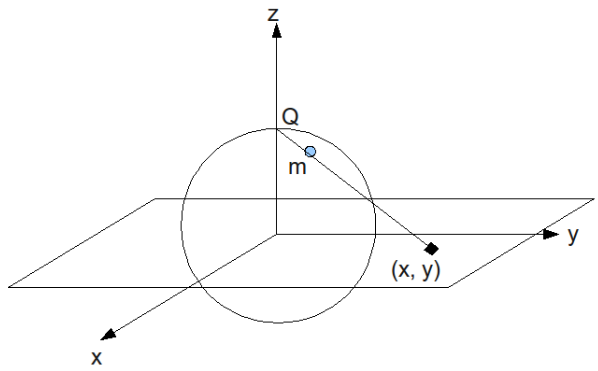

Figure 10: A representation of stereographic projection with an example tracing of a point "m".

We now look again at the local coordinates of the unit sphere S^n but this time using stereographic projection. Let Q be the point (0, 0, 0,... 1) and P be the point (0, 0, 0,... -1) (i.e. north and south poles respectively). We can then define two charts as follows in Fig. 11:

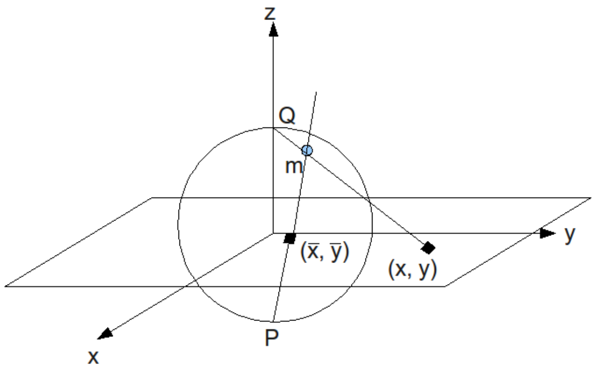

Figure 11: A demonstration of the use of two charts in stereographic projection.

That is, we have one chart due to tracing from the north pole and one chart due to tracing from the south pole. Note that at P or Q, stereographic projection fails at that point for the chart which maps using that point. So, if (y_1, y_2, y_n, y_{n+1}) is a point in S^n, we get for our charts from the north and south poles respectively:

$$x^1 = \frac{y_1}{1-y_{n+1}}, \bar{x}^1 = \frac{y_1}{1+y_{n+1}}$$
$$...$$
$$x^n = \frac{y_n}{1-y_{n+1}}, \bar{x}^n = \frac{y_n}{1+y_{n+1}}$$

As usual, we can invert these maps to solve for global coordinates in terms of local coordinates $x^i, \bar{x}^i$ as follows:

$$y_1 = \frac{2x^1}{r^2+1}, y_1 = \frac{2\bar{x}^1}{1+\bar{r}^2}$$
$$...$$
$$y_n = \frac{2x^n}{r^2+1}, y_n = \frac{2\bar{x}^n}{1+\bar{r}^2}$$
$$y_{n+1} = \frac{r^2-1}{r^2+1}, y_{n+1} = \frac{1-\bar{r}^2}{1+\bar{r}^2}$$

Here, the values r^2 and $\bar{r}^2$ are defined as:

$$r^2 = \sum_i x^i x^i = (x^1)^2 + (x^2)^2 + ...(x^n)^2$$

$$\bar{r}^2 = \sum_i \bar{x}^i \bar{x}^i = (\bar{x}^1)^2 + (\bar{x}^2)^2 ... + (\bar{x}^n)^2$$

We can then use these global and local coordinates to arrive at a change of coordinate transformation between our two stereographic projection maps:

$$\begin{aligned} x^1 &= \frac{y_1}{1-y_{n+1}} \\ &= \frac{\frac{2\bar{x}^1}{1+\bar{r}^2}}{1-\frac{1-\bar{r}^2}{1+\bar{r}^2}} \\ &= \frac{2\bar{x}^1(1+\bar{r}^2)}{(1+\bar{r}^2)2\bar{r}^2} \\ &= \frac{\bar{x}^1}{\bar{r}^2} \end{aligned}$$

This is exaclty the same form for $x^i, (i = 1, 2, 3...n)$. Via the same methodology we can also obtain the reverse transformation between maps, which is:

$$\bar{x}^i = \frac{x^i}{r^2}$$

In summary, for the example of stereographic projection of a unit sphere we have just found:

- Local coordinates in terms of global coordinates for both maps (both the projections from P and Q)
- Global coordinates in terms of local coordinates
- The transformations between the map given via projection from P and the map given via projection from Q

3.4 Scalar Fields

A smooth scalar field on a smooth manifold "M" is simply a smooth, real valued mapping of some scalar property of the manifold using a chart. This mapping is written as:

$$\phi : M \to E_1$$

That is, some function which is the scalar field ϕ associates to each point "m" on he surface of "M" a unique scalar value $\phi(m)$. This also applies to subsets of M, "U", however if the scalar field for U differs for that at M it is called a *local scalar field*.
If Φ is a scalar field on M and x^i are local coordinates of a chart **x**, then we can express Φ as a smooth function ϕ (ok, so not the best notation) of these local coordinates. This is effectively expressing a scalar field in terms of global coordinates (Φ) and local coordinates (ϕ). Similarly to the considerations of 3.2, if we have a second chart $\bar{\mathbf{x}}$ we will get a second function $\bar{\phi}$ which, for these scalar fields, must satisfy $\phi = \bar{\phi}$ at each point on the manifold. This is a consequence of our requirement that there be no preferential coordinate systems, so once all details have been re-mapped onto the surface they should all agree.

3.4.1 Example: The local Scalar Field

The most obvious candidate for local scalar fields are the coordinate functions themselves. If U is a subset (also referred to as a coordinate neighborhood) and **x** = $\{x^i\}$ is a chart on U, then the maps x^i are themselves *scalar fields*.
Sometimes we may wish to specify a scalar field purely in terms of its local paramaters (as opposed to global); that is, by specifying various functions ϕ instead of some single function Φ (as we do when we break a manifold up using a chart). However, we cannot just specify this collection of ϕ any way we want, since they must collaborate to give a value to each point of the manifold independently of local coordinates. That is, overlapping local coordinate systems must agree on values for the scalar field.
We can express this formally by considering a point P on our manifold ($P \in M$) with coordinates x^h and $\bar{x}^j$ in the charts **x** and $\bar{\mathbf{x}}$ respectively. The transformation rule requires that these local coordinates agree as to the value of the scalar field at this point *upon transformation*

$$\bar{\phi}(\bar{x}^j) = \phi(x^h)$$

$$\bar{\phi}(\bar{x}^j(x^h)) = \phi(x^h)$$

i.e. upon relation using the coordinate transformations we should find the same values for the scalar field, or $\bar{x}^j = \bar{x}^j(x^h)$.

4 Tangent Vectors and the Tangent Space

In the previous section we looked at scalar fields on smooth manifolds. We now look at vectors on smooth manifolds. In order to achieve this we first look at the idea of a *smooth path* in M.

4.1 Smooth Paths

A smooth path on the smooth manifold "M" is a smooth map defined on an open segment of the real line. Put simply it is a 1 dimensional map (a line) upon the manifold which covers some portion of the real line $\mathbb{R}$ (or E_1). Formally, it is the map:

$$\mathbf{r} : J \to M$$

where: J is some open interval of the real line, M represents the manifold and **r** is the vector valued mapping function $\mathbf{r}(t) = (y_1(t), y_2(t), ...y_s(t))$. Obviously, we require that the paramaterisation of the line "t" be in J ($t \in J$).
r is said to be a smooth line through a point on the surface $m \in M$ if $\vec{r}(t_o) = m$ for some $t_o \in J$. That is, if some point t_o along our paramaterised smooth path is at m.
We can specify a path in M which goes through m by its coordinates:

$$y_i = y_i(t)$$

Where the point m would be specified by $y_i(t_o)$. Furthermore, since the ambient and local coordinates are functions of one another, we can also express a path in terms of its local coordinates via:

$$x^i = x^i(t)$$

4.1.1 Example: Path segment ("latitude curve") through any fixed point "p" on the unit sphere

Consider some point on the surface of a unit sphere of n+1 global coordinates which we will denote as $p = (p_1, p_2...p_n, p_{n+1})$. This latitude curve path is then specified by:

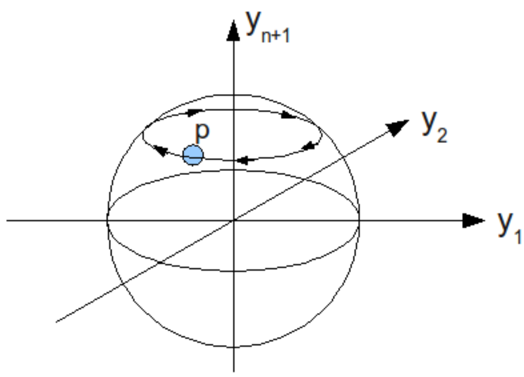

Figure 12: Path segment through "p" on the unit sphere

$$\begin{aligned} y_1 &= p_1 \cos(t) - p_2 \sin(t) \\ y_2 &= p_1 \sin(t) + p_2 \cos(t) \\ y_3 &= p_3 \end{aligned}$$

$$\cdots$$
$$y_{n+1} = p_{n+1}$$

This is a smooth path on S^n since the trigonometric functions have an infinite number of derivatives.

4.2 Tangent Vectors

Formally, a tangent vector at a point "m" on a maninfold "M" in some r-dimensional Euclidean space ($m \in M \subset E_r$) is a vector $\vec{v}$ in E_r of the form $\vec{v} = y'(t_o)$ for a path $y = y(t)$ in M through m with $y(t_o) = m$. Put more simply, it is a vector at some point along a path on the manifold which passes through m, *at* $y(t_o)$.

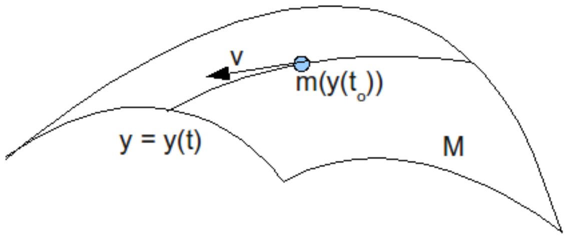

Figure 13: A representation of a tangent vector for a path through a point m on M

4.2.1 Example: The Paraboloid of Rotation

Let M be the surface $y_3 = {y_1}^2 + {y_2}^2$, i.e. the parabloid of rotation, which is paramaterised by:

$$y_1 = x^1, y_2 = x^2, y_3 = (x^1)^2 + (x^2)^2$$

This manifold is illustrated in Fig 14

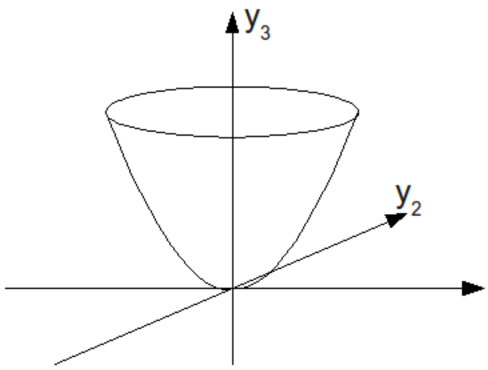

Figure 14: The paraboloid of rotation "M"'

This correspondance between the local and global coordinaes implies a single chart where U = M and extends over all x^1, x^2.

$$U = M : x^1, x^2$$

This is because we have two local coordinates x^1, x^2 which paramaterize the manifold fully (see the equations above).
To specify a tengent vector, we first need to specify a path in M which must satisfy the equation of the surface (otherwise it is not on M!). e.g if we choose:

$$y_1 = t\sin(t)$$

$$y_2 = tcos(t)$$

$$\begin{aligned} y_3 &= t^2\sin^2(t) + t^2\cos^2(t) \\ &= t^2(\sin^2(t) + \cos^2(t)) \\ &= t^2 \end{aligned}$$

This gives a spiralling path:

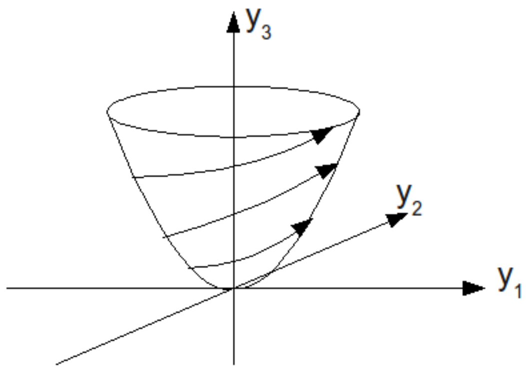

Figure 15: An illustration of the spiralling path of the parabloid of rotation M

Then, to obtain the tangent vector along the path we have to find the *tangent vector field* (see shortly) along the path by taking the derivatives:

$$(\frac{dy_1}{dt}, \frac{dy_2}{dt}, \frac{dy_3}{dt}) = (t\cos(t) + \sin(t), cos(t) - t\sin(t), 2t)$$

What we have here is the field of tangent vectors along the path. To then get a tangent vector we evaluate this *at some fixed point*.
Now, we already know that we can express global coordinates in terms of local ones. We can therefore arrive at an expression for the path in terms of local coordinates:

$$x^1 = y_1 = t\sin(t)$$

$$x^2 = y_2 = t\cos(t)$$

So we can get another form of the tangent vector in terms of the local coordinates using the same derivative of the line with respect to the paramaterization:

$$(\tfrac{dx^1}{dt}, \tfrac{dx^2}{dt}) = (t\cos(t) + \sin(t), cos(t) - t\sin(t))$$

To reiterate, this is also thought of as a tangent vector, but given in terms of local coordinates. We will move on to look at the relationship between these two forms of tangent vector shortly later, in section 4.3.3.

4.3 Algebra of Tangent Vectors

4.3.1 Addition and Scalar Multiplication

When summing or performing scalar multiplication of tangent vectors, the resulting object is also a tangent vector. However, these operations are not necessarily straightforward actions of summing or scalar multiplying paths on our manifold "M". This is because we defined tangent vectors using using paths on "M", so we have to ensure that the vectors resulting from these operations *also produce paths in M*.
However, we *can* add paths by using some charts as follows:
We choose a chart $\overline{\mathrm{x}}$ at $m \in M$ such that the chart has $x(m) = 0$, i.e. the point m is the origin in out chart(this is purely for convenience). Subsequently, the paths $x(f(t))$ and $x(g(t))$ on the chart give two paths through the origin in coordinate space. We may now add these paths or

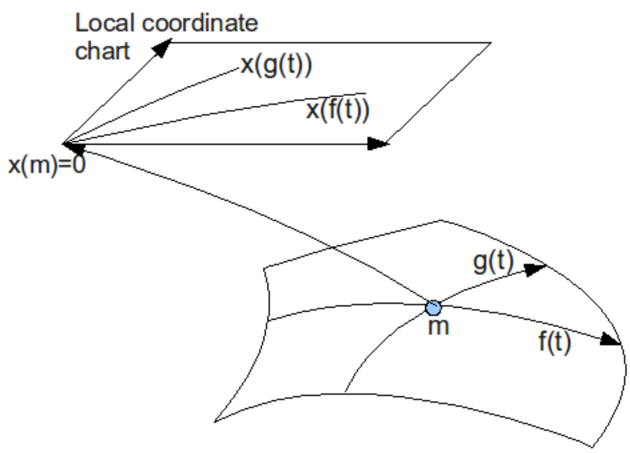

Figure 16: Mapping paths passing through m on M to a chart with m as the origin

scalar multiply them without leaving our local coordinate space. The chart map can then be used to pass the results back to the manifold M. More formally, for addition of two paths $g(t), f(t)$ we would write:

$$f(t) + g(t) = x^{-1}(x(f(t)) + x(g(t)))$$

So the sum of the paths on the manifold is the inverse mapping of the sum of the paths on the chart. Also, for scalar multiplicaiton:

$$\lambda f(t) = x^{-1}(\lambda x(f(t)))$$

Again, the scalar multiple of the path is the inverse mapping of the scalar multiple on the chart. All that we have found so far is the sum or scalar product of two paths. To get the tangent vectors themselves we are requried to follow the same steps as in the example of paraboloid of rotation above. Namely, we have to take the derivates with respect to the path paramater "t" to get the vector field and evaluate this at a point (m) to get the tangent vector there. i.e.

$$\frac{d}{dt}(f(t) + g(t)) = \frac{dx^{-1}(x(f(t))+x(g(t)))}{dt}$$

and

$$\frac{d}{dt}(\lambda f(t)) = \frac{dx^{-1}(\lambda x(f(t)))}{dt}$$

4.3.2 Tangent Space and Vector Space

If M is an n-dimensional manifold, and $m \in M$ (m is a point on M) then the **tangent space at m** is the set T_m of all tangent vectors at m. That is, if we run a large (infinite) number of paths through m spanning all trajectories and represent the associated tangent vectors at m collectively, we arrive at the tangent space.

4.3.3 Local and Global Descriptions of the Tangent Vector

Now we return to examine the fact that we have two ways of describing the coordinates of a tangent vector at a point $m \in M$:

1. Writing the path as $y_i = y_i(t)$ we get the "s" dimensional ambient coordinates of the tangent vector:

$$\vec{y}'(t_o) = (\frac{dy_1}{dt}, ... \frac{dy_s}{dt})_{t=t_o}$$

2. Using some local chart **x** at m, we get the local coordinates of the tangent vector:

$$\mathbf{x}'(t_o) = (\frac{dx^1}{dt}, ... \frac{dx^n}{dt})_{t=t_o}$$

This is all as was stated in the previous 2 sections, but how do we relate these two descriptions of the tangent vector? In general we relate the $\frac{dx^i}{dt}$ and the $\frac{dy^i}{dt}$ via the chain rule.

$$\frac{dy_1}{dt} = \frac{\partial y_1}{\partial x^1}\frac{\partial x^1}{\partial t} + \frac{\partial y_1}{\partial x^2}\frac{\partial x^2}{\partial t} + ... \frac{\partial y_1}{\partial x^n}\frac{\partial x^n}{\partial t}$$

and so on for $\frac{dy_2}{dt}, \frac{dy_3}{dt}$.... Therefore, we can recover the ambient tangent vector coordinates from the local ones. In other words, the local vector coordinates completely define the tangent vector. As a generalization, if the tangent vector "V" has local coordinates $(v^1, v^2, ...v^n)$ and ambient coordinates $(v_1, v_2, ...v_s)$ then they are related by the formulae:

$$v_i = \sum_{k=1}^{n} \frac{\partial y_i}{\partial x^k} v^k$$

To get the amibent coordinates from the local, and:

$$v^i = \sum_{k=1}^{s} \frac{\partial x^i}{\partial y_k} v_k$$

to get the local cooridnates from the ambient. Where, to clarify, these $v_i = \frac{dy_i}{dt}$ and the $v^i = \frac{dx^i}{dt}$. Note, that although we cannot sum or scalar multiply simply the *paths*, to take the sums or scalar multiples of tangent vectors, we are allowed to take the corresponding sums and scalar multiples of the coordinates. In other words:

$$(v + w)^i = v^i + w^i$$

and

$$(\lambda v)^i = \lambda v^i$$

This is what we would expect for vectors in ambient coordinates. The reason we can now do this is that when working in local coordinates we are automatically contstrained to the surface of the manifold. So we dont have to worry about the resulting vectors not belonging to paths in M.
Note, from here on, Einstein's summation convention will be used. That is, that summation over repeated indices in expressions is implied. As an example, the right hand side of the relations between global and local tangent vectors can be rewritten:

$$\sum_{k=1}^{n} \frac{\partial y_i}{\partial x^k} v^k = \frac{\partial y_i}{\partial x^k} v^k$$

and

$$\sum_{k=1}^{n} \frac{\partial x^i}{\partial y_k} v_k = \frac{\partial x^i}{\partial y_k} v_k$$

Since the k index is repeated.

4.4 Example, And Another Formulation Of The Tangent Vector

- Consider M = E_n (n dimensional Euclidean space). For such a space our $\frac{\partial y_i}{\partial x^k} = 1$ and as such, our relation between local and global vectors in section 4.3.3 yields:

  $$v_k = \tfrac{\partial y_i}{\partial x^k} v^k = v^k$$

  Therefore for Euclidean space tangent vectors are the same as "ordinary" vectors.

- In general, if we have a local coordinate system near some point $m \in M$ then we can also obtain a path in terms of global coordinate paramaterization $y_i(t)$ via:

  $$x^j = \begin{cases} t + constant & j = i \\ constant & j \end{cases}$$

  Where the constants are chosen so that the local path $x^j(t_o)$ corresponds to $m \in M$. Then, to view our global path $y_i(t)$ as a function of "t" we apply the parametric equations $y_i = y_i(x^j)$.
  The associated tangent vector at the point t = t_o expressed in a (to me, unusual) form independent of global or local coordinates is:

  $$\tfrac{\partial}{\partial x^i}$$

  It has local coordinates:

  $$\begin{aligned} v^j &= (\frac{dx^j}{dx^i}) \\ &= (\frac{dx^j}{dt}) \\ &= \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} = \delta^j_i \end{aligned}$$

Where δ_i^j is the Kronecker delta. This results occurs because local coordinates should be independent of one another (as, for example our familiar x, y, z axes are). We can then get the ambient cooridnates via the conversion found earlier:

$$v_j = \frac{\partial y_j}{\partial x^k} v^k = \frac{\partial y_j}{\partial x^k} \delta_i^k = \frac{\partial y_j}{\partial x^i}$$

This is again, why we generally refer to the tangent vector as $\frac{\partial}{\partial x^i}$, since it applies equally to the global or local coordinates but is not inherently specified.

Note that everything in the previous consideration is evaluated at t_o and that the nature of the path itself has now disappeared from the definition of the tangent vector!

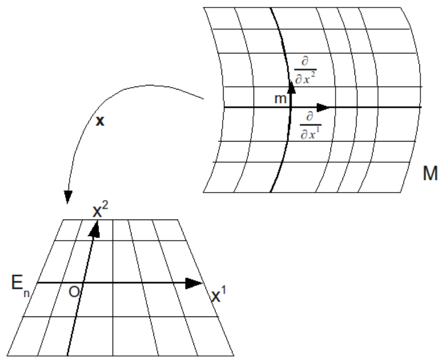

Figure 17: Graphical representation of tangent vectors in our new definition irrespective of path taken

In this formulation, there is a linear one to one correspondance between tangent vectors at m and the "normal" vectors of n dimensional Euclidean space. In other words, tangent space *looks* like E_n.

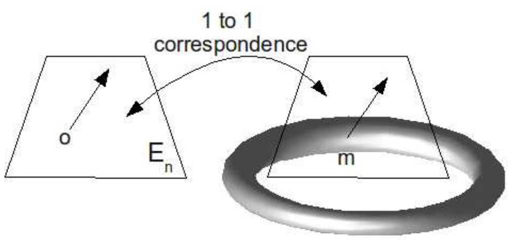

Figure 18: There is a one to one correspondance between tangent vectors at the surface and "normal" vectors in n dimensional Euclidean space.

4.5 Proof Of 1 to 1 Correspondance Between Tangent And "Normal" Vectors

Let T_m be the set of tangent vectors at m (i.e. the tangent space) and define some function "F" which will map this T_m onto n dimensional Euclidean space:

$$F : T_m \rightarrow E_n$$

This mapping will work by assigning to some typical tangent vector its n local coordinates (as we have done throughout).
We also define an inverse function "G" for mapping vectors in Euclidean space onto the tangent plane:

$$G : E_n \rightarrow T_m$$

This inverse is then:

$$G(v^i) = \sum_i v^i \frac{\partial}{\partial x^i}$$

$$v^i \tfrac{\partial}{\partial x^i}$$

These functions are represented graphically in figure 19.

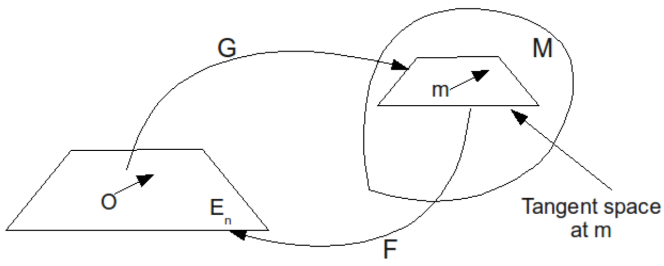

Figure 19: Demonstrating the functions F and G and their relation between the tangent plane at m and the Euclidean plane.

Then, we look to verify that F and G are inverses as follows:

$$F(G(v^i)) = F(v^i \tfrac{\partial}{\partial x^i})$$

However, we have seen that locally for tangent vectors that $\frac{\partial x^j}{\partial x^i} = \delta^j_i$. Substituting this in gives, for the j'th coordinate:

$$F(G(v))^j = v^i \delta^j_i = v^j$$

Which is the j^{th} coordinate of v. Conversely, for the inverse funtion:

$$G(F(w)) = w^i \tfrac{\partial}{\partial x^i}$$

We want to check the ambient coordinates of these results to see if they are the same, because if two vectors have the same ambient coordinates then they are certainly the same vector. We already saw how to find the ambient coordinates of these sums in section 4.3.3. so, applying the same method:

$$G(F(w))_j = w^i \tfrac{\partial y_j}{\partial x^i} = w_j$$

So "F" and its inverse "G" produce local and global tangent vector descriptions respectively. That is, there is one to one correspondence with no "baggage".

5 Covariant And Contravariant Vector Fields

In the previous section we looked at the definiton of tangent vectors. The question which now arises is; how can we relate the coordinates of a given tangent vector in one chart, to those of the same vector in another chart? Once again we use the chain rule:

$$\frac{d\bar{x}^i}{dx^j} = \frac{d\bar{x}^i}{dx^j}\frac{dx^j}{dt}, \text{ or } \bar{v}^i = \frac{d\bar{x}^i}{dx^j}v^j$$

That is, a tangent vector through $m \in M$ is a collection of numbers $v^i = \frac{dx^i}{dt}$ for each chart at m, where the quantities for on echart are related to those for another according to:

$$\bar{v}^i = \frac{d\bar{x}^i}{dx^j}v^j$$

From this, we take the definition of a contravariant vector:

5.1 Contravariant Vectors and Contravariant Vector Fields

A contravariant vector at $m \in M$ is a collection of quantities which transform according to the above expression. That is, a contravariant vector is *just* a tangent vector. Thus, by extension, a contrvariant vector field "V" on M associates with each chart "**x**", a collection of real-valued coordinate functions V^i of local coordinates $(x^1, x^2, ...x^n)$ such that evaluating V^i at any point gives a vector at that point.
The same applies to subsets of M ($U \subset M$) only the vector field domain is restricted to that subset.
So, the transformation rule for all contravariant vector fields is:

$$\bar{V}^i = \frac{d\bar{x}^i}{dx^j}V^j$$

Where now $V^j = V^j(x^1, x^2, ...x^n)$ and $\bar{V}^i = \bar{V}^i(\bar{x}^1, \bar{x}^2, ...\bar{x}^n)$. Obviously this transformation can only apply where charts overlap. Further note that this transformation is reversible (i.e. we can perform $V^j \to \bar{V}^i$ and subsequently $\bar{V}^i \to V^j$). Lastly, if, as was seen in the previous section, V is a smooth contravariant vector field on M then:

$$V = V^i\frac{\partial}{\partial x^i}$$

5.1.1 Example

In section 3 we defined the vectors $\frac{\partial}{\partial x^i}$ as each point $m \in M$ having n vectors $\frac{\partial}{\partial x^i}, (i = 1, 2, 3...n)$ where the i^{th} vector is found by taking the derivative of the path at that point. E.g. for the path x^j (also local coord x^j)

$$\frac{\partial x^j}{\partial x^i} = \begin{cases} t + constant & i = j \\ constant & i \neq j \end{cases}$$

Where the constants are chosen to force the path through m at some value of path paramaterisation $x(t_o)$, hence:

$$(\frac{\partial}{\partial x^i})^j = \delta_i^j$$

Are the coordinates of the vector in chart **x**. That is, in some chart we can force the tangent vectors to be along local coordinate axes.
For these same coordinates in some other, general, chart $\overline{\text{x}}$ are then:

$$\frac{\partial \bar{x}^j}{\partial x^i} = \frac{\partial \bar{x}^k}{\partial x^i}\frac{\partial \bar{x}^j}{\partial x^k}$$
$$\frac{\partial \bar{x}^j}{\partial x^i} = \delta_i^k\frac{\partial \bar{x}^j}{\partial x^k}$$

We define the local vector fields $(\frac{\partial}{\partial x^i})^j = \frac{\partial \bar{x}^j}{\partial x^i}$. This defines a local vector field on a domain U. It is the constant field of unit vectors pointing in the x^i direction. Note that $\frac{\partial}{\partial x^i}$ *is a field*, not the x^{ith} coordinate of the field.
So the $\frac{\partial}{\partial x^i}$ vector field is a unit vector field along all coordinates. It is a generalisation of the way that we relate to contravariant vectors in different local coordinate systems. Note that locally for some specific $\frac{\partial}{\partial x^i}$ it has constant value (given that the local tangent space is Euclidean) however, it will differ under the use of some other chart $\overline{\text{x}}$, for which $\frac{\partial \bar{x}^j}{\partial x^i} \neq constant$ generally.

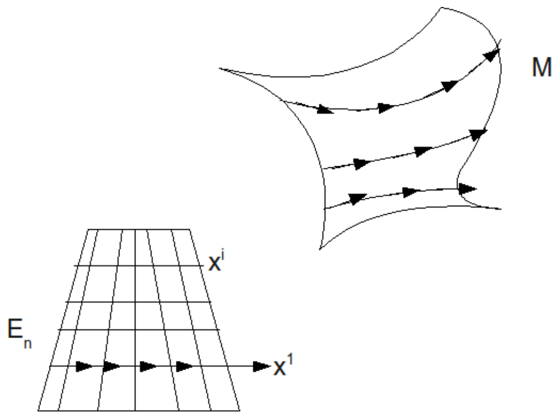

Figure 20: The local vector field along E_1 taken from M.

5.2 Patching Together Local Contavariant Vector Fields

The previous example was specifically for a *local* vector field. Now we look to extend this to consider the entire manifold. In order to do this we have to set the vector field equal to zero near to the edge of its coordinate "patch" (remember that they overlap) so that, again, through the use of many charts we cover the entire manifold.
To achieve this, we draw a disc at x(m) (on the chart x) of radius "r" that is completely contained within the contravariant space. The vector field spanning the whole of M is then:

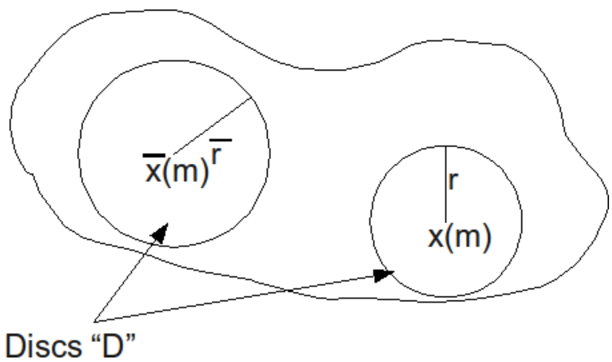

Figure 21: Vector field discs on M.

$$w(p) = \begin{cases} \frac{\partial}{\partial x^j} e^- R^2 & w(p) \in D \\ 0 & otherwise \end{cases}$$

Where $R = \frac{|x(p)-x(m)|}{r-|x(p)-x(m)|}$. That is, if w(p) is in the disc D about some point x(m) of its contrvariant vector space then we use that local space to define the vector field there, otherwise we say it is zero in that local space and move on trying others until we *do* find one that is local. We can do this because as we leave the local coordinate system, the $\frac{\partial}{\partial x^i}$ vanish (tend to zero). Notice that x(m), the vector field, agrees with $\frac{\partial}{\partial x^i}$ at that point and varies smoothly away from that point.

5.3 Covariant Vector Fields

Now we look back at scalar fields on M. If a part of the scalar field has a chart $\bar{x}$ we can analyse the *gadient of the scalar field* which is a locally defined vector field $\frac{\partial \phi}{\partial x^i}$. We transform between this vector field in different charts as follows:

$$\frac{\partial\phi}{\partial\bar{x}^i} = \frac{\partial\phi}{\partial x^j}\frac{\partial x^j}{\partial\bar{x}^i}$$

or we can set $c_j = \frac{\partial\phi}{\partial x^j}$ and $\bar{c}_i = \frac{\partial\phi}{\partial\bar{x}^i}$. We can then define the **covariant vector field** C on M, which associates with each chart **x** a collection of n smooth functions $C_i(x^1, x^2, ...x^n)$ which satisfies the covariant vector transformation rule:

$$\bar{C}_i = C_j\frac{\partial x^j}{\partial\bar{x}^i}$$

Note that for contravariant objects the indices are upper (V^i) and for covariant objects the indices are lower (C_i). I can see the reason for this being the position in the definition of the vector of the local coordinate term:

$$V^i = \frac{\partial x^i}{\partial t}$$
$$C_i = \frac{\partial\phi}{\partial x^i}$$

Generally, as present, we assume scalar and vector fields to be smooth. We saw earlier that geometrically, contravariant vectors are tangent to the manifold... geometrically... *what are* covariant vectors?

Firstly, a one form (or smooth cotangent vector field) on M (or some subset $U \subset M$) is a function "F" that assigns to each smooth contravariant vector field V on M (or U) a smooth scalar field F(V) which has the following properties:

$$F(V+W) = F(V) + F(W)$$
$$F(\alpha V) = \alpha F(V)$$

Where V and W are contravariant vector fields and α is some scalar. We say that there is a one to one correspondence between covariant vector fields on M (or U) and one forms on M (or U). Therefore we can think of covariant vector fields as one forms.

In short a covariant vector field is a transformation of the contravariant field onto the surface of the manifold.

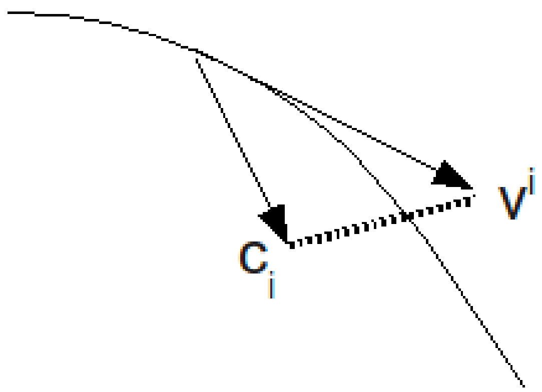

Figure 22

5.3.1 Proof that we can consider C_i as a one fom

If we denote our set of one forms "F" on M (a single one form may not be enough in a similar way to our requiring a number of charts to describe a manifold). We call such a collection a family. We can subsequently talk about a "family" of covariant vector fields "C" on M. Now, we define some function Φ which maps C onto F.

$$\Phi : C \rightarrow F$$

by

$$\Phi(C_i)(V^j) = C_k V^k$$

Now this $C_k V^k$ is in fact:

$$C_k V^k = \frac{\partial \Phi}{\partial x^k}\frac{\partial x^k}{\partial t} = \frac{\partial \Phi}{\partial t}$$

This quantity is some scalar value, which satisfies our description of a one form above.
The inverse function we then define as ψ

$$\Psi : F \rightarrow C$$

by

$$(\Psi(F))_i = F(\frac{\partial}{\partial x^i})$$

To verfiy the validity of this transformation, we need to check if this is a smooth covariant vector field (i.e. are the local components smooth functions of local coordinates and do they transform correctly?). The smoothness comes from the fact that $F\frac{\partial}{\partial x^i}$ is a smooth scalar field (as in our definition of covariant vector field) and hence a smooth function of local coordinates. The transformation can be verified by trying for two charts **x** and $\bar{\mathbf{x}}$:

$$F(\frac{\partial}{\partial x^i}) = F(\frac{\partial x^j}{\partial \bar{x}^i}\frac{\partial}{\partial x^j})$$
$$= \frac{\partial x^j}{\partial x^i}F(\frac{\partial}{\partial x^j})$$

This is exactly the same form as the transformation rule for covariant vectors ($\bar{C}_i = \frac{\partial x^j}{\partial \bar{x}^i}C_j$). We can therefore think of C as one forms.

5.3.2 Example

Let ϕ be some scalar field. Its ambient gradient, $\nabla\phi$ (which we know well) is:

$$\nabla\phi = \frac{\partial \phi}{\partial y_1} + \frac{\partial \phi}{\partial y_2} + ...\frac{\partial \phi}{\partial y_s}$$

This is neither covariant or contravariant (generally), since it is the gradient of the scalar field upon the surface, rather than being a surface bound representation of contravariant vectors. However, we can use this to obtain a covariant vector field as follows:
Choose V as a contravariant vector field, then rate of change of ϕ along V is:

$$V.\nabla\phi$$

This action assigns to every contravariant vector field, the scalar field :

$$F(V) = V.\nabla\phi$$

which tells us how fast ϕ is changing along V. The coodinates corresponding to the covariant vector field are then:

$$F(\frac{\partial}{\partial x^i}) = (\frac{\partial}{\partial x^i}).\nabla\phi$$
$$= [\frac{\partial y_1}{\partial x^i} + \frac{\partial y_2}{\partial x^i} + ...\frac{\partial y_j}{\partial x^i}].[\frac{\partial \phi}{\partial y_1} + \frac{\partial \phi}{\partial y_2} + ...\frac{\partial \phi}{\partial y_s}]$$
$$= \frac{\partial \phi}{\partial x^i}$$

Which is what we first considered when looking at covariant vector fields.
We can in fact, generalize the above. If Σ is any smooth vector field of M, then the operation $V.\Sigma$ results in a transformation from a smooth tangent field to a smooth scalar field. Therefore, a one form on M with local coordinates is given by applying the linear function to the charts $\frac{\partial}{\partial x^i}$:

$$C_i = \frac{\partial}{\partial x^i}.\Sigma$$

In the previous example $\Sigma = \nabla\phi$, but to reiterate, any smooth vector field will also do. Note that this operation depends only on the tangential component of Σ.
If V is any tangent (contravariant) field, then we can refer to the generalized equation given above to obtain an associated covariant field. The coordinates of this field are not the same as those of V. To find them, we write $V = V^i(\frac{\partial}{\partial x^i})$ and use this in place of Σ in the expression for C_j to give:

$$C_j = \frac{\partial}{\partial x^j}.V^i\frac{\partial}{\partial x^i} = V^i\frac{\partial}{\partial x^j}\cdot\frac{\partial}{\partial x^i}$$

The dot product in the above expression is not necessarily orthogonal (i.e. not necessarily δ_{ij}. We thus define certain functions:

$$g_{ij} = \frac{\partial}{\partial x^j}\cdot\frac{\partial}{\partial x^i}$$

i.e.

$$C_j = g_{ij}V^i$$

Note that the upper and lower indices on the right hand side of the above expression "cancel" just leaving a lower index. We will move on to look at these g_{ij} As a final note, for any *pair* of covariant *or* contravariant vector field, the following operations are permitted:

$$(V+W)^i = V^i + W^i \ (\alpha V)^i = \alpha V^i$$

These actions convert a set of vector fields into a vector space. Note that we cannot expect to obtain a vector field through addition of a covariant field and a contravariant field.

6 Tensor Fields

We previously looked at vector fields on manifolds, now we move on to look at the notion of tensors on smooth manifolds. At present (for simplicity) we will consider 3 dimensional Euclidean space E_3. Consider two vector fields on E_3, **V** = (V_1, V_2, V_3) and **W** = (W_1, W_2, W_3). The *Tensor Poduct* of these vector fields consist of the nine quantities:

$$\begin{pmatrix} V_1W_1 & V_1W_2 & V_1W_3 \\ V_2W_1 & V_2W_2 & V_2W_3 \\ V_3W_1 & V_3W_2 & V_3W_3 \end{pmatrix}$$

How do such quantities transform? We need to consider the thee cases of: V_iW_j = contravariant, V_iW_j = covariant and the mixed case. Therefore, let V and W represent contravariant vector fields and C and D be covariant vector fields. We then get the following possibilities for tansformation:

1. **Two contravariant vector fields:**

$$\bar{V}^i\bar{W}^j = \frac{\partial \bar{x}^i}{\partial x^k}V^k\frac{\partial \bar{x}^j}{\partial x^m}W^m \quad = \quad \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial \bar{x}^j}{\partial x^m}V^kW^m$$

2. **Covariant vector field and contravariant vector field:**

$$\bar{V}^i\bar{C}_j = \tfrac{\partial \bar{x}^i}{\partial x^k}\tfrac{\partial x^m}{\partial \bar{x}^j}V^kC_m$$

3. **Two covariant vector fields:**

$$\bar{C}_i\bar{D}_j = \tfrac{\partial x_k}{\partial \bar{x}^i}\tfrac{\partial x_m}{\partial \bar{x}^j}C_kD_m$$

These product fields are known as tensors. (1) is type (2,0), (2) is type (1, 1) and (3) is type (0, 2). In short the "type" is (no. of contravariant indices, no. of covariant indices). A tensor field on an n-dimensional smooth manifold M, associates with each chart **x** a collection of n^2 smooth functions which satisfy their appopriate transformation rules:
Transformation Rules:

1. Type (2, 0):

$$\bar{T}^{ij} = \tfrac{\partial \bar{x}^i}{\partial x^k}\tfrac{\partial \bar{x}^j}{\partial x^m}T^{km}$$

   This is the transformation between contravariant rank 2 tensors.

2. Type(1, 1):

$$\bar{T}^i_j = \tfrac{\partial \bar{x}^i}{\partial x^k}\tfrac{\partial x^m}{\partial \bar{x}^j}T^k_m$$

   This is the transformation between mixed contravariant rank 1 and covariant rank 1 tensors.

3. Type (0, 2):

$$\bar{T}_{ij} = \tfrac{\partial x^k}{\partial \bar{x}^i}\tfrac{\partial x^m}{\partial \bar{x}^j}T_{ij}$$

   This is the transformation between covariant rank 2 tensors.

A tensor field of the type (1, 0) is just a contravariant vector field, similarly a tensor field of type (0, 1) is a covariant vector field. A type (0, 0) is simply a scalar field. Note that we can add and scalar multiply tensor fields in a similar way to vector fields, for example, if A and B are both type (1, 2) tensors, then their sum is given by:

$$(A+B)^c_{ab} = A^c_{ab} + B^c_{ab}$$

As an example, the Kronecker delta is actually a Tensor

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

it is a type (1, 1) tensor, where:

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

where, as a reminder, the x^i and x^j are the i^{th}/j^{th} local coordinates. This transforms according to:

$$\bar{\delta}^i_j = \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^k}{\partial x^m}\frac{\partial x^m}{\partial \bar{x}^j}$$
$$= \frac{\partial \bar{x}^i}{\partial x^k}\delta^k_m\frac{\partial x^m}{\partial \bar{x}^j}$$

By comparison with the transformation rules above we see that this is a tensor field of type (1, 1). Note that since $\delta^i_j = \bar{\delta}^i_j$ and $\delta^i_j = \delta^i_j$, therefore δ^i_j is a *symmetric* tensor. So, given some point "p" on the manifold and a chart **x** at p, then this tensor assigns n^2 quantities δ^i_j which is the identity matrix $\frac{\partial x^i}{\partial x^j}$ that is independent of the chart we chose.
Interpreting this object: In section 4, we saw that covariant vector fields convert contravariant vector fields into scalars, we shall see that a type (1, 1) tensor converts contravariant fields to other contravariant fields. This particular tensor δ^j_i does very little, if we put in a specific vector field V, and you obtain the same vector field as was input as the result. This δ^j_i is an identiy transformation.
We can make new tensor fields out of existing tensors in various ways, for example:

$$M^i_{jk}N^{pq}_{ns}$$
$$M^i_{jk}N^{jk}_{rs}$$

The upper product results in a type (3, 4) tensor, while the lower results in a type (1, 2) tensor since the j, k indices cancel. As an example, consider the transformation rules for type (1, 1) and (2, 0) tensors and take the product of them:

$$\bar{T}^i_j = \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^m}{\partial \bar{x}^j}T^k_m$$

And

$$\bar{T}^{ij} = \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial \bar{x}^j}{\partial x^m}T^{km}$$

$$\begin{aligned}\bar{T}^j_i\bar{T}^{ij} &= \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^m}{\partial \bar{x}^j}\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial \bar{x}^j}{\partial x^m}T^k_mT^{km}\\ &= \left(\frac{\partial \bar{x}^i}{\partial x^k}\right)^2T^k_mT^{km}\end{aligned}$$

$$(\bar{T}^i)^2 = \left(\frac{\partial \bar{x}^i}{\partial x^k}\right)^2(T^k)^2$$

$$\bar{T}^i = \frac{\partial \bar{x}^i}{\partial x^k}T^k$$

And we recover the contravariant tensor transformation rule (i.e. we *have* a contravariant tensor!).
If X is some contravariant vector field, then the functions $\frac{\partial X^i}{\partial x^j}$ do not define a tensor. To verify, check the transformation rule:

$$\frac{\partial \bar{X}^i}{\partial \bar{x}^j} = \frac{\partial}{\partial \bar{x}^j}\left(X^k\frac{\partial \bar{x}^i}{\partial x^k}\right)$$

Then, we can rewrite this as:

$$\frac{\partial \bar{x}^i}{\partial \bar{x}^j} = \frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial}{\partial x^h}\left(X^k\frac{\partial \bar{x}^i}{\partial x^k}\right)$$

And differentiate by parts

$$u = X^k \qquad v = \frac{\partial \bar{x}^i}{\partial x^k}$$
$$\frac{\partial u}{\partial x^h} = \frac{\partial X^k}{\partial x^h} \quad \frac{\partial v}{\partial x^h} = \frac{\partial^2 \bar{x}^i}{\partial x^h \partial x^k}$$

$$\frac{\partial \bar{X}^i}{\partial \bar{x}^j} = \frac{\partial X^k}{\partial x^h}\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^h}{\partial \bar{x}^j} + X^k \frac{\partial^2 \bar{x}^i}{\partial x^h \partial x^k}$$

This transformation gives some extra term on the right hand side which violates our transformation rules.
Generally, if we are given a smooth local function g_{ij} with the property that for every pair of contravariant vector fields X^i and Y^j, the smooth functions $g_{ij}X^iY^j$ determine a scalar field, then the g_{ij} determine a smooth tensor field of type (0, 2).

6.1 Proof of the above statement

If $\bar{g}_{ij}\bar{X}^i\bar{Y}^j = g_{hk}X^hY^k ...eqn(1)$, then we require, since $g_{hk} = \frac{\partial}{\partial x^h} \cdot \frac{\partial}{\partial x^k}$ that:

$$\bar{g}_{ij}\bar{x}^i\bar{y}^j = \bar{g}_{ij}X^hY^k \frac{\partial \bar{x}^i}{\partial x^h}\frac{\partial \bar{x}^j}{\partial x^k} ...eqn(2)$$

We then equate the right hand sides of both (1) and (2) to obtain:

$$g_{hk}X^hY^k = \bar{g}_{ij}\frac{\partial \bar{x}^i}{\partial x^h}\frac{\partial \bar{x}^j}{\partial x^k} ...(3)$$

If we can find a way to remove the X^hY^k then we would arrive at the covariant transformation rule. If we think back to section 4, where we discussed how a covariant field can be applied to an entire manifold, we saw that evaluating at the point $x(m)$, then we would get:

$$g_{hk} = \bar{g}_{ij} = \frac{\partial \bar{x}^i}{\partial x^h}\frac{\partial \bar{x}^j}{\partial x^k}$$

At m. In addition, we can follow the section 4 example further by saying:

$$X^i(m) = \begin{cases} 1 & i = h \\ 0 & i \neq h \end{cases}, Y^j(m) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

Finally, substituting these into equation (3) gives the type (0, 2) covariant tensor tansformation rule:

$$g_{hk} = \bar{g}_{ij}\frac{\partial \bar{x}^i}{\partial x^h}\frac{\partial \bar{x}^j}{\partial x^k}$$

6.2 The Metric Tensor

We had defined a set of quantities g_{ij} by:

$$g_{ij} = \frac{\partial}{\partial x^j} \cdot \frac{\partial}{\partial x^i}$$

Now, if X^i and Y^i are any contravariant fields on M, then $\mathbf{X} \cdot \mathbf{Y}$ is scalar, and:

$$\mathbf{x}.\mathbf{y} = X^i \frac{\partial}{\partial x^i}.Y^j \frac{\partial}{\partial x^j} = g_{ij}X^iY^j$$

This g_{ij} is a type (0, 2) tensor. We call this tensor the *metric tensor* inherited from the embedding of M in E_s

7 Riemannian Manifolds

7.1 The Inner Product

The inner product is a generalization of the familiar dot product. In vector space it multiplies vectors together to form scalars.
A *smooth inner product* on a manifold M is a smooth function $<,>$, which associates with each pair of contravariant vector fields X, Y a scalar $< X, Y >$, which has the following properties:

- Symmetry: $< x, y >=< y, x >$
- Bilinearity: $< ax, by >= ab < x, y >$
- Non Degeneracy: if $< x, y >= 0$ for all y then x=0

Such an object is known as a symmetric bilinear form. A manifold endowed with a smooth inner product is known as a *Riemannian manifold*. Thus, if **x** is a chart and p is some point in the domain of **x** then:

$$< x, y >= x^i y^j < \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} >$$

Where $< x, y >$ is the scalar field, x^i, y^j are contravariant vectors and $< \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} >$ is a smooth inner product which we write as g_{ij}. These g_{ij} are the type(0,2) fundamental or metric tensors of a Riemannian manifold.

7.1.1 Examples:

- If $M = E_n$, we already saw that $g_{ij} = \delta_{ij}$
- **The Minkwoski metric:**

$$g_{ij} = G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}$$

Where c is the speed of light. We call this manifold *flat Minkowski space* M^4. In Minkowski space the length of vectors is different to that of Euclidean space. The length is determined by the metric...
In 3D Euclidean space:

$$d(x, y) = \left[(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2\right]^{1/2}$$

Whereas in Minkowski 4-space:

$$d(x, y) = \left[(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2 - c^2(y_4 - x_4)\right]^{1/2}$$

There are therefore some major differences between Euclidean and Minkowski space. In Euclidean space, the set of all points a distance r from a point in E_3 is just a sphere of radius r. In Minkowski space, the set of all points a distance r from a point in M^4 is a hyperbolic surface. There is also an interesting difference at $r = 0$ where in Euclidean space we get a point and in M^4 we get a cone. This cone is called a light cone.

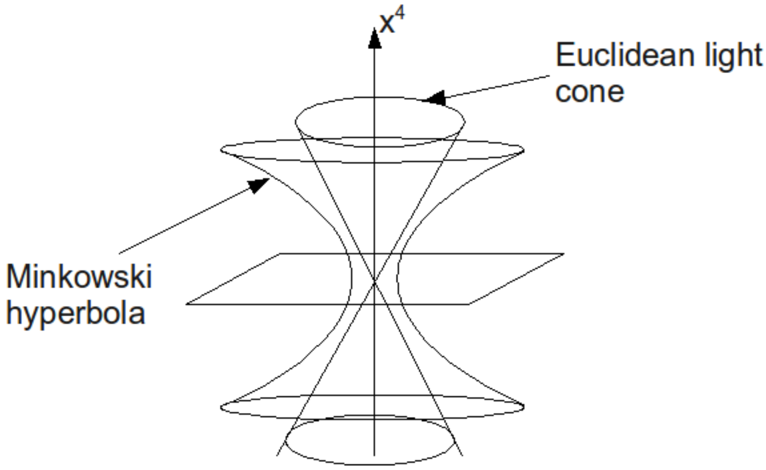

Figure 23: Euclidean light cone and minkowski hyperbola schematic.

We can use the inner product to find the metric of a manifold. For example, if we wish to find the metric for our usual example, a sphere of radius r and local coordinates $x^1 = \theta$ and $x^2 = \phi$. In order to obtain the metric we need to find the inner product of the basis vectors $\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}$ in the ambient space E_s.

$$\text{j}^{th}\text{ambient coordinate} = \frac{\partial y^i}{\partial x^i}$$

Where, for our sphere:

$$\begin{aligned} y_1 &= r\sin(x^1)\cos(x^2) \\ y_2 &= r\sin(x^1)\sin(x^2) \\ y_3 &= r\cos(x^1) \end{aligned}$$

So, finding our collection of ambient coordinates, $\frac{\partial y_1}{\partial x^1}, \frac{\partial y_2}{\partial x^1}, \frac{\partial y_3}{\partial x^1}$

$$\begin{aligned} \frac{\partial}{\partial x^1} &= \left(r\cos(x^1)\cos(x^2), r\cos(x^1)\sin(x^2), -r\sin(x^1)\right) \\ \frac{\partial}{\partial x^2} &= \left(-r\sin(x^1)\sin(x^2), r\sin(x^1)\cos(x^2), 0)\right) \end{aligned}$$

Now, remembering that $(<,>)$ is an extension of the dot product, we analyse g_{ij}...

$$
\begin{aligned}
g_{11} &= < \tfrac{\partial}{\partial x^1}, \tfrac{\partial}{\partial x^1} > \\
g_{11} &= r^2 < (\cos(x^1)\cos(x^2), \cos(x^1)\sin(x^2), -\sin(x^1)), (\cos(x^1)\cos(x^2), \cos(x^1)\sin(x^2), -\sin(x^1)) > \\
&= r^2 \left[\cos^2(x^1)(\cos^2(x^2) + \sin^2(x^2)) + \sin^2(x^1)\right] \\
&= r^2 \left[\cos^2(x^1) + \sin^2(x^1)\right] \\
&= r^2
\end{aligned}
$$

$$
\begin{aligned}
g_{22} &= < \tfrac{\partial}{\partial x^2}, \tfrac{\partial}{\partial x^2} > \\
&= r^2 < (-\sin(x^1)\sin(x^2), \sin(x^1)\cos(x^2), 0), (-\sin(x^1)\sin(x^2), \sin(x^1)\cos(x^2), 0) > \\
&= r^2 \left[\sin^2(x^1)\sin^2(x^2) + \sin^2(x^1)\cos^2(x^2)\right] \\
&= r^2 \left[\sin^2(x^1)(\sin^2(x^2) + \cos^2(x^2))\right] \\
&= r^2 \sin^2(x^1)
\end{aligned}
$$

$$
\begin{aligned}
g_{12} = g_{21} &= < \tfrac{\partial}{\partial x^1}, \tfrac{\partial}{\partial x^2} > \\
&= r^2 < (\cos(x^1)\cos(x^2), \cos(x^1)\sin(x^2), -\sin(x^1)), (-\sin(x^1)\sin(x^2), \sin(x^1)\cos(x^2), 0) > \\
&= r^2 \left[-\cos(x^1)\cos(x^2)\sin(x^1)\sin(x^2) + \cos(x^1)\sin(x^2)\sin(x^1)\cos(x^2) + 0\right] \\
&= r^2 [0] \\
&= 0
\end{aligned}
$$

$$
\therefore g_{ij} = \begin{pmatrix} r^2 & 0 \\ 0 & r^2\sin^2(x^1) \end{pmatrix}
$$

7.2 Diagonalizing The Metric

If we let G be the matrix of g_{ij} in a local coordinate system, evaluated at some point p on a Riemannian manifold. From our rules, the scalar field resulting from an inner product is symmetric. It then follows, from study of matrices in other modules, that there is some invertible matrix $P_{ij} = P_{ji}$:

$$
PGP^T = \begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & ... \\ ... & 0 & 0 & \pm 1 \end{pmatrix}
$$

Where P^T is the transpose of P. The sequence $(\pm 1, \pm 1, \pm 1, \pm 1)$ is called the signature of the metric at p.

- Euclidean E_3 signature: (1, 1, 1)
- Minkowski M^4 signature: (1, 1, 1, -1)

We can now define change of coordinates via:

$$
x^i = P_{ij}\bar{x}^j
$$

then $\frac{\partial x^i}{\partial \bar{x}^j = P_{ij}}$ and so:

$$
\begin{aligned}
g_{ij} &= \tfrac{\partial x^a}{\partial \bar{x}^i} g_{ab} \tfrac{\partial x^b}{\partial \bar{x}^j} = P_{ia} g_{ab} P_{jb} \\
& P_{ia} g_{ab} (P^T)_{bj} = (PGP^T)_{ij}
\end{aligned}
$$

with, at the point p:

$$
g_{ij} = \begin{pmatrix} \pm 1 & 0 & 0 & ... \\ 0 & \pm 1 & & \\ 0 & & & \\ ... & & & \pm 1 \end{pmatrix}
$$

$\therefore$ for the metric, the unit base vectors $e_i = \frac{\partial}{\partial \bar{x}^i}$ are orthogonal; that is:

$$< e_i, e_j >= \pm\delta_{ij}$$

Note that in our initial conditions we stated that the smooth inner product is non degenerate, i.e. if $< x, y >= 0$ for all y then x = 0. This is equivalent to requiring that the $det(g_{ij}) \neq 0$. That is, we don't want any singularities (the determinant gives an area enclosed by some points, we don't want this to be zero!).

7.3 The Square Norm

If X is a contravariant vector field of M, then the *square norm* of X is given by:

$$\parallel X \parallel^2 =< X, X >= g_{ij}X^iX^j$$

If $\parallel X \parallel^2 <, >, = 0$ we call X timelike, spacelike and null respectively. If X is not spacelike, then we can define:

$$\parallel X \parallel = \sqrt{\parallel X \parallel^2} = (g_{ij}X^iX^j)^1/2$$

Given that $< X, X >$ is a scalar field, $\parallel X \parallel$ is also a scalar field.

7.4 Arc length

A path C on a manifold which is paramaterised as $x = x(t)$ is non-null if:

$$\parallel \frac{dx^i}{dt} \parallel^2 = g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} \neq 0$$

in which case it is always timelike or spaceline. If C is a non null path in M, then we define its length as follows...
Break the path into segments S, each of which lie in some coordinate neighborhood and define the length of S by:

$$L(a, b) = \int_a^b \left[\pm g_{ij}\frac{dx^i}{dt}.\frac{dx^j}{dt}\right]^1 /2dt$$

where the $\pm$ is chosen as +1 for spacelike and -1 for timelike curves. Equivalently:

$$L(a, b) = \int_a^b \sqrt{\parallel \frac{dx^i}{dt} \parallel^2} dt = \int_a^b \parallel \frac{dx^i}{dt} \parallel dt$$

In differential form, we are defining the *arc-length* as:

$$ds^2 = \pm g_{ij}dx^idx^j$$

This definition is independent of the chart chosen since the resultant value is a scalar.

8 Covariant Differentiation

When considering a parallel vector field on a manifold, we check if it is parallel by following a path and taking the derivative of the vector field iwth respect to the paramaterised path t. In rectilinear coordinates we should get zero, however this may not be the case in curvilinear coordinates, where the vector field may change direction as we move along the curved coordinate axis. i.e. if X^j is a field, we check for parallelism by taking $\frac{dX^j}{dt}$ along the path $x^i = x^i(t)$... however, there are some issues...

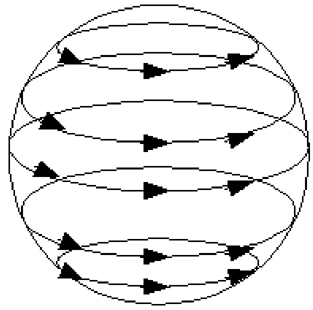

Figure 24: A parallel field on a sphere.

Geometric Issue: If we consider a parallel field on a sphere, The field is circulating and therefore non-constant, $\frac{dX}{dt} \neq 0$, which does not result in parallelism! (However, the projection of $\frac{dX}{dt}$ parallel to the manifold *does* vanish - see later).

Algebraic Issue: Since $\bar{X}^j = \frac{\partial \bar{X}^j}{\partial X^h} X^h$, we get by the product rule:

$$\frac{d\bar{X}^j}{dt} = \frac{\partial^2 \bar{X}^j}{\partial X^k \partial X^h} \frac{dX^k}{dt} + \frac{d\bar{X}^j}{dX^h} \frac{dX^h}{dt}$$

This shows that if the second derivative does not vanish, $\frac{d\bar{X}^j}{dt}$ does not transform a vector field. We therefore cannot check for parallelism.
The projection of $\frac{dX}{dt}$ along M will be called a *covariant derivative* of X w.r.t t and is written $\frac{DX}{dt}$.

8.1 Working Towards The Covariant Derivative

8.1.1 Projection Onto The Tangent Space

For a manifold embedded in s-dimensional Euclidean space with a metric g as consequence (of the embedding), a vector V in E_s has a projection πV to a local vector T_m with coordinates:

$$(\pi V)^i = g^{ik}(V.\frac{\partial}{\partial x^k}) \tag{1}$$

Where g^{ik} is the matrix inverse of g_{ij} and $g_{ij} = \frac{\partial}{\partial x^1}\frac{\partial}{\partial x^2}$ as before. I am going to take this as a given for now.

8.1.2 Christoffel Symbols

The partial derivatives of the metric in terms of its ambient coordinates are, using the product rule:

$$\begin{aligned} \frac{\partial}{\partial x^p} g_{qr} &= \frac{\partial}{\partial x^p}\left[\frac{\partial y_s}{\partial x^q}\frac{\partial y_s}{\partial x^r}\right] \\ &= \frac{\partial^2 y_s}{\partial x^p \partial x^q}\frac{\partial y_s}{\partial x^r} + \frac{\partial^2 y_s}{\partial x^r \partial x^p}\frac{\partial y_s}{\partial x^q} \end{aligned}$$

We write this using “comma notation”:

$$g_{qr,p} = y_{s,pq} y_{s,r} + y_{s,rp} y_{s,q}$$

Now, by cycling through the indices q, r and p (permuting them) we get 3 equations in total:

$$g_{qr,p} = y_{s,pq}y_{s,r} + y_{s,rp}y_{s,q}$$
$$g_{rp,q} = y_{s,qr}y_{s,p} + y_{s,pq}y_{s,r}$$
$$g_{pq,r} = y_{s,rp}y_{s,q} + y_{s,qr}y_{s,p}$$

We can then solve for $y_{s,pq}y_{s,r}$ by adding the first two equations and subtracting the first.

$$g_{qr,p} + g_{rp,q} = 2y_{s,pq}y_{s,r} + y_{s,rp}y_{s,q} + y_{s,qr}y_{s,p}$$
$$g_{qr,p} + g_{rp,q} - g_{pq,r} = 2y_{s,pq}y_{s,r} + \not{y}_{s,rp}y_{s,q} + \not{y}_{s,qr}y_{s,p} - \not{y}_{s,rp}y_{s,q} - \not{y}_{s,qr}y_{s,p}$$

We then make the following definition of the Christoffel symbol by rearranging for $y_{s,pq}y_{s,r}$:

- **Christoffel Symbol Of The First Kind:**

$$[pq,r] = 1/2[g_{qr,p} + g_{rp,q} - g_{pq,r}]$$

- **Christoffel Symbol Of The Second Kind:**

$$\left\{ {i \atop pq} \right\} = g^{ir}[pq,r] = \frac{1}{2}g^{ir}[g_{qr,p} + g_{rp,q} - g_{pq,r}]$$

These quantities are *not* tensors, instead they transform as follows:

- **Transformation law for first kind:**

$$[hk,l] = [ri,j]\frac{\partial \bar{x}^r}{\partial x^h}\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial \bar{x}^j}{\partial x^l} + \bar{g}_{ij}\frac{\partial^2 \bar{x}^i}{\partial x^h \partial x^k}\frac{\partial \bar{x}^j}{\partial x^l}$$

- **Transformation law for second kind:**

$$\left\{ {p \atop hk} \right\} = \left\{ {t \atop ri} \right\}\frac{\partial x^p}{\partial x^t}\frac{\partial \bar{x}^r}{\partial x^h}\frac{\partial \bar{x}^i}{\partial x^k} + \frac{\partial x^p}{\partial \bar{x}^t}\frac{\partial^2 \bar{x}^t}{\partial x^h \partial x^k}$$

We can now use this to work towards an expression for the covariant derivative. By its definition as the projection of the derivative of the vector field onto the manifold, we get:

$$\frac{DX}{dt} = \pi\frac{dX}{dt}$$

Which, via equation 1, has local coordinates:

$$\frac{DX^i}{dt} = g^{ir}\left(\frac{dX}{dt}\cdot\frac{\partial}{\partial x^r}\right)$$

The bracekted term is evaluated using global coordinates. $\frac{dX}{dt}$ has global coordinates given by:

$$\frac{dX}{dt} = \frac{d}{dt}\left(X^p\frac{\partial y_s}{\partial x^p}\right)$$
$$= \frac{dX^p}{dt}\frac{\partial y_s}{\partial x^p} + X^p\frac{\partial^2 y_s}{\partial x^p \partial x^q}\frac{dx^q}{dt}$$

Then, substituting this into the bracket, we get:

$$\frac{dX}{dt}\cdot\frac{\partial y_s}{\partial x^r} = \frac{dX^p}{dt}\frac{\partial y_s}{\partial x^p}\frac{\partial y_s}{\partial x^r} + X^p\frac{\partial^2 y_s}{\partial x^p \partial x^q}\frac{\partial y_s}{\partial x^r}\frac{\partial x^q}{\partial t}$$
$$= \frac{dX^p}{dt}g_{pr} + X^p[pq,r]\frac{dx^p}{dt}$$

So, putting this back into our expression for covariant derivative:

$$\frac{DX^i}{dt} = g^{ir}\left(\frac{dX}{dt}\cdot\frac{\partial}{\partial x^r}\right)$$
$$= g^{ir}\left(\frac{dX^p}{dt}g_{pr} + X^p[pq,r]\frac{dx^q}{dt}\right)$$

We get

$$\frac{DX^i}{dt} = \delta^i_p\frac{dX^p}{dt} + X^p\left\{ {i \atop pq} \right\}\frac{dx^q}{dt}$$
$$\frac{DX^i}{dt} = \frac{dX^i}{dt} + X^p\left\{ {i \atop pq} \right\}\frac{dx^q}{dt}$$

This is the expression for the *covariant derivative*. For a general Riemannian manifold, it is a contravariant vector. This allows us, for a field of vectors of constant length, to check if it is parallel by asking $\frac{DX^i}{dt} = 0$?

8.2 The Covariant Partial Derivative

If we re-write the expression for covariant derivative as:

$$\frac{DX^i}{dt} = \frac{\partial X^i}{\partial x^q}\frac{dx^q}{dt} + \left\{ {i \atop pq} \right\} X^p \frac{dx^q}{dt}$$
$$= \left[\frac{\partial X^i}{\partial x^q} + \left\{ {i \atop pq} \right\} X^p \right] \frac{dx^q}{dt}$$

The quantity in brackets is thus responsible for the conversion of the vector $\frac{dx^q}{dt} \rightarrow \frac{DX^i}{dt}$. It is a type (1,1) tensor which we call the q^{th} covariant partial derivative of X^i...

- **Covariant partial derivative of X^i:**

$$X^i_{|q} = \frac{\partial X^i}{\partial x^q} + \left\{ {i \atop pq} \right\} X^p$$

We now see where the name "covariant derivative" comes from, since the covariant partial derivative is always using the covariant index! Similarly, for global coordinates we get

- **Covariant partial derivative of Y_p**

$$Y_{p|q} = \frac{\partial Y_p}{\partial x^q} - \left\{ {i \atop pq} \right\} Y_p$$

9 The Riemann Curvature Tensor

For a vector field X^j, we showed in the previous section that a parallel vector field of constant length M must satisfy $\frac{DX^j}{dt} = 0$ for any path in M. The vector field X^j is parallel along a curve if it satisifes:

$$\frac{DX^j}{dt} = \frac{dX^j}{dt} + \Gamma^j_{ih} X^i \frac{dx^h}{dt} = 0$$

For some specified curve. Here we are writing the Christoffel symbol $\left\{ {j \atop ih} \right\}$ as Γ^j_{ih}

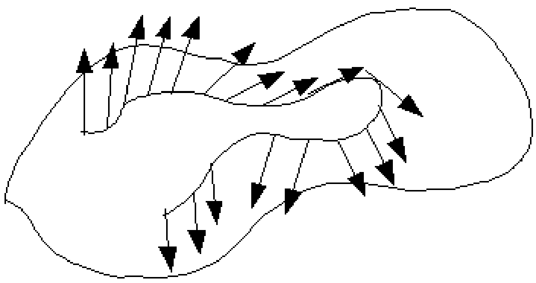

Figure 25: A parallel vector field along a curve C on some manifold M.

If X^j is parallel along the curve, which has paramaterization with domain [a,b] and corresponding points α and β on M, then, since:

$$\frac{dX^j}{dt} = -\Gamma^j_{ih} X^i \frac{dx^h}{dt} \qquad \text{... (I)}$$

we can integrate to obtain

$$X^j(\beta) = X^j(\alpha) - \int_a^b \Gamma^j_{ih} X^i \frac{dx^h}{dt} \qquad \text{... (II)}$$

If $X^j(\alpha)$ is any vector at the point $\alpha \in M$ and if the curve C is any path from α to β in M, then the *parallel transport* of $X^j(\alpha)$ along C is the vector $X^j(\beta)$ given by the solution to equn I with initial conditions given by $X^j(\alpha)$.

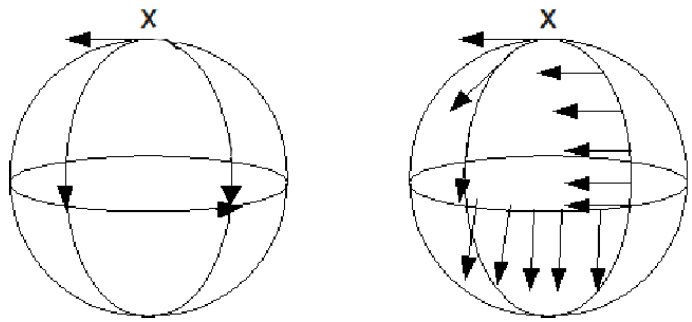

Figure 26: Parallel transport on the surface of a sphere from a starting vector.

9.1 Working Towards The Curvature Tensor

First, we ask under what conditions parallel transport is independent of the path taken? If this is the case then $\frac{DX^j}{dt} = 0$ and so equation I can be used to create a parallel vector field of constant length on M (see previous section on covariant derivative).
We investigate this by taking some fixed vector $V = X^j(a)$ and parallel transporting it round a small rectangle, where we call the first two coordinates of the start point:

$$x^1(a) = r, x^2(a) = s$$

We then choose δs and δr to be so small that the following paths are written within the coordinate neighborhood in question:

$$C_1 : x^j(t) = \begin{cases} x^i(a) & i \neq 1,2 \\ r + t\delta r & i = 1 \\ s & i = 2 \end{cases}$$
$$C_2 : x^j(t) = \begin{cases} x^i(a) & i \neq 1,2 \\ r + \delta r & i = 1 \\ s + t\delta s & i = 2 \end{cases}$$
$$C_3 : x^j(t) = \begin{cases} x^i(a) & i \neq 1,2 \\ r + (1-t)\delta r & i = 1 \\ s + \delta s & i = 2 \end{cases}$$
$$C_4 : x^j(t) = \begin{cases} x^i(a) & i \neq 1,2 \\ r & i = 1 \\ s + (1-t)\delta s & i = 2 \end{cases}$$

These paths are shown in figure 27 Now, if we parallel transport $X^j(a)$ along C, we must have via equation II:

$$X^j(b) = X^j(a) - \int_0^1 \Gamma^j_{ih} X^i \frac{dx^h}{dt} dt$$

Where we translate the vector from the origin to $C_1 = 1$. This is thus:

$$X^j(b) = X^j(a) - \int_0^1 \Gamma^j_{i1} X^i dr dt$$

The integrand term $\Gamma^j_{i1} X^i$ varies as a function of the paramaterizing variable t on C_1. However, if the path is a small one, the integral is approximately equal to the value of the integrand at the midpoint of the path segment...

$$\begin{aligned} X^j(b) &= X^j(a) - \Gamma^j_{i1} X^i(middleOfC_1)\delta r \\ &\approx X^j(a) - (\Gamma^j_{i1} X^i(a) + 0.5\frac{\partial}{\partial x^1}(\Gamma^j_{i1} X^i)\delta r)\delta r \end{aligned}$$

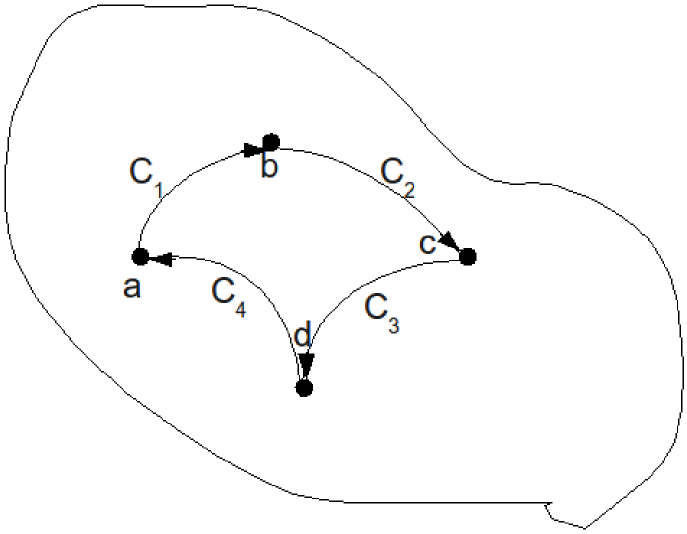

Figure 27: Path segments which form a small rectangle for the study of parallel transport.

Where the partial derivative is evaluated at the point a. Similarly:

$$\begin{aligned} X^j(c) &= X^j(b) - \int_0^1 \Gamma^j_{i2} X^i \delta s dt \\ &\approx X^j(b) - (\Gamma^j_{i2} X^i(middleOfC_2))\delta s \\ &\approx X^j(b) - (\Gamma^j_{i2} X^i(a) + \frac{\partial}{\partial x^1}(\Gamma^j_{i2})\delta r + 0.5\frac{\partial}{\partial x^2}(\Gamma^j_{i2} X^i)\delta s)\delta s \end{aligned}$$

Where, again, all partial derivatives are evaluated at point a. Continuing with the rest of the path...

$$\begin{aligned} X^j(d) &= X^j(c) - \int_0^1 \Gamma^j_{i1} X^i \delta r dt \\ &\approx X^j(c) - (\Gamma^j_{i1} X^i(middleOfC_3))\delta r dt \\ &\approx X^j(c) - (\Gamma^j_{i1} X^i(a) + 0.5\frac{\partial}{\partial x^1}(\Gamma^j_{i1} X^i)\delta r + \frac{\partial}{\partial x^2}(\Gamma^j_{i1} X^i)\delta s)\delta r \end{aligned}$$

And the vector arrives back at the point a according to:

$$\begin{aligned} X^{*j}(a) &= X^j(d) + \int_0^1 \Gamma^j_{i2} X^i \delta s dt \\ &\approx X^j(d) + (\Gamma^j_{i2} X^i(middleOfC_4))\delta r \\ &\approx X^j(d) + (\Gamma^j_{i2} X^i(a) + 0.5\frac{\partial}{\partial x^1}(\Gamma^j_{i2} X^i)\delta s)\delta s \end{aligned}$$

Where $X^{*j}(a)$ is the new vector at the point a. We can then arrive at the total change in vector:

$$X^{*j}(a) - X^j(a) = \delta X^j \approx (\frac{\partial}{\partial x^2}(\Gamma^j_{i1} X^i) - \frac{\partial}{\partial x^1}(\Gamma^j_{i2} X^i))\delta r \delta s$$

Now, analyze partial derivatives using the product rule:

$$\delta X^j = (X^i \frac{\partial}{\partial x^2}\Gamma^j_{i1} + \Gamma^j_{i1}\frac{\partial}{\partial x^2} X^i - X^i \frac{\partial}{\partial x^1}\Gamma^j_{i2} - \Gamma^j_{i2}\frac{\partial}{\partial x^1} X^i)\delta r \delta s \quad \text{... (III)}$$

Next, given a chain rule formula:

$$\frac{DX^j}{dt} = X^j_{1h}\frac{dx^h}{dt}$$

Since RHS must equal zero, this implies that $X^j_{1h} = 0$ for all p and k. Since the $\frac{dx^h}{dt}$ are non-zero, this means that:

$$\frac{\partial X^j}{\partial x^h} + \Gamma^j_{ih} = 0$$

so that

$$\frac{\partial X^j}{\partial x^h} = -\Gamma^j_{ih} X^i$$

We now sub these into III to obtain:

$$\delta X^j \approx (X^i \frac{\partial}{\partial x^2}\Gamma^j_{i1} - \Gamma^j_{i1}\Gamma^i_{p2}X^p - X^i \frac{\partial}{\partial x^1}\Gamma^j_{i2} + \Gamma^j_{i2}\Gamma^i_{p1}X^p)\delta r \delta s$$

Where everything in the brackets is evaluated at a. Now change dummy indices in 1st and 3rd terms to obtain:

$$\delta X^j = (\ \Gamma^j_{p1} - \Gamma^j_{i1}\Gamma^i_{p2} -\ \Gamma^j_{p2} + \Gamma^j_{i2}\Gamma^i_{p1})X^p \delta r \delta s$$

This formula has the form:

$$\delta X^j \approx R^j_{p12} X^p \delta r \delta s$$

Where the quantity R^j_{p12} is known as the *curvature tensor*:

$$R^a_{bcd} = (\Gamma^i_{bc}\Gamma^a_{id} - \Gamma^i_{bd}\Gamma^a_{ic} + \frac{\partial \Gamma^a_{bc}}{\partial x^d} - \frac{\partial \Gamma^a_{bd}}{\partial x^c})$$

With $R^a_{bcd} = -R^a_{bdc}$, i.e. it is asymmetric with respect to the last two covariant indices. The condition that parallel transport be independent of path is that the curvature tensor vanishes. A mamifold with zero curvature is called *flat*.

9.2 Ricci And Einstein Tensors

The **Ricci tensor** is written as

$$R_{ab} = R^i abi = g^{ij} R_{ajbi}$$

Where we can raise the indices of any tensor, thus giving:

$$R^{ab} = g^{ai} g^{bj} R_{ij}$$

The **Einstein tensor** is written as:

$$G^{ab} = R^{ab} - \frac{1}{2} g^{ab} R$$

Where R is the **Ricci scalar** and is given by:

$$R = g^{ab} R_{ab} = g^{ab} g^{cd} R_{abcd}$$

These quantities will be used in upcoming sections.

10 The Stress and Relativistic Stress-Energy Tensors

10.1 The Classical Stress Tensor

The classical stress tensor measures internal forces that parts of a medium exhert on other parts of itself (even in equilibrium forces are exherted). For example a surface element separating components in a body, these components will be exherting a force on one another through the surface element ΔS, which in equilibrium cancel. More formally, we define a surface element vector with magnitude ΔS and direction $\hat{n}$ perpendicular to the surface element $\Delta \vec{S} = \Delta S \hat{n}$. Associated with this surface element is a vector representing the force exherted by the fluid behind the surface element (i.e. the fluid on the opposite side to $\hat{n}$).
The force per unit area (i.e. pressure) is given by:

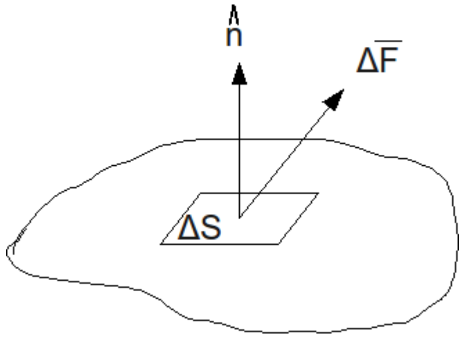

Figure 28: The components used in constructing the stress tensor

$$\vec{T}(\hat{n}) = \lim_{\Delta S \to 0} \frac{\Delta \vec{F}}{\Delta S}$$

T is a function of $\hat{n}$, however if we do not consider a unit vector, we can equally define $\vec{T}$ of some arbitrary vector $\vec{v}$ via $\vec{T}(\frac{\vec{v}}{|\vec{v}|})$ and multiply the RHS by $|\vec{v}|$. Therefore, for general $\vec{v}$ perpendicular to ΔS:

$$\vec{T}(\hat{n}) = \lim_{\Delta S \to 0} \frac{\Delta \vec{F}}{\Delta S} |\vec{v}|$$

$\vec{T}$ operates on vector fields to give new vector field and we call it the *stress tensor*. For example, the a-component of stress on the b-interface is T^{ab}.

10.2 The Relativistic Stress-Energy Tensor

We now look to generalize the stress tensor to 4-dimensional space. We choose to work the 4D manifold with the (1,1,1,-1) signature, i.e. Minkowski space. Note that in the rest of this section time is redifined such that c = 1 (i.e. it is the time for light to travel "one spatial unit"). So, we begin with our definition of the stress-tensor:

$$\vec{T}(\hat{n}) = \frac{\Delta \vec{F}}{\Delta S}$$

for a surface element. Firstly, we get rid of unit vectors since they do not appear in Minkowski space.

$$\vec{T}(\hat{n}\Delta S) = \Delta \vec{F}$$

the above is the total force across the surface element. Now multiply both sides by a time coordinate (x^4) increment:

$$\vec{T}(\hat{n}\Delta S)\Delta x^4 = \Delta \vec{F} \Delta x^4 = \Delta \vec{p}$$

Where p is the 3-momentum. This works well in 3 dimensions. In other words:

$$\vec{T}(\hat{n}\Delta V) = \Delta \vec{p} \text{ or } \vec{T}(\Delta V) = \Delta \vec{p}$$

Where V is the volume in Euclidean 4-space. However, the generalization to 4 dimensions is tougher...
We can replace the 3-momentum by the 4-momentum $\vec{P}$ easily enough. We then fix the RHS by noting that $n\Delta\vec{S}\Delta x^4$ is 3D quantity (e.g. $\Delta\vec{S} \to x^1, x^2, \&t \to x^4$) we then re-write these as:

$$(\triangle V)_i = \epsilon_{ijkl}\Delta x^j \Delta x^k \Delta x^l$$

which gives

$$\vec{T}(\Delta\vec{V}) = \Delta\vec{P}$$

Where ϵ_{ijkl} is the Levi-Civita tensor, which comprises the determinant of the set of vectors:

$$\epsilon_{ijkl} = det(\vec{D}_i\vec{D}_j\vec{D}_k\vec{D}_l)$$

where $\vec{D}$ is the matrix of $\frac{\partial x^k}{\partial \bar{x}^l}$. However, in order for this equation to be accurate we require that $\Delta\vec{V}$ be very small (in terms of coordinates) and we therefore re-write in differential form:

$$\vec{T}(h^3\Delta\vec{V}) = \lim_{h\to 0}\vec{P}(h)$$

and

$$\vec{T}(\Delta\vec{V}) = \lim_{h\to 0}\frac{\vec{P}(h)}{h^3}$$

This converts a covariant vector $\Delta\vec{V}$ to a contravariant vector field $\vec{P}$, it is the relativistic stress-energy tensor. The stress-energy tensor in a comoving frame of some particle in a perfect fluid (that is no viscosity or heat conduction) is:

$$T = \begin{bmatrix} \text{p} & 0 & 0 & 0 \\ 0 & \text{p} & 0 & 0 \\ 0 & 0 & \text{p} & 0 \\ 0 & 0 & 0 & \rho \end{bmatrix}$$

where ρ is the total energy density (or the energy per unit volume) along the time axis. We can then use the fact that the particle's 4-velocity in its own frame is $u = (0001)$ or:

$$u^a u^b = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the metric tensor for Minkowski space:

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \text{-1} \end{bmatrix}$$

to get:

$$T^{ab} = (\rho + p)u^a u^b + pg^{ab}$$

This is the stress-energy tensor for a perfect fluid (no viscosity or heat conduction) as measured in a comoving frame.
Finally, since energy and momentum are conservec we get:

$$\vec{\nabla}\cdot\vec{T} = 0, (\vec{\nabla}\cdot\vec{T})^j = T^{jk}_{k} = 0$$

This is Einstein's conservation law.

11 The Einstein Field Equations

I am well aware of Newtonian gravity, with gravitational potential ϕ:

$$\vec{F} = \vec{\bigtriangledown}\phi \qquad \text{(I)}$$

and Poisson's equation:

$$\bigtriangledown^2\phi = 4\pi\rho G \qquad \text{(II)}$$

where ρ is density and G is the graviational constant.
We look to find the relativistic version of Poisson's equation.

- **First step:** Generalize mass density to energy density, i.e. use stress-energy tensor instead T instead of ρ.
- **Second step:** What about ϕ? ϕ affects the trajectory of particles, however we also saw in PHYM432 that the metric is what determines the trajectory of particles (geodesics), thus we generalize ϕ as g. Now the idea of a gravitational "force" $\vec{\bigtriangledown}\phi$ is replaced by a geometric construct.
- **Third Step:** Modify the operator $\vec{\bigtriangledown}$ to some, as yet unknown 2nd order differential operator Δ. We can therefore generate equation (I) to:

  $$\Delta g^{**} = kT^{**}$$

  where k is some constant. In a comoving frame Δg is a linear combination of g^{ab}_{ij}, g^{ab}_{i} and g^{ab} and must be symmetric (since T is). We have already seen such objects in the form of Ricci tensors $R^{ab}, g^{ab}R$ as well as the metric tensor g^{ab}. We therefore select a candidate solution:

  $$R^{ab} + \mu g^{ab}R + \Lambda g^a b = kT^{ab} \qquad \text{(III)}$$

  We now apply the conservation law $T^{ab}_{|a} = 0$ giving:

  $$R^{ab} + \mu g^{ab}R)_{|b} = 0 \qquad \text{(a)}$$

  since $g^{ab}_{|b} = 0$.
  Then, since Einstein tensor in section 9.2 was:

  $$(R^{ab} - 1/2g^{ab}R) = G^{ab}$$

  and

  $$G^{ab}_{|b} = 0 = (R^{ab} - 1/2g^{ab}R)_{|b} \qquad \text{(b)}$$

  (a) - (b) then gives:

  $$\begin{aligned}&(R^{ab} + \mu g^{ab}R - R^{ab} + 1/2g^{ab}R)_{|b} = 0\\&(\mu + 1/2)g^{ab}R_{|b} = 0\\&= (\mu + 1/2)R_{|b} = 0\end{aligned}$$

  which implies that generally $\mu = -1/2$. Therefore equation (II) becomes:

$$G^{ab} + \Lambda g^{ab} = kT^{ab}$$

Finally given that we require that these equations reduce to Newton's law, we necessarily require that $k = 8\pi$

$$G^{ab} + \Lambda g^{ab} = 8\pi T^{ab}$$

These are Einstein's field equations. The constant Λ is the cosmological constant. Originally it was set $\Lambda = 0$ but its actual value is not certain at present. Have now reached the stage where further application of this equation is covered in PHYM432.

References

[1] S. Waner, *Introduction to Differential Geometry and General Relativity*, 2002

[2] B. F. Schutz, *A First Course in General Relativity*, 1986, Cambridge University Press

[3] I. R. Kenyon, *General Relativity*, 2004, Oxford University Press

[4] D. W. Henderson, *Differential Geometry: A Geometric Introduction*, 1998, Prentice Hall

[5] M. Nakahara, *Geometry, Topology and Physics*, 1990, Institute of Physics