PMATH365 — Differential Geometry

Classnotes for Winter 2019

by

Johnson Ng

BMath (Hons), Pure Mathematics major, Actuarial Science Minor

University of Waterloo

Table of Contents

List of Definitions

List of Theorems

Preface

This course is a post-requisite of MATH 235/245 (Linear Algebra II) and AMATH 231 (Calculus IV) or MATH 247 (Advanced Calculus III). In other words, familiarity with vector spaces and calculus is expected.

The course is spiritually separated into two parts. The first part shall be called **Exterior Differential Calculus**, which allows for a natural, metric-independent generalization of **Stokes' Theorem**, **Gauss's Theorem**, and **Green's Theorem**. Our end goal of this part is to arrive at Stokes' Theorem, that renders the **Fundamental Theorem of Calculus** as a special case of the theorem.

The second part of the course shall be called in the name of the course: **Differential Geometry**. This part is dedicated to studying geometry using techniques from differential calculus, integral calculus, linear algebra, and multilinear algebra.

Part I

Exterior Differential Calculus

1 Lecture 1 Jan 07th

1.1 Linear Algebra Review

Definition 1 (Linear Map)

Let V, W be finite dimensional real vector spaces. A map $T : V \to W$ is called ***linear*** *if $\forall a, b \in \mathbb{R}$, $\forall v \in V$ and $\forall w \in W$,*

$$T(av + bw) = aT(v) + bT(w).$$

We define L(U, W) to be the set of all linear maps from V to W.

Note

- *Note that $L(U, W)$ is itself a finite dimensional real vector space.*
- *The structure of the vector space $L(V, W)$ is such that $\forall T, S \in L(V, W)$, and $\forall a, b \in \mathbb{R}$, we have*

  $$aT + bS : V \to W$$

  and

  $$(aT + bS)(v) = aT(v) + bS(v).$$

- *A special case: when $W = V$, we usually write*

  $$L(V, W) = L(V),$$

  and we call this the ***space of linear operators on V***.

Now suppose $\dim(V) = n$ for some $n \in \mathbb{N}$. This means that there exists a basis $\{e_1, \ldots, e_n\}$ of V with n elements.

Definition 2 (Basis)

A basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ of an n-dimensional vector space V is a subset of V where

1. *$\mathcal{B}$ **spans** V, i.e. $\forall v \in V$*

$$v = \sum_{i=1}^{n} v^i e_i.$$ [1]

2. *$e_1, \ldots, e_n$ are linearly independent, i.e.*

$$v^i e_i = 0 \implies v^i = 0 \text{ for every } i.$$

[1] We shall use a different convention when we write a linear combination. In particular, we use v^i to represent the i^{th} coefficient of the linear combination instead of v_i. Note that this should not be confused with taking powers, and should be clear from the context of the discussion.

Note

We shall abusively write

$$v^i e_i = \sum_i v^i e_i.$$

Again, this should be clear from the context of the discussion.

The two conditions that define a basis implies that any $v \in V$ can be expressed as $v^i e_i$, where $v^i \in \mathbb{R}$.

Definition 3 (Coordinate Vector)

*The n-tuple $(v^1, \ldots, v^n) \in \mathbb{R}^n$ is called the **coordinate vector** $[v]_{\mathcal{B}} \in \mathbb{R}^n$ of v with respect to the basis $\mathcal{B} = \{e_1, \ldots, e_n\}$.*

Note

It is clear that the coordinate vector $[v]_{\mathcal{B}}$ is dependent on the basis $\mathcal{B}$. Note that we shall also assume that the basis is "ordered", which is somewhat important since the same basis (set-wise) with a different "ordering" may give us a completely different coordinate vector.

Example 1.1.1

Let $V = \mathbb{R}^n$, and $\hat{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, where 1 is the i^{th} compoenent of $\hat{e}_1$. Then

$$\mathcal{B}_{\text{std}} = \{\hat{e}_1, \ldots, \hat{e}_n\}$$

is called the **standard basis** of $\mathbb{R}^n$.

Note

Let $v = (v^1, \ldots, v^n) \in \mathbb{R}^n$. Then

$$v = v^1 \hat{e}_1 + \ldots v^n \hat{e}_n.$$

So $\mathbb{R}^n \ni [v]_{\mathcal{B}_{\text{std}}} = v \in V = \mathbb{R}^n$.

This is a privilege enjoyed by the n-dimensional vector space $\mathbb{R}^n$.

Now if we choose a **non-standard basis** of $\mathbb{R}^n$, say $\tilde{\mathcal{B}}$, then $[v]_{\tilde{\mathcal{B}}} \neq v$.

Note

It does not make sense to ask if a standard basis exists for an arbitrary space, as we have seen above. A geometrical way of wrestling with this notion is as follows:

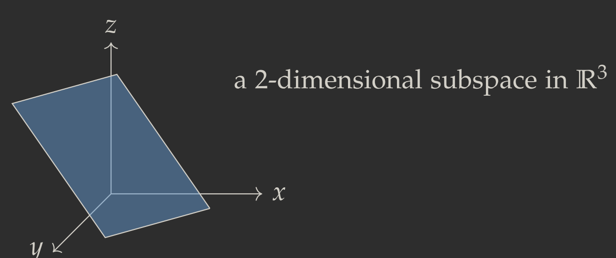

Figure 1.1: An arbitrary 2-dimensional subspace in a 3-dimensional space

While the subspace is embedding in a vector space of which has a standard basis, we cannot establish a "standard" basis for this 2-dimensional subspace. In laymen terms, we cannot tell which direction is up or down, positive or negative for the subspace, without making assumptions.

However, since we are still in a finite-dimensional vector space, we can still make a connection to a Euclidean space of the same dimension.

Definition 4 (Linear Isomorphism)

Let V be n-dimensional, and $\mathcal{B} = \{e_1, \ldots, e_n\}$ be some basis of V. The map

$$v = v^i e_i \mapsto [v]_{\mathcal{B}}$$

*from V to $\mathbb{R}^n$ is a **linear isomorphism** of vector spaces.*

Exercise 1.1.1

Prove that the said linear isomorphism is indeed linear and bijective[2].

[2] i.e. we are right in calling it linear and being an isomorphism

Note

*Any n-dimensional real vecotr space is isomorphic to $\mathbb{R}^n$, but not **canonically** so, as it requires the knowledge of the basis that is arbitrarily chosen. In other words, a different set of basis would give us a different isomorphism.*

1.2 Orientation

Consider an n-dimensional vector space V. Recall that for any linear operator $T \in L(V)$, we may associate a real number $\det(T)$, called the **determinant** of T, such that T is said to be **invertible** iff $\det(T) \neq 0$.

Definition 5 (Same and Opposite Orientations)

Let

$$\mathcal{B} = \{e_1, \ldots, e_n\} \quad \text{and} \quad \tilde{\mathcal{B}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$$

be two ordered bases of V. Let $T \in L(V)$ be the linear operator defined by

$$T(e_i) = \tilde{e}_i$$

for each $i = 1, 2, \ldots, n$. This mapping is clearly invertible, and so

$\det(T) \neq 0$, and T^{-1} is also linear, such that $T^{-1}(\tilde{e}_i) = e_i$, for each i.

We say that $\mathcal{B}$ and $\tilde{\mathcal{B}}$ determine the ***same orientation*** *if $\det(T) > 0$, and we say that they determine the* ***opposite orientations*** *if $\det(T) < 0$.*

Note

- *This notion of orientation only works in real vector spaces, as, for instance, in a complex vector space, there is no sense of "positivity" or "negativity".*
- *Whenever we talk about same and opposite orientation(s), we are usually talking about 2 sets of bases. It makes sense to make a comparison to the standard basis in a Euclidean space, and determine that the compared (non-)standard basis is "positive" (same direction) or "negative" (opposite), but, again, in an arbitrary space, we do not have this convenience.*

Exercise 1.2.1 (A1Q1)

Show that any n-dimensional real vector space V admits exactly 2 orientations.

Example 1.2.1

On $\mathbb{R}^n$, consider the standard basis

$$\mathcal{B}_{\text{std}} = \{\hat{e}_1, \ldots, \hat{e}_n\}.$$

The orientation determined by $\mathcal{B}_{\text{std}}$ is called the **standard orientation** of $\mathbb{R}^n$.

1.3 Dual Space

Definition 6 (Dual Space)

Let V be an n-dimensional vector space. Then $\mathbb{R}$ is a 1-dimensional real vector space. Thus we have that $L(V, \mathbb{R})$ is also a real vector space [3]. The

[3] Note that $L(V, \mathbb{R})$ is also finite dimensional since both the domain and codomain are finite dimensional.

dual space *V^* of V is defined to be*

$$V^* := L(V, \mathbb{R}).$$

Let $\mathcal{B}$ be a basis of V. For all $i = 1, 2, \ldots, n$, let $e^i \in V^*$ such that

$$e^i(e_j) = \delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

This δ^i_j is known as the **Kronecker Delta**.

In general, we have that for every $v = v^j e_j \in V$, where $v^i \in \mathbb{R}$, by the linearity of e^i, we have

$$e^i(v) = e^i(v^j e_j) = v^j e^i(e_j) = v_j \delta^i_j = v^i.$$

So each of the e^i, when applied on v, gives us the i^{th} component of $[v]_{\mathcal{B}}$, where $\mathcal{B}$ is a basis of V, in particular

$$v = v^i e_i, \text{ where } v^i = e^i(v). \tag{1.1}$$

2 Lecture 2 Jan 09th

2.1 Dual Space (Continued)

Proposition 1 (Dual Basis)

The set

$$\mathcal{B}^* := \left\{ e^1, \ldots, e^n \right\}$$

[1] *is a basis of* V^**, and is called the* ***dual basis*** *of* $\mathcal{B}$*, where* $\mathcal{B}$ *is a basis of* V*. In particular,* $\dim V^* = n = \dim V$.

[1] Note that the e^i's are defined as in the last part of the last lecture.

Proof

$\mathcal{B}^*$ spans V^* Let $\alpha \in V^*$. Let $v = v^j e_j \in V$, where we note that

$$\mathcal{B} = \{e_i\}_{i=1}^n .$$

We have that

$$\alpha(v) = \alpha(v^j e_j) = v^j \alpha(e_j).$$

Now for all $j = 1, 2, \ldots, n$, define $\alpha_j = \alpha(e_j)$. Then

$$\alpha(v) = \alpha_j v^j = \alpha_j e^j(v),$$

which holds for all $v \in V$. This implies that $\alpha = \alpha_j e^j$, and so $\mathcal{B}^*$ spans V^*.

$\mathcal{B}^*$ is linearly independent Suppose $\alpha_j e^j = 0 \in V^*$. Applying $\alpha_j e^j$ to each of the vectors e_k in $\mathcal{B}$, we have

$$\alpha_j e^j(e_k) = 0(e_k) = 0 \in \mathbb{R}$$

and

$$\alpha_j e^j(e_k) = \alpha_j \delta_k^j = \alpha_k.$$

By A1Q2, we have that $a_k = 0$ for all $k = 1, 2, \ldots, n$, and so $\mathcal{B}^*$ is linearly independent. □

Remark

Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis of V, with dual space $\mathcal{B}^* = \{e^1, \ldots, e^n\}$. Then the map $T : V \to V^*$ such that

$$T(e_i) = e^i$$

*is a vector space isomorphism. And so we have that $V \simeq V^*$, but not **cannonically** so since we needed to know what the basis is in the first place.*

We will see later that if we impose an **inner product** on V, then it will induce a canonical isomorphism from V to V^*.

Definition 7 (Natural Pairing)

The function

$$\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R}$$

given by

$$\langle \alpha, v \rangle \mapsto \alpha(v)$$

*is called a **natural pairing** of V^* and V.*

Note

A natural pairing is bilinear, i.e. it is linear in α and linear in v, which means that

$$\langle \alpha, t_1 v_1 + t_2 v_2 \rangle = t_1 \langle \alpha, v_1 \rangle + t_2 \langle \alpha, v_2 \rangle$$

and

$$\langle t_1 \alpha_1 + t_2 \alpha_2, v \rangle = t_1 \langle \alpha_1, v \rangle + t_2 \langle \alpha_2, v \rangle,$$

respectively.

Proposition 2 (Natural Pairings are Nondegenerate)

For a finite dimensional real vector space V, a natural pairing is said to be nondegenerate if

This is A1Q2.

$$\forall v \in V \ \ \langle \alpha, v \rangle = 0 \iff \alpha = 0$$

and

$$\forall \alpha \in V^* \ \ \langle \alpha, v \rangle = 0 \iff v = 0.$$

Example 2.1.1

Fix a basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ of V. Given $T \in L(V)$, there is an associated $n \times n$ matrix $A = [T]_{\mathcal{B}}$ defined by

$$T(e_i) = A_{\underset{\text{row index}}{i}}^{\overset{\text{column index}}{j}} e_j.$$

In particular,

$$A = \overbrace{\begin{bmatrix} [T(e_1)]_{\mathcal{B}} & \ldots & [T(e_n)]_{\mathcal{B}} \end{bmatrix}}^{\text{block matrix}}$$

and

$$A_i^k = e^k(T(e_i)) = \langle e^k, T(e_i) \rangle.$$

Definition 8 (Double Dual Space)

The set

$$V^{**} = L(V^*, \mathbb{R})$$

is called the ***double dual space****.*

Proposition 3 (The Space and Its Double Dual Space)

Let V be a finite dimensional real vector space and V^{**} be its double dual space. There exists a linear map ξ such that

$$\xi : V \to V^{**}$$

Proof

Let $v \in V$. Then $\xi(v) \in V^{**} = L(V^*, \mathbb{R})$, i.e. $\xi(v) : V^* \to \mathbb{R}$. Then

As messy as this may seem, this is really a follow your nose kind of proof.
Since we are proving that a map exists, we need to construct it. Since $\xi : V \to V^{**} = L(V^*, \mathbb{R})$, for any $v \in V$, we must have $\xi(v)$ as some linear map from V^* to $\mathbb{R}$.

for any $\alpha \in V^*$,

$$(\xi(v))\,(\alpha) \in \mathbb{R}.$$

Since $\alpha \in V^*$, i.e. $\alpha : V \to \mathbb{R}$, and α is linear, let us define

$$\xi(v)(\alpha) = \alpha(v).$$

To verify that $\xi(v)$ is indeed linear, notice that for any $t, s \in \mathbb{R}$, and for any $\alpha, \beta \in V^*$, we have

$$\begin{aligned} \xi(v)(t\alpha + s\beta) &= (t\alpha + s\beta)(v) \\ &= t\alpha(v) + s\beta(v) \\ &= t\xi(v)(\alpha) + s\xi(v)(\beta). \end{aligned}$$

It remains to show that ξ itself is linear: for any $t, s \in \mathbb{R}$, any $v, w \in V$, and any $\alpha \in V^*$, we have

$$\begin{aligned} \xi(tv + sw)(\alpha) = \alpha(tv + sw) &= t\alpha(v) + s\alpha(w) \\ &= t\xi(v)(\alpha) + s\xi(v)(\alpha) \\ &= [t\xi(v) + s\xi(w)](\alpha) \end{aligned}$$

by addition of functions. □

Proposition 4 (Isomorphism Between The Space and Its Dual Space)

The linear map in Proposition 3 is an isomorphism.

Proof

From Proposition 3, ξ is linear. Let $v \in V$ such that $\xi(v) = 0$, i.e. $v \in \ker(\xi)$. Then by the same definition of ξ as above, we have

$$0 = (\xi(v))(\alpha) = \alpha(v)$$

for any $\alpha \in V^*$. By Proposition 2, we must have that $v = 0$, i.e. $\ker(\xi) = \{0\}$. Thus by Proposition A.2, ξ is injective.

Now, since

$$V^{**} = L(V^*, \mathbb{R}) = L(L(V, \mathbb{R}), \mathbb{R}),$$

we have that

$$\dim(V^{**}) = \dim(V^*) = \dim(V).$$

Thus, by the Rank-Nullity Theorem [2], we have that ξ is surjective.

[2] See Appendix A.1, and especially Proposition A.3.

□

The above two proposition shows to use that we may identify V with V^{**} using ξ, and we can gleefully assume that $V = V^{**}$.

Consequently, if $v \in V = V^{**}$ and $\alpha \in V^*$, we have

$$\alpha(v) = v(\alpha) = \langle \alpha, v \rangle. \tag{2.1}$$

2.2 Dual Map

Definition 9 (Dual Map)

Let $T \in L(V, W)$, where V, W are finite dimensional real vector spaces. Let

$$T^* : W^* \to V^*$$

be defined as follows: for $\beta \in W^*$, we have $T^*(\beta) \in V^*$. Let $v \in V$, and so $(T^*(\beta))(v) \in \mathbb{R}$ [3]. From here, we may define

$$(T^*(\beta))(v) = \beta(T(v)).$$

The map T^* is called ***the dual map***.

[3] It shall be verified here that $T^*(\beta)$ is indeed linear: let $v_1, v_2 \in V$ and $c_1, c_2 \in \mathbb{R}$. Indeed

$$\begin{aligned} &T^*(\beta)(c_1 v_1 + c_2 v_2) \\ &\quad = c_1 T^*(\beta)(v_1) + c_2 T^*(\beta)(v_2) \end{aligned}$$

Exercise 2.2.1

Prove that $T^* \in L(W^*, V^*)$, i.e. that T^* is linear.

Proof

Let $\beta_1, \beta_2 \in W^*$, $t_1, t_2 \in \mathbb{R}$, and $v \in V$. Then

$$\begin{aligned} T^*(t_1\beta_1 + t_2\beta_2)(v) &= (t_1\beta_1 + t_2\beta_2)(Tv) \\ &= t_1\beta_1(Tv) + t_2\beta_2(Tv) \\ &= t_1 T^*(\beta_1)(v) + t_2 T^*(\beta_2)(v). \end{aligned}$$

□

Note

Note that in Definition 9, our construction of T^* is canonical, i.e. its construction is independent of the choice of a basis.

Also, notice that in the language of pairings, we have

$$\langle T^* \beta, v \rangle = (T^*(\beta))(v) = \beta(T(v)) = \langle \beta, T(v) \rangle,$$

where we note that

$$\begin{gathered} T^*(\beta) \in V^* \quad v \in V \\ \beta \in W^* \quad T(v) \in W. \end{gathered}$$

3 Lecture 3 Jan 11th

3.1 Dual Map (Continued)

Note

*Elements in V^* are also called **co-vectors**.*

Recall from last lecture that if $T \in L(V, W)$, then it induces a dual map $T^* \in L(W^*, V^*)$ such that

$$(T^*\beta)(v) = \beta(T(v)).$$

Proposition 5 (Identity and Composition of the Dual Map)

Let V and W be finite dimensional real vector spaces.

1. *Supppose $V = W$ and $T = I_V \in L(V)$, then*

$$(I_V)^* = I_{V^*} \in L(V^*).$$

2. *Let $T \in L(V, W)$, $S \in L(W, U)$. Then $S \circ T \in L(V, U)$. Moreover,*

$$L(U^*, V^*) \ni (S \circ T)^* = T^* \circ S^*.$$

Proof

1. Observe that for any $\beta \in V^*$, and any $v \in V$, we have

$$((I_V)^*(\beta))(v) = \beta((I_V)(v)) = \beta(v).$$

Therefore $(I_V)^* = I_{V^*}$.

2. Observe that for $\gamma \in U^*$ and $v \in V$, we have

$$\begin{aligned}((S \circ T)^*(\gamma))(v) &= \gamma((S \circ T)(v)) \\ &= \gamma(S(T(v))) \\ &= S^*(\gamma T(v)) \\ &= (T^* \circ S^*)(\gamma)(v),\end{aligned}$$

and so $(S \circ T)^* = T^* \circ S^*$ as required.

$\square$

Let $T \in L(V)$, and the dual map $T^* \in L(V^*)$. Let $\mathcal{B}$ be a basis of V, with the dual basis $\mathcal{B}^*$. We may write

$$A = [T]_{\mathcal{B}} \text{ and } A^* = [T^*]_{\mathcal{B}^*}.$$

Note that

$$T(e_i) = A_i^j e_j \text{ and } T^*(e^i) = (A^*)_j^i e^j.$$

Consequently, we have

$$\langle e^k, T(e_i) \rangle = A_i^k \text{ and } \langle T^*(e^i), e_k \rangle = (A^*)_k^i.$$

From here, notice that by applying $e_k \in V = V^{**}$ to both sides, we have

$$(A^*)_k^i = e_k(T^*(e^i)) = \langle T^*(e^i), e_k \rangle \overset{(*)}{=} \langle e^i, T(e_k) \rangle = A_k^i.$$

Thus A^* is the transpose of A, and

$$[T^*]_{\mathcal{B}^*} = [T]_{\mathcal{B}}^t \tag{3.1}$$

where M^t is the transpose of the matrix M.

3.1.1 *Application to Orientations*

Let $\mathcal{B}$ be a basis of V. Then $\mathcal{B}$ determines an orientation of V. Let $\mathcal{B}^*$ be the dual basis of V^*. So $\mathcal{B}^*$ determines an orientation for V^*.

Example 3.1.1

Suppose $\mathcal{B}$ and $\tilde{\mathcal{B}}$ determines the same orientation of V. Does it follow that the dual bases $\mathcal{B}^*$ and $\tilde{\mathcal{B}}^*$ determine the same orientation of V^*?

Proof

Let

$$\mathcal{B} = \{e_1, \ldots, e_n\} \qquad \tilde{\mathcal{B}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$$
$$\mathcal{B}^* = \left\{e^1, \ldots, e^n\right\} \qquad \tilde{\mathcal{B}}^* = \left\{\tilde{e}^1, \ldots, \tilde{e}^n\right\}$$

Let $T \in L(V)$ such that $T(e_i) = \tilde{e}_i$. By assumption, $\det T > 0$. Notice that

$$\delta^i_j = \tilde{e}^i(\tilde{e}_j) = \tilde{e}^i(Te_j) = (T^*(\tilde{e}^i))(e_j),$$

and so we must have $T^*(\tilde{e}^i) = e^i$. By Equation (3.1), we have that

$$\det T^* = \det T > 0$$

as well. This shows that $\mathcal{B}^*$ and $\tilde{\mathcal{B}}^*$ determines the same orientation.

3.2 The Space of k-forms on V

Definition 10 (k-Form)

*Let V be an ndimensional vector space. Let $k \geq 1$. A **k-form** on V is a map*

$$\alpha : \underbrace{V \times V \times \ldots \times V}_{k \text{ times}} \to \mathbb{R}$$

such that

1. *(**k-linearity / multi-linearity**) if we fix all but one of the arguments of α, then it is a linear map from V to $\mathbb{R}$; i.e. if we fix*

   $$v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k \in V,$$

   then the map

   $$u \mapsto \alpha(v_1, \ldots, v_{j-1}, u, v_{j+1}, \ldots, v_k)$$

   is linear in u.

2. *(**alternating property**) α is **alternating** (aka **totally skewed-symmetric**) in its k arguments; i.e.*

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = \alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k).$$

Example 3.2.1

The following is an example of the second condition: if $k = 2$, then $\alpha : V \times V \to \mathbb{R}$. Then $\alpha(v, w) = -\alpha(w, v)$.

If $k = 3$, then $\alpha : V \times V \times V \to \mathbb{R}$. Then we have

$$\begin{aligned} \alpha(u, v, w) &= -\alpha(v, u, w) = -\alpha(w, v, u) = -\alpha(u, w, v) \\ &= \alpha(v, w, u) = \alpha(w, u, v). \end{aligned}$$

Note

Note that if $k = 1$, then condition 2 is vacuous. Therefore, a 1-form of V is just an element of $V^* = L(W, \mathbb{R})$.

Remark (Permutations)

*From the last example, we notice that the 'sign' of the value changes as we permute more times. To be precise, we are performing **transpositions** on the arguments [1], i.e. we only swap two of the arguments in a single move. Here are several remarks about permutations from group theory:*

[1] See PMATH 347.

- *A permutation σ of $\{1, 2, \ldots, k\}$ is a bijective map.*
- *Compositions of permutations results in a permutation.*
- *The set S_k of permutations on the set $\{1, 2, \ldots, k\}$ is called a **group**.*
- *There are $k!$ such permutations.*
- *For each transposition, we may assign a **parity** of either -1 or 1, and the parity is determined by the number of times we need to perform a transposition to get from $(1, 2, \ldots, k)$ to $(\sigma(1), \sigma(2), \ldots, \sigma(k))$. We usually denote a parity by $\operatorname{sgn}(\sigma)$.*

The following is a fact proven in group theory: let $\sigma, \tau \in S_k$. Then

$$\begin{aligned} \operatorname{sgn}(\sigma \circ \tau) &= \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau) \\ \operatorname{sgn}(\mathrm{id}) &= 1 \\ \operatorname{sgn}(\tau) &= \operatorname{sgn}(\tau^{-1}). \end{aligned}$$

Using the above remark, we can rewrite condition 2 as follows:

Note (Rewrite of condition 2 for Definition 10)

α is alternating, i.e.

$$\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma) \cdot \alpha(v_1, \ldots, v_k),$$

where $\sigma \in S_k$.

Remark

If α is a k-form on V, notice that

$$\alpha(v_1, \ldots, v_k) = 0$$

if ***any 2 of the arguments are equal****.*

4 Lecture 4 Jan 14th

4.1 The Space of k-forms on V (Continued)

Definition 11 (Space of k-forms on V)

*The **space of k-forms on V**, denoted as $\bigwedge^k (V^*)$, is the set of all k-forms on V, made into a vector space by setting*

$$(t\alpha + s\beta)(v_1, \ldots, v_k) := t\alpha(v_1, \ldots, v_k) + s\beta(v_1, \ldots, v_k),$$

for $\alpha\beta \in \bigwedge^k (V^*)$, $t, s \in \mathbb{R}$.

Note

By convention, we define $\bigwedge^0 (V^*) = \mathbb{R}$. The reasoning shall we shown later.

Note

By the note on page 28, observe that $\bigwedge^1 (V^*) = V^*$.

Proposition 6 (A k-form is equivalently 0 if its arguments are linearly dependent)

Let α be a k-form. Then if $v_1, \ldots, v_k$ are linearly dependent, then

$$\alpha(v_1, \ldots, v_k) = 0.$$

Proof

Suppose one of the $v_1, \ldots, v_k$ is a linear combination of the rest of the other vectors; i.e.

$$v_j = c_1 v_1 + \ldots + c_{j-1} v_{j-1} + c_{j+1} v_{j+1} + \ldots + c_k v_k.$$

Then since α is multilinear, and by the last remark in Chapter 3, we have

$$\alpha(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_k) = 0.$$

$\square$

Corollary 7 (k-forms of even higher dimensions)

$\bigwedge^k (V^*) = \{0\}$ *if* $k > n = \dim V$.

Proof

Any set of $k > n$ vectors is necessarily linearly dependent. $\square$

Note

Corollary 7 implies that $\bigwedge^k (V^*)$ *can only be non-trivial for* $0 \leq k \leq n = \dim V$.

4.2 *Decomposable k-forms*

There is a simple way to construct a k-form on V using k-many 1-forms from V, i.e. k-many elements from V^*. Let $\alpha^1, \ldots, \alpha^k \in V^*$. Define a map

$$\alpha^1 \wedge \ldots \wedge \alpha^k : \underbrace{V \times V \times \ldots \times V}_{k \text{ copies}} \to \mathbb{R}$$

by

$$\left(\alpha^1 \wedge \ldots \wedge \alpha^k\right)(v_1, \ldots, v_k) := \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \alpha^{\sigma(1)}(v_1)\alpha^{\sigma(2)}(v_2) \ldots \alpha^{\sigma(k)}(v_k). \tag{4.1}$$

We need, of course, to verify that the above formula is, indeed, a k-form. Before that, consider the following example:

Example 4.2.1

If $k = 2$, we have

$$\left(\alpha^1 \wedge \alpha^2\right)(v_1, v_2) = \alpha^1(v_1)\alpha^2(v_2) - \alpha^2(v_1)\alpha^1(v_2).$$

and if $k = 3$, we have

$$\begin{aligned}\left(\alpha^1 \wedge \alpha^2 \wedge \alpha^3\right)(v_1, v_2, v_3) &= \alpha^1(v_1)\alpha^2(v_2)\alpha^3(v_3) + \alpha^2(v_1)\alpha^3(v_2)\alpha^1(v_1) \\ &\quad + \alpha^3(v_1)\alpha^1(v_2)\alpha^2(v_3) - \alpha^1(v_1)\alpha^3(v_2)\alpha^2(v_3) \\ &\quad - \alpha^2(v_1)\alpha^1(v_1)\alpha^3(v_3) - \alpha^3(v_1)\alpha^2(v_2)\alpha^1(v_3).\end{aligned}$$

Now consider a general case of k. It is clear that Equation (4.1) is k-linear: if we fix any one of the arguments, then Equation (4.1) is reduced to a linear equation.

For the alternating property, let $\tau \in S_k$. WTS

$$\left(\alpha^1 \wedge \ldots \wedge \alpha^k\right)\left(v_{\tau(1)}, \ldots, v_{\tau(k)}\right) = (\operatorname{sgn} \tau)\left(\alpha^1 \wedge \ldots \wedge \alpha^k\right)(v_1, \ldots, v_k).$$

Observe that

$$\begin{aligned}&\left(\alpha^1 \wedge \ldots \wedge \alpha^k\right)\left(v_{\tau(1)}, \ldots, v_{\tau(k)}\right) \\ &= \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \alpha^{\sigma(1)}\left(v_{\tau(1)}\right) \ldots \alpha^{\sigma(k)}\left(v_{\tau(k)}\right) \\ &= \sum_{\sigma \in S_k} \left(\operatorname{sgn} \sigma \circ \tau^{-1}\right)(\operatorname{sgn} \tau)\, \alpha^{\left(\sigma \circ \tau^{-1}\right)(\tau(1))}\left(v_{\tau(1)}\right) \ldots \alpha^{\left(\sigma \circ \tau^{-1}\right)(\tau(k))}\left(v_{\tau(k)}\right) \\ &= (\operatorname{sgn} \tau) \sum_{\sigma \circ \tau^{-1} \in S_k} \left(\operatorname{sgn} \sigma \circ \tau^{-1}\right) \alpha^{\left(\sigma \circ \tau^{-1}\right)(1)}(v_1) \ldots \alpha^{\left(\sigma \circ \tau^{-1}\right)(k)}(v_k) \\ &= (\operatorname{sgn} \tau) \sum_{\sigma \in S_k} \alpha^{\sigma(1)}(v_1) \ldots \alpha^{\sigma(k)}(v_k) \quad \because \text{ relabelling} \\ &= (\operatorname{sgn} \tau)\left(\alpha^1 \wedge \ldots \alpha^k\right)(v_1, \ldots, v_k),\end{aligned}$$

as claimed.

Definition 12 (Decomposable k-form)

The k-form as discussed above is called a ***decomposable k-form****, which for ease of reference shall be re-expressed here:*

$$\left(\alpha^1 \wedge \ldots \wedge \alpha^k\right)(v_1, \ldots, v_k) := \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \alpha^{\sigma(1)}(v_1)\alpha^{\sigma(2)}(v_2) \ldots \alpha^{\sigma(k)}(v_k).$$

Note

Not all k-forms are decomposable. If $k = 1, n-1$ and n, but not for $1 < k < n-1$.

In A1Q5(c), we will show that there exists a 2-form in $n = 4$ that is not decomposable.

Proposition 8 (Permutation on k-forms)

Let $\tau \in S_k$. Then

$$\alpha^{\tau(1)} \wedge \ldots \wedge \alpha^{\tau(k)} = (\operatorname{sgn} \tau)\alpha^1 \wedge \ldots \wedge \alpha^k$$

Proof

Firstly, note that $\operatorname{sgn} \tau = \operatorname{sgn} \tau^{-1}$. Then for any $(v_1, \ldots, v_k) \in V^k$, we have

$$\begin{aligned}
&\alpha^{\tau(1)} \wedge \ldots \wedge \alpha^{\tau(k)}(v_1, \ldots, v_k) \\
&= \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\alpha^{\sigma\circ\tau(1)}(v_1) \ldots \alpha^{\sigma\circ\tau(k)}(v_k) \\
&= \sum_{\sigma\circ\tau S_k} (\operatorname{sgn} \sigma \circ \tau)\left(\operatorname{sgn} \tau^{-1}\right) \alpha^{\sigma\circ\tau(1)}(v_1) \ldots \alpha^{\sigma\circ\tau(k)}(v_k) \\
&= (\operatorname{sgn} \tau) \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\alpha^{\sigma(1)}(v_1) \ldots \alpha^{\sigma(k)}(v_k) \\
&= (\operatorname{sgn} \tau)(\alpha^1 \wedge \ldots \wedge \alpha^k).
\end{aligned}$$

This completes our proof. □

Proof for Proposition 9 is in A1.

Proposition 9 (Alternate Definition of a Decomposable k-form)

Another way we can define a decomposable k-form is

$$(\alpha^1 \wedge \ldots \wedge \alpha^k)(v_1, \ldots, v_k) = \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \alpha^1(v_{\sigma(1)}) \ldots \alpha^k(v_{\sigma(k)}).$$

Theorem 10 (Basis of $\Lambda^k(V^*)$)

Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis of V, a n-dimensional real vector space, and the dual basis $\mathcal{B}^* = \{e^1, \ldots, e^n\}$ of V^*. THen the set

$$\left\{ e^{j_1} \wedge \ldots \wedge e^{j_k} \,\middle|\, 1 \leq j_1 < j_2 < \ldots < j_k \leq n \right\}$$

is a basis of $\Lambda^k(V^*)$.

Corollary 11 (Dimension of $\Lambda^k(V^*)$)

The dimension of $\Lambda^k(V^*)$ is $\binom{n}{k} = \binom{n}{n-k}$, which is also the dimension of $\Lambda^{n-k}(V^*)$. This also works for $k = 0$ [1]*.*

[1] This is why we wanted $\Lambda^0(V^*) = \mathbb{R}$.

Proof (Theorem 10)

Firstly, let α be an arbitrary k-form, and let $v_1, \ldots, v_k \in V$. We may write

$$v_i = v_i^j e_j,$$

where $v_i^j \in \mathbb{R}$. Then

$$\begin{aligned} \alpha(v_1, \ldots, v_k) &= \alpha\left(v_1^{j_1} e_{j_1}, \ldots, v_k^{j_k} e_{j_k}\right) \\ &= v_1^{j_1} \ldots v_k^{j_k} \alpha(e_{j_1}, \ldots, e_{j_k}) \end{aligned}$$

by multilinearity and totally skew-symmetry of α, where $j_i \in \{1, \ldots, n\}$. Let

$$\alpha(e_{j_1}, \ldots, e_{j_k}) = \alpha_{j_1, \ldots, j_k}, \tag{4.2}$$

represent the scalar. Then

$$\begin{aligned} \alpha(v_1, \ldots, v_k) &= \alpha_{j_1, \ldots, j_k} v_1^{j_1} \ldots v_k^{j_k} \\ &= \alpha_{j_1, \ldots, j_k} e^{j_1}(v_1) \ldots e^{j_k}(v_k). \end{aligned}$$

Now since $\alpha_{j_1,\ldots,j_k}$ is totally skew-symmetric, $\alpha = 0$ if any of the j_k's are equal to one another. Thus we only need to consider the terms where the j_k's are distinct. Now for any set of $\{j_1, \ldots, j_k\}$, there exists a unique $\sigma \in S_k$ such that σ rearranges the j_i's so that $j_1, \ldots, j_k$ is strictly increasing. Thus

$$\begin{aligned}
\alpha(v_1, \ldots, v_k) &= \sum_{j_1 < \ldots < j_k} \sum_{\sigma \in S_k} \alpha_{j_{\sigma 1()}, \ldots, \sigma(k)} e^{j_{\sigma(1)}}(v_1) \ldots e^{j_{\sigma(k)}}(v_k) \\
&= \sum_{j_1 < \ldots < j_k} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \alpha_{j_1,\ldots,j_k} e^{j_{\sigma(1)}}(v_1) \ldots e^{j_{\sigma(k)}}(v_k) \\
&= \sum_{j_1 < \ldots < j_k} \alpha_{j_1,\ldots,j_k} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) e^{j_{\sigma(1)}}(v_1) \ldots e^{j_{\sigma(k)}}(v_k) \\
&= \underbrace{\sum_{j_1 < \ldots < j_k} \alpha_{j_1,\ldots,j_k} \left(e^{j_1} \wedge \ldots \wedge e^{j_k} \right)}_{\alpha} (v_1, \ldots, v_k).
\end{aligned}$$

Thus we have that

$$\alpha = \sum_{j_1 < \ldots < j_k} \alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k}. \tag{4.3}$$

Hence $e^{j_1} \wedge \ldots \wedge e^{j_k}$ spans $\Lambda^k(V^*)$.

Now suppose that

$$\sum_{j_1 < \ldots < j_k} \alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k}$$

is the zero element in $\Lambda^k(V^*)$. Then the scalar in Equation (4.2) must be 0 for any $j_1, \ldots, j_k$. Thus

$$\left\{ e^{j_1} \wedge \ldots \wedge e^{j_k} \;\middle|\; 1 \leq j_1 < j_2 < \ldots < j_k \leq n \right\}$$

is linearly independent. $\square$

5 Lecture 5 Jan 16th

5.1 Decomposable k-forms Continued

There exists an equivalent, and perhaps more useful, expression for Equation (4.3), which we shall derive here. Sine $\alpha_{j_1,\ldots,j_k}$ and $e^{j_1} \wedge \ldots \wedge e^{j_k}$ are both totally skew-symmetric in their k indices, and since there are $k!$ elements in S_k, we have that

$$\begin{aligned}
\frac{1}{k!}\alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k} &= \frac{1}{k!} \sum_{\substack{j_1,\ldots,j_k \\ \text{distinct}}} \alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k} \\
&= \frac{1}{k!} \sum_{j_1 < \ldots < j_k} \sum_{\sigma \in S_k} \alpha_{\sigma(j_1),\ldots,\sigma(j_k)} e^{\sigma(j_1)} \wedge \ldots \wedge e^{\sigma(j_k)} \\
&= \frac{1}{k!} \sum_{j_1 < \ldots < j_k} \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma) \alpha_{j_1,\ldots,j_k} (\operatorname{sgn}\sigma) e^{j_1} \wedge \ldots \wedge e^{j_k} \\
&= \frac{1}{k!} \sum_{j_1 < \ldots < j_k} \sum_{\sigma \in S_k} \alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k} \quad [1] \\
&= \sum_{j_1 < \ldots < j_k} \alpha_{j_1,\ldots,j_k} e^{j_1} \wedge \ldots \wedge e^{j_k}.
\end{aligned}$$

The major advantage of the expression with $\frac{1}{k!}$ is that all k indices $j_1, \ldots, j_k$ are summed over all possible values $1, \ldots, n$ instead of having to start with a specific order.

[1] Note that $(\operatorname{sgn}\sigma)(\operatorname{sgn}\sigma) = 1$.

5.2 Wedge Product of Forms

Definition 13 (Wedge Product)

Let $\alpha \in \Lambda^k(V^*)$ and $\beta \in \Lambda^l(V^*)$. We define $\alpha \wedge \beta \in \Lambda^{k+l}(V^*)$ as *follows. Choose a basis $\mathcal{B}^* = \left\{ e^1, \ldots, e^k \right\}$ of V^*. Then we may write*

$$\alpha = \frac{1}{k!} \alpha_{i_1, \ldots, i_k} e^{i_1} \wedge \ldots \wedge e^{i_k} \quad \beta = \frac{1}{l!} \beta_{j_1, \ldots, j_l} e^{j_1} \wedge \ldots \wedge e^{j_l}.$$

We define the ***wedge product*** *as*

$$\begin{aligned} \alpha \wedge \beta &:= \frac{1}{k! l!} \alpha_{i_1, \ldots, i_k} \beta_{j_1, \ldots, j_l} e^{i_1} \wedge \ldots \wedge e^{i_k} \wedge e^{j_1} \wedge \ldots \wedge e^{j_l} \\ &= \sum_{i_1 < \ldots < i_k} \sum_{j_1 < \ldots < j_l} \alpha_{i_1, \ldots, i_k} \beta_{j_1, \ldots, j_k} e^{i_1} \wedge \ldots \wedge e^{i_k} \wedge e^{j_1} \wedge \ldots \wedge e^{j_l}. \end{aligned}$$

One can then question if this definition is well-defined, since it appears to be reliant on the choice of a basis. In A1Q4(a), we will show that this defintiion of $\alpha \wedge \beta$ is indeed well-defined. In particular, one can show that we may express $\alpha \wedge \beta$ in a way that does not involve any of the basis vectors $e^1, \ldots, e^n$.

Definition 14 (Degree of a Form)

For $\alpha \in \Lambda^k(V^*)$, we say that α has ***degree*** *k, and write $|\alpha| = k$.*

Note

By our definition of a wedge product above, we have that

$$|\alpha \wedge \beta| = |\alpha| + |\beta| .$$

Note that since a 0-form lies in $\Lambda^k(V^*)$ for all k, we let $|k|$ be anything / undefined.

Remark

1. *$\alpha \wedge \beta$ is linear in α and linear in β by its definition, i.e. for any $t_1, t_2 \in \mathbb{R}$, $\alpha_1, \alpha_2 \in \Lambda^k(V^*)$, and any $\beta \in \Lambda^l(V^*)$,*

$$(t_1 \alpha_1 + t_2 \alpha_2) \wedge \beta = t_1(\alpha_1 \wedge \beta) + t_2(\alpha_2 \wedge \beta),$$

   and a similar equation works for linearity in β.

2. *The wedge product is associative; this follows almost immediately from its construction.*

3. *The wedge product is not commutative. In fact, if $|\alpha| = k$ and $|\beta| = l$, then*
$$\beta \wedge \alpha = (-1)^{kl} \alpha \wedge \beta. \tag{5.1}$$
We call this property of a wedge product ***graded commutative***, ***super commutative*** *or* ***skewed-commutative***.

Note that this also means that ***even degree forms commute*** *with any form.*

Also, note that if $|\alpha|$ is odd, then $\alpha \wedge \alpha = 0$.

Example 5.2.1

Let $\alpha = e^1 \wedge e^3$ and $\beta = e^2 + e^3$. Then

$$\begin{aligned} \alpha \wedge \beta &= \left(e^1 \wedge e^3\right) \wedge \left(e^2 + e^3\right) \\ &= e^1 \wedge e^3 \wedge e^2 + e^1 \wedge e^3 \wedge e^3 \\ &= -e^1 \wedge e^2 \wedge e^3 + 0 \\ &= -e^1 \wedge e^2 \wedge e^3. \end{aligned}$$

Example 5.2.2

Suppose $\alpha^1, \ldots, \alpha^k$ are linearly dependent 1-forms on V. Then $\alpha^1 \wedge \ldots \wedge \alpha^k = 0$.

Proof

Suppose at least one of the α^j is a linear combination of the rest, i.e.

$$\alpha^j = c_1\alpha^1 + \ldots + c_{j-1}\alpha^{j-1} + c_{j+1}\alpha^{j+1} + \ldots + c_k\alpha^k.$$

Since all of the α^i's are 1-forms, we will have $\alpha^i \wedge \alpha^i$ in the wedge product, and so our result follows from our earlier remark. $\square$

Example 5.2.3

Let $\alpha = \alpha_i e^i, \beta = \beta_j e^j \in V^*$. Then

$$\begin{aligned}\alpha \wedge \beta &= \alpha_i \beta_j e^i \wedge e^j \\ &= \frac{1}{2}\alpha_i \beta_j e^i \wedge e^j + \frac{1}{2}\alpha_i \beta_j e^i \wedge e^j \\ &= \frac{1}{2}\alpha_i \beta_j e^i \wedge e^j - \frac{1}{2}\alpha_j \beta_i e^i \wedge e^j \\ &= \frac{1}{2}(\alpha_i \beta_j - \alpha_j \beta_i) e^1 \wedge e^j \\ &= \frac{1}{2}(\alpha \wedge \beta)_{ij} e^i \wedge e^j,\end{aligned}$$

where $(\alpha \wedge \beta)_{ij} = \alpha_i \beta_j - \alpha_j \beta_i$.

We shall prove the following in A1Q6.

Exercise 5.2.1

Let $\alpha = \alpha_i e^i \in V^*$, and

$$\eta = \frac{1}{2}\eta_{jk} e^j \wedge e^k \in \Lambda^2(V^*).$$

Show that

$$\alpha \wedge \eta = \frac{1}{6!}(\alpha \wedge \eta)_{ijk} e^i \wedge e^j \wedge e^k,$$

where

$$(\alpha \wedge \eta)_{ijk} = \alpha_1 \eta_{jk} + \alpha_j \eta_{ki} + \alpha_k \eta_{ij}.$$

5.3 Pullback of Forms

For a linear map $T \in L(V, W)$, we have seen its induced dual map $T^* \in L(W^*, V^*)$. We shall now generalize this dual map to k-forms, for $k > 1$.

Definition 15 (Pullback)

Let $T \in L(V, W)$. For any $k \geq 1$, define a map

$$T^* : \Lambda^k(W^*) \to \Lambda^k(V^*),$$

called the ***pullback****, as such: let $\beta \in \Lambda^k(W^*)$, and define $T^*\beta \in \Lambda^k(V^*)$ such that*

$$(T^*\beta)(v_1, \ldots, v_k) := \beta(T(v_1), \ldots, T(v_k)).$$

Note

It is clear that $T^*\beta$ is multilinear and alternating, since T itself is linear, and β is multilinear and alternating.

The pullback has the following properties which we shall prove in A1Q8.

Proposition 12 (Properties of the Pullback)

1. *The map $T^* : \Lambda^k(W^*) \to \Lambda^k(V^*)$ is linear, i.e. $\forall \alpha, \beta \in \Lambda^k(W^*)$ and $s, t \in \mathbb{R}$,*
$$T^*(t\alpha + s\beta) = tT^*\alpha + sT^*\beta. \tag{5.2}$$

2. *The map T^* is compatible iwth the wedge product operation in the following sense: if $\alpha \in \Lambda^k(W^*)$ and $\beta \in \Lambda^l(W^*)$, then*
$$T^*(\alpha \wedge \beta) = (T^*\alpha) \wedge (T^*\beta).$$

Part II

The Vector Space $\mathbb{R}^n$ as a Smooth Manifold

6 Lecture 6 Jan 18th

6.1 The space $\Lambda^k(V)$ of k-vectors and Determinants

Recall that we identified V with V^{**}, and so we may consider $\Lambda^k(V) = \Lambda^k(V^{**})$ as the space of k-linear alternating maps

$$\underbrace{V^* \times V^* \times \ldots \times V^*}_{k \text{ copies}} \to \mathbb{R}.$$

Consequently (to an extent), the elements of $\Lambda^k(V)$ are called **k-vectors**. A k-vector is an alternating k-linear map that takes k **covectors** (of 1-forms) to $\mathbb{R}$.

Example 6.1.1

Let $\{e_1, \ldots, e_n\}$ be a basis of V with the dual basis $\{e^1, \ldots, e^n\}$, which is a basis of V^*. Then any $\mathcal{A} \in \Lambda^k(V^*)$ can be written uniquely as

$$\mathcal{A} = \sum_{i_1 < \ldots < i_k} \mathcal{A}^{i_1, \ldots, i_k} \, e_{i_1} \wedge \ldots \wedge e_{i_k}$$

where

$$\mathcal{A}^{i_1, \ldots, i_k} = \mathcal{A}\left(e^{i_1}, \ldots, e^{i_k}\right).$$

We also have that

$$\mathcal{A} = \frac{1}{k!} \mathcal{A}^{i_1, \ldots, i_k} \, e^{i_1} \wedge \ldots \wedge e^{i_k}.$$

Note

Note that

$$\dim \Lambda^k(V) = \frac{n!}{k!(n-k)!}.$$

Definition 16 (k^{th} Exterior Power of T)

Let $T \in L(V, W)$. Then T induces a linear map

$$\Lambda^k(T) \in L\left(\Lambda^k(V), \Lambda^k(W)\right),$$

defined as

$$\left(\Lambda^k T\right)(v_1 \wedge \ldots \wedge v_k) = T(v_1) \wedge \ldots \wedge T(v_k),$$

*where $v_1, \ldots, v_k$ are decomposable elements of $\Lambda^k(V)$, and then extended by linearity to all of $\Lambda^k(V)$. The map $\Lambda^k T$ is called the **k^{th} exterior power of T**.*

Note

Consider the special case of when $W = V$ and $k = n = \dim V$. Then $T \in L(V)$ induces a linear operator $\Lambda^n(T) \in L(\Lambda^n(V))$. It is also noteworthy to point out that any linear operator on a 1-dimensional vector space is just scalar multiplication.

Furthermore, notice that in the above special case, we have

$$\dim \Lambda^n(V) = \binom{n}{n} = 1.$$

Definition 17 (Determinant)

*Let $\dim V = n$ and $T \in L(V)$. We have that $\dim \Lambda^n(V) = 1$. Then $\Lambda^n T \in L(\Lambda^n(V))$ is a scalar multiple of the identity. We denote this scalar multiple by $\det T$, and call it the **determinant** of T, i.e.*

$$\Lambda^n(T)\mathcal{A} = (\det T) I A$$

for any $\mathcal{A} \in \Lambda^n(V)$, where I is the identity operator.

Note

We should verify that this 'new' definition of a determinant agrees with *the 'classical' definition of a determinant.*

Proof

Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis of V, and let $A = [T]_{\mathcal{B}}$ be the $n \times n$ matrix of T wrt the basis $\mathcal{B}$. So $T(e_i) = A_i^j e_j$. Then $\{e_1 \wedge \ldots \wedge e_n\}$ is a basis of $\Lambda^n(V)$, and

$$\begin{aligned}
(\Lambda^n T)(e_1 \wedge \ldots \wedge e_n) &= T(e_1) \wedge \ldots \wedge T(e_n) \\
&= A_1^{i_1} e_{i_1} \wedge \ldots \wedge A_n^{i_n} e_{i_n} \\
&= A_1^{i_1} A_2^{i_2} \ldots A_n^{i_n} \, e_{i_1} \wedge \ldots \wedge e_{i_n} \\
&= \sum_{\substack{i_1, \ldots, i_n \\ \text{distinct}}} A_1^{i_1} \ldots A_n^{i_n} \, e_{i_1} \wedge \ldots \wedge e_{i_n} \\
&= \sum_{\sigma \in S_n} A_1^{\sigma(1)} \ldots A_n^{\sigma(n)} \, e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(n)} \\
&= \sum_{\sigma \in S_n} A_1^{\sigma(1)} \ldots A_n^{\sigma(n)} \, (\operatorname{sgn} \sigma) e_1 \wedge \ldots \wedge e_n \\
&= \left(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) A_1^{\sigma(1)} \ldots A_n^{\sigma(n)} \right) (e_1 \wedge \ldots \wedge e_n) \\
&= \left(\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^n A_i^{\sigma(i)} \right) (e_1 \wedge \ldots \wedge e_n).
\end{aligned}$$

We observe that we indeed have

$$\det T = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^n A_i^{\sigma(i)}.$$

□

6.2 *Orientation Revisited*

Now that we have this notion, we may finally clarify to ourselves what an orientation is without having to rely on roundabout methods as before.

Basically, we now have a more mathematical way of saying 'pick a direction and consider it as the positive direction of V, and that'll be our orientation'.

Definition 18 (Orientation)

Let V be an n-dimensional real vector space. Then $\Lambda^n(V)$ is a 1-dimensional real vector space. An ***orientation*** *on V is defined as a* ***choice*** *of a non-*

zero element $\mu \in \Lambda^n(V)$, up to positive scalar multiples.

Note

For any two such orientations μ and $\tilde{\mu}$, we have that $\tilde{\mu} = \lambda\mu$ for some non-zero $\lambda \in \mathbb{R}$, and by using the definition of having the same orientation, we say that $\mu \sim \tilde{\mu}$ if $\lambda > 0$ and $\mu \not\sim \tilde{\mu}$ if $\lambda < 0$.

Exercise 6.2.1

Check that Definition 18 agrees with Definition 5. (Hint: Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis of V and let $\mu = e_1 \wedge \ldots \wedge e_n$.)

6.3 Topology on $\mathbb{R}^n$

We shall begin with a brief review of some ideas from multivariable calculus.

We know that $\mathbb{R}^n$ is an n-dimensional real vector space. It has a canonical **positive-definite** **inner product**, aka the **Euclidean inner product**, or the **dot product**: given $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, we have

$$x \cdot y = \sum_{i=1}^{n} x^i y^i = \delta_{ij} x^i y^j.$$

The following properties follow from above: for any $t, s \in \mathbb{R}$ and $x, y, w \in \mathbb{R}^n$,

- $(tx + sy) \cdot w = t(x \cdot w) = s(y \cdot w)$;
- $x \cdot (ty + sw) = t(x \cdot y) + t(x \cdot w)$;
- $x \cdot y = y \cdot x$;
- (positive definiteness) $x \cdot x \geq 0$ with $x \cdot x = 0 \iff x = 0$;
- (Cauchy-Schwarz Ineq.) $-\|x\|\,\|y\| \leq x \cdot y \leq \|x\|\,\|y\|$, i.e.

$$x \cdot y = \|x\|\,\|y\| \cos\theta$$

where $\theta \in [0, \pi]$.

Definition 19 (Distance)

*The **distance** between $x, y \in \mathbb{R}^n$ is given as*

$$\operatorname{dist}(x, y) = \|x - y\|.$$

Note (Triangle Inequality)

Note that the triangle inequality holds for the distance function[1]: for any $x, z \in \mathbb{R}^n$, for any $y \in \mathbb{R}^n$,

$$\operatorname{dist}(x, z) \leq \operatorname{dist}(x, y) + \operatorname{dist}(y, z).$$

[1] See also PMATH 351

Definition 20 (Open Ball)

Let $x \in \mathbb{R}^n$ and $\varepsilon > 0$. The open ball of radius ε centered at x is

$$B_\varepsilon(x) = \{y \in \mathbb{R}^n \mid \operatorname{dist}(x, y) < \varepsilon\}.$$

*A subset $U \subseteq \mathbb{R}^n$ is called **open** if $\forall x \in U$, $\exists \varepsilon > 0$ such that*

$$B_\varepsilon(x) \subseteq U.$$

Example 6.3.1

- $\emptyset$ and $\mathbb{R}^n$ are open.
- If U and V are open, so is $U \cap V$.
- If $\{U_\alpha\}_{\alpha \in A}$ is open, so is $\bigcup_{\alpha \in A} U_\alpha$.

7 Lecture 7 Jan 21st

7.1 Topology on $\mathbb{R}^n$ (Continued)

Definition 21 (Closed)

*A subset $F \subseteq \mathbb{R}^n$ is **closed** if its complement $\mathbb{R}^n \setminus F =: F^C$ is open.*

Warning

A subset does not have to be either open or closed. Most subsets are neither.

Note

- *Arbitrary intersections of closed sets is closed.*
- *Finite unions of closed sets is closed.*

Note (Notation)

We call

$$\overline{B}_\varepsilon(x) := \{y \in \mathbb{R}^n \mid \|x - y\| \leq \varepsilon\}$$

the closed ball of radius ε centered at x.

Definition 22 (Continuity)

*Let $A \subseteq \mathbb{R}^n$. Let $f : A \to \mathbb{R}^m$, and $x \in A$. We say that f is **continuous** at x if $\forall \varepsilon > 0, \exists \delta > 0$ such that*

$$f(B_\delta(x) \cap A) \subseteq B_\varepsilon(f(x)).$$

*We say that f is **continuous** on A if $\forall x \in A$, f is continuous on x.*

Proposition 13 (Inverse of a Continuous Map is Open)

For a proof, see PMATH 351.

Let $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$. Then f is continuous on A iff whenever $V \subseteq \mathbb{R}^m$ is open, $f^{-1}(V) = A \cap U$ for some $U \subseteq \mathbb{R}^n$ is open.

Definition 23 (Homeomorphism)

Let $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$. Let $B = f(A)$. We say that f is a homeomorphism of A onto B if $f : A \to B$

- *is a bijection;*
- *and $f^{-1} : B \to A$ is continuous on A and B, respectively.*

7.2 Calculus on $\mathbb{R}^n$

Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ be a continuous map. Also, let

$$x = (x^1, \ldots, x^n) \in \mathbb{R}^n \text{ and } y = (y^1, \ldots, y^m) \in \mathbb{R}^m.$$

Then the **component functions** of f are defined by

$$y^k = f^k(x^1, \ldots, x^n), \text{ where } y = (y^1, \ldots, y^m) = f(x) = f(x^1, \ldots x^n).$$

Thus $f = (f^1, \ldots, f^m)$ is a collection of m-real-valued functions on $U \subseteq \mathbb{R}^n$.

Definition 24 (Smoothness)

*Let $x_0 \in U$. We say that f is **smooth** (or C^∞, or **infinitely differentiable**) if all **partial derivatives** of each component function f^k exists* *and are* ***continuous*** *at* x_0*. I.e., if we let* $\frac{\partial}{\partial x^i} = \partial_i$ *denote the operator of partial differentiation in the* x^i *direction, then*

$$\partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f^k$$

exists and is continuous at x_0*, for all* $k = 1, \dots, n$*, and all* $\alpha_i \geq 0$*.*

Definition 25 (Diffeomorphism)

Let $U \subseteq \mathbb{R}^n$ *be open,* $f : U \to \mathbb{R}^m$*, and* $V = f(U)$*. We say* f *is a* ***diffeomorphism*** *of* U *onto* V *if* $f : U \to V$ *is bijective*[1]*, smooth, and that its inverse* f^{-1} *is smooth.*

We say that U *and* V *are* ***diffeomorphic*** *if such a diffeomorphism exists.*

[1] A function that is **not injective** may not have a surjection from its image.

Note

A diffeomorphism preserves the 'smoothness of a structure', i.e. the notion of calculus is the same for diffeomorphic spaces.

Example 7.2.1

If $f : U \to V$ is a diffeomorphism , then $g : V \to \mathbb{R}$ is smooth iff $g \circ f : U \to \mathbb{R}$ is smooth.

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[rr, bend left, "g \circ f"] & V \arrow[r, "g"] & \mathbb{R} \end{tikzcd}$$

Figure 7.1: Preservation of smoothness via diffeomorphisms

Note

A diffeomorphism is also called a ***smooth reparameterization*** *(or just a* ***parameterization*** *for short).*

Definition 26 (Differential)

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *be a smooth mapping, and* $x_0 \in U$*. The* ***differential*** *of* f *at* x_0*, denoted* $(df)_{x_0}$*, is a linear map* $(\mathrm{D} f)_{x_0} : \mathbb{R}^n \to$

$\mathbb{R}^m$, or an $m \times n$ real matrix, given by

$$(\mathrm{D} f)_{x_0} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \dots & \frac{\partial f^1}{\partial x^n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1}(x_0) & \dots & \frac{\partial f^m}{\partial x^n}(x_0) \end{pmatrix},$$

where the notation (x_0) means evaluation at x_0, and the $(i,j)^{th}$ entry of $(\mathrm{D} f)_{x_0}$ is $\frac{\partial f^i}{\partial x^j}(x_0)$. $(\mathrm{D} f)_{x_0}$ is also called the ***Jacobian*** *or* ***tangent map*** *of f at x_0.*

Note (Change of notation)
We changed the notation for the ***differential*** *on Feb 3rd to using $\mathrm{D} f$. The old notation was df.*

Proposition 14 (Differential of the Identity Map is the Identity Matrix)

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be the identity mapping $f(x) = x$. Then $(\mathrm{D} f)_{x_0} = I_n$, the $n \times n$ matrix, then for any $x_0 \in U$.

Proof
Since $f(x) = x$, since $x \in \mathbb{R}^n$, we may consider the function f as

$$f(x) = I_n x = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n \end{pmatrix}.$$

Then it follows from differentiation that

$$(\mathrm{D} f)_{x_0} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

and it does not matter what x_0 is. $\square$

Note
In multivariable calculus, we learned that if f is ***smooth*** *at x_0* [2]*, then*

$$\underset{m\times 1}{f(x)} = \underset{m\times 1}{f(x_0)} + \underset{m\times n}{(\mathrm{D} f)_{x_0}} \underset{n\times 1}{(x - x_0)} + \underset{m\times 1}{Q(x)},$$

[2] Back in multivariable calculus, just being C^1 at x_0 is sufficient for being smooth

where $Q : U \to \mathbb{R}^m$ satisfies

$$\lim_{x \to x_0} \frac{Q(x)}{\|x - x_0\|} = 0.$$

Note

Note that when $n = m = 1$, the existence of the differential of a continuous real-valued function $f(x)$ at a real number $x_0 \in U \subseteq \mathbb{R}$ is the same of the usual derivative $f'(x)$ at $x = x_0$. In fact, $f'(x_0) = (\mathrm{D} f)_{x_0} = \frac{df}{dx}(x_0)$.

Theorem 15 (The Chain Rule)

Let

$$\begin{aligned} f &: U \subseteq \mathbb{R}^n \to \mathbb{R}^m \\ g &: V \subseteq \mathbb{R}^m \to \mathbb{R}^p, \end{aligned}$$

be two smooth maps, where U, V are open in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and and such that $V = f(U)$. Then the composition $g \circ f$ is also smooth. Further, if $x_0 \in U$, then

$$(D(g \circ f))_{x_0} = (\mathrm{D} g)_{f(x_0)} (\mathrm{D} f)_{x_0}. \tag{7.1}$$

7.3 Smooth Curves in $\mathbb{R}^n$ and Tangent Vectors

We shall now look into tangent vectors and the tangent space at every point of $\mathbb{R}^n$. We need these two notions to construct objects such as vector fields and **differential forms**. In particular, we need to consider these objects in multiple abstract ways so as to be able to generalize these notions in more abstract spaces, particularly to **submanifolds** of $\mathbb{R}^n$ later on.

Plan We shall first consider the notion of **smooth curves**, which we shall simply call a curve, and shall always (in this course) assume curves as smooth objects. We shall then use **velocities** of curves to define **tangent vectors**.

Definition 27 (Smooth Curve)

Let $I \subseteq \mathbb{R}$ be an open interval. A smooth map $\phi : I \to \mathbb{R}^n$ is called a ***smooth curve****, or* ***curve****, in $\mathbb{R}^n$. Let $t \in I$. Then each of its component functions $\phi^k(t)$ in $\phi(t) = (\phi^1(t), \ldots, \phi^n(t))$ is a smooth real-valued function of t.*

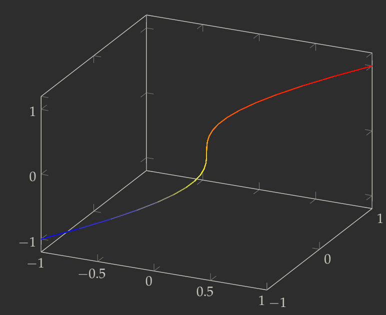

Figure 7.2: A curve in $\mathbb{R}^3$

Example 7.3.1

Let $a, b > 0$. Consider $\phi : I \to \mathbb{R}^3$ given by

$$\phi(t) = (a \cos t, a \sin t, bt).$$

Since each of the components are smooth[3], we have that ϕ itself is also smooth. The shape of the curve is as shown in Figure 7.3.

[3] **Wait**, do we actually consider bt smooth when it's only C^1, in this course?

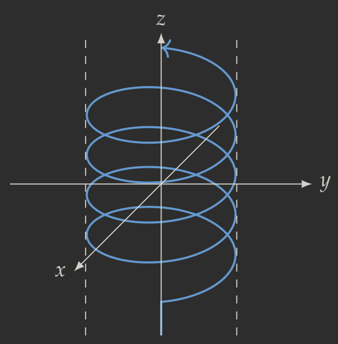

Figure 7.3: Helix curve

8 Lecture 8 Jan 23rd

8.1 Smooth Curves in $\mathbb{R}^n$ and Tangent Vectors (Continued)

Definition 28 (Velocity)

*Let $\phi : I \to \mathbb{R}^n$ be a curve. The **velocity** of the curve ϕ at the point $\phi(t_0) \in \mathbb{R}^n$ for $t_0 \in I$ is defined as*

$$\phi'(t_0) = (d\phi)_{t_0} \in \mathbb{R}^{n\times 1} \simeq \mathbb{R}^n.$$

Note

*$\phi'(t_0) = (d\phi)_{t_0}$ is the **instantaneous rate of change** of ϕ at the point $\phi(t_0) \in \mathbb{R}^n$.*

Example 8.1.1

From the last example, we had $\phi(t) = (a\cos t, a\sin t, bt)$ for $a, b > 0$. Then

$$\phi'(t) = (-a\sin t, a\cos t, b)$$

Let $t_0 = \frac{\pi}{2}$. Then the velocity of ϕ at

$$\phi\left(\frac{\pi}{2}\right) = (0, a, \frac{b\pi}{2})$$

is

$$\phi'\left(\frac{\pi}{2}\right) = (-a, 0, b).$$

Definition 29 (Equivalent Curves)

Let $p \in \mathbb{R}^n$. Let $\phi : I \to \mathbb{R}^n$ and $\psi : \tilde{I} \to \mathbb{R}^n$ be two smooth curves in $\mathbb{R}^n$ such that both the open intervals I and $\tilde{I}$ contain 0. We say that ϕ is ***equivalent at p to*** *ψ, and denote this as*

$$\phi \sim_p \psi,$$

iff

- $\phi(0) = \psi(0) = p$, *and*
- $\phi'(0) = \psi'(0)$.

Note

In other words, $\phi \sim_p \psi$ iff both ϕ and ψ passes through p at $t = 0$, and have the same velocity at this point.

Example 8.1.2

Consider the two curves

$$\phi(t) = (\cos t, \sin t) \text{ and } \psi(t) = (1, t),$$

where $t \in \mathbb{R}$.

Notice that at $p = (1, 0)$, i.e. $t = 0$, we have

$$\phi'(0) = (0, 1) \text{ and } \psi'(0) = (0, 1).$$

Thus

$$\phi \sim_p \psi.$$

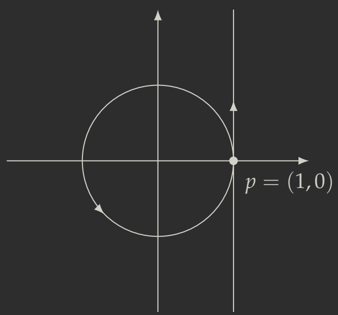

Figure 8.1: Simple example of equivalent curves in Example 8.1.2

Proposition 16 (Equivalent Curves as an Equivalence Relation)

$\sim_p$ is an equivalence relation.

Exercise 8.1.1

Proof of Proposition 16 is really straightforward so try it yourself.

Definition 30 (Tangent Vector)

A ***tangent vector*** *to* $\mathbb{R}^n$ *at* p *is a vector* $v \in \mathbb{R}^n$*, thought of as '****emanating****' from* p*, is in a one-to-one correspondence with an equivalence class*

$$[\phi]_p := \left\{ \psi : I \to \mathbb{R}^n \mid \psi \sim_p \phi \right\}.$$

Definition 31 (Tangent Space)

The ***tangent space*** *to* $\mathbb{R}^n$ *at* p*, denoted* $T_p\left(\mathbb{R}^n\right)$ *is the set of all equivalence classes* $[\phi]_p$ *wrt* $\sim_p$*.*

Now if $\phi : I \to \mathbb{R}^n$ is a smooth curve in $\mathbb{R}^n$ with $0 \in I$, and $\phi'(0) = v \in \mathbb{R}^n$, then we write v_p to denote the element in $T_p\left(\mathbb{R}^n\right)$ that it represents.

Proposition 17 (Canonical Bijection from $T_p(\mathbb{R}^n)$ to $\mathbb{R}^n$)

There exists a canonical bijection from $T_p(\mathbb{R}^n)$ *to* $\mathbb{R}^n$*. Using this bijection, we can equip the tangent space* $T_p(\mathbb{R}^n)$ *with the structure of a real* n*-dimensional real vector space.*

Proof

Let $v_p = [\phi]_p \in T_p(\mathbb{R}^n)$, where $v = \phi'(0) \in \mathbb{R}^n$, for any $\phi \in [\phi]_p$. Let $\gamma_{v_p} : \mathbb{R} \to \mathbb{R}^n$ by

$$\gamma_{v_p}(t) = (p + tv) = (p^1 + tv^1, p^2 + tv^2, \ldots, p^n + tv^n).$$

It follows by construction that γ_{v_p} is smooth, $\gamma_{v_p}(0) = p$, and $\gamma'_{v_p}(0) = v$. Thus $\gamma_{v_p} \sim_p \phi$. In particular, we have $[\gamma_{v_p}]_p = [\phi]_p = v_p \in T_p(\mathbb{R}^n)$. In fact, notice that γ_{v_p} is the straight line through p in the direction of v.

Now consider the map $T_p : \mathbb{R}^n \to T_p(\mathbb{R}^n)$, given by

$$T_p(v) = [\gamma_{v_p}]_p.$$

In other words, we defined the map T_p to send a vector $v \in \mathbb{R}^n$

to the **equivalence class of all smooth curves passing through p with velocity v at p**. Note that since γ_{v_p} has a 'dependency' on v, it follows that T_p is indeed a bijection.

We now get a vector space structure on $T_p(\mathbb{R}^n)$ from that of $\mathbb{R}^n$ by letting T_p be a linear isomorphism, i.e. we set

$$a[\phi]_p + b[\psi]_p = T_p\left(aT_p^{-1}([\phi]_p) + bT_p^{-1}([\psi]_p)\right)$$

for all $a, b \in \mathbb{R}$ and all $[\phi]_p, [\psi]_p \in T_p(\mathbb{R}^n)$. $\square$

Note

Another way we can say the last line in the proof above is as follows: if $v_p, w_p \in T_p(\mathbb{R}^n)$ and $a, b \in \mathbb{R}$, then we define $av_p + bw_p = (av + bw)_p$.

In other words, looking at the tangent vectors at p is similar to looking at the tangents vectors at the origin 0.

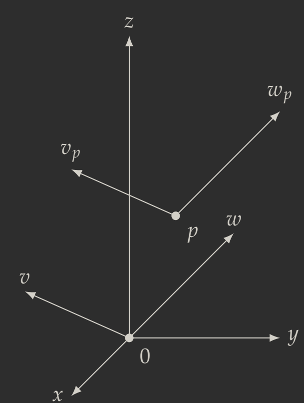

Figure 8.2: Canonical bijection from $T_p(\mathbb{R}^n)$ to $\mathbb{R}^n$

Note

The fact that there is a canonical isomorphism between $\mathbb{R}^n$ and the equivalence classes wrt $\sim_p$ is a pheonomenon that is particular to $\mathbb{R}^n$.

For a k-dimensional ***submanifold*** *M of $\mathbb{R}^n$, or more generally, for an abstract smooth k-dimensional manifold M, and a point $p \in M$, it is true that we can still define $T_p(M)$ to be the set of equivalence classes of curves wrt to some 'natural' equivalence relation. However, there is no canonical representation of each equivalence class, and so $T_p(M) \simeq \mathbb{R}^k$, but not canonically so.*

9 Lecture 9 Jan 25th

9.1 Derivations and Tangent Vectors

Recall the notion of a **directional derivative**.

Definition 32 (Directional Derivative)

Let $p, v \in \mathbb{R}^n$. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be smooth, where U is an open set that contains p (i.e. an open nbd of p). The ***directional derivative*** *of f at p in the direction of v, denoted $v_p f$, is defined as*

$$v_p f = \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}. \tag{9.1}$$

Remark

The above limit may or may not exist given an arbitrary f, p and v. However, since we're working exclusively with smooth functions, this limit will always exist for us.

Note

By definition, we may think of $v_p f \in \mathbb{R}$ as the instantaneous rate of change of f at the point p as we 'move in the direction of' the vector v.

Remark

In multivariable calculus, one may have seen this definition with the additional condition that v is a unit vector. We do not have that restriction here.

Also, note that we have deliberately used the same notation v_p that we *used for elements of $T_p(\mathbb{R}^n)$, which seems awkward, but it shall be clarified in* *Corollary 20**.*

Example 9.1.1

In the special case of when $v = \hat{e}_i$, where $\hat{e}_i$ is the ith standard basis vector. Then we have

$$(\hat{e}_i)_p f = \lim_{t \to 0} \frac{f(p + t\hat{e}_i) - f(p)}{t} = \frac{\partial f}{\partial x^i}(p) = (f \circ \gamma_{v_p})'(p)$$

for the directional derivative of f at p in the $\hat{e}_i$ direction. This is precisely the partial derivative of f in the x^i direction at the point $p \in \mathbb{R}^n$.

Theorem 18 (Linearity and Leibniz Rule for Directional Derivatives)

Let $p \in \mathbb{R}^n$, and let f, g be smooth real-valued functions defined on open neighbourhoods of p. Let $a, b \in \mathbb{R}$. Then

1. *(**Linearity**) $v_p(af + bg) = av_p f + bv_p g$;*
2. *(**Leibniz Rule** / **Product Rule**) $v_p(fg) = f(p)v_p g + g(p)v_p f$.*

Proof

Proven on A2Q2.

Recall that given $p, v \in \mathbb{R}^n$, we denote γ_{v_p} as the curve $\gamma_{v_p}(t) = p + tv$, which is the straight line passing through p with constant velocity v. Thus we mmay rewrite Equation (9.1) as

$$v_p f = \lim_{t \to 0} \frac{f(\gamma_{v_p}(t)) - f(\gamma_{v_p}(0))}{t} = (f \circ \gamma_{v_p})'(0), \tag{9.2}$$

where $f \circ \gamma_{v_p} : \mathbb{R} \to \mathbb{R}$ is smooth as it is a composition of smooth functions.

Theorem 19 (Canonical Directional Derivative, Free From the Curve)

Suppose that $\phi \sim_p \psi$ are two curves on $\mathbb{R}^n$. Let $f : U \to \mathbb{R}$ where U is an open neighbourhood of p. Then

$$(f \circ \phi)'(0) = (f \circ \psi)'(0).$$

Proof

By the chain rule,

$$(f \circ \phi)'(0) = (\mathrm{D}(f \circ \phi))_0 = (\mathrm{D}\, f)_{\phi(0)} (\mathrm{D}\,\phi)_0 = (\mathrm{D}\, f)_{\phi(0)} \phi'(0),$$

and a similar expression holds for ψ. Our desired result follows from the definition of $\sim_p$. $\square$

Corollary 20 (Justification for the Notation $v_p f$)

Let $[\phi]_p \in T_p\mathbb{R}^n$. It follows that

$$v_p f = (f \circ \gamma_{v_p})'(0) = (f \circ \phi)'(0)$$

by Equation (9.2).

Remark

With that, we have established that tangent vectors give us directional derivatives in a way compatible with the characterization of $T_p\mathbb{R}^n$ as equivalence classes wrt $\sim_p$.

Now the fact that Equation (9.1) depends only on the values of f in some open neighbourhood of p motivates us towards the following definition.

Definition 33 ($f \sim_p g$)

Let $p \in \mathbb{R}^n$. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ and $g : V \subseteq \mathbb{R}^n \to \mathbb{R}$ be smooth where U and V are both open neighbourhoods of p. We say that $f \sim_p g$ if $\exists W \subseteq U \cap V$ such that $f \restriction_W = g \restriction_W$. That is, $f \sim_p g$ iff f and g agree at all points sufficiently closde to p.

Note

It is clear from Equation (9.1) that if $f \sim_p g$, then $f(p) = g(p)$ and $v_p f = v_p g$, i.e. f and g agree at p and all possible directional derivatives at p of f and g also agree with each other.

Proposition 21 ($\sim_p$ for Smooth Functions is an Equivalence Relation)

The relation $\sim_p$ on the set of smooth real-valued functions defined on some open neighbourhood of p is an equivalence relation.

Exercise 9.1.1

Prove Proposition 21.

Of course, what else is there to talk about an equivalence relation if not for its equivalence class?

Definition 34 (Germ of Functions)

*An equivalence class of $\sim_p$ is called a **germ of functions** at p. The set of all such equivalence classes is dentoed C_p^∞, called the **space of germs** at p.*

Note

Suppose $f : U \to \mathbb{R}$, where U is an open neighbourhood of p. Then it is clear that $[f]_p = [f \restriction_V]_p$ for any open neighbourhood V of p if $V \subseteq U$.

We can define the structure of a real vector space on C_p^∞ as follows. Let $[f]_p, [g]_p \in C_p^\infty$, where the functions

$$f : U \to \mathbb{R} \text{ and } g : V \to \mathbb{R}$$

represent $[f]_p$ and $[g]_p$, respectively. Also, let $a, b \in \mathbb{R}$. Then we define

$$a[f]_p + b[g]_p = [af + bg]_p, \tag{9.3}$$

where $af + bg$ is restricted to the open neighbourhood $U \cap V$ of p on which both f and g are defined.

We need to show that Equation (9.3) is well-defined. Well suppose $f \sim_p \tilde{f}$ and $g \sim_p \tilde{g}$. Then what we need to show is

$$(af + bg) \sim_p (a\tilde{f} + b\tilde{g}).$$

Since $f \sim_p \tilde{f}$ and $g \sim_p \tilde{g}$, we have that

$$\tilde{f} : \tilde{U} \to \mathbb{R} \text{ and } \tilde{g} : \tilde{V} \to \mathbb{R}.$$

Then, in particular, there exists $W \subseteq U \cap \tilde{U}$ and $Y \subseteq V \cap \tilde{V}$ such that

$$f \restriction_W = \tilde{f} \restriction_W \text{ and } g \restriction_Y = \tilde{g} \restriction_Y .$$

Then $Z = W \cap Y$ is an open neighbourhood of p and thus we must have

$$af + bg = a\tilde{f} + b\tilde{g}$$

on Z. Thus Equation (9.3) is true and C_p^∞ is indeed a vector space.

Further, we can even define a **multiplication** on C_p^∞ by setting

$$[f]_p[g]_p = [fg]_p. \tag{9.4}$$

Example 9.1.2

Check that Equation (9.4) is well-defined.

Proposition 22 (Linearity of the Directional Derivative over the Germs of Functions)

Let $v_p \in T_p\mathbb{R}^n$. Then the map $v_p : C_p^\infty \to \mathbb{R}$ defined by $[f]_p \mapsto v_p[f]_p = v_p f$ is well-defined. This map is also linear in the sense that

$$v_p(a[f]_p + b[g]_p) = av_p[f]_p + bv_p[g]_p.$$

Moreover, this map satisfies Leibniz's rule:

$$v_p([f]_p[g]_p) = f(p)v - p[g]_p + g(p)v_p[f]_p.$$

Proof

Our desired result follows almost immedaitely from Definition 33 and Theorem 18. □

10 Lecture 10 Jan 28th

10.1 Derivations and Tangent Vectors (Continued)

Recall Corollary 20.

Definition 35 (Derivation)

*A **derivation** at p is a linear map $\mathcal{D} : C_p^\infty \to \mathbb{R}$ satisfying the additional property that*

$$\mathcal{D}([f]_p[g]_p) = f(p)\mathcal{D}[g]_p + g(p)\mathcal{D}[f]_p.$$

Remark

Proposition 22 tells us that any tangent vector $v_p \in T_p\mathbb{R}^n$ is a derivation, so the set of derivations is not trivial.

Proposition 23 (Set of Derivations as a Space)

Let Der_p be the set of all derivations at p. Then this is a subset of the vector space $L(C_p^\infty, \mathbb{R})$. In fact, Der_p is a linear subspace.

Proof

We shall prove this in A2Q3.

This is likely surprising seeing that we just introduced yet another definition but there are actually no other derivations at p aside from the tangent vectors at p. In fact, any derivation must be a directional differentiation wrt to some tangent vector $v_p \in T_p\mathbb{R}^n$. Before we can show this, observe the following.

First Let us describe a tangent vector v_p as a derivation at p in terms of the standard basis. Let $\mathcal{B} = \{\hat{e}_1, \ldots, \hat{e}_n\}$ be the standard basis of $\mathbb{R}^n$. Then

$$\{(\hat{e}_1)_p, \ldots, (\hat{e}_n)_p\}$$

is a basis of $T_p\mathbb{R}^n$, which is called the standard basis of $T_p\mathbb{R}^n$. It is the image of $\mathcal{B}$ under the canonical isomorphism

$$T_p : \mathbb{R}^n \to T_p\mathbb{R}^n.$$

Recall from Example 9.1.1 that

$$(\hat{e}_k)_p f = \frac{\partial f}{\partial x^k}(p).$$

As a linear map, we can write

$$(\hat{e}_k)_p = \frac{\partial}{\partial x^k}\bigg|_p. \tag{10.1}$$

Let $v \in \mathbb{R}^n$ be expressed as $v = v^i \hat{e}_i$, in terms of the standard basis. By the chain rule, we have

$$\begin{aligned} v_p f &= (f \circ \gamma_{v_p})'(0) = (\mathrm{D} f)_{\gamma_{v_p}(0)} (\mathrm{D}\, v_p)_0 \\ &= (df)_p v = \frac{\partial f}{\partial x^i}(p) v^i = v^i \frac{\partial}{\partial x^i}\bigg|_p f. \end{aligned}$$

From Equation (10.1), we can write the above as

$$v_p = v^i (\hat{e}_i)_p,$$

which we see is indeed the image of $v = v^i \hat{e}_i$ under the linear isomorphism T_p. Henceforth, we will often express tangent vectors at p in the above form, using linear combinations of the operators $(\hat{e}_i)_p = \frac{\partial}{\partial x^i}\big|_p$.

Second Consider the smooth function $x^j : \mathbb{R}^n \to \mathbb{R}$ given by

$$x^j(q) = q^j,$$

for all $q = (q^1, \dots, q^n) \in \mathbb{R}^n$. So as a function of $x^1, \dots, x^n$ we have

$$x^j(x^1, \dots, x^n) = x^j, \tag{10.2}$$

which is smooth. Let $v_p = v^i \frac{\partial}{\partial x^i}\Big|_p$. Then

$$v_p x^j = v^i \frac{\partial}{\partial x^i}\Big|_p x^j = v^i \delta_i^j = v^j.$$

Thus, we deduced that

$$v_p = v^i \frac{\partial}{\partial x^i}\Big|_p, \text{ where } v^i = v_p x^i. \tag{10.3}$$

Remark

Compare Equation (10.3) and Equation (1.1) and notice the similarity of their v^i's. We shall look into why this is the case later on.

Lemma 24 (Derivations Annihilates Constant Functions)

Let $\mathcal{D}_p$ be a derivation at p. Then $\mathcal{D}$ annihilates constant functions, i.e. if $f(q) = c \in \mathbb{R}$ for all $q \in \mathbb{R}^n$, then $\mathcal{D}_p f = 0$.

Proof

First, consider the constant function $1 : \mathbb{R}^n \to \mathbb{R}$ given by $q \mapsto 1$. Note that $1 \cdot 1 = 1$. By Leibniz's Rule, we have

$$\mathcal{D}_p(1) = \mathcal{D}_p(1 \cdot 1) = 1(p)\mathcal{D}_p 1 + 1(p)\mathcal{D}_p 1 = 2\mathcal{D}_p(1).$$

It follows that $\mathcal{D}_p(1) = 0$.

Now let f be a constant function. Then $f = c1$ for some $c \in \mathbb{R}$. It follows by linearity that

$$\mathcal{D}_p f = \mathcal{D}_p(c1) = c\mathcal{D}_p 1 = 0.$$

$\square$

Theorem 25 (Derivations are Tangent Vectors)

Let $\mathcal{D}_p$ be a derivation at p. Then $\mathcal{D}_p = v_p$ for some $v_p \in T_p\mathbb{R}^n$.

Consequently, $\mathrm{Der}_p = T_p\mathbb{R}^n$.

Proof

Note that if there exists a v_p such that $\mathcal{D}_p = v_p$, then we must have $v_p = v^i \frac{\partial}{\partial x^i}\Big|_p$ with coefficients

$$v^i = v_p x^j = \mathcal{D}_p x^j.$$

In particular, we can show that

$$\mathcal{D}_p = (\mathcal{D}_p x^i) \frac{\partial}{\partial x^i}\Big|_p.$$

Let f be a smooth function defined in an open neighbourhood of p. By the **integral form of Taylor's Theorem**, for $x = (x^1, \ldots, x^n)$ sufficiently close to p, we can write

$$f(x) = f(p) + \frac{\partial f}{\partial x^i}\Big|_p^{(} x^i - p^i) + g_i(x)(x^i - p^i),$$

where the functions $g_i(x)$ satisfy $g_i(p) = 0$. More succinctly,

$$f = f(p) + \frac{\partial f}{\partial x^i}\Big|_p (x^i - p^i) + g_i \cdot (x^i - p^i), \tag{10.4}$$

where x^i is the function $x^i(x) = x^i$ as in Equation (10.2), and p^i and $f(p)$ are constant functions. Apply $\mathcal{D}_p$ to Equation (10.4). By the linearity and Leibniz's rule, both of which are satisfied by $\mathcal{D}_p$, and Lemma 24, we get

$$\begin{aligned}
\mathcal{D}_p f &= \mathcal{D}_p \left(f(p) + \frac{\partial f}{\partial x^i}\Big|_p (x^i - p^i) + g_i \cdot (x^i - p^i) \right) \\
&= 0 + \frac{\partial f}{\partial x^i}\Big|_p \mathcal{D}_p (x^i - p^i) + \mathcal{D}_p (g_i \cdot (x^i - p^i)) \\
&= \frac{\partial f}{\partial x^i}\Big|_p (\mathcal{D}_p x^i + 0) + g_i(p) \mathcal{D}_p (x^i - p^i) + (x^i - p^i)(p) \mathcal{D}_p (g_i) \\
&= (\mathcal{D}_p x^i) \frac{\partial}{\partial x^i}\Big|_p f + 0 + 0 = \left((\mathcal{D}_p x^i) \frac{\partial}{\partial x^i}\Big|_p \right) f.
\end{aligned}$$

Since f was arbitrary, it follows that $\mathcal{D}_p = (\mathcal{D}_p x^i) \frac{\partial}{\partial x^i}\Big|_p$, which is what we desired. $\square$

Remark

From Section 7.3 and Section 9.1, a tangent vector $v_p \in T_p\mathbb{R}^n$ can be considered in any one of the following three ways:

1. *as a vector $v \in \mathbb{R}^n$, enamating from the point $p \in \mathbb{R}^n$;*
2. *as a unique equivalence class of curves through p;*
3. *as a unique derivation at p.*

The three different viewpoints are useful in their own ways, and we will be alternating between these ideas as we go forward.

10.2 Smooth Vector Fields

The idea of a vector field on $\mathbb{R}^n$ is the assignment of a tangent vector at p for every $p \in \mathbb{R}^n$. A smooth vector field is where we attach these tangent vectors to every point in a **smoothly varying way**.

Definition 36 (Tangent Bundle)

The ***tangent bundle*** *of $\mathbb{R}^n$ is defined as*

$$T\mathbb{R}^n = \bigcup_{p \in \mathbb{R}^n} T_p\mathbb{R}^n.$$

Remark

For us, the tangent bundle is just a set, but it is a very important mathematical object which shall be studied in later courses (PMATH 465).

Definition 37 (Vector Field)

A ***vector field*** *on $\mathbb{R}^n$ is a map $X : \mathbb{R}^n \to T\mathbb{R}^n$ such that $X(p) \in T_p\mathbb{R}^n$ for all $p \in \mathbb{R}^n$. We shall always denote $X(p)$ by X_p.*

Let $\{\hat{e}_1, \ldots, \hat{e}_n\}$ be the standard basis of $\mathbb{R}^n$. We have seen that $\{(\hat{e}_1)_p, \ldots, (\hat{e}_n)_p\}$ is a basis of $T_p\mathbb{R}^n$. We can think of each $\hat{e}_i$ as a vector field, where $\hat{e}_i(p) = (\hat{e}_i)_p$. We call these the **standard vector**

fields on $\mathbb{R}^n$. Recall that we wrote that

$$(\hat{e}_k) = \frac{\partial}{\partial x^k},$$

which means that $(\hat{e}_k)_p = \left.\frac{\partial}{\partial x^k}\right|_p$. Henceforth, we shall write the standard vector fields on $\mathbb{R}^n$ as $\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\}$.

Now it follows that for any vector field X on $\mathbb{R}^n$, since $X_p \in T_p\mathbb{R}^n$, we can write

$$X_p = X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p,$$

where each $X^i : \mathbb{R}^n \to \mathbb{R}$. More succinctly,

$$X = X^i \frac{\partial}{\partial x^i}.$$

The functions $X^i : \mathbb{R}^n \to \mathbb{R}$ are called the **component functions of the vector field** X wrt the standard vector fields.

We are now ready to define smoothness of a vector field.

Definition 38 (Smooth Vector Fields)

Let X be a vector field on $\mathbb{R}^n$. Then $X = X^i \frac{\partial}{\partial x^i}$ for some uniquely determined function $X^i : \mathbb{R}^n \to \mathbb{R}$. We say that X is ***smooth*** *if X^i is smooth for every i. We write $X^i \in C^\infty(\mathbb{R}^n)$.*

Remark

In multivariable calculus, a smooth field on $\mathbb{R}^n$ is a smooth map $X : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$X(p) = (X^1(p), \ldots, X^n(p)),$$

i.e. we could say that $X = (X^1, \ldots, X^n)$ is an n-tuple of smooth functions on $\mathbb{R}^n$.

Note that this view is particular to $\mathbb{R}^n$ due to the canonical isomorphism between $T_p\mathbb{R}^n$ and $\mathbb{R}^n$ for all $p \in \mathbb{R}^n$.

A Review of Earlier Contents

A.1 Rank-Nullity Theorem

Definition A.1 (Kernel and Image)

Let V and W be vector spaces, and let $T \in L(V, W)$. The ***kernel*** *(or* ***null space****) of T is defined as*

$$\ker(T) := \{v \in V \mid Tv = 0\},$$

i.e. the set of vectors in V such that they are mapped to 0 under T.

The ***image*** *(or* ***range****) of T is defined as*

$$\mathrm{Img}(T) = \{Tv \mid v \in V\},$$

that is the set of all images of vectors of V under T.

It can be shown that for a linear map $T \in L(V, W)$, $\ker(T)$ and $\mathrm{Img}(T)$ are subspaces of V and W, respectively. As such, we can define the following:

Definition A.2 (Rank and Nullity)

Let V, W be vector spaces, and let $T \in L(V, W)$. If $\ker(T)$ and $\mathrm{Img}(T)$ are finite-dimensional [1], then we define the ***nullity*** *of T as*

$$\mathrm{nullity}(T) := \dim \ker(T),$$

[1] In this course, this is always the case, since we are only dealing with finite dimensional real vector spaces.

and the ***rank*** *of T as*

$$\operatorname{rank}(T) := \dim \operatorname{Img}(T).$$

Note

From the action of a linear transformation, we observe that the ***larger the nullity, the smaller the rank****. Put in another way, the more vectors are sent to* 0 *by the linear transformation, the smaller the range.*

Similarly, the larger the rank, the smaller the nullity.

This observation gives us the Rank-Nullity Theorem.

Theorem A.1 (Rank-Nullity Theorem)

Let V and W be vector spaces, and $T \in L(V, W)$. If V is finie-dimensional, then

$$\operatorname{nullity}(T) + \operatorname{rank}(T) = \dim(V).$$

From the Rank-Nullity Theorem, we can make the following observations about the relationships between injection and surjection, and the nullity and rank.

Proposition A.2 (Nullity of Only 0 and Injectivity)

Let V and W be vector spaces, and $T \in L(V, W)$. Then T is injective iff $\operatorname{nullity}(T) = \{0\}$.

Surjection and injectivity come hand-in-hand when we have the following special case.

Proposition A.3 (When Rank Equals The Dimension of the Space)

Let V and W be vector spaces of equal (finite) dimension, and let $T \in L(V, W)$. TFAE

1. *T is injective;*
2. *T is surjective;*
3. $\operatorname{rank}(T) = \dim(V)$.

Note that the proof for Proposition A.3 requires the understanding that $\ker(T) = \{0\}$ implies that $\operatorname{nullity}(T) = 0$. See this explanation on Math SE.

Bibliography

Friedberg, S. H., Insel, A. J., and Spence, L. E. (2002). *Linear Algebra*. Pearson Education, 4th edition.

Karigiannis, S. (2019). *PMATH 365: Differential Geometry (Winter 2019)*. University of Waterloo.

Index