

# Differential Geometry

*A Summary of Loring W. Tu's An Introduction to Manifolds, Second Edition*  
*Ari Feiglin*

## Contents

<b>1</b>	<b>Euclidean Spaces</b>	<b>2</b>
1.1	Taylor's Theorem . . . . .	2
	Taylor's Theorem with Remainder . . . . .	2
1.2	Tangent Vectors . . . . .	3
1.3	The Exterior Algebra of Multivectors . . . . .	6
1.4	Solutions to Exercises . . . . .	13
1.5	Differential Forms on $\mathbb{R}^n$ . . . . .	18

# 1 Euclidean Spaces

## 1.1 Taylor's Theorem

We begin with a result which you may have shown in calculus three, but which nevertheless I am including here. First we define

**Definition 1.1.1:**

A set  $S \subseteq \mathbb{R}^n$  is **star-shaped** with respect to  $p \in S$  if for every point  $x \in S$ , the line connecting  $p$  to  $x$  is contained within  $S$ .

Thus, it is obvious that every convex set is star-shaped with respect to every point in it. In particular, balls are star-shaped.

**Theorem 1.1.2 (Taylor's Theorem with Remainder):**

If  $\mathcal{U} \subseteq \mathbb{R}^n$  is open and star-shaped with respect to  $p$ , then for every function  $f \in C^\infty(\mathcal{U})$ , there exist  $C^\infty(\mathcal{U})$  functions  $g_1, \dots, g_n$  such that

$$f(x) = f(p) + \sum_{i=1}^n (x^i - p^i) \cdot g_i(x), \quad x \in \mathcal{U}$$

and

$$g_i(p) = \frac{\partial f}{\partial x^i}(p)$$

**Proof:**

For  $x \in \mathcal{U}$  we can focus on the line segment from  $p$  to  $x$ ,  $p + t(x - p)$ . Then by the chain rule,

$$\frac{d}{dt} f(p + t(x - p)) = \nabla f(p + t(x - p)) \cdot (x - p) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(p + t(x - p)) \cdot (x^i - p^i)$$

Integrating this from  $t = 0$  to  $t = 1$  gives, by the fundamental theorem of calculus,

$$f(p + t(x - p)) \Big|_{t=0}^1 = \sum_{i=1}^n (x^i - p^i) \cdot \int_0^1 \frac{\partial f}{\partial x^i}(p + t(x - p)) dt$$

And thus we get

$$f(x) = f(p) + \sum_{i=1}^n (x^i - p^i) \cdot \int_0^1 \frac{\partial f}{\partial x^i}(p + t(x - p)) dt$$

So if we define

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(p + t(x - p)) dt$$

Then  $g_i$  is in  $C^\infty(\mathcal{U})$  since it is the integral of a  $C^\infty(\mathcal{U})$  function, and for every  $x \in \mathcal{U}$ ,

$$f(x) = f(p) + \sum_{i=1}^n (x^i - p^i) \cdot g_i(x)$$

and finally we have that

$$g_i(p) = \int_0^1 \frac{\partial f}{\partial x^i}(p) dt = \frac{\partial f}{\partial x^i}(p)$$

as required. ■

In the case of  $n = 1$ , then

$$f(x) = f(p) + (x - p)f_1(x)$$

where  $f_1(p) = f'(p)$ . But since  $f_1(x)$  is also in  $C^\infty$ , we can apply Taylor's theorem on it as well and get  $f_1(x) =$

$f_1(p) + (x - p)f_2(x)$ . Continuing inductively, we have  $f_i(x) = f_i(p) + (x - p)f_{i+1}(x)$ , and so for any  $n$  we have

$$\begin{aligned} f(x) &= f(p) + (x - p)f_1(p) + (x - p)^2f_2(p) + \cdots + (x - p)^{n-1}f_{n-1}(p) + (x - p)^nf_n(x) \\ &= \sum_{k=0}^{n-1} (x - p)^k f_k(p) + (x - p)^n f_n(x) \end{aligned}$$

where  $f_0 = f$ . Differentiating this  $m$  times gives

$$f^{(m)}(x) = \sum_{k=0}^{n-1} k \cdots (k - m + 1) \cdot (x - p)^{k-m} f_k(p) + ((x - p)^n f_n(x))^{(m)}$$

So if  $m < n$ , we get that the  $m$ th derivative of  $(x - p)^n f_n(x)$  is zero when  $x = p$ , and so we get

$$f^{(m)}(p) = m! \cdot f_m(p)$$

Thus we have that  $f_k(p) = \frac{f^{(k)}(p)}{k!}$  and so

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(p)}{k!} \cdot (x - p)^k + (x - p)^n f_n(x)$$

Which agrees with the first  $n$  coefficients of  $f$ 's Taylor series, and  $(x - p)^n f_n(x)$  is its remainder.

## 1.2 Tangent Vectors

Let us define a few basic mathematical concepts.

### Definition 1.2.1:

Suppose  $p \in \mathbb{R}^n$  is a point, then for any two neighborhoods of  $p$ ,  $\mathcal{U}$  and  $\mathcal{V}$ , and function  $f: \mathcal{U} \rightarrow \mathbb{R}$  and  $g: \mathcal{V} \rightarrow \mathbb{R}$ , we say that  $f$  and  $g$  are equivalent if there exists another neighborhood of  $p$ ,  $\mathcal{W} \subseteq \mathcal{U} \cap \mathcal{V}$  such that  $f$  and  $g$  agree on  $\mathcal{W}$ :  $f|_{\mathcal{W}} = g|_{\mathcal{W}}$ . This is denoted  $(f, \mathcal{U}) \sim_p (g, \mathcal{V})$ , and  $\sim_p$  is obviously an equivalence relation. The equivalence class of a function under this relation is called the **germ** of  $f$  at  $p$ .

Let us reduce the functions which we focus on to be  $C^\infty(\mathbb{R}^n)$  functions, then we denote the partition as

$$C_p^\infty(\mathbb{R}^n) = C^\infty(\mathbb{R}^n) / \sim_p$$

### Definition 1.2.2:

An algebra over a field  $\mathbb{F}$  (or a  $\mathbb{F}$ -algebra) is a  $\mathbb{F}$ -vector space  $A$  such which is equipped with a multiplication function

$$\times: A \times A \rightarrow A$$

Such that multiplication satisfies

- (1) Associativity:  $a \times (b \times c)$
- (2) Distributivity on the left and right:  $a \times (b + c) = (a \times b) + (a \times c)$  and  $(a + b) \times c = (a \times c) + (b \times c)$
- (3) Homogeneity: for every  $r \in \mathbb{F}$ ,  $r(a \times b) = (ra) \times b = a \times (rb)$

Thus an algebra is an vector space which is also an rng satisfying homogeneity. Trivial examples of algebras are fields, as they are algebras over themselves. Another example is  $C_p^\infty(\mathbb{R}^n)$ : it is a real algebra whose multiplication function is the product of functions.

### Definition 1.2.3:

Let  $p \in \mathbb{R}^n$  be a point in  $n$ -dimensional Euclidean space. A **tangent vector** to  $p$  is a vector which originates at  $p$ .

Formally we define  $p$ 's tangent space to be

$$T_p\mathbb{R}^n = \{(v, p) \mid v \in \mathbb{R}^n\}$$

which forms a real vector space where addition and scalar multiplication are defined

$$(v, p) + (u, p) = (v + u, p), \quad \alpha(v, p) = (\alpha v, p)$$

Obviously  $T_p\mathbb{R}^n \cong \mathbb{R}^n$ , so this definition may seem redundant. We will nevertheless use it, because in the future we will generalize this definition to manifolds. Now, recall the definition of the directional derivative of a function  $f$  at  $p$  in the direction  $v \in T_p\mathbb{R}^n$ :

$$D_v f|_p = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t} = \nabla f|_p \cdot v$$

Here  $v$  is a tangent vector to  $p$ , this is mathematically the same as just taking  $v \in \mathbb{R}^n$  for now. The directional derivative satisfies

$$D_v(f \cdot g)|_p = D_v(f)|_p \cdot g(p) + f(p) \cdot D_v(g)|_p$$

So we can generalize this:

**Definition 1.2.4:**

Let  $p \in \mathbb{R}^n$ . A linear map  $D: C_p^\infty(\mathbb{R}^n)$  which satisfies Leibniz's rule:

$$D(f \cdot g) = D(f) \cdot g(p) + f(p) \cdot D(g)$$

is called a **point derivation** at  $p$ . We define  $D_p(\mathbb{R}^n)$  to be the set of all point derivations at  $p$ .

So  $D_v|_p$  is a point derivation at  $p$  for every  $v \in T_p\mathbb{R}^n$ . And in general, if  $D$  is a point derivation at  $p$ , then

$$D(1) = D(1 \cdot 1) = D(1) \cdot 1 + 1 \cdot D(1) = 2D(1)$$

so  $D(1) = 0$ . And since  $D$  is linear, for any constant  $c$ ,  $D(c) = cD(1) = 0$ .

Now, notice that we can define a the function

$$\varphi: T_p\mathbb{R}^n \rightarrow D_p(\mathbb{R}^n), \quad v \mapsto D_v|_p$$

It turns out that this is actually a linear isomorphism.

**Theorem 1.2.5:**

$\varphi$  is a linear isomorphism,  $T_p\mathbb{R}^n \cong D_p(\mathbb{R}^n)$ .

**Proof:**

$\varphi$  is linear since  $D_{v+u} = D_v + D_u$  and  $D_{\alpha v} = \alpha D_v$  (think of the dot product). And if  $\varphi(v) = 0$  then this would mean that for every  $f \in C_p^\infty(\mathbb{R}^n)$ ,  $D_v f|_p = 0$ . Let us take  $f(x) = x^i$ , the projection of  $x$  onto its  $i$ th coordinate, then

$$0 = D_v x^i|_p = \nabla x^i|_p \cdot v = e_i \cdot v = v^i$$

meaning  $v = 0$ , so  $\text{Ker}(\varphi) = 0$  meaning  $\varphi$  is injective.

Now we must show that  $\varphi$  is surjective. Let  $(f, \mathcal{U}) \in C_p^\infty(\mathbb{R}^n)$  then by **Taylor's Theorem with Remainder** there exist  $C^\infty(\mathcal{U})$  functions  $g_1, \dots, g_n$  such that

$$f(x) = f(p) + \sum_{k=1}^n (x^k - p^k) g_k(x)$$

where  $g_k(p) = \frac{\partial f}{\partial x^k}(p)$ . Now suppose  $D$  is a point derivation at  $p$ , then by linearity

$$D(f) = D(f(p)) + \sum_{k=1}^n D((x^k - p^k) g_k(x))$$

Since  $f(p)$  is a constant, we have  $D(f(p)) = 0$ . And since  $D$  is a point derivation we get that this is equal to

$$D(f) = \sum_{k=1}^n D(x^k - p^k)g_k(p) + (p^k - p^k)D(g_k(x))$$

Since  $D(x^k - p^k) = D(x^k) - D(p^k) = D(x^k)$  we get

$$D(f) = \sum_{k=1}^n D(x^k)g_k(p) = \sum_{k=1}^n \frac{\partial f}{\partial x^k}(p) \cdot D(x^k)$$

And so if we define  $v = (D(x^1), \dots, D(x^n))$ , then we get that

$$D(f) = \nabla f|_p \cdot v = D_v f|_p$$

Thus we have that  $D = D_v|_p$ , meaning that  $\varphi(v) = D$  so  $\varphi$  is indeed surjective. ■

This isomorphism is natural, and so from now on we will identify  $T_p\mathbb{R}^n$  with  $D_p\mathbb{R}^n$ . While  $T_p\mathbb{R}^n$  is much more geometric than  $D_p\mathbb{R}^n$ ,  $D_p\mathbb{R}^n$  generalizes easier. Also, note that under  $\varphi$ ,

$$e_i \mapsto D_{e_i}|_p = \frac{\partial}{\partial x^i}|_p$$

And so the canonical basis of  $T_{\mathbb{R}}^n$  (which we're now identifying with  $D_p\mathbb{R}^n$ ) is

$$\left\{ \frac{\partial}{\partial x^1}|_p, \dots, \frac{\partial}{\partial x^n}|_p \right\}$$

So for a tangent vector  $v = \sum_{i=1}^n v^i e_i$ , we write instead

$$v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}|_p$$

#### Definition 1.2.6:

If  $\mathcal{U}$  is an open subset of  $\mathbb{R}^n$ , a **vector field** on  $\mathcal{U}$  is a function

$$X: \mathcal{U} \longrightarrow \bigcup_{p \in \mathcal{U}} T_p\mathbb{R}^n$$

where for every  $p \in \mathcal{U}$ ,  $X(p) = X_p \in T_p\mathbb{R}^n$ . Thus

$$X_p = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x^i}|_p$$

So  $a_i$  forms a function  $a_i: \mathcal{U} \longrightarrow \mathbb{R}$ , and we write

$$X = \sum_{i=1}^n a_i \frac{\partial}{\partial x^i}$$

If every  $a_i$  is in  $C^\infty(\mathcal{U})$ , then we say that  $X$  is also  $C^\infty$ . We define  $\mathfrak{X}(\mathcal{U})$  to be the set of all  $C^\infty$  vector fields on  $\mathcal{U}$ .

Notice that the sum of two  $C^\infty$  vector fields is another  $C^\infty$  vector fields, and the product of a  $C^\infty$  vector field by a  $C^\infty(\mathcal{U})$  function is still a  $C^\infty$  vector field. Meaning that if  $X$  and  $Y$  are both in  $\mathfrak{X}(\mathcal{U})$ , then so is  $X + Y$ . And if  $f \in C^\infty(\mathcal{U})$ , then  $f \cdot X$  is still in  $\mathfrak{X}(\mathcal{U})$ . Thus  $\mathfrak{X}(\mathcal{U})$  is a real vector space, and a module over  $C^\infty(\mathcal{U})$ .

Suppose

$$X = \sum_{i=1}^n a_i \frac{\partial}{\partial x^i}$$

Now, since  $X_p \in T_p\mathbb{R}^n$ , it acts on  $C^\infty$  functions:

$$X_p(f) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x^i} f \Big|_p$$

So we can define a function

$$Xf: \mathcal{U} \longrightarrow C^\infty, \quad (Xf)_p = X_p(f)$$

This means that

$$Xf = \sum_{i=1}^n a_i \cdot \frac{\partial f}{\partial x^i}$$

If  $X$  is  $C^\infty$ , then  $Xf$  is in  $C^\infty(\mathcal{U})$ . Obviously  $X(f+g) = Xf + Xg$  and  $X(\alpha f) = \alpha X(f)$ , so a vector fields in  $\mathfrak{X}(\mathcal{U})$  can be viewed as linear operators over  $C^\infty(\mathcal{U})$  mapping  $f$  to  $Xf$ .

Notice that since  $X_p$  is a point derivation at  $p$ :

$$X_p(f \cdot g) = X_p(f) \cdot g(p) + f(p) \cdot X_p(g)$$

and so we have that

$$X(f \cdot g) = X(f) \cdot g + f \cdot X(g)$$

We can generalize this idea

**Definition 1.2.7:**

Let  $A$  be an algebra over  $\mathbb{F}$ , then a **derivation** of  $A$  is an  $\mathbb{F}$ -linear operator  $D$  such that for every  $a, b \in A$ :

$$D(ab) = D(a)b + aD(b)$$

If  $D_1$  and  $D_2$  are two derivations of  $A$ , so is  $D_1 + D_2$ , and if  $\alpha \in \mathbb{F}$  then  $\alpha D_1$  is also a derivation. Thus the set of all derivations of  $A$  is an  $\mathbb{F}$ -vector space, which we denote  $\text{Der}(A)$ .

This means that every  $C^\infty$  vector field over  $\mathcal{U}$  is a derivation of  $C^\infty(\mathcal{U})$  in the sense that it can be viewed as a map  $f \mapsto Xf$ . More formally,

$$\varphi(X) = (f \mapsto Xf)$$

is a linear transformation from  $\mathfrak{X}(\mathcal{U})$  to  $\text{Der}(C^\infty(\mathcal{U}))$ . It turns out that this function is actually an isomorphism (injectivity is simple, surjectivity is a little harder; we will prove this later). So we can identify tangent vectors with point derivations, we can identify vector fields (functions of tangent vectors) with derivations (functions of point derivations). That is, we have the following correspondences:

$$\begin{aligned} T_p\mathbb{R}^n &\longleftrightarrow D_p\mathbb{R}^n \\ \mathfrak{X}(\mathcal{U}) &\longleftrightarrow \text{Der}(C^\infty(\mathcal{U})) \end{aligned}$$

### 1.3 The Exterior Algebra of Multivectors

Let us recall a few definitions from linear algebra:

**Definition 1.3.1:**

Let  $V$  be a finite dimensional vector space, then its **dual space**, denoted  $V^*$ , is defined to be the set of all linear functionals on  $V$  (linear transformations from  $V$  to  $\mathbb{F}$ ). If  $B = \{e_1, \dots, e_n\}$  is a basis for  $V$ , its **dual basis** is defined to be  $B^* = \{\alpha^1, \dots, \alpha^n\}$  where  $\alpha^i$  is the linear functional defined by

$$\alpha^i(e_j) = \delta_j^i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

That is,  $\alpha^i(v)$  maps to  $e_i$ 's coefficient in  $v$ 's representation under  $B$ . The dual basis, being true to its name, is a basis for  $V^*$  and so  $V \cong V^*$ .

**Definition 1.3.2:**

Let  $V$  and  $W$  be vector spaces, then a function

$$f: V^k \longrightarrow W$$

is **multilinear** if  $f$  is linear in every argument:

$$f(\dots, \alpha v + \beta w, \dots) = \alpha f(\dots, v, \dots) + \beta f(\dots, w, \dots)$$

A multilinear function from  $V^k$  is also called  **$k$ -linear**.

For example, inner products are 2-linear (bilinear), and the determinant (viewed as a function taking  $n$  inputs) is  $n$ -linear.

**Definition 1.3.3:**

Let  $f$  be a function in  $k$  arguments, and  $\sigma$  be a permutation in  $S_k$ . Then we define the function  $\sigma f$  by

$$(\sigma f)(v_1, \dots, v_k) = f(v_{\sigma(1)}, \dots, v_{\sigma(k)})$$

A function  $f$  is **symmetric** if for every  $\sigma \in S_k$ ,  $\sigma f = f$  (permuting the order of the arguments doesn't change the result). And a function  $f$  is **alternating** if for every  $\sigma \in S_k$ ,  $\sigma f = \text{sgn}(\sigma)f$ .

It is not hard to see that

$$\tau(\sigma f) = (\tau\sigma)f$$

this is since

$$(\tau(\sigma f))(v_1, \dots, v_n) = (\sigma f)(v_{\tau(1)}, \dots, v_{\tau(n)})$$

Now, setting  $w_i = v_{\tau(i)}$  we have

$$= (\sigma f)(w_1, \dots, w_n) = f(w_{\sigma(1)}, \dots, w_{\sigma(n)}) = f(v_{\tau(\sigma(1))}, \dots, v_{\tau(\sigma(n))}) = (\tau\sigma)f(v_1, \dots, v_n)$$

This means that  $(\sigma, f) \mapsto \sigma f$  defines a group action of  $S_k$  on a set of functions in  $k$  arguments.

**Proposition 1.3.4:**

A function  $f$  is alternating if and only if for every transposition  $\tau$  (meaning a permutation which only swaps two values),  $\tau f = -f$ . If  $f$  is a linear transformation, then  $f$  is alternating if and only if whenever one of its arguments are repeated, it is equal to zero (meaning  $f(\dots, v, \dots, v, \dots) = 0$ ).

**Proof:**

If  $f$  is alternating, since the sign of every transposition is  $-1$ , it is true by definition that  $\tau f = -f$ . Now, if for every transposition  $\tau$ ,  $\tau f = -f$  we must show that  $f$  is alternating. Let  $\sigma$  be a permutation, so there exist transpositions  $\tau_1, \dots, \tau_k$  such that  $\sigma = \tau_1 \cdots \tau_k$ . This means that  $\text{sgn}(\sigma) = \text{sgn}(\tau_1) \cdots \text{sgn}(\tau_k) = (-1)^k$ . So for every  $v_1, \dots, v_n$  we get

$$\sigma f(v_1, \dots, v_n) = f(v_{\tau_1 \cdots \tau_k(1)}, \dots, v_{\tau_1 \cdots \tau_k(n)})$$

Setting  $u_i = v_{\tau_2 \cdots \tau_k(i)}$  we get that this is equal to

$$= f(u_{\tau_1(1)}, \dots, u_{\tau_1(n)}) = \tau_1 f(u_1, \dots, u_n) = -f(u_1, \dots, u_n) = -f(v_{\tau_2 \cdots \tau_k(1)}, \dots, v_{\tau_2 \cdots \tau_k(n)})$$

Continuing inductively on  $k$  we get

$$= (-1)^k f(v_1, \dots, v_k) = \text{sgn}(\sigma)f(v_1, \dots, v_k)$$

as required.

Obviously if  $f$  is alternating and  $v_i = v_j$  for  $i \neq j$ , then let  $\tau = (i, j)$  be a transposition then

$$\tau f(v_1, \dots, v_n) = f(v_{\tau(1)}, \dots, v_{\tau(n)}) = f(v_1, \dots, v_n)$$

But at the same time since  $f$  is alternating,  $\tau f = -f$  and so we get

$$f(v_1, \dots, v_n) = -f(v_1, \dots, v_n)$$

and so  $f(v_1, \dots, v_n) = 0$  as required.

And if  $f$  is multilinear and repeating an argument gives zero, then let  $\tau = (i, j)$  (suppose  $i < j$ ) be a transposition. For every  $v_1, \dots, v_n \in V$  notice that

$$f(v_1, \dots, v_n) + \tau f(v_1, \dots, v_n) = f(\dots, v_j, \dots, v_i, \dots) + f(\dots, v_i, \dots, v_j, \dots)$$

Now notice that by multilinearity

$$\begin{aligned} f(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\ = f(\dots, v_i, \dots, v_i, \dots) + f(\dots, v_i, \dots, v_j, \dots) + f(\dots, v_j, \dots, v_i, \dots) + f(\dots, v_j, \dots, v_j, \dots) \end{aligned}$$

Since  $f$  is zero when an argument is repeated, we get that this is equal to

$$= f(\dots, v_i, \dots, v_j, \dots) + f(\dots, v_j, \dots, v_i, \dots)$$

Thus we have that

$$f(v_1, \dots, v_n) + \tau f(v_1, \dots, v_n) = f(\dots, v_i + v_j, \dots, v_i + v_j, \dots)$$

Since the right hand repeats an argument, it is equal to zero and so we get

$$\tau f(v_1, \dots, v_n) = -f(v_1, \dots, v_n)$$

So we have shown that for every transposition  $\tau$ , we have  $\tau f = -f$ . We showed that this is equivalent to being an alternating function, as required. ■

#### Definition 1.3.5:

Let us define

- (1)  $L_k(V)$  to be the set of all  $k$ -linear functionals over  $V$ ,
- (2)  $S_k(V)$  to be the set of all symmetric  $k$ -linear functionals over  $V$ , and
- (3)  $A_k(V)$  to be the set of all alternating  $k$ -linear functionals over  $V$ .

These spaces are obviously linear (ie. vector spaces).

For example, real inner products are symmetric, while determinants are alternating. When  $k = 0$ , then 0-linear functionals are simply scalars and so

$$L_0(V) = S_0(V) = A_0(V) = \mathbb{F}$$

and when  $k = 1$ ,  $S_1 = \{\text{id}\}$  and so

$$L_1(V) = S_1(V) = A_1(V) = V^*$$

#### Definition 1.3.6:

Elements of  $L_k(V)$  are called **tensors** (specifically,  $k$ -**tensors**). Elements of  $A_k(V)$  are called  $k$ -**covectors**. In the case that  $k = 1$ , they are also called simply **covectors**, and if  $k > 1$  **multicovectors**.

Now, if  $f$  is a  $k$ -linear transformation then let us define new  $k$ -linear transformations

$$Sf = \sum_{\sigma \in S_k} \sigma f, \quad Af = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma f$$

Notice that  $Sf$  is symmetric:

$$\tau(Sf) = \sum_{\sigma \in S_k} (\tau\sigma) f$$

and since  $\sigma \mapsto \tau\sigma$  is bijective, this is equal to

$$= \sum_{\sigma \in S_k} \sigma f = Sf$$

And similarly  $Af$  is alternating:

$$\tau(Af) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) (\tau\sigma) f = \text{sgn}(\tau) \sum_{\sigma \in S_k} \text{sgn}(\tau\sigma) (\tau\sigma) f = \text{sgn}(\tau) Af$$



If  $f$  is already symmetric then

$$Sf = \sum_{\sigma \in S_k} \sigma f = \sum_{\sigma \in S_k} f = k! \cdot f$$

and

$$Af = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma f = \left( \sum_{\sigma \in S_k} \text{sgn}(\sigma) \right) f$$

and since for  $k > 1$  there is an equal number of even and odd permutations, the sum of the signs is equal to zero. So this is equal to zero.

And if  $f$  is alternating then

$$Af = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma f = \sum_{\sigma \in S_k} f = k! \cdot f, \quad Sf = 0$$

### Definition 1.3.7:

Suppose  $V$  is a vector space, and  $f$  and  $g$  is a  $k$ -linear and  $\ell$ -linear functional respectively. We define their **tensor product** to be a  $k + \ell$ -linear functional defined by

$$(f \otimes g)(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+\ell}) = f(v_1, \dots, v_k) \cdot g(v_{k+1}, \dots, v_{k+\ell})$$

It is obvious that  $f \otimes g$  is indeed  $k + \ell$ -linear. For example, if  $E = \{e_1, \dots, e_n\}$  is an orthonormal basis for  $V$  and  $B^* = \{\alpha^1, \dots, \alpha^n\}$  is its dual basis then the inner product of the vector space is given by

$$\langle v, w \rangle = \sum_{i=1}^n \alpha^i(v) \cdot \alpha^i(w) = \sum_{i=1}^n (\alpha^i \otimes \alpha^i)(v, w)$$

Now, notice that the tensor product of two alternating functions need not be alternating (for example if  $\sigma = (1, 2)$  then  $f \otimes g(v_1, v_2) = f(v_2)g(v_1)$  which need not have any relation with  $f \otimes g(v_1, v_2)$ ). So we also define the following

### Definition 1.3.8:

Let  $f$  be an alternating  $k$ -linear functional, and  $g$  be an alternating  $\ell$ -linear functional, both over the same vector space. Then we define their **wedge product** (also called the **exterior product**) to be an alternating  $k + \ell$ -linear functional defined by

$$f \wedge g = \frac{1}{k!\ell!} A(f \otimes g)$$

Notice then that

$$f \wedge g(v_1, \dots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) \cdot f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \cdot g(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)})$$

Notice that for every  $\sigma \in S_{k+\ell}$ ,  $\tau_1 \in S_k$  and  $\tau_2 \in S_\ell$  (though we will view  $\tau_2$  as permuting the indexes  $\{k+1, \dots, k+\ell\}$ ), we have

$$\text{sgn}(\sigma \tau_1 \tau_2) (\sigma \tau_1 \tau_2) f \otimes g = \text{sgn}(\sigma \tau_1 \tau_2) \sigma (\tau_1 f) \otimes (\tau_2 g)$$

since  $\tau_2$  permutes only the last  $\ell$  indexes and thus does not affect  $f$ , and similarly  $\tau_1$  does not affect  $g$ . Since  $f$  and  $g$  are alternating, we have that this is equal to

$$= \text{sgn}(\sigma \tau_1 \tau_2) \text{sgn}(\tau_1) \text{sgn}(\tau_2) \sigma (f \otimes g) = \text{sgn}(\sigma) \sigma (f \otimes g)$$

So for every  $\sigma \in S_{k+\ell}$ , there exist  $|S_k \times S_\ell| = k!\ell!$  other permutations which give the same factor in the sum (permutations of the form  $\sigma \tau_1 \tau_2$  as above). This is a reasoning for dividing the wedge product by  $k!\ell!$ , to get rid of the repeated factors. If we define an equivalence  $\sigma \sim \sigma'$  if and only if  $\sigma'$  is of the form  $\sigma \tau_1 \tau_2$  for  $\tau_1 \in S_k$  and  $\tau_2 \in S_\ell$ , then we have an equivalence relation on  $S_{k+\ell}$ , and by above

$$f \wedge g = \sum_{\sigma \in S_{k+\ell} / \sim} \text{sgn}(\sigma) \sigma (f \otimes g)$$

Since equivalent permutations give the same factor. Now, every equivalence class has a permutation  $\sigma$  where  $\sigma(1) < \dots < \sigma(k)$  and  $\sigma(k+1) < \dots < \sigma(k+\ell)$ , this is called a  $(k, \ell)$ -shuffle. We know this since we can define  $\tau_1$  to reorder  $\{1, \dots, k\}$  in order to ensure that the values become ascending, and similar for  $\tau_2$ . Two  $(k, \ell)$ -shuffles cannot be equivalent and so we get the following

**Proposition 1.3.9:**

$$f \wedge g = \sum_{\substack{\sigma \text{ is a} \\ (k, \ell)\text{-shuffle}}} \text{sgn}(\sigma) \sigma(f \otimes g)$$

There are  $\binom{k+\ell}{k}$  shuffles: choose  $k$  values for  $\sigma(1), \dots, \sigma(k)$ , this defines  $\sigma(1), \dots, \sigma(k)$ , and since  $\sigma$  is a shuffle this also defines  $\sigma(k+1), \dots, \sigma(k+\ell)$ . Since there are  $\binom{k+\ell}{k}$  ways to choose the first  $k$  values, this means there are  $\binom{k+\ell}{k}$  shuffles. So instead of summing over  $k!\ell!$  factors, we can sum over  $\binom{k+\ell}{k}$  in order to compute the wedge product. The wedge product is also obviously distributive in both arguments, this follows from the fact that  $f \otimes (g+h) = f \otimes g + f \otimes h$ , and  $A(f+g) = A(f) + A(g)$ . It is also obviously homogeneous, again since  $f \otimes (c \cdot g) = c \cdot (f \otimes g)$  and  $A(cf) = c \cdot A(f)$ .

**Proposition 1.3.10:**

The wedge product is both left and right distributive,

$$f \wedge (g + h) = f \wedge g + f \wedge h, \quad (f + g) \wedge h = f \wedge h + g \wedge h$$

and homogeneous in both arguments

$$(cf) \wedge g = f \wedge (cg) = c(f \wedge g)$$

**Proposition 1.3.11:**

The wedge product is antisymmetric: if  $f$  is a  $k$ -covector and  $g$  is an  $\ell$ -covector then

$$f \wedge g = (-1)^{k\ell} g \wedge f$$

**Proof:**

Let us define the following permutation:

$$\tau = \begin{pmatrix} 1 & \cdots & \ell & \ell+1 & \cdots & \ell+k \\ k+1 & \cdots & k+\ell & 1 & \cdots & k \end{pmatrix}$$

Thus for every  $\sigma \in S_{k+\ell}$ , we have that for  $1 \leq i \leq k$ :  $\sigma(i) = \sigma\tau(\ell+i)$ , and for  $1 \leq i \leq \ell$ :  $\sigma(k+i) = \sigma\tau(i)$ . Thus

$$\begin{aligned} A(f \otimes g)(v_1, \dots, v_{k+\ell}) &= \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \cdot g(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}) \\ &= \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) g(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(\ell)}) \cdot f(v_{\sigma\tau(\ell+1)}, \dots, v_{\sigma\tau(\ell+k)}) \\ &= \text{sgn}(\tau) \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma\tau) g(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(\ell)}) \cdot f(v_{\sigma\tau(\ell+1)}, \dots, v_{\sigma\tau(\ell+k)}) \\ &= \text{sgn}(\tau) A(g \otimes f) \end{aligned}$$

Now, let us try to compute  $\tau$ . We can do this computing the number of inversions of  $\tau$ . Notice that all of the inversions of  $\tau$  are of the form  $\{i, \ell+j\}$  where  $1 \leq i \leq \ell$  and  $1 \leq j \leq k$ . Thus there are  $k\ell$  inversions, and so  $\text{sgn}(\tau) = (-1)^{k\ell}$  meaning

$$A(f \otimes g) = (-1)^{k\ell} A(g \otimes f)$$

And dividing both sides by  $k!\ell!$  yields

$$f \wedge g = (-1)^{k\ell} g \wedge f \quad \blacksquare$$

An immediate corollary of this is that if  $f$  is a covector (meaning 1-covector), then  $f \wedge f = 0$ . This is as  $f \wedge f = -f \wedge f$  by the proposition above, and so  $f \wedge f = 0$ . This is obviously true for all  $2k+1$ -covectors.

**Corollary 1.3.12:**

If  $f$  is a covector, then  $f \wedge f = 0$ .

We will now establish the associativity of the wedge product. To do so, we will first prove a lemma:

**Lemma 1.3.13:**

If  $f$  is a  $k$ -covector and  $g$  is an  $\ell$ -covector, then

- (1)  $A(A(f) \otimes g) = k!A(f \otimes g)$
- (2)  $A(f \otimes A(g)) = \ell!A(f \otimes g)$

**Proof:**

- (1) We will prove this directly,

$$A(A(f) \otimes g) = \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) \sigma(A(f) \otimes g) = \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) \sigma \left( \left( \sum_{\tau \in S_k} \text{sgn}(\tau) \tau f \right) \otimes g \right)$$

Since  $(f_1 + f_2) \otimes g = f_1 \otimes g + f_2 \otimes g$ , this is equal to

$$= \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) \sigma \left( \sum_{\tau \in S_k} \text{sgn}(\tau) \tau(f) \otimes g \right)$$

And it is similarly obvious that  $\sigma(f_1 + f_2) = \sigma(f_1) + \sigma(f_2)$ , and so

$$= \sum_{\sigma \in S_{k+\ell}} \sum_{\tau \in S_k} \text{sgn}(\sigma\tau) \sigma(\tau(f) \otimes g)$$

Since  $\tau \in S_k$ ,  $\tau(f \otimes g) = \tau(f) \otimes g$  and so this is equal to

$$= \sum_{\sigma \in S_{k+\ell}} \sum_{\tau \in S_k} \text{sgn}(\sigma\tau) \sigma\tau(f \otimes g)$$

Now, for every  $\mu \in S_{k+\ell}$  there are  $k!$  ways to write  $\mu$  as  $\sigma\tau$  for  $\sigma \in S_{k+\ell}$  and  $\tau \in S_k$  (define  $\tau$ , then  $\sigma$  is defined by  $\mu\tau^{-1}$ ). So for every  $\mu \in S_{k+\ell}$  the above sum sums over the factor  $\text{sgn}(\mu)\mu(f \otimes g)$  exactly  $k!$  times, and so it is equal to

$$= k! \cdot \sum_{\mu \in S_{k+\ell}} \text{sgn}(\mu) \mu(f \otimes g) = k!A(f \otimes g)$$

as required.

- (2) Here we can utilize the fact that  $A(f \otimes g) = (-1)^{k\ell}A(g \otimes f)$  and what we just proved:

$$A(f \otimes A(g)) = (-1)^{k\ell}A(A(g) \otimes f) = (-1)^{k\ell}\ell!A(g \otimes f) = \ell!A(f \otimes g) \quad \blacksquare$$

We now establish the wedge product's associativity:

**Proposition 1.3.14:**

The wedge product is associative: if  $f$  is a  $k$ -covector,  $g$  an  $\ell$ -covector, and  $h$  an  $m$ -covector, then

$$(f \wedge g) \wedge h = f \wedge (g \wedge h)$$

**Proof:**

Notice that

$$(f \wedge g) \wedge h = \frac{1}{(k+\ell)!m!} A((f \wedge g) \otimes h) = \frac{1}{(k+\ell)!m!} \cdot \frac{1}{k!\ell!} A(A(f \otimes g) \otimes h)$$

By the above lemma, this is equal to

$$= \frac{(k+\ell)!}{k!\ell!m!} A(f \otimes g \otimes h)$$

Similarly, we get

$$f \wedge (g \wedge h) = \frac{1}{k!(\ell+m)!} A(f \otimes (g \wedge h)) = \frac{1}{k!(\ell+m)!m!} A(f \otimes A(g \otimes h)) = \frac{1}{k!\ell!m!} A(f \otimes g \otimes h)$$

Thus we have shown that  $(f \wedge g) \wedge h = f \wedge (g \wedge h)$ , as required. ■

An important note here is that we have just shown that

$$f \wedge g \wedge h = \frac{1}{k!\ell!m!} A(f \otimes g \otimes h)$$

And inductively we can extend this: if for  $1 \leq i \leq n$ ,  $f^i$  is a  $d_i$ -covector then

$$f^1 \wedge \cdots \wedge f^n = \frac{1}{(d_1)! \cdots (d_n)!} A(f^1 \otimes \cdots \otimes f^n)$$

Now, recall that the determinant is an alternating linear functional, meaning it is a multicovector. It turns out that there is a simple relation between the wedge product of covectors and the determinant.

**Proposition 1.3.15:**

Suppose  $\alpha^1, \dots, \alpha^k$  are covectors (meaning 1-covectors, linear functionals) over  $V$ , then if  $v_1, \dots, v_k$  are vectors in  $V$  then

$$(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \dots, v_k) = \det(\alpha^i(v_j))_{i,j=1}^k$$

The notation  $(a_{ij})_{i,j=1}^k$  means a  $k \times k$  matrix whose coefficient at the  $i, j$ th position is  $a_{ij}$ .

**Proof:**

We know that

$$(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \dots, v_k) = A(\alpha^1 \otimes \cdots \otimes \alpha^k)(v_1, \dots, v_k) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \alpha^1(v_{\sigma(1)}) \cdots \alpha^k(v_{\sigma(k)}) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \cdot \prod_{i=1}^k \alpha^i(v_{\sigma(i)})$$

This is, by definition, the determinant of the matrix  $(\alpha^i(v_j))_{i,j=1}^k$ , as required. ■

Now, let  $B = (e_1, \dots, e_n)$  be a basis for  $V$  and  $B^* = (\alpha^1, \dots, \alpha^n)$  be its dual basis. For a multi-index  $I = (i_1, \dots, i_k)$  (a tuple where  $1 \leq i_j \leq n$ ), let us define

$$e_I = (e_{i_1}, \dots, e_{i_k}), \quad \alpha^I = \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$$

Suppose  $f \in L_k(V)$  is a linear functional (this is true for multilinear functions in general), then  $f$  is determined uniquely by its value on  $e_I$  for all multi-indexes of length  $k$ . This is as every  $v \in V$  can be written as a linear combination of  $e_i$ s. Now, if  $f$  is alternating then it is determined uniquely by its value on ascending multi-indexes (meaning multi-indexes where  $i_1 \leq i_2 \leq \cdots \leq i_k$ ), as for every multi-index  $I$  there exists a permutation  $\sigma$  such that  $\sigma I = (i_{\sigma(1)}, \dots, i_{\sigma(k)})$  is ascending and

$$f(e^{\sigma I}) = f(\sigma e^I) = \text{sgn}(\sigma) f(e^I)$$

so it is sufficient to know  $f$ 's value on  $e^{\sigma I}$  in order to attain its value on  $e^I$ . Now, for multi-indexes which are not *strictly* ascending, there exist indexes  $i_\alpha = i_\beta$  where  $\alpha \neq \beta$  and so  $f(e_I) = 0$  as an alternating linear function is equal to zero when any of its arguments are repeated.

Now, suppose  $I$  and  $J$  are strictly ascending multi-indexes of length  $k$ , then notice that

$$\alpha^I(e_J) = \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}(e_{j_1}, \dots, e_{j_k}) = \det(\alpha^{i_\ell}(e_{j_m}))_{\ell,m}$$

If  $I \neq J$ , then there exists an  $i_\ell$  which is in  $I$  but not in  $J$  which means that for every  $j_m$ ,  $\alpha^{i_\ell}(e_{j_m}) = 0$ , so this determinant becomes zero. If  $I = J$  then

$$\alpha^{i_\ell}(e_{j_m}) = \delta_m^\ell = \begin{cases} 1 & \ell = m \\ 0 & \ell \neq m \end{cases}$$

So the matrix is the identity matrix, meaning the determinant is one. So we have shown

$$\alpha^I(e_J) = \delta_J^I = \begin{cases} 1 & I = J \\ 0 & I \neq J \end{cases}$$

Now utilizing this we can show

**Theorem 1.3.16:**

The set  $S = \{\alpha^I \mid I \text{ is a strictly ascending multi-index of length } k\}$  is a basis for the space of  $k$ -covectors,  $A_k(V)$ .

**Proof:**

We will first show that  $S$  is linearly independent. Suppose there exists a zeroing linear combination of  $S$ , which is of the form

$$0 = \sum_{I \in S} c_I \alpha^I$$

then let  $J$  be a strictly ascending multi-index of length  $k$ , (ie.  $J \in S$ ), then composing both sides with  $e_J$  gives

$$0 = \sum_{I \in S} c_I \alpha^I(e_J) = \sum_{I \in S} c_I \delta_J^I = c_J$$

so for every  $J \in S$ ,  $c_J = 0$ . This means that  $S$  is indeed linearly independent.

Now, if  $f$  is a  $k$ -covector, then we claim that

$$f = \sum_{I \in S} f(e_I) \alpha^I$$

(Think about why this is the only possibility.) Let  $g$  be the sum on the right. Since every  $k$ -covector is determined by its values on  $e_J \in S$ , it is sufficient to show that  $f(e_J) = g(e_J)$  for every  $J \in S$ . And so

$$g(e_J) = \sum_{I \in S} f(e_I) \alpha^I(e_J) = \sum_{I \in S} f(e_I) \delta_J^I = f(e_J)$$

as required. So the span of  $S$  is indeed  $A_k(V)$ , satisfying the conditions for it to be a basis. ■

Notice that a strictly ascending multi-index is determined entirely by its elements, without taking into account their order (because their order is defined). So there are  $\binom{n}{k}$  strictly ascending multi-indexes of length  $k$ , meaning

$$\dim A_k(V) = \binom{n}{k}$$

where  $\dim V = n$ . Notice that this means  $A_k(V) = 0$  for  $k > n$  (as there cannot be a strictly ascending multi-index which uses indexes within  $\{1, \dots, n\}$  but has a length of  $k$ ).

## 1.4 Solutions to Exercises

**Exercise 1.4.1:**

Let  $e_1, \dots, e_n$  be a basis for a vector space  $V$ , and  $\alpha^1, \dots, \alpha^n$  be its dual basis. Suppose  $(g_{ij})$  is an  $n \times n$  matrix. Define a bilinear function  $f: V \times V \rightarrow \mathbb{R}$  by

$$f(v, w) = \sum_{i,j=1}^n g_{ij} v^i w^j$$

for  $v = \sum v^i e_i$  and  $w = \sum w^j e_j$  in  $V$ . Describe  $f$  in terms of the tensor products of  $\alpha^i$  with  $\alpha^j$ .

Notice that  $v^i = \alpha^i(v)$  and  $w^j = \alpha^j(w)$  and so

$$f(v, w) = \sum_{i,j=1}^n g_{ij} \alpha^i(v) \alpha^j(w) = \sum_{i,j=1}^n g_{ij} \alpha^i \otimes \alpha^j(v, w) = \left( \sum_{i,j=1}^n g_{ij} \alpha^i \otimes \alpha^j \right)(v, w)$$

#### Exercise 1.4.2:

- (1) Let  $V$  be an  $n$  dimensional vector space and  $f: V \rightarrow \mathbb{R}$  a nonzero linear functional. Show that  $\dim \text{Ker } f = n-1$ . A linear subspace of  $V$  of dimension  $n-1$  is called a *hyperplane* in  $V$ .
- (2) Show that a nonzero linear functional on a vector space  $V$  is determined up to a multiplicative constant by its kernel, a hyperplane in  $V$ .

- (1) Notice that  $f(V) \neq 0$  and so  $f(V)$  must be a linear subspace of  $f$ 's codomain,  $\mathbb{R}$ . Since  $\mathbb{R}$  is of dimension one, this means  $f(V) = \mathbb{R}$  (proving this directly is also very simple). Thus  $\dim f(V) = 1$ . Recall the rank-nullity theorem from linear algebra: for any linear transformation  $f$

$$\dim \text{Ker } f + \dim f(V) = \dim V$$

Thus

$$\dim \text{Ker } f + 1 = n \implies \dim \text{Ker } f = n - 1$$

as required.

- (2) Suppose  $f$  and  $g$  both have the same kernels. Since their kernels have a dimension of  $n-1$ , their shared kernel has a basis of  $v_1, \dots, v_{n-1}$ . We can extend this to a basis of  $V$ ,  $v_1, \dots, v_{n-1}, u$ . Now, we cannot have that  $f(u) = 0$  or  $g(u) = 0$  as linear transformations are determined by their images of a basis, and  $f$  and  $g$  are non-zero. So let us define  $c = \frac{f(u)}{g(u)}$  (recall these are real numbers, so we can divide). Then for every  $w \in V$ , we can write it as a linear combination

$$w = \sum_{i=1}^{n-1} w^i v_i + w^n u$$

And so

$$f(w) = w^n f(u) = w^n c g(u)$$

And

$$g(w) = w^n g(u)$$

Thus we have that for every  $w \in V$ ,  $g(w) = c f(w)$ , as required.

This answer was quite explicit, but the solution should be obvious to anyone who has taken linear algebra, and should not require this level of explanation.

#### Exercise 1.4.3:

Let  $V$  be a vector space with a basis  $e_1, \dots, e_n$  whose dual basis is  $\alpha^1, \dots, \alpha^n$ . Show that

$$S = \{ \alpha^{i_1} \oplus \dots \oplus \alpha^{i_k} \mid 1 \leq i_1, \dots, i_k \leq n \}$$

is a basis for  $L_k(V)$ . In particular, this means that  $\dim L_k(V) = n^k$ .

Let us reuse the notation

$$\alpha^I = \alpha^{i_1} \oplus \dots \oplus \alpha^{i_k}, \quad I = (i_1, \dots, i_k)$$

Suppose  $I$  and  $J$  are two multi-indexes of length  $k$ , then

$$\alpha^I(e_J) = \alpha^{i_1}(e_{j_1}) \dots \alpha^{i_k}(e_{j_k})$$

and so if  $I = J$  then  $\alpha^{i_\ell}(e_{j_\ell}) = \alpha^{i_\ell}(e_{i_\ell}) = 1$  for every  $\ell$ , meaning  $\alpha^I(e_J) = 1$ . If  $I \neq J$  then there exists an index  $j_\ell \in J$  which is not in  $I$ , which means  $\alpha^{i_\ell}(e_{j_\ell}) = 0$  and so  $\alpha^I(e_J) = 0$ . Thus we have

$$\alpha^I(e_J) = \delta_J^I$$

Let us now show that  $S$  is linearly independent. So suppose

$$0 = \sum_{I \in [n]^k} c_I \alpha^I$$

where the sum is over all  $k$ -length multi-indexes. Then for every multi-index  $J$ , we have

$$0 = \sum_I c_I \alpha^I(e_J) = \sum_I c_I \delta_J^I = c_J$$

so for every  $J$ ,  $c_J = 0$ , so  $S$  is indeed linearly independent.

Now we will show that  $S$  spans  $L_k(V)$ . Suppose  $f \in L_k(V)$ , then we claim

$$f = \sum_I f(e_I) \alpha^I$$

Since  $f$  is determined entirely by its values on  $e_I$ , and we have

$$\left( \sum_I f(e_I) \alpha^I \right)(e_J) = \sum_I f(e_I) \alpha^I(e_J) = \sum_I f(e_I) \delta_J^I = f(e_J)$$

there is indeed equality.

#### Exercise 1.4.4:

Let  $f$  be a  $k$ -tensor, prove that  $f$  is alternating if and only if  $f$  changes signs whenever successive arguments are swapped:

$$f(v_1, \dots, v_i, v_{i+1}, \dots, v_k) = -f(v_1, \dots, v_{i+1}, v_i, \dots, v_k)$$

See **proposition 1.3.4** (note that permutations can be written as the product of transpositions of successive indexes).

#### Exercise 1.4.5:

Let  $f$  be a  $k$ -tensor, then  $f$  is alternating if and only if  $f(v_1, \dots, v_k) = 0$  whenever there exist  $i \neq j$  such that  $v_i = v_j$ .

See **proposition 1.3.4**.

#### Exercise 1.4.6:

Let  $V$  be a vector space,  $f$  a  $k$ -covector, and  $g$  an  $\ell$ -covector. Show that for scalars  $a, b \in \mathbb{R}$ ,  $(af) \wedge (bg) = (ab)f \wedge g$ .

This is simple. It is sufficient to show that  $A((af) \otimes (bg)) = abA(f \otimes g)$ . We know  $(af) \otimes (bg) = ab(f \otimes g)$ , so it is sufficient to show that  $A(af) = aA(f)$  for functions  $f$ . This is pretty immediate:

$$A(af) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma(af) = a \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma f = aA(f)$$

#### Exercise 1.4.7:

Suppose two sets of covectors on a vector space  $V$ :  $\beta^1, \dots, \beta^k$  and  $\gamma^1, \dots, \gamma^k$  are related by

$$\beta^i = \sum_{j=1}^k a_{jk}^i \gamma^j$$

for  $i = 1, \dots, k$ . Show then that

$$\beta^1 \wedge \dots \wedge \beta^k = \det(a_{jk}^i) \gamma^1 \wedge \dots \wedge \gamma^k$$

We will show this directly

$$\beta^1 \wedge \dots \wedge \beta^k = \left( \sum_{j_1=1}^k a_{j_1}^1 \gamma^{j_1} \right) \wedge \dots \wedge \left( \sum_{j_k=1}^k a_{j_k}^k \gamma^{j_k} \right) = \sum_{j_1, \dots, j_k=1}^k a_{j_1}^1 \dots a_{j_k}^k \gamma^{j_1} \wedge \dots \wedge \gamma^{j_k}$$

If any  $j_\ell = j_m$ , then  $\gamma^{j_1} \wedge \dots \wedge \gamma^{j_k} = 0$  and so we can sum over unique  $j_1, \dots, j_k$  which is equivalent to summing over  $S_k$ :

$$\sum_{\sigma \in S_k} a_{\sigma(1)}^1 \dots a_{\sigma(k)}^k \gamma^{\sigma(1)} \wedge \dots \wedge \gamma^{\sigma(k)}$$

Now, we know that for transpositions of successive indexes

$$\gamma^{\tau(1)} \wedge \dots \wedge \gamma^{\tau(k)} = -\gamma^1 \wedge \dots \wedge \gamma^k$$

since  $\gamma^i \wedge \gamma^j = -\gamma^j \wedge \gamma^i$ , so by **proposition 1.3.4** we have

$$\gamma^{\sigma(1)} \wedge \dots \wedge \gamma^{\sigma(k)} = \text{sgn}(\sigma) \gamma^1 \wedge \dots \wedge \gamma^k$$

And so we have that

$$\beta^1 \wedge \dots \wedge \beta^k = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \cdot \prod_{i=1}^k a_{\sigma(i)}^i \gamma^1 \wedge \dots \wedge \gamma^k = \det(a_j^i)_{ij} \cdot \gamma^1 \wedge \dots \wedge \gamma^k$$

as required.

#### Exercise 1.4.8:

Let  $f$  be a  $k$ -covector over  $V$ . Suppose two sets of vectors  $u_1, \dots, u_k$  and  $v_1, \dots, v_k$  in  $V$  are related by

$$u_j = \sum_{i=1}^k a_j^i v_i$$

for  $j = 1, \dots, k$ . Show that

$$f(u_1, \dots, u_k) = \det(a_j^i)_{ij} f(v_1, \dots, v_k)$$

Again we will show this directly

$$f(u_1, \dots, u_k) = f\left(\sum_{i_1=1}^k a_1^{i_1} v_{i_1}, \dots, \sum_{i_k=1}^k a_k^{i_k} v_{i_k}\right) = \sum_{i_1, \dots, i_k=1}^k a_1^{i_1} \dots a_k^{i_k} \cdot f(v_{i_1}, \dots, v_{i_k})$$

Since if  $i_\ell = i_m$  for any  $\ell \neq m$  we get  $f(v_{i_1}, \dots, v_{i_k}) = 0$ , we can sum over distinct  $i_1, \dots, i_k$ . This is equivalent to summing over  $S_k$ :

$$\begin{aligned} &= \sum_{\sigma \in S_k} a_{\sigma(1)}^1 \dots a_{\sigma(k)}^k f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \sum_{\sigma \in S_k} \prod_{i=1}^k a_{\sigma(i)}^i \sigma f(v_1, \dots, v_k) \\ &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{i=1}^k a_{\sigma(i)}^i f(v_1, \dots, v_k) = \det(a_j^i)_{ij} f(v_1, \dots, v_k) \end{aligned}$$

#### Exercise 1.4.9:

Let  $V$  be a vector space of  $n$  dimensions. Prove that if an  $n$ -covector  $\omega$  vanishes on a basis  $e_1, \dots, e_n$  for  $V$ , then  $\omega$  is the zero covector on  $V$ .

Let  $v_1, \dots, v_n \in V$  then there exist coefficients  $a_i^j$  such that

$$v_i = \sum_{j=1}^n a_i^j e_j$$

Since  $e_1, \dots, e_n$  is a basis. By above, we have

$$\omega(v_1, \dots, v_n) = \det(a_j^i)_{ij} \omega(e_1, \dots, e_n) = 0$$

as required.



**Exercise 1.4.10:**

Let  $\alpha^1, \dots, \alpha^k$  be 1-covectors on  $V$ . Show that  $\alpha^1 \wedge \dots \wedge \alpha^k \neq 0$  if and only if  $\alpha^1, \dots, \alpha^n$  are linearly independent in the dual space  $V^*$ .

Suppose  $\alpha^1, \dots, \alpha^k$  are linearly independent in  $V^*$  and so they can be extended to  $\alpha^1, \dots, \alpha^n$  which is a basis for  $V^*$ . Every basis of the dual space is a dual basis for some basis of  $V$ . So suppose  $\alpha^1, \dots, \alpha^n$  is the dual basis of  $e_1, \dots, e_n$  then this means that  $\alpha^i$  is the dual vector of  $e_i$  and so

$$\alpha^1 \wedge \dots \wedge \alpha^k(e_1, \dots, e_k) = 1$$

so  $\alpha^1 \wedge \dots \wedge \alpha^k \neq 0$ .

Now suppose

$$\alpha^1 \wedge \dots \wedge \alpha^k \neq 0$$

So there exist  $v_1, \dots, v_k$  such that

$$\alpha^1 \wedge \dots \wedge \alpha^k(v_1, \dots, v_k) \neq 0$$

$v_1, \dots, v_k$  must be linearly independent as otherwise we could take  $u_1, \dots, u_k$  linearly independent vectors which span a space containing  $v_1, \dots, v_k$ . Then by above the wedge product on  $v_1, \dots, v_k$  would equal  $\det A$  times the wedge product on  $u_1, \dots, u_k$ , where  $A$  is the transition matrix from  $u_1, \dots, u_k$  to  $v_1, \dots, v_k$  which is singular since  $v_1, \dots, v_k$  are linearly dependent and so its determinant is zero, meaning the wedge product on  $v_1, \dots, v_k$  is zero, in contradiction.

So we can extend  $v_1, \dots, v_k$  to a basis of  $V$ ,  $v_1, \dots, v_n$  and if we take  $\alpha^i$  to be the dual covector of  $v_i$  for  $i > k$  relative to this basis. Then we claim

$$\alpha^1 \wedge \dots \wedge \alpha^n(v_1, \dots, v_n) = \alpha^1 \wedge \dots \wedge \alpha^k(v_1, \dots, v_k) \neq 0$$

Let us prove this. Firstly, for  $i > k$  we have  $\alpha^i(v_j) = \delta_j^i$  as we defined them as the dual covectors of  $v_i$  relative to the basis  $v_1, \dots, v_n$ . And for  $j > k$  and  $i \leq k$  we have that  $\alpha^i(v_j) = 0$  since otherwise  $\alpha^i$  would have two linearly independent vectors not in its kernel (since we know  $\alpha^i(v_\ell) \neq 0$  for some  $\ell \leq k$ ). But we know that the kernel of  $\alpha^i$  has a dimension of  $n - 1$ , which contradicts this.

So any  $\sigma \in S_n$  which isn't in  $S_k$  would have  $\sigma(\alpha^1 \otimes \dots \otimes \alpha^k)(v_1, \dots, v_n) = 0$ . Thus we have

$$\begin{aligned} \alpha^1 \wedge \dots \wedge \alpha^n(v_1, \dots, v_n) &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \cdot \alpha^1(v_{\sigma(1)}) \dots \alpha^k(v_{\sigma(k)}) \cdot \alpha^{k+1}(v_{k+1}) \dots \alpha^n(v_n) \\ &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma(\alpha^1 \otimes \dots \otimes \alpha^k)(v_1, \dots, v_k) = \alpha^1 \wedge \dots \wedge \alpha^k(v_1, \dots, v_k) \end{aligned}$$

as required.

So now let us take a basis for  $V$ ,  $B = (e_1, \dots, e_n)$ . And its dual basis  $B^* = (\omega^1, \dots, \omega^n)$ . So there exist  $a_j^i$  such that

$$\alpha^i = \sum_{j=1}^n a_j^i \omega^j$$

And so we get, by a previous question,

$$\alpha^1 \wedge \dots \wedge \alpha^n = \det(a_j^i)_{ij} \omega^1 \wedge \dots \wedge \omega^n$$

Since  $\alpha^1 \wedge \dots \wedge \alpha^n \neq 0$  this means that  $\det(a_j^i)_{ij} \neq 0$ , so the matrix  $(a_j^i)_{ij}$  is invertible. Since the matrix represents the transition matrix from  $\alpha^1, \dots, \alpha^n$  to the basis  $\omega^1, \dots, \omega^n$ , this means that  $\alpha^1, \dots, \alpha^n$  is linearly independent. Therefore  $\alpha^1, \dots, \alpha^k$  is linearly independent, as required.

**Exercise 1.4.11:**

Let  $\alpha$  be a nonzero 1-covector and  $\gamma$  a  $k$ -covector on a finite dimensional vector space  $V$ . Show that  $\alpha \wedge \gamma = 0$  if and only if  $\gamma = \alpha \wedge \beta$  for some  $(k-1)$ -covector  $\beta$  on  $V$ .

Obviously if  $\gamma = \alpha \wedge \beta$ , then  $\alpha \wedge \gamma = \alpha \wedge \alpha \wedge \beta = 0$ . So we must show the converse as well.

Since  $\alpha$  is a nonzero 1-covector, we can extend it to a basis  $\alpha^1, \dots, \alpha^n$  of  $V^*$  (let  $\alpha^1 = \alpha$ ). This must be the dual basis of some basis  $e_1, \dots, e_n$  in  $V$ . This problem is trivial for  $k = n$  as the only  $n$ -covectors are of the form  $c \cdot \alpha^1 \wedge \dots \wedge \alpha^n$  ( $A_n(V)$  has a dimension of one).

Otherwise we know that the ascending wedge products of these covectors forms a basis for  $A_k(V)$ . Thus there exist  $c_I$  where

$$\gamma = \sum_I c_I \alpha^I$$

where the sum is over strictly ascending multi-indexes  $I$  of length  $k$ . This means that

$$\alpha \wedge \gamma = \sum_I c_I \alpha^1 \wedge \alpha^I$$

So if 1 is an element of  $I$ , then this means that  $\alpha^1 \wedge \alpha^I = 0$ . And so we have

$$\alpha \wedge \gamma = \sum_{1 \notin I} c_I \alpha^1 \wedge \alpha^I$$

So let  $J$  be an ascending multi-index of length  $k$  which doesn't include 1, and let us define  $J'$  to be equal to  $J$  where we add a 1 to the beginning. Then  $\alpha^1 \wedge \alpha^I(e_{J'}) = \delta_J^I$  and so

$$\alpha \wedge \gamma(e_{J'}) = \sum_{1 \notin I} c_I \alpha^1 \wedge \alpha^I(e_{J'}) = \sum_{1 \notin I} c_I \delta_J^I = c_J$$

Since  $\alpha \wedge \gamma = 0$ , we have  $c_J = 0$  for every multi-index  $J$  which doesn't have a 1.

Given a multi-index  $I$  which contains 1, let us define  $I'$  to be the rest of the multi-index (if  $I = (1, i_2, \dots, i_k)$  then  $I' = (i_2, \dots, i_k)$ ). Since for multi-indexes  $I$  which don't contain 1,  $c_I = 0$ , we can obtain  $\gamma$  by summing over multi-indexes which contain 1. Thus

$$\gamma = \sum_{1 \in I} c_I \alpha^I = \sum_{1 \in I} c_I \alpha^1 \wedge \alpha^{I'} = \alpha^1 \wedge \sum_{1 \in I} c_I \alpha^{I'}$$

And so if we define

$$\beta = \sum_{1 \in I} c_I \alpha^{I'}$$

we get that

$$\gamma = \alpha \wedge \beta$$

as required.

## 1.5 Differential Forms on $\mathbb{R}^n$

### Definition 1.5.1:

The **cotangent space** of  $\mathbb{R}^n$  at a point  $p$  is defined to be the dual of the tangent space at  $p$ , ie.  $T_p^* \mathbb{R}^n$ . This is the set of all linear functionals  $T_p^* \mathbb{R}^n \rightarrow \mathbb{R}$ . An **covector field** (or a **differential 1-form**) on an open subset  $\mathcal{U}$  of  $\mathbb{R}^n$  is a function

$$\omega: \mathcal{U} \rightarrow \bigcup_{p \in \mathcal{U}} T_p^* \mathbb{R}^n$$

where for every  $p \in \mathcal{U}$ ,  $\omega(p) = \omega_p \in T_p^* \mathbb{R}^n$ . That is,  $\omega_p$  is a linear functional  $T_p \mathbb{R}^n \rightarrow \mathbb{R}$ .

### Definition 1.5.2:

Let  $f \in C^\infty(\mathcal{U})$ , then we define its *differential* to be the differential 1-form  $df$ , defined by, for  $p \in \mathcal{U}$  and  $X_p \in T_p \mathbb{R}^n$ ,

$$df_p(X_p) = X_p f$$

Now, recall that the set  $\left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$  forms a basis for  $T_p \mathbb{R}^n$ . Now, the differential of  $x^i$  satisfies

$$(dx^i)_p \left( \frac{\partial}{\partial x^j} \Big|_p \right) = \frac{\partial}{\partial x^j} x^i \Big|_p = \delta_j^i$$

This means that the dual basis of  $\left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$  is

$$\{(dx^1)_p, \dots, (dx^n)_p\}$$

So this forms a basis for the cotangent space  $T_p^*\mathbb{R}^n$ .

This means that if  $\omega$  is a differential 1-form on  $\mathcal{U}$ , for every  $p \in \mathcal{U}$  there exist coefficients  $a_i(p) \in \mathbb{R}$  such that

$$\omega_p = \sum_{i=1}^n a_i(p)(dx^i)_p$$

So we can write

$$\omega = \sum_{i=1}^n a_i dx^i$$

where  $a_i$  is a function  $\mathcal{U} \rightarrow \mathbb{R}$ . We say that a differential 1-form is  $C^\infty$  if every  $a_i$  is in  $C^\infty(\mathcal{U})$ . Since the differential  $df$  is a differential 1-form, we have that

$$df = \sum_{i=1}^n a_i dx^i$$

And so

$$df\left(\frac{\partial}{\partial x^j}\right) = \sum_{i=1}^n a_i dx^i\left(\frac{\partial}{\partial x^j}\right) = \sum_{i=1}^n a_i \delta_j^i = a_j$$

Now, by definition

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

Which means

$$a_j = \frac{\partial f}{\partial x^j}$$

### Proposition 1.5.3:

If  $f$  is a  $C^\infty(\mathcal{U})$  function, then

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i$$

### Definition 1.5.4:

A differential  $k$ -form  $\omega$  on an open subset  $\mathcal{U}$  of  $\mathbb{R}^n$  is a function

$$\omega: \mathcal{U} \rightarrow \bigcup_{p \in \mathcal{U}} A_k(T_p \mathbb{R}^n), \quad \omega_p \in A_k(T_p \mathbb{R}^n)$$

Since  $A_1(T_p \mathbb{R}^n) = T_p^* \mathbb{R}^n$ , this agrees with the definition of a differential 1-form. Notice that differential 0-forms are just functions from  $\mathcal{U}$  to  $\mathbb{R}$

Suppose  $\mu$  is a differential  $k$ -form, then for every  $p \in \mathbb{R}^n$ , we showed in the previous section that  $A_k(T_p \mathbb{R}^n)$  has a basis

$$\{dx_p^{i_1} \wedge \cdots \wedge dx_p^{i_k} \mid i_1 < \cdots < i_k\} = \{dx_p^I \mid I \text{ is an ascending multi-index of length } k\}$$

And this means that there exist coefficients  $a_I(p)$  where

$$\omega_p = \sum_I a_I(p) dx_p^I$$

Where the sum is over ascending multi-indexes of length  $k$ . And so

$$\omega = \sum_I a_I dx^I$$

where  $a_I$  is a function  $\mathcal{U} \rightarrow \mathbb{R}$ . Similar to before, a differential  $k$ -form is  $C^\infty$  if every function  $a_I$  is in  $C^\infty(\mathcal{U})$ .

**Definition 1.5.5:**

We define the space of all  $C^\infty$  differential  $k$ -forms over  $\mathcal{U}$  as  $\Omega^k(\mathcal{U})$

Since differential 0-forms are functions from  $\mathcal{U}$  to  $\mathbb{R}$ , we have  $\Omega^0(\mathcal{U}) = C^\infty(\mathcal{U})$ . Also, notice that  $\Omega^k(\mathcal{U})$  is a vector space. We can turn it into an algebra by defining the wedge product (or exterior product) of two differential forms. Suppose  $\omega$  is a  $k$ -form, and  $\tau$  is an  $\ell$ -form on an open set  $\mathcal{U}$ , we define their wedge product to be a  $k + \ell$ -form to be

$$(\omega \wedge \tau)_p = \omega_p \wedge \tau_p, \quad p \in \mathcal{U}$$

This means that if

$$\omega = \sum_I a_I dx^I, \quad \tau = \sum_J a_J dx^J$$

then

$$\omega \wedge \tau = \sum_{I,J} a_I b_J dx^I \wedge dx^J$$

Where  $I$  is an ascending  $k$ -multi-index, and  $J$  is an ascending  $\ell$ -multi-index. (This of course needs a proof, but it is pretty simple.) Now, if  $I$  and  $J$  are not disjoint then  $dx^I \wedge dx^J = 0$ , so this sum can be written as

$$\omega \wedge \tau = \sum_{I,J \text{ disjoint}} a_I b_J dx^I \wedge dx^J$$

Thus the wedge product of two  $C^\infty$  forms is also  $C^\infty$ , meaning the wedge product defines a bilinear form

$$\wedge: \Omega^k(\mathcal{U}) \times \Omega^\ell(\mathcal{U}) \longrightarrow \Omega^{k+\ell}(\mathcal{U})$$

We know that the wedge product of multivectors is anticommutative and associative and this carries over to the wedge product of differential forms.

Notice that the wedge product with a 0-form  $f$  is simple:

$$(f \wedge \omega)_p = f(p) \wedge \omega_p = f(p) \omega_p$$

And so  $f \wedge \omega = f\omega$ .

**Note:**

Recall that if  $\{V_k\}_{k=0}^\infty$  is a sequence of disjoint  $\mathbb{F}$ -vector spaces, then their **direct sum**,

$$V = \bigoplus_{k=0}^\infty V_k$$

is defined to be the vector space

$$V = \left\{ \sum_{k=0}^\infty a_k v_k \left| \text{For every } k, v_k \in A_k, \text{ and all but a finite number of the coefficients } a_k \text{ are zero} \right. \right\}$$

**Definition 1.5.6:**

An  $\mathbb{F}$ -algebra  $A$  is **graded** if it can be written as the infinite direct sum

$$A = \bigoplus_{k=0}^\infty A_k$$

where  $A_k$  is a vector space over  $\mathbb{F}$  such that  $A_k A_\ell \subseteq A_{k+\ell}$  (the product of an element in  $A_k$  with an element in  $A_\ell$  is in  $A_{k+\ell}$ ). A graded algebra is **anticommutative** (or **graded-commutative**) if for every  $a_k \in A_k$  and  $a_\ell \in A_\ell$ ,

$$a_k a_\ell = (-1)^{k\ell} a_\ell a_k$$

Nonzero elements of  $A_k$  are called **homogeneous elements of degree  $k$** , and a homomorphism of graded algebras is a homomorphism of algebras which preserves degrees.

For example, for a real vector space  $V$  of dimension  $n$ , we define

$$A_*(V) = \bigoplus_{k=0}^{\infty} A_k(V)$$

this is a graded algebra under the wedge product, called the *Grassman algebra* of multivectors over  $V$  (or the *exterior algebra* over  $V$ ). The Grassman algebra, as we showed in the previous section, is anticommutative, and its homogeneous elements of degree  $k$  are  $k$ -covectors.

Similarly we define

$$\Omega^*(\mathcal{U}) = \bigoplus_{k=0}^{\infty} \Omega^k(\mathcal{U})$$

This is also an anticommutative graded algebra over  $\mathbb{R}$  under the wedge product. Since multiplying a  $C^\infty$   $k$ -form by a  $C^\infty(\mathcal{U})$  function is still a  $C^\infty$   $k$ -form,  $\Omega^k(\mathcal{U})$  is both a real vector space and a module over the ring  $C^\infty(\mathcal{U})$ . And therefore  $\Omega^*(\mathcal{U})$  is also a module over  $C^\infty(\mathcal{U})$ .

Now, if  $\omega$  is a  $C^\infty$  1-form and  $X$  is a  $C^\infty$  vector field on an open set  $\mathcal{U}$ , we define the function  $\omega(X)$ :

$$\omega(X)_p = \omega_p(X_p)$$

So if

$$\omega = \sum a_i dx^i, \quad X = \sum b^j \frac{\partial}{\partial x^j}$$

Then we get

$$\omega(X) = \left( \sum a_i dx^i \right) \left( \sum b^j \frac{\partial}{\partial x^j} \right) = \sum_{i,j} a_i b^j dx^i \left( \frac{\partial}{\partial x^j} \right) = \sum_{i,j} a_i b^j \delta_j^i = \sum_i a_i b^i$$

This shows that since  $a_i$  and  $b^j$  are in  $C^\infty(\mathcal{U})$ ,  $\omega(X)$  is also a  $C^\infty(\mathcal{U})$  function. So a  $C^\infty$  1-form defines a map

$$\mathfrak{X}(\mathcal{U}) \longrightarrow C^\infty(\mathcal{U})$$

This map is linear over the ring  $C^\infty(\mathcal{U})$ . Obviously  $\omega(X + Y) = \omega(X) + \omega(Y)$ , and for  $f \in C^\infty(\mathcal{U})$  and  $X \in \mathfrak{X}(\mathcal{U})$ :

$$(\omega(fX))_p = \omega_p(f(p)X_p) = f(p) \cdot \omega_p(X_p) = (f\omega(X))_p$$

so  $\omega(fX) = f\omega(X)$ , meaning  $\omega$  is indeed linear over  $C^\infty(\mathcal{U})$ .

Similarly, if  $\omega$  is a  $C^\infty$   $k$ -form we can define a  $k$ -linear map

$$\mathfrak{X}(\mathcal{U}) \times \cdots \times \mathfrak{X}(\mathcal{U}) \longrightarrow C^\infty(\mathcal{U}), \quad (X_1, \dots, X_k) \mapsto \omega(X_1, \dots, X_k)$$

Where

$$\omega(X_1, \dots, X_k)_p = \omega_p(X_1|_p, \dots, X_k|_p)$$

#### Definition 1.5.7:

The **exterior derivative** of a  $C^\infty$  differential  $k$ -form on an open set  $\mathcal{U}$  is defined as follows:

- (1) The exterior derivative of a 0-form, which is a function  $f \in C^\infty(\mathcal{U})$  is defined to be its differential, and so

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i$$

- (2) For  $k \geq 1$  if  $\omega = \sum_I a_I dx^I$  is in  $\Omega^k(\mathcal{U})$  then we define its exterior derivative is defined as

$$d\omega = \sum_I da_I \wedge dx^I = \sum_I \left( \sum_j \frac{\partial a_I}{\partial x^j} dx^j \right) \wedge dx^I = \sum_I \sum_j \frac{\partial a_I}{\partial x^j} dx^j \wedge dx^I \in \Omega^{k+1}(\mathcal{U})$$

So for example if  $\omega$  is the 1-form  $fx + gdy$  where  $f$  and  $g$  are  $C^\infty(\mathbb{R}^2)$  functions, then

$$d\omega = df \wedge dx + dg \wedge dy = (f_x dx + f_y dy) \wedge dx + (g_x dx + g_y dy) \wedge dy = f_y dy \wedge dx + g_x dx \wedge dy = (g_x - f_y) dx \wedge dy$$

Notice how the exterior derivative is linear. We can generalize this notion of an exterior derivative to graded algebras.

**Definition 1.5.8:**

Suppose  $A = \bigoplus_{k=0}^{\infty} A_k$  is a graded algebra over some field  $\mathbb{F}$ . An **antiderivation** of  $A$  is a linear map  $D: A \rightarrow A$  such that for every  $a_k \in A^k$  and  $a_\ell \in A^\ell$ :

$$D(a_k a_\ell) = D(a_k) a_\ell + (-1)^k a_k D(a_\ell)$$

If there exists an integer  $m$  such that  $D$  maps  $A^k$  to  $A^{k+m}$  for all  $k$ ,  $D$  is termed an **antiderivation of degree  $m$** . For  $k < 0$ , we view  $A_k = 0$  and so the degree of an antiderivation may be negative.

**Proposition 1.5.9:**

- (1) Exterior derivation, as a map  $d: \Omega^*(\mathcal{U}) \rightarrow \Omega^*(\mathcal{U})$  is an antiderivation of degree 1. Meaning for a  $k$ -form  $\omega$  and an  $\ell$ -form  $\tau$ ,

$$d(\omega \wedge \tau) = (d\omega) \wedge \tau + (-1)^k \omega \wedge (d\tau)$$

- (2) Exterior derivation is nilpotent of degree 2,  $d^2 = 0$ .

- (3) If  $f$  is a  $C^\infty$  0-form and  $X \in \mathfrak{X}(\mathcal{U})$  is a  $C^\infty$  vector field over  $\mathcal{U}$ , then  $(df)(X) = Xf$ .

**Proof:**

- (1) So we are trying to prove

$$d(\omega \wedge \tau) = (d\omega) \wedge \tau + (-1)^k \omega \wedge (d\tau)$$

Both sides of this equation are linear in  $\omega$  and  $\tau$ , and so we can prove this for the case that  $\omega = f dx^I$  and  $\tau = g dx^J$ . We have

$$d(\omega \wedge \tau) = d(f g dx^I \wedge dx^J) = d(fg) \wedge dx^I \wedge dx^J = \sum_i \frac{\partial f g}{\partial x^i} dx^i \wedge dx^I \wedge dx^J$$

Using Leibniz's law for the derivative of products

$$= \sum_i \frac{\partial f}{\partial x^i} g dx^i \wedge dx^I \wedge dx^J + \sum_i f \frac{\partial g}{\partial x^i} dx^i \wedge dx^I \wedge dx^J$$

Now let us expand the right hand side,

$$\begin{aligned} (d\omega) \wedge \tau + (-1)^k \omega \wedge (d\tau) &= df \wedge dx^I \wedge g dx^J + (-1)^k f dx^I \wedge dg \wedge dx^J \\ &= \sum_i \frac{\partial f}{\partial x^i} g dx^i \wedge dx^I \wedge dx^J + (-1)^k \sum_i f \frac{\partial g}{\partial x^i} dx^I \wedge dx^i \wedge dx^J \end{aligned}$$

Now, since  $I$  has a length of  $k$ , we have  $dx^I \wedge dx^i = (-1)^k dx^i \wedge dx^I$  so this is equal to

$$\sum_i \frac{\partial f}{\partial x^i} g dx^i \wedge dx^I \wedge dx^J + \sum_i f \frac{\partial g}{\partial x^i} dx^i \wedge dx^I \wedge dx^J = d(\omega \wedge \tau)$$

as required.

- (2) Again, since  $d$  is a linear operator over  $\Omega^*(\mathcal{U})$ , it is sufficient to show that  $d^2(\omega) = 0$  for  $\omega = f dx^I$ . We know

$$d(d(f dx^I)) = d\left(\sum_i \frac{\partial f}{\partial x^i} dx^i \wedge dx^I\right) = \sum_i d\left(\frac{\partial f}{\partial x^i}\right) \wedge dx^i \wedge dx^I = \sum_i \sum_j \frac{\partial^2 f}{\partial x^i \partial x^j} dx^j \wedge dx^i \wedge dx^I$$

If  $i = j$ , then  $dx^j \wedge dx^i = 0$  and so we can sum over distinct  $i, j$ s. But notice that

$$\frac{\partial^2 f}{\partial x^i \partial x^j} dx^j \wedge dx^i \wedge dx^I = -\frac{\partial^2 f}{\partial x^j \partial x^i} dx^i \wedge dx^j \wedge dx^I$$

And so pairing the sum over  $i, j$  with  $j, i$ , we get zero. So the sum is equal to zero, as required.

- (3) This is the definition of  $df$ . ■

What is significant about these three properties is that they uniquely determine the exterior derivative. What this means

is that instead of giving an explicit definition for the exterior derivative, we could've axiomatized it with these above properties.

**Theorem 1.5.10:**

The exterior derivative is uniquely defined by the three properties from the above proposition. Meaning it is the unique antiderivation of degree 1 on the graded algebra  $\Omega^*(\mathcal{U})$  which is nilpotent of degree 2, such that for every 0-form  $f \in \Omega^0(\mathcal{U})$  and  $X \in \mathfrak{X}(\mathcal{U})$ ,  $(df)(X) = Xf$ .

**Proof:**

The above proposition already showed that the exterior derivative satisfies all these properties. We must now show that an operator satisfying these properties is the exterior derivative. So suppose  $D$  is a degree-2 nilpotent antiderivation of degree 1 on  $\Omega^*(\mathcal{U})$ , where  $(Df)(X) = Xf$  for vector fields  $X$ .

Since  $(Df)(X) = Xf = (df)(X)$  for every vector field  $X$ , we must have that  $Df = df$  for 0-forms  $f \in \Omega^0(\mathcal{U})$ . This is since this is the definition of the differential of  $f$ . Thus this means

$$D(dx^i) = D(Dx^i) = D^2(x^i) = 0$$

Inductively we claim that  $D(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = 0$ . The induction step relies on  $D$  being an antiderivation:

$$D(dx^{i_1} \wedge \cdots \wedge dx^{i_{k+1}}) = D(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \wedge dx^{i_{k+1}} + (-1)^k dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge D(dx^{i_{k+1}})$$

and since  $D(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = 0$  by our inductive assumption and  $D(dx^{i_{k+1}}) = 0$ , we have that this is indeed 0.

Now, since  $D$  is a linear operator over  $\Omega^*(\mathcal{U})$ , it is sufficient to show that  $D(\omega) = d\omega$  for  $\omega = f dx^I$ . Since  $D$  is an antiderivation and  $f$  is a 0-form,

$$D(\omega) = D(f dx^I) = D(f) \wedge dx^I + f \wedge D(dx^I) = df \wedge dx^I = d\omega$$

And so  $D = d$  on  $\Omega^*(\mathcal{U})$ , as required. ■

Notice that we didn't even need to assume that  $D$  is an antiderivation of degree 1, just that it is an antiderivation.

**Definition 1.5.11:**

A  $k$ -form  $\omega$  on  $\mathcal{U}$  is **closed** if  $d\omega = 0$ .  $\omega$  is **exact** if there exists a  $(k-1)$ -form  $\tau$  such that  $\omega = d\tau$ . Since  $d(d\tau) = 0$ , exact forms are closed.

**Definition 1.5.12:**

A sequence of vector spaces  $\{V^k\}_{k=0}^\infty$  equipped with linear maps  $d_k: V^k \longrightarrow V^{k+1}$  such that  $d_{k+1} \circ d_k = 0$  is called a **differential complex** or a **cochain complex**. This is denoted

$$V^0 \xrightarrow{d_0} V^1 \xrightarrow{d_1} V^2 \xrightarrow{d_2} \cdots$$

(Cochain complexes were discussed briefly in group theory.)

Notice that  $\{\Omega^k(\mathcal{U})\}_{k=0}^\infty$  along with the exterior derivative  $d$  forms a cochain complex

$$0 \rightarrow \Omega^0(\mathcal{U}) \xrightarrow{d} \Omega^1(\mathcal{U}) \xrightarrow{d} \Omega^2(\mathcal{U}) \xrightarrow{d} \cdots$$

The closed forms make up precisely the kernel of  $d$ , and the exact forms make up precisely the image of  $d$ .

Let us now demonstrate the usefulness of the exterior derivative, and its relation with commonly used notions from vector calculus. Recall that an *vector-valued function* on an open set  $\mathcal{U} \subseteq \mathbb{R}^3$  is a function

$$\mathbf{F}: \mathcal{U} \longrightarrow \mathbb{R}^3$$

So  $\mathbf{F}$  assigns to every  $p \in \mathcal{U}$  a vector in  $\mathbb{R}^3 \cong T_p \mathbb{R}^3$ . So we can view  $\mathbf{F}$  as a vector field on  $\mathcal{U}$ . We write the function  $\mathbf{F}$  in terms of its coordinates:

$$\mathbf{F} = \begin{pmatrix} P \\ Q \\ R \end{pmatrix}$$

Now, recall the three operators on  $C^\infty$  vector-valued functions: the gradient, curl, and divergence operators. These are operators

$$\{\text{scalar functions}\} \xrightarrow{\text{grad}} \{\text{vector functions}\} \xrightarrow{\text{curl}} \{\text{vector functions}\} \xrightarrow{\text{div}} \{\text{scalar functions}\}$$

These are defined by

$$\begin{aligned} \text{grad} f &= \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} f = \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} \\ \text{curl} \begin{pmatrix} P \\ Q \\ R \end{pmatrix} &= \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} P \\ Q \\ R \end{pmatrix} = \begin{pmatrix} R_y - Q_z \\ -(R_x - P_z) \\ Q_x - P_y \end{pmatrix} \\ \text{div} \begin{pmatrix} P \\ Q \\ R \end{pmatrix} &= \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} P \\ Q \\ R \end{pmatrix} = P_x + Q_y + R_z \end{aligned}$$

1-forms on  $\mathcal{U}$  are linear combinations of  $dx$ ,  $dy$ , and  $dz$ , and so we can identify 1-forms with vector fields on  $\mathcal{U}$  via

$$Pdx + Qdy + Rdz \longleftrightarrow \begin{pmatrix} P \\ Q \\ R \end{pmatrix}$$

Similarly 2-forms on  $\mathcal{U}$  can also be identified with vector fields by

$$Pdy \wedge dz + Qdz \wedge dx + Rdx \wedge dy \longleftrightarrow \begin{pmatrix} P \\ Q \\ R \end{pmatrix}$$

And 3-forms can be identified with scalar functions

$$f dx \wedge dy \wedge dz \longleftrightarrow f$$

The exterior derivative of a 0-form is

$$d(f) = f_x dx + f_y dy + f_z dz \longleftrightarrow \begin{pmatrix} \partial f / \partial x \\ \partial f / \partial y \\ \partial f / \partial z \end{pmatrix} = \text{grad} f$$

So the exterior derivative of a 0-form, which are scalar functions, corresponds to the gradient of the scalar function. The exterior derivative of a 1-form is

$$d(Pdx + Qdy + Rdz) = (R_y - Q_z)dy \wedge dz - (R_x - P_z)dz \wedge dx + (Q_x - P_y)dx \wedge dy$$

which corresponds to the curl operator. And the exterior derivative of a 2-form is

$$d(Pdy \wedge dz + Qdz \wedge dx + Rdx \wedge dy) = (P_x + R_y + Q_z)dx \wedge dy \wedge dz$$

which corresponds to the divergence operator.

So we have shown the identifications

$$\begin{array}{ccccccc} \Omega^0(\mathcal{U}) & \xrightarrow{d} & \Omega^1(\mathcal{U}) & \xrightarrow{d} & \Omega^2(\mathcal{U}) & \xrightarrow{d} & \Omega^3(\mathcal{U}) \\ \uparrow \cong & & \uparrow \cong & & \uparrow \cong & & \uparrow \cong \\ C^\infty(\mathcal{U}) & \xrightarrow{\text{grad}} & \mathfrak{X}(\mathcal{U}) & \xrightarrow{\text{curl}} & \mathfrak{X}(\mathcal{U}) & \xrightarrow{\text{div}} & C^\infty(\mathcal{U}) \end{array}$$

Since the exterior derivative is nilpotent of degree two, we have shown two basic facts about the gradient, curl, and divergence operators:

$$\text{curl}(\text{grad} f) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{div} \left( \text{curl} \begin{pmatrix} P \\ Q \\ R \end{pmatrix} \right) = 0$$

Now it turns out that in  $\mathbb{R}^n$ , closed and exact forms are the same. So a vector function  $\mathbf{F}$  is the gradient of a scalar function if and only if its curl is zero. This is because  $\mathbf{F}$  being the gradient of a scalar function means it is an exact 1-form in the identification above, which is if and only if it is closed (curl zero).



# Index

Algebra:  $3^*$   
Alternating function:  $7^*$   
Antiderivation:  $21^*$   
C-infty: 19, 19  
Closed differential form:  $23^*$   
Cochain complex:  $23^*$   
Cotangent space:  $18^*$   
Covector field:  $18^*$   
Covectors:  $8^*$   
Derivation:  $6^*$   
Differential: 18  
Differential complex:  $23^*$   
Differential form:  $18^*$ ,  $19^*$   
Direct sum:  $20^*$   
Directional derivative: 4  
Dual basis:  $6^*$   
Dual space:  $6^*$   
Exact differential form:  $23^*$   
Exterior algebra:  
    See also: Grassman algebra  
Exterior derivative:  $21^*$   
Exterior product:  $9^*$   
    See also: Wedge product  
Germ:  $3^*$   
Graded algebra:  $20^*$   
Graded-commutative:  $20^*$   
Grassman algebra: 21  
Multicovectors:  $8^*$   
Multilinear function:  $6^*$   
Point derivation:  $4^*$   
Star-shaped:  $2^*$   
Symmetric function:  $7^*$   
Tangent space:  $3^*$   
Tangent vector:  $3^*$   
Taylor's theorem with remainder: 2  
Tensor product:  $9^*$   
Tensors:  $8^*$   
Vector field:  $5^*$   
Wedge product:  $9^*$ , 20