MA224 Geometry

Based on the Lectures of Prof Simms

Typeset by Josh Tobin

Contents

Introduction

This is a set of notes for MA224 Geometry, as taught by Professor David Simms in 2008/2009. It mostly consists of the notes given in class, however in parts I've added comments and examples. In other parts I've removed comments or examples, but usually only because they're difficult to typeset (and only when they aren't strictly necessary; I'd like to think that everything required for the exams is included, at least in the sections that these notes cover). Most of the examples are taken from problem sets given during the course (these are labelled as such). This set of notes was written mainly so that I didn't have to use my messy handwritten notes when revising. So at times I might assume things, or use abbreviations or notation that I use myself. There are also going to be a lot of mistakes. It also should be noted that I basically learned everything I currently know about LaTeX while typing these notes. So, in parts, the style is inconsistent (and sometimes downright ugly!). Finally, the second half of the course is mostly missing. In short, your mileage may vary.

Chapter 1

Prerequisites

This chapter covers the basics that will be required for all of the later chapters. This only includes topics that were covered in class however. So although a good knowledge of Linear Algebra is definitely needed here, it is not covered. Everything needed was covered in 114 though.

We will start with the Einstein Convention. This is nothing more than a different type of notation, one that is standard to physicists (and not so standard to mathematicians). This is one of the bigger road-blocks to getting a feel for 224. Without spending time to understand what the notation means, all you can really do is move symbols around a page. You won't be able to see what's really going on. So it is worth spending some time getting familiar with it, and how it relates to the notation you are familiar with from 114.

1.1 The Einstein Convention

At its heart, the Einstein Convention is nothing more than a different way of writing summations. For example, the following is a common summation in Linear Algebra:

$$\sum_{i=1}^{n} a_i u_i$$

For instance, this could be the sum of components $a_1, a_2, \cdots, a_n$ of a vector with respect to the basis $u_1, u_2, \cdots, u_n$. Note that we use *sub*-scripts here - that is, the indices appear at the bottom right of the a. We could instead use *super*-scripts - where the indices appear at the *top* right. For example, a^i instead of a_i. The reason for this is that the former is easily confused with a to the power of i. However, assume for the moment that we are using superscripts for the components. So the previous equation becomes:

$$\sum_{i=1}^{n} a^i u_i$$

Now note how the upper i has a matching lower i. This is called a **repeated index**. The Einstein convention is the following: wherever we see a repeated index, a summation is implied. That is:

$$a^i u_i = a^0 u_0 + a^1 u_1 + \cdots + a^n u_n = \sum_{i=1}^{n} a^i u_i$$

We do this for two reasons. Firstly, we will at times want to talk about summations within summations within summations. Writing three nested Σs is a pain, and so the Einstein Convention allows us to be a lot more concise. Secondly, in some ways this convention is actually more natural - it allows us to see certain things at a glance, which Σ would not (this sentence is neccasserily vague; it will make more sense later on). Note that the Einstein Convention does not allow us to specify what range we are summing between (whereas the Σ notation does). We depend on context for this - in the above example, we want to sum over every basis vector, so we sum from 1 to n.

Let's say we want to express the product of two $n \times n$ matrices, A and B, using the Einstein Convention. Usually we write:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Where a_{ij}, b_{ij}, c_{ij} are the components of A, B and AB respectively. To use the Einstein Convention, we need to have upper indices as well as lower indices. So instead of a_{ij} we use a^i_j. Note that here the is denote the rows and the js denote the columns. The choice here is not completely arbitrary, it fits in with a more general convention that will be elaborated on later. Now, the summation becomes:

$$c^i_j = \sum_{k=1}^{n} a^i_k b^k_j = a^i_k b^k_j$$

Example: *(224 Excercises 1, Q.9)* If α^i_j stands for the matrix A, what does $\alpha^i_k \alpha^k_l \alpha^l_j$ stand for?

Answer: Well, notice that the first two terms match the expansion for A^2. That suggests the following approach: let $\beta^i_l = \alpha^i_k \alpha^k_l$, where β^i_l are the components of the matrix given by the first two terms, which we know is A^2. Then the whole expression is $\beta^i_l \alpha^l_j$ which is the matrix product of A^2 and A, which is A^3.

Where an index is not repeated, for example α_i, we often take this to refer to $\alpha_0, \alpha_1, \cdots, \alpha_n$. For example, we might talk of the basis u_i. Of course, u_i is really just one vector. But clearly we are referring to the basis $u_0, u_1, \cdots, u_n$. There is an ambiguity here - when we say u_i we could be referring to the single vector u_i or the entire basis $u_0, u_1, \cdots, u_n$. We again depend on context to determine which of the two interpretations we want. The same applies for more than one non-repeated index. So α^i_j could be taken to mean $\alpha^0_0, \alpha^0_1, \cdots, \alpha^0_n, \alpha^1_0, \cdots, \alpha^n_n$. This justifies talking about α^i_j as the components of a matrix.

Example: *(224 Excercises 1, Q.10)* If α^i_j stands for the matrix A, what does $\alpha^i_k\alpha^k_l\alpha^l_i$ stand for?

Answer: This looks very similar to the last example. In fact, it is the same expression, just the first and last indices are now equal. We know that $\alpha^i_k\alpha^k_l\alpha^l_j$ are the components of A^3. Let $B = A^3$. Then the components of B are $\beta^i_j = \alpha^i_k\alpha^k_l\alpha^l_j$. So $\alpha^i_k\alpha^k_l\alpha^l_i = \beta^i_i$.

What is β^i_i? Expanding, we get
$\beta^i_i = \sum_{i=1}^{n} \beta^i_i = \mathrm{Tr}(B) = \mathrm{Tr}(A^3)$

An important symbol is the Kronecker Delta, defined as follows:

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

Note that the matrix whose components are δ^i_j is the identity matrix. So therefore, if we have a matric A and its inverse B, who have components α^i_j and β^i_j respectively, then:

$$\alpha^i_k\beta^k_j = \delta^i_j = \beta^i_k\alpha^k_j$$

Above, the delta function has one upper index and one lower index. This doesn't have to be the case: δ_{ij} and δ^{ij} are all both equally acceptable. This can be useful where we want to sum, but we don't have a matching upper/lower index pair. The following example shows this.

Example: *(224 Excercises 1, Q.8)* What does $\delta_{ij}\alpha^i\beta^j$ stand for?

Answer: Any term where $i \neq j$ will disappear, by the definition of the Kronecker Delta. So we are just left with a sum over the values where $i = j$. That is:
$\delta_{ij}\alpha^i\beta^j = \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}\alpha^i\beta^j = \sum_{i=1}^{n} \alpha^i\beta^i$

One final thing: letting one index equal another is called *contraction*. So, for example, if we have a matrix with components α^i_j, then contracting we get α^i_i, which we know is the trace. So contracting the components of a matrix gives us the trace. If we have another matrix with components β^i_j, then we can take $\alpha^i_k\beta^l_j$ and contract the k with the l, giving $\alpha^i_k\beta^k_j$ - the matrix product AB.

1.2 Some Linear Algebra

Throughout this section we will take M to be an n dimensional vector space, over a field K.

Remember that if we have a basis u_i and a new basis w_i, then each old basis vector u_j is a linear combination of w_i:

$$u_j = p^i_j w_i$$

We call P the transition matrix (where P is the matrix with components p^i_j). Let Q (with components q^i_j) be the inverse of P. Then:

$$\begin{array}{rcl} PQ & = I = & QP \\ p^i_k q^k_j & = \delta^i_j = & q^i_k p^k_j \\ u_j & = & p^i_j w_i \\ w_j & = & q^i_j u_i \end{array}$$

If $x \in M$ has components/coordinates α^i with respect to u_i: $x = \alpha^i u_i$, then $x = \alpha^j u_j = \alpha^j p^i_j w_i$. So if we look at α^i as a column vector X, then the new components are given by the column vector PX. Therefore, when we're changing the basis of a vector, the components of the vector undergo multiplication by P.

So we can change the components of a vector from one basis to another. What about a linear oeprator? Well, let's say that T is a linear operator with components α^i_j with respect to u_i. Then:

$$\begin{array}{rcl} Tw_j = Tq^l_j u_l & = & q^l_j T u_l \\ & = & q^l_j \alpha^k_l u_k \\ & = & q^l_j \alpha^k_l p^i_k w_i \\ & = & p^i_k \alpha^k_l q^l_j w_i \\ & = & PAQ w_i \end{array}$$

Therefore the matrix of T with respect to the new basis is PAQ $(= PAP^{-1})$. Thus on change of basis, the new matrix is gotten from the old α^i_j by contracting the upper index (called the *contravariant* index) with p^i_j and contracting the lower index (called the *covariant* index) with q^i_j.

A scalar valued linear map $f : M \to K$ is called a linear form on M. We write $\langle f, x \rangle$ to denote $f(x)$. Call $\beta_j = \langle f, u_j \rangle$ the j^{th} component of f with respect to the basis u_i. If $x \in M$ has components α^i with respect to u_i then:

$$\langle f, x \rangle = \left\langle f, \alpha^i u_i \right\rangle = \alpha^i \langle f, u_i \rangle = \alpha^i \beta_j$$

This is the matrix product of the row vector β_i and the column vector α^i.

If w_i is a new basis.

$$w_j = q^i_j w_i$$

Then f, in this new basis, has components:

$$\langle f, w_j \rangle = \left\langle f, q^i_j u_i \right\rangle = q^i_j \langle f, u_i \rangle$$

So new components of f are gotten from old components by contraction with q^i_j.

Recall that a bilinear form is a map $B : M \times M \to K$, that is linear in both variables. So $B(\alpha x, y) = \alpha B(x, y) = B(x, \alpha y)$ and $B(x_1 + x_2, y_1 + y_2) = B(x_1, y_1) + B(x_1, y_2) + B(x_2, y_1) + B(x_2, y_2)$. Now, say B is a bilinear form and:

$$B(u_i, u_j) = g_{ij}$$

We say that B has components g_{ij} with respect to the basis u_i. If $x \in M$ has components α^i and $y \in M$ has components β^i, then:

$$B(x, y) = B(\alpha^i u_i, \beta^j u_j) = \alpha^i \beta^j B(u_i, u_j) = \alpha^i \beta^j g_{ij}$$

So we contract the components of x with the first index of g_{ij} and we contract the components of y with the second index. It would be nice to look at this in terms of matrix multiplication. We have a slight problem in that the indices of g are both lower. If we instead look on g_{ij} as the components of a matrix, then we get g^i_j. In this case, we need to change α^i to α_i - so we look at it as a row vector instead of a column vector. This gives:

$$\alpha_i g^i_j \beta^j = X^t G Y$$

But, I can hear you say, surely it wouldn't matter if we took g^j_i instead of g^i_j. Well, in this case we get:

$$\beta_j g^j_i \alpha^i = Y^t G^t X$$

This is the transpose of the previous equation. But as both are in fact scalars, and we know that the transpose of a scalar is a scalar, they are actually the same!

Now let's look at changing the basis of a bilinear form. Say w_i is a new basis, then B has components

$$B(w_i, w_j) = B(q^k_i u_k, q^l_j u_j) = q^k_i q^l_j B(u_k, u_l) = q^k_i q^l_j g_{kl} = Q^t G Q$$

Note that, as above, we've taken g_{ij} to be the matrix g^i_j, in order to get a matrix interpretation for the change of components. For kicks, let's do as we did above, and instead see what happens if we take g^j_i. Well, in this case we get:

$$q^k_i q^j_l g^l_k = Q G^t Q^t$$

But this is not (in general) equal to the expression above, $Q^t G Q$! Is this a contradiction? While it might appear so at first, what we've done in the second case is find the components g^j_i, whereas in the first case we found the components of g^i_j. So the should be the transpose of the first (and it is). So all is well.

What we've done in this section is go through several common operators in linear algebra, express each in terms of the Einstein Convention, and then show how we can bring it from one basis into another. It would be nice to have a table of the required conversion in each case. Oh well.

1.3 Dual Spaces

The set of all linear forms on a vector space M (with dimension n over a field K, as always), is called the **dual space** on M. We denote it as M^*. We can then define addition and scalar multiplication for the dual space. If $f, g \in M^*$, define:

$$\langle f + g, x \rangle = \langle f, x \rangle + \langle g, x \rangle$$
$$\langle \alpha f, x \rangle = \alpha \langle f, x \rangle$$

So, in other words, $\langle ., . \rangle$ is bilinear. We can also see that the above definitions make M^* a vector space.

Now, define $u^i : M \to K$ by:

$$\left\langle u^i, \alpha^j u_j \right\rangle = \alpha^i$$

we say that u^i is the i^{th} coordinate function on M with respect to the basis u_i. Note that $\left\langle u^i, u_j \right\rangle = \delta^i_j$. Also notice that this is a linear form (check this).

So we have that $x = \left\langle u^i, x \right\rangle u_i$ for $x \in M$. We now want to show that $f = \langle f, u_i \rangle u^i$ for $f \in M^*$. This would tell us that u^i is a basis of M^*.

Theorem:
Let u_i be a basis for an n-dimensional vector space M. Let M^* be the dual space. Then the coordinate functions u^i form a basis for M^* and $f = \langle f, u_i \rangle u^i$, for each $f \in M^*$. Thus M^* has the same dimension n, and f has components $\langle f, u_i \rangle$ with respect to the basis u^i.

Proof:
(i) Generate (span):
First we show that it generates M^*. Let $f \in M^*$, and let $\langle f, u_j \rangle = \beta_j$. Then

$$\left\langle B_i u^i, u_j \right\rangle = \beta_i \left\langle u^i, u_j \right\rangle = \beta_i \delta^i_j = \beta_j = \langle f, u_j \rangle$$

Therefore the linear forms $\beta_i u^i$ and f take the same values of each basis vector, so they are equal. So $f = \langle f, u_i \rangle u^i$.

(ii) Linear Independence:
$\alpha_i u^i = 0$
$\Rightarrow \left\langle \alpha_i u^i, u_j \right\rangle = 0$
$\Rightarrow \alpha_i \left\langle u^i, u_j \right\rangle = 0$
$\Rightarrow \alpha_i \delta^i_j = 0$
$\Rightarrow \alpha_j = 0$ ■

So we know that M^* is the dual of M. What is the dual of M^*? As it happens, it's M. In order to show this, we prove that $M^{**} = M$.

Theorem:
M^{**} and M are equal.
Proof:
Consider the map that takes $M \to M^{**}$, defined by $x \to \langle ., x \rangle$. Here, $\langle ., x \rangle$ denotes the linear form on M^* whose value on $f \in M^*$ is $\langle f, x \rangle$.

This map is linear and injective (check this; it is injective as it maps only 0 to 0). We know that dim$M$ = dimM^*, and similarly dimM^* = dimM^{**}. So the dimension of M is the same as that of M^{**}, and therefore the map is bijective. So we have an isomorphism between M and M^{**}. Note that we're trying to prove that these two spaces are equal, which is stronger than isomorphism. What we are looking for is in fact a **natural** isomorphism.

We have an isomorphism of M with M^{**} which allows us to identify x with $\langle ., x \rangle$. Thus each $f \in M^*$ is a function $x \to \langle f, x \rangle$. And each $x \in M$ is a function $f \to \langle f, x \rangle$. ■

So u_i is a basis of M and u^i is the basis of dual M^*, with $\left\langle u^i, u_j \right\rangle = \delta^i_j$. We say that u_i is a basis of M dual to u^i. We also have:

$$x \in M,\ x = \langle u^i, x \rangle u_i$$
$$f \in M^*,\ f = \langle f, u_i \rangle u^i$$

So we have a symmetry between M and M^*, called duality.

1.4 Scalar Products

Let M be a vector space over a field K, of finite dimension n. A scalar product is a bilinear form $(.|.)$, which maps $M \times M \to K$, $(x, y) \mapsto (x|y)$, and is:

1. symmetric, $(x|y) = (y|x)$
2. non-degenerate, $(x|y) = 0$, $\forall y$ iff $x = 0$

If u_i is a basis for M, then $(u_i|u_j) = g_{ij}$ are the components with respect to u_i. By symmetry, we have $g_{ij} = g_{ji}$, so the matrix $G = (g_{ij})$ is symmetric. So $G^t = G$.

If $x = \alpha^i u_i$, $y = \beta^i u_i$, then $(x|y) = \left(\alpha^i u_i | \beta^j u_j\right) = \alpha^i \beta^j g_{ij}$. Note that this is the same thing we got for bilinear forms. Repeating what we did there, we take $X = \alpha^i$, $Y = \beta^j$ and get $(x|y) = X^t G Y$.

The map $M \to M^*$ defined by $x \to (x|.)$ is linear and injective (check this; it is injective as non-degeneracy implies that only 0 is mapped to 0). Therefore the map is bijective. If x has components α^i with respect to u_i, we denote by α_i the components of $(x|.)$ with respect to u^i. This map is called 'lowering the index', and its inverse is called 'raising the index'.

α^i are called the contravariant components of x. $\alpha_i = (x|u_i)$ are called the covariant components of x. If x has contravariant components $\alpha^i = \left\langle u^i, x \right\rangle$ then x has covariant components α_i, and:

$$\alpha_i = (x|u_i) = (\alpha^j u_j | u_i) = \alpha^j (u_j|u_i) = \alpha^j g_{ji} = g_{ij} \alpha^j$$

Therefore $\alpha^i = g^{ij} \alpha_j$, where g^{ij} is the inverse of g_{ij}. So contract with g_{ij} to lower an index, and contract with g^{ij} to raise an index.

As an example, take normal Euclidean geometry, with the usual inner product (the dot product). Note that the components of the dot product are δ^i_j. This is because standard dot product is orthonormal. So therefore to lower an index we contract with δ^i_j:

$$\alpha_i = \delta^i_j \alpha^j = \alpha^i$$

So in normal Euclidean geometry, the covariant and contravariant components are the same with respect to the usual inner product.

Chapter 2

Tensors

This chapter introduces the idea of a tensor, which builds on the idea of a linear form. Here, instead of linear forms, we'll be looking at *multi*linear forms - called tensors (of which linear and bilinear forms are special cases). Imagine it as linear algebra on steroids. For this entire section, we have an n dimensional vector space M over a field K, with dual space M^*.

2.1 Tensors

A tensor is a scalar valued function, so it maps something to an element of K. What it maps from is a little more complicated, and a lot more arbitrary (in the sense that there is a lot of choice, not in the sense that tensors were defined like this for shits and giggles): $T : M_1 \times M_2 \times \cdots \times M_k \to K$, where each M_i is either M or M^*. If it is of this form, and if it is linear in each variable (seperately, like a bilinear form) then it is a tensor. We say it is of degree k.

If (say), we have $T : M \times M^* \times M \to K$ is a tensor and u_i a basis for M. Then the array of n^3 scalars $\alpha^j_{ik} = T(u_i, u^j, u_k)$ are called the components of T with respect to the basis u_i. Note that, in general, a tensor of degree k will have n^k components.

Now we're going to repeat what we did for linear forms - try to see how components change as we move from one basis to another. For simplicity, we only do this for the specific tensor defined in the previous paragraph, but it's pretty obvious from this what to do in the general case. So, we know that α^j_{ik} are the components with respect to the basis u_i. Say we have a new basis w_i with transition matrix P (and with $Q = P^{-1}$ as usual).

So $w^i = p^i_j u^j$ and $w_j = q^i_j u_i$. So, the new components of T are $T(w_i, w^j, w_k) = T(q^r_i u_r, p^j_s u^s, q^t_k u_k) = q^r_i p^j_s q^t_k \alpha^s_{rt}$. Therefore, to get new components, we contract each old lower/covariant index with Q and each upper/contravariant index with P.

We say two tensors are of the same type when their domain is the same. So if $T_1 : M \times M^* \to K$, $T_2 : M \times M^* \to K$ and $T_3 : M^* \times M^* \to K$, then T_1 and T_2 are of the same type, but T_3 is not. The set of all tensors of the same type form a vector space of degree n^k over K, where $(S + T)(x_1, \cdots, x_k) = S(x_1, \cdots, x_k) + T(x_1, \cdots, x_k)$ and $(\alpha T)(x_1, \cdots, x_k) = T(\alpha x_1, x_2, \cdots, x_k) = T(x_1, \alpha x_2, \cdots, x_k) = \cdots = T(x_1, x_2, \cdots, \alpha x_k)$. For example, the space of tensors of type $M \times M^* \times M \to K$ is mapped isomorphically onto the vector space K^{n^3} by $T \to T(u_i, u^j, u_k)$ (that is, a tensor is mapped to its components with respect to some chosen basis).

Examples:

1. $f \in M^*$, $f : M \to K$ is a tensor with components $\alpha_i = \langle f, u_i \rangle$.
2. $x \in M$, $\langle ., x \rangle : M^* \to K$, $f \to \langle f, x \rangle$ is a tensor with components $\alpha^i = \langle u^i, x \rangle$.
3. If $T : M \to M$ is a linear map, we identify it with $T : M^* \times M \to K$, $(f, x) \to \langle f, Tx \rangle$ with components $\langle u^i, Tu_j \rangle = \alpha^i_j$, which is the matrix of T with respect to u_i.
4. If $B : M \times M \to K$ is a bilinear map, B is a tensor with components $\alpha_{ij} = B(u_i, u_j)$.

Let $M_1 \times M_2 \times \cdots \times M^r \times \cdots \times M^s \times \cdots \times M_k \to K$ be a tensor over M with r^{th} index upper ($M^r = M^*$) and s^{th} index lower ($M^s = M$). Then we define a new tensor of degree $k-2$, denoted by S, by the following rule:

$$S : M_1 \times M_2 \times \cdots \times M_{r-1} \times M_{r+1} \times \cdots \times M_{s-1} \times M_{s+1} \times \cdots \times M_k \to K$$
$$S(x_1, \cdots, x_{k-2}) = T(x_1, \cdots, u^i, \cdots, u_i, \cdots, x_{k-2})$$

We call S the contraction of T with respect to the r^{th} and s^{th} indices.

Theorem:
Contraction is well-defined, independent of choice of basis.
Proof:
If w_i is another basis, then:

$$\begin{aligned}
&T(x_1, \cdots, w^i, \cdots, w_i, \cdots, x_{k-2}) \\
&= T(x_1, \cdots, p^i_k u^k, \cdots, q^l_i u_l, \cdots, x_{k-2}) \\
&= p^i_k q^l_i T(x_1, \cdots, u^k, \cdots, u_l, \cdots, x_{k-2}) \\
&= \delta^l_k T(x_1, \cdots, u^k, \cdots, u_l, \cdots, x_{k-2}) \\
&= T(x_1, \cdots, u^k, \cdots, u_k, \cdots, x_{k-2}) \qquad \blacksquare
\end{aligned}$$

If, say, T has components α^{jk}_{il}, then the different possible contractions are:

(i) α^{kj}_{ik}

(ii) α^{jk}_{ik}

(iii) α^{ik}_{il}

(iv) α^{ji}_{ik}

Contracting (ii) and (iii) give α^{ik}_{ik}; contracting (i) and (iv) give α^{ji}_{ij}. These are both scalars.

If $S : M_1 \times \cdots \times M_k \to K$ and $T : M_{k+1} \times \cdots \times M_l \to K$ are two tensors over K. We define the tensor product $S \otimes T : M_1 \times \cdots \times M_l \to K$ given by:

$$S \otimes T(x_1, \cdots, x_l) = S(x_1, \cdots, x_k)T(x_{k+1}, \cdots, x_l)$$

For example, say we have $S : M \times M^* \times M \to K$ and $T : M^* \times M \to K$. Then

$$S \otimes T : M \times M^* \times M \times M^* \times M \to K$$

where

$$S \otimes T(x, f, y, g, z) = S(x, f, y)T(g, z)$$

If S has components $\alpha^j_{ik} = S(u_i, u^j, u_k)$, and T has components $\beta^i_j = T(u^i, u_j)$, then $S \otimes T$ has components $\alpha^j_{ik}\beta^l_s$

Tensors satisfy algebraic laws such as:

- $R \otimes (S + T) = R \otimes S + R \otimes T$
- $(R + S) \otimes T = R \otimes T + S \otimes T$
- $(\alpha S) \otimes T = \alpha(S \otimes T) = S \otimes (\alpha T)$
- $(R \otimes S) \otimes T = R \otimes (S \otimes T)$

The first three of these show that tensor product is bilinear. The last shows that it is associative. In general, it is not commutative. For example, matrix multiplication is in general not commutative.

Note that if $f \in M^*$, $x \in M$, $g \in M^*$ then $f \otimes x \otimes g : M \times M^* \times M \to K$ is given by $\langle f, y\rangle \langle h, x\rangle \langle g, z\rangle$. Thus $f \otimes x \otimes g \in M^* \otimes M \otimes M^*$ (where $M^* \otimes M \otimes M^*$ is the space of all tensors of type $M \times M^* \times M \to K$).

Theorem:
Let u_i be a basis for M, then $u^i \otimes u_j \otimes u^k$ is a basis for $M^* \otimes M \otimes M^*$. Also, if $T \in M^* \otimes M \otimes M^*$ has components $\alpha^j_{ik} = T(u_i, u^j, u_k)$ then $T = \alpha^j_{ik}$. That is, α^j_{ik} are the components of T with respect to $u^i \otimes u_j \otimes u^k$.
Proof:
(i) Let T have components α^j_{ik}, then $T = \alpha^j_{ik} u^i \otimes u_j \otimes u^k$ because components of RHS are
$\alpha^j_{ik} u^i \otimes u_j \otimes u^k(u_r, u^s, u_t)$
$= \alpha^j_{ik} \langle u^i, u_r\rangle \langle u^s, u_j\rangle \langle u^k, u_t\rangle$
$= \alpha^j_{ik} \delta^i_r \delta^s_j \delta^k_t$
$= \alpha^s_{rt} = T(u_r, u^s, u_t)$ as required.

(ii) We know that $u^i \otimes u_j \otimes u^k$ generate $M^* \otimes M \otimes M^*$ by (i). Also:
$\alpha^j_{ik} u^i \otimes u_j \otimes u^k = 0$
$\Rightarrow$ the components of $u^i \otimes u_j \otimes u^k$ are 0
$\Rightarrow \alpha^s_{rt} = 0$ by (i)
$\Rightarrow$ they are linearly independent. ■

2.2 Skew-Symmetric Tensors

A bijective map $\sigma : \{1, 2, \cdots, r\} \to \{1, 2, \cdots, r\}$ is called a permutation of degree r. We often write $\sigma(i)$ as σ_i. The group of all permutations of degree r is denoted by $\mathcal{S}_r$.

Let $\mathcal{T}^r M$ denote the space $M^* \otimes \cdots \otimes M^*$ (r times). That is, the space of all tensors of type $M \times M \cdots \times M \to K$ (with M written r times). If $T \in \mathcal{T}^r M$, then with respect to the basis u_i, T has components $T(u_{i_1}, \cdots, u_{i_r}) = \alpha_{i_1 \cdots i_r}$. For each permutation $\sigma \in \mathcal{S}_r$ and each $T \in \mathcal{T}^r M$, we define $\sigma.T \in \mathcal{T}^r M$ by:

$$(\sigma.T)(x_1, \cdots, x_r) = T(x_{\sigma(1)}, \cdots, x_{\sigma(r)})$$

We call $\sigma.T$ 'T permuted by sigma'.

Theorem:
The group $\mathcal{S}_r$ acts on $\mathcal{T}^r M$ by linear transformations.

(i) $\sigma.(\alpha T + \beta S) = \alpha\sigma.T + \beta\sigma.S$

(ii) $\sigma(\tau.T) = (\sigma\tau).T$

(iii) $1.T = T$

Proof:

(i) We prove this in two parts. Firstly:

$$\begin{aligned}
&\sigma.(\alpha T)(x_1, \cdots, x_k)\\
&= \sigma.T(\alpha x_1, x_2, \cdots, x_k)\\
&= T(x_{\sigma(1)}, \cdots, \alpha x_1, \cdots, x_{\sigma(k)})\\
&= \alpha T(x_{\sigma(1)}, \cdots, x_{\sigma(k)})\\
&= \alpha(\sigma.T)
\end{aligned}$$

So $\sigma.(\alpha T) = \alpha(\sigma.T)$. Also:

$$\begin{aligned}
&\sigma.(T + S)(x_1, \cdots, x_k)\\
&= (T + S)(x_{\sigma(1)}, \cdots, x_{\sigma(k)})\\
&= T(x_{\sigma(1)}, \cdots, x_{\sigma(k)}) + S(x_{\sigma(1)}, \cdots, x_{\sigma(k)})\\
&= (\sigma.T)(x_1, \cdots, x_k) + (\sigma.S)(x_1, \cdots, x_k)
\end{aligned}$$

So $\sigma.(T + S) = \sigma.T + \sigma.S$. These two facts immediately imply the result we want.

(ii) $$\begin{aligned}
&\sigma.(\tau.T)(x_1, \cdots, x_k)\\
&= (\tau.T)(x_{\sigma(1)}, \cdots, x_{\sigma(k)})\\
&= T(x_{\tau\sigma(1)}, \cdots, T_{\tau\sigma(k)})\\
&= (\sigma\tau)T(x_1, \cdots, x_k)
\end{aligned}$$

(iii) $$\begin{aligned}
&(1.T)(x_1, \cdots, x_k)\\
&= T(x_{1(1)}, \cdots, x_{1(k)})\\
&= T(x_1, \cdots, x_k) \qquad \blacksquare
\end{aligned}$$

An important idea is the parity of a permutation. This is just the parity of the of the number of transpositions in the permutation. We also have the following symbol, called the sign of a permutation:

$$\varepsilon^\sigma = \begin{cases} 1 & \text{if } \sigma \text{ is an even permutation} \\ -1 & \text{if } \sigma \text{ is an odd permutation} \end{cases}$$

We can show (i) $\varepsilon^{\sigma\tau} = \varepsilon^\sigma \varepsilon^\tau$, (ii) $\varepsilon^1 = 1$ and (iii) $\varepsilon^{\sigma^{-1}} = \varepsilon^\sigma$.

We say that $T \in \mathcal{T}^r M$ is skew-symmetric (or, more commonly, alternating or anti-symmetric) if $\sigma.T = \varepsilon^\sigma T$, $\forall \sigma \in \mathcal{S}_r$. That is $T(x_{\sigma(1)}, \cdots, x_{\sigma(k)}) = \varepsilon^\sigma T(x_1, \cdots, x_k)$. Which is to say, the components satisfy:

$$\alpha_{i_{\sigma(1)} \cdots i_{\sigma(k)}} = \varepsilon^\sigma \alpha_{i_1 \cdots i_k}$$

For example, if $T \in \mathcal{T}^3 M$ is skew-symmetric, then $T(x,y,z) = -T(y,x,z) = T(y,z,x) = -T(z,y,x) = T(z,x,y) = -T(x,z,y)$. This implies $T(x,x,z) = -T(x,x,z) \Rightarrow T(x,x,z) = 0$. Note that here we assume that $k+k=0$ implies $k=0$ in our field K. We also have the following relation involving the components of T:

$$\alpha_{ijk} = -\alpha_{jik} = \alpha_{jki} = -\alpha_{kji} = \alpha_{kij} = -\alpha_{ikj}$$

It follows that if $T \in \mathcal{T}^k M$ is skew-symmetric with components $\alpha_{i_1 \cdots i_k}$ with respect to the basis u_i, then:

- $\alpha_{i_1 \cdots i_k} = 0$ if $i_1, \cdots, i_k$ are not all distinct.
- If we know $\alpha_{i_1 \cdots i_k}$ for all increasing sequences then we know $\alpha_{i_1 \cdots i_k}$ for all sequences.
- If S is skew-symmetric with components $\beta_{i_1 \cdots i_k}$ and $\alpha_{i_1 \cdots i_k} = \beta_{i_1 \cdots i_k}$ for all increasing sequences then $T = S$.

Recall that the dimension of M is n. Let $T \in \mathcal{T}^n M$ be skew-symmetric. Then with respect to any basis u_i, T is determined by the single component $T(u_1, \cdots, u_n)$. Suppose $T \neq 0$. Then:

(i) $x_1, \cdots, x_n$ are linearly independent $\Rightarrow x_1, \cdots, x_n$ are a basis for $M \Rightarrow T(x_1, \cdots, x_n) \neq 0$.

(ii) If they are linearly independent, then (say) $x_1 = \alpha^2 x_2 + \cdots + \alpha^n x_n$.
$\Rightarrow T(x_1, \cdots, x_n) = T(\alpha^2 x_2 + \cdots + \alpha^n x_n, x_2, \cdots, x_n)$
$= \alpha^2 T(x_2, x_2, \cdots, x_n) + \cdots + \alpha^n T(x_n, x_2, \cdots, x_n) = 0$

Thus $x_1, \cdots, x_n$ are linearly independent iff $T(x_1, \cdots, x_n) \neq 0$.

Theorem:
Let $T \in \mathcal{T}^n M$, where M has dimension n and T is skew-symmetric. Let $S : M \to M$ be a linear operator with matrix A with respect to some basis u_i, then:

$$T(Sx_1, \cdots, Sx_n) = (\det A) T(x_1, \cdots, x_n)$$

(which confirms that $\det A$ is independent of choice of basis, so we can define $\det S = \det A$).
Proof:
Both sides are skew-symmetric tensors as functions of $x_1, \cdots, x_n$. So it is completely determined by the values it takes on any basis, say u_i. Then

$T(Su_1, \cdots, Su_n)$
$= T(\alpha_1^{i_1} u_{i_1}, \cdots, \alpha_1^{i_n} u_{i_n})$
$= \alpha_1^{i_1} \alpha_2^{i_2} \cdots \alpha_n^{i_n} T(u_{i_1}, \cdots, u_{i_n})$
$= \varepsilon_{i_1, i_2, \cdots, i_n} \alpha_1^{i_1} \alpha_2^{i_2} \cdots \alpha_n^{i_n} T(u_1, \cdots, u_n)$
$= (\det A) T(u_1, \cdots, u_n)$ ■

For K^n we have the usual basis $e_1 = (1,0,0,\cdots,0), e_2 = (0,1,0,\cdots), \cdots, e_n = (0,0,0,\cdots,1)$. We define the determinant tensor D of degree n on K^n to be the unique skew-symmetric tensor such that $D(e_1, \cdots, e_n) = 1$. So D has components $D(e_{i_1}, e_{i_2}, \cdots, e_{i_n}) = \varepsilon_{i_1, i_2, \cdots, i_n}$. If A is an $n \times n$ matrix over K, then $A : K^n \to K^n$. Ae_j is then the j^{th} column of A. Denote this by a_j. Then

$$D(a_1, \cdots, a_n) = D(Ae_1, \cdots, Ae_n) = (\det A) D(e_1, \cdots, e_n) = \det A$$

We have the following properties:

1. $\det AB = D(AB_1, \cdots, AB_n) = (\det A)D(B_1, \cdots, B_n) = (\det A)(\det B)$

2. A is invertible $\Leftrightarrow$ $A : K^n \to K^n$ is bijective $\Leftrightarrow$ $Ae_1, \cdots, Ae_n$ are linearly independent $\Leftrightarrow$ $\det A \neq 0$.

3. If $x \in K^n$ has components α^i with respect to e_i, then:

$$D(e_1, \cdots, x, \cdots, e_n) = D(e_1, \cdots, \alpha^s e_s, \cdots, e_n) = \alpha^i D(e_1, \cdots, e_n) = \alpha^i$$

(where x is inserted in the i^{th} position).

2.3 Wedge Product

If $S \in \mathcal{T}^s M$ and $T \in \mathcal{T}^t M$, we define the wedge product (or exterior product) of S and T by:

$$S \wedge T = \frac{1}{s!t!} \sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^\sigma \sigma.(S \otimes T)$$

For example, if $S, T \in M^*$, with components α_i and β_i respectively. Then $S \wedge T = S \otimes T - T \otimes S$, and has components:

$$S \wedge T[u_i, u_j] = S(u_i)T(u_j) - T(u_i)S(u_j) = \alpha_i\beta_j - \beta_i\alpha_j =$$

$$\begin{vmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{vmatrix}.$$

We have the following properties:

1. $S \wedge T$ is skew-symmetric

2. $(R + S) \wedge T = R \wedge T + S \wedge T$

3. $R \wedge (S + T) = R \wedge S + R \wedge T$

4. $(\alpha S) \wedge T = \alpha(S \wedge T) = S \wedge (\alpha T)$

5. $(R \wedge S) \wedge T = R \wedge (S \wedge T)$

6. $R_1 \wedge R_2 \wedge \cdots \wedge R_k = \dfrac{(r_1 + \cdots + r_k)!}{r_1! \cdots r_k!} \mathcal{A}(R_1 \otimes \cdots \otimes R_k)$

7. $S \wedge T = (-1)^{st} T \wedge S$ (super-symmetry)

If s is odd, we call S odd, and if s is even then we call S even. Super-symmetry also implies the following properties:

1. $S \wedge T = T \wedge S$ if either S or T has even degree.

2. $S \wedge T = -T \wedge S$ if both S or T has odd degree.

3. $S \wedge S = 0$ if S has odd degree.

4. $T_1 \wedge \cdots \wedge S \wedge \cdots \wedge S \wedge \cdots \wedge T_k = 0$ if S has odd degree.

5. If $x_1, \cdots, x_r \in M$ and $i_1, \cdots, i_r$ are selected from $\{1, 2, \cdots, r\}$ then:

$$x_{i_1} \wedge \cdots \wedge x_{i_r} = \varepsilon_{i_1 \cdots i_r} x_1 \wedge \cdots \wedge x_r$$

Example: *(224 Excercises 4, Q.1)* Let A be an $n \times n$ matrix with columns $c_1, \cdots, c_n$. Show that

$$c_1 \wedge \cdots \wedge c_n = \det A e_1 \wedge \cdots \wedge e_n$$

Let $x = (x^1, \cdots, x^n)$. Show that the equation $Ax = b$ is equivalent to:

$$x^1 c_1 + \cdots + x^n c_n = b$$

Solve the equation for the scalar x^n by taking the wedge product with $c_1 \wedge \cdots \wedge c_{n-1}$ (Cramer's Rule).

Answer:

(i) We have

$$\begin{aligned} c_1 \wedge \cdots \wedge c_n &= (c_1^{i_1} e_{i_1} \wedge \cdots \wedge c_n^{i_n} e_{i_n}) \\ &= (c_1^{i_1} \cdots c_n^{i_n})(e_{i_1} \wedge \cdots \wedge e_{i_n}) \\ &= \varepsilon_{i_1 \cdots i_n}(c_1^{i_1} \cdots c_n^{i_n}) e_1 \wedge \cdots \wedge e_n \\ &= \det A e_1 \wedge \cdots \wedge e_n \end{aligned}$$

(ii)

$$(c_1 c_2 \cdots c_n) \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = c_1 x^1 + c_2 x^2 \cdots c_n x^n$$

So therefore $Ax = b \Leftrightarrow c_1 x^1 + c_2 x^2 \cdots c_n x^n = b$.

(iii) We get:

$$\begin{aligned} x^1 c_1 \wedge \cdots \wedge c_{n-1} \wedge c_1 + \cdots + x^n c_1 \wedge \cdots \wedge c_n &= c_1 \wedge \cdots \wedge c_{n-1} \wedge b \\ 0 + 0 + \cdots + \det A x^n e_1 \wedge \cdots \wedge e_n &= \det A_n e_1 \wedge \cdots \wedge e_n \\ x^n &= \frac{\det A_n}{\det A} \end{aligned}$$

Where A_n is the matrix with columns $c_1, \cdots, c_{n-1}, b$.

Example: *(224 Excercises 4, Q.2)* Let $x_1, \cdots, x_r \in M$. Show that $x_1, \cdots, x_r$ are linearly independent iff

$$x_1 \wedge \cdots \wedge x_r \neq 0$$

Answer: Note that the solution given here in previous versions of these notes is incorrect. To prove it properly, note that if they are not linearly independent we can easily reduce it to the wedge product of two equal terms, which is 0. To prove the converse, prove the contrapositive (if they are not linearly independent, the wedge product is not 0). This can be shown by extending the vectors to a basis, and then using the lemma that appears at the end of this section. This example really should be moved, but it stays here for now.

Denote by $M^{(r)}$ the space of all skew-symmetric tensors of degree r, and type $M \times M \cdots \times M \to K$, over a finite dimensional vector space M (where $\dim M = n$). u_i is a basis of M, and u^i is a basis of M^*. Will show:

$u^i \wedge u^j : M \times M \to K$, $i < j$, basis for $M^{(2)}$
$u^i \wedge u^j \wedge u^k : M \times M \times M \to K$, $i < j < k$, basis for $M^{(3)}$
$\vdots$

For example, $\dim M = 4$:

$M^{(1)} = M^*$, basis is u^1, u^2, u^3, u^4
$M^{(2)}$, basis is $u^1 \wedge u^2, u^1 \wedge u^3, u^1 \wedge u^4, u^2 \wedge u^3, u^2 \wedge u^4, u^3 \wedge u^4$
$M^{(3)}$, basis is $u^1 \wedge u^2 \wedge u^3, u^1 \wedge u^2 \wedge u^4, u^1 \wedge u^3 \wedge u^4, u^2 \wedge u^3 \wedge u^4$
$M^{(4)}$, basis is $u^1 \wedge u^2 \wedge u^3 \wedge u^4$

The linear operator $\mathcal{A} : \mathcal{T}^r M \to \mathcal{T}^r M$ defined by:

$$\mathcal{A}T = \frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \varepsilon^\sigma \sigma.T$$

This is called the skew-symmetriser (and it converts each tensor to a skew-symmetric one). Note that in many books this is called the alternator. As such, $\mathcal{A}T$ is sometimes written $\mathrm{Alt}(T)$.

Theorem:
$\mathcal{A}T$ is skew-symmetric.
Proof:
Let τ be a permutation of degree r.

$\tau(\mathcal{A}T)$
$= \tau.[\frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \varepsilon^\sigma \sigma.T]$
$= \varepsilon^\tau [\frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \varepsilon^\sigma \varepsilon^\tau (\tau\sigma).T]$
$= \varepsilon^\tau [\frac{1}{r!} \sum_{\sigma' \in \mathcal{S}_r} \varepsilon^{\sigma'} \sigma'.T]$ (where $\sigma' = \tau\sigma$)
$= \varepsilon^\tau.(\mathcal{A}T)$ ■

For example, let $T \in \mathcal{T}^3 M$ with components α_{ijk}. Then

$$\mathcal{A}T(x,y,z) = \frac{1}{6}[T(x,y,z) - T(x,z,y) + T(y,z,x) - T(y,x,z) + T(z,x,y) - T(z,y,x)]$$

$\mathcal{A}T$ has components

$$\beta_{ijk} = \tfrac{1}{6}[\alpha_{ijk} - \alpha_{ikj} + \alpha_{jki} - \alpha_{jik} + \alpha_{kij} - \alpha_{kji}]$$

Theorem:
Let $S \in \mathcal{T}^s M$, $T \in \mathcal{T}^t M$, then

(i) $\mathcal{A}[(\mathcal{A}S) \otimes T] = \mathcal{A}[S \otimes T] = \mathcal{A}[S \otimes (\mathcal{A}T)]$

(ii) $\mathcal{A}(S \otimes T) = (-1)^{st}\mathcal{A}(T \otimes S)$

Proof:

(i) Note that if $\tau \in \mathcal{S}_s$ then

$[(\tau.S) \otimes T](x_1, \cdots, x_s, x_{s+1}, \cdots, x_{s+t})$
$= (\tau.S)(x_1, \cdots, x_s)T(x_{s+1}, \cdots, x_{s+t})$
$= S(x_{\tau(1)}, \cdots, \_\tau(s))T(x_{s+1}, \cdots, x_{s+t})$
$= S(x_{\tau'(1)}, \cdots, x_{\tau'(s)})T(x_{\tau'(s+1)}, \cdots, x_{\tau'(s+t)})$
Where

$$\tau' = \begin{pmatrix} 1 & 2 & \cdots & s & s+1 & \cdots & s+t \\ \tau(1) & \tau(2) & \cdots & \tau(s) & s+1 & \cdots & s+t \end{pmatrix} \in \mathcal{S}_{s+t}$$

$= [\tau'.(S \otimes T)](x_1, \cdots, x_{s+t})$
$\therefore (\tau.S) \otimes T = \tau'.(S \otimes T)$

Now, $\mathcal{A}[(\mathcal{A}S) \otimes T] = \dfrac{1}{(s+t)!} \displaystyle\sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^\sigma \sigma.[(\mathcal{A}S) \otimes T]$

$$= \frac{1}{(s+t)!} \sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^\sigma [(\frac{1}{s!} \sum_{\tau \in \mathcal{S}_s} \varepsilon^\tau.\tau S) \otimes T]$$
$$= \frac{1}{s!} \sum_{\tau \in \mathcal{S}_s} \frac{1}{(s+t)!} \sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^{\sigma\tau'} (\sigma\tau').(S \otimes T)$$
$$= \frac{1}{s!} \sum_{\tau \in \mathcal{S}_s} \mathcal{A}[S \otimes T]$$
$$= (\frac{1}{s!})(s!)\mathcal{A}[S \otimes T]$$
$$= \mathcal{A}[S \otimes T]$$

(ii) Let $\tau \in \mathcal{S}_{s+t}$ such that:

$$\tau = \begin{pmatrix} 1 & \cdots & s & s+1 & \cdots & s+t \\ t+1 & \cdots & t+s & 1 & \cdots & t \end{pmatrix} \in \mathcal{S}_{s+t}$$

Note that $\varepsilon^t = (-1)^{st}$. Then $[\tau(S \otimes T)](x_1, \cdots, x_{s+t})$

$= [S \otimes T](x_{t+1}, \cdots, x_{t+s}, x_1, \cdots, x_t)$
$= S(x_{t+1}, \cdots, x_{t+s})T(x_1, \cdots, x_t)$
$= T(x_1, \cdots, x_t)S(x_{t+1}, \cdots, x_{t+s})$
$= [T \otimes S](x_1, \cdots, x_{t+s})$
$\therefore \tau.(S \otimes T) = T \otimes S$
$$\therefore \mathcal{A}(S \otimes T) = \frac{1}{(s+t)!} \sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^{\sigma\tau} (\sigma\tau)(S \otimes T)$$
$$= \varepsilon^\tau \frac{1}{(s+t)!} \sum_{\sigma \in \mathcal{S}_{s+t}} \varepsilon^\sigma \sigma.(T \otimes S)$$
$(-1)^{st}\mathcal{A}(T \otimes S)$ ■

Recall $S \wedge T = \dfrac{1}{s!t!}\displaystyle\sum_{\sigma \in \tau} \varepsilon^{\tau}\sigma.(S \otimes T)$. Therefore $S \wedge T = \dfrac{(s+t)!}{s!t!}\mathcal{A}(S \otimes T)$. Also recall that $M^{(r)}$ is the space of all skew-symmetric tensors of degree r. What is the dimension of $M^{(r)}$? We'll build up to this now, first by proving the following lemma. Thanks Amy for the proof :)

Lemma:
If $1 \leq i_1 < \cdots < i_r \leq n$ and $1 \leq j_1 < \cdots < j_r \leq n$, then:

$$u^{i_1} \wedge u^{i_2} \wedge \cdots \wedge u^{i_r}[u_{j_1}, u_{j_2}, \cdots, u_{j_r}] = \begin{cases} 1 & \text{if } i_1 = j_1, \cdots, i_r = j_r \\ 0 & \text{otherwise} \end{cases}$$

Proof:

$$\begin{aligned} u^{i_1} \wedge \cdots \wedge u^{i_r}[u_{j_1}, \cdots, u_{j_r}] &= n!\mathcal{A}(u^{i_1} \otimes \cdots \otimes u^{i_r})[u_{j_1}, \cdots, u_{j_r}] \\ &= \sum_{\sigma \in \mathcal{S}_r} \varepsilon^{\sigma} u^{i_1} \wedge \cdots \wedge u^{i_r}[u_{j_1}, \cdots, u_{j_r}] \\ &= \sum_{\sigma \in \mathcal{S}_r} \varepsilon^{\sigma} \delta^{i_1}_{j_{\sigma(1)}} \cdots \delta^{i_r}_{j_{\sigma(r)}} \end{aligned}$$

This clearly is 1 if $i_1 = j_1, \cdots, i_r = j_r$, and 0 otherwise. ■

So, for example, take M to have dimension 4, then we have:

1. $u^1 \wedge u^2[u_1, u_2] = 1$
2. $u^1 \wedge u^2[u_2, u_1] = -1$ (this follows from skew-symmetry)
3. $u^1 \wedge u^2[u_1, u_1] = 0$
4. $u^1 \wedge u^2[u_2, u_3] = 0$

Anyway, this lemma will be helpful in proving the following theorem, which answers our question about the dimension of $M^{(r)}$.

Theorem:
Let $\dim M = n$, u_i be a basis. Then:

(i) $M^{(r)} = 0$ if $r > n$.

(ii) The tensor $\sum_{i_1<\cdots<i_r} \alpha_{i_1\cdots i_r} u^{i_1} \wedge \cdots \wedge u^{i_r}$ has components $\alpha_{i_1\cdots i_r}$ for $i_1 < \cdots < i_r$.

(iii) $\{u^{i_1} \wedge \cdots \wedge u^{i_r}\}$, $i_1 < \cdots < i_r$ is a basis for $M^{(r)}$. So $\dim M^{(r)} = \dfrac{n!}{r!(n-r)!}$.

Proof:

(i) Let $T \in M^{(r)}$ have components $\alpha_{i_1\cdots i_r}$ with respect to u_i. If $r > n$, then $i_1, \cdots, i_r$ are not all distinct. Therefore $\alpha_{i_1\cdots i_r} = 0$, thus $T = 0$.

(ii) The tensor

$$\sum_{i_1<\cdots<i_r} \alpha_{i_1\cdots i_r} u^{i_1} \wedge \cdots \wedge u^{i_r} \tag{2.1}$$

has components $j_1, j_2, \cdots, j_r$ for $j_1 < \cdots < j_r$. Now

$$\sum_{i_1<\cdots<i_r} \alpha_{i_1\cdots i_r} u^{i_1} \wedge \cdots \wedge u^{i_r}[u_{j_1}, \cdots, u_{j_r}] = \alpha j_1 \cdots j_r \text{ by lemma.}$$

(iii) Let $T \in M^{(r)}$ have components $\alpha_{i_1\cdots i_r}$ for $i_1 < \cdots < i_r$, then T is as in equation (2.1) above. $\{u^{i_1} \wedge \cdots \wedge u^{i_r}\}, i_1 < \cdots < i_r$ are linearly independent, since if $T = 0$ then the components of T are all 0, that is $\alpha_{i_1\cdots i_r}$. ■

So, we have:

(i) If $T \in M^{(r)} T = \sum_{i_1<\cdots<i_r} \alpha_{i_1\cdots i_r} u^{i_1} \wedge \cdots \wedge u^{i_r}$.

(ii) If $T \in M_{(r)} T = \sum_{i_1<\cdots<i_r} \alpha^{i_1\cdots i_r} u_{i_1} \wedge \cdots \wedge u_{i_r}$.

So we have $M^{(1)} = M^*$, $M_{(1)} = M$, $M^{(0)} = K = M_{(0)}$.

Chapter 3

Pull-Back, Push-Forward and Orientation

Categories are an abstraction of other mathematical structures (such as groups and vector spaces). We use categeries in relation to the "Pull-Back" and "Push-Forward" functors. During lectures, pictures were often used (what were referred to as 'commutative diagrams'). I haven't copied them here yet. Instead, have a look at the Mathsoc wiki.

3.1 Categories

A category is a collection of objects, and a collection of morphisms. Each morphism maps from one unique object to another. For each object, we have a morphism from that object to itself, called the identity (and denoted 1_M, if M is the object in question). Finally, we have an associative binary operation called composition over the morphisms, such that is f maps from A to B, and g maps from B to C, then gf maps from A to C.
Examples:

1. Category set where objects are sets and morphisms are maps of sets.
2. Category K-Vect where objects are vector spaces over a field K, and morphisms are linear maps.
3. Category group whose objects are groups and morphisms are homomorphisms.

A functor F from a category $\mathcal{C}$ to a category $\mathcal{D}$ assigns:

1. To each object $\mathcal{M}$ in $\mathcal{C}$, an object $F(M)$ in $\mathcal{D}$.
2. To each morphism $T : M \to N$ in $\mathcal{C}$, assigns a morphism

$$T_* : F(M) \to F(N) \text{ where } F \text{ is covariant.}$$
$$T^* : F(M) \leftarrow F(N) \text{ where } F \text{ is contravariant.}$$

   Such that

   (a) If $T : L \to M$, $S : M \to N$ then $(ST)_* = S_*T_*$ if F is covariant, and $(ST)^* = T^*S^*$ where F is contravariant.

(b) $(1_M)_* = 1_{F(M)}$ (if F is covariant), or $(1_M)^* = 1_{F(M)}$ (if F is contravariant).

Theorem:
For each linear operator $T : M \to N$ of K-vector spaces M, N define the pull-back (or transpose) $T^* : M^* \leftarrow N^*$ of T by:

$$\langle T^* f, x\rangle = \langle f, Tx\rangle$$

Then this is a contravariant functor from category K-Vect to category K-Vect.
Proof:

(i) T^* is a linear operator.

(ii) Let $T : L \to M$, $S : M \to N$ and $(ST) : L \to N$. Say we have $f \in N^*$. Then $(ST)^* f \in L$. We have:

$$\begin{aligned} \langle (ST)^* f, x\rangle &= \langle f, (ST)x, \rangle \\ &= \langle S^* f, Tx\rangle \\ &= \langle T^* S^* f, x\rangle \, \forall x \in L \end{aligned}$$

Therefore $(ST)^* f = T^* S^* f$ for all $f \in N^*$. Hence $(ST)^* = T^* S^*$ as required.

(iii) $\langle (1_M)^* f, x, \rangle = \langle f, 1_M x\rangle = \langle f, x\rangle$, $\forall x \in M$
Therefore $(1_M)^* f = f$ for all $f \in M^*$ as required. ■

3.2 Properties of Pull-Back and Push-Forward

Note that in the finite dimensional case, the category of K-vector spaces and linear maps, we have a duality between M and M^*: M is the dual of M^* and M^* is the dual of M. So let's say that $T : M \to N$ has pull-back $T^* : M^* \leftarrow N^*$. Then T^* has pull-back $T_* : M \to N$. We have that:

$$\langle f, T_* x\rangle = \langle T^* f, x\rangle = \langle f, Tx\rangle, \forall f \in N^*$$

Therefore $T_* x = Tx$, $\forall x \in M$, and so $T_* = T$. What about the more general case $T : M \otimes \cdots \otimes M \to N \otimes \cdots \otimes N$? The following theorem answers these questions.

Theorem:
To each linear operator $T : M \to N$ on finite dimensional K-vector spaces M, N define the push-forward

$$T_* : M \otimes M \otimes \cdots \otimes M \to N \otimes N \otimes \cdots N \text{ (each term appearing } r \text{ times)}$$

And the pull-back

$$T^* : M^* \otimes M^* \otimes \cdots \otimes M^* \leftarrow N^* \otimes N^* \cdots N^*$$

By:

$$(T_* S)(f^1, \cdots, f^r) = S(T^* f^1, \cdots, T^* f^r) \text{ for } f^1, \cdots, f^r \in N^*$$

And

$$(T^* S)(x_1, \cdots, x_r) = S(T_* x_1, T_* x_2, \cdots, T_* x_r) \text{ for } x_1, \cdots, x_r \in M$$

Then the push-forward is a covariant functor and the pull-back is a contravariant functor from finite dimensional K-Vect to finite dimensional K-Vect.
Proof:
Let $T : L \to M$, $U : M \to N$, $(UT) : L \to N$, be a commutative diagram. Then:

$$\begin{aligned} T^* : L^* \otimes \cdots \otimes L^* &\leftarrow M^* \otimes \cdots \otimes M^* \\ U^* : N^* \otimes \cdots \otimes N^* &\rightarrow M^* \otimes \cdots \otimes M^* \\ (UT)^* : N^* \otimes \cdots \otimes N^* &\rightarrow L^* \otimes \cdots \otimes L^* \end{aligned}$$

Then:

$[(UT)^* S](x_1, \cdots, x_r)$, $x_i \in L$, $S \in N^* \otimes \cdots \otimes N^*$
$= S[UTx_1, \cdots, UTx_r]$
$= (U^* S)[Tx_1, \cdots, Tx_r]$
$= [T^*(U^* S)](x_1, \cdots, x_r)\ \forall x_i \in L, \forall S$

So $(UT)^* = T^* U^*$ as required. Therefore pull-back is a contravariant functor. Similarly for push-forward. ■

Theorem:

(i) $T^*(R \otimes S) = (T^*R) \otimes (T^*S)$, that is pull-back preserves tensor product.

(ii) S is skew-symmetric $\Rightarrow T^*S$ is skew-symmetric.

(iii) $T^*(R \wedge S) = (T^*R) \wedge (T^*S)$

And similarly for push-forward.
Proof:
Let $T : M \to N$, $R \in N^* \otimes \cdots \otimes N^*$ (r times) and $S \in N^* \otimes \cdots \otimes N^*$ (s times). Then:

(i)

$$\begin{aligned} [T^*(R \otimes S)](x_1, \cdots, x_{r+s}) &= (R \otimes S)[Tx_1, \cdots, Tx_{r+s}] \\ &= R(Tx_1, \cdots, Tx_r)S(Tx_{r+1}, \cdots, Tx_{r+s}) \\ &= (T^*R)(x_1, \cdots, x_r)(T^*S)(x_{r+1}, \cdots, x_{r+s}) \\ &= (T^*R \otimes T^*S)[x_1, \cdots, x_{r+s}] \qquad \forall x_i \end{aligned}$$

Therefore $T^*(R \otimes S) = (T^*R) \otimes (T^*S)$, as required.

(ii) Let $\sigma \in \mathcal{S}_s$, then:

$$\begin{aligned} \sigma.(T^*S)(x_1, \cdots, x_s) &= \sigma S(Tx_1, \cdots, Tx_s) \\ &= S(Tx_{\sigma(1)}, \cdots, Tx_{\sigma(s)})\mathrm{sgn}(\sigma) \\ &= \mathrm{sgn}(\sigma)[T^*S](x_{\sigma(1)}, \cdots, x_{\sigma(s)}) \end{aligned}$$

Therefore T^*S is skew-symmetric as required.

(iii)

$$\begin{aligned} T^*[R \wedge S] &= T^*[\frac{1}{r!s!} \sum_{\sigma \in \mathcal{S}_{r+s}} \varepsilon^\sigma \sigma.(R \otimes S)] \\ &= \frac{1}{r!s!} \sum_{\sigma \in \mathcal{S}_{r+s}} \varepsilon^\sigma \sigma.(T^*R \otimes T^*S)] \qquad \text{(by(ii))} \\ &= (T^*R) \wedge (T^*S) \qquad \blacksquare \end{aligned}$$

Let $\mathrm{dim}M = n$. Then:

$$\begin{aligned} T : M &\to M \\ T_* : M_{(n)} &\to M_{(n)} \\ T^* : M^{(n)} &\leftarrow M^{(n)} \end{aligned}$$

$$\mathrm{dim}M_{(r)} = \frac{n!}{r!(n-r)!}$$

So $\mathrm{dim}M_{(n)} = 1 = \mathrm{dim}M^{(n)}$. On the 1-dimensional spacve $M_{(n)}$, push-forward is just multiplication by a scalar. The same is true for pull-back on $M^{(n)}$. If $S \in M^{(n)}$, then:

$$(T^*S)(x_1, \cdots, x_s) = S(Tx_1, \cdots, Tx_s) = \det(T)S(x_1, \cdots, x_s)$$

As S is skew-symmetric. Therefore pull-back of S is just $\det T$ times S. Hence T^* on $M^{(n)}$ is multiplication by the determinant of T.

Also, if u_i is a basis of M and T has matrix α_j^i with respect to u_i, then:

$$\begin{aligned} T_*(u_1 \wedge u_2 \wedge \cdots \wedge u_n) &= (T_*u_1) \wedge (T_*u_2) \wedge \cdots \wedge (T_*u_n) \\ &= (\alpha_1^{i_1} u_{i_1}) \wedge (\alpha_2^{i_2} u_{i_2}) \wedge \cdots \wedge (\alpha_n^{i_n} u_{i_n}) \\ &= (\det T) u_1 \wedge u_2 \wedge \cdots \wedge u_n \end{aligned}$$

Therefore T_* on $M_{(n)}$ is also multiplication by $\det T$.

3.3 Orientation

In the previous section, we saw that for a basis u_i and some vectors $x_1, x_2, \cdots, x_n$ we have:

$$x_1 \wedge x_2 \wedge \cdots \wedge x_n = (\det X) u_1 \wedge u_2 \wedge \cdots \wedge u_n$$

What if we have two bases: $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$? Well, in that case we have:

$$u_1 \wedge u_2 \wedge \cdots \wedge u_n = (\det U) v_1 \wedge v_2 \wedge \cdots \wedge v_n$$

We know that $\det \mathrm{U} \neq 0$ (as it is a basis). We say that $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ have the same orientation if $\det \mathrm{U} > 0$, and the opposite orientation otherwise. It is easily seen that 'same orientation' is an equivalence relation on the set of all bases in M. The two equivalence classes are naturally the positive and negative bases. For $\mathbb{R}^n$, the usual basis is designated positive in what is called the 'usual' orientation of $\mathbb{R}^n$.

Theorem:
Let M be a real oriented vector space of finite dimension n with a (symmetric and non-degenerate) scalar product. We call $u_1, u_2, \cdots, u_n$ a standard basis for M if it is positively oriented and

$$(u_i|u_j) = \begin{pmatrix} \pm1 & 0 & \cdots & 0 \\ 0 & \pm1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm1 \end{pmatrix}$$

(Recall that by Sylvester's Theorem the number of plus/minus signs is uniquely determined by $(.|.)$).

The n-form

$$\mathrm{vol} = u^1 \wedge u^2 \wedge \cdots \wedge u^n$$

is independent of choice of standard basis $u_1, u_2, \cdots, u_n$ and is called the volume form on M. Thus, for any standard basis, we have:

$$\mathrm{vol}(u_1, \cdots, u_n) = 1$$

If $w_1, \cdots, w_n$ is any positively oriented basis then

$$\mathrm{vol} = \sqrt{|\det(w_i|w_j)|}\, w^1 \wedge \cdots \wedge w^n$$

Thus

$$\mathrm{vol}(w_1, \cdots, w_n) = \sqrt{|\det(w_i|w_j)|}$$

Proof:
Fix a standard basis $u_1, \cdots, u_n$. Let $w_1, \cdots, w_n$ be any positively oriented basis. Then we have:

$$u^i = p^i_j w^j$$

Where P is the transition matrix from w_i to u_i.
$\therefore u^1 \wedge \cdots \wedge u^n = (\mathrm{detP}) w^1 \wedge \cdots \wedge w^n$, $\mathrm{detP} > 0$
Now,

$$P^t \begin{pmatrix} \pm1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pm1 \end{pmatrix} P = G$$

Where $G = (g_{ij})$ is a matrix with $g_{ij} = (w_i|w_j)$.

$$\begin{aligned} \therefore \pm(\mathrm{detP})^2 &= \mathrm{detG} \\ \therefore \mathrm{detP} &= \sqrt{|\mathrm{detG}|} \\ \therefore \mathrm{vol} &= u^1 \wedge \cdots \wedge u^n \\ &= \sqrt{|\mathrm{detG}|}\, w^1 \wedge \cdots \wedge w^n \end{aligned}$$

And if $w_1, \cdots, w_n$ is a standard basis, the above works out to be exactly

$$w^1 \wedge \cdots \wedge w^n \qquad \blacksquare$$

Note that the above theorem gives us that given any positively oriented basis w_i, vol has components $\sqrt{|\det(w_i|w_j)|}\varepsilon_{i_1,\cdots,i_n}$.

3.4 Hodge Star Operator

Let M be a scalar, oriented vector space of dimension n with a (symmetric, non-degenerate) scalar product $(.|.)$. Then, for each $0 \leq r \leq n$ define the Hodge Star Operator:

$$* : M^{(r)} \to M^{(n-r)}$$

By

$$(*w)(v_1, \cdots, v_{n-r}) = \frac{1}{r!} w(u_{i_1}, \cdots, u_{i_r}) \mathrm{vol}(u^{i_1}, \cdots, u^{i_r}, u^{i_{r+1}}, \cdots, u^{i_n})(u_{i_{r+1}}|v_1) \cdots (u_{i_n}|v_{n-r})$$

Theorem:
Let $u_1, \cdots, u_n$ be a standard basis for M. That is

$$(u_i|u_i) = s_i = \pm 1$$

Then

$$*u^1 \wedge \cdots \wedge u^r = s_{r+1} \cdots s_n . u^{r+1} \wedge \cdots \wedge u^n$$

Proof:
Let $w = u^1 \wedge \cdots \wedge u^r$.
Then $w(u_1, \cdots, u_r) = 1$ and all other components of w are 0 except for $r!$ permutations of $u_1, \cdots, u_r$.
$\therefore *w(v_1, \cdots, v_{n-r}) = \frac{1}{r!} r! \mathrm{vol}(u^1, \cdots, u^r, u^{i_{r+1}}, \cdots, u^{i_n})(u_{i_{r+1}}|v_1) \cdots (u_{i_n}|v_{n-r})$
$\therefore *w(u_{r+1}, \cdots, u_n) = \mathrm{vol}(u^1, \cdots, u^r, u^{r+1}, \cdots, u^n)(u_{r+1}|u_{r+1}) \cdots (u_n|u_n)$
$\therefore *w(u_{r+1}, \cdots, u_n) = s_{r+1} \cdots s_n$
And all the other components of $*w$ are 0 except the permutations of $u_{r+1}, \cdots, u_n$ as required. ■

For example, if M is a 3 dimensional oriented Euclidean space, and u_1, u_2, u_3 are positively oriented orthonormal, that is $(u_i|u_i) = 1$, $s_i = 1$. Then:
$*u^1 = u^2 \wedge u^3$
$*u^2 = u^3 \wedge u^1$ (as u^2, u^3, u^1 are positive)

$x = \alpha^1 u_1 + \alpha^2 u_2 + \alpha^3 u_3$
$y = \beta^1 u_1 + \beta^2 u_2 + \beta^3 u_3$

$$x \wedge y = \det \begin{pmatrix} \alpha^2 & \alpha^3 \\ \beta^2 & \beta^3 \end{pmatrix} u_2 \wedge u_3 + \det \begin{pmatrix} \alpha^3 & \alpha^1 \\ \beta^3 & \beta^1 \end{pmatrix} u_3 \wedge u_1 + \det \begin{pmatrix} \alpha^1 & \alpha^2 \\ \beta^1 & \beta^2 \end{pmatrix} u_1 \wedge u_2$$

$$*x \wedge y = \det \begin{pmatrix} \alpha^2 & \alpha^3 \\ \beta^2 & \beta^3 \end{pmatrix} u_1 + \det \begin{pmatrix} \alpha^3 & \alpha^1 \\ \beta^3 & \beta^1 \end{pmatrix} u_2 + \det \begin{pmatrix} \alpha^1 & \alpha^2 \\ \beta^1 & \beta^2 \end{pmatrix} u_3 = x \times y$$

Chapter 4

Continuity and Differentiability

It should be noted that Chris Blair has some excellent notes on this section on his webpage. Basically everything given by Professor Simms verbatim, LaTeXed up. Anyway, this chapter firstly covers continuity (and the basic topology needed to express things), in order to describe differentiability, which makes up the final part of this chapter.

4.1 Continuity

The main goal of this chapter is to describe differentiation of linear operators. It helps a bit to look at what we're trying to do, before launching into continuity. We can already find the derivative of a function $f : \mathbb{R} \to \mathbb{R}$. That is:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

We can't do the same here because we can't divide by h (as h is a vector). There are also some other complications. In the $\mathbb{R} \to \mathbb{R}$ case, when we say $h \to 0$ we obviously mean that h gets arbitrarily small. But what does this mean in a vector space? We're going to need to have some notion of the distance between vectors, and of 'small' regions around vectors.

So, we define a norm (or length function) on the vector space M. A norm is a function $||.|| : M \to \mathbb{R}$ such that:

(i) $||x + y|| \leq ||x|| + ||y||$ (the triangle inequality)

(ii) $||\alpha x|| = |\alpha|.||x||$, where $\alpha \in \mathbb{R}$

(iii) $||x|| \geq 0$

(iv) $||x|| = 0$ if and only if $x = 0$

Now that we have a way of measuring distance, we can describe small regions. If V is a subset of a normed space M, then we call V open if for all $a \in V$, there exists $\delta > 0$ such that $||x - a|| < \delta \Rightarrow x \in V$. In fact, no matter what norm we choose, the open sets will be the same (in a finite dimensional real vector space). This is assumed, no proof was given in class.

Another way to look at open sets is the following... Talk about balls here. Then continuity. Include Theorems 35-38 (see Chris Blair's notes) and proofs...

4.2 Differentiability

Now we're ready to define the derivative. There is one big difference between the derivative of a linear operator that maps between vector spaces, and the usual derivative for functions that map from $\mathbb{R} \to \mathbb{R}$. It arises from the fact that unlike in $\mathbb{R}$, we can usually move in several different directions in vector space. So we can't describe the rate of change of a function in a vector space with a single scalar. The rate of change is going to depend on the direction we're moving in. So instead we define the derivative at a point to be another linear operator. So if $f : M \to N$, where M and N are vector spaces, then for each $a \in M$, $f'(a)$ is a linear operator that maps from $M \to N$, and $[f'(a)](h)$ measures the rate of change of f *in the direction of the vector* h. Note that $[f'(a)](h)$ is more often written as $f'(a)h$. Its important to remember that this doesn't mean multiplication of $f'(a)$ and h.

Anyway, on with the definition: let M, N be real or complex vector spaces, with V, W open subsets of M and N respectively. Let $f : V \to W$. We say that f is differentiable at $a \in M$ if there exists a linear operator

$$f'(a) : M \to N$$

called the derivative of f at a, such that

$$f(a + h) = f(a) + f'(a)h + \phi(h)$$

where $f'(a)h$ is a linear approximation to the change in f when a changes by h, and $\phi(h)$ is a remainder term such that

$$\frac{||\phi(h)||}{||h||} \to 0 \text{ as } ||h|| \to 0$$

The following theorem gives us that the derivative is unique.

Theorem:
Let M, N be vector spaces, with V, W open sets in M and N respectively. Let $f : V \to W$ be differentiable at $a \in V$. Then the derivative $f'(a)$ is uniquely determined by the formula

$$\begin{aligned} f'(a)h &= \lim_{t \to 0} \frac{f(a + th) - f(a)}{t} \\ &= \frac{d}{dt} f(a + th)\Big|_{t=0} \end{aligned}$$

Proof:

$f(a + th) = f(a) + f'(a)th + \phi(th)$ (by definition)

$\Rightarrow \left|\left|\dfrac{f(a + th) - f(a)}{t} - f'(a)h\right|\right| = \dfrac{\phi(th)}{|t|} = \dfrac{\phi(th)}{||th||}||h||$

The RHS tends to 0 as $t \to 0$, so the LHS must also tend to 0, which gives us

$$f'(a)h = \lim_{t \to 0} \frac{f(a + th) - f(a)}{t} \qquad \blacksquare$$

Let $f : V \to \mathbb{R}$, where V is an open subset of $\mathbb{R}^n$.

$$(x_1, x_2, \cdots, x_n) \mapsto f(x_1, x_2, \cdots, x_n)$$

f is called a real valued function of n independent real variables. We also define:

$$\begin{aligned}\frac{\partial f}{\partial x^j}(a) &= \frac{\partial f}{\partial x^j}(a_1, a_2, \cdots, a_n)\\ &= \lim_{t\to 0}\frac{f(a_1,\cdots,a_j+t,\cdots,a_n)}{t}\\ &= \lim_{t\to 0}\frac{f(a+te_j)}{t}\\ &= \frac{d}{dt}f(a+te_j)\Big|_{t=0}\end{aligned}$$

This is the directional derivative in the direction of the basis vector e_j, and is called the partial derivative of f at a with respect to the j^{th} coordinate function x^j. We use the partial derivative to describe the derivative of a linear operator.

Theorem:
Let $V \subset \mathbb{R}^n$ be open, and let $f : V \to \mathbb{R}^m$, where $f(x) = (f^1(x), f^2(x), \cdots, f^m(x))$. Then the derivative of f

$$f'(a) : \mathbb{R}^n \to \mathbb{R}^m$$

is the $m \times n$ matrix

$$f'(a) = \Big(\frac{\partial f^i(a)}{\partial x^j}\Big),\ i = 1, \cdots, m,\ j = 1, \cdots, n$$

Proof:
Consider the j^{th} column of $f'(a)$. We have:

$$\begin{aligned}f'(a)e_j &= \frac{d}{dt}f(a+te_j)\Big|_{t=0}\\ &= \frac{\partial f(a)}{\partial x^j}\\ &= (\frac{\partial f^1(a)}{\partial x^j}, \cdots, \frac{\partial f^m(a)}{\partial x^j})\end{aligned}$$

This is true for each column, hence result. ■

Next, we look at the chain rule for functions on finite dimensional real or complex vector spaces. But before we can get into this, we need to define a norm that acts on operators. Let $T : M \to N$, a linear operator between finite dimensional normed vector spaces. Then we define

$$||T|| = \sup_{||u||=1} ||Tu||$$

This is called (slightly prematurely) the operator norm of T. The following theorem shows us that it is indeed a norm.

Theorem:
The operator norm satisfies:

(i) $||S+T|| \leq ||S|| + ||T||$

(ii) $||\alpha T|| = |\alpha|||T||$

(iii) $||Tx|| \leq ||T||||x||$

(iv) $||ST|| \leq ||S||||T||$

(v) $||T|| \geq 0$

(vi) $||T|| = 0 \Leftrightarrow T = 0$

Proof:
These all should be straightforward enough (at least, Professor Simms thought so, enough to leave out most of the proofs). For example:

(iii) $||Tx|| = ||x||\left|\left|T\frac{x}{||x||}\right|\right| \leq ||x||||T||$, as $\frac{x}{||x||}$ is a unit vector.

(iv)

$$\begin{aligned} ||ST|| = \sup_{||u||=1} ||STu|| &\leq \sup_{||u||=1} ||S||.||Tu|| \qquad \text{by(iii)} \\ &\leq \sup_{||u||=1} ||S||.||T||.||u|| \\ &= ||S||.||T|| \end{aligned}$$

Theorem: (Chain Rule)
Let U, V, W be open subsets of finite dimensional vector spaces. Let $g : U \to V$, $f : V \to W$, $f \circ g : U \to W$. If g is differentiable at a and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a and

$$(f \circ g)' = f'(g(a))g'(a)$$

Proof:
We know that f and g are differentiable, so we have that:

$f(a+h) = f(a) + f'(a)h + \psi(h)$
$g(a+h) = g(a) + g'(a)h + \phi(h)$

Where as $||h|| \to 0$ we have $\dfrac{||\psi(h)||}{||h||} \to 0$ and $\dfrac{||\phi(h)||}{||h||} \to 0$. Then:

$$\begin{aligned}
f(g(a+h)) &= f(g(a) + g'(a)h + \phi(h)) \\
&= f(g(a) + (g'(a)h + \phi(h))) \\
&= f(g(a)) + f'(g(a))(g'(a)h + \phi(h))) + \psi(g'(a)h + \phi(h))) \\
&= f(g(a)) + f'(g(a))g'(a)h + f'(g(a))\phi(h) + \psi(g'(a)h + \phi(h)))
\end{aligned}$$

Now, take $y = g'(a) + \phi(h)$. So the above becomes:

$$f(g(a+h)) = f(g(a)) + f'(g(a))g'(a)h + f'(g(a))\phi(h) + \psi(y)$$

So $f'(g(a))g'(a)h$ is the linear term, and $f'(g(a))\phi(h) + \psi(y)$ is the remainder term. We just need to show that the remainder term goes to 0 when we divide by $||h||$ and let $||h|| \to 0$.

$$\begin{aligned}
\frac{\Big|\Big|f'(g(a))\phi(h) + \psi(y)||}{||h||} &= \frac{||f'(g(a))\phi(h) + ||y||.\frac{\psi(y)}{||y||}\Big|\Big|}{||h||} \\
&\leq \Big|\Big|f'(g(a))\Big|\Big|\frac{||\phi(h)||}{||h||} + \frac{||y||}{||h||}.\Big|\Big|\frac{\psi(y)}{||y||}\Big|\Big|
\end{aligned}$$

Now, we have:

$y = g'(a)h + \phi(h)$
$\Rightarrow ||y|| \leq ||g'(a)||.||h|| + ||\phi(h)||$
$\Rightarrow \dfrac{||y||}{||h||} \leq ||g'(a)|| + \dfrac{||\phi(h)||}{||h||}$
$\therefore ||h|| \to 0 \Rightarrow ||y|| \to 0$

So:

$$||h|| \to 0 \Rightarrow \Big|\Big|f'(g(a))\Big|\Big|\frac{||\phi(h)||}{||h||} \to 0$$

And also:

$$||h|| \to 0 \Rightarrow ||y|| \to 0 \Rightarrow \frac{||y||}{||h||}.\Big|\Big|\frac{\psi(y)}{||y||}\Big|\Big| \to 0$$

Therefore $||h|| \to 0 \Rightarrow \dfrac{||f'(g(a))\phi(h) + \psi(y)||}{||h||} \to 0$. Thus

$$(f \circ g)'(a) = f'(g(a))g'(a) \qquad \blacksquare$$

Now we start to relate continuity and differentiability. We start with the following theorems.

Theorem:
Let M, N be finite dimensional real vector spaces, with V an open subset of M. Then f is differentiable at $a \Rightarrow f$ is continuous at a.
Proof:
Let f be differentiable at a. Then:

$$f(a+h) = f(a) + f'(a)h + \phi(h)$$

Where $\frac{||\phi(h)||}{||h||} \to 0$ as $||h|| \to 0$. Note that

$$||f(a+h) - f(a)|| \leq ||f'(a)h||.||h|| + ||\phi(h)||$$

The RHS goes to 0 as $h \to 0$, so the LHS must too. Thus we have continuity.
■

Theorem:
Let V be an open subset of $\mathbb{R}^n$, and $f : V \to \mathbb{R}$. That is, f is a real valued function of n independent variables. Then f has a continuous derivative $\Leftrightarrow$ $\dfrac{\partial f}{\partial x^j}$ exists and is continuous for each j.

Proof:
We only prove this for the $n = 2$ case. The other cases are similar.

- If f has a continuous derivative, then

$$f' = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$$

  Therefore each component exists and is continuous, hence $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ exist and are continuous.

- Conversely, assume that $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ are continuous. Then

$$f(a+h) = f(a) + \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] h + \left(f(a+h) - f(a) - \left[\tfrac{\partial f}{\partial x}, \tfrac{\partial f}{\partial y}\right] h\right)$$

  Let $\phi(h) = f(a+h) - f(a) - \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] h$. Then we just need to show that $\dfrac{||\phi(h)||}{||h||} \to 0$ as $||h|| \to 0$. Now:

$$\begin{aligned}
\phi(h) &= f(a+h) - f(a) - \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] h \\
&= f(a+h) - f(a) - \frac{\partial f}{\partial x} h^1 - \frac{\partial f}{\partial y} h^2 \\
&= f(a+h) - f(a+h^2 e_2) - \frac{\partial f}{\partial x} h^1 + f(a+h^2 e_2) - f(a) - \frac{\partial f}{\partial y} h^2 \\
&= \frac{\partial f}{\partial x}(m_1) h^1 - \frac{\partial f}{\partial x}(a) h^1 + \frac{\partial f}{\partial y}(m_2) h^2 - \frac{\partial f}{\partial y}(a) h^2 (\text{byMVT})
\end{aligned}$$

  Thus

$$\frac{||\phi(h)||}{||h||} \leq \frac{||h^1||}{||h||}.\left|\left|\frac{\partial f}{\partial x}(m_1) - \frac{\partial f}{\partial x}(a)\right|\right| + \frac{||h^2||}{||h||}.\left|\left|\frac{\partial f}{\partial y}(m_2) - \frac{\partial f}{\partial y}(a)\right|\right|$$

  The RHS tends to 0 as $||h|| \to 0$, so f' exists, and is $\left[\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}\right]$. ■

Let M, N be finite dimensional real vector spaces, and $f : V \to N$, where V is an open subset of M. We say that f is differentiable if it is differentiable at each point, and we have the function f', where $x \mapsto f'(x)$. We call f' the derivative of f. If also f' is continuous on V, we say that f is C^1. If f' is also differentiable, we call its derivative f''. If f'' is continuous, we say that f is C^2. In general, if the r^{th} derivative exists and is continuous, we say that f is C^r. If f is C^r for all r, we say that f is C^∞ (or equivalently we say that f is smooth).

Now, let's say that we have V an open subset of $\mathbb{R}^n$, with $f : V \to \mathbb{R}$. Then f is $C^1 \Leftrightarrow \dfrac{\partial f}{\partial x^j}$ exists and is continuous for each j. Similarly, f is $C^2 \Leftrightarrow \dfrac{\partial^2 f}{\partial x^i \partial x^j}$ exist and are continuous. Does it matter what order we take derivatives? That is to say, is $\dfrac{\partial^2 f}{\partial x^i \partial x^j} = \dfrac{\partial^2 f}{\partial x^j \partial x^i}$?

Theorem:
If V is an open set in $\mathbb{R}^n$ and $f : V \to \mathbb{R}$, where f is C^2. Then

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i} \forall i, j$$

Proof:
We just consider the case where $n = 2$. So we want to show that

$$\frac{\partial}{\partial x}.\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}.\frac{\partial f}{\partial x}$$

Consider any point (a, b) in V. As V is open, we can find a rectangle $[a, a+h] \times [b, b+k]$ contained in V. A picture would be helpful here, but haven't gotten round to that yet. Check the mathsoc wiki and the old 211/221 notes. Anyway, so we have a rectangle in V, whose lower left corner is (a, b) and whose upper right corner is $(a + h, b + k)$, where $h \neq 0, k \neq 0$.
Define $g(x) = f(x, b + k) - f(x, b)$. This is just the value of f along the top of the rectangle, minus the value along the bottom. Then:

$$\begin{aligned}
f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) &= g(a+h) - g(a) \\
&= g'(c).h \qquad \text{byMVT} \\
&= h[\frac{\partial f}{\partial x}(c, b+k) - \frac{\partial f}{\partial x}(c, b)] \\
&= hk\frac{\partial f}{\partial y \partial x}(c, d) \qquad \text{byMVT}
\end{aligned}$$

Similarly, if we define $h(y) = f(a + h, y) - f(a, y)$ and repeat the above steps, we get

$$f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) = kh\tfrac{\partial f}{\partial x \partial y}(c', d')$$

Thus

$$\begin{aligned}
hk\frac{\partial f}{\partial y \partial x}(c, d) &= kh\frac{\partial f}{\partial x \partial y}(c', d') \\
\frac{\partial f}{\partial y \partial x}(c, d) &= \frac{\partial f}{\partial x \partial y}(c', d')
\end{aligned}$$

We now let $k \to 0$, $h \to 0$. Thus $c \to a$, $c' \to a$, $d \to b$, $d' \to b$. As the partial derivatives are continuous, the value at the limit is equal to the limit of the value. Thus we get

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b), \forall (a, b) \in V \qquad \blacksquare$$

Let M, V be vector spaces with V, W open subsets respectively. Let $f : M \to V$. We say that f is a C^r-diffeomorphism if f is bijective, and if both f and f^{-1} are C^r. The Inverse Function Theorem tells us that under certain conditions we can find local inverses for f. But before we can prove the Inverse Function Theorem, we need the Mean Value Theorem for vector valued functions. This allows us to control the distance between $f(x)$ and $f(y)$, if we know something about f' in the interval between x and y.

Theorem: (Mean Value Theorem for vector valued functions)
Let f be C^1. Let $x, y \in V$, such that

$$tx + (1-t)y \in V,\ \forall 0 \leq t \leq 1$$

And:

$$||f'(tx + (1-t)y)|| \leq k$$

Then:

$$||f(x) - f(y)|| \leq k||x - y||$$

Proof:

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \Big(\frac{d}{dt} f(tx + (1-t)y)\Big) dt \\ &= \int_0^1 f'\Big(tx + (1-t)y\Big).(x - y) dt \end{aligned}$$

Therefore

$$||f(x) - f(y)|| \leq \int_0^1 k.||x - y|| dt = k||x - y|| \qquad \blacksquare$$

Now we can look at the inverse function theorem. This is probably the hardest proof on the course. I've divided it up into 5 parts, but this is more or less arbitrary (I found this particular division easiest to understand, but in class he used 4 parts, and Chris Blair's notes uses 3).

Theorem: (Inverse Function Theorem)
Let M, N be finite dimensional real or complex vector spaces, with V open in M. Let $f : V \to N$ be a C^r function, and $a \in V$ such that

$$f'(a) : M \to N$$

is invertible, then there exists an open neighbourhood W of a such that

$$f : W \to f(W)$$

is a C^r diffeomorphism onto open $f(W)$ in N.
Proof:

(i) **Initial reduction:**
Let T be the inverse of $f'(a)$. Let

$$F(x) = Tf(x+a) - Tf(a)$$

Then we have the following:

$$\begin{aligned} F(0) &= 0 \\ F'(x) &= Tf'(x+a) \\ F'(0) &= Tf'(a) = 1_M \end{aligned}$$

We will show that F maps an open neighbourhood U of 0 onto an open neighbourhood $F(U)$ by a C^r diffeomorphism. It will follow that f maps $U + a$ onto $f(U + a)$ by a C^r diffeomorphism.

(ii) **Define domain B on which F is a homeomorphism:**
Choose a closed ball B, centre 0, with $r > 0$ such that for all $x \in B$:

$$\begin{aligned} ||1_M - F'(x)|| &\leq \frac{1}{2} \\ \det F'(x) &\neq 0 \end{aligned}$$

(We can do this because F' is continuous). Then

$$\begin{aligned} ||x-y|| - ||F(x) - F(y)|| &\leq ||(x - F(x)) - (y - F(y))|| \\ &= ||(1_M - F)x - (1_M - F)y|| \\ &\leq \frac{1}{2}||x-y|| \qquad \text{(MVT)} \end{aligned}$$

Therefore $||F(x) - F(y)|| \geq \frac{1}{2}||x - y|| \Rightarrow F$ is injective. So

$$F : B \to F(B)$$

is bijective. Also, $F(x) \to F(y) \Rightarrow x \to y$ so F^{-1} is continuous. So

$$F : B \to F(B)$$

is a homeomorphism.

(iii) **Show that $\frac{1}{2}B \subseteq F(B)$:**
Let $a \in \frac{1}{2}B$. Put $g(x) = x - F(x) + a$. Then

$$\therefore ||g'(x)|| = ||1 - F'(x)|| \leq \tfrac{1}{2} \forall x \in B$$

So

$$||g(x) - g(y)|| \leq \tfrac{1}{2}||x - y|| \text{ by the MVT}$$

Therefore g is a contraction mapping. Also (noting that $g(0) = a$):

$$\begin{aligned} ||g(x)|| &= ||g(x) - g(0) + a|| \\ &\leq ||g(x) - g(0)|| + ||a|| \\ &\leq \frac{1}{2}||x|| + ||a|| \\ &\leq \frac{1}{2}r + \frac{1}{2}r = r \end{aligned}$$

So g maps B into itself and is contracting.

Choose $x_0 \in B$. Put $x_{n+1} = g(x_n)$. (x_n) has a limit point b. Now, by continuity

$$g(b) = g(\lim x_n) = \lim g(x_n) = \lim x_{n+1} = b$$

So b is a fixed point. Therefore $g(b) = b - F(b) + a = b$. So $F(b) = a$, thus $a \in F(b)$. So now

1. $F : B \to F(B)$ is a homeomorphism.
2. $\tfrac{1}{2}B \subseteq F(B)$
3. $||F(x) - F(y)|| \leq \tfrac{1}{2}||x - y||$

(iv) **Show F is a C^1 diffeomorphism:**
Let B^0 be the interior of B and let

$$U = B^0 \cap F^{-1}(\tfrac{1}{2}B^0)$$

Then U is open and $F : U \to F(U)$ is a homeomorphism of open U onto open $F(U)$. F is a C^r function, so we just need to show $G : F(U) \to U$ is C^r, where $G = F^{-1}$.

Let $G(x) = y$, $G(x + h) = y + l$. Then

$$F(y + l) = F(y) + Sl + \phi(l)$$

Where $S = F'(y)$ and $\dfrac{||\phi(l)||}{||l||} \to 0$ as $||l|| \to 0$.

Also

$$\begin{aligned} ||F(x) - F(y)|| &\geq \frac{1}{2}||x - y|| \\ ||F(y + l) - F(y)|| &\geq \frac{1}{2}||l|| \\ ||x + h - x|| &\geq \frac{1}{2}||l|| \\ ||h|| &\geq \frac{1}{2}||l|| \end{aligned}$$

So

$F(y + l) = F(y) + Sl + \phi(l)$
$\Rightarrow x + h = x + Sl + \phi(l)$
$\Rightarrow Sl = -\phi(l) + h$
$\Rightarrow l = S^{-1}h - S^{-1}\phi(l)$

So $G(x + h) = y + l = G(x) + S^{-1}h - S^{-1}\phi(l)$. Now:

$$\frac{||S^{-1}\phi(l)||}{||h||} \leq ||S^{-1}||.\frac{||\phi(l)||}{||l||}.\frac{||l||}{||h||}$$

We also have that as $||h|| \to 0$, $||l|| \to 0$ so $\frac{||\phi(l)||}{||l||} \to 0$. And we know that $\frac{||l||}{||h||} \leq 2$. So therefore:

$$||h|| \to 0 \Rightarrow \frac{||S^{-1}\phi(l)||}{||h||} \to 0$$

So G is differentiable with derivative S^{-1}.

(v) **Show that G is C^r:**
We now know that G is differentiable at x and

$$G'(x) = S^{-1} = [F'(y)]^{-1} = [F'(G(x))]^{-1}$$

So if G is C^s for some $0 \leq s < r$ then G' is the composition of the C^s functions G, F' and the inverse function, and so G' is C^s, which means that G is C^{s+1}, for all $0 \leq s < r$. Therefore G is C^r. ■

Chapter 5

Manifolds

This (unfinished) chapter defines what a manifold is, and how we can determine if something is a manifold (by the Implicit Function Theorem). Then we look at how to find derivatives of functions defined on manifolds.

5.1 What is a Manifold?

The title of this section is misleading in that it suggests I'm going to explain exacly what a manifold is, which I'm not. Instead, I'm going to waffle about what a manifold is loosely speaking, then give the formal definition.

A mainfold is a 'surface' which is 'locally' 'like' $\mathbb{R}^n$ (where n is constant for every point on the manifold, which we refer to as the rank of the mainfold). That is, if we take any point in the manifold, there is an open set around this which is homeomorphic to $\mathbb{R}^n$. Remember that, roughly speaking, two spaces are homeomorphic if we can stretch or bend one to get the other. So in $\mathbb{R}^2$, an open square is homeomorphic to an open circle. Also, any open set in $\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$. A few examples:

1. Take the surface of a sphere in $\mathbb{R}^3$. Then take a small open ball around any point on the sphere. The points in this open ball that are on the sphere form a bent disc. If we unbend this, it will just be an open disc from $\mathbb{R}^2$ - which we know is homeomorphic to $\mathbb{R}^2$. So we can see that the surface of a sphere in $\mathbb{R}^3$ is a manifold of rank 2.

2. Take $\mathbb{R}^n$. Then any open set in $\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$. So $\mathbb{R}^n$ is a manifold of rank n.

Anyway, before we can define a manifold, we need to define a coordinate system. Let X be a topological space and V be open in X. A sequence $(y^1, \cdots, y^n)$ of real valued functions on V is called an n-dimensional coordinate system on X with domain V if

$$\begin{aligned} y : V &\to y(V) \subset \mathbb{R}^n \\ x &\mapsto y(x) = (y^1(x), \cdots, y^n(x)) \end{aligned}$$

We also need to know that on a manifold the different coordinate systems agree on the points in the intersections of their domains. Let $y = (y^1, \cdots, y^n)$ with domain V and $z = (z^1, \cdots, z^n)$ with domain W be n-dimensional coordinate systems. We say that y and z are C^r-compatible if

$$\begin{aligned} y^i &= F^i(z^1, \cdots, z^n) \\ z^i &= G^i(y^1, \cdots, y^n) \end{aligned}$$

where F^i is C^r on $z(V \cap W)$ and G^i is C^r on $y(V \cap W)$.

Finally we can define a manifold. A topological space X is a C^r-manifold if a collection of mutually C^r-compatible coordinate systems is given whose domain cover X. Each coordinate system is a called a chart, and the collection is called an atlas.

We need some way to check if some equation describes a manifold. This is provided by the Implicit Function Theorem.

Theorem: (Implicit Function Theorem)
Let $f = (f^1, \cdots, f^l)$ be C^r real-valued functions on an open set V of $\mathbb{R}^n$, and let

$$X = \{x \in V | f(x) = 0\}$$

Let $a \in X$ such that $\text{rank} f'(a) = l$. Then there exists an open neighbourhood U of a in X such that $x^{l+1}, \cdots, x^n$ form an $n-l$ dimensional coordinate system with domain U and $x^1, \cdots, x^l$ are C^r functions of $x^{l+1}, \cdots, x^n$ on U.
Proof:
Assume the first l columns are linearly independent (we can do this without loss of generality). Put

$$F = (f^1, \cdots, f^l, x^{l+1}, \cdots, x^n).$$

Then:

$$F' = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \dots & \frac{\partial f^1}{\partial x^l} & \frac{\partial f^1}{\partial x^{l+1}} & \dots & \frac{\partial f^1}{\partial x^n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^l}{\partial x^1} & \dots & \frac{\partial f^l}{\partial x^l} & \frac{\partial f^l}{\partial x^{l+1}} & \dots & \frac{\partial f^l}{\partial x^n} \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

Clearly:

$$\det F' = \det \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \dots & \frac{\partial f^1}{\partial x^l} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^l}{\partial x^1} & \dots & \frac{\partial f^l}{\partial x^l} \end{pmatrix}$$

We assumed that the first l columns were linearly independent at a, so $\det F' \neq 0$ at a.

By the inverse function theorem, F maps an open neighbourhood W of a in $\mathbb{R}^n$ onto an open set $F(W)$ by a C^r-diffeomorphism. Let $G = F^{-1}$.

Put $U = X \cap W$, an open neighbourhood of a in X. In X (and hence in U), the first l coordinates are 0. Therefore $x^{l+1}, \cdots, x^n$ map U homeomorphically onto open $F(U)$ in $\mathbb{R}^{n-l}$. So $x^{l+1}, \cdots, x^n$ form an $(n-l)$ dimensional coordinate system on X, domain U.

Also if $G = (G^1, \cdots, G^n)$, then on U:

$$\begin{aligned} x^1 &= G^1(0, 0, \cdots, 0, x^{l+1}, \cdots, x^n) \\ &\vdots \\ x^l &= G^l(0, 0, \cdots, 0, x^{l+1}, \cdots, x^n) \end{aligned}$$

Each of these are C^r functions, as required. ■

An immediate corollary to this is that if

$$X = \{x \in V| \text{ rank of } f'(x) = l \text{ and } f(x) = 0\}$$

then X is a C^r $(n-l)$ dimmensional manifold.

5.2 Tangent Vectors

A real vector function f is called smooth at a if its domain U is an open neighbourhood of a and there exists coordinates y^i at a such that

$$f = F(y^1, \cdots, y^n)$$

on an open neighbourhood of a with F being C^∞. We then write:

$$\frac{\partial F}{\partial y^j}(a) = \frac{\partial F}{\partial x^j}(y^1(a), \cdots, y^n(a)) = \frac{d}{dt}F(y^1(a), \cdots, y^j(a) + t, \cdots, y^n(a))|_{t=0}$$