

Contents

Chapter 0: Banach Spaces	3
Introduction	4
Banach Spaces	4
Vector Spaces	5
Enumeration of lists	5
Chapter 1: Multilinear maps	7
Bilinear maps	8
k -linear maps	9
Chapter 2: Differentiation	12
The derivative	13
Chapter 3: Integration	17
Introduction	18
Integration of step mappings	18
Product of step mappings	19
Regulated mappings	20
Fundamental Theorem of Calculus	21
Mean value theorems	22
Chapter 4: Higher order derivatives	24
Introduction	25
The second derivative	25
The p -th derivatives	29
Chapter 1: Manifolds	32
Introduction	33
The structure of a manifold	33
Morphisms between manifolds	35
Tangent spaces	37
Curves	39
Splitting	40
Submanifolds	41
Vector Bundles	43

Chapter 2: Coordinates	47
Introduction	48
Derivations	48
Boundary	49

Chapter 0: Banach Spaces

Introduction

This section is quite incomplete, and all over the place. I have been meaning to put all the notation/terminology I am going to use in this section. Please skip to the Chapter 1 for now.

Banach Spaces

A *Banach space* is a normed vector space that is Cauchy-complete under the usual metric induced by its norm.

If E and F are Banach spaces over \mathbb{R} . We will denote the norms on E , and F by single lines, so

$$|x| = \|x\|_E \quad \text{and} \quad |y| = \|y\|_F \quad \forall x \in E, y \in F$$

$\mathcal{L}(E, F)$ will denote the space of linear maps between E and F . In the category of Banach spaces, the space of morphisms are called *toplinear morphisms* - or *CLMs* (*continuous linear maps*); which we will denote by $L(E, F)$ for toplinear morphisms between E and F .

We use $\|\cdot\|_{L(E, F)}$ or $\|\cdot\|$ to denote the operator norm, depending on how much emphasis we wish to place on $L(E, F)$. Recall,

$$\begin{aligned} \|\varphi\|_{L(E, F)} &= \inf \left\{ A \geq 0, |\varphi(x)| \leq A|x| \forall x \in E \right\} \\ &= \sup \left\{ |\varphi(x)|, x \in E, |x| = 1 \right\} \end{aligned}$$

By the open mapping theorem: any continuous surjective linear map is an open map. Hence invertible elements in $L(E, F)$ are naturally called *toplinear isomorphisms*. If $\varphi \in L(E, F)$ such that φ preserves the norm between the Banach Spaces, that is for every $x \in E$, $|x| = |\varphi(x)|$ then we call φ an *isometry*, or a *Banach space isomorphism*. If E_1 and E_2 are Banach spaces, we will use the usual *product norm* $(x_1, x_2) \mapsto \max(|x_1|, |x_2|)$.

- We say a map F is *between* the spaces X and Y if $F : X \rightarrow Y$.
- $\mathcal{L}(V^k, W)$ denotes the space of k -linear maps from V to W that are not necessarily continuous.

Proposition 2.1: Hahn Banach Theorem (Geometric Form)

Let E be a Banach space, A and B are closed disjoint subsets of E . Assuming one of the two is compact, then there exists a *clf* λ which *strictly separates* A and B .

$$A \subseteq [\lambda \leq \alpha - \varepsilon] \quad \text{and} \quad B \subseteq [\lambda \geq \alpha + \varepsilon] \tag{1}$$

for $\alpha \in \mathbb{R}$ and $\varepsilon > 0$.

Definition 2.1: Product of Banach Spaces

Let E_1, \dots, E_k be Banach spaces over \mathbb{R} . The Cartesian product of (E_1, \dots, E_k) is denoted by $\prod_i^k E_i$.

It is again a Banach space with the norm

$$(x_1, \dots, x_k) \mapsto |(x_1, \dots, x_k)| = \sup_{1 \leq i \leq k} |x_i| \quad (2)$$

Vector Spaces

Let V be any vector space over \mathbb{R} or \mathbb{C} , and $\{v_\alpha\} \subseteq V$, the symbol $\sum^\wedge v_\alpha$ refers to a partially specified object which is any **finite** linear combination of the elements of $\{v_\alpha\}$. If the cardinality of $\{v_\alpha\}$ is finite,

$$\sum^\wedge v_\alpha = \sum^\wedge v_{\underline{k}} \text{ for some } k \geq 1. \quad (3)$$

where eq. (3) should be interpreted as eq. (4)

$$\sum^\wedge v_{\underline{k}} = \sum_{i=\underline{k}} c^i v_i \quad (4)$$

for some $c^i \in \mathbb{R}$ or \mathbb{C} where $i = \underline{k}$.

Definition 3.1: x is essentially in W_1

If V is the vector space direct sum of W_1 and W_2 , a vector $x \in W_1$ is *essentially in W_i* if it is invariant under the canonical projection of $\pi_i V \rightarrow W_i$. That is,

$$\pi_i(x) = x$$

equivalently, the x is expressed as the linear combination of $x + 0 \in W_1 \oplus W_2$.

Composition of maps: If $f : E \rightarrow F$ and $g : F \rightarrow G$ are maps between Banach spaces, we write gf to mean $g \circ f$.

Enumeration of lists

We use the following notation to simplify computations concerning multilinear maps. Let E and F be sets, elements $v_1, \dots, v_k \in E$, and a map $f : E \rightarrow F$.

- Listing individual elements: $v_{\underline{k}}$ means v_1, \dots, v_k as separate elements.
- Creating a k -list: $(v_{\underline{k}}) = (v_1, \dots, v_k) \in \prod E_{j \leq k}$ if $v_i \in E_i$ for $i = \underline{k}$.
- Double indices: $(v_{\underline{n_k}}) = (v_{\underline{n_k}}) = (v_{n_1}, \dots, v_{n_k})$, and

$$(v_{\underline{n_k}}) \neq (v_{n_{(1, \dots, k)}})$$

- Closest bracket convention:

$$(v_{(n_{\underline{k}})}) = (v_{(n_1, \dots, n_k)}) \quad \text{and} \quad (v_{n_{(\underline{k})}}) = (v_{n_{(1, \dots, k)}})$$

- Underlining 0 means it is iterated 0 times:

$$(v_{\underline{0}}, a, b, c) = (a, b, c)$$

- Skipping an index:

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

for $i = \underline{k}$.

- Applying f to a particular index:

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

Of course, if $i = 1$, then the above expression reads $(f(v_1), v_2, \dots, v_k)$ by the $\underline{0}$ interpretation.

- In any list using this 'underline' notation, we can find the size of a list by summing over all the underlined terms, and the number of terms with no underline.
- If $\wedge : E \times E \rightarrow F$ is any associative binary operation,

$$\bigcirc(\wedge)(v_{\underline{k}}) = v_1 \wedge \dots \wedge v_k$$

Remark 4.1: Preview of exterior calculus

We can write the formula for the determinant of a $\mathbb{R}^{k \times k}$ matrix in this notation. Suppose $a_i \in \mathbb{R}$, and $b_i \in \mathbb{R}^{k-1}$ for $i = \underline{k}$.

$$M = \begin{bmatrix} a_1 & \dots & a_k \\ | & & | \\ b_1 & \dots & b_k \\ | & & | \end{bmatrix}$$

The determinant of M is a linear combination of determinants of $k-1$ -sized matrices, given in terms of the columns of b

$$\det(M) = \sum_{i=\underline{k}} (-1)^{i-1} a_i \det(b_{\underline{i-1}}, b_{i+\underline{k-i}})$$

In general, the 'hats' that we will use are left-associative. Meaning

$$\hat{x} = \hat{x} \quad \text{and} \quad \tau_y f^\wedge = \widehat{(\tau_y f)}$$

Chapter 1: Multilinear maps

Bilinear maps

Definition 1.1: Bilinear map

A map $\varphi : E_1 \times E_2 \rightarrow F$, where F is also a Banach space, is said to be *bilinear* if

$$\varphi(x, \cdot) : E_2 \rightarrow F \quad \text{and} \quad \varphi(\cdot, y) : E_1 \rightarrow F$$

are linear for every $x \in E_1$ and $y \in E_2$.

Proposition 1.1: Continuity criterion of a bilinear map

Let E_1, E_2, F be Banach spaces, a bilinear map $m : E_1 \times E_2 \rightarrow F$ is continuous if and only if there exists a $C \geq 0$, where

$$|m(x, y)| \leq C|x||y| \quad (5)$$

Proof. Suppose such a C exists, fix a convergent sequence $(x_n, y_n) \rightarrow (x, y)$ in $E_1 \times E_2 = E$. Because the projection maps are continuous, this means $x_n \rightarrow x$ and $y_n \rightarrow y$. Using inspiration from the proof where $x_n y_n \rightarrow xy$, where

$$x_n(y_n - y) + (x_n - x)y = x_n y_n - xy \quad x, y, x_n, y_n \in \mathbb{R}$$

Using the inspiration, and replacing multiplication in \mathbb{R} with the bilinear map m , we have:

$$\begin{aligned} m(x_n, y_n - y) + m(x_n - x, y) &= m(x_n, y_n) - m(x, y) \\ |m(x_n, y_n) - m(x, y)| &\leq C[|x_n| \cdot |y_n - y| + |x_n - x| \cdot |y|] \rightarrow 0 \end{aligned}$$

Conversely, if m is continuous, then it is continuous at the origin $(0, 0) = 0$. There exists a δ where $|(x, y)| \leq \delta$ implies $|m(x, y)| \leq 1$. Now, if $x, y \neq 0$ are elements in E , we normalize so that (x, y) has length δ

$$|(x|x|^{-1}\delta, y|y|^{-1}\delta)| = \delta|(x|x|^{-1}, y|y|^{-1})| = \delta$$

So that $|m(x|x|^{-1}\delta, y|y|^{-1}\delta)| \leq 1$, using bilinearity of m :

$$|m(x, y)| \leq \delta^{-2}|x| \cdot |y|$$

Setting $\delta^{-2} = C$ finishes the proof (notice if either x or y is 0, then m is trivially 0 and the inequality holds). ■

Proposition 1.2: $L(E_1, E_2; F)$ is isomorphic to $L(E_1, L(E_2, F))$

For each bilinear map $\omega \in L(E_1, E_2; F)$, there exists a unique map $\varphi_\omega \in L(E_1, L(E_2, F))$ such that $|\omega| = |\varphi_\omega|$; such that for every $(x, y) \in E_1 \times E_2$, $\omega(x, y) = \varphi_\omega(x)(y)$.

Proof. Let $\varphi_\omega : E_1 \rightarrow L(E_2, F)$ be the unique map such that $\varphi_\omega(x)(y) = \omega(x, y)$. Proposition 1.1 shows that $\varphi_\omega(x)$ is a continuous linear map into F at each x , and $|\varphi_\omega(x)| \leq |\omega||x|$. This holds for an arbitrary x , and $\varphi_\omega(\cdot)$ is clearly linear, hence $|\varphi_\omega| \leq |\omega|$. Reversing the roles of ω and φ shows proves the other

estimate.

The rule as outlined above is linear in ω ; and it is not hard to see $\varphi : L(E_1, E_2; F) \rightarrow L(E_1, L(E_2, F))$ is an injection. By the open mapping theorem, the proposition is proven if φ is a surjection. Fix $\theta \in L(E_1, L(E_2, F))$, define a map $\omega : E_1 \times E_2 \rightarrow F$ such that $\omega(x, \cdot) = \theta(x)(\cdot)$. So that ω is linear in its second argument. To show ω is linear in its first: fix a linear combination $A = \sum^\wedge x$ in E_1 , and $y \in E_2$.

$$\omega(A, y) = \theta(\sum^\wedge x)(y) = \sum^\wedge \theta(x)(y) = \sum^\wedge \omega(x, y)$$

Continuity follows from Equation (5), and $\varphi_\omega = \theta$ as needed. ■

k -linear maps

Definition 2.1: k -linear maps

Let $E_{\underline{k}}, F$ be Banach spaces. A map $\varphi : \prod E_{\underline{k}}$ is k -linear if for every $i = \underline{k}$, $v_i \in E_i$,

$$\varphi(\cdot \frac{i-1}{i}, v_i, \cdot \frac{k-i}{k-i}) : \bigoplus (E_{i-1}, E_{i+k-i}) \rightarrow F \text{ is } (k-1)\text{-linear}$$

A k -linear *symmetric* map between Banach spaces E, F is a map $A \in \mathcal{L}(E^k, F)$ such that for every k -permutation $\theta \in S_{\underline{k}}$,

$$A(v_{\underline{k}}) = A(v_{\theta(\underline{k})})$$

The following theorem should give confidence to the notation we have adopted to use.

Proposition 2.1: Continuity criterion of k -linear maps

Let $E_{\underline{k}}$ and F be Banach spaces, a k -linear map $\varphi : \prod E_{\underline{k}} \rightarrow F$ is continuous iff there exists a $C > 0$, such that for every $x_i \in E_i$, $i = \underline{k}$

$$|\varphi(x_{\underline{k}})| \leq C \prod |x_{\underline{k}}|$$

Proof. Suppose φ is continuous, then it is continuous at the origin. Picking $\varepsilon = 1$ induces a $\delta > 0$ such that for $|(x_{\underline{k}})| \leq \delta$, $|\varphi(x_{\underline{k}})| \leq 1$. The usual trick of normalizing an arbitrary vector $(x_{\underline{k}}) \in \prod E_{\underline{k}}$ does the job:

$$\left| \varphi(x_{\underline{k}} \cdot |x_{\underline{k}}|^{-1} \cdot \delta) \right| \leq 1 \implies |\varphi(x_{\underline{k}})| \leq \delta^{-k} \prod |x_{\underline{k}}|$$

Conversely, fix a sequence (indexed by n , in k elements in the product space $\prod E_{\underline{k}}$), so

$$(x_{\underline{n}}^k) \rightarrow (x_{\underline{k}}^k) \text{ as } n \rightarrow +\infty \tag{6}$$

To proceed any further, we need to prove an important equation that decomposes a difference in φ .

$$\varphi(b^{\underline{k}}) - \varphi(a^{\underline{k}}) = \sum_{i=\underline{k}} \varphi(b^{\underline{i}-1}, \Delta_i, a^{\underline{i}+k-i}) \tag{7}$$

where $(b^{\underline{k}})$ and $(a^{\underline{k}})$ are elements in $\prod E_{\underline{k}}$, and $\Delta_i = b^i - a^i$ for $i = \underline{k}$. The proof is in the following note, which is in more detail than usual - to help the reader ease into the new notation.

Note 2.1

We proceed by induction, and eq. (7) follows by setting $m = k$ in

$$\varphi(a^{\underline{k}}) = \varphi(b^{\underline{m}}, a^{m+k-\underline{m}}) - \sum_{i=\underline{m}} \varphi(b^{i-1}, \Delta_i, a^{i+k-i}) \quad (8)$$

Base case: set $m = 1$, by definition of k -linearity (def. 2.1) of φ . Since $a^1 = b^1 - \Delta_1$,

$$\varphi(a^{\underline{k}}) = \varphi(b^1 - \Delta_1, a^{1+k-1}) = \varphi(b^1, a^{1+k-1}) - \varphi(\Delta_1, a^{1+k-1})$$

Induction hypothesis: suppose eq. (8) holds for a fixed m . Since $a^{m+1} = b^{m+1} - \Delta_{m+1}$,

$$\begin{aligned} \varphi(a^{\underline{k}}) &= \varphi(b^{\underline{m}}, a^{m+k-\underline{m}}) - \sum_{i=\underline{m}} \varphi(b^{i-1}, \Delta_i, a^{i+k-i}) \\ &= \varphi(b^{\underline{m}}, a^{m+1}, a^{(m+1)+k-(m+1)}) - \sum_{i=\underline{m}} \varphi(b^{i-1}, \Delta_i, a^{i+k-i}) \\ &= \varphi(b^{m+1}, a^{(m+1)+k-(m+1)}) - \varphi(b^{m+1}, \Delta_{m+1}, a^{(m+1)+k-(m+1)}) - \sum_{i=\underline{m}} \varphi(b^{i-1}, \Delta_i, a^{i+k-i}) \end{aligned}$$

and this proves eq. (7)

We substitute $a^i = x^i$, and $b^i = x_n^i$ for $i = \underline{k}$, and eq. (7) becomes eq. (9)

$$\varphi(x_n^{\underline{k}}) - \varphi(x^{\underline{k}}) = \sum_{i=\underline{k}} \varphi(x_n^{i-1}, x_n^i - x^i, x^{i+k-i}) \quad (9)$$

Then the triangle inequality reads

$$\begin{aligned} \left| \varphi(x_n^{\underline{k}}) - \varphi(x^{\underline{k}}) \right| &\leq \sum_{i=\underline{k}} \left| \varphi(x_n^{i-1}, x_n^i - x^i, x^{i+k-i}) \right| \\ &\leq \sum_{i=\underline{k}} |\varphi| \cdot \bigoplus \left(x_n^{i-1}, \Delta_i, x^{i+k-i} \right) \\ &\leq \sum_{i=\underline{k}} |\varphi| \cdot |x_n^i - x^i| \bigoplus \left(x_n^{i-1}, x^{i+k-i} \right) \\ &\lesssim_n |\varphi| \sup_{i=\underline{k}} |x_n^i - x^i| \rightarrow 0 \end{aligned}$$

where we identify the product $\bigoplus(v^{\underline{k}})$ with the product of their norms $\bigoplus(|v^{\underline{k}}|)$. ■

Remark 2.1: Currying isomorphism

The k -linear variant of prop. 1.2 holds. We will use but not prove this fact.

Remark 2.2: k -linear maps from the same space

We denote the space of k -linear maps from E into F by $L(E_{\underline{k}}; F) = L(E^{\underline{k}}, F) = L^k(E, F)$. *Tensors* on

E are k -linear maps from the product space of E into \mathbb{R} , by replacing F with \mathbb{R} .

Chapter 2: Differentiation

The derivative

Definition 1.1: Open sets and neighbourhoods

If U is an open subset of a topological space X , we denote this by $U \subseteq X$. If U is a *neighbourhood* of a point $p \in X$, we write $p \in U$.

We do not require neighbourhoods to be open sets; rather, we say U is a neighbourhood of p when the interior of U contains p .

Definition 1.2: Little o

A real-valued function in a real variable defined for all t sufficiently small is said to be $o(t)$ if $\lim_{t \rightarrow 0} o(t)/t = 0$. A map $\psi : U \rightarrow F$ where $U \subseteq E$ contains 0 in E , is said to be $o(h)$ if $|\psi(h)|/|h| \rightarrow 0$ as $h \rightarrow 0$ in E .

Definition 1.3: Differentiability

Let $f : E \rightarrow F$ be a map, replacing E and F by their open subsets if necessary. We say f is *differentiable* at $x \in E$ when there exists a **continuous linear map on E** : $\lambda \in L(E, F)$ such that

$$f(x + h) = f(x) + \lambda h + o(h) \quad \text{for sufficiently small } h \quad (10)$$

The role $o(h)$ plays here is a map from $U \rightarrow F$, where U is some neighbourhood of 0.

Proposition 1.1: Basic properties of the derivative

If f is differentiable at x , then the λ in eq. (10) is unique. We write $f'(x) = Df(x) = \lambda$ as in ?? . Furthermore, if $f'(x)$ and $g'(x)$ exist, then $(f + g)'(x) = f'(x) + g'(x)$ as linear maps, similar for scalar multiplication.

Proof. Suppose $\lambda_i \in L(E, F)$ are both derivatives of f at x . Then,

$$\begin{cases} f(x + h) = f(x) + \lambda_1(h) + o(h) \\ f(x + h) = f(x) + \lambda_2(h) + o(h) \end{cases}$$

And $(\lambda_1 - \lambda_2)(h) = o(h) = \varphi(h) \cdot |h|$, where $\varphi(h) \rightarrow 0$ as $h \rightarrow 0$. Using the operator norm, we see that

$$\|\lambda_1 - \lambda_2\|_{L(E, F)} \leq |\varphi(h)| \rightarrow 0$$

This proves uniqueness. Suppose f and g are differentiable at x , denote $\lambda_f = f'(x)$ (resp. $g'(x)$). The definition of def. 1.3 reads

$$\begin{aligned} f(x + h) + g(x + h) &= (f(x) + g(x)) + (\lambda_f(h) + \lambda_g(h)) + o(h) + o(h) \\ (f + g)(x + h) &= (f + g)(x) + (\lambda_f + \lambda_g)(h) + o(h) \end{aligned} \quad (11)$$

since eq. (11) satisfies eq. (10), the proof is complete. ■

Proposition 1.2: Chain rule

Let E, F, G be Banach spaces. If $f \in C^1(E, F)$, $g \in C^1(F, G)$, for every $x \in E$,

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x) \quad (12)$$

Proof. Since f is differentiable at x , $f(x + h) = f(x) + f'(x)(h) + o_1(h)$, (resp. for g , $o_2(h)$). Set $k(h) = f(x + h) - f(x)$, and

$$\begin{aligned} g(f(x + h)) &= g(f(x)) + g'(f(x))(k(h)) + o_2(k(h)) \\ &= g(f(x)) + g'(f(x))(f'(x)(h) + o_1(h)) + o_2(k(h)) \\ (g \circ f)(x + h) &= (g \circ f)(x) + g'(f(x)) \circ f'(x)(h) + g'(f(x))(o_1(h)) + o_2(k(h)) \\ (g \circ f)(x + h) &= (g \circ f)(x) + g'(f(x)) \circ f'(x)(h) + o(h) \end{aligned}$$

because $|A(o_1(h))| \leq |A||o_1(h)|$ for all $A \in L(E, F)$; and $o(k(h)) = o(h)$ for every continuous $k : E \rightarrow F$ such that $k(h) \rightarrow 0$ as $h \rightarrow 0$. ■

Proposition 1.3: Derivatives of CLMs

If $\lambda \in L(E, F)$, then $\lambda \in C^1(E, F)$ and $D\lambda(x) = \lambda$ for every $x \in E$. Furthermore, if $f \in C^1(E, F)$, and $\nu \in L(F, G)$, then the composition $\nu \circ f$ is in $C^1(E, G)$, and $(\nu \circ f)'(x) = \nu \circ f'(x)$ for every $x \in E$.

Proof. See $\lambda(x + h) = \lambda(x) + \lambda(h) + 0$ at every $x \in E$. Using the chain rule (prop. 1.2) proves the second claim. ■

Proposition 1.4: Product rule in k variables

Let $m : \prod F_{\underline{k}} \rightarrow G$ be a k -linear map between Banach spaces $F_{\underline{k}}$ and G . Suppose $f_i \in C^1(E, F_i)$ with $i = \underline{k}$, writing

$$m(f_{\underline{k}})(x) = m(f_{\underline{k}}(x)) \quad (13)$$

then $m(f_{\underline{k}})$ is in $C^1(E, G)$ and for every $y \in E$,

$$Dm(f_{\underline{k}})(x)(y) = \sum_{i=\underline{k}} m(f_{\underline{i}-1}(x), Df_i(x)(y), f_{i+\underline{k}-i}(x)) \quad (14)$$

Proof. Let x be fixed. Equation (14) is proven if we show eq. (15)

$$m(f_{\underline{k}})(x + h) = m(f_{\underline{k}})(x) + \left(\sum_{i=\underline{k}} m(f_{\underline{i}-1}(x), Df_i(x)(h), f_{i+\underline{k}-i}(x)) \right) + o(h) \quad (15)$$

and for sufficiently small h we have

$$f_i(x + h) - f_i(x) = Df_i(x)(h) + o(h^i) \quad (16)$$

We will use the difference formula in eq. (8), with the following substitutions

$$f_i(x + h) = b^i \quad f_i(x) = a^i \quad (17)$$

$$Df_i(x)(h) = c^i \quad o(h^i) = \varepsilon^i \quad (18)$$

$$f_i(x + h) - f_i(x) = c^i + \varepsilon^i \quad \Delta^i = o(h^i) + c^i \quad (19)$$

With these substitutions, the equation we want to prove (eq. (14)) becomes eq. (20)

$$m(b^{\underline{k}}) - m(a^{\underline{k}}) = \left(\sum_{i=\underline{k}} m(a^{i-1}, c^i, a^{i+k-i}) \right) + o(h) \quad (20)$$

Starting from eq. (8),

$$m(b^{\underline{k}}) - m(a^{\underline{k}}) = \sum_{i=\underline{k}} m(b^{i-1}, \Delta^i, a^{i+k-i})$$

We can expand each term, if $i = \underline{k}$,

$$m(b^{i-1}, \Delta^i, a^{i+k-i}) = m(b^{i-1}, c^i, a^{i+k-i}) + m(b^{i-1}, o(h^i), a^{i+k-i}) \quad (21)$$

Let us study the first term in eq. (21), and with i held fixed, define

$$m_i(z^{i-1}) = m(z^{i-1}, c_i, a^{i+k-i}) \quad (22)$$

Expanding the first term within eq. (21), and because m_i as defined in eq. (22) is $i-1$ -linear (because it is a k -linear map with $k - (i-1)$ variables held constant); we use eq. (8) again.

$$m_i(b^{i-1}) = \left(\sum_{j=\underline{k}} m_i(b^j, \Delta^j, a^{j+(i-1)-j}) \right) + m_i(a^{i-1}) \quad (23)$$

Unboxing the last term in eq. (23) using the definition of m_i reads

$$m(b^{i-1}, \Delta^i, a^{i+k-i}) = m(a^{i-1}, c^i, a^{i+k-i}) + \sum_{j=i-1} m_i(b^j, \Delta^j, a^{j+(i-1)-j}) \quad (24)$$

We wish to remove all of the b^i 's. Since $\Delta^i = c^i + \varepsilon^i$ (eq. (19)), we have

$$\begin{aligned} m(b^{\underline{k}}) - m(a^{\underline{k}}) &= \sum_{i=\underline{k}} m(b^{i-1}, c^i, a^{i+k-i}) + m(b^{i-1}, \varepsilon^i, a^{i+k-i}) \\ &= \left(\sum_{i=\underline{k}} m_i(b^{i-1}) \right) + \sum_{i=\underline{k}} m(b^{i-1}, \varepsilon^i, a^{i+k-i}) \\ &= \left(\sum_{i=\underline{k}} m_i(a^{i-1}) + \sum_{j=i-1} m_i(b^{j-1}, \Delta^j, a^{j+(i-1)-j}) \right) + \sum_{i=\underline{k}} m(b^{i-1}, \varepsilon^i, a^{i+k-i}) \\ &= \left(\sum_{i=\underline{k}} m_i(a^{i-1}) \right) + \sum_{\substack{i=\underline{k} \\ j=i-1}} m_i(b^{j-1}, \Delta^j, a^{j+(i-1)-j}) + \sum_{i=\underline{k}} m(b^{i-1}, \varepsilon^i, a^{i+k-i}) \end{aligned} \quad (25)$$

The last term within eq. (25) is $o(h)$, since it is a linear combination of $o(h^i)$'s.

$$\left| \sum_{i=\underline{k}} m(b^{i-1}, \varepsilon^i, a^{i+k-i}) \right| \lesssim_{m,a,b} |o(h)| \quad (26)$$

Each summand in the second last term in eq. (25) is $o(h)$ as well, as

$$\begin{aligned}
 |m_i(b^{j-1}, \Delta^j, a^{j+(i-1)-j})| &\leq |m_i| \left(\prod (b^{j-1}, \Delta^j, a^{j+(i-1)-j}) \right) \\
 &\leq |m| \cdot \left(\prod (c^i, a^{i+k-i}) \right) \left(\prod (b^{j-1}, \Delta^j, a^{j+(i-1)-j}) \right) \\
 &\lesssim_{m,a,b} \sup_{\substack{i=k \\ j=i-1}} |c^i| \cdot |\Delta^j| \\
 &\lesssim_{m,a,b} \sup_{\substack{i=k \\ j=i-1}} |Df_i(x)(h)| \cdot |f_j(x+h) - f_j(x)| \\
 &\lesssim_{m,a,b} |Df_i(x)||h| \sup_{\substack{i=k \\ j=i-1}} |\Delta^j| \\
 &\lesssim_{m,a,b} |o(h)|
 \end{aligned} \tag{27}$$

for the second last estimate we used $\Delta^j \rightarrow 0$. Therefore the second term in eq. (25) is $o(h)$, and eq. (15) is proven. Therefore $m(f_k)$ is differentiable at x . Continuity of $Dm(f_k)$ follows from the fact that

$$Dm(f_k)(x) = \sum_{i=k} m(f_{i-1}(x), Df_i(x)(\cdot), f_{i+k-i}(x)) \tag{28}$$

and each of the summands eq. (28) can be broken down as the product of the compositions shown in eqs. (29) and (30)

$$x \mapsto (f_{i-1}(x), f_{i+k-i}(x)) \mapsto m(f_{i-1}(x), \cdot, f_{i+k-i}(x)) \tag{29}$$

$$x \mapsto Df_i(x)(\cdot) \tag{30}$$

which are continuous from E to $L(E, F)$. ■

Chapter 3: Integration

Introduction

This chapter will be on the integration of *regulated* mappings, the space of which are precisely the uniform closure of rectangle functions. from a compact interval. We will go through some of the elementary results, and prove the Fundamental Theorem.

Integration of step mappings

Definition 2.1: Partition on $[a, b]$

Let $I = [a, b]$ be a compact interval. An N -partition P on I is a list of $N + 1$ elements in $[a, b]$, which are assumed to be well ordered as in $p_0 \leq p_1 \leq \dots \leq p_N$.

$$P = (a = p_0, p_1, \dots, p_N = b) \quad \text{or} \quad P = (p_0, \underline{p_N}) \quad (31)$$

The space of partitions on I will be denoted by I_p .

As per usual, we have *common refinements of partitions*, given two partitions P and Q on the same compact interval $I = [a, b]$, where P is defined as in eq. (31), and $Q = (q_0, \underline{q_N})$ similarly. The common refinement of P and Q is another partition R on I which contains all of the elements in $P \cup Q$.

- Given a partition P of size N represented as $P = (p_0, \underline{p_N})$, the cells of P are indexed using their rightmost points.
- The interval (p_{i-1}, p_i) is denoted as $\text{cell}(p_i)$, and
- the *length* of the i th cell: $|\text{cell } p_i| = |p_i - p_{i-1}|$.
- If we want to sequence the cells of P based on their right endpoints, it is expressed as $\text{cell}(P) = (\text{cell}(p_{\underline{N}}))$.
- Note that these cells do not form a disjoint union of I .

Remark 2.1: Assume all intervals are compact

For the rest of this chapter, we assume all intervals are compact and of the form $I = [a, b]$. If P, Q, R are partitions, their elements will be represented by p_i , (resp. r_i, q_i).

Definition 2.2: Step mapping

A step mapping on $I = [a, b]$ is a vector space of maps from I to a Banach space E over \mathbb{R} . It is equipped with the supremum norm, and its elements are denoted by Σ ,

$$\Sigma = \left\{ f : [a, b] \rightarrow E, \text{ there exists a } N\text{-partition } P \in I_p, \{v_{\underline{N}}\} \subseteq E \text{ such that } f|_{(p_{i-1}, p_i)} = v_i \forall i = \underline{N} \right\} \quad (32)$$

If $f \in \Sigma$, we denote its norm by $\|f\|_u = \sup_{x \in I} |f(x)|$.

Definition 2.3: Integration on Σ

If $f \in \Sigma$ and is of the form inside the set-builder notation in eq. (32), we define the integral of f by

$$\int_a^b f = \sum_{i=\underline{N}} (p_i - p_{i-1})v_i \quad (33)$$

Remark 2.2: Distinguishing between intervals I, J

If I and J are compact intervals, we distinguish the step mappings from I and J by Σ_I and Σ_J .

We now state some definition and properties of eq. (33) which we will not prove.

Proposition 2.1: Properties of the integral on Σ

Let I and J be intervals, $f, f_{\underline{k}} \in \Sigma_I$, and $g \in \Sigma_J$.

- The integral is linear, that is

$$\int \sum^{\wedge} f_{\underline{k}} = \sum^{\wedge} \int f_{\underline{k}} \quad (34)$$

- The integral over $[b, a]$ is *defined* to be the negative of eq. (33):

$$\int_a^b f = - \int_b^a f \quad (35)$$

- The integral is domain-additive, if $b = c$, then

$$\int_a^b f + \int_c^d g = \int_a^d (f + g) \quad (36)$$

where we identify $(f + g)$ to be the step mapping in $\Sigma_{[a,d]}$ whose restriction I (resp. J) agree with f (resp. g).

Product of step mappings

Let $E_{\underline{k}}$ be Banach spaces, and $I = [a, b]$ a fixed compact interval. Let E refer to the product space $\prod E_{\underline{k}}$, which is equipped with the supremum norm as outlined in def. 2.1

$$\Sigma_i = \left\{ f_i : I \rightarrow E_i, f_i \text{ is a step mapping.} \right\}$$

There are two ways of defining the space of step-mappings from I into E eqs. (37) and (38). Using a combinatorial argument with common refinements, it is not hard to see the two are subsets of each other.

$$\Sigma_E^1 = \left\{ f : I \rightarrow E, \text{proj}_i f \in \Sigma_i \forall i = \underline{k} \right\} \quad (37)$$

$$\Sigma_E^2 = \left\{ f : I \rightarrow E, f \text{ is a step mapping.} \right\} \quad (38)$$

And since the product space E is toplinearly isomorphic to its external direct sum, $E_1 \times \cdots \times E_k$, the integral over $\Sigma_E = \Sigma_E^1 = \Sigma_E^2$ is defined to be

$$\int_a^b f = \left(\int_a^b \text{proj}_{\underline{k}} f \right) = \left(\int_a^b \text{proj}_1 f, \dots, \int_a^b \text{proj}_k f \right) \quad (39)$$

Regulated mappings

Definition 4.1: Regulated mappings

Let I be a compact interval. A mapping from I into E is *regulated* if it is the uniform limit of step mappings. We denote the space of regulated mappings by $\overline{\Sigma}_I$ or $\overline{\Sigma}$.

Proposition 4.1: Continuity implies a regulated mapping

Every continuous function $f : I \rightarrow E$ is the uniform limit of step mappings in $\Sigma_I = \Sigma$.

Proof. Let $f \in C(I, E)$, the continuity of f is uniform; given $\varepsilon > 0$ there exists $\delta > 0$ where $|y - x| < \delta$ implies $|f(y) - f(x)| < \varepsilon$. δ induces a smallest integer $n \geq 1$ such that $p_n = a + n\delta > b$. Define $p_0 = a$ and $p_i = a + i\delta$, relabelling $p_n = b$, we see that $P = (p_0, p_n)$ is a partition.

We construct a step mapping by sampling values of f . Set $g|_{\text{cell}(p_i)} = f(p_i)$, $g(a) = f(a)$, $g(p_i) = f(p_i)$. Defining the endpoints is necessary, and g still remains a member of Σ_I by eq. (32). Each $x \in I \setminus P$ belongs in some $\text{cell}(p_i)$, of which $|p_i - x| < \delta$, and $g(x) = f(p_i)$ implies $|g(x) - f(x)| < \delta$. If x is in P , then $g(x) = f(x)$, and $\|f - g\|_u \leq +\varepsilon$. ■

Proposition 4.2: Integration of regulated mappings

Let $f : I \rightarrow E$ be continuous, if $\{f_n\} \subseteq \Sigma$ converges uniformly to f , then $\{\int_a^b f_n\}$ is Cauchy in E , whose limit we *define* to be $\int_a^b f$ — the integral of f . Furthermore,

1. For any regulated mapping $f : I \rightarrow E$,

$$\left| \int_a^b f \right| \leq \int_a^b |f| \leq (b - a) \|f\|_u \quad (40)$$

2. The integral on $\overline{\Sigma}$ (resp. $\overline{\Sigma}_I, \overline{\Sigma}_J$) satisfies all of the properties in prop. 2.1.

Proof. Let f be a step mapping on E , we wish to show eq. (40) holds. If f is induced by some n -partition P ,

$$\int_a^b f = \sum_{i=\underline{n}} |\text{cell}(p_i)| f(p_i) \leq \sum_{i=\underline{n}} |\text{cell}(p_i)| \|f(p_i)\| = \int_a^b |f| \quad (41)$$

The integral in eq. (41) should be interpreted as a Riemann integral on \mathbb{R} , and eq. (42) is immediate:

$$\int_a^b |f| \leq |b - a| \|f\|_u \quad (42)$$

Next, let $\{f_n\}_{n \geq 1}$ be a sequence of step mappings in I which converges uniformly to $f \in \bar{\Sigma}$. Equation (42) tells us the sequence of integrals is uniformly Cauchy, as

$$\left| \int_a^b f_m - \int_a^b f_n \right| \leq |b - a| \|f_m - f_n\|_u \quad (43)$$

Hence $\int_a^b f$ is well defined, eq. (40) and the properties listed in prop. 2.1 follow upon taking limits. ■

Proposition 4.3: Integration and clms

Let E and F be Banach spaces, and $\lambda \in L(E, F)$. For a fixed interval I , denote the space of step mappings from I to E (resp. F) by Σ_E (resp. Σ_F), and regulated mappings similarly. If $\{f_n\} \subseteq \Sigma_E$ converges uniformly to $f \in \bar{\Sigma}_E$, then $\{\lambda f_n\} \rightarrow \lambda f$ uniformly in $\bar{\Sigma}_F$. Moreover,

$$\lambda \left(\int_a^b f \right) = \int_a^b \lambda f \quad (44)$$

Proof. The map λ is Lipschitz between E and F , and it descends into a map between the vector spaces Σ_E and Σ_F by composition. If f is a step mapping, and $f|_{\text{cell}(p_i)} = v_i$ for $i = \underline{k}$; the composition of f with λ is again a step mapping $\lambda f|_{\text{cell}(p_i)} = \lambda v_i$.

It is not hard to see $\|\lambda f\|_u \leq |\lambda| \|f\|_u$, and

- λ is Lipschitz between E and F ,
- λ , when viewed as a map between Σ_E and Σ_F , is Lipschitz.

Computing the integral of $\lambda f \in \Sigma_F$,

$$\int_a^b \lambda f = \sum_{i=\underline{k}} |\text{cell}(p_i)| \lambda v_i = \lambda \left(\sum_{i=\underline{k}} |\text{cell}(p_i)| v_i \right) = \lambda \int_a^b f$$

proves eq. (44) for step mappings, and the general case follows from continuity. ■

Fundamental Theorem of Calculus

Proposition 5.1

Let I be a compact interval, and $f : I \rightarrow E$ be regulated. Defining $\varphi : I \rightarrow E$ as the *integral of f with basepoint a*

$$\varphi(t) = \int_a^t f \quad (45)$$

Then φ is differentiable where f is continuous, and if $t_0 \in I$ is such a point:

$$(D\varphi)(t_0) = f(t_0) \quad (46)$$

Remark 5.1: Identifications

The left hand side in eq. (46) should be thought of as a clm in $L(\mathbb{R}, E)$. We identify the point $f(t_0)$ as the map $t \mapsto t \cdot f(t_0)$.

Proof. Suppose f is continuous at t_0 . For all h sufficiently small, set $\varepsilon(h) = \sup_{|t-t_0| \leq h, t \in I} |f(t) - f(t_0)|$ as the modulus of continuity; where $\varepsilon(h) \rightarrow 0$ as $h \rightarrow 0$. Applying the well-known technique of estimating the integrand $f(t) = [f(t) - f(t_0)] + f(t_0)$, we have

$$\begin{aligned} \varphi(t_0 + h) - \varphi(t_0) &= \int_{t_0}^{t_0+h} f(t) dt \\ &= f(t_0) \cdot h + \int_{t_0}^{t_0+h} [f(t) - f(t_0)] dt \end{aligned} \quad (47)$$

The last term within eq. (47) is $o(h)$, and the proof is complete. ■

Mean value theorems

If $\lambda \in L(E, F)$, and $x \in E$, we write $\lambda \dot{x} = x \dot{\lambda}$. If $t \in \mathbb{R}$, and we want to think of x as the map $t \mapsto tx$, we will write $t \cdot x = x \cdot t = tx$ to emphasize the role that x plays. The duality pairing between $L(E, F) \times E \rightarrow F$ is bilinear and continuous. For any regulated mapping $\alpha : I \rightarrow L(E, F)$,

$$\int_a^b \alpha(t) \cdot x dt = \left(\int_a^b \alpha(t) dt \right) \cdot x \quad (48)$$

Furthermore, if $f \in C^1(I, E)$, we use the notation $f'(t)$ to refer to $Df(t)$; and we identify $f'(t)$ with an element in E ; while $Df(t)$ should be thought of as a mapping in $L(\mathbb{R}, E)$.

Lemma 6.1: Constant curves

If $\alpha \in C^1(I, E)$, $\alpha' = 0$, iff α is constant.

Proof. Suppose α' vanishes, and assume for contradiction there exists points $t_0 < t_1$ in I such that $\alpha(t_0) \neq \alpha(t_1)$. Hahn Banach gives us a clf $\lambda \in L(E, \mathbb{R})$ that strictly separates the two points. See prop. 2.1 for a refresher. The ordinary derivative of $\lambda \circ f$ is 0 everywhere which implies $\lambda \circ f$ is constant. The converse is trivial. ■

Lemma 6.2: FTC 2

Let $f \in C^1(I, E)$, then

$$f(b) - f(a) = \int_a^b f'(t) dt \quad (49)$$

where the integrand in eq. (49) is — rigourously speaking — a map $\mathbb{R} \rightarrow L(\mathbb{R}, E)$, but we treat $f'(t) \in E$.

Proof. Throughout this proof, we will treat $f' : \mathbb{R} \rightarrow E$. Because f' is continuous everywhere, it is regulated. Define $\varphi(t) = \int_a^t f'(t) dt$, by eq. (46):

$$\varphi'(t) - f'(t) \equiv 0$$

By lem. 6.1, it suffices to show $(\varphi - f)(t) = f(a)$ at any point $t \in [a, b]$. Take $t = a$, and $(\varphi(a) - f(a)) = 0$, so that

$$\varphi(t) = f(t) + f(a)$$

and eq. (49) follows. ■

Remark 6.1: Usefulness of FTC 2

lem. 6.2 is most useful when $[a, b] = [0, 1]$, and the f is a curve interpolating between a C^1 function evaluated two different points, as in prop. 6.1.

Proposition 6.1: MVT 1

Let $U \subseteq E$ and $x \in U$, $y \in E$. If the line segment $L = \{x + ty, 0 \leq t \leq 1\}$ is also contained in U (draw a picture), then eq. (50) holds.

$$f(x + y) = f(x) + \int_0^1 Df(x + ty)y dt = \left(\int_0^1 Df(x + ty) dt \right) \cdot y \quad (50)$$

Proof. The curve $g(t) = f(x + ty)$ is composed of $f \circ l(t)$, for $l(t) = x + ty$. It has derivative

$$g'(t) = Df(x + ty) \circ l'(t) = Df(x + ty) \circ (y \in L(\mathbb{R}, E))$$

By lem. 6.2, $g(1) - g(0) = \int_0^1 Df(x + ty) \cdot y dt$. Given $g(1) - g(0) = f(x + y) - f(x)$, the proof is complete. ■

Chapter 4: Higher order derivatives

Introduction

We start with the definition of $C^p(E, F)$. Let E and F be Banach Spaces, if $p \geq 1$ is an integer, we define the class C^p to be the set of maps which are p times differentiable, and $D^p f \in C(E, X)$, where

$$X = L(E, L(E, L(E, \dots F))) \text{ } p \text{ times } \xLeftrightarrow{\mathcal{L}} L(E^p, F)$$

Sometimes we replace E with an open subset $U \subseteq E$ if necessary, and we write $f \in C(U, F)$ if $D^p \in C(U, X)$. Note, even if $f \in C^1(U, F)$, Df is still a map from U into $L(E, F)$.

We will prove two major results in this section.

- The structure of the derivative $D^p f$, in particular, if $f \in C^p(E, F)$, then $D^p f(x)$ is a *symmetric multilinear map* in p arguments.
- Taylor's Theorem

The second derivative

Proposition 2.1: Product rule in 2 variables

Let E_1 , E_2 and F be Banach spaces, if $\omega : E_1 \times E_2 \rightarrow F$ is bilinear and continuous, then ω is differentiable, and for every $(x_1, x_2) \in E_1 \times E_2$, $(v_1, v_2) \in E_1 \times E_2$,

$$D\omega(x_1, x_2)(v_1, v_2) = \omega(x_1, v_2) + \omega(v_1, x_2)$$

Furthermore, $D^2\omega(x, y) = D\omega \in L(E^2, F)$, and $D^3\omega = 0$.

Proof. By the definition of ω , using the familiar interpolation method

$$\omega(x_1 + h_2, x_2 + h_2) = \omega(x_1, x_2) + \omega(x_1, h_2) + \omega(h_1, x_2) + \omega(h_1, h_2)$$

by continuity of ω , the last term (which we wish to make $o(h)$):

$$|\omega(h_1, h_2)| \leq \|\omega\| \cdot |(h_1, h_2)|^2$$

so that $\omega(h_1, h_2) = o(h)$, and $D\omega(x_1, x_2)$ exists and is continuous, and is given by the *linear map* $\omega(x_1, \cdot) + \omega(\cdot, x_2)$. The rest of the proof follows, if it is not immediately obvious then read the following note.

Note 2.1

Write $E = E_1 \times E_2$ for convenience. The linear map $A = D\omega(x_1, x_2)$ takes arguments E into F , consider the projections π_1 and π_2 , and $v \in E_1 \times E_2$, then

$$A(v) = \omega(x_1, \pi_1 v) + \omega(\pi_2 v, x_2)$$

We can view $A(x) = D\omega(x_1, x_2) \in L(E, F)$. It is clear that A is linear in x , if we fix $v \in E$,

$$A(x + y, v) = \omega(\pi_1(x + y), \pi_2 v) + \omega(\pi_1 v, \pi_2(x + y)) = A(x, v) + A(y, v)$$

and similarly for scalar multiplication. Hence $DA(x) = A \in L(E, L(E, F))$ and $D^2 A(x) = D^3 \omega = 0$.

Our next result is the following, which states that if $f : U \rightarrow F$ where $U \subseteq E$, and $Df, DDf = D^2f$ exists and are continuous maps from U into $L(E, F)$ and $L(E, L(E, F))$ respectively, then $D^2f(x)$ is a *symmetric bilinear map*. The proof is non-trivial, and relies on computing the 'Lie Bracket':

$$D^2f(x)(v, w) - D^2f(x)(w, v)$$

Which we will prove is equal to 0 for every $x \in U$, and $v, w \in E$.

Proposition 2.2: Second derivative is symmetric

Let $f \in C^2(U, F)$, where $U \subseteq E$ with the possibility that $U = E$. For every point $x \in U$, the *second derivative* $D^2f(x)$ is bilinear and symmetric.

Proof. Fix $x \in U \ni B(r) + x \subseteq U$. We restrict our attention to vectors $v, w \in E$ where $|v|, |w| < r2^{-1}$ for now, so that the

$$\{x, x + w, x + v, x + v + w\} \subseteq U$$

We will denote the following quantity by Δ

$$\Delta = f(x + w + v) - f(x + w) - f(x + v) + f(x)$$

By rearranging terms, we see that Δ can be approximated in two ways:

- Postponing the discussion about the the domain of y , set $g(y) = f(y + v) - f(y)$ is C^2 , and

$$\Delta = g(x + w) - g(x) \tag{51}$$

- Again, for y sufficiently close to x , define $h(y) = f(y + w) - f(y)$, and

$$\Delta = h(x + v) - h(x) \tag{52}$$

- To find the domain for y , an easy argument using the Triangle inequality gives us $g, h \in C^2(B(r2^{-1}) + x, F)$,
- Leaving the computations of h as an exercise, we compute Dg , recall the shift map $y \mapsto y + v$ commutes with D , and

$$Dg(y) = D(\tau_{-v}f)(y) - Df(y) = Df(y + v) - Df(y) \tag{53}$$

Using MVT twice, once on Equation (51) (the line segment $x + tw$, $0 \leq t \leq 1$ is contained in the domain of g), and another time on Equation (53) (with $y = x + tw$ in the integrand). We obtain:

$$\begin{aligned} \Delta &= g(x + w) - g(x) \\ &= \int_0^1 Dg(x + tw) \cdot w dt \\ &= \int_0^1 \int_0^1 D^2f(x + tw + sv) \cdot v ds dt \cdot w \\ &= \int_0^1 \int_0^1 D^2f(x + tw + sv) ds dt \cdot v \cdot w \end{aligned}$$

We can rewrite the application of v then w by $\cdot(v, w)$, and using the approximation $D^2f(x + tw + sv) \cdot (v, w) = D^2f(x) \cdot (v, w) + \delta_1(tw, sv)$. Integrating over s, t gives

$$\Delta = D^2f(x) \cdot (v, w) + \int_0^1 \int_0^1 \delta_1(tw, sv) ds dt$$

Note 2.2

The error term δ_1 in the integrand is given by

$$\delta_1(tw, sv) = D^2f(x + tw + sv)(v, w) - D^2f(x)(v, w)$$

for v, w sufficiently small and $0 \leq s, t \leq 1$.

A similar argument for h shows that $\Delta = D^2f(x) \cdot (w, v) + \int_0^1 \int_0^1 \delta_2(tw, sv) ds dt$. Combining the two together, the following holds for all v, w sufficiently small:

$$D^2f(x) \cdot (v, w) - D^2f(x) \cdot (w, v) = \int_0^1 \int_0^1 \delta_1(tw, sv) ds dt - \int_0^1 \int_0^1 \delta_2(tw, sv) ds dt \quad (54)$$

To show the right hand side is 0, we will need the following note.

Note 2.3

We wish to show the RHS of Equation (54) is 0. We begin by controlling the RHS and show that it is super-bilinear; meaning it shrinks after than the product $|v||w|$. Then, we will prove a lemma which will show the only bilinear map that satisfies this property is the 0 map.

- For $j = 1, 2$, relabel $\delta = \delta_j$ for convenience. We can use the L^1 inequality, to obtain the estimate

$$\left| \int_0^1 \int_0^1 \delta(tw, sv) ds dt \right| \leq \int_0^1 \int_0^1 |\delta(tw, sv)| ds dt \quad (55)$$

- $\delta(tw, sv)$ is controlled by $|D^2f(x + tw + sv) - D^2f(x)||v||w|$. Take $y = tw + sv$, then $|y| \leq |tw| + |sv|$. Hence,

$$|\delta_j| \leq |D^2f(x + tw + sv) - D^2f(x)||v||w| \quad (56)$$

- Let A denote the span of w, v for scalars $s, t \in [0, 1]$. In symbols,

$$A = \left\{ tw + sv, s, t \in [0, 1] \right\}$$

A is clearly compact, and the continuity of D^2f means

$$R(v, w, \delta) = \sup_{y \in A} |D^2f(x + y) - D^2f(x)| \text{ is finite, and } \lim_{(v, w) \rightarrow 0} R(v, w, \delta) = 0 \quad (57)$$

See remark 2.1 for a generalization of this argument.

- Relabel $R(v, w)$ to be the maximum across $R(v, w, \delta_1)$ and $R(v, w, \delta_2)$.

- Combining Equations (55) to (57), we obtain the following bound on Equation (54)

$$\begin{aligned}
 \left| D^2 f(x) \cdot (v, w) - D^2 f(x) \cdot (w, v) \right| &\leq \left| \iint \delta_1(tw, sv) ds dt - \iint \delta_2(tw, sv) ds dt \right| \\
 &\leq \iint |\delta_1| ds dt + \iint |\delta_2| ds dt \\
 &\leq |v||w|R(v, w)
 \end{aligned} \tag{58}$$

The following Lemma gives a useful criterion to check when a multilinear map is identically 0.

Lemma 2.1

Let E be a Banach space, and $k \geq 1$ be an integer. If $\lambda \in L(E^k, F)$ and there exists another map $\theta : E^k \rightarrow F$ (defined perhaps on an open neighbourhood of the origin), such that

$$|\lambda(u_{\underline{k}})| \leq |\theta(u_{\underline{k}})| \cdot \prod |u_{\underline{k}}|$$

for all $(u_{\underline{k}})$ sufficiently small. And $\lim_{(u_{\underline{k}}) \rightarrow 0} \theta(u_{\underline{k}}) = 0$, then, $\lambda = 0$.

Proof. Fix arbitrary $(u_{\underline{k}}) \in E^k$, for $s > 0$ sufficiently small, the left hand side of the equation reads

$$|s|^k |\lambda(u_{\underline{k}})| \leq |\theta(su_{\underline{k}})| \cdot |s|^k \prod |u_{\underline{k}}|$$

The rest of the argument is Archimedean: divide by $|s|^k$ and send $s \rightarrow 0$ (while paying attention to the term with θ): perhaps after relabelling $v_s = su_{\underline{k}}$ for sufficiently small s , then $|\theta(v_s)| \rightarrow 0$ as $s \rightarrow 0$. ■

■

Remark 2.1: Compact linear combinations

Generalization of the "compact linear combination" argument used above. Let $(t_{\underline{k}}) \subseteq \mathbb{C}^k$ or \mathbb{R}^k , and vectors $v_{\underline{k}} \in E$. Suppose further $(t_{\underline{k}}) \subseteq A$ is compact in \mathbb{C}^k or \mathbb{R}^k . It is clear that if $y = t_i v^i \in E$, where the summation convention is in effect. Then,

$$|y| \lesssim_A |(v^{\underline{k}})|_{E^k}$$

Now, fix a continuous function $f \in C(E, F)$, we can approximate the maximum error over all such y

$$\sup_{y \in B} |f(x + y) - f(x)| < \varepsilon \quad \forall |y| \lesssim_A |(v^{\underline{k}})| < \delta$$

where

$$B = \left\{ \sum t_i v^i, (t_{\underline{k}}) \subseteq A, (v^{\underline{k}}) \in E^k \right\}$$

The p -th derivatives

If f is p times differentiable, and $f, Df, D^2f, \dots, D^p f$ are all continuous, then we say $f \in C^p(E, F)$ (replacing E with an open subset of E if necessary).

Proposition 3.1

If $f \in C^p(E, F)$, then $D^p f(x)$ is symmetric for every $x \in E$. (Replace E with an open set if necessary).

Proof. The main proof proceeds as follows. We will use induction on p , with $p = 2$ serving as the base case. Our induction hypothesis is that for every $f \in C^{p-1}(E, F)$, for every permutation $\beta \in S_{p-1}$, at every point $x \in E$, for every possible choice of $p-1$ vectors $(v_2, \dots, v_p) = (v_{1+\underline{p-1}})$,

$$D^{p-1}f(x)(v_{1+\underline{p-1}}) = D^{p-1}f(x)(v_{1+\beta(\underline{p-1})})$$

To prove the assertion for p , it suffices to show $D^p f(x)(v_p)$ is invariant under transpositions of indices; since the transpositions generate S_p . Furthermore, the transpositions in S_p are generated by

- the transposition $(1, 2, \dots) \mapsto (2, 1, \dots)$ where the omitted indices are held fixed, and
- the transpositions which leave the first index fixed:

$$(1, 1 + \underline{p-1}) \mapsto (1, 1 + \beta(\underline{p-1}))$$

where $\beta \in S_{p-1}$

so it suffices to prove invariance under those two types of transpositions. Let $g = D^{p-2}f$, so $g \in C^2(E, L(E^{p-2}, F))$. Because the application of vectors (currying) on a multilinear map $A \in L(E^p, F)$ is associative, illustrated as follows:

$$(A \cdot v_1) \cdot v_2 = A \cdot (v_1, v_2) = A(v_1, v_2, \cdot) \in L(E^{p-2}, F)$$

Then, let $\lambda : L(E^{p-2}, F) \rightarrow F$ be the evaluation map at $(v_3, \dots, v_p) = (v_{2+\underline{p-2}})$. Using the base case on $D^{p-2}f = g \in C^2(E, L(E^{p-2}, F))$,

$$(D^2g)(x)(v_1, v_2) = (D^2g)(x)(v_2, v_1) \implies \lambda((D^2g)(x)(v_1, v_2)) = \lambda((D^2g)(x)(v_2, v_1))$$

But λ is the map that *applies* the rest of the vectors, and

$$(D^2g)(x)(v_1, v_2) \cdot (v_{2+\underline{p-2}}) = (D^2g)(x)(v_2, v_1) \cdot (v_{2+\underline{p-2}}) \quad (59)$$

Since D commutes with continuous linear maps (and λ is continuous because $(v_{2+\underline{p-2}})$ is fixed),

$$\lambda(D^2(D^{p-2}f)) = D(\lambda(D(D^{p-2}f))) = D(D\lambda \circ D^{p-2}f) = D^2(\lambda \circ D^{p-2}f) \quad (60)$$

Substituting Equation (59) for the rightmost hand side of Equation (60) gives the result.

Note 3.1

There are no magic 'identifications' being made here. To be perfectly clear, for each $x \in E$, $g(x)$ is an element in $L(E^{p-2}, F)$, and $(D^2g)(x) \in L(E^2, L(E^{p-2}, F))$. Evaluating g at a point x gives a bilinear map that takes values in the Banach space $L(E^{p-2}, F)$.

For the second case, beginning from the induction hypothesis. If θ is a p -permutation that leaves the first coordinate unchanged, then there exists a unique $p-1$ -permutation $\beta \in S_{p-1}$ such that

$$\begin{aligned} (\theta(\underline{p})) &= (1, \theta(1 + \underline{p-1})) \\ &= (1, 1 + \beta(\underline{p-1})) \end{aligned} \tag{61}$$

Using a similar argument as the first case, set $g = D^{p-1}f$ and $\lambda, \lambda' \in L(E^{p-1}, F)$ to be the evaluation maps of $(v_1, v_{1+\underline{p-1}}) = (v_{\underline{p}})$ and $(v_1, v_{1+\beta(\underline{p-1})})$ respectively. Rehearsing the same proof as before:

$$\begin{aligned} (D^p f)(x)(v_{\underline{p}}) &= D(\lambda D^{p-1} f)(x)(v_1) && \text{Equation (60)} \\ &= D(\lambda' D^{p-1} f)(x)(v_1) && \text{ind. hyp.} \\ &= (D^p f)(x)(v_{\theta(\underline{p})}) && \text{Equation (60)} \end{aligned}$$

This proves the induction step, and the proof is complete. ■

Before stating and proving Taylor's Theorem, an important remark on the 'postcomposition' of linear maps. Summarized in the following note.

Note 3.2

Let $f \in C^p(E, F)$, and $\lambda \in L^p(F, G)$. λ induces a map between $L(E^p, F)$ and $L(E^p, G)$ by postcomposing any multi-linear map $A \in L(E^p, F)$ by λ . Denoting this map by λ_* ,

$$\lambda_* : L(E^p, F) \rightarrow L(E^p, G)$$

It is clear λ_* is linear and continuous. And its action on A , evaluated at $(v_{\underline{p}}) \in E^p$ is given by

$$\lambda_*(A) \in L(E^p, G) \quad (\lambda_*(A))(v_{\underline{p}}) = \lambda(A(v_{\underline{p}})) = (\lambda \circ A)(v_{\underline{p}})$$

Now, recall that for $p = 1$

$$[D(\lambda \circ f)](x) = \lambda[(Df)(x)]$$

To simplify the notation, we want to 'move' the evaluation x outside of the brackets, and somehow write $x \mapsto \lambda[(Df)(x)]$ as one map between E and $L(E, G)$. We further *identify* λ as this map, so that

$$[D(\lambda \circ f)](x) = \lambda = (\lambda \circ Df)(x)$$

Dropping the x from the expression, for $p \geq 2$ *assuming a similar formula holds*, then we write $[D^p(\lambda \circ f)] = \lambda_* \circ D^p f$. We make a final identification, of $\lambda = \lambda_*$ (thereby conflating the two different maps, the first is a map from E to F , the second is a map from $L(E^p, F)$ into $L(E^p, G)$).

Proposition 3.2: Linear maps commute with D^p

If $p \geq 2$, $f \in C^p(E, F)$, $\lambda \in L(F, G)$, then

$$D^p(\lambda \circ f) = \lambda \circ D^p f$$

Where we have identified λ as the same map that acts on $L(E^p, F)$ to produce another map in $L(E^p, G)$,

and suppressed the point x .

Proof. Use induction on p . ■

Proposition 3.3: C^p is closed under composition

If $f \in C^p(E, F)$, and $g \in C^p(F, G)$, then $g \circ f \in C^p(E, G)$.

Proof. Postponed. ■

Proposition 3.4: Taylor's Formula

Let $f \in C^p(U, F)$, where $U \overset{\circ}{\subseteq} E$. For $x \in U$ and $y \in E$ such that $L = \{x + ty, 0 \leq t \leq 1\}$ is contained in U , then

$$f(x + y) = f(x) + \left(\sum_{i=\underline{p-1}} \frac{D^i f(x) \cdot (y^{(i)})}{(p-1)!} \right) + R_p \quad (62)$$

where $\cdot(y^{(i)})$ denotes the application of y , consecutively for i times. The remainder R_p is given by eq. (63)

$$R_p = \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p f(x + ty) dt \cdot (y^{(p)}) \quad (63)$$

Furthermore, we include the p th term in the series using eq. (64)

$$f(x + y) = f(x) + \sum_{i=\underline{p}} \frac{D^i f(x) \cdot (y^{(i)})}{i!} + \theta(y) \quad (64)$$

where θ is defined for small y , and $o(|y|^p)$.

$$|\theta(y)| \leq \sup_{0 \leq t \leq 1} \frac{|D^p f(x + ty) - D^p f(x)|}{p!} |y|^p \quad (65)$$

Proof. Postponed. ■

Chapter 1: Manifolds

Introduction

We begin with some terminology concerning maps between Banach spaces. Let E and F be Banach spaces over \mathbb{R} .

- $\mathcal{L}(E, F)$, (resp. $L(E, F)$) = Linear (resp. toplinear) maps between E and F ,
- $\text{Topliso}(E, F)$ = toplinear isomorphisms between E and F ,
- $\text{Laut}(E)$ = toplinear automorphisms on E

Remark 1.1: Laut is open in L(E,E)

The space of toplinear automorphisms is open in the strong topology in the space of toplinear endomorphisms.

The structure of a manifold

It is fruitful to *construct* the manifold rather than *define* it. We also insist on working with open sets of Banach spaces instead coordinate functions as our primary data.

We will be working in the category of C^p Banach spaces (all Banach spaces are assumed to be over \mathbb{R}). Its morphisms are C^p morphisms: the maps which are continuously p -times differentiable (but not necessarily linear). Note that if $p \geq 0$, every toplinear morphism is a C^p morphism, and every toplinear isomorphism is a C^p isomorphism. However, a bijective C^p morphism is usually not a C^p isomorphism.

Definition 2.1: Chart

Let X be a non-empty set. A *chart on X modelled on a Banach space E* is a tuple (U, φ) , such that $U \subseteq X$, $\varphi(U) = \tilde{U}$ is an *open* subset of E , and φ is a bijection onto \tilde{U} .

Definition 2.2: Compatibility

Let (U, φ) and (V, ψ) be charts on X modelled on E , they are called C^p compatible if $U \cap V = \emptyset$, or

- $\varphi(U \cap V)$ and $\psi(U \cap V)$ are *both* open subsets of E , and
- the *transition map* $\psi \circ \varphi^{-1} : \varphi(U \cap V) \rightarrow \psi(U \cap V)$ is a C^p isomorphism between open subsets of E .

It should be clear that compatibility is an equivalence relation on the space of charts of X (that are modelled on E).

Definition 2.3: Atlas

Let X be a non-empty set. A C^p *atlas on X modelled on E* is a pairwise C^p compatible collection of charts $\{(U_\alpha, \varphi_\alpha)\}$ whose union over the domains cover X .

Remark 2.1: Omissions

If the *model space* E is implied, we will not explicitly reference it. When operating 'within category', we might refer to two charts as *compatible* or *smoothly compatible*, implying they are C^p compatible. This comes from the perspective that, in the context of C^p manifolds, any smoothness exceeding C^p is deemed sufficiently smooth for our purposes.

Let X be a non-empty set, equipped with a C^p atlas $\{(U_\alpha, \varphi_\alpha)\}$ modelled on E . If α and β both index the atlas, we write $U_{\alpha\beta} = U_\alpha \cap U_\beta$.

Suppose $U_{\alpha\beta}$ is non-empty. Then, (by definition) the images $\varphi_\alpha(U_{\alpha\beta})$, $\varphi_\beta(U_{\alpha\beta})$ are *both* open subsets of E , and we will denote the transition map by

$$\varphi_\beta \circ \varphi_\alpha^{-1} = \varphi_{\beta\alpha^{-1}} : \varphi_\alpha(U_{\alpha\beta}) \rightarrow \varphi_\beta(U_{\alpha\beta}) \quad (66)$$

If $p \in (U, \varphi)$, we write \hat{p} for $\varphi(p)$ if there is no room for ambiguity. From Definitions 2.2 and 2.3, the compatibility relation on charts descends into a compatibility relation on the space of atlases, whose properties are summarized in the following note.

Note 2.1: Descent of an equivalence relation

Let Ω be a non-empty set with an associated equivalence relation \sim . This relation \sim induces another equivalence relation on the set containing all subsets of equivalence classes from Ω . Suppose A and B are subsets of the equivalence classes $[A]$ and $[B]$ respectively. The condition $A \sim B$ holds if and only if for all elements x in A and y in B , $x \sim y$.

This is equivalent to stating that the union $A \cup B$ lies entirely within some equivalence class, and further, that $[A] \sim [B]$. The class $[A]$ represents the largest subset of Ω that is entirely contained within a single equivalence class (namely $[A]$ itself) and contains A as a subset.

Definition 2.4: Structure determined by an atlas

The maximal atlas that contains \mathcal{A} as a subset is called the C^p *structure determined by* \mathcal{A} . This maximal atlas is unique, by note 2.1.

Definition 2.5: Manifold

A C^p *manifold modelled on* E is a non-empty set X with a C^p structure modelled on E . We sometimes refer to the manifold as the smooth structure, rather than the set X itself. Man^p refers to the *category of* C^p *manifolds*.

Proposition 2.1: E is a manifold

Let $p \geq 1$. The identity map $\text{id}_E : E \rightarrow E$ defines an atlas on E , which determines a structure called the *standard* C^p *structure on* E or *standard structure on* E if the class of morphisms is understood.

We will call (E, id_E) the *standard chart*, or the *global chart* on E .

Proposition 2.2: Topology is unique on a manifold

Let X be a manifold modelled on E , it has a unique topology such that the domain for each chart in its smooth structure is open, and each chart is a homeomorphism onto its range (with respect to the subspace topology of E).

Proof. We offer a sketch of the proof. Fix a chart (U, φ) , it is clear that U has to be in the topology of X , and because $\varphi : U \rightarrow \hat{U}$ is required to be a homeomorphism, we duplicate all the open sets in \hat{U} by using the inverse image through φ . The collection of all such inverse images form a sub-basis, thus defines a unique topology as is well known.

There is an alternate way of thinking about this 'induced topology'. Given a chart domain, there exists a unique coarsest topology such that all charts with the same chart domain are homeomorphisms onto their images. We can stitch these weak topologies together to form a ambient topology on X , as the chart domains cover X . ■

Remark 2.2: Not necessarily Hausdorff

The topology generated is not necessarily Hausdorff, nor second countable. So X may not admit partitions of unity, but for our current purposes we will work with this general definition. Because of the uniqueness of the topology, we sometimes refer to the topology as being part of the *structure* of the manifold.

Proposition 2.3: Open subsets of manifolds

If U is an open subset of a C^p manifold X , then U is a C^p manifold whose structure is determined by the atlas

$$\left\{ (V, \varphi) \text{ in the structure of } X, \text{ where } V \subseteq U \right\} \quad (67)$$

Proof. The smooth structure of X includes all possible restrictions to open sets; hence the set in eq. (67) defines an atlas, and a unique structure by def. 2.4. ■

Morphisms between manifolds

Definition 3.1: C^p morphisms between manifolds

Let X and Y be C^p manifolds over the spaces E and F . A map $F : X \rightarrow Y$ is a morphism in Man^p if for every $p \in X$, there exists charts (U, φ) in X and (V, ψ) in Y such that the image $F(U)$ is contained in V , and the conjugation of F with respect to the two charts is C^p smooth between open subsets of Banach spaces.

$$F_{U,V} \triangleq \psi F \varphi^{-1} \in C^p(\hat{U}, \hat{V}) \quad (68)$$

The map defined in eq. (68) is called the *coordinate representation of F* with respect to the charts

$(U, \varphi), (V, \psi).$

Remark 3.1: Identifying charts with their domains

The scenario in eq. (68) occurs so often that we decide to simply write

$$F_{U,V} = \psi F \varphi^{-1} \quad (69)$$

to mean there exists charts $(U, \varphi), (V, \psi)$ in the structure of X, Y such that

$$F(U) \subseteq V \quad (70)$$

Consistent with our notation for identifying a chart by its chart domain, we write

$$F_{U,V}(\hat{p}) = (\psi F \varphi^{-1})(\hat{p}) \quad (71)$$

for any morphism $F \in \text{Mor}(X, Y)$, charts that satisfy eq. (70). We refer to the map in eq. (71) as a *coordinate representation of F about p* .

Definition 3.1 may leave one unsatisfied. A common question that comes to mind is: why do we require the image $F(U)$ be contained in another chart domain in Y ? There are two reasons.

1. First, it is easily verified that the C^p maps between open subsets of Banach spaces satisfy the usual functorial properties in its category. The definition of smoothness between Banach spaces is a purely local one, and it is defined between open subsets; and recall: **every chart domain U in a manifold X corresponds to an open subset $\hat{U} \subseteq E$ in the model space**. The necessity that $F(U)$ must be contained in a single chart domain of Y is a relic of the original definition.
2. Second, suppose f is a map between E and F , and the restriction of f onto a family of open subsets $U_\alpha \subseteq E$ is C^p for $p \geq 0$. If $\{U_\alpha\}$ is an open cover for E , then f is continuous. Proposition 3.1 shows this equally holds for manifolds.

Proposition 3.1

Every C^p morphism between manifolds is a continuous map, and the composition of C^p morphisms is again a morphism.

Proof. The first claim follows immediately from eq. (68), since p is arbitrary, choose any neighbourhood W of $F(p)$, by shrinking this neighbourhood, it suffices to assume it is a subset of the chart domain V . The charts on X and Y are homeomorphisms, and unwinding the formula shows that $F|_U = \psi^{-1} F_{U,V} \varphi$, so that

$$U \cap F^{-1}(W) = (F|_U)^{-1}(W) \text{ is open in } X$$

To prove the second, let X_3 be manifolds modelled over E_3 , and F_1, F_2 is smooth between X_i such that $F_2 \circ F_1$ makes sense. Since F_1 is smooth, there a pair of charts $(U_i, \varphi_i) \in X_i$ for $i = 1, 2$ about each $p \in X_1$ such that F_{1U_1, U_2} is C^p between open subsets.

$F_2(F_1(p))$ induces another pair of charts $(V_i, \psi_i) \in X_i$ for $i = 2, 3$. Since F_2 is smooth, it is continuous. $F_1^{-1} \circ F_2^{-1}(V_3)$ is open in X_1 , and we can shrink all of our charts so that $F_2 F_1(U_1)$ is contained in

V_3 . Finally, because C^p morphisms between open subsets of Banach spaces is closed under composition, $F_{U_1 \cap F_1^{-1} F_2^{-1}(V_3), V_3}$ is smooth. ■

Remark 3.2: Concluding remarks

Manifolds hereinafter will be assumed of class C^p , where $p \geq 0$. If (U, φ) is a chart in the structure of X , we will simply say (U, φ) is in X ; or (U) is in X .

Tangent spaces

The next question that we will address is taking derivatives of smooth maps between manifolds. There is no reason to demand C^p smoothness between maps, or even a C^p category of manifolds if we cannot borrow something more other than the morphisms on open sets. **In this section, all manifolds will be of class C^p for $p \geq 1$.**

Suppose U is an open subset of E and $f : U \rightarrow Y$ is C^p for $p \geq 1$. The derivative $Df(x)$ is a linear map $E \rightarrow F$, not from U to F (U might not even be a vector space). This suggests the 'derivative' of a morphism $F : X \rightarrow Y$ between manifolds can in some sense be interpreted as the *ordinary derivative* of its coordinate representation $DF_{U,V}(\hat{p})$, adhering to our principle of using open sets.

But there is a problem with this 'derivative': it gives different values for different charts. With infinitely many charts in X and Y , this definition becomes useless. To see this, let X be a manifold modelled on E and $p \in X$. If $f : X \rightarrow Y$ is a morphism, and (U_1, φ_1) , (U_2, φ_2) are charts defined about p such that the representations $f_{U_1,V}$ and $f_{U_2,V}$ are morphisms. Writing $p_i = \varphi_i(p)$, $U_{1,2} = U_1 \cap U_2$ and

$$\varphi_{1,2} = \varphi_2 \varphi_1^{-1} : \varphi_1(U_{1,2}) \rightarrow \varphi_2(U_{1,2}) \quad (72)$$

(because the map in eq. (72) goes from the domain U_1 to U_2), a simple computation yields eq. (73).

$$\begin{aligned} Df_{U_1,V}(p_1)(v) &= D(\psi f \varphi_2^{-1} \varphi_2 \varphi_1^{-1})(p_1)(v) \\ &= Df_{U_2,V}(p_2) \left(D\varphi_{1,2}(p_1)(v) \right) \\ &= Df_{U_2,V}(p_2) \circ D\varphi_{1,2}(p_1) \cdot (v) \end{aligned} \quad (73)$$

where $\cdot(v)$ denotes the evaluation at $v \in E$, and is assumed to be left associative over composition. The computation in eq. (73) suggests that interpreting the derivative by pre-conjugation is dependent on the chart being used to interpret the derivative. In fact, $D\varphi_{1,2}(p_1)$ can be replaced with any toplinear isomorphism on E (relabel $\varphi_2 = A\varphi_1$ where A is any linear automorphism on E), so the right hand side of eq. (73) can be interpreted as $Df_{U_2,V}(p_2)(w)$ where w is any vector in E .

Definition 4.1: Concrete tangent vector

Suppose $k \geq 1$, X a C^k -manifold on E , and $p \in X$. If (U, φ) is any chart containing p , for each $v \in E$ we call (U, φ, p, v) a *concrete tangent vector at p* that is *interpreted* with respect to the chart (U, φ) . The disjoint union of concrete tangent vectors, as shown in eq. (74)

$$T_{(U, \varphi, p)} X = \bigcup_{v \in E} \{(U, \varphi, p, v)\} \cong E \quad (74)$$

is called the *concrete tangent space at p* interpreted with respect to (U, φ) ; and it inherits a TVS structure from E .

Fix a point p in a manifold X . Suppose (U_i, φ_i) are charts containing p , from eq. (73) there exists a natural (toplinear) isomorphism between the concrete tangent spaces, namely

$$(U_1, \varphi_1, p, v_1) \sim (U_2, \varphi_2, p, v_2) \quad \text{iff} \quad v_2 = D\varphi_{1,2}(p_1)(v_1) \quad (75)$$

where $p_i = \varphi_i(p)$. The right member of eq. (75) is the derivative of a transition map — which is a toplinear automorphism on E . Hence $D\varphi_{1,2}(p_1)$ defines a toplinear isomorphism between $T_{(U_1, \varphi_1, p)}X$ and $T_{(U_2, \varphi_2, p)}X$. With this, we define the primary object of our study.

Definition 4.2: Tangent vector

A *tangent vector* (or an *abstract tangent vector*) at p is defined as an equivalence class of concrete tangent vectors at p , under the relation in eq. (75).

Definition 4.3: Tangent space

The *tangent space* at p , denoted by T_pX is the set of all tangent vectors at p . It is toplinearly isomorphic to the model space E .

Definition 4.4: Differential of a morphism

Let X and Y be modelled on the spaces E and F . If f be a morphism between X and Y , the *differential of f at p* is the unique linear map denoted by

$$df(p) = df_p : T_pX \rightarrow T_{f(p)}Y \quad (76)$$

Whose action is characterized by the following:

- if (U, φ) and (V, ψ) are any pair of charts that satisfy the morphism condition in eq. (68) about p ,
- if $v \in T_pX$ is represented by (U, φ, p, \hat{v}) ,
- then $df(p)(v) \in T_{f(p)}Y$ is represented by $(V, \psi, f(p), Df_{U,V}(\hat{p})(\hat{v}))$

Note 4.1: Interpretation using co-product

There is another way of interpreting the construction above. Each concrete tangent space is toplinearly isomorphic to E , the projection maps onto $\{p\}$ and E can be glued together using the universality of the coproduct, where $\{p\}$ is interpreted as a 0-dimensional vector space. The construction of T_pM follows by invoking the property of the quotients.

Remark 4.1: Omission of chart in concrete representation

If p is a point on a manifold X , $v \in T_p M$, we sometimes say (U, \hat{v}) , or \hat{v} is an interpretation of v if it is clear (U, φ) is a chart in X . If $X = E$, we will *identify* v with its concrete representation in the *standard chart* (E, id_E) . The standard representation of a tangent vector is written with a bar on top: \bar{w} is the *standard representation*, or *standard interpretation* of w .

Furthermore, we also write (\hat{p}, \hat{v}) where $\hat{p} = \varphi(p)$ and (U, φ, p, \hat{v}) .

Remark 4.2: Morphisms between C^k , C^p manifolds

Let X be a C^k -manifold, and Y a C^p manifold, where $k, p \geq 0$. A morphism between X and Y is a map $f : X \rightarrow Y$ such that each point $p \in X$ admits a coordinate representation

$$f_{U,V} \in C^{\min(p,k)}(\hat{U}, \hat{V}) \quad (77)$$

If $\min(p, k) \geq 1$, then we define its differential as in def. 4.4.

Curves

In the previous section, we motivated the definition of $T_p X$ using the computation of the derivative of a morphism from X . Dually, the tangent space allows us compute the derivatives of morphisms into X in a coordinate independent manner.

Let $J_\varepsilon = (-\varepsilon, +\varepsilon)$ be an open interval in \mathbb{R} . Viewing J_ε as a manifold, the morphisms $\gamma : J_\varepsilon \rightarrow X$ are *curves in X* and $\gamma(0)$ is called the *starting point* of γ .

Definition 5.1: Velocity of a curve

Let γ be a curve in X and $t \in J_\varepsilon$. The *velocity* at t , denoted by $\gamma'(t)$ — is the tangent vector with representation $d_{J_\varepsilon, V} \gamma(\bar{1})$; where $(J_\varepsilon, \bar{1})$ is a concrete tangent vector in $T_t J_\varepsilon$.

Proposition 5.1: Tangent vectors are velocities

Let p be a point on a manifold X . For every tangent vector $v \in T_p X$, there exists a curve starting at p whose velocity is v .

Proof. Find a chart (U) in X where $\hat{p} = 0$. Such a chart exists, because translations and dilations are C^p isomorphisms. If the tangent vector v has interpretation \hat{v} in U , there exists $\varepsilon > 0$ so small that the range of $\hat{\gamma}$, as defined eq. (78), lies in \hat{U}

$$\hat{\gamma} : J_\varepsilon \rightarrow \hat{U} \quad \gamma(t) = \int_0^t \hat{v} dt \quad (78)$$

$\hat{\gamma}$ is a curve in \hat{U} starting at \hat{p} with velocity \hat{v} . Defining γ as the composition of $\hat{\gamma}$ with the chart inverse finishes the proof. ■

Splitting

Recall: if W is a vector space and W_1, W_2 are linear subspaces of V . W_2 is the vector space complement of W_1 (resp. with the indices reversed) if

$$W_1 + W_2 = W, \quad \text{and} \quad W_1 \cap W_2 = 0$$

Definition 6.1: Splitting in E

A linear subspace E_1 splits in E if both E_1 and its vector space complement E_2 are closed, and the addition map $\theta : E_1 \times E_2 \rightarrow E$ given by

$$\theta(x, y) = x + y \quad \text{is a toplinear isomorphism.}$$

We sometimes refer to the vector space complement of W_1 as its *linear complement*.

Remark 6.1: Every linear subspace splits in finite dimensions.

Every finite dimensional or finite codimensional linear subspace of E splits. If E is finite dimensional, then every linear subspace splits.

Definition 6.2: Splitting in $L(E, F)$

A continuous, injective linear map $\lambda \in L(E, F)$ *splits* iff its range splits in F . Alternatively, λ splits iff there exists a toplinear isomorphism $\alpha : F \rightarrow F_1 \times F_2$ such that λ composed with α induces a toplinear isomorphism from E onto $F_1 \times 0$ — which we identify with F_1 .

For the next few definitions, X and Y will be manifolds.

Definition 6.3: Immersion

A morphism $f \in \text{Mor}(X, Y)$ is an *immersion* at a point $p \in X$ if there exists a coordinate representation about $f_{U,V}$ such that

$$Df_{U,V}(\hat{p}) \quad \text{is injective and splits.} \tag{79}$$

The morphism f is called an immersion if eq. (79) holds at every p .

Definition 6.4: Submersion

A morphism $f \in \text{Mor}(X, Y)$ is an *submersion* at a point $p \in X$ if there exists a coordinate representation about $f_{U,V}$ such that

$$Df_{U,V}(\hat{p}) \quad \text{is surjective and its kernel splits.} \tag{80}$$

The morphism f is called an submersion if eq. (80) holds at every p .

Remark 6.2: Similarity to Matrix Canonical Forms

Suppose X and Y are finite dimensional and if f is an immersion (resp. submersion), there exist a coordinate representation about each $p \in X$ such that eq. (81) (resp. eq. (82)) holds.

$$D\hat{f}(\hat{p}) = \begin{bmatrix} \text{id}_{m \times m} \\ 0_{n-m \times m} \end{bmatrix} \quad (81)$$

$$D\hat{f}(\hat{p}) = \begin{bmatrix} \text{id}_{n \times n} & 0_{n \times m-n} \end{bmatrix} \quad (82)$$

Definition 6.5: Embedding

A morphism $f \in \text{Mor}(X, Y)$ is an *embedding* if it is a immersion and it is a homeomorphism onto its range.

Definition 6.6: Toplinear subspace

Let E be a Banach space, a *toplinear subspace* (of E) is a closed linear subspace E_1 which splits in E .

Submanifolds

Before we state the definition of a submanifold, it is important to recapitulate the construction of a manifold X .

1. Given a non-empty set X and an atlas modelled on a space E .
2. The purpose of each chart in the atlas is to borrow open subsets $\hat{U} \stackrel{\circ}{\subseteq} E$. If we single out a single chart, **the construction is entirely topological**. It is of little importance *how* the individual chart domains U are mapped onto \hat{U} ,
3. Each chart is in **bijection with its range**, which is an open subset of E , and
4. the transition maps $\varphi_{\beta\alpha^{-1}}$ are **morphisms between open subsets of E** .

In the spirit of borrowing definitions and properties from existing objects, it makes (functoral) sense a submanifold S should be modelled a linear subspace of E_1 of E . The natural charts we can borrow from the structure of X are those with the 'other coordinates' muted. If (U, φ) is a chart whose domain intersects S , the restriction of φ onto $U \cap S$ should be in bijection with an open subset of E_1 .

$$\varphi(S \cap U) = \hat{U}_1 \times ?, \quad \hat{U}_1 \stackrel{\circ}{\subseteq} E_1 \quad (83)$$

There is a problem with eq. (83) however, φ is a C^p isomorphism onto \hat{U} ; not onto open subsets of the product space $E_1 \times E_2$. An easy fix to this would be to require E_1 **to split in E** (and shrinking U using a basis argument). Let α be a C^p isomorphism between E and $E_1 \times E_2$. Equation (83) becomes

$$\alpha\varphi(S \cap U) = \hat{U}_1 \times a_2 \quad \text{where} \quad \hat{U}_1 \stackrel{\circ}{\subseteq} E_1 \text{ and } a_2 \in E_2 \quad (84)$$

Identifying \hat{U} with $\alpha(\hat{U})$, and requiring $U_1 \times a_2$ to be in $\alpha(\hat{U})$, we arrive at the following definition.

Definition 7.1: Submanifold

Let X be a manifold, and S a subset of X . We call S a *submanifold* of X if there exist split subspaces E_1, E_2 of E ; such that, every $p \in S$ is contained in the domain of some chart (U, φ) in X . Where

$$\varphi : U \rightarrow \hat{U} \cong \hat{U}_1 \times \hat{U}_2, \quad \text{where} \quad U_i \stackrel{\circ}{\subseteq} E_i \quad i = 1, 2 \quad (85)$$

and there exists an element $a_2 \in \hat{U}_2$

$$\varphi(U \cap S) = \hat{U}_1 \times a_2 \quad (86)$$

We call a chart satisfying eqs. (85) and (86) a *slice chart* of S ; to simplify what follows, we write $\varphi^i = \text{proj}_i \varphi$ for $i = 1, 2$ for any slice chart (U) . Given that proj_i is a morphism between open subsets of Banach spaces, φ^i is again a morphism. In particular, φ^1 is in bijection from $U^s = U \cap S$ onto \hat{U}_1 ; the latter being an open subset of E_1 . To show S is indeed a manifold it remains to show the collection of charts in eq. (87) forms a C^p atlas modelled E_1 , which we will prove in prop. 7.1

$$\mathcal{A} = \left\{ (U^s, \varphi^1), (U, \varphi) \text{ is a slice chart of } S \right\} \quad (87)$$

Proposition 7.1: Structure of a submanifold

If S is a submanifold of X , eq. (87) defines a C^p atlas over the space E_1 . The manifold S has a topology that coincides with the subspace topology, and the inclusion map $\iota_S : S \rightarrow X$ is a morphism and a homeomorphism onto its range.

Proof. Each of the charts in eq. (87) is in bijection with an open subset of E_1 . Let $(U_\alpha^s, \varphi_\alpha^1)$ and $(U_\beta^s, \varphi_\beta^1)$ be overlapping charts. Writing $U_{\alpha\beta}^s = U_\alpha^s \cap U_\beta^s$ as usual, and the transition map $\varphi_{\beta\alpha}^1 = \varphi_\beta^1(\varphi_\alpha^1)^{-1}$ from $\varphi_\alpha^1(U_{\alpha\beta}^s)$ to $\varphi_\beta^1(U_{\alpha\beta}^s)$. Equation (86) tells us there exists $a_2 \in \hat{U}_{2,\alpha}$ and $b_2 \in \hat{U}_{2,\beta}$, that can help us recover the original chart. Identifying a_2 (resp. b_2) with the constant function ($p \mapsto a_2$) for $p \in U_\alpha^s$, we get eq. (88).

$$\varphi_\alpha^1 \times a_2 = \varphi_\alpha|_{U_\alpha^s} \quad (88)$$

(resp. $\varphi_\beta^1 \times b_2 = \varphi_\beta|_{U_\beta^s}$). The transition map is given by

$$\varphi_\beta^1 \circ (\varphi_\alpha^1)^{-1} = \text{proj}_{1,\beta} \varphi_\beta|_{U_\beta^s} (\varphi_\alpha|_{U_\alpha^s})^{-1} (\text{proj}_{1,\alpha}|_{U_\alpha^s})^{-1} \quad (89)$$

We can combine the two middle terms into $\varphi_\beta \varphi_\alpha^{-1}|_{U_{\alpha\beta}^s} = \varphi_{\beta\alpha}^{-1}|_{U_{\alpha\beta}^s}$. Which is a C^p isomorphism, because the domain (resp. codomain) of $\varphi_\alpha(U_{\alpha\beta}^s)$ (resp. β) is given by eq. (86). Suppressing the restrictions onto $U_{\alpha\beta}^s$, we have

$$\varphi_\alpha(U_{\alpha\beta}^s) = (\hat{U}_{1,\alpha} \cap \hat{U}_{1,\beta}) \times a_2 \quad \text{and} \quad \varphi_\beta(U_{\alpha\beta}^s) = (\hat{U}_{1,\alpha} \cap \hat{U}_{1,\beta}) \times b_2$$

The middle term in eq. (89) then becomes

$$\varphi_{\beta\alpha}^{-1} = (\varphi_{\beta\alpha}^1, (a_2 \mapsto b_2))$$

Which is a C^p isomorphism. The other terms in eq. (89) are either projections or products of isomorphisms with constant functions, therefore eq. (87) forms an atlas.

Let us use $\iota_S : S \rightarrow X$ to represent the inclusion map and consider a fixed point $p \in S$. It is always possible to identify a slice chart (U, φ) for the structure of X that contains $p = \iota_S(p)$. By definition of the atlas on S , eq. (87), this induces a 'truncated' chart (U^s, φ^1) .

Observing that $\iota_S(U^s) = \iota_S(U \cap S)$ lies within (U, φ) , the morphism criteria in eq. (68) is satisfied. Computing the coordinate representation of ι_S , we obtain eq. (90).

$$(\iota_S)_{U^s, U} = \varphi \iota_S (\varphi^1)^{-1} = \text{id}_{\hat{U}_1} \times a_2 \quad (90)$$

Equation (90) shows that the coordinate representation of ι_S is a local isomorphism. Since the inclusion map is a bijection and continuous, and the coordinate representation of ι_S^{-1} is simply the inverse eq. (90); ι_S^{-1} is a morphism and therefore continuous. ■

Remark 7.1: Pairs of slice charts

If p is a point on a submanifold S , we refer to a *pair of slice charts* containing p as the pair (U^s, φ^1) and (U, φ) in the structure of S and X .

Proposition 7.1 shows every point $p \in S$ is in the domain of a pair of slice charts. The inclusion map ι_S has coordinate representation eq. (90). Computing its ordinary derivative we obtain eq. (91).

$$D(\iota_S)_{U^s, U}(\hat{p}) : T_{(U^s, \varphi^1, p)} \longrightarrow T_{(U, \varphi, p)} \quad \text{and} \quad D(\iota_S)_{U^s, U}(\hat{p}) = \text{id}_{E_1} \times 0 \quad (91)$$

which is a toplinear morphism between concrete tangent spaces and has a simple representation of 'adding zeroes' at the end of a vector $\hat{v} \in E_1$.

Definition 7.2: Exterior tangent space of S

The *exterior tangent space* of a point $p \in S$ is the image of $T_p S$ under $d\iota_S(p)$,

$$T_p^{\text{ext}} S = d\iota_S(p)(T_p S) \quad (92)$$

which is a toplinear subspace of $T_p X$.

Vector Bundles

Our goal this section is to construct a vector bundle of a manifold X , which is the proper setting to study vector / covector fields, and later different forms.

Let X be a class C^p manifold modelled on a space E , and F another Banach space. Suppose for each p , the set W_p is toplinearly isomorphic to F at for each p , then we call W_p an *F-fiber* at p . The set-theoretic coproduct of all such W_p as in eq. (93) is called a *coproduct of F-fibers modelled over X*.

$$W = \coprod_{p \in X} W_p \quad \text{comes with} \quad \pi : W \rightarrow X, \quad \pi^{-1}(p) = W_p \quad (93)$$

It turns out the natural way of making W a manifold would be to steal open sets from *both* E and F — in this case, sets of the form $\hat{U} \times F$. We sometimes write \tilde{U} instead of $\pi^{-1}(U)$ for brevity, and \tilde{p} in place of $\pi^{-1}(p)$. The next few definitions will ring a few bells.

Definition 8.1: Local trivialisation

Let W be as in eq. (93). A *local trivialisation* of W is a tuple (\tilde{U}, Φ) , such that the diagram in fig. 1 commutes, and

- $U \subseteq X$ is open in X , and for each $p \in U$,
- $\Phi|_{\tilde{p}}$ is in bijection with $W_p = F$.

Definition 8.2: Compatibility between trivialisations

Let (\tilde{U}, Φ) and (\tilde{V}, Ψ) be local trivialisations of W , they are called C^k -compatible if $U \cap V = \emptyset$, or both of the following hold:

- for each $p \in U \cap V$ — the restriction of $\Psi \circ \Phi^{-1}$ onto the fiber of p — $(\Psi \circ \Phi^{-1})|_{\tilde{p}}$ is in $\text{Laut}(F, F)$, and
- the map $\theta : U \cap V \rightarrow L(F, F)$ as defined by eq. (94), is a C^k morphism into the Banach space $L(F, F)$.

$$\theta(p) = (\Psi \circ \Phi^{-1})|_{\tilde{p}} \quad (94)$$

(equivalently, we can require θ be C^k as a map into $\text{Laut}(F)$).

Definition 8.3: Trivialisation covering

Let W be a coproduct of F -fibers over X . A C^k *trivialisation covering* of W is a collection of pairwise C^k -compatible local trivialisations $\{(\tilde{U}_\alpha, \Phi_\alpha)\}$ where $\{U_\alpha\}$ is an open cover of X .

Definition 8.4: Vector bundle

Let X be a C^p manifold over E , and let F be a Banach space. An F -*vector bundle of rank k* is a coproduct of F -fibers modelled over a manifold X equipped with a **maximal C^k trivialisation covering**.

Remark 8.1: Maximality of trivialisation covering

One can easily verify the compatibility condition defines an equivalence relation, thus any C^k -trivialisation covering *determines* a maximal one.

Remark 8.2: k vs. p

A vector bundle W over X can be of a different class than X , a morphism between C^k and C^p manifolds are precisely the maps whose coordinate representation about every point is $C^{\min(p,k)}$; see remark 4.2 for details.

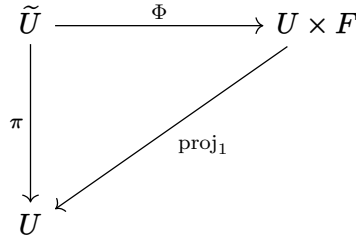


Figure 1: Local Trivialisation

The above definitions calls for some commentary, our end goal is to make an arbitrary vector bundle W a C^p manifold. Open sets will still be our primary 'topological' data. But in order to make W as compatible to X as possible, the eventual manifold structure we will put on W will **embed the structure of X into W** . This is the same argument as in the submanifold case but with the roles of X and S reversed.

Suppose we have a structure on W , then $X = \bigcup_{p \in X} \{p\} \times 0$ is a submanifold of the W — since E splits in the product space $E \times F$. Let us motivate a couple of the requirements above.

- Definition 8.1: 1) U is required to be open because W needs to inherit part of the topology, and we wish to 'steal' all of the charts in E whose domain is a subset of U , 2) the second bullet point implies that **each Φ is in bijection with $\Phi(\tilde{U}) = U \times F$, which is open in $E \times F$** .
- Definition 8.2: 1) the overlap restricts to a toplinear isomorphism on each fiber because **it allows us to quotient out the effects of the trivialisation transitions**, by rehearsing the same 'coproduct and quotient' argument in Definitions 4.1 to 4.3. 2) the requirement that θ as defined in eq. (66) be a C^p morphism is inherited from the fact that, if $p \in (U_i, \varphi_i)$ for $i = 1, 2$. Then $\varphi_2 \circ \varphi_1^{-1}$ is a C^p isomorphism between $\varphi_1(U_{1,2})$ and $\varphi_2(U_{1,2})$; **whose derivative is a C^{p-1} map into $\text{Laut}(E)$ that encodes the transformation between the concrete tangent spaces**. In the notation of eq. (72), this means

$$x \mapsto D\varphi_{1,2}(x) \quad \text{is in } C^{p-1}(\hat{U}_{1,2}, \text{Laut}(E))$$

In fact, the tangent bundle is a C^{p-1} vector bundle (modelled on E) over X .

Suppose W is an F -vector bundle over X with the trivialisation covering $\{(\tilde{U}^\alpha, \Phi_\alpha)\}$. For each α , we can cover U^α using chart domains $(U_\beta^\alpha, \varphi_\beta^\alpha)$ in X — without loss of generality, we can assume $U_\beta^\alpha \subseteq U^\alpha$ by restricting the chart domain and relabelling.

Similar to the construction of the induced atlas of a submanifold, given a 'piece' of the original manifold X — **instead of dropping the coordinates that correspond to E_2 , we add an F -component to construct a bijection with an open subset of $E \times F$** . This is shown in eq. (95)

$$\tilde{\varphi}_\beta^\alpha : \tilde{U}_\beta^\alpha \longrightarrow \hat{U}_\beta^\alpha \times F \quad \text{defined by} \quad \tilde{\varphi}_\beta^\alpha = (\varphi_\beta^\alpha \times \text{id}_F) \circ \Phi_\alpha \quad (95)$$

Remark 8.3: Hats and wiggles

Here, \tilde{U}_β^α should be interpreted as the inverse image of the open set U_β^α through π . Similarly, \hat{U}_β^α is

the image of U_β^α through φ_β^α .

The collection of charts in eq. (96) cover W with their chart domains, and each chart is in bijection with an open subset of $E \times F$.

$$\mathcal{A} = \left\{ (\tilde{U}_\beta^\alpha, \tilde{\varphi}_\beta^\alpha), (\tilde{U}^\alpha, \Phi_\alpha) \text{ is in the trivialisation covering of } W. \right\} \quad (96)$$

Proposition 8.1: Structure of a Vector Bundle

Let X be a C^p manifold modelled over E . If W is a C^k vector bundle modelled on F over the manifold X , then W is a C^k manifold modelled on the product space $E \times F$. Furthermore:

1. The *canonical projection* $\pi : W \rightarrow X$ is a morphism and a submersion.
- 2.

Chapter 2: Coordinates

Introduction

In the previous chapters, a chart (U, φ) was often equated with its domain. We will now express a concrete tangent vector as (\hat{p}, \hat{v}) , omitting any reference to the chart or its domain.

Let X be a manifold and F a Banach space. Consider a morphism $f \in \text{Mor}(X, F)$ and fix a point $p \in X$, and write $q = f(p)$. By adopting the canonical interpretation \bar{w} for a tangent vector $w \in T_q F$ (as discussed in remark 4.1), we

- reinterpret the differential at p df_p as a linear map from $T_p X$ to F ,
- always use the standard chart (id_F, F) so that $\hat{f} = f_{U, F}$.

In this context, morphisms into \mathbb{R} almost serve as test functions in the framework of distribution theory. This requires a definition.

Definition 1.1: Function on X

Let X be a manifold of class C^p over \mathbb{R}^n for $n, p \geq 1$. A *function* on X is a morphism $f : X \rightarrow \mathbb{R}$, where \mathbb{R} should be interpreted as a manifold. We denote the commutative ring of functions on X by $C^p(X, \mathbb{R})$ or $C^p(X)$. If U is an open subset of X , its functions are denoted by $C^p(U, \mathbb{R})$ or $C^p(U)$.

For the rest of this chapter, assume all manifolds to be C^p -manifolds over \mathbb{R}^n , where $n, p \geq 1$.

Derivations

Let E and F be Banach spaces and $U \subseteq E$, suppose f is a morphism from U to F . If p is a point in U , $Df(p)$ is of course a linear map from E to F ; this suggests a natural pairing $\hat{\mathcal{D}}$ of f with and $(p, v) \in U \times E$ as shown in eq. (97).

$$\hat{\mathcal{D}} : (U \times E) \times C^p(U, F) \longrightarrow F : \quad ((p, v), f) \mapsto Df(p)(v) \in F \quad (97)$$

Suppose $F = \mathbb{R}$ and denote pointwise multiplication on \mathbb{R} by m . The above pairing trivially satisfies the product rule displayed in eq. (98).

$$Dm(f_{\underline{k}})(p)(v) = \sum_{i=\underline{k}} m(f_{i-1}(p), Df_i(p)(v), f_{i+k-i}(p)) \quad (98)$$

where $f_{\underline{k}} \in C^p(U, \mathbb{R})$. Next, if f is a function (from a manifold X) defined on an open neighbourhood U of p . If $v \in T_p X$, the commentary in the introduction suggests a 'duality pairing' between f and (p, v) in the form of eq. (99).

$$\mathcal{D} : (U \times E) \times C^p(U, F) \longrightarrow F : \quad \mathcal{D}((p, v), f) = df_p(v) \quad (99)$$

By definition of the differential df_p , the right hand side of eq. (99) is representation independent, hence

$$\mathcal{D}((p, v), f) = D\hat{f}(\hat{p})(\hat{v}), \quad \text{where the right member is an ordinary derivative} \quad (100)$$

for any representation (\hat{p}, \hat{v}) , \hat{f} . We also see that $\mathcal{D}((p, v), f) = \hat{\mathcal{D}}((\hat{p}, \hat{v}), \hat{f})$, which shows functions defined on U are dual to $T_p X$ for each $p \in U$. We will make this notion precise when we introduce covectors.

Definition 2.1: Derivation at p

A *derivation at p* is a **linear functional** v on $C^p(U, \mathbb{R})$, where U is any neighbourhood of p ; such that for $\underline{f}_k \in C^p(U)$, eq. (101) holds.

$$v(m(\underline{f}_k)) = \sum_{i=\underline{k}} m(\underline{f}_{i-1}(x), v(\underline{f}_i), \underline{f}_{i+k-i}(x)) \quad (101)$$

We will denote the space of derivations at p by $\mathcal{D}_p(X)$, and if $v \in \mathcal{D}_p(X)$, we say v *derives* f for any function f defined about p .

We have shown every tangent vector is a derivation, since the product rule descends from eq. (98) and its computation in coordinates in eq. (100). If X is finite-dimensional, prop. 2.1 shows derivations at a point $p \in X$ are uniquely represented by a tangent vector.

Proposition 2.1: $T_p X$ is isomorphic to $\mathcal{D}_p(X)$

Let p be a point on a manifold X , then its tangent space is isomorphic to the vector space of derivations. If (\hat{p}, \hat{v}) is a concrete tangent vector, its derivation of f computed using eq. (100).

Proof. Postponed. ■

Boundary