Contents

Chapter 0: Preliminaries

This section serves to recall a few results from [1, 2, 6, 4, 5], as well as to define the symbols and notation we will use.

Vector Spaces

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$.

- We say V is a $\mathbb{R}$-vector space or $\mathbb{C}$-vector space. If V is already understood to be a vector space, we say V is $\mathbb{R}$ or $\mathbb{C}$.
- We say V is *finite dimensional* whenever V admits a finite ordered basis. In this case, we say V is a *finite-dimensional vector space* (hereinafter abbreviated as FDVS).
- If $\{v_\alpha\} \subseteq V$, the symbol $\sum^\wedge v_\alpha$ refers to a partially specified object representing any **finite** sum of the $\{v_\alpha\}$ defined in eq. (1).

$$\sum^{\wedge} v_\alpha \in \left\{ \sum v_{i \leq k}, \quad k \leq \operatorname{card}\{v_\alpha\} \right\} \tag{1}$$

- If V is $\mathbb{C}$ (resp. $\mathbb{R}$), a *finite linear combination* (hereinafter abbreviated as FLC) of $\{v_\alpha\}$ is a partially specified object defined in eq. (2).

$$\sum_{\substack{\mathbb{C} \\ \text{(resp. } \mathbb{R})}}^{\wedge} v_\alpha \in \left\{ \sum_{i \leq k} c^i v_i, \quad k \leq \operatorname{card}\{v_\alpha\},\ c^i \in \mathbb{C} \text{ (resp. } \mathbb{R}) \right\} \tag{2}$$

- If V is a $\mathbb{C}$ vector space, a *real linear combination* of the subset $\{v_\alpha\}$ is a partially specified object, denoted by $\sum_{\mathbb{R}}^\wedge v_\alpha$. It is defined in eq. (2) by viewing V as a vector space over $\mathbb{R}$.

Let V and W be vector spaces over the same field $\mathbb{R}$ or $\mathbb{C}$.

- The *convex combination* of two elements $x_1, x_2 \in V$ is the linear combination

$$\mathfrak{c}_t(x_1, x_2) = x_1 + t(x_2 - x_1) \quad t \in [0, 1]$$

- A mapping $f : V \to \mathbb{R}$ is *convex* whenever

$$\mathfrak{c}_t(x_1, x_2) \leq c_t(f(x_1), f(x_2)) \quad \forall t \in [0, 1],\ x_1, x_2 \in V$$

- A mapping $f : V \to \mathbb{R}$ is a *subadditive* whenever 'adding inside is less than adding outside'. That is, $f \circ \sum^\wedge \leq \sum^\wedge \circ f$.
- A mapping $f : V \to W$ is *linear* whenever FLCs commute with f. This is written as $\sum_K^\wedge \circ f = f \circ \sum_K^\wedge$.

It is useful to have the following generalization when V and W are vector spaces over different base fields.

- If V is a $\mathbb{C}$-vector space and W a $\mathbb{R}$-vector space, a mapping $f : V \to W$ is said to be *linear* whenever $\sum_{\mathbb{R}}^\wedge$ commutes with f. In symbols, $f \circ \sum_{\mathbb{R}}^\wedge = \sum_{\mathbb{R}}^\wedge \circ f$.

If V is the vector space direct sum of W_1 and W_2, a vector $x \in W_i$ is *essentially in* W_i if it is invariant under the canonical projection of $\pi_i V \to W_i$. That is, $\pi_i(x) = x$. Equivalently, the element $x \in V$ is expressed as the linear combination of $x + 0 \in W_1 \oplus W_2$.

Enumeration of lists

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

- Individual elements: $v_{\underline{k}}$ means $v_1, \ldots, v_k$ as separate elements.
- Creating a k-list: $(v_{\underline{k}}) = (v_1, \ldots, v_k) \in \prod E_{j \leq k}$ if $v_i \in E_i$ for $i = \underline{k}$.
- Nested indices: $(v_{\underline{n_k}}) = (v_{n_{\underline{k}}}) = (v_{n_1}, \ldots, v_{n_k})$, and $(v_{\underline{n_k}}) \neq (v_{n_{(1,\ldots,k)}})$
- Closest bracket: $(v_{(n_{\underline{k}})}) = (v_{(n_1,\ldots,n_k)})$ and $(v_{n_{(\underline{k})}}) = (v_{n_{(1,\ldots,k)}})$
- Underlining 0 = empty: $(v_{\underline{0}}, a, b, c) = (a, b, c)$
- Skipping an index: $(v_{\underline{i-1}}, v_{i+\underline{k-i}}) = (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ for $i = \underline{k}$.
- Applying f to an element: $(v_{\underline{i-1}}, f(v_i), v_{i+\underline{k-i}}) = (v_1, \ldots, v_{i-1}, f(v_i), v_{i+1}, \ldots, v_k)$. Of course, if $i = 1$, then the above expression reads $(f(v_1), v_2 \ldots, v_k)$ 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 $\wedge : E \times E \to F$ is any associative binary operation: $\bigwedge(v_{\underline{k}}) = v_1 \wedge \cdots \wedge v_k$.

Example 2.1: 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 $i = \underline{k}$.

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

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

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

Metric Vector Spaces

Let V be a vector space over $\mathbb{R}$.

- A *bilinear form* $\omega : V \times V \to \mathbb{R}$ is a 2-tensor on V. (does not mean alternating, contrary to the notion of a differential form)
- A bilinear form on V is
  - *symmetric* if $\omega(x, y) = \omega(y, x)$ for all x, y.
  - *skew-symmetric* or *anti-symmetric* if $\omega(x, y) = (-1)\omega(y, x)$ for all x, y.

  – *alternating* if $\omega(x,x) = 0$ for all x.

The last two conditions are equivalent. Let V be a vector space over $\mathbb{R}$ with a bilinear form ω, then

- V is called a(n) *orthogonal geometry* (resp. *symplectic geometry*) if ω is symmetric (resp. alternating).
- V is called a *metric vector space* (hereinafter abbreviated as MVS) if it is an orthogonal or a symplectic geometry.

Matrices and bilinear forms

Definition 4.1: Matrix of bilinear form

If $B = (b_1, \ldots, b_n)$ is an ordered basis for V, we define the *matrix representation of* ω by

$$\mathcal{M}(\omega) = (a_{ij}) = (\omega(b_i, b_j))$$

Let $A = (a_{ij})$ be a matrix on V with respect to some basis $B = (b_{\underline{n}})$ it is clear that A induces a bilinear form, on V through $A(x,y) = [x]_B^T A[y]_B$, where $[\cdot]_B$ denotes the canonical isomorphism $V \cong \mathbb{R}^n$ with respect to the basis B.

$$[x]_B^T A[y]_B = \begin{bmatrix} x^1 & \ldots & x^n \end{bmatrix} A \begin{bmatrix} y^1 \\ \vdots \\ y^n \end{bmatrix}$$

for $x = x^i b_i$ and $y = y^j b_j$. Moreover,

$$A[x]_B = \begin{bmatrix} A(b_1, x) \\ \vdots \\ A(b_n, x) \end{bmatrix}$$ is a *column* vector whose entries are given by applying x on the second coordinate

and

$$[x]_B^T A = \begin{bmatrix} A(x, b_1) & \cdots & A(x, b_n) \end{bmatrix}$$ is a *row* vector whose entries are given by applying x on the first coordinate

Let A_B be the matrix representation of ω with respect to the B, if C is another basis on V, then how do we compute A_C? The answer is simple, recall for any vector $x \in V$, $x = x_B^i b_i$ and $x = x_C^j c_j$, then

$$[x]_B = M_{C,B}[x]_C \quad \text{for some matrix of an automorphism } M_{C,B}$$

$\omega(x,y) = [x]_B^T A_B [y]_B = ([x]_C^T M_{C,B}^T) A_B (M_{C,B}[y]_C) = [x]_C^T A_C [y]_C$, then

$$M_{C,B}^T A_B M_{C,B} = A_C \tag{3}$$

We can describe this relation between the two matrices A_B and A_C by the following

Definition 4.2: Congruent matrices

Two matrices M and N are said to be *congruent*, if there exists an invertible matrix P for which

$$P^T M P = N$$

Congruence is an equivalence relation on the space of matrices, and the equivalence classes over congruence are called *congruence classes*.

Orthogonality

For this section, (V, ω) will denote a finite dimensional metric vector space.

- A vector $x \in V$ is orthogonal to another vector $y \in V$, written $x \perp y$, if $\omega(x, y) = 0$.
- If V is an orthogonal or symplectic geometry then $\perp$ is a symmetric relation. If E is a subset of V, we denote the *orthogonal complement of* E by $E^{\perp} \triangleq \left\{ v \in V,\ v \perp E \right\}$

Let V be a metric vector space.

- A nonzero vector $x \in V$ is *isotropic*, or *null* if $\omega(x, x) = 0$
- V is *isotropic* if it contains at least one isotropic vector.
- V is *anisotropic* or *nonisotropic* if for every $x \in V$, $\omega(x, x) = 0 \implies x = 0$,
- V is *totally isotropic* or *symplectic* if $\omega(x, x) = 0$ for every vector $x \in V$.

The first bullet point above is about vectors in V, while the others are properties of V.

- A vector $x \in V$ is called *degenerate* if $x \perp V$, that is, $\forall y \in V$, $\omega(x, y) = 0$
- The *radical* of V, denoted by $\text{rad}(V) = V^{\perp}$ is the set of all degenerate vectors in V.
- V is *singular* or *degenerate* if $\text{rad}(V) \neq \{0\}$,
- V is *non-singular* or *non-degenerate* if $\text{rad}(V) = \{0\}$,
- V is *totally singular*, if $\text{rad}(V) = V$.

To summarize the above:

- V is isotropic if there exists a non-zero isotropic vector, meaning $\omega(x, x) = 0$, for some $x \neq 0$,
- V is degenerate if there exists a degenerate vector, $x \perp V$.

Lemma 5.1: Characterisation of bilinear forms

Let V be a finite dimensional vector space over $\mathbb{R}$ and let ω be a bilinear form on V. Then, the following properties of the matrix representation of ω with respect to ordered bases of V are invariant under congruence.

- non-singularity,
- symmetry,
- skew symmetry

If (ω_{ij}) is its induced matrix representation relative to any ordered basis then,

- ω is non-singular iff (ω_{ij}) is non-singular.

- ω is symmetric iff (ω_{ij}) is symmetric.
- ω is skew-symmetric iff (ω_{ij}) is skew-symmetric.

Proposition 5.1: Riesz Representation Theorem

Let (V, ω) be a nonsingular metric vector space, the map $x \mapsto x \lrcorner \omega \in V^*$ defined by

$$x \lrcorner \omega = \omega(x, \cdot), \quad \text{and} \quad (x \lrcorner \omega)(y) = \omega(x, y), \quad \forall y \in V$$

is a linear isomorphism from V to V^*.

Let (V, ω) and (W, η) be metric vector spaces. An *isometry* $\tau \in L(V, W)$ is a linear isomorphism that preserves the bilinear form.

$$\omega(u, v) = \eta(\tau u, \tau v)$$

Definition 5.1: Orthogonal, symplectic groups

Let V be a nonsingular metric vector space. If V is an orthogonal (resp. symplectic) geometry, the set of all isometries on V is called the *orthogonal (resp. symplectic) group on V*. It is a group under composition, and is denoted by $\mathcal{O}(V)$ (resp. $\mathrm{Sp}(V)$).

Remark 5.1: Assume all metric vector spaces are non-singular

We assume all MVS are non-singular unless specified otherwise.

Linear Algebra

The space of $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m \times n}$. If $A \in \mathbb{R}^{m \times n}$, we denote its entries by $[A]_{ij}$ for $i = \underline{m}$ and $j = \underline{n}$. Conversely, we define the matrix A with entries a_{ij} by writing $A = (a_{ij})$

Let $(e_{\underline{n}})$ be the standard ordered basis of $\mathbb{R}^n$, and $(\varepsilon^{\underline{n}})$ be its induced dual basis. If $A \in \mathbb{R}^{n \times n}$ and $a_{ij} = [A]_{ij}$, A defines a covariant 2-tensor (also denoted by a) in eq. (4).

$$a(e_i, e_j) = a_{ij} \in \mathbb{R} \quad \text{extended by linearity} \tag{4}$$

With this, we denote the tensor product of between ε^i and ε^j by $\varepsilon^i \otimes \varepsilon^j \in L(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R})$. Recall,

$$(\varepsilon^i \otimes \varepsilon^j)(e_k, e_l) = \delta^i_k \delta^j_l = \delta^{(i,j)}_{(k,l)} \quad \text{extended by linearity} \tag{5}$$

We can write $A \cong a = \sum_{i,j=\underline{n}} a_{ij} \varepsilon^i \otimes \varepsilon^j$. If $v = \sum_{i=\underline{n}} v^i e_i$, then

$$Av = a(\cdot, v) = \sum_{i=\underline{n}} \varepsilon^i a_{ij} v^j$$

If $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$, we can write $a_{ij} = \langle e_i, Ae_j \rangle$. , and if we allow ourselves to write $e^i = e_i$, then

$$Av = \sum_{i=\underline{n}} e^i \langle e^i, Av \rangle = \sum_{i=\underline{n}} e^i a_{ij} v^j$$

If $x_i \in \mathbb{R}^n$ for $i = \underline{n}$, we denote the matrix with x_i as columns by $(x_{\underline{n}})$, and its determinant by $\det\left(x_{\underline{n}}\right)$.

Summation Notation

We will still be working in $\mathbb{R}^n$, and let $a, A, v, w, (e_i), (\varepsilon^i)$ be as in the previous section. The *summation convention* is compact way of writing matrix (tensor) multiplication in coordinates, summed up in the following sentence.

> Upper and lower indices are paired together and summed over the dimension of the vector space.

This mantra however gives little motivation as to why it is a good piece of notation, nor does it actually give any new insights into tensor/exterior algebra.

Some advice for understanding summation notation.

- **Summation notation is not about summation**, it is a way of representing the linear-algebraic (partial or full) evaluation of bilinear forms (or tensors) in terms of matrix (tensor) coefficients.

$$\langle v, Aw\rangle = v^i w^j a_{ij}$$

  Observe that v lies in the first coordinate of the inner product on the left hand side of the equation, and that on the right hand side we see that its coefficients are paired with the first index of the matrix entries a_{ij}, similarly for w.

- Indices are paired **vertically**, other than that the **numerical** information is in **horizontal placement** of the indices that are being summed over. i.e:

$$a^{opq}_{ijkl} v^j w^l t_q = \sum_{j,l,q=\underline{n}} a^{opq}_{ijkl} v^j w^l t_q$$

We will list a few examples.

The first (and only) coordinate is summed over.

- Dot Product: $v \cdot w = v_i w^i = \sum_{i=\underline{n}} v^i w^i$.
- Vector basis expansion: $v = v^i e_i$ if $v = \sum_{i=\underline{n}} v^i e_i$.
- Covector basis expansion: $B = b_i \varepsilon^i$ if $B = \sum_{i=\underline{n}} b_i \varepsilon^i$.

Both indices in the matrix entries are summed over.

- Full Evaluation of a Bilinear Form: $a(v, w) = v^i w^j a_{ij} = \langle v, Aw\rangle_{\mathbb{R}^n}$.
- Transposition = permutation of indices in entries: $a(w, v) = v^i w^j a_{ji} = \Big\langle w,\ Av \Big\rangle_{\mathbb{R}^n}$
- Matrix Inverse = raise indices of entries: $A^{-1} = (a^{ij})$. Where $AA^{-1} = (a_{ij})(a^{kl})$ is equal to

$$a_{ij} a^{jl} \varepsilon^i \otimes \varepsilon^l = \delta^i_l \varepsilon^{(i,l)}$$

Partial pairing of indices. (Again: focus on the position.)

- Ordinary matrix multiplication: $Av = a_{ij} v^j e^i$ Here, A is a linear operator on $\mathbb{R}^n$.

- Partial evaluations of a bilinear form a.

$$a(\cdot, v) = Av = v^j a_{ij} \varepsilon^i \quad \text{or} \quad a(v, \cdot) = v^T A = v^i a_{ij} \varepsilon^j$$

- $\hat{a}(v)$ is given by eq. (6) in coordinates.

$$\hat{a}(v) = \sum_{i=\underline{n}} \nu_i \varepsilon^i = a_{ji} v^j \varepsilon^i \tag{6}$$

In eq. (6), the coefficients (v_i) get paired with the first index in a_{ji}, which represents multiplication of A 'from the left' — which is *precisely* what the matrix transpose does.

Example 7.1: Advanced Example

Let F be a $(0, k)$-tensor on $\mathbb{R}^n$. If $I = (i_{\underline{k}})$ is a multi-index with entries $1 \leq i_j \leq n$ for $j = \underline{k}$, we write

$$F_I = F(e_{i_{\underline{k}}}) = F(e_{i_1}, \ldots, e_{i_k})$$

as the number obtained by evaluating F at the basis vectors $e_{i_{\underline{k}}}$. F is then the linear combination of F_I and $\varepsilon^I = \bigotimes \varepsilon^{(i_{\underline{k}})}$, in summation convention

$$F = F_I \varepsilon^I = F_{(i_{\underline{k}})} \bigotimes \varepsilon^{(i_{\underline{k}})}$$

Let $L = (l_{\underline{p}})$ where $p \leq k$ be a multi-index whose entries satisfy the condition two paragraphs above, and are increasing (this means $l_1 < \cdots < l_p$). We wish to compute the partial evaluation of F when the l_pth argument is held at $x_{l_p} \in \mathbb{R}^n$. Intuitively, the result should be a $(0, k-p)$-tensor on $\mathbb{R}^n$. However, we will need one more piece of notation that describes the partial 'evaluation' or multi-indices.

Let I be a k-index and L an increasing p-index with entries $1 \leq l_r \leq k$ for $r = \underline{p}$ and $p \leq k$. We define the *contraction* of I in L to be a $(k-p)$-index — $L \lrcorner I$ — to be I but with the entries at $l_{\underline{p}}$ removed.

For example: if $L = (1, 2)$, then $L \lrcorner I = (i_{2+\underline{k-2}})$ is the same multi-index but with its first two entries removed.

Finally, we see that

$$x_L \lrcorner F = \textcircled{\scriptsize{Π}} (x_{l_{r=\underline{p}}}^{i_{l_r}}) F_I \varepsilon^{L \lrcorner I}$$

Musical Isomorphisms

Let $n \geq 1$ be a non-negative integer. Let $a : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a bilinear form that makes $(\mathbb{R}^n, a)$ a metric vector space.

If a has matrix representation $A = (a_{ij})$, $v = \sum_{i=\underline{n}} v^i e_i$, and $w = \sum_{i=\underline{n}} w^i e_i$, then

$$v \lrcorner a(w) = a(v, w) \quad \text{by definition}$$

Since $v \lrcorner a$ is an element in the dual space, it can be written as $\sum_{k=\underline{n}} \nu_k \varepsilon^k$. We wish to compute the coefficients ν_k, and the left hand side becomes

$$\sum_{k=\underline{n}} \nu_k \varepsilon^k \left(\sum_{i=\underline{n}} w^i e_i \right) = \sum_{i,k=\underline{n}} \nu_k w^i = \Big\langle (\nu_k),\ (w^i) \Big\rangle_{\mathbb{R}^n}$$

Similarly, the RHS reads

$$a(v,w) = \sum_{i,j=\underline{n}} v^i w^j a_{ij} = \Big\langle (v^i),\ (a_{ij})(w^j) \Big\rangle_{\mathbb{R}^n}$$

We see that $(\nu_k) = A^T(v_j)$. We now define

Definition 8.1: Musical isomorphism $\breve{a}$

Let $(\mathbb{R}^n, a)$ be a MVS, every vector $v \in \mathbb{R}^n$ induces a covector, denoted by $\breve{a}(v)$ such that the bilinear form a *becomes* the evaluation map.

$$\breve{a}(v) = v \lrcorner a \quad \text{such that} \quad \breve{a}(v)(w) = a(v,w)\, \forall w \in \mathbb{R}^n \tag{7}$$

The mapping $\breve{a}$ is called a *musical isomorphism*. We sometimes write $\breve{a}(v) = v^\flat$ if the ambient MVS interpretation is understood.

We can write eq. (7) in coordinates by appealing to the *summation convention*.

Conversely, by the Riesz Representation Theorem (see prop. 5.1), covector $B \in (\mathbb{R}^n)^*$ can be uniquely identified with a vector $b \in \mathbb{R}^n$. Since a is non-singular, its matrix representation (a_{ij}) is non-singular as well,

Definition 8.2: Musical isomorphism $\hat{a}$

Let $(\mathbb{R}^n, a)$ be a MVS, every covector $f \in \mathbb{R}^{n*}$ induces a vector, denoted by $\hat{a}(f) \in \mathbb{R}^n$ such that a becomes the evaluation map.

$$a\Big(\hat{a}(f),\ v\Big) = f(v) \quad \forall v \in \mathbb{R}^n \tag{8}$$

We sometimes write $\hat{a}(f) = f^\wedge$.

We can compute the musical isomorphisms in coordinates. Let (a_{ij}) be the matrix representation of a with matrix inverse (a^{ij}), then

$$\breve{a}(v) = a_{ij} v^j \varepsilon^i \quad \text{and} \quad \hat{a}(f) = a^{ij} f_j e_i$$

Let a be a non-singular bilinear form on a $\mathbb{R}$-FDVS V.

- we call $\breve{a} : V \to V^*$ the *flat map* of a, where $\breve{a}(x) = a(x, \cdot)$. We sometimes write $\breve{x} = \breve{a}(x)$, and
- we call $\hat{a} : V^* \to V$ the *sharp map* of a, where $\hat{a}(f)$ is a vector in V that satisfies $a(\hat{a}(f), x) = f(x)$. Equivalently, $\breve{a}$ is the two-sided inverse of $\breve{a}$. We sometimes write $\hat{f} = \hat{a}(f)$.

If f is a covector in V, **one thinks of the bilinear form a as being the evaluation map**, since $a(\hat{f}, x) = f(x)$.

Exterior Algebra

Let V be a n-dimensional $\mathbb{R}$-vector space with ordered basis $(e_{\underline{n}})$ and its induced dual basis $(\varepsilon^{\underline{n}})$. We begin with some semantics.

- If $k \geq 1$ is an integer, $\mathcal{T}^k(V) = \{f : V^k \to \mathbb{R},\ f \text{ is } k\text{-linear.}\}$. We refer to $\mathcal{T}^k$ as the space of *k-covariant tensors on V.*
- If $k \geq 1$ is an integer, $\Lambda^k(V) = \{f \in \mathcal{T}^k(V),\ f \text{ is alternating.}\}$. This means, if σ is in the k-permutation group, then $f(v_{\sigma(\underline{k})}) = \operatorname{sgn}(\sigma) f(v_{\underline{k}})$. We refer to $\Lambda^k(V)$ as the space of *alternating k-vectors on V.*
- If $k = 0$ then $\mathcal{T}^k(V) = \Lambda^k(V) = \mathbb{R}$.

The covectors $\varepsilon^{\underline{n}}$ are sometimes referred to as *elementary covectors.* We are ready to define the wedge product and discuss its properties.

- We define the *wedge product* between covectors ε^i and ε^j as the alternating 2-tensor that satisfies

$$\varepsilon^i \wedge \varepsilon^j(x, y) = \det\left(\begin{bmatrix} \varepsilon^i(x) & \varepsilon^i(y) \\ \varepsilon^j(x) & \varepsilon^j(y) \end{bmatrix}\right) \quad \forall x, y \in V$$

- If $I = (i_{\underline{k}})$ is a k-multi-index (or k-index for short), with entries in $\{\underline{n}\}$, we denote the wedge product between the k elementary covectors $\varepsilon^{i_1}, \dots, \varepsilon^{i_k}$ by

$$\varepsilon^I = \bigwedge(\varepsilon^{i_{\underline{k}}}) = \varepsilon^{i_1} \wedge \varepsilon^{i_2} \wedge \cdots \wedge \varepsilon^{i_k}$$

  which is an alternating k-tensor, whose action on vectors $v_{\underline{k}} \in V$ is defined by

$$\varepsilon^I(v_{\underline{k}}) = \det\left(\begin{bmatrix} \varepsilon^{i_1}(v_1) & \cdots & \cdots & \varepsilon^{i_1}(v_k) \\ \varepsilon^{i_2}(v_1) & \cdots & \cdots & \varepsilon^{i_2}(v_k) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \varepsilon^{i_k}(v_1) & \cdots & \cdots & \varepsilon^{i_k}(v_k) \end{bmatrix}\right)$$

- If I and J are k and l indices with entries in $\{\underline{n}\}$, the wedge product of ε^I and ε^J is defined to be

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{(i_{\underline{k}}, j_{\underline{l}})}$$

The space of k-covariant tensors on V (resp. alternating k-forms on V) form a $\mathbb{R}$-vector space of dimension n^k (resp. $\binom{n}{k}$).

- If $k \geq 1$, the following is a linear basis for $\mathcal{T}^k(V)$.

$$\mathcal{B}^k_{\mathcal{T}} = \left\{\bigotimes(\varepsilon^I),\ I \text{ is a } k\text{-index, with entries in } \{\underline{n}\}\right\} \tag{9}$$

  That is, every element in $\mathcal{T}^k(V)$ is the FLC of elements in $\mathcal{B}^k_{\mathcal{T}}$.

- A k-covariant tensor $f \in \mathcal{T}^k(V)$ is said to be *decomposable* if it is the k-tensor product of covectors (not necessarily elementary). That is,

$$f = \bigotimes(\alpha_{\underline{k}}) \quad \alpha_{\underline{k}} \in V^*$$

- If $k \geq 1$, the following is a linear basis for $\Lambda^k(V)$.

$$\mathcal{B}^k_\Lambda = \left\{ \bigwedge(\varepsilon^I),\ I \text{ is a } k\text{-index, with entries in } \{\underline{n}\} \right\} \tag{10}$$

  That is, every element in $\Lambda^k(V)$ is the FLC of elements in $\mathcal{B}^k_\Lambda$.

With this, we can extend the wedge product by (multilinearity) to arbitrary alternating forms. An alternating k-tensor is *decomposable* if it is the k wedge product of k covectors (that are not necessarily elementary).

For example, if $\omega = \bigwedge(\omega_{\underline{k}}) \in \Lambda^k(V)$ and $\eta = \bigwedge(\eta_{\underline{l}}) \in \Lambda^l(V)$, the wedge product between the two alternating tensor is a $k+l$ alternating tensor where

$$(\omega \wedge \eta)(v_{\underline{k}}, y_{\underline{l}}) = \bigwedge(\omega_{\underline{k},\eta_{\underline{l}}})(v_{\underline{k}}, y_{\underline{l}}) = \det\left(\begin{bmatrix} \begin{bmatrix} \omega_1(v_1) & \cdots & \omega_1(v_k) \\ \vdots & \vdots & \vdots \\ \omega_k(v_1) & \cdots & \omega_k(v_k) \end{bmatrix} & \begin{bmatrix} \omega_1(y_1) & \cdots & \omega_1(y_l) \\ \vdots & \vdots & \vdots \\ \omega_k(y_1) & \cdots & \omega_k(y_l) \end{bmatrix} \\ \begin{bmatrix} \eta_1(v_1) & \cdots & \eta_1(v_k) \\ \vdots & \vdots & \vdots \\ \eta_l(y_1) & \cdots & \eta_l(y_l) \end{bmatrix} & \begin{bmatrix} \eta_1(y_1) & \cdots & \eta_1(y_l) \\ \vdots & \vdots & \vdots \\ \eta_l(y_1) & \cdots & \eta_l(y_l) \end{bmatrix} \end{bmatrix}\right) \tag{11}$$

The block matrices on the off diagonal in eq. (11) are not square unless $k = l$. If that is the case, then the wedge product admits the obvious simplification. Let ω and η be alternating k and l tensors respectively (not necessarily decomposable) where $k, l \geq 1$. The following properties of the wedge product are derived entirely from eq. (11).

- Anticommutativity: $\eta \wedge \omega = (-1)^{kl}\omega \wedge \eta$. Sketch of proof: Assume all tensors involved are decomposable, and swap the columns in eq. (11). Extend by multilinearity.
- Bilinearity and associativity. Sketch of proof: Assume all tensors involved are decomposable, and compare the matrices that are used in the determinants. Extend by multilinearity.
- Interior multiplication: if $v_1 \in V$, we define the the alternating $k+l-1$ form $\iota_{v_1}(\omega \wedge \eta) \in \Lambda^{k+l-1}(V)$ by placing v_1 into the first argument of $\omega \wedge \eta$,

$$\iota_{v_1}(\omega \wedge \eta)(v_{1+\underline{k+l-1}}) = (\omega \wedge \eta)(v_{\underline{k+l}}) \quad \forall v_{1+\underline{k+l-1}} \in V$$

  satisfies $(\iota_{v_1}\omega) \wedge \eta + (-1)^k \omega \wedge (\iota_{v_1}\eta)$. Sketch: Assume decomposable, and use exmp. 2.1. Extend by multilinearity.

Vector and Covector Fields

Differential Forms

Let M be a n-dimensional $\mathbb{R}$-manifold of class C^p where $p \geq 2n+1$. A *differential form of rank k* (or k-forms for short) is a smooth section of the vector bundle $\Lambda^k(M) = \coprod_{x \in M} \Lambda^k(T_x M)$. The previous section shows that $\dim(\Lambda^k(M)) = \binom{n}{k}$. The space of k-forms on M is denoted by $\Omega^k(M)$

The *differential* or the covector field of a function $f \in C^p(M)$ is a $(0,1)$ tensor field on M, denoted by

$$df \in \mathfrak{X}^*(M) \quad \text{where} \quad df(p)(v) = \sum_{i=\underline{n}} \frac{\partial f}{\partial x^i}\bigg|_p v^i$$

for any tangent vector $v = v^i e_i \in T_pM$.

If $\alpha \in \Omega^k(M)$, we define the *exterior derivative* of α to be a section of the $k+1$ alternating tensor bundle. In coordinates, if $\alpha = \sum' \alpha_I dx^I$, then

$$d\alpha = d\alpha_I \wedge dx^I$$

Measure theory

Let $(X, \mathcal{M}, \mu)$ be a measure space. A measurable function is a measurable mapping $f : X \to \mathbb{C}$. we often write $f \in \mathcal{M}$ in an abuse of notation.

- $\mathcal{L}^+(X, \mu)$ for non-negative measurable functions, and $L^+(X, \mu)$ its quotient space.
- If $p \in [1, +\infty)$, $\mathcal{L}^p(X, \mu)$ for the 'L^p' functions and $L^p(X, \mu)$ its quotient space.

$$\mathcal{L}^p(X, \mu) = \left\{ f \in \mathcal{M},\ \int |f|^p < +\infty \right\} \quad \text{resp.} \quad L^p(X, \mu) = \mathcal{L}^p(X, \mu)/\text{equality a.e}$$

- In the context of L^p theory, we say p is *usual* if $p \in [1, +\infty)$, and p is *reflexive* whenever L^p is (meaning $1 < p < \infty$).
- A measurable function $\phi \in \mathcal{M}$ is *simple* whenever its range is a finite subset of $\mathbb{C}$.

$$\Sigma = \left\{ f \in \mathcal{M},\ f \text{ is simple.} \right\} \subseteq \mathcal{M}/\text{equality a.e}$$

- We denote the non-negative (resp. p-integrable) simple functions by $\Sigma^+ = \Sigma \cap L^+$ (resp. $\Sigma^p = \Sigma \cap L^p$).
- If E is a measurable set, the indicator on E is denoted by χ_E.

Some notation regarding the L^p spaces.

- Let $f \in L^p$ for usual p, we denote the L^p norm of f by $\|f\|_p$.
- If $p = +\infty$, then L^p is the space of measurable functions with finite essential supremum which we denote by $\|\cdot\|_\infty$.
- For clarity, we sometimes write $\|f\|_{L^p}$ instead of $\|f\|_p$.

Remark 12.1: Assumption of almost everywhere

In any measure theoretical setting, when we say f is 'L^p', we mean $f \in L^p$ unless otherwise stated. We also identify $\mathbb{B}_X$ with its quotient space.

Topology

Let X be a topological space, and $E \subseteq X$.

- The topological interior of $E \subseteq X$ is denoted by E^o.
- The topological closure of E is denoted by $\overline{E}$.
- If U is an open subset of X, we write $U \overset{\circ}{\subseteq} X$.
- A neighbourhood of a point $p \in X$ is a subset $U \subseteq X$ (not necessarily open) where $p \in U^o$.
- A subset of X is *precompact* whenever its closure is compact.
- If $A, B \subseteq X$, we say A *hides* in B whenever $\overline{A} \subseteq B^o$.

Furthermore,

- $\mathbb{B}_X$ = Borel σ-algebra of X.
- $C(X)$ = continuous, complex valued functions from X.
- $BC(X)$ = bounded, continuous complex-valued functions. It is endowed with the *uniform norm* in eq. (12).

$$f \mapsto \|f\|_u = \sup |f| \tag{12}$$

All functions hereinafter are assumed to be complex valued. The following function spaces are subsets of $BC(X)$, and inherits the same norm as in eq. (12).

- $C_c(X)$ = continuous functions with compact support.
- $C_0(X)$ = continuous vanishing functions, whose elements are defined in eq. (13).

$$C_0(X) = \left\{ f \in C(X), \text{ for } \varepsilon > 0 \text{ exists compact } K, \sup_{K^c} |f| \leq \varepsilon \right\} \tag{13}$$

- $UBC(X)$ = uniformly continuous functions, whenever X is a metric space.

We also use the following shorthand when discussing partitions of unity and bump functions. Let E be any subset of X, and $f \in C(X, [0,1])$.

- $E \lesssim f$ whenever $f = 1$ on E.
- $f \lesssim E$ whenever $\operatorname{supp}(f) \subseteq E$.

Partitions of Unity

Let X be a topological space. A (continuous) *partition of unity* on X is family of continuous functions $\{\varphi_\alpha\} \subseteq C(X, [0,1])$ where $\sum \varphi_\alpha \equiv 1$ and whose supports form a *locally finite* collection of subsets. That is, every point $p \in X$ admits a neighbourhood U such that U intersects finitely many of $\operatorname{supp}(\varphi)_\alpha$.

If $\{U_\alpha\}$ be an open cover of X, we say a partition of unity $\{\varphi_\alpha\}$ is *subordinate to* $\{U_\alpha\}$ whenever $\varphi_\alpha \lesssim U_\alpha$. We often place additional requirements on $\{\varphi_\alpha\}$, e.g $\{\varphi_\alpha\}$ is a compactly supported partition of unity whenever $\{\varphi_\alpha\} \subseteq C_c(X, [0,1])$.

X is said to *admit partitions of unity* of class C^p (resp. C_c) whenever every open cover $\{U_\alpha\}$ of X has a partition of unity $\{\varphi_\alpha\} \subseteq C^p(X, [0,1])$ (resp. C_c) subordinate to $\{U_\alpha\}$.

LCH Spaces

A topological space X is *locally compact* if every point $p \in X$ admits a compact neighbourhood. We say X is LCH (or X is a LCH space) whenever it is locally compact and Hausdorff. We sum up some useful facts about LCH spaces.

- Let K be compact and $K \subseteq U \mathring{\subseteq} X$. There exists a function $f \in C_c(X, [0,1])$ where $K \lesssim f \lesssim U$. Because of this, when we write $A \lesssim g \lesssim B$, it is convenient to assume that $g \in C_c(X, [0,1])$ whenever A is compact.
- With K, U being the same as above, every continuous function $f \in C(K)$ admits a compactly supported extension whose support hides in U. That is, there exists $\widetilde{f} \in C_c(U)$ such that $\widetilde{f} \lesssim U$, and $\widetilde{f}|_K = f$.
- LCH spaces are paracompact on their compact sets. If K is compact in X, for every finite open cover $\{U_j\}$ of K, there exists a continuous partition of unity on K subordinate to this open cover.

Banach Spaces

- A *normed vector space* (hereinafter abbreviated as NVS) is a vector space X with a norm $p \mapsto |p|$. We always use $|\cdot|$ to refer to the endowed norm of a NVS. A *Banach space* is a Cacuchy-complete NVS.
- An *inner product space* (hereinafter abbreviated as IPS) is a vector space X over K with an inner product $(x, y) \mapsto \langle x, y \rangle \in K$. It is also an NVS with the norm $|x| = \langle x, x \rangle^{1/2}$. A *Hilbert space* is a Cauchy-complete IPS.
- If X is a IPS, its inner product will always be denoted by $\langle \cdot, \cdot \rangle_X$ or $\langle \cdot, \cdot \rangle$ when it is unambigious to do so.

Let X be a Banach space over $K = \mathbb{R}$ or $\mathbb{C}$.

- The *dual* (or the dual space) of X is the Banach space of toplinear mappings into the base field K. We usually denote it by X^* or X'. The *bidual* of X is X^{**}.
- X is *reflexive* whenever it is toplinearly isomorphic to its bidual.

- The *weak topology* on X refers to the coarsest topology on X that makes the evaluation maps $\{\langle f, \cdot\rangle\}_{f\in X^*}$ continuous. Where,

$$\langle f, \cdot\rangle : X \to \mathbb{R} \quad \text{and} \quad \langle f, \cdot\rangle(x) = f(x)$$

- The *weak-$*$ topology* on X^* refers to the coarsest topology on X^* that makes the evaluation maps $\{\langle \cdot, x\rangle\}_{x\in X}$ continuous.
- The *duality pairing* between X and X^* is always denoted by $\langle \cdot, \cdot\rangle_X$ where elements in X are placed in the right hand side of the bracket.

Let X and Y be Banach spaces over $K = \mathbb{R}$ or $\mathbb{C}$.

- We say a map F is *between* the spaces X and Y if $F : X \to Y$.
- $\mathcal{L}(V^K, W)$ denotes the space of k-linear maps from V to W that are not necessarily continuous.
- $\mathcal{L}(X, Y)$ will denote the space of linear maps between X and Y.
- In the category of Banach spaces, the space of morphisms are called *toplinear morphisms* - or *CLM*s (*continuous linear maps*); which we will denote by $L(X, Y)$ for toplinear morphisms bewteen X and Y.
- 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 that,

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

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 *CLM*s (*continuous linear maps*); which we will denote by $L(E, F)$ for toplinear morphisms bewteen E and F.

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* $(x_1, x_2) \mapsto \max(|x_1|, |x_2|)$.

Proposition 16.1: 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.

$$A \subseteq [\lambda \leq \alpha - \varepsilon] \quad \text{and} \quad B \subseteq [\lambda \geq \alpha + \varepsilon] \quad \text{for all } \alpha \in \mathbb{R} \text{ and } \varepsilon > 0. \tag{14}$$

Definition 16.1: 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

$$(x_1, \dots, x_k) \mapsto |(x_1, \dots, x_k)| = \sup_{1 \leq i \leq k} |x_i| \tag{15}$$

The following are natural generalizations of Banach spaces.

- A *topological vector space* (hereinafter abbreviated as TVS) is a vector space X over a field $K = \mathbb{R}$ or $\mathbb{C}$ such that the addition map $A(x, y) = x + y$ and the scalar multiplication map $m(k, x) = kx$ is continuous.
- A TVS is *locally convex* if it admits a basis of convex sets.
- A *Frechet Space* is a Cauchy-complete (in terms of Cauchy nets), Hausdorff TVS whose topology is defined by a countable family of seminorms.

Let X and Y be TVS whose topologies are defined by the families (not necessarily countable) $\{p_\alpha\}$ and $\{q_\beta\}$ of seminorms. A linear mapping $F : X \to Y$ is toplinear if and only if

for each β, there exists **finitely many** $(\alpha_{\underline{k}})$ and a constant $C > 0$ such that

$$q_\beta(F(x)) \leq C \sum p_{\alpha_{\underline{k}}}(x) \quad \forall x \in X$$

Functions on Euclidean Space

We turn to the case where $X = \mathbb{R}^n$, where we adopt the following terminology.

- L^1_{loc} = quotient space of locally integrable functions. If $f \in L^1_{loc}$ then $f\chi_K \in L^1$ for every bounded measurable K.
- $C^k = C^k(\mathbb{R}^n)$ the space of k times continuously differentiable functions, where $k \geq 0$.
- $C^k_0 = C_0 \cap C^k$ for $k \geq 0$. It is endowed with the norm in eq. (16) that makes it a Banach space.

$$f \mapsto \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_u \tag{16}$$

- $C^\infty = C^\infty(\mathbb{R}^n)$ = smoothly differentiable, complex-valued functions.
- $C^\infty_c = C^\infty_c(\mathbb{R}^n)$ = compactly supported smooth functions.

Let $E \subseteq \mathbb{R}^n$ be any subset.

- $C^\infty_c(E) = \left\{ f \in C^\infty_c(\mathbb{R}^n),\ \operatorname{supp}(f) \subseteq E \right\}$ = compactly supported smooth functions whose support is contained within E.

- $\mathcal{S}$ = Schwartz space, defined in eq. (17) is the space of *rapidly decreasing* smooth functions.

$$\mathcal{S} = \left\{ f \in C^\infty,\ \|f\|_{(N,\alpha)} < +\infty \text{ for all } N, \alpha \right\} \tag{17}$$

  where $\|f\|_{(N,\alpha)} = \sup_x (1+|x|)^N |\partial^\alpha f(x)|$.

- $C_s^\infty = C_s^\infty(\mathbb{R}^n)$ = *slowly increasing* smooth functions, defined in eq. (18).

$$C_s^\infty = \left\{ f \in C^\infty,\ |\partial^\alpha f(x)| \lesssim_\alpha (1+|x|)^{N_\alpha} \right\} \tag{18}$$

We write $\mathcal{E} = C^\infty$ (resp. $\mathcal{E}(E)$ for $E \subseteq \mathbb{R}^n$). If K is a compact subset of $\mathbb{R}^n$, then $\mathcal{E}(K)$ is a Frechet Space with the norms in eq. (19)

$$\phi \mapsto \|\partial^\alpha \phi|_K\|_u = \|\partial^\alpha \phi\|_u \tag{19}$$

where α ranges over all multi-indices of length n.

We write $\mathcal{D} = C_c^\infty$ (resp. $\mathcal{D}(E)$ for $E \subseteq \mathbb{R}^n$) and recall that $\mathcal{D}$ is equipped with the canonical LF topology which means it locally borrows the open sets of $\mathcal{E}(K)$ where $K \subseteq U$ is compact.

- A sequence $\{\phi_j\} \subseteq \mathcal{D}(U)$ where $U \mathring{\subseteq} \mathbb{R}^n$ converges to some $\phi \in \mathcal{D}$ whenever $\{\phi_j\} \subseteq \mathcal{D}(K)$ for compact K and $\phi_j \to \phi$ in $\mathcal{D}(K)$.
- A linear mapping $F : \mathcal{D}(U) \to Y$ where Y is a Banach space is continuous whenever $F|_{\mathcal{E}(K)} : \mathcal{E}(K) \to Y$ is toplinear for every compact $K \subseteq U$.
- A linear mapping $F : \mathcal{D}(U) \to \mathcal{D}(U')$ where $U' \mathring{\subseteq} \mathbb{R}^n$ is continuous, if the restriction of F onto $\mathcal{E}(K)$ (where $K \subseteq U$ compact) has range

$$F(\mathcal{E}(K)) \subseteq \mathcal{E}(K') \quad K' \text{ compact, } K' \subseteq U'$$

  and $F|_{\mathcal{E}(K)}$ is toplinear.

Definition 17.1: Slowly increasing sequences

$C_s(\mathbb{Z}^n)$ is the space of slowly increasing sequences with domain $\mathbb{Z}^n$,

$$C_s(\mathbb{Z}^n) = \left\{ g : \mathbb{Z}^n \to \mathbb{C},\ |g(k)| \lesssim_g (1+|k|)^N,\ N \in \mathbb{N}^+ \right\}$$

Fourier Transforms

The *Fourier Transform of a function* (a.e class, or pointwise) f is defined by the integral in eq. (20).

$$\mathcal{F}f(\zeta) = \hat{f}(\zeta) = \int_{\mathbb{R}^n} f(x) E_{-\zeta}(x) dx \quad \zeta \in \mathbb{R}^n \tag{20}$$

where $E_{-\zeta}(x) = e^{-2\pi i \langle \zeta, x \rangle}$.

- The integral in eq. (20) converges whenever $f \in L^p$ for $1 \le p \le 2$.

- $\|\hat{f}\|_q \leq \|f\|_p$ for $1 \leq p \leq 2$ and q conjugate to p.
- In particular, $\|\hat{f}\|_2 = \|f\|_2$.

The *periodic* Fourier Transform of a measurable function $f : \mathbb{T}^n \to \mathbb{C}$ is the map in eq. (124).

$$\mathcal{F}f(k) = \int_{\mathbb{T}^n} f(x)E_{-k}(x)dx \quad k \in \mathbb{Z}^n \tag{21}$$

If f is in $L^2(\mathbb{T}^n)$, eq. (124) simplifies to

$$\mathcal{F}f(k) = \hat{f}(k) = \langle f, E_k \rangle_{L^2(\mathbb{T}^n)}$$

We list some properties of Fourier Transforms on L^1, Let $f, g \in L^1(\mathbb{R}^n)$.

- Translations: $(\tau_y f)^\wedge(\zeta) = E_{-y}(\zeta)\hat{f}(\zeta)$, and $(\tau_y \hat{f})(\zeta) = (E_y f)^\wedge(\zeta)$.
- If $M \in GL(n)$, then $(f \circ M)^\wedge = \hat{f} \circ M^{-T}$, where M^{-T} is the inverse of the adjoint map.
- Convolutions: $(f * g)^\wedge = \hat{f}\hat{g}$
- Riemann Lebesgue Lemma: If $f \in L^1$, then $\hat{f} \in C_0$.

An important property of $\mathcal{F}$ is that it diagonalizes differentiation.

- Integrability transforms into regularity: $x^\alpha f \in L^1$ for $|\alpha| \leq k$, then $\hat{f} \in C_0^k$,
- Multiplication by coordinate functions transforms into differentiation: $\partial^\alpha \hat{f} = [(-2\pi i x)^\alpha f]^\wedge$, whenever the previous condition is satisfied.
- Regularity transforms into integrability: $f \in C_0^k$, and $\partial^\alpha \in C_0 \cap L^1$ for all $|\alpha| \leq k-1$, then $\zeta^\alpha \hat{f} \in L^1$.
- Differentiation transforms into multiplication by coordinate functions: $(\partial^\alpha f)^\wedge(\zeta) = (2\pi i \zeta)^\alpha \hat{f}(\zeta)$, whenever the previous condition is satisfied.

The *inverse Fourier Transform* is the integral in eq. (22)

$$\mathcal{F}^{-1}f(x) = \check{f}(x) = \hat{f}(-x) = \int_{\mathbb{R}^n} f(\zeta)E_x(\zeta)d\zeta \tag{22}$$

We also have the following isomorphisms.

- $\mathcal{F}$ is a linear automorphism on $\mathcal{S}$
- $\mathcal{F}$ is a unitary isomorphism on $L^2(\mathbb{R}^n)$.
- $\mathcal{F}$ is a unitary isomorphism between $L^2(\mathbb{T}^n)$ and $l^2(\mathbb{Z}^n, \mathbb{C})$.

Distributions

Let U be an open subset of $\mathbb{R}^n$.

A *distribution on* U is a continuous linear functional $F : \mathcal{D}(U) \to \mathbb{R}$ such that $\lim\langle F, \phi_j\rangle_{\mathcal{D}} = \langle F, \lim \phi_j\rangle_{\mathcal{D}}$.

The space of distributions on U is denoted by $\mathcal{D}'(U)$ and has the weak-$*$ topology, where $\lim F_n = F$ if and only if

$$\lim\langle F_n, \phi\rangle_{\mathcal{D}} = \langle \lim F_n, \phi\rangle_{\mathcal{D}} \quad \forall \phi \in \mathcal{D}$$

We also define several operations on $\mathcal{D}'$. Let $F \in \mathcal{D}'$ and $\phi \in \mathcal{D}$.

- Differentiation: $\langle \partial^\alpha F, \phi\rangle_{\mathcal{D}} = (-1)^{|\alpha|}\langle F, \partial^\alpha \phi\rangle_{\mathcal{D}}$
- Multiplication by Smooth Functions: If $g \in \mathcal{E}$, we define $\langle Fg, \phi\rangle_{\mathcal{D}} = \langle F, g\phi\rangle_{\mathcal{D}}$.
- Translation: Let $y \in \mathbb{R}^n$, and $\langle \tau_y F, \phi\rangle_{\mathcal{D}} = \langle F, \tau_{-y}\phi\rangle_{\mathcal{D}}$.
- Reflection: If $\widetilde{F}$ is the reflection of F about the origin, its action on ϕ is $\langle \widetilde{F}, \phi\rangle_{\mathcal{D}} = \langle F, \widetilde{\phi}\rangle_{\mathcal{D}}$ — where $\widetilde{\phi}(x) = \phi(-x)$.
- Convolutions: We define a new pointwise function $(F * \phi)$ such that $(F * \phi)(x) = \langle F, \tau_x \widetilde{\phi}\rangle_{\mathcal{D}}$.

Remark 19.1: Reflections and Translations

The function $\tau_x\widetilde{\phi}$ should be interpreted as the translation of the reflection of ϕ.

$$\tau_x\widetilde{\phi} = \tau_x(\widetilde{\phi}) = \tau_x(y \mapsto \phi(-y)) = y \mapsto \Big(y - x \mapsto \phi(-(y-x))\Big) = y \mapsto \phi(x-y)$$

Lemma 19.1: Density Lemma

The following inclusions are toplinear and dense.

- If $(X, \mathcal{M}, \mu)$ is any measure space, $\Sigma_1 \subseteq L^p$ for usual p, and $\Sigma \subseteq L^\infty$ is dense.
- If X is LCH and μ is a Radon measure on X, $C_c(X \subseteq L^p(\mu)$ for usual p.
- Lusin's Theorem. If $(X, \mathcal{M}, \mu)$ is LCH and Radon, every $f \in \mathbb{B}_X$ can be uniformly approximated by $\phi \in C_c(X)$ with $\phi = f$ on A^c and $\mu(A) < +\varepsilon$
- Stone's Theorem. The complex exponentials, $(E_k)_{k\in\mathbb{Z}^n}$ where

$$E_k(x) = \exp(2\pi i\langle k, x\rangle) \quad \forall x \in \mathbb{T}^n$$

form a dense subset of $C(\mathbb{T}^n)$ (uniformly) and in $L^2(\mathbb{T}^n)$.

In the theory of distributions, we have the following

- $\mathcal{D} \subseteq \mathcal{S} \subseteq L^p$ for usual p on $\mathbb{R}^n$,

- $\mathcal{D} \subseteq \mathcal{S} \subseteq C_0$, and
- $\mathcal{D} \subseteq \mathcal{E}$.

More facts about L^p spaces

- If $(X, \mathcal{M}, \mu) = (\mathbb{R}^n, \mathbb{B}, \mu)$, translation is continuous in L^p for usual p. That is,

$$\lim_{y \to 0} \|\tau_y f - f\|_p = 0$$

Chapter 1: Manifolds

Introduction

We serve to give the reader the shortest introduction to manifold theory. This and the subsequent two chapters are loosely based on [3, 4], and the symbols E, F will always denote Banach spaces, and all Banach spaces are assumed to be over $\mathbb{R}$. We sometimes say E (resp. F) is a space for brevity, and

- $\mathcal{L}(E, F)$ = linear maps between E and F,
- $L(E, F)$ = toplinear (continuous and linear) maps between E and F,
- $\mathrm{Topliso}(E, F)$ = toplinear isomorphisms between E and F,
- $\mathrm{Laut}(E)$ = toplinear automorphisms on E, which form a strongly open subset of $L(E, E)$.

We will be working in the category of C^p Banach spaces — where $p \geq 0$. The morphisms in the category of $\mathrm{Ban}_{\mathbb{R}}$ are called C^p morphisms, which are p-times continuously differentiable functions.

Definition 1.1: Morphisms between open subsets of Banach spaces

Let E and F be Banach spaces, and $U \subseteq E$, $V \subseteq F$ be open subsets. A mapping $f : E \to F$ is of class C^p if $f \in C(E, F)$ and eq. (23) holds.

$$D^{(i)} f : E \to L^i(E, F) \quad \text{exists and is continuous for} \quad i = \underline{p} \tag{23}$$

$C^p(E, F)$ denotes the vector space of C^p mappings between E and F. Sometimes, we restrict our attention to *open subsets* of E and F, in this case: $f \in C^p(U, V)$ if $f \in C(U, V)$ and eq. (24) holds.

$$D^{(i)} f : U \to L^i(E, F) \quad \text{exists and is continuous for} \quad i = \underline{p} \tag{24}$$

We sometimes write C^p for $C^p(E, F)$ when it is clear. A C^p *isomorphism* is a bijective C^p morphism whose inverse is also a morphism.

Remark 1.1: Implicit assumption

In eq. (24) we assumed that $f(U) \subseteq V$. This is a non-trivial part of the definition of C^p morphisms between E and F, we will come back to this in def. 3.1.

Let f_1 and f_2 be mappings, and X a non-empty set.

- We say they are *composable* if either one of $f_2 \circ f_1$ or $f_1 \circ f_2$ makes sense.
- We also write $f_2 f_1$ to refer to $f_2 \circ f_1$ if there is no ambiguity.
- If $U \subseteq X$ and $V \subseteq Y$, and $f : U \to V$ is a bijection — meaning $f(U) = V$ and f is injective, we say f is a bijection between U and V.
- With regards to inverse image notation, we allow ourselves to write

  $$f_2^{-1} \circ f_1^{-1} \quad \text{is the same as} \quad f_2^{-1} f_1^{-1}$$

  and inversion is never left associative.

  $$f_2 f_1^{-1} = f_2 \circ f_1^{-1} \neq (f_2 \circ f_1)^{-1}$$

Composable C^p mappings are functors in the category of open subsets between Banach spaces. Few basic facts about C^p morphisms:

- If f is a toplinear mapping between E and F, then $f \in C^p(E, F)$ for all $p \geq 0$.
- If f is a bijective toplinear mapping, then it is a C^p isomorphism for all $p \geq 0$.
- However, a bijective C^p morphism need not be a C^p isomorphism.

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.

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 $U \subseteq X$, $\varphi(U) = \hat{U}$ is an *open* subset of E, and φ is a bijection onto $\hat{U}$.

Definition 2.2: Compatibility

Let (U, φ) and (V, ψ) be charts on X modelled on E, they are called C^p compatible (for $p \geq 0$) if $U \cap V = \varnothing$, or both of the following hold

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

Definition 2.3: Atlas

Let X be a non-empty set and $p \geq 0$. 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.

We will assume hereinafter that atlases are of class C^p for $p \geq 0$. Let X be a non-empty set, equipped with an atlas $\{(U_\alpha, \varphi_\alpha)\}$ modelled on a space E. Suppose α, and β both index the atlas.

- We write $\hat{U}_\alpha$ to refer to $\varphi_\alpha(U_\alpha)$, and
- $\hat{p} = \varphi_\alpha(p)$ for $p \in U_\alpha$ when it is clear which chart we are using.
- $U_{\alpha\beta} = U_\alpha \cap U_\beta$, and if $U_{\alpha\beta} \neq \varnothing$: the *transition map from α to β* is defined in eq. (25).

$$\varphi_{\alpha\beta} \triangleq \varphi_\beta|_{U_{\alpha\beta}} \circ \left(\varphi_\alpha|_{U_{\alpha\beta}}\right)^{-1} : \varphi_\alpha(U_{\alpha\beta}) \to \varphi_\beta(U_{\alpha\beta}) \tag{25}$$

- We often suppress the restrictions of the two charts in the composition, and eq. (25) reads

$$\varphi_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1} = \varphi_\beta \varphi_\alpha^{-1} \tag{26}$$

Remark 2.1: Omissions of C^p

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. We also say C^p for C^p where $p \geq 0$.

Given that compatibility is an equivalence relation on the set of all charts on X that are modelled on E, it should not be surprising it descends into an equivalence relation among atlases. This is condensed in note 2.1.

Note 2.1: Descent of an equivalence relation

Let Ω be a non-empty set with an associated equivalence relation $\sim$. Suppose $A_i \subseteq \Omega$ is also a subset of the equivalence class $[A_i]$ where $i = \underline{2}$. We say the $A_1 \sim A_2$ if any of the following equivalent statements hold.

1. For every $(x, y) \in A_1 \times A_2$, we have $x \sim y$.
2. There exists $x \in A_i$, where $x \sim y$ for all $y \in A_{3-i}$.
3. $A_1 \cup A_2$ is a subset of an equivalence class over $\Omega / \sim$.
4. $A_j \subseteq [A_i]$ for $i, j = \underline{2}$.

It is not hard to see this defines an equivalence relation. And $[A_i]$ represents the largest superset of A_i that is contained within a single equivalence class.

Definition 2.4: Structure determined by an atlas

Let $\mathcal{A}$ be an atlas on X, the maximal atlas containing $\mathcal{A}$ is called the C^p *structure determined by* $\mathcal{A}$.

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 refer to E as the *model space* of X.

Proposition 2.1: E is a manifold

The identity id_E defines an atlas on E, which determines a C^p structure called the *standard structure of* E for $p \geq 0$. We call (E, id_E) the *standard chart* on E.

Proposition 2.2: Topology is unique on a manifold

Let X be a C^p manifold modelled on E, it induces 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 in the subspace topology.

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 $\varphi : U \to \hat{U}$ 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 constructing the above topology. It is well known of the existence of a unique coarsest topology on a chart domain U where all charts (V, φ) whose domains intersect U — when restricted onto U — are homeomorphisms onto their ranges. Stitching the weak topologies together, we obtain an ambient topology on X. ■

Remark 2.2: Not necessarily Hausdorff

The topology generated by prop. 2.2 is not necessarily Hausdorff, nor second countable. So a manifold 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.

Remark 2.3: Omission of model space

For any of the objects we have defined in this section, that depend upon a model space or a morphism class (i.e C^p), we will say 'X is a manifold', rather than X is a manifold of class C^p modelled over E when it is convenient to do so. If the model space E is infinite (resp. finite) dimensional, we say X is infinite (resp. finite) dimensional. And a reminder: C^p should always be interpreted with $p \geq 0$.

Proposition 2.3: Open subsets of manifolds

Let U be an open subset of a manifold X, then U is a manifold whose structure is determined by the atlas $\mathcal{A}$ in eq. (27).

$$\mathcal{A} = \Big\{ (V, \varphi) \text{ in the structure of } X, \text{ where } V \subseteq U \Big\} \tag{27}$$

Proof. The structure of X includes all possible restrictions to open sets; hence $\mathcal{A}$ in eq. (27) is an atlas, and a unique structure by def. 2.4. ■

Morphisms between manifolds

Definition 3.1: Morphisms between manifolds

A mapping $f : X \to Y$ between manifolds is a *morphism* (a C^p morphism to be precise) if for every $p \in X$, there exist charts $(U, \varphi) \in X$ and $(V, \psi) \in Y$ such that 1) the image $f(U)$ is contained in a the chart domain V, and 2)

$$f_{U,V} \triangleq \psi \circ f \circ \varphi^{-1} \in C^p(\hat{U}, \hat{V}) \quad \text{in the sense of def. 1.1.} \tag{28}$$

The map $f_{U,V}$ as defined in eq. (28) is called the *coordinate representation of f* with respect to the charts $(U, \varphi), (V, \psi)$.

Remark 3.1: Identifying X with its structure

If (U, φ) is a chart in the structure of X, we will simply say (U, φ) is in X.

Remark 3.2: Identifying charts with their domains

The scenario in eq. (28) occurs so often that we decide to simply write

$$f_{U,V} = \psi f \varphi^{-1} \tag{29}$$

to mean there exists charts (U, φ), $(V \psi)$ in the structure of X, Y with

$$f(U) \subseteq V \tag{30}$$

Consistent with the notation of putting hats on objects borrowed or pulled back from the model spaces, we write $\hat{f} = f_{U,V}$. Equation (31) gives an example of this.

$$\hat{f}(\hat{p}) = f_{U,V}(\hat{p}) = f_{U,V}(\varphi(p)) \tag{31}$$

for any morphism $f \in \mathrm{Mor}(X, Y)$, and charts that satisfy eq. (30). We refer to the map in eq. (31) as a *coordinate representation of f about p*, with the inference that $p \in (U, \varphi)$.

Definition 3.1 may leave one unsatisfied. Why do we require the image $f(U)$ be contained in another chart domain in Y? There are two reasons.

1. 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.

2. The definition of smoothness between open subsets of Banach spaces (see def. 1.1) is a purely local one. And let us recall: every chart domain U in a manifold X corresponds to an open subset $\hat{U} \subseteq E$ in the model space, and see remark 1.1 as well. Hence, **the necessity that the image $f(U)$ is contained in a single chart domain of Y is a relic of the original definition**. The astute reader will also see that the openness requirement of $\psi(U \cap V)$, and $\varphi(U \cap V)$ in def. 2.2 is completely natural as well, since C^p morphisms are defined between open subsets of Banach spaces.

Proposition 3.1: Properties of morphisms between manifolds

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 is proven if we show f is locally continuous. Using Equation (28), 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 homeomoprhisms, and unwinding the formula shows that $f|_U = \psi^{-1} f_{U,V} \varphi$, so that

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

To prove the second, let $X_{\underline{3}}$ be manifolds modelled over $E_{\underline{3}}$, 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 $(f_1)_{U_1,U_2}$ 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 3.3: 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 $f : X \to Y$ such that each point $p \in X$ admits a coordinate representation

$$f_{U,V} \in C^{\min(p,k)}(\hat{U}, \hat{V}) \tag{32}$$

If $\min(p, k) \geq 1$, then we define its differential as in def. 4.4 by treating both X and Y as $C^{\min(k,p)}$ manifolds.

Tangent spaces

In this section, all manifolds will be of class C^p for $p \geq 1$. 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.

Suppose U is an open subset of E and $f : U \to Y$ is C^p. The derivative $Df(x)$ is a linear map $E \to F$, not from U to F (U might not even be a vector space). This suggests the 'derivative' of a morphism $F : X \to Y$ 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 $f : X \to Y$ 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 $p_i = \varphi_i(p)$, $U_{12} = U_1 \cap U_2$ and

$$\varphi_{12} = \varphi_2 \varphi_1^{-1} : \varphi_1(U_{12}) \to \varphi_2(U_{12}) \tag{33}$$

(because the map in eq. (33) goes from the domain U_1 to U_2), a simple computation yields eq. (34).

$$\begin{aligned} Df_{U_1,V}(p_1)(v) &= D(\psi f \varphi_2^{-1} \varphi_2 \varphi_1^{-1})(p_1)(v) = Df_{U_2,V}(p_2)\Big(D\varphi_{12}(p_1)(v)\Big) \\ &= Df_{U_2,V}(p_2) \circ D\varphi_{12}(p_1) \cdot (v) \end{aligned} \tag{34}$$

where $\cdot(v)$ denotes the evaluation at $v \in E$, and is assumed to be left associative over composition. The computation in eq. (34) suggests that interpreting the derivative by pre-conjugation is dependent on the chart being used to interpret the derivative. In fact, $D\varphi_{12}(p_1)$ 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. (34) 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. (35)

$$T_{(U,\varphi,p)}X = \bigcup_{v \in E} \{(U, \varphi, p, v)\} \cong E \tag{35}$$

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. (34) there exists a natural (toplinear) isomorphism between the concrete tangent spaces, namely

$$(U_1, \varphi_1, p, v_1) \sim (U_2, \varphi_2, p, v_2) \quad \text{iff} \quad v_2 = D\varphi_{12}(p_1)(v_1) \tag{36}$$

where $p_i = \varphi_i(p)$. The right member of eq. (36) is the derivative of a transition map — which is a toplinear automorphism on E. Hence $D\varphi_{12}(p_1)$ 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. (36).

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, and fix $p \in X$. We define a linear map, called the *differential of f at p* shown in eq. (37).

$$df(p) : T_pX \to T_{f(p)}Y \tag{37}$$

Whose action on tangent vectors is characterized by

- if (U, φ) and (V, ψ) are any pair of charts that satisfy the morphism condition in eq. (28) about p, and suppose
- $v \in T_pM$ is represented by $\Big(U, \varphi, p, \hat{v}\Big)$
- then $df(p)(v) \in T_{f(p)}Y$ is represented by $\Big(V, \psi, f(p), Df_{U,V}(\hat{p})(\hat{v})\Big)$

Alternatively, the diagram shown in fig. 1 commutes. We also write $df_p = df(p)$.

$$\begin{tikzcd} T_pX \arrow[r] \arrow[d, "df(p)"] & T_{(U,\varphi,p)}X \arrow[d, "Df_{U,V}(\hat{p})"] \\ T_{f(p)}Y \arrow[r] & T_{(V,\psi,f(p))}Y \end{tikzcd}$$

Figure 1: Differential of a morphism

Velocities

In the previous section, we motivated the definition of T_pX 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.

Definition 5.1: Curve

Let $J_\varepsilon = (-\varepsilon, +\varepsilon)$ be an open interval in $\mathbb{R}$ containing the origin. Proposition 2.3 tells us J_ε is a manifold. A morphism $\gamma : J_\varepsilon \to X$ is called a *curve in* X, and $\gamma(0)$ is called the *starting point of* γ.

Remark 5.1: Omission of chart in concrete representation

If p is a point on a manifold X, and $v \in T_pX$ is represented by $(U, \varphi, p, \hat{v})$, we write

$$(U, \hat{v}) = (\hat{p}, \hat{v}) = \hat{v} = (U, \varphi, p, \hat{v}) \tag{38}$$

Remark 5.2: Standard representation of tangent vectors

If X is an open subset of E, and $p \in X$, we identify a tangent vector $v \in T_pX$ by its *standard representation.* Instead of using a $\hat{v}$, we use $\overline{v}$.

$$(X, \mathrm{id}_X, p, \overline{v}) = (X, \overline{v}) = (X, \hat{v}) \quad \text{is a representation of } v \in T_pX \tag{39}$$

Definition 5.2: Velocity of a curve

Let γ be a curve in X and $t \in J_\varepsilon$. We denote the *velocity* of a curve γ at $t = t_0$ by $\gamma'(t_0)$; which is defined in eq. (40).

$$\gamma'(t_0) = \left[D\gamma_{J_\varepsilon,V}(t_0)(\overline{1})\right] \tag{40}$$

where $(J_\varepsilon, \mathrm{id}_{J_\varepsilon}, t_0, \overline{1})$ is a concrete tangent vector within $T_{t_0}J_\varepsilon$.

Equation (40) might seem arbitrary at first, but we must emphasize that this is the most natural interpretation of a velocity, encodes much of the geometric information of a tangent vector. We will revisit this topic when we discuss exterior differentiation.

Proposition 5.1: Tangent vectors are velocities

Let p be a point on a manifold X. For every tangent vector $v \in T_pX$, 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. (41), lies in $\hat{U}$

$$\hat{\gamma} : J_\varepsilon \to \hat{U} \quad \gamma(t) = \int_0^t \hat{v} dt \tag{41}$$

$\hat{\gamma}$ 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

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

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

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 $\theta : E_1 \times E_2 \to E$ given by

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

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.

Every finite dimensional or finite codimensional linear subspace of E splits. And if E itself is finite dimensional, then every linear subspace of E splits. An alternative definition of def. 6.2 is as follows: an map $\lambda \in L(E, F)$ splits iff there exists a toplinear isomorphism $\theta : F \to F_1 \times F_2$ such that λ composed with α induces a toplinear isomorphism from E onto $F_1 \times 0$ — which we identify with F_1.

If E and F are finite dimensional (so $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$ respectively), def. 6.2 refers to the familiar matrix canonical form in eq. (42), and defs. 6.3 and 6.4 can be seen as infinite-dimensional analogues of eq. (42).

$$A_{\text{injective}} = \begin{bmatrix} \text{id}_{m\times m} \\ 0_{n-m\times m} \end{bmatrix} \qquad A_{\text{surjective}} = \begin{bmatrix} \text{id}_{n\times n} & 0_{n\times m-n} \end{bmatrix} \tag{42}$$

Definition 6.3: Immersion

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

$$Df_{U,V}(\hat{p}) \quad \text{is injective and splits.} \tag{43}$$

The morphism f is called an immersion if eq. (43) holds at every p.

Definition 6.4: Submersion

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

$$Df_{U,V}(\hat{p}) \quad \text{is surjective and its kernel splits.} \tag{44}$$

The morphism f is called a submersion if eq. (44) holds at every p.

Definition 6.5: Embedding

A morphism $f \in \operatorname{Mor}(X, Y)$ is an *embedding* if it is an immersion and 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 $\hat{U} \overset{\circ}{\subseteq} E$. 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 $\varphi_{\alpha\beta} = \varphi_\beta \varphi_\alpha^{-1}$ are **morphisms between open subsets of E**.

If $(U, \varphi) \in X$ is a chart whose domain intersects S, the question then becomes: Is it possible to modify (U, φ) so that it becomes a chart modelled on E_1? If we restrict φ onto $U \cap S$, its range is still an open subset of E. We can assume $\varphi(U \cap S) \subseteq E$ is constant on the linear complement of E_1, that way $\varphi|_{U\cap S}$ will be a bijection.

The range of the restricted chart is still a subset of E, and not E_1. An easy fix to this would be to require **E_1 to split in E** (and shrinking U using a basis argument). Let θ be a toplinear isomorphism between E and $E_1 \times E_2$, and we obtain eq. (45).

$$\theta\varphi(S \cap U) = \hat{U}_1 \times a_2 \quad \text{where} \quad \hat{U}_1 \overset{\circ}{\subseteq} E_1 \text{ and } a_2 \in E_2. \tag{45}$$

Identifying $\hat{U}$ with $\theta(\hat{U})$, and requiring $U_1 \times a_2$ to be in $\theta(\hat{U})$, 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

$$\varphi : U \to \hat{U} \cong \hat{U}_1 \times \hat{U}_2, \quad \text{where} \quad \hat{U}_i \mathring{\subseteq} E_i \quad \text{for} \quad i = \underline{2}, \tag{46}$$

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

$$\varphi(U \cap S) = \hat{U}_1 \times a_2. \tag{47}$$

We call a chart satisfying eqs. (46) and (47) a *slice chart* of S; to simplify what follows, we write $\varphi^i = \mathrm{proj}_i \varphi$ for $i = \underline{2}$ 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 a 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. (48) forms a C^p atlas modelled E_1, which we will prove in prop. 7.1

$$\mathcal{A} = \left\{ (U^s, \varphi^s) = (U^s, \varphi^1),\ (U, \varphi) \text{ is a slice chart of } S \right\}. \tag{48}$$

Proposition 7.1: Structure of a submanifold

If S is a submanifold of X, eq. (48) defines a C^p atlas over the space E_1. The manifold S has a topology that coincides with the subspace topology. Furthermore, the inclusion map $\iota_S : S \to X$ is a morphism and an embedding.

Proof. Each of the charts in eq. (48) is in bijection with an open subset of E_1. Let $(U^s_\alpha, \varphi^s_\alpha)$ and $(U^s_\beta, \varphi^s_\beta)$ be overlapping charts in $\mathcal{A}$. Using θ as our toplinear isomorphism from E onto $E_1 \times E_2$ as usual.

- By eq. (46), $(U^s_\alpha, \varphi^s_\alpha)$ is induced by a chart $(U_\alpha, \varphi_\alpha) \in X$,

$$\varphi_\alpha : U_\alpha \to \hat{U}_\alpha \mathring{\subseteq} E \quad \text{which splits into} \quad \theta(\hat{U}_\alpha) = \hat{U}^s_\alpha \times \hat{U}_{2,\alpha}$$

  such that $\hat{U}^s_\alpha \mathring{\subseteq} E_1$ and $\hat{U}_{2,\alpha} \mathring{\subseteq} E_2$. Similarly for β as well.

- There exists elements $a_2 \in \hat{U}_{2,\alpha}$, (resp. $b_2 \in \hat{U}_{2,\beta}$) where

$$\theta\varphi_\alpha(U^s_\alpha) = \hat{U}^s_\alpha \times a_2 \quad \text{resp. } \beta.$$

Note 7.1

Let us define $U^s_{\alpha\beta} = U^s_\alpha \cap U^s_\beta$, we will show lem. 7.1.

Lemma 7.1

Both $\varphi^s_\alpha(U^s_{\alpha\beta})$ and $\varphi^s_\beta(U^s_{\alpha\beta})$ are open subsets of E_1.

Proof of lem. 7.1. We can factor $U^s_{\alpha\beta} = (U^s \cap U_\alpha) \cap U_{\alpha\beta}$, and because φ_α is a bijection, we have

$\varphi^s_\alpha(U^s_{\alpha\beta}) = \mathrm{proj}_1\, \theta\Big(\varphi_\alpha(U^s \cap U_\alpha) \cap \varphi_\alpha(U_{\alpha\beta})\Big).$

θ and proj_1 are both open maps, and because $W \triangleq \varphi_\alpha(U_{\alpha\beta})$ is open in E: $\theta(\varphi_\alpha(U^s \cap U_\alpha) \cap W)$ splits into a subset of $\hat{U}^s_\alpha \times a_2$,

$$\mathrm{proj}_1\, \theta(\varphi_\alpha(U^s \cap U_\alpha) \cap W) = \mathrm{proj}_1(\text{Open subset of } E_1 \times a_2)$$

which is open in E_1. ■

The diagram in fig. 2 provides a summary.

$$\begin{tikzcd}
U^s_{\alpha\beta} \arrow[r, "\varphi_\alpha"] & \varphi_\alpha(U^s_{\alpha\beta}) \arrow[r, "\theta"] \arrow[d, "\varphi_{\alpha\beta}"] & \varphi_\alpha(U^s_{\alpha\beta})_1 \times a_2 \arrow[r, "\mathrm{proj}_1"] \arrow[d, "\theta\varphi_{\alpha\beta}\theta^{-1}"] & \varphi^s_\alpha(U^s_{\alpha\beta}) \\
U^s_{\alpha\beta} \arrow[r, "\varphi_\beta"] & \varphi_\beta(U^s_{\alpha\beta}) \arrow[r, "\theta"] & \varphi_\beta(U^s_{\alpha\beta})_1 \times b_2 \arrow[r, "\mathrm{proj}_1"] & \varphi^s_\beta(U^s_{\alpha\beta})
\end{tikzcd}$$

Figure 2: Overlap of slice charts

Identifying a_2 (resp. b_2) with the constant function $(p \mapsto a_2)$ for $p \in U^s_\alpha$, we get eq. (49).

$$\varphi^s_\alpha \times a_2 = \theta \circ \varphi_\alpha \quad \text{resp. } \beta \tag{49}$$

Suppressing the restrictions onto domains, the transition map is given by the composition of maps in eq. (50).

$$\varphi^s_\beta \circ (\varphi^s_\alpha)^{-1} = \mathrm{proj}_1\, \theta\varphi_\beta\varphi_\alpha^{-1}\theta^{-1}\,\mathrm{proj}_1^{-1} : \varphi^s_\alpha(U^s_{\alpha\beta}) \to \varphi^s_\beta(U^s_{\alpha\beta}) \tag{50}$$

which is clearly a bijection. It suffices to show eq. (50) is a morphism between open subsets of E_1. Let $a_2 : \varphi^s_\alpha(U^s_{\alpha\beta}) \to \hat{U}_{2,\alpha}$, which is the constant function a_2 and hence a morphism.

The product $(\mathrm{id}_{\varphi^s_\alpha(U^s_{\alpha\beta})} \times a_2) = \mathrm{proj}_1^{-1}$ is a morphism into $\varphi^s_\alpha(U^s_{\alpha\beta}) \times \hat{U}_{2,\alpha}$. The inverse of θ is an open morphism, and the terms $\varphi_\beta\varphi_\alpha^{-1}$ combine into the transition map $\varphi_{\alpha\beta}$ in X (up to a restriction on an open set). Equation (50) then reads

$$\varphi^s_\beta \circ (\varphi^s_\alpha)^{-1} = \mathrm{proj}_1\, \theta\varphi_{\alpha\beta}\theta^{-1}(\mathrm{id}_{\varphi^s_\alpha(U^s_{\alpha\beta})} \times a_2) \tag{51}$$

which is a morphism between open subsets. Reversing the roles of α, β shows that eq. (50) is an isomorphism. Therefore the collection of charts in eq. (48) forms an atlas of S.

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

Observing that $\iota_S(U^s) = \iota_S(U \cap S)$ lies within (U, φ), the morphism criteria in eq. (28) is satisfied. Computing the coordinate representation of ι_S, we obtain eq. (52).

$$(\iota_S)_{U^s,U} = \varphi\iota_S(\varphi^s)^{-1} = \mathrm{id}_{\hat{U}_1} \times a_2 \tag{52}$$

Equation (52) 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. (52); ι_S^{-1} is a morphism and therefore continuous. The manifold topology of S coincides with its subspace topology.

At last, the inclusion map ι_S has coordinate representation eq. (52). Computing its ordinary derivative we obtain eq. (53).

$$D(\iota_S)_{U^s,U}(\hat{p}) : T_{(U^s,\varphi^s,p)} \longrightarrow T_{(U,\varphi,p)} \quad \text{and} \quad D(\iota_S)_{U^s,U}(\hat{p}) = \mathrm{id}_{E_1} \times 0 \tag{53}$$

which is a toplinear morphism between concrete tangent spaces and has a simple representation of 'adding zeroes' (see def. 6.2) at the end of a vector $\hat{v} \in E_1$ — which is to say: **the differential of ι_S is injective and splits in E**. Therefore ι_S is an embedding. ■

Remark 7.1: Pairs of slice charts

Proposition 7.1 shows every point $p \in S$ is in the domain of a slice chart in S, and the domain of the chart in X that induces the slice chart — whose inclusion map satisfies eqs. (52) and (53). 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.

Definition 7.2: Exterior tangent space of S

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

$$T_p^{\mathrm{ext}} S = d\iota_S(p)(T_pS), \tag{54}$$

which is a toplinear subspace of T_pX.

Chapter 2: Vector Bundles

Vector Bundles

Let X be a class C^p manifold modelled on a space E, and F another Banach space. Our goal in this section is to construct the vector bundle of a manifold, which has the following desirable properties.

- The vector bundle W embeds X into itself as a submanifold.
- At each point $p \in X$, we attach a F space structure exclusive to each p like the tangent space T_pX.
- W locally isomorphic to the product space $U \times F$, where $U \overset{\circ}{\subseteq} X$, and
- a subset of the morphisms $A : X \to W$ locally resemble morphisms $U \to U \times F$.

Definition 1.1: Coproduct of fibers

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. (55) is called a *coproduct of F-fibers modelled over X*.

$$W = \coprod_{p \in X} W_p \quad \text{comes with} \quad \pi : W \to X, \quad \pi^{-1}(p) = W_p \tag{55}$$

where π is a surjection onto X and is called the *canonical projection*.

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 $\widetilde{U}$ instead of $\pi^{-1}(U)$ for brevity, and $\widetilde{p}$ in place of $\pi^{-1}(p)$. The next few definitions should feel familiar.

Definition 1.2: Local trivialisation

Let W be as in eq. (55). A *local trivialisation* of W is a tuple $(\widetilde{U}, \Phi)$, such that the diagram in fig. 3 commutes, and

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

Definition 1.3: Compatibility between trivialisations

Let $(\widetilde{U}, \Phi)$ and $(\widetilde{V}, \Psi)$ be local trivialisations of W, they are called C^k-compatible if $U \cap V = \varnothing$, 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 — $(\Psi \circ \Phi^{-1})|_{\widetilde{p}}$ is a toplinear isomorphism, and
- the map $\theta : U \cap V \to L(F, F)$ as defined by eq. (56), is a C^k morphism into the Banach space $L(F, F)$.

$$\theta(p) = (\Psi \circ \Phi^{-1})|_{\widetilde{p}} \tag{56}$$

(equivalently, we can require θ be a C^k morphism into the open submanifold $\mathrm{Laut}(F)$).

Note: we assume that $0 \leq k \leq p$.

Definition 1.4: 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 $\{(\widetilde{U}_\alpha, \Phi_\alpha)\}$ where $\{U_\alpha\}$ is an open cover of X.

Definition 1.5: 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 over X* is a coproduct of F-fibers modelled over X equipped with a **maximal C^k trivialisation covering**.

Remark 1.1: 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 1.2: Omissions for vector bundles

We say W is a *bundle over* X when it is unambiguous to do so.

$$\begin{tikzcd} \widetilde{U} \arrow[r, "\Phi"] \arrow[d, "\pi"'] & U \times F \arrow[dl, "\mathrm{proj}_1"] \\ U & \end{tikzcd}$$

Figure 3: Local Trivialisation

The above definitions calls for some commentary, our end goal is to make an arbitrary rank C^k vector bundle W a C^k manifold. Open sets will still be our primary topological data. To ensure that W is as similar to X as possible, the eventual manifold structure we will put on W will **embed the structure of X into W**. We are repeating (essentially) the same argument as in the submanifold case but with the roles of X and the submanifold S reversed.

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

Definition 1.2
- U is required to be open because W inherits part of the topology, and hence the charts in E whose domain is a subset of U,
- The second requirement implies **each Φ is in bijection with $\Phi(\widetilde{U}) = U \times F$, which is open in $E \times F$**, which will allow us to construct bijections with open subsets of the model space $E \times F$. Furthermore, eq. (57) below holds for an arbitrary $V \subseteq X$.

$$\Phi|_{\pi^{-1}(U \cap V)} \quad \text{is a bijection onto } U \cap V \times F \tag{57}$$

Definition 1.3

- 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.
- The requirement that the mapping eq. (56) is a morphism is because we wish to **have control over the smoothness of morphisms $X \to W$.**

Suppose W is an F-vector bundle over X with the trivialisation covering $\{(\widetilde{U}^\alpha, \Phi_\alpha)\}$. For each α, we can cover U^α using chart domains $(U^\alpha_\beta, \varphi^\alpha_\beta)$ in X — without loss of generality, we can assume $U^\alpha_\beta \subseteq U^\alpha$ 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. (58)

$$\widetilde{\varphi}^\alpha_\beta : \widetilde{U}^\alpha_\beta \longrightarrow \hat{U}^\alpha_\beta \times F \quad \text{defined by} \quad \widetilde{\varphi}^\alpha_\beta = \left(\varphi^\alpha_\beta \times \mathrm{id}_F\right) \circ \Phi_\alpha|_{\widetilde{U}^\alpha_\beta} \tag{58}$$

Remark 1.3: Hats and wiggles

Here, $\widetilde{U}^\alpha_\beta$ should be interpreted as the inverse image of the open set U^α_β through π. Similarly, $\hat{U}^\alpha_\beta$ is the image of U^α_β through φ^α_β.

The collection of charts in eq. (59) cover W with their chart domains, and each chart is in bijection with an open subset of $E \times F$.

$$\mathcal{A} = \left\{(\widetilde{U}^\alpha_\beta,\ \widetilde{\varphi}^\alpha_\beta),\ (\widetilde{U}^\alpha, \Phi_\alpha) \text{ is in the trivialisation covering of } W.\right\} \tag{59}$$

Proposition 1.1: 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* $\pi : W \to X$ is a morphism and a submersion.
2. X is C^k isomorphic to a submanifold of W

Proof. Suppose we are given two charts in eq. (59), $(\widetilde{U}^{\alpha_1}_{\beta_1})$, and $(\widetilde{U}^{\alpha_2}_{\beta_2}, \widetilde{\varphi}^{\alpha_2}_{\beta_2})$. We first prove that $\widetilde{\varphi}^{\alpha_1}_{\beta_1}(\widetilde{U}^{\alpha_1}_{\beta_1} \cap \widetilde{U}^{\alpha_2}_{\beta_2})$ is open in $E \times F$.

$$\begin{aligned}\widetilde{\varphi}^{\alpha_1}_{\beta_1}(\widetilde{U}^{\alpha_1}_{\beta_1} \cap \widetilde{U}^{\alpha_2}_{\beta_2}) &= \left[(\varphi^{\alpha_1}_{\beta_1} \times \mathrm{id}_F) \circ \Phi_{\alpha_1}\right](\widetilde{U}^{\alpha_1}_{\beta_1} \cap \widetilde{U}^{\alpha_2}_{\beta_2}) = \left[(\varphi^{\alpha_1}_{\beta_1} \times \mathrm{id}_F) \circ \Phi_{\alpha_1}\right](\pi^{-1}(U^{\alpha_1}_{\beta_1} \cap U^{\alpha_2}_{\beta_2})) \\ &= (\varphi^{\alpha_1}_{\beta_1} \times \mathrm{id}_F)\Big((U^{\alpha_1}_{\beta_1} \cap U^{\alpha_2}_{\beta_2}) \times F\Big) && \text{by eq. (57).}\end{aligned}$$

Suppressing restrictions and computing the chart transistions in eq. (60),

$$\widetilde{\varphi}^{\alpha_2}_{\beta_2}\left(\widetilde{\varphi}^{\alpha_1}_{\beta_1}\right)^{-1} = (\varphi^{\alpha_2}_{\beta_2} \times \mathrm{id}_F) \circ \Phi_{\alpha_2}\Phi^{-1}_{\alpha_1} \circ \left((\varphi^{\alpha_1}_{\beta_1})^{-1} \times \mathrm{id}_F\right). \tag{60}$$

which is clearly a bijection. And it is not hard to see that eq. (60) can be factored into

$$\widetilde{\varphi}^{\alpha_2}_{\beta_2}\left(\widetilde{\varphi}^{\alpha_1}_{\beta_1}\right)^{-1}(x,v) = \left(\varphi^{\alpha_1\alpha_2}_{\beta_1\beta_2}(x),\ \left[\theta\circ(\varphi^{\alpha_1}_{\beta_1})^{-1}\right](x)(v)\right). \tag{61}$$

for any $x \in \varphi^{\alpha_1}_{\beta_1}(U^{\alpha_1\alpha_2}_{\beta_1\beta_2})$ and $v \in F$. **From eq. (61), it should now be clear why we demand $k \leq p$** . The mapping in the second coordinate within eq. (61) can be reduced to a composition with the evaluation map $\mathbf{E} : \mathrm{Laut}(F) \times F \to F$.

$$\left[\theta\circ(\varphi^{\alpha_1}_{\beta_1})^{-1}\right](x)(v) = \mathbf{E}\circ\left([\theta\circ(\varphi^{\alpha_1}_{\beta_1})^{-1}]\times\mathrm{id}_F\right). \tag{62}$$

Since $\mathbf{E}$ is continuous and bilinear, eq. (62) and hence eq. (60) describes a C^k mapping between open subsets of Banach spaces. It is a morphism, and reversing the roles of the two charts proves its inverse is again a morphism.

To prove π is a submersion, recall W is the set-theoretic disjoint union of F-fibers. Every element in W can be represented by $(x,v) \in X \times F$. **We will identify elements of W as elements in $X \times F$. However, this is not a manifold isomorphism**.

Fix $(x,v) \in W$, it is in the domain of some chart $(\widetilde{U}^\alpha_\beta, \widetilde{\varphi}^\alpha_\beta)$. The π-image of the chart domain is $\pi\pi^{-1}(U^\alpha_\beta) = U^\alpha_\beta$ because π is surjective. Using eq. (58) and the diagram found in fig. 3, the coordinate representation of π becomes

$$\begin{aligned}
\pi_{(\widetilde{U}^\alpha_\beta, U^\alpha_\beta)} &= \varphi^\alpha_\beta\circ\pi\circ\Phi^{-1}_\alpha\circ\left((\varphi^\alpha_\beta)^{-1}\times\mathrm{id}_F\right)\\
&= \varphi^\alpha_\beta\circ\mathrm{proj}_1\circ\left((\varphi^\alpha_\beta)^{-1}\times\mathrm{id}_F\right)\\
&= \mathrm{proj}_1(\mathrm{id}_{\hat{U}^\alpha_\beta}\times\mathrm{id}_F)
\end{aligned} \tag{63}$$

We can differentiate both sides of eq. (63) and if we write $\hat{U} = \hat{U}^\alpha_\beta$, we obtain eq. (64).

$$D\,\mathrm{proj}_1(\mathrm{id}_{\hat{U}}\times\mathrm{id}_F)(x,v) = \mathrm{proj}_1 \in L(E\times F; E) \quad \forall x\in\hat{U},\ v\in F \tag{64}$$

which means π submersion.

Finally, the subset $X \times 0 \subseteq W$ is easily shown to be a submanifold of W, and is isomorphic to X by dropping the F coordinate and retracing the argument we made in constructing the structure of W. ■

Remark 1.4: Pair of bundle charts

If X is a manifold and W a vector bundle over X, the charts realizing the representations of π in eqs. (63) and (64) are called *bundle charts*.

Definition 1.6: Section of a vector bundle

Let W be a bundle over a manifold X. A *section* of W is a morphism $\sigma \in \mathrm{Mor}(X, W)$ such that the diagram in fig. 4a commutes, which is synonymous with $\pi\sigma = \mathrm{id}_X$. A *local section* of W is a morphism $\sigma : U \to W$ where $U \overset{\circ}{\subseteq} X$ is viewed as a submanifold and $\pi\sigma = \mathrm{id}_U$. We denote the sections of W by

$\Gamma(W)$.

$$\Gamma(W) = \left\{\sigma : X \to W,\ \sigma \text{ is a section of } W.\right\} \tag{65}$$

The *zero section* of W is the section $\sigma(p) = 0 \in W_p$ for every $p \in X$. If σ is a section of W, $\operatorname{supp}(\sigma)$ refers to the *support* of σ, and is defined in eq. (66).

$$\operatorname{supp}(\sigma) = \overline{\{p \in X,\ \sigma(p) \neq 0\}} \tag{66}$$

Remark 1.5: Bundle coordinates

Let X and W be as in def. 1.6, and suppose σ is a section on W. Using a pair of bundle charts, $(U) \in X$ and $(\widetilde{U}) \in W$, we define the *bundle coordinates of* σ

$$\sigma_{U,\widetilde{U}} = \widetilde{\varphi} \circ \sigma \circ \varphi^{-1} \tag{67}$$

expanding the induced chart on W within eq. (67) reads

$$\sigma_{U,\widetilde{U}} = (\varphi \times \mathrm{id}_F) \circ \Phi \circ \sigma \circ \varphi^{-1} \tag{68}$$

Refer to the diagram in fig. 4b. We will always use bundle charts when discussing the coordinate representation of a section, and we write

$$\sigma_U = \sigma_{U,\widetilde{U}} = \hat{\sigma}$$

Sections are precisely the morphisms into W whose coordinate representation resembles that of a graph: $\hat{\sigma} : \hat{U} \to \hat{U} \times F$ and because of this: we identify $\hat{\sigma}(\hat{p}) = (\hat{p}, v)$ with $v \in F$.

$$\begin{tikzcd}
W \arrow[d, "\pi"] \\
X \arrow[u, bend left, "\sigma"]
\end{tikzcd}$$

(a) Section of a bundle

$$\begin{tikzcd}
\widetilde{U} \arrow[r, "\Phi"] \arrow[d, "\pi"] & U \times F \arrow[dl, "\mathrm{proj}_1"] \arrow[d, "\widetilde{\varphi}"] \\
U \arrow[u, bend left, "\sigma"] \arrow[d, "\varphi"] & \hat{U} \times F \\
\hat{U} \arrow[ur, "\sigma_{U,\widetilde{U}}"] &
\end{tikzcd}$$

(b) Local coordinates of a bundle section

Figure 4: Diagrams for bundle section and its local representation

Functorality

Let X and E_i be Banach spaces for $i = \underline{2}$. We discussed the difference in the role that a toplinear mapping $f \in L(E_1, E_2)$ plays in pushing points from $E_1 \to E_2$, and the role it plays from pushing *incoming maps* with source X from $L(X, E_1) \to L(X, E_2)$ by composing the incoming map with itself.

$$f : E_1 \to E_2 \quad \text{and} \quad f_c : L(X, E_1) \to L(X, E_2) \tag{69}$$

where the c within f_c stands for composition. This is summarized in fig. 5

$$\begin{tikzcd}
X \arrow[r, "\lambda"] \arrow[dr, "\lambda_c(f)"'] & E_2 \arrow[d, "f"] \\
 & E_2
\end{tikzcd}$$

Figure 5: Functorality through post-composition

The two maps f and f_c often thought of as the same, and we can identify f with two separate actions. Here, it is not so obvious why we need the concept of functorality, a closer look at the **different roles that the same mathematical object can play** will surely motivate the above discussion.

Let E_i be Banach spaces, $f \in C^p(E_1, E_2)$ and $\lambda \in L(E_2, E_3)$.

- The derivative of f is a continuous map $Df : E_1 \to L(E_1, E_2)$.
- The derivative of λ (now 'identified with $\lambda \in C^p(E_2, E_3)$), is continuous constant map from E_2 to $L(E_2, E_3)$

$$D(\lambda \circ f) = \lambda \circ Df(x) \tag{70}$$

but what if we have a multi-linear map whose destination is E_1, and what about symmetric/alternating multi-linear maps, continuous maps, C^p morphisms? Should we let f take on all of those roles as well? Should we identify f with its adjoint as well? This is the first of the many problems.

The problem becomes even clearer when we look at maps between F-fibers. Fix a manifold X and F_i-bundles W^i over X for $i = \underline{2}$. Suppose $A : X \to W^1$ is a section on W^1, and $\lambda \in L(F_1, F_2)$.

At each point $p \in X$, our linear map λ can be identified with the linear map that acts between the fibers.

$$\lambda_p : W_p^1 \to W_p^2$$

which is toplinear hence a morphism. Our main problem is concerned with the conditions under which the composition λA — as defined in eq. (71) — is a morphism.

$$\lambda A : X \to W^2 \quad \text{and} \quad (\lambda A)(p) = \lambda_p(A_p) \in W_p^2 \tag{71}$$

Under what conditions can a morphism take on additional roles? The mapping λ_p on each tangent space is a C^p morphism in the manifold sense, and the morphisms that preserve the C^p smoothness of sections are called VB morphisms. Which we will define after some more category theory.

For the remainder of this section, let C_1 and C_2 be categories. We denote the objects of C_1 by E_i, and the objects of C_2 by F_i.

Definition 2.1: Functor

A correspondence θ between C_1 and C_2 is called a *functor* — which we denote by $\theta : C_1 \rightrightarrows C_2$ — if all of the following rules satisfied.

Ob1: θ maps objects in C_1 to objects in C_2. We write

$$\forall E \in Ob(C_1), \quad E^{\theta} = \theta(E) \in Ob(C_2) \tag{72}$$

Mor1: θ associates morphisms in C_1 to morphisms in C_2 that respects Ob1.

$$\forall f \in \mathrm{Mor}_{C_1}(E_1, E_2), \quad \theta(f) \in \mathrm{Mor}_{C_2}(E_1^{\theta}, E_2^{\theta}) \tag{73}$$

Mor2: Identity is associated with identity: $\theta(\mathrm{id}_E) = \theta(\mathrm{id}_{\theta(E)})$.

Mor3: Commutes with inversion: $\theta(f^{-1}) = \theta(f)^{-1}$ if the inverse of f exists.

Mor4: Functoral: $\theta(g \circ f) = \theta(g) \circ \theta(f)$.

Note 2.1: The Hom_X functor

We continue our discussion from fig. 5. Recall $\mathrm{Ban}_{\mathbb{R}}$ is the category of Banach spaces over $\mathbb{R}$, and we will refer to toplinear morphisms as morphisms for brevity. If X is an object in $\mathrm{Ban}_{\mathbb{R}}$, the Hom_X *functor* is a covariant functor between $\mathrm{Ban}_{\mathbb{R}}$ and $L(X, \cdot)$ — the space of toplinear mappings with source X such that

1. to each $E_i \in \mathrm{Ob}(\mathrm{Ban}_{\mathbb{R}})$ $\mathrm{Hom}_X(E_i) = L(X, E_i)$ — **the space of incoming morphisms with source X**, and
2. to each morphism $f \in L(E_1, E_2)$ another morphism between $L(X, E_1)$ and $L(X, E_2)$ — denoted by $(\mathrm{Hom}_X)f$.
3. The functor Hom_X converts **outgoing morphisms from E_1 to the redirection morphism of incoming morphisms with source X**.

Notice this is precisely what the diagram in fig. 5 describes.

Proposition 2.1: Hom_X functor is a functor

The Hom_X functor as defined in note 2.1 is a covariant functor.

Proof. Postponed. ■

Note 2.2: The tangent space functor

Let X be a C^1 manifold, we call the tuple (p, X) for $p \in X$ the *centering of X centered at p*. The category of pointed manifolds, denoted by Man_*. Its objects consist of all centerings across C^1 manifolds, and the morphisms in Man_* are called *pointed morphisms*.

If (q, Y) is another object in Man_*, a pointed morphism between (p, X) and (q, Y) is a tuple (p, f); where f is a manifold morphism between X and Y and $f(p) = q$. We sometimes write $f_p = (p, f)$ when it is clear.

The *tangent space functor*, denoted by $T : \mathrm{Man}_* \rightrightarrows \mathrm{Ban}_{\mathbb{R}}$ is a covariant functor where

- we define $T(p, X) = T_p X$ that takes a pointed C^1 manifold to its tangent space, and
- to each pointed C^1 morphism $f_p \in \mathrm{Mor}_{\mathrm{Man}_*}((p, X)\,(q, Y))$ we associate with the toplinear mapping

$$T f_p = df(p) : T_p X \to T_q Y \tag{74}$$

Proposition 2.2: Tangent space functor is a functor

The tangent space functor as defined in note 2.2 is a covariant functor.

Proof. Postponed. ■

We leave the verification that T satisfies the properties in def. 2.1 as an exercise. Dual to the concept of a functor is that of the cofunctor, which — for our purposes — captures the idea of the toplinear adjoint.

Definition 2.2: Cofunctor

Let C_1 and C_2 be categories. A correspondence $\eta : C_1 \leftrightarrows C_2$ is called a *cofunctor* (or a contravariant functor) if all of the following rules are satisfied.

Ob: η maps objects in C_1 to objects in C_2. We write $E^\lambda = \lambda(E) \in \mathrm{Ob}(C_2)$ for every $E \in \mathrm{Ob}(C_1)$.

Mor1: η associates morphisms in C_1 to morphisms in C_2 that respects Ob1.

$$\forall f \in \mathrm{Mor}_{C_1}(E_1, E_2), \quad \eta(f) \in \mathrm{Mor}_{C_2}(E_2^\eta, E_1^\eta) \tag{75}$$

Mor2: Identity is associated with identity: $\eta(\mathrm{id}_E) = \eta(\mathrm{id}_{\eta(E)})$.

Mor3: Commutes with inversion: $\eta(f^{-1}) = \eta(f)^{-1}$ if the inverse of f exists.

Mor4: Cofunctoral: $\eta(g \circ f) = \eta(f) \circ \eta(g)$.

Remark 2.1: Cofunctors are opposite to functors

The cofunctor η reverses the arrows a morphism f. Refer to fig. 6 for a comparison between eq. (75) and eq. (73).

Note 2.3: The Hom^X cofunctor

Let $X \in \mathrm{Ob}(\mathrm{Ban}_{\mathbb{R}})$, it defines a cofunctor from $\mathrm{Ban}_{\mathbb{R}} \leftrightarrows L(\cdot, X)$ where $L(\cdot, X)$ is the space of toplinear mappings whose destination is X.

1. to each $E_i \in \mathrm{Ob}(\mathrm{Ban}_{\mathbb{R}})$ $\mathrm{Hom}^X(E_i) = L(E_i, X)$ — **the space of outgoing morphisms with destination X**, and
2. to each morphism $f \in L(E_1, E_2)$ another morphism between $L(E_1, X)$ and $L(E_2, X)$ — denoted by $(\mathrm{Hom}^X)f$.
3. The functor Hom^X converts **outgoing morphisms from E_1 to the precomposition morphism which acts on morphisms with destination X**.

$$\begin{tikzcd}
\eta(E_1) & E_1 \arrow[l, "\eta"'] \arrow[r, "\theta"] \arrow[d, "f"] & \theta(E_1) \arrow[d, "\theta(f)"] \\
\eta(E_2) \arrow[u, "\eta(f)"] & E_2 \arrow[l, "\eta"'] \arrow[r, "\theta"] & \theta(E_2)
\end{tikzcd}$$

Figure 6: Functor θ vs. cofunctor η comparison

Tangent Bundle

Definition 3.1: Tangent bundle

Let X be a C^p manifold over E, the *tangent bundle* is a E-vector bundle of rank $p-1$, denoted by TX, and

$$TX = \coprod_{x \in X} T_x X.$$

The construction of the tangent bundle is outlined in note 3.1.

Note 3.1: Construction of the Tangent Bundle

Let X be a C^p manifold with $p \geq 1$, so that the tangent space at every point is defined. If $p \in (U_i, \varphi_i)$ for $i = 1, 2$. Then φ_{12} is a C^p isomorphism between $\varphi_1(U_{12})$ and $\varphi_2(U_{12})$; **whose derivative is a C^{p-1} map into** $\mathrm{Laut}(E)$ **that encodes the transformation between the concrete tangent spaces**. In the notation of eq. (33), this means

$$x \mapsto D\varphi_{12}(x) \quad \text{is in } C^{p-1}(\hat{U}_{12}, \mathrm{Laut}(E))$$

In fact, the tangent bundle $TX \triangleq \coprod_{p \in X} T_p X$ is a C^{p-1} vector bundle (modelled on E) over X. If (U, φ) is a chart in X, it induces a local trivialisation on TX by taking each tangent vector $v \in T_p X$ to its concrete representation $(p, \hat{v}) \in X \times E$.

$$\Phi : \widetilde{U} \to U \times E \quad \text{and} \quad \Phi(v) = (p, \hat{v}) \tag{76}$$

where $(U, \varphi, p, \hat{v})$ is a concrete representation of $v \in T_p X$.

Similarly, we have the cotangent bundle which is modelled on toplienar dual of the tangent spaces of X.

Definition 3.2: Cotangent bundle

Let X be a C^p manifold over E, the *cotangent bundle* is a E-vector bundle of rank $p-1$, denoted by $T * X$, and

$$T^*X = \coprod_{x \in X} T_x X^*,$$

where $T_x X$ is toplinearly isomorphic to X, and $T_x X^*$ its toplinear dual.

Chapter 3: 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 $f \in \mathrm{Mor}(X, F)$ and fix a point $p \in X$, and write $q = f(p)$. By adopting the canonical interpretation $\overline{w}$ for a tangent vector $w \in T_q F$ (as discussed in remark 5.1), 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 $f : X \to \mathbb{R}$, 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 $C^p(U)$.

Exterior Derivative

Let X be a manifold, and $f \in C^p(X)$. If $\gamma : (-\delta, +\delta) \to X$ is a curve starting at $x_0 \in X$ with velocity v, the composition $f \circ \gamma$ is a morphism. Let us write

$$F : (-\delta, +\delta) \to \mathbb{R}, \quad F(\varepsilon) = f \circ \gamma(\varepsilon) - f \circ \gamma(0) \tag{77}$$

Suppose we wish to measure the rate at which f moves in the direction of v, then we can simply take the derivative of eq. (77). We define the *exterior derivative of f at x_0*, denoted by $df(x_0) : T_{x_0} X \to \mathbb{R}$ by eq. (78)

$$df(x_0)(v) = DF(0)(\overline{1}) \quad \text{where} \quad F = f \circ \gamma \tag{78}$$

for any curve starting at x_0 with velocity v.

Let E be a Banach space, and suppose ω is a k-form on E, and $x_0 \in E$ with $k+1$ tangent vectors $v_{\underline{k+1}}$. The parallelopiped defined by the $k+1$ vectors is

$$P_{x_0}(v_{\underline{k+1}}) = \left\{ x_0 + \sum_{i=\underline{k+1}} t_i v_i,\ 0 \leq t_i \leq 1\ \forall i = \underline{k+1} \right\}$$

As with eq. (77), we can integrate over the boundary defined by $P_{x_0}(v_{\underline{k+1}})$, and obtain a new function. We can shrink each v_i by ε_i, and we define

$$P_{x_0}^{\varepsilon_{\underline{k+1}}}(v_{\underline{k+1}}) = \left\{ x_0 + \sum_{i=\underline{k+1}} t_i v_i,\ 0 \leq t_i \leq \varepsilon_i\ \forall i = \underline{k+1} \right\} \tag{79}$$

(Note: Perhaps after shrinking the domain of F, here we should replace everything by their coordinate representations).

$$F : (-\delta, +\delta)^{k+1} \to \mathbb{R}, \quad F(\varepsilon_{\underline{k+1}}) = \int_{\partial P_{x_0}^{\varepsilon_{\underline{k+1}}}(v_{\underline{k+1}})} \omega \tag{80}$$

We define the exterior derivative of a k-form by the map $DF(0)(1^{(k+1)})$.

Derivations

For the rest of this chapter, assume all manifolds to be C^p-manifolds over $\mathbb{R}^n$, where $n, p \geq 1$. Let E and F be Banach spaces and $U \overset{\circ}{\subseteq} E$, 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. (81).

$$\hat{\mathcal{D}} : (U \times E) \times C^p(U, F) \longrightarrow F : \quad \Big((p, v), f\Big) \mapsto Df(p)(v) \in F \tag{81}$$

Suppose $F = \mathbb{R}$ and denote pointwise multiplication on $\mathbb{R}$ by m. The above pairing trivially satisfies the product rule displayed in eq. (82).

$$Dm(f_{\underline{k}})(p)(v) = \sum_{i=\underline{k}} m(f_{\underline{i-1}}(p), Df_i(p)(v), f_{i+\underline{k-i}}(p)) \tag{82}$$

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 $v \in T_pX$, the commentary in the introduction suggests a 'duality pairing' between f and (p, v) in the form of eq. (83).

$$\mathcal{D} : (U \times E) \times C^p(U, F) \longrightarrow F \quad \text{and} \quad \mathcal{D}\Big((p, v), f\Big) = df_p(v) \tag{83}$$

By definition of the differential df_p, the right hand side of eq. (83) is representation independent, hence

$$\mathcal{D}((p, v), f) = D\hat{f}(\hat{p})(\hat{v}), \quad \text{where the right member is an ordinary derivative} \tag{84}$$

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_pX for each $p \in U$. We will make this notion precise when we introduce covectors.

Definition 3.1: 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 $f_{\underline{k}} \in C^p(U)$, eq. (85) holds.

$$v\Big(m(f_{\underline{k}})\Big) = \sum_{i=\underline{k}} m(f_{\underline{i-1}}(x), v(f_i), f_{i+\underline{k-i}}(x)) \tag{85}$$

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. (82) and its computation in coordinates in eq. (84). If X is finite-dimensional, prop. 3.1 shows derivations at a point $p \in X$ are uniquely represented by a tangent vector.

Proposition 3.1: T_pX 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. (84).

Proof. Postponed. ■

Boundary

Definition 4.1: Subsets of the upper half-plane

Let $n \geq 1$, the *upper half plane* of $\mathbb{R}^n$ is the subset $\mathbb{H}^n = [x^n \geq 0]$. Its topological interior (resp. boundary) is denoted by $\operatorname{Int}\mathbb{H}^n = [x^n > 0]$ (resp. $\partial\mathbb{H}^n = [x^n = 0]$).

We wish to define a notion of the boundary for a manifold. Instead of modelling X on open subsets of $\mathbb{R}^n$, we consider open subsets of both $\mathbb{R}^n$ and $\mathbb{H}^n$. To wit, we start by generalizing the notion of C^p smoothness between **arbitrary subsets of** $\mathbb{R}^n$. Let $S \subseteq \mathbb{R}^n$ be a subset of $\mathbb{R}^n$, recall that:

> A function $f : S \to \mathbb{R}^m$ is said to be continuous on S, if $f^{-1}(U)$ is relatively open in S for every $U \overset{\circ}{\subseteq} \mathbb{R}^m$.

The take-away intuitinon for this is that manifolds with boundary are supposedly used to model geometric objects that are suddenly 'cut off'. In the case of the upper half-plane, this manifests in a sudden stop in the n-th coordinate. The morphisms we seek are the ones which are ordinary morphisms but whose domains 'cut off'.

> Let $F : S \to S'$ be a mapping between arbitrary subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. It is a C^p morphism, whenever each point $p \in S$ admits a C^p extension to a neighbourhood containing p.

With this, a C^p isomorphism between subsets of $\mathbb{R}^n$ is a bijective C^p morphism whose inverse is also a C^p morphism.

Definition 4.2: Boundary chart

Let X be a non-empty subset. A *boundary chart* on X modelled on $\mathbb{R}^n$ is a tuple (U, φ), such that $U \subseteq X$, $\varphi(U) = \hat{U}$ is an open subset of $\mathbb{H}^n$, and φ is a bijection onto $\hat{U}$.

To distinguish between this new definition of the previous one, the word *chart* will always refer to the charts defined in Chapter 1. If we wish to be precise, we will use the word *interior chart*.

Definition 4.3: Compatibility between boundary and interior charts

Let (U, φ) and (V, ψ) be boundary or interior charts of X, modelled on $\mathbb{R}^n$ or $\mathbb{H}^n$. They are called C^p*-compatible* (for $p \geq 0$) if $U \cap V = \varnothing$, or both of the following hold:

- $\varphi(U \cap V)$ and $\psi(U \cap V)$ are *both* open subsets of $\mathbb{R}^n$ or $\mathbb{H}^n$.
- the *transition map* $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ is a C^p isomorphism between open subsets of $\mathbb{R}^n$ or $\mathbb{H}^n$.

Definition 4.4: Boundary atlas, structure with boundary

Let X be a non-empty set and $p \geq 0$. A C^p *boundary atlas* on X modelled on $\mathbb{R}^n$ is a pairwise C^p-compatible collection of charts $\{(U_\alpha, \varphi_\alpha)\}$ whose union over the domains cover X.

If $\mathcal{A}$ is a boundary atlas of X, the maximal boundary atlas containing $\mathcal{A}$ is called the C^p *boundary* *structure determined by* $\mathcal{A}$. A C^p *manifold with boundary* modelled on $\mathbb{R}^n$ is a a non-empty set X with a C^p boundary structure modelled on $\mathbb{R}^n$.

We also have the following terminology for points in a manifold with boundary X. Let $p \in X$,

- p is called a *boundary point* whenever it is in the domain of a boundary chart (U, φ), **and** $\varphi(p) \in \partial\mathbb{H}^n$,
- p is called an *interior point* otherwise.

Remark 4.1

It can be shown (using deRham cohomology, see Chapter 16-17 [4]) that a point on a manifold cannot be a boundary point and an interior point at the same time.

Chapter 4: Symplectic Geometry

Primer on Differential Forms

Remark 1.1: Finite-dimensional Manifolds

We assume all manifolds modelled over $\mathbb{R}^n$ $(n \geq 1)$ are of class C^∞, and are equipped with Hausdorff, second-countable topologies.

Let M be a manifold modelled on $\mathbb{R}^n$.

- $\mathfrak{X}(M) = (C^\infty(M))$ module of vector fields on M,
- $\mathfrak{X}^*(M)$ = module of covector fields on M,
- $\mathcal{T}^{(j,k)}(M)$ = module of j-contravariant, k-covariant tensor fields on M.
- $\mathcal{T}^k(M) = \mathcal{T}^{(0,k)}(M)$.
- $\Omega^k(M)$ = module of k-forms on M.

Note 1.1: Covariant and Contravariant Tensors

We recall that if V is a $\mathbb{R}$-vector space, a j-contravariant, k-covariant tensor on V — denoted by F — is a $(j+k)$ linear mapping that takes j-covectors, and k-vectors to a real number. In symbols,

$$F : (V^*)^j \times V^k \to \mathbb{R} \quad \text{is multilinear.}$$

We denote the space of (j,k) tensors on V by $\mathcal{T}^{(j,k)}(V)$. The space of $(0,0)$ tensors on V is identified with $\mathbb{R}$ — as it depends on 0 arguments.

If V is finite dimensional, then $V = \mathcal{T}^{(1,0)}(V)$, and $V^* = \mathcal{T}^{(0,1)}(V)$. Similarly, $\mathfrak{X}(M) = \mathcal{T}^{(1,0)}(M)$, $\mathfrak{X}^*(M) = \mathcal{T}^{(0,1)}(M)$ and $C^\infty(M) = \mathcal{T}^{(0,0)}(M)$.

If N is another manifold and $u : M \to N$ a morphism,

- for every $f \in C^\infty(N)$, the *pullback* through u is the precomposition $u^* f = f \circ u \in C^\infty(M)$, and
- for every $A \in \mathfrak{X}^*(N)$, the *tensor field pullback* through u is the precomposition. Ita is defined by

  $$(u^*A)(p)(v) = A[u(p)][du(p)(v)], \quad \text{where the square brackets are for readability.}$$

  For a general $A \in \mathcal{T}^k N$, we have

  $$(u^*A)(p)(v_{\underline{k}}) = A[u(p)][du(p)(v_{\underline{k}})], \quad \text{for an arbitrary } p \in M,\ v_k \in T_pM.$$

- If u is a diffeomorphism, we define the *vector field pullback* of a vector field $Y \in \mathfrak{X}(N)$ by

  $$(u^*Y)(p) = du^{-1}\Big(Y_{u(p)}\Big) = (du^{-1} \circ Y \circ u)(p)$$

We recall a few facts from differential geometry.

- If $f \in C^\infty(M)$, the *exterior derivative* of f is the covector field df with coordinate representation

  $$df(p) = \sum_{i=\underline{n}} \frac{\partial f}{\partial x^i} dx^i.$$

- If $A \in \Omega^k(M)$, the *exterior derivative* of A is a $k+1$ form that is defined by its local coordinate representation.
- The exterior derivative d commutes with the tensor field pullback. That is, for every $A \in \Omega^k(N)$, $u^*(dA) = du^*A$.

Definition 1.1: Exterior Derivative in Local Coordinates

Let M be a manifold modelled on $\mathbb{R}^n$, and $A \in \Omega^k(M)$. If (x^i) are the local coordinates in some open subset $U \mathring{\subseteq} M$, A can be written as the tensor product of dual basis vectors (dx^i).

$$A = \sum_J{}' A_J dx^J \tag{86}$$

where $\sum'$ refers to an increasing sum taken over k-indices. We define the *exterior derivative of* A by the $k+1$ form in local coordinates

$$dA = d\left(\sum_J{}' A_J dx^J\right) = \sum_J{}' dA_J \wedge dx^J. \tag{87}$$

Unboxing the differential of A_J and the wedge product, eq. (87) becomes:

$$dA = \sum_J{}' \sum_{i=\underline{n}} \frac{\partial A_J}{\partial x^i} dx^i \wedge dx^J = \sum_J{}' \sum_{i=\underline{n}} \frac{\partial A_J}{\partial x^i} dx^{(i,J)}. \tag{88}$$

Example 1.1: Exterior Derivative in Coordinates

Let $M = \mathbb{R}^3 \setminus \{0\}$, we will use the standard coordinates (x, y, z) on M.

1. $f(x, y, z) = (x^2 + y^2)^{1/2}$ = scalar valued function.
2. $A(x, y, z) = (y - x)dz - zdy$ = covector field.
3. $B(x, y, z) = f(x, y, z)dx \wedge dy$ = 2-form.

Exterior Derivative of f:

$$df(x, y, z) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = \frac{xdx + ydy}{f(x, y, z)}$$

Exterior Derivative of A:

$$\begin{aligned} dA(x, y, z) &= d(y - x) \wedge dz + d(-z) \wedge dy \\ &= dy \wedge dz - dx \wedge dy - dz \wedge dy = 2dy \wedge dz - dx \wedge dy \end{aligned}$$

Exterior Derivative of B:

$$dB(x, y, z) = df(x, y, z) \wedge (dx \wedge dy) = \frac{xdx + ydy}{f(x, y, z)} \wedge (dx \wedge dy) = 0$$

Remark 1.2: Exterior Derivative on Banach Manifolds

If X is a Banach space, which is also a Banach manifold of class C^k for $k \geq 1$, the exterior derivative of C^k function f is a C^{k-1} covector field whose evaluation at $p \in X$ coincides with the Frechet Derivative $Df(p)$. Recall that $Df(p)$ is the unique linear map that satisfies

$$f(p+v) = f(p) + Df(p)(v) + o(|v|).$$

Remark 1.3: Closed, and exact differential forms

Let $A \in \Omega^k(M)$ be a k-form on manifold M.

- It is *closed* whenever $dA = 0$, and
- is *exact* whenever $A = dB$ where $B \in \Omega^{k-1}(M)$.

Remark 1.4: Poincare's Lemma

A subset $S \subseteq \mathbb{R}^n$ is said to be *star-shaped* if there exists some $a \in S$ where $\{a + (b-a)[0,1]\} \subseteq S$ for every $b \in S$. That is, the straight line segment between a and every point S is contained in S.

Poincare's Lemma states that, if U is an open, star-shaped subset of $\mathbb{R}^n$, then every closed form is exact.

Remark 1.5: Line integral

Let $\gamma : [0,L] \to M$ where M is a smooth manifold. For any smooth 1-form λ on M, the integral of γ over λ is the integral

$$\int_\gamma \lambda = \int_{[0,L]} \gamma^* \lambda = \int_0^L \lambda(\gamma(t))(\mathring{\gamma}(t))dt.$$

Example 1.2: Line integral in coordinates

Let $\gamma(t) = (\cos(2\pi t), \sin(2\pi t), 0)$ for $t \in [0,1]$, and the covector field

$$A(x,y,z) = \frac{xdy - ydx}{x^2+y^2} \quad \forall (x,y) \neq 0.$$

Suppressing the trigonometric arguments, the line integral of A over γ is given by

$$\int_\gamma A = \int_0^1 A[\gamma(t)][\dot{\gamma}(t)]dt = \int_0^1 \frac{\cos dy - \sin dx}{\cos^2 + \sin^2}(2\pi(-\sin, \cos, 0))dt$$

which gives

$$\int_\gamma A = 2\pi \int_0^1 \cos\cos - \sin(-\sin)dt = 2\pi.$$

Standard Symplectic Form

We begin the case in $\mathbb{R}^2$. The *standard symplectic form* on $\mathbb{R}^2$ is the bilinear form represented by the matrix (with repect to the standard basis) in eq. (89).

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{89}$$

The following note summarizes several properties of J.

Note 2.1: Properties of the standard symplectic form on $\mathbb{R}^2$

If $x, y \in \mathbb{R}^2$, eq. (89) defines a pairing $\omega_0 \in \Omega^2(\mathbb{R}^2)$ between x and y. Where $\omega_0(x, y) = \langle x, Jy \rangle_{\mathbb{R}^2}$. An easy computation in coordinates will show that

$$\omega_0(x, y) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \det\left(\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \right) = \det(x, y) \tag{90}$$

Furthermore,

- J is non-singular and skew-symmetric, and $J^{-1} = (-1)J$.
- ω is non-singular and skew-symmetric, it is a non-degenerate 2-form on $\mathbb{R}^{2n}$ by lem. 5.1.
- Left multiplication by a vector $v = (v^1, v^2)$ reads and $\omega_0(v, \cdot) = v^1\varepsilon^2 + (-1)v^2\varepsilon^1$.
- Right multiplication by v: by skew-symmetry of J reads: $\omega_0(\cdot, v) = (-1)v^1\varepsilon^2 + v^2\varepsilon^1$.

Definition 2.1: Standard symplectic form

Let $n \geq 1$, the *standard symplectic form* is the bilinear form defined by the matrix representation in eq. (91).

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \tag{91}$$

The matrix in eq. (91) induces a bilinear pairing, which we will denote by $\omega_0 \in \Omega^2(\mathbb{R}^{2n})$. Its defining property is that it computes the sum of n 2×2 determinants, as shown in eq. (92).

$$\omega_0(x, y) = \langle x, y \rangle_{\omega_0} = \sum_{i=\underline{n}} \det\left(\begin{bmatrix} x^i & y^i \\ x^{n+i} & y^{n+i} \end{bmatrix} \right) \tag{92}$$

We can rewrite ω_0 using the language of differential forms:

$$\omega_0 = \sum_{i=\underline{n}} \varepsilon^i \wedge \varepsilon^{n+i}. \tag{93}$$

The properties outlined in note 2.1 all hold for $\mathbb{R}^{2n}$. Moreoever, ω_0 is exact, as one can verify that if $\lambda = \sum x^i dx^{n+i}$ with the sum taken over $\underline{n}$, then $d\lambda = \omega_0$. Recall if $p, v \in \mathbb{R}^{2n}$,

$$\lambda(p) = \sum p^i dx^{n+i} \quad \text{and} \quad \lambda(p)(v) = \sum p^i v^{n+i}.$$

Remark 2.1: Alternate Symplectic Structure

Some texts use eq. (94), or $J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$. The following decomposition is called the *maximal hyperbolic decomposition* of $\mathbb{R}^{2n}$, see [6] Chapter 13. We will return to this later when we discuss periodic solutions on ellipsoids.

$$J = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{bmatrix} \tag{94}$$

Note 2.2: Computations with the standard symplectic form

We call the standard symplectic form given in eq. (91) in terms of the Kronecker delta. A moment's thought will show that $J = [\delta_{(i,j-n)} - \delta_{(i,j+n)}]_{ij} = [\delta_{(n+i,j)} - \delta_{(i-n,j)}]_{ij}$. Left multiplication by a vector $v = (v^{\underline{2n}})$ yields

$$\begin{aligned} \omega_0(v, \cdot) &= \langle v, J\cdot\rangle_{\mathbb{R}^{2n}} = v^i[\delta_{(i,j-n)} - \delta_{(i,j+n)}]\varepsilon^j \\ &= \sum_{i=\underline{2n}} v^i\varepsilon^{i+n} - v^i\varepsilon^{i-n} = \sum_{i=\underline{n}} v^i\varepsilon^{i+n} - v^{i+n}\varepsilon^i \end{aligned}$$

Right multiplication then give us

$$\omega_0(\cdot, v) = \langle \cdot, Jv\rangle_{\mathbb{R}^{2n}} = \sum_{i=\underline{n}} (-1)v^i\varepsilon^{i+n} + v^{i+n}\varepsilon^i.$$

Symplectic Manifolds

We introduce a more differential geometric viewpoint, and work with arbitrary symplectic structures.

Definition 3.1: Symplectic Manifold

A *symplectic manifold* is a manifold M modelled on $\mathbb{R}^{2n}$ (for $n \geq 1$), equipped with a **closed, non-degenerate 2-form ω**. We sometimes refer the tuple (M, ω) as the *symplectic structure*.

Definition 3.2: Symplectomorphism

Let (M, ω) and (N, η) be symplectic manifolds of dimension $2m$ and $2n$ respectively. A mapping $u : M \to N$ is a *symplectomorphism* (or is symplectic as an adjective) whenever it preserves the symplectic structure under the tensor pullback. That is,

$$u^*\eta = \omega, \quad \text{which means} \quad \omega(p)(v_1, v_2) = \eta(u(p))\Big(du(p)[v_1],\, du(p)[v_2]\Big) \quad \forall p \in M,\, v_{\underline{2}} \in T_pM.$$

An embedding that is a symplectomorphism is called a *symplectic embedding*.

Example 3.1: Symplectomorphism on $\mathbb{R}^{2n}$

Let $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be smooth, we say φ is a *symplectomorphism* (or φ is symplectic as an adjective) whenever it preserves ω. That is,

$$\langle D\varphi(x)(v_1), D\varphi(x)(v_2)\rangle_{\omega_0} = \omega_0(v_1, v_2) \quad \forall x, v_1, v_2 \in \mathbb{R}^{2n},$$

where $D\varphi(x)$ refers to the Jacobian matrix of φ evaluated at $x \in \mathbb{R}^{2n}$. If φ is a C^∞ diffeomorphism and φ and its inverse are symplectomorphisms, we call φ a *symplectic diffeomorphism* or a *symplectic isomorphism*.

Definition 3.3: Symplectic action on (M, ω)

If (M, ω) is a symplectic manifold, we write

$$\langle v_1, v_2\rangle_{\omega(p)} = \omega(p)(v_1, v_2) \quad \text{for every } p \in M, \text{ and } v_1, v_2 \in T_pM.$$

Given an interval $\mathcal{I} \subseteq \mathbb{R}$, the *symplectic pairing* between two curves is defined to be $A(\gamma, \eta) = 2^{-1}\int_{\mathcal{I}}\langle \mathring{\gamma}, \eta\rangle_{\omega_0}$, for every $\gamma, \eta \in C^\infty(\mathcal{I}, M)$. The *symplectic action* on a curve γ is

$$A(\gamma) = A(\gamma, \gamma) = 2^{-1}\int_{\mathcal{I}}\langle \mathring{\gamma}, \gamma\rangle_\omega$$

Remark 3.1

Symplectomorphisms are volume preserving. If $\varphi : M \to N$ is a symplectomorphism, then the determinant of the Jacobian matrix (with respect to any pair of charts) is 1.

Lemma 3.1: Symplectomorphisms from $\mathbb{R}^{2n}$ preserve the symplectic action

If $\varphi : (\mathbb{R}^{2n}, \omega_0) \to (M, \eta)$, then $A(\varphi \circ \gamma) = A(\gamma)$ for every $\gamma \in \Omega$.

Proof. Because the exterior derivative commutes with the tensor pullback, we have

$$d(\lambda - \varphi^*\lambda) = d\lambda - \varphi^*(d\lambda) = 0$$

whence $\lambda - \varphi^*\lambda$ is a closed differential form. Using eq. (102), we see that

$$A(\varphi \circ \gamma) - A(\gamma) = \int_\gamma \varphi^*\lambda - \int_\gamma \lambda = \int_\gamma (\varphi^*\lambda - \lambda)$$

It follows from Poincare's Lemma that the right hand vanishes, since it is the closed curve over an exact form. ■

Corollary 3.1

If $\varphi : (M, \omega) \to (N, \eta)$ is a symplectomorphism, and $H^1_{dR}(M) = 0$, then $A(\varphi \circ \gamma) = A(\gamma)$ for every curve γ in M.

Hamiltonian Vector Fields

We begin with some well known sign conventions for area.

> Fix any two vectors $v_1, v_2 \in \mathbb{R}^n$, we say that positive area opens anti-clockwise from v_1. (Draw a picture).

Because of this, we refer to $\det(v_1, v_2)$ as the *area spanned from* v_1 *to* v_2 if $v_1, v_2 \in \mathbb{R}^2$, and we call $\omega_0(v_1, v_2)$ the *symplectic area from* v_1 *to* v_2. Moreoever, if S is a compact region in $\mathbb{R}^n$ whose (topological or manifold) boundary ∂S can be traversed by a continuous curve $\gamma : [0, 1] \to \mathbb{R}^n$.

> We say γ is positively oriented (with respect to the *Stokes' orientation*), whenever the region S lies to the left of γ at every point.

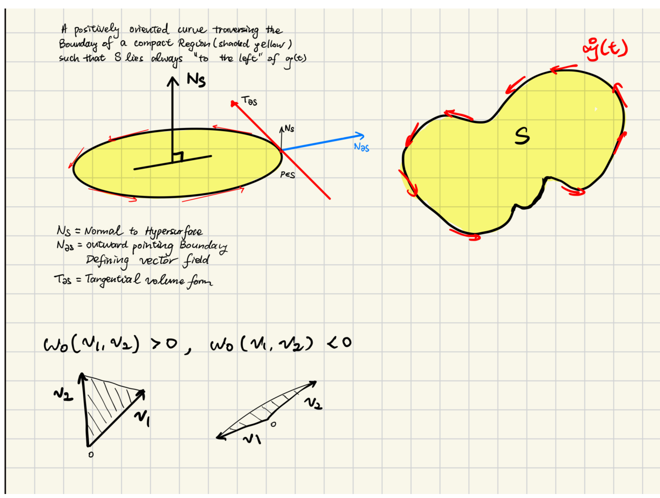

Figure 7: Illustrations of area sign conventions

Remark 4.1: Positive Gradient Flow

Let $H \in C^\infty(\mathbb{R}^{2n})$, the *positive gradient flow* of H is the vector field ∇H such that at every point $p \in \mathbb{R}^{2n}$, and $v_p \in T_p\mathbb{R}^{2n}$:

> The **angle between** $\nabla H(p)$ **and** v_p **is equal to** $DH(p)(v_p)$, where
>
> $$H(p + v_p) = H(p) + DH(p)(v_p) + o(|v_p|) \quad \text{for sufficiently small } v. \tag{95}$$

By the 'angle' we refer to the Euclidean inner product which takes on values in $\mathbb{R}$ instead of in $[-\pi, +\pi]$.

Moreover, the *Euclidean gradient* of H in coordinates is given by

$$\nabla H = (\partial_{\underline{2n}} H) \in \mathfrak{X}(\mathbb{R}^{2n}).$$

Definition 4.1: Hamiltonian Flow

The *Hamiltonian flow* of H is the vector field X_H such that at every point $p \in \mathbb{R}^{2n}$, and $v_p \in T_p\mathbb{R}^{2n}$:

The **symplectic area from $X_H(p)$ to v_p is equal to $DH(p)(v_p)$.**

More precisely, the Hamiltonian flow of H is defined by the sharpening the covector field of H: $X_H = \omega_0^{\wedge}(dH)$, such that

$$\omega_0(X_H(p), v_p) = dH(p)(v_p) \quad \text{for all } p \in \mathbb{R}^{2n},\ v_p \in T_p\mathbb{R}^{2n}. \tag{96}$$

In Euclidean space, X_H has a simple structure, and is related to the ∇H by a factor of J.

Lemma 4.1: Hamiltonian Flows in Euclidean Space

The Hamiltonian flow of $H \in C^\infty(\mathbb{R}^{2n}, \mathbb{R})$, X_H has matrix representation which satisfies

$$X_H = J\nabla H. \tag{97}$$

Proof. Let p and v_p be arbitrary, it follows from the definition of X_H that

$$\langle X_H(p), v_p\rangle_{\omega_0} = \langle X_H(p), Jv_p\rangle_{\mathbb{R}^{2n}} = dH(p)(v_p) = \langle \nabla H(p), v_p\rangle_{\mathbb{R}^{2n}}.$$

Notice that J is skew-symmetric, so we can move J over to the other side of the bracket at the cost of a minus sign, hence:

$$\langle (-1)JX_H(p), v_p\rangle_{\mathbb{R}^{2n}} = \langle \nabla H(p), v_p\rangle_{\mathbb{R}^{2n}}.$$

The proof is complete upon seeing that $(-1)J = J^{-1}$. ∎

Remark 4.2

If X_H is a Hamiltonian flow, we sometimes refer to its integral curves as *solutions*. If a solution is periodic, it is called an *orbit* of X_H.

Statement of the Weinstein's Conjecture

Definition 5.1: Closed submanifold

Let M be a manifold modelled on $\mathbb{R}^n$. A submanifold $S \subseteq M$ is said to be *closed* whenever it is a compact subset of M, and the S is a manifold without boundary.

Remark 5.1: Stokes' Theorem

Let M be a manifold with boundary modelled on $\mathbb{R}^{2n}$, for any compactly supported $(n-1)$ form ω:

The integral of $d\omega$ over M is equal to the integral of ω over ∂M. In symbols,

$$\int_M d\omega = \int_{\partial M} \omega.$$

If S is a closed submanifold of M, and ω a $(n-1)$-form, an immediate corollary is that

$$\int_S d\omega = \int_{\partial S} \omega = 0.$$

Definition 5.2: Regular hypersurface

A *regular hypersurface* on a smooth manifold M is a subset $S = f^{-1}(c)$ where $f \in C^\infty(M, \mathbb{R})$, and $df(p) \neq 0$ for every $p \in S$. We call f the *defining function* of S which admits a natural manifold structure that makes S a submanifold of M.

Definition 5.3: Energy surface

An *energy surface* is a compact, regular hypersurface of a symplectic manifold (M, ω).

We conclude this section by stating Weinstein's conjecture on $\mathbb{R}^{2n}$.

Does every energy surface on $(\mathbb{R}^{2n}, \omega_0)$ admit a non-degenerate, closed orbit?

A more abstract reformulation of the conjecture is given below.

Given an energy surface S, does its line bundle $\mathcal{L}(S)$ admit a closed characteristic?

Where $\mathcal{L}(S) = \Big\{(x, v) \in TS,\ v \in \operatorname{rad}(\omega_0(p))\Big\}$ and a *closed characteristic* of S refers to a 1-dimensional submanifold that is diffeomorphic to the 1-sphere.

Proposition 5.1: WC Reduction 1 — Independence of Hamiltonian

Let S be a compact, regular hypersurface on a symplectic manifold (M, ω_0). If $F, G \in C^\infty(M)$ are defining functions of S such that

$$S = F^{-1}(c) = G^{-1}(c'),$$

where

$$dF(x) \neq 0 \quad \text{and} \quad dG(x) \neq 0 \quad \forall p \in S.$$

Then, there exists a $\rho \in C_c^\infty(M, \mathbb{R})$ that satisfy the following

- For every $x \in S$, $\rho(x) \neq 0$ and $dF(x) = \rho(x) dG(x)$.
- For any $x \in S$, let $\varphi_x(s) = \varphi(s, x)$ and $\theta_x(t) = \theta(t, x)$ denote the integral curves starting at x of X_F and X_G. The smooth function α constructed by solving the IVP in eq. (98) relates the two flows by its reparametrization.

$$\frac{\mathrm{d}\alpha}{\mathrm{d}s} = \rho(\varphi_x(s)) \quad \alpha(0) = 0 \tag{98}$$

By reparametrization we mean that $\varphi_x(s) = \theta_x(\alpha(s))$ for all s whenever either side is defined.

- The periodic orbits of X_F and X_G on S correspond bijectively.
- For any $x \in S$, φ_x is a non-degenerate periodic orbit if and only if $\theta_x \circ \alpha$ is.

Proof. Both F and G are global defining functions of the submanifold S, if $p \in S$ is arbitrary, the exterior tangent space coincides precisely with $\operatorname{Ker} dF(p) = \operatorname{Ker} dG(p) = T_p^{\text{ext}}(S)$. (Lee 5.38, 5.40). Since $T_p^{\text{ext}} S$ has dimension 1, there exists a suitably chosen coordinate chart (U, ζ) about p such that $dz \in \mathfrak{X}^*(\mathbb{R})$ spans the coordinate representation of $\mathfrak{X}^*(T_p(S))$, and there exists smooth functions u_F and u_G where

$$\zeta(dF(q)) = u_F(q)dz \quad \text{and} \quad \zeta(dG(q)) = u_G(q)dz \quad \text{locally.}$$

This uniquely defines ρ on a neighbourhood of p (by definition of the abstract tangent space), we can assume ρ is compactly supported by appealing to Urysohn's Lemma for smooth manifolds.

The symplectic form is $C^\infty(M)$-linear, hence $X_F = \rho X_G$ on a precompact neighbourhood of S. Given a point $x \in S$, we see that

$$\varphi_x(s) = X_F(\varphi_x(s)) = \rho(\varphi_x(s)) X_G(\varphi_x(s)).$$

Using eq. (98), we can define a smooth function α (because ρ is smooth). Using the chain rule, and suppressing $\varphi_x(s)$:

$$\frac{\mathrm{d}}{\mathrm{d}s} \theta_x(\alpha(s)) \bigg|_s = \rho X_G = X_F$$

So that $\theta_x \circ \alpha$ is an integral curve of X_F starting at x, and must be equal to φ_x by uniqueness. Next,

- if $\theta_x(\alpha(s))$ is a periodic orbit of X_G, it follows that $\varphi_x(s)$ is a periodic orbit of X_F; and
- because ρ is either strictly positive or negative, $\varphi_x(s)$ is a critical point of X_F iff $\theta_x(\alpha(s))$ is a critical point of X_G.

At last, if $X_F = \rho X_G$ about S, then $X_G = \rho^{-1} X_F$. Let $\beta(x,t) = \int_0^t \rho(x,u)^{-1} du$, and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi_x(\beta(t)) \bigg|_t = \rho^{-1} X_F = X_G,$$

and rehearsing the same argument we had for α completes the proof. ■

Symplectic Action

We return to a more abstract-analytic perspective. Let $(X, \mathcal{M}, \mu)$ be a measure space, suppose $\gamma, \eta : X \to \mathbb{R}^{2n}$ is an L^2 function, in the sense that it is L^2 in each coordinate. Holder's inequality tells us that

$$\langle \gamma, \eta \rangle_{\omega_0} = 2^{-1} \int_X \langle \gamma(x), \eta(x) \rangle_{\omega_0} d\mu(x) \quad \text{converges absolutely.}$$

With this, we can extend the symplectic form ω_0 to L^2 mappings into $\mathbb{R}^{2n}$. The following is a natural function space to consider.

Definition 6.1: Loop Space

We define the space of *loops* as the function space $\Omega = C^\infty(S^1, \mathbb{R}^{2n})$. It is equipped with the *symplectic pairing*, which is denoted by $A : \Omega \times \Omega \to \mathbb{R}$;and $A(\cdot, \cdot)$ is defined by the integral in eq. (99).

$$A(x, y) = \langle \mathring{x}, y \rangle_{\omega_0} = 2^{-1} \int_{S^1} \langle \mathring{x}(t), y(t) \rangle_{\omega_0} dt \quad \forall x, y \in \Omega. \tag{99}$$

Equation (99) can be rewritten explicitly as the sum of half-determinants, which we now give

$$\langle x, y\rangle_{\omega_0} = 2^{-1}\int_{S^1}\sum_{i=\underline{n}}\det\left(\begin{bmatrix}\dot{x}_i & y_i\\ \dot{x}_{n+i} & y_{n+i}\end{bmatrix}\right). \tag{100}$$

A simple application of Holder's inequality will show that, for every $x, y \in \Omega$, we have

$$|\langle x, y\rangle_{\omega_0}| \leq 2^{-1}\sum_{i=\underline{n}}\|y_i\dot{x}_{n+i}\|_{L^2} + \|y_{n+i}\dot{x}_i\|_{L^2}.$$

Definition 6.2: Symplectic Action

The *symplectic action* (on closed curves) is a mapping $A : \Omega \to \mathbb{R}$ which **computes the area swept by the curve**. For an arbitrary loop $\gamma \in \Omega$, its action $A(\gamma)$ is given by eq. (101):

$$A(\gamma) = 2^{-1}\int_{S^1}\langle\mathring{\gamma}, \gamma\rangle_{\omega_0} = 2^{-1}\int_{S^1}\sum i = \underline{n}\det\left(\begin{bmatrix}\mathring{\gamma}^i & \gamma^i\\ \mathring{\gamma}^{n+i} & \gamma^{n+i}\end{bmatrix}\right)dt \tag{101}$$

Lemma 6.1: Symplectic Action in terms of λ

The symplectic action, as defined in eq. (101) on Ω is also given by the line integral over λ.

$$A(\gamma) = \int_\gamma \lambda \tag{102}$$

Proof. We expand **??** to obtain

$$A(\gamma) = 2^{-1}\int_{S^1}\sum_{i=\underline{n}}\gamma^i\mathring{\gamma}^{n+i} - \mathring{\gamma}^i\gamma^{n+i} \tag{103}$$

Notice that the first term $\int_{S^1}\sum_{i=\underline{n}}\gamma^i\mathring{\gamma}^{n+i}$ is equal to $2^{-1}\int_\gamma\lambda$. Indeed,

$$\int_\gamma\lambda = \int_0^1\lambda(\gamma(t))(\mathring{\gamma}(t))dt = \int_{S_1}\sum_{i=\underline{n}}\gamma^i\mathring{\gamma}^{n+i}$$

Using integration by parts, the integral over the each of the second terms in eq. (103) evaluates to

$$2^{-1}\int_{S^1}\sum_{i=\underline{n}}\mathring{\gamma}^i\gamma^{n+i} = 2^{-1}\sum_{i=\underline{n}}\gamma^i\gamma^{n+i}\Big|_{\partial S^1} - 2^{-1}\int_{S^1}\sum_{i=\underline{n}}\mathring{\gamma}^{n+i}\gamma^i$$

The boundary terms disappear since γ is periodic, and we notice that the left hand side of eq. (103) is the sum of $2^{-1}\int_\gamma\lambda + 2^{-1}\int_\gamma\lambda$, and the proof is complete. ∎

Remark 6.1: General loops with period L

More generally, if we have two loops of period L eq. (101) suggests that we can descend ω_0 to an even larger space.

$$\Omega_{[0,L]} = C^\infty(\mathbb{R}/L\mathbb{Z}, \mathbb{R}^{2n}) \tag{104}$$

with $A_{[0,L]}(\gamma,\eta) = A(\gamma(Lt),\eta(Lt))$ which evaluates to

$$A_{[0,L]}(\gamma,\eta) = \frac{1}{2L}\int_0^L \langle \mathring{\gamma}(t),\eta(t)\rangle_{\omega_0} dt \tag{105}$$

Note 6.1: L^2 descent of bilinear forms

The argument in this section about descending symplectic (resp. orthogonal) geometries onto L^2 functions *into* the space is one of the reasons why L^2 functions are of such importance. To recapitulate:

- Given a bilinear form υ on $\mathbb{R}$, we can extend it to $\mathbb{R}^{2n}$ for $n \geq 1$ using a 'hyperbolic decomposition' similar to eq. (94).
- This bilinear form on $\mathbb{R}^{2n}$ descends into a bilinear form on the space of L^2 1-periodic loops from S^1 into $\mathbb{R}^{2n}$,
- Since every loop with period L admits a 1-periodic representation, υ further descends to a bilinear form on $L^2([0,L],\mathbb{R}^{2n})$. Whose action is defined by the integral

$$\langle \gamma,\eta\rangle_\upsilon = \frac{1}{2L}\int_0^L \langle \gamma(t),\eta(t)\rangle_\upsilon dt$$

Proposition 6.1: WC Reduction 2 — Independence of the Symplectic Structure

Suppose (M,ω) and (N,η) are symplectic manifolds modelled on $\mathbb{R}^{2n}$, and $u : M \to N$ is a symplectomorphism.

> To every function $F \in C^\infty(N)$, the vector field pullback of the Hamiltonian flow of F is equal to the Hamiltonian flow of its pullback through u.

More precisely, if $X_F = \eta^\wedge(dF)$ and $u^*F = F \circ u$, we claim that

$$\omega^\wedge(d(u^*F)) = u^*(\eta^\wedge(dF)) \quad \text{where} \quad u^*(\eta^\wedge(dF)) = du^{-1} \circ X_F \circ u. \tag{106}$$

If γ is an integral curve of $X_{F\circ u}$, then $u \circ \gamma$ is an integral curve of X_F, and if $\varphi(s,x)$ and $\theta(t,y)$ denote the flows of $X_{F\circ u}$ and X_F, they relate to each other by u-conjugation as in eq. (107)

$$u \circ \varphi^t = \theta^t \circ u. \tag{107}$$

Proof. Let F be fixed, and write $A = dF$. Recall $d(F\circ u) = dF \circ du$. It suffices to show that

$$\omega^\wedge(u^*A) = u^*(\eta^\wedge A). \tag{108}$$

We will show the left and right hand sides are equal at every tangent space. Given $p \in M$, we write

$$X_p = \omega^\wedge(u^*A)(p) \quad \text{and} \quad Y_p = u^*(\eta^\wedge A)(p).$$

If $Z_p \in T_pM$ is arbitrary, we compute $\omega(p)(X_p,Z_p)$ and $\omega(p)(Y_p,Z_p)$ and the proof is complete upon showing equality. Now,

$$\omega(p)\Big(X_p,Z_p\Big) = \eta(u(p))\Big(du(p)(X_p),\, du(p)(Z_p)\Big) = A(u(p))\Big(du(p)(Z_p)\Big).$$

Using the same technique of exchanging $\omega(p)$ with $\eta(u(p))$, we get

$$\omega(p)\Big(Y_p, Z_p\Big) = \eta(u(p))\Big(du(p)(Y_p),\, du(p)(Z_p)\Big),$$

since $Y_p = (du^{-1} \circ \eta^\wedge A \circ u)(p)$, we obtain

$$\omega(p)\Big(Y_p, Z_p\Big) = \eta(u(p))\Big(\eta^\wedge A(u(p)),\, du(p)(Z_p)\Big),$$

which implies $X_p = Y_p$. This proves the first claim, and

$$du \circ X_{F\circ u} = X_F \circ du.$$

Next, if γ is an integral curve of $X_{F\circ u}$, then $\mathring{\gamma}(t) = X_{F\circ u}(\gamma(t))$ implies

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} u \circ \gamma(t)\Big|_t &= du(\gamma(t))(X_{F\circ u}) = du \circ X_{F\circ u}\Big|_{\gamma(t)} \\ &= (X_F \circ u)\Big|_{\gamma(t)} = X_F\Big|_{u\circ\gamma(t)}\end{aligned}$$

and $u \circ \gamma$ is an integral curve of X_F. Equation (108) is proven upon realizing that X_F and $X_{F\circ u}$ are u-related (Theorem 9.13 of [4]). ■

Corollary 6.1

Let M, N and u be as in Proposition 6.1, the orbits of X_F correspond bijectively to the orbits of $X_{F\circ u}$ through conjugation; and the same for the non-degenerate orbits.

Quadratic Forms

Let $n \geq 1$ be fixed, a *quadratic form* on $\mathbb{R}^n$ is a mapping $q : \mathbb{R}^n \to \mathbb{R}$ where $q(rx) = |r|^2 q(x)$ for every $r \in \mathbb{R}$ and $x \in \mathbb{R}^n$ and

$$\langle x, y\rangle_q = q(x+y) - [q(x) + q(y)] \quad \text{is a symmetric bilinear form.}$$

If q is a quadratic form, it is *positive definite* whenever $\langle\cdot,\cdot\rangle_q$ is. We denote the set of positive definite quadratic forms by $\mathbb{P}$. The following proposition shows that every $q \in \mathbb{P}$ can be symplectically diagonalized on $\mathbb{R}^{2n}$.

Proposition 7.1: Symplectic diagonalization of positive definite quadratic forms

If $q \in \mathbb{P}$, there exists a linear mapping $\varphi \in \mathrm{Sp}(n)$ such that $q \circ \varphi$ takes on the form:

$$q \circ \varphi(x) = \sum_{i=\underline{n}} \frac{x_i^2 + x_{n+i}^2}{r_i^2} \quad \text{where} \quad 0 < r_1 \leq r_2 \leq \cdots \leq r_n. \tag{109}$$

We call eq. (109) the *normal form* of q.

Proof. Postponed for now. ■

Definition 7.1: Associated open ellipsoid

Let q be a positive definite quadratic form on $\mathbb{R}^{2n}$; its *asscoiated open ellipsoid* is the subset $\mathcal{E}_q = [q < 1]$. If q is given in normal coordinates,

$$\mathcal{E}_q = \left\{ x \in \mathbb{R}^{2n},\ \sum_{i=\underline{n}} r_j^{-2}(x_j^2 + x_{n+j}^2) < 1 \right\}, \quad \text{and} \quad \partial\mathcal{E}_q = \{x \in \mathbb{R}^{2n},\ q(x) = 1\}.$$

Orbits on Ellipsoids

We see that reparametrization by a diffeomorphism α does not affect the **magnitude** of the symplectic action. Indeed, if γ is a closed characteristic, and α a reparametrization, then

$$A(\gamma) = \int_\gamma \lambda = \pm \int_{\gamma\circ\alpha} \lambda = \pm A(\gamma \circ \alpha).$$

In the previous chapter, we have also proven that symplectomorphisms on $\mathbb{R}^{2n}$ leave the action invariant; so it makes sense to speak of the action across closed characteristics on an energy surface (on $\mathbb{R}^{2n}$). Our main result in this section concerns the **non-constant, periodic solutions** of Hamiltonian flows on the boundaries of ellipsoids.

Let X be a vector field on a manifold M.

- A *solution* to X is a mapping $\gamma : \mathcal{I} \to M$ where $\mathring{\gamma}(t) = X(\gamma(t)$ at every t in the open interval $\mathcal{I}$.
- An *orbit* of X is a non-constant solution.
- A solution γ, to X is said to be of p-type, or a p-solution whenever it is periodic; and respectively for orbits of X.

Let (M, ω) be a symplectic manifold, and S an energy surface.

- A *smooth defining function* of S is a function $F \in C^\infty(M, \mathbb{R})$ such that S is a regular level set of F. We sometimes say F is a *defining function* of S when it is implicit.
- A *characteristic* of S is the image $\gamma(\mathcal{I})$ where γ is a solution of the Hamiltonian flow of a defining function of S.
- Given a characteristic of S, $\gamma(\mathcal{I})$. It is *closed* whenever γ is of p-type, and is *non-degenerate* whenever γ is an orbit.

We will prove Proposition 8.1 in a few steps.

Proposition 8.1: Action on the boundary of ellipsoids

Let q be a positive definite quadratic form, (not necessarily normal with respect to standard coordinates), then:

$$\pi r_1^2 = \inf\left\{ |A(\gamma)|,\ \gamma \text{ is a non-degenerate closed characteristic of } \partial\mathcal{E}_q. \right\},$$

and the infimum is attained.

Step 1: Orbits of X_q on $\partial\mathcal{E}_q$

Suppose $q = \sum_{i=\underline{n}}(x_i^2 + x_{n+i}^2)$ is in normal form. Every non-constant solution γ on $\partial\mathcal{E}_q$ is given by

$$\gamma(t) = \sum_{i=\underline{n}}[c_{\lambda_i}\gamma_i(0) + s_{\lambda_i}\gamma_{n+i}(0)]e_i + [-s_{\lambda_i}\gamma_i(0) + c_{\lambda_i}\gamma_{n+i}(0)]e_{n+i},$$

or equivalently: $\gamma(t) = \sum_{i=\underline{n}} \exp\big(\lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t\big) \left[\begin{smallmatrix} \gamma_i(0) \\ \gamma_{n+i}(0) \end{smallmatrix}\right]$ with $\left[\begin{smallmatrix} e_i & e_{n+i} \end{smallmatrix}\right]$ suppressed.

Proof of Step 1. Making the substitution, $\lambda_i = 2{r_i}^{-2}$, we can rewrite eq. (109), which is convenient when we compute the gradient of $q \circ \varphi$.

$$q \circ \varphi(x) = 2^{-1} \sum_{i=\underline{n}} \lambda_i (x_i^2 + x_{n+i}^2) \quad \text{where} \quad 0 < \lambda_n \le \lambda_{n-1} \le \cdots \le \lambda_1. \tag{110}$$

Suppose q is given by the right hand side of eq. (110), then $\nabla q(x) = (\lambda_1 x_1, \dots, \lambda_n x_n, \lambda_1 x_{n+1}, \dots, \lambda_n x_{2n})$; or

$$\nabla q(x) = \sum_{i=\underline{n}} \lambda_i (x_i e_i + x_{n+i} e_{n+i}) \quad \text{and} \quad X_q(x) = \sum_{i=\underline{n}} \lambda_i (x_{n+i} e_i - x_i e_{n+i})$$

Before proceeding any further, we will need to work out some computations involving symplectic bases (see [6]).

Note 8.1: Hyperbolic decompositions

Given a vector $x = (x_1, \dots, x_n, x_{n+1}, \dots, x_{2n})$, we can write x in terms of the symplectic bases:

$$x = \sum_{i=\underline{n}} x_i e_i + x_{n+i} e_{n+i} = \sum_{i=\underline{n}} \begin{bmatrix} e_i & e_{n+i} \end{bmatrix} \begin{bmatrix} x_i \\ x_{n+i} \end{bmatrix}.$$

Where J acts on x in a convenient manner:

$$Jx = \sum_{i=\underline{n}} \begin{bmatrix} e_i & e_{n+i} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ x_{n+i} \end{bmatrix} = \sum_{i=\underline{n}} \begin{bmatrix} e_i & e_{n+i} \end{bmatrix} \begin{bmatrix} x_{n+i} \\ -x_i \end{bmatrix}.$$

Now, suppose we are given a Hamiltonian $F \in C^\infty(\mathbb{R}^{2n}, \mathbb{R})$, its Hamiltonian flow given by

$$X_F = J\nabla F = \sum_{i=\underline{n}} \begin{bmatrix} e_i & e_{n+i} \end{bmatrix} \begin{bmatrix} \partial_{n+i} F \\ (-1)\partial_i F \end{bmatrix}.$$

If $\gamma(t) = (\gamma_1, \dots, \gamma_{2n})$ is a solution to X_q it must satisfy $\mathring{\gamma} = X_q(\gamma)$. Suppressing $\left[\begin{smallmatrix} e_i & e_{n+i} \end{smallmatrix}\right]$ within the summation, it follows that

$$\sum \mathring{\gamma}_i e_i + \mathring{\gamma}_{n+i} e_{n+i} = \sum \lambda_i \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} \gamma_i \\ \gamma_{n+i} \end{smallmatrix}\right] \quad \text{where } i = \underline{n}$$

Comparing coefficients, we see that for $j = \underline{n}$, and $z_j(t) = (\gamma_j, \gamma_{n+j}) \in \mathbb{R}^2$:

$$\mathring{z}_j = \lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] z_j \quad \text{implies} \quad z_j(t) = \exp\big(\lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t\big) z_j(0).$$

Each z_j has eigenvalues $\pm i\lambda_j$, where $i = \sqrt{-1}$ in this context. Computing $\exp\big(\lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t\big)$ with the eigenvalues gives us

$$\exp\big(\lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t\big) = \begin{bmatrix} \cos(\lambda_j t) & \sin(\lambda_j t) \\ -\sin(\lambda_j t) & \cos(\lambda_j t) \end{bmatrix}, \quad \text{which has period } 2\pi\lambda_j^{-1} = \pi r_j^2.$$

Note 8.2: Matrix exponentials

Write $J_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, and $J = J_{2n}$ is still the standard symplectic form on $\mathbb{R}^{2n}$, then

$$\exp(\lambda J_2 t) = \begin{bmatrix} \cos(\lambda t) & \sin(\lambda t) \\ -\sin(\lambda t) & \cos(\lambda t) \end{bmatrix} = \cos(\lambda t)\,\mathrm{id}_{\mathbb{R}^2} + \sin(\lambda t) J_2; \tag{111}$$

and for the general case with the substitution $c_\lambda(t) = \cos(\lambda t)$, (resp. $s_\lambda(t)$):

$$\exp(\lambda J t) = \begin{bmatrix} c_\lambda\,\mathrm{id}_{\mathbb{R}^n} & s_\lambda\,\mathrm{id}_{\mathbb{R}^n} \\ -s_\lambda\,\mathrm{id}_{\mathbb{R}^n} & c_\lambda\,\mathrm{id}_{\mathbb{R}^n} \end{bmatrix} = \cos(\lambda t)\,\mathrm{id}_{\mathbb{R}^{2n}} + \sin(\lambda t) J. \tag{112}$$

Suppose $\gamma(0) = (\gamma_{\underline{2n}}(0)) \in \partial\mathcal{E}_q$, then the integral curve generated by $\gamma(0)$ is the sum of the eigenmodes $\pm\lambda_i$:

$$\gamma(t) = \sum_{i=\underline{n}} \exp\big(\lambda_j \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t\big) \left[\begin{smallmatrix} \gamma_i(0) \\ \gamma_{n+i}(0) \end{smallmatrix}\right], \quad \text{with } \left[\begin{smallmatrix} e_i & e_{n+i} \end{smallmatrix}\right] \text{ suppressed.}$$

We note that: because $0 < r_1 \leq r_2 \leq \cdots \leq r_n$, the minimum period of a non-degenerate orbit is πr_1^2. ■

Step 2: Action of solutions of X_q on $\partial\mathcal{E}_q$

Let $q = \sum_{i=\underline{n}} r_i^{-2}(x_i^2 + x_{n+i}^2)$ be a positive definite ellipsoid in normal form on $(\mathbb{R}^{2n}, \omega_0)$, then

$$\pi r_1^2 = \inf\Big\{ |A(\gamma)|,\ \gamma \text{ is a non-constant, periodic integral curve of } X_q \text{ on } \partial\mathcal{E}_q. \Big\}.$$

Proof of Step 2. Suppose q is in normal form, and let $\gamma(t)$ be such a curve on $\partial\mathcal{E}_q$, and define

- $z_i = \gamma_i(0)$ for $i = \underline{2n}$,
- $\vec{z}_i = (z_i, z_{n+i})$ for $i = \underline{n}$,
- $J_2 = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$.

Notice that $\frac{\mathrm{d}}{\mathrm{d}t} e^{\lambda_i J_2 t} = \lambda_i J_2 e^{\lambda_i J_2 t}$, and from Step 1, we compute the integrand of $A(\gamma)$:

$$\begin{aligned} 2^{-1}\langle \mathring{\gamma}(t), \gamma(t)\rangle_{\omega_0} &= 2^{-1} \sum \Big\langle \lambda_i J_2 e^{\lambda_i \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t} \vec{z}_i, \quad J_2 e^{\lambda_i \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] t} \vec{z}_i \Big\rangle_{\mathbb{R}^{2n}} \\ &= 2^{-1} \sum \lambda_i |\vec{z}_i|^2 = q(\gamma(t)) = 1 \end{aligned}$$

Hence, $A(\gamma) = L$ where L is the period of γ, and is minimized whenever $L = \pi r_1^2$. For an arbitrary q, there exists a linear symplectic mapping φ such that $\varphi \circ q$ is in normal form. Since φ preserves the symplectic action, this proves Step 2. ■

Step 3: Action of arbitrary Hamiltonian flows on $\partial\mathcal{E}_q$

Let $q \in \mathbb{P}$, and F define the boundary of $\mathcal{E}_q$, then

$$\pi r_1^2 = \inf\left\{|A(\gamma)|,\ \gamma \text{ is a non-constant, periodic solution to } X_F \text{ on } \partial\mathcal{E}_q.\right\}.$$

Proof of Step 3. From Proposition 5.1, we know that the orbits of X_q and X_F on $\partial\mathcal{E}_q$ correspond one-to-one with each other. Given an orbit, γ of X_F, the composition $\gamma \circ \alpha$ is an orbit of X_q for some diffeomorphism α. Therefore, $A(\gamma) = A(\gamma \circ \alpha) = \pm A(\gamma)$ and the proof is complete. ■

Chapter 5: Symplectic Capacities

Introduction

We will discuss a class of correspondences from the set of all symplectic manifolds to $[0,+\infty]$ — similar to the total measure or volume of a space — but are preserved under symplectomorphisms. First, we recall two very special symplectic manifolds, the *open r-ball* and the *open r-cylinder*:

$$B(r)=\left\{x\in\mathbb{R}^{2n},\ \sum_{i=\underline{n}}x_i^2+x_{n+i}^2=|x|^2<r^2\right\}\quad\text{and}\quad Z(r)=\left\{x\in\mathbb{R}^{2n},\ x_1^2+x_{n+1}^2<r^2\right\};$$

which are both equipped with the standard symplectic form ω_0.

Definition 1.1: Symplectic capacity

A *symplectic capacity* $\mathfrak{C}$ is a function that assigns to each symplectic manifold (M,ω): a number $\mathfrak{C}(M,\omega)\in[0,+\infty]$ satisfying the following properties

1. Monotonicity: Given two symplectic manifolds (M,ω) and (N,η) **of the same dimension**, if (M,ω) embeds symplectically into (N,η), then $\mathfrak{C}(M,\omega)\leq\mathfrak{C}(N,\eta)$.
2. Conformality: If $\alpha\neq 0$ is a real number, then $\mathfrak{C}(M,\alpha\omega)=|\alpha|\mathfrak{C}(M,\omega)$.
3. Non-triviality: The capacities of $B(1)$ and $Z(1)$ are equal to π, **across all n**.

Proposition 1.1: Gromov's Width

If (M,ω) is a symplectic manifold modelled on $\mathbb{R}^{2n}$, its *Gromov's width* is the number

$$\operatorname{Gromov}(M,\omega)=\sup\left\{\pi r^2,\ \begin{array}{l}\text{There exists a symplectic embed-}\\ \text{ding of }(B(r),\omega_0)\hookrightarrow(M,\omega).\end{array}\right\}.$$

Gromov's width is a symplectic capacity, and it is minimal:

$$\mathfrak{C}(M,\omega)\geq\operatorname{Gromov}(M,\omega)\quad\text{for every symplectic manifold}(M,\omega).$$

Proposition 1.2: Darboux's Theorem

Let (M,ω) be a symplectic manifold modelled on $\mathbb{R}^{2n}$. At every point $p\in M$, there exists a chart $\varphi:U\to\hat{U}$ where its **inverse** satisfies

$$(\varphi^{-1})^*\omega=\omega_0.$$

Corollary 1.1: Gromov's width is non-negative

For every symplectic manifold (M,ω), the set $\{\pi r^2,\ B(r)\hookrightarrow(M,\omega)$ symplectically.$\}$ is non-empty.

Chapter 1: Sobolev spaces

Introduction

In this chapter, we will introduce a special subset of tempered distributions on $\mathbb{R}^n$ called the *Sobolev spaces*, denoted by H_s for $s \in \mathbb{R}$. What is nice about H_s is that the are Hilbert spaces, and the *periodic Sobolev spaces* over $\mathbb{R}$ are used in the proof of the *non-triviality of the special symplectic capacity* c_0.

Lemma 1.1: Slowly increasing lemma

The space of slowly increasing function is closed under pointwise multiplication. If $f_{\underline{k}} \in C_s^\infty$, then $m(f_{\underline{k}}) \in C_s^\infty$ as well.

Moreoever, the multiplication map by C_s^∞ on $\mathcal{S}$ is toplinear. That is, for any $g \in C_s^\infty$ the map

$$m_g : \mathcal{S} \to \mathcal{S} \quad m_g(\phi) = g\phi \quad \text{is an endomorphism on } \mathcal{S}$$

Proof. Fix a non-negative integer p and a multi-index β with $|\beta| = p$. We remark that for any $g \in C^\infty$,

$$|D^{e_j} g(x)| \le |Dg(x)||e_j| = |Dg(x)|$$

where $Dg(x)$ is the Frechet derivative of g evaluated at x. A simple induction on the order of the multi-index will show that

$$|D^\beta g(x)| \lesssim |D^p g(x)|$$

Back to the main proof, since m is k-linear we have

$$|D^p m(f_{\underline{k}})| = \left| \sum_{|\alpha|=p} \left(\frac{p!}{\alpha!}\right) m(D^{\alpha_{i=\underline{k}}} f_i(x)) \right| \lesssim_{p,m} \sum_{|\alpha|=p} \text{(II)}(D^{\alpha_{i=\underline{k}}} f_i(x))$$

$$\lesssim_{p,m,f_{\underline{k}}} (1+|x|)^N$$

where $\text{(II)}(x_{\underline{k}}) = \text{(II)}(|x_{\underline{k}}|)$ for readability. Therefore $D^\beta m(f_{\underline{k}})$ is slowly increasing as well.

Next, we show m_g maps into $\mathcal{S}$. Indeed, fix $\phi \in \mathcal{S}$ and a multi-index α, and $N \in \mathbb{N}^+$. we can obtain $M_\alpha \in \mathbb{N}^+$, such that

$$|\partial^\beta g(x)| \lesssim_{g,\alpha} (1+|x|)^{M_\alpha} \quad \forall |\beta| \le |\alpha|$$

It follows from the product rule that

$$\partial^\alpha (g\phi) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} (\partial^\beta g)(\partial^\gamma \phi) \implies |\partial^\alpha (g\phi)| \lesssim_{\alpha,g} \sum^{\wedge} |\partial^\gamma \phi| (1+|x|)^{M_\alpha}$$

Multiplying by $(1+|x|)^N$ on both sides of the estimate yields

$$(1+|x|)^N |\partial^\alpha (g\phi)| \lesssim_{\alpha,g} \sum^{\wedge} |\partial^\gamma \phi|(1+|x|)^{M_\alpha+N} \lesssim_{\alpha,g} \sum^{\wedge} \|\phi\|_{(N+M_\alpha,\gamma)} < +\infty$$

The continuity of the multiplication map follows immediately from the definition of continuity between TVS, as the right hand side of the last estimate is a finite sum of seminorms in $\mathcal{S}$.

∎

Lemma 1.2: Multiplication by Sobolev factor is an automorphism

The map $A_s : \mathbb{R}^n \to \mathbb{R}$ defined by $A_s(x) = (1 + |x|^2)^{s/2}$ is slowly increasing for all $s \in \mathbb{R}$. Moreoever, $\Lambda_s : f \mapsto [(1 + |\zeta|^2)^{s/2}\hat{f}]^\vee$ is an automorphism on $\mathcal{S}$ and $\mathcal{S}'$, with inverse $\Lambda_s^{-1} = \Lambda_{-s}$.

Proof. We postpone the proof for $A_s \in C_s^\infty$ for now, but proving the rest of the Lemma. The Fourier Transform on $\mathcal{S}'$ is an on $\mathcal{S}'$, hence the first and last step below are continuous.

$$f \mapsto \hat{f} \mapsto (1 + |\zeta|^2)^{s/2}\hat{f} \mapsto \Lambda_s f$$

Since $A_s(\zeta) = (1 + |\zeta|^2)^{s/2} \in C_s^\infty$, it suffices to show that multiplication of a tempered distribution by an arbitrary slowly increasing function $g \in C_s^\infty$ is continuous. Fix a sequence $F_n \to F$ in $\mathcal{S}'$, and Lemma 1.1 tells us that

$$\langle gF_n, \phi\rangle = \langle F_n, g\phi\rangle \longrightarrow \langle F, g\phi\rangle = \langle gF, \phi\rangle$$

and Λ_s is toplinear. Observe that the action of the the distribution $\Lambda_s F$ on $\mathcal{S}$ is **defined by precomposing the duality pairing** by $\Lambda_s : \mathcal{S} \to \mathcal{S}$. In symbols, $\langle \Lambda_s F, \phi\rangle = \langle F, \Lambda_s \phi\rangle$. The inverse of Λ_s on $\mathcal{S}$ is Λ_{-s}, this is because

$$\begin{aligned}\Lambda_{-s}\Lambda_s\phi = \Lambda_{-s}\Big(((1 + |\zeta|^2)^{s/2}\hat{\phi})^\vee\Big) &= \Big[(1 + |\zeta|^2)^{-s/2}\Big((1 + |\zeta|^2)^{s/2}\hat{\phi}\Big)\Big]^\vee \\ &= \phi^{\wedge\vee} = \phi\end{aligned}$$

Reversing the roles of $-s$ and s proves $\Lambda_s^{-1} = \Lambda_{-s}$. This however implies $\langle \Lambda_{-s}\Lambda_s F, \phi\rangle = \langle F, \Lambda_{-s}\Lambda_s\phi\rangle$ and the proof is complete. ■

Note 1.1: Estimates on multi-index powers

Lemma 1.3: Multi-index power bounded by polynomial factor

For any $N \in \mathbb{N}^+$, if β is a multi-index with $|\beta| \le N$, then

$$|x^\beta| \le (1 + |x|)^N \tag{113}$$

Proof. Expanding the left and right hand sides of eq. (113), we have

$$\Big|\prod_{i\le n} x_i^{\beta_i}\Big| = \prod_{i \le n} |x_i|^{\beta_i} \quad \text{and} \quad \left(1 + \left(\sum_{i\le n} |x_i|^2\right)^{1/2}\right)^N \tag{114}$$

The estimate in eq. (114) clearly holds for $N = 0$. Suppose it holds for $N - 1$, if β is a multi-index with order $|\beta| \le N - 1$, then

$$|x^\beta| \le (1 + |x|)^{N-1} \le (1 + |x|)^N$$

Assume $|\beta| = N - 1$, and we multiply by any coordinate $|x_i|$, to the left and right hand sides of the equation, and see that

$$|x_i||x^\beta| \le |x_i|(1 + |x|)^{N-1} \le (1 + |x|)^N$$

■

Lemma 1.4: Multi-index power bounded by Sobolev factor

If α is any multi-index, then

$$|x^\alpha| \leq (1+|x|^2)^{|\alpha|/2} \tag{115}$$

Proof. We will use induction on the order of α. If $|\alpha| = 0$, both sides are equal to 1. Assume it holds for $|\alpha| = 0, 1, \dots, N-1$, and for any multi-index β with order $|\beta| = N$, there exists a (not necessarily unique) multi-index α of order $|\alpha| = N-1$, such that $\beta - \alpha = e_j$. By induction hypothesis, we see that

$$|x^\alpha| = \prod_{j \leq n} |x_j^{\alpha_j}| \leq (1+|x|^2)^{|\alpha|/2}$$

Multiplying by the coordinate function $|x_j|$, similar to the previous proof.

$$|x^\beta| = |x_j||x^\alpha| \leq |x_j|(1+|x|^2)^{|\alpha|/2}$$

Since the "L^1 cosine" is less than 1, meaning $|x_j|(1+|x|^2)^{-1/2} \leq 1$, and this implies $|x^\beta| \leq (1+|x|^2)^{|\beta|/2}$. ■

Remark 1.1

To show the L^1 cosine inequality, the projection map $x \mapsto x_j$ is norm-decreasing; as $\mathbb{R}^n$ is a Hilbert space with the standard inner product. And $|x|^2 \leq 1+|x|^2$.

L^2 Sobolev Spaces

Definition 2.1: L^2 Sobolev Space

If $s \in \mathbb{R}$, the *Sobolev space* H_s is the subspace of tempered distributions

$$H_s = \left\{ f \in \mathcal{S}',\ \Lambda_s f \in L^2(dx) \right\}$$

The claim $\Lambda_s f \in L^2$ should be interpreted with respect to the ambient topology of $\mathcal{S}'$. *There exists a L^2 function, which we will also denote by $\Lambda_s f$ that realizes the duality pairing $\langle \Lambda_s f, \cdot \rangle$*, if $\phi \in \mathcal{S}$ is arbitrary, then

$$\langle \Lambda_s f, \phi \rangle_{\mathcal{S}} = \langle f, \Lambda_s \phi \rangle_{\mathcal{S}} = \langle \Lambda_s f, (\iota\phi) \rangle_{L^2}$$

where $\iota : \mathcal{S} \to L^2$ is the inclusion map.

We call $\Lambda_s f$ the L^2 *representative of* f. We sometimes omit the inclusion map when it leaves no room for ambiguity. A moment's thought will show that the L^2 representative of f is unique a.e., as elements of the reflexive L^p spaces are completely determined by their duality pairing with elements in its conjugate space L^q where $p^{-1} + q^{-1} = 1$.

Lemma 2.1

If $s \in \mathbb{R}$, the mapping $\Lambda_s : H_s \to L^2(\mathbb{R}, dx)$ is toplinear and injective.

Proof. It is clear the zero distribution is represented by the zero function in L^2. so $\Lambda_s 0 = 0$. Next, let $f, g \in H_s$, and $\Lambda_s f$ and $\Lambda_s g$ be their L^2 representatives. Fix $\phi \in \mathcal{S}$,

$$\begin{aligned}\langle \Lambda_s(f+g), \phi \rangle_{L^2} &= \langle f+g, \Lambda_s \phi \rangle_{\mathcal{S}} = \langle f, \Lambda_s \phi \rangle_{\mathcal{S}} + \langle g, \Lambda_s \phi \rangle_{\mathcal{S}} \\ &= \langle \Lambda_s f, \phi \rangle_{L^2} + \langle \Lambda_s g, \phi \rangle_{L^2} = \langle \Lambda_s f + \Lambda_s g, \phi \rangle_{L^2}\end{aligned}$$

Homogeneity is proven in a similar manner. Suppose $f \in H_s$ and $\Lambda_s f = 0$ in $L^2 \cong (L^2)^*$. Let ϕ be any Schwartz function, precomposing the L^2 pairing with the inclusion map $\iota : \mathcal{S} \to L^2$ yields

$$\langle \Lambda_s f, \phi \rangle_{\mathcal{S}} = \langle \Lambda_s f, \iota\phi \rangle_{L^2} = 0$$

Therefore $\Lambda_s f = 0$ in $\mathcal{S}'$ as well. But Λ_s is an automorphism on the space of Tempered Distributions, so that

$$\Lambda_s f = 0 \iff f \in \operatorname{Ker}(\Lambda_s) \iff f = 0$$

and the proof is complete. ■

We can borrow the algebraic and topological structure from L^2 by pulling back the inner product, lem. 2.1 tells us the L^2 representative is injective, so eq. (116) defines an inner product.

$$\langle f, g \rangle_{(s)} = \langle \Lambda_s f, \Lambda_s g \rangle_{L^2(dx)} = \langle (\Lambda_s f)^\wedge, (\Lambda_s g)^\wedge \rangle_{L^2(d\zeta)} = \int_{\mathbb{R}^n} \hat{f}(\zeta)\overline{\hat{g}(\zeta)}(1+|\zeta|^2)^s d\zeta \tag{116}$$

Lemma 2.2

If $s \in \mathbb{R}$, then the Fourier Transform $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is a unitary isomorphism between H_s and $L^2(\mu_s)$, where

$$d\mu_s(\zeta) = (1+|\zeta|^2)^s d\zeta$$

Proof. If $F \in H_s$, then $(\Lambda_s \hat{F})^\wedge = A_s \hat{F}$ is in $L^2(d\zeta)$. This defines $A_s \hat{F}$ as a pointwise function, and because A_s is slowly decreasing with multiplicative inverse A_{-s}, $\hat{F}$ is a pointwise function as well, and $g = \hat{F} \in L^2(\mu_s)$ as needed.

Conversely, given some $g \in L^2(\mu_s)$ we simply set $F = \breve{g}$ and bijectivity follows. We can absorb the factor of $(1+|\zeta|^2)^s$ into the measure within the right hand side of eq. (116). This gives

$$\langle f, g \rangle_{(s)} = \langle \hat{f}, \hat{g} \rangle_{L^2(\mu_s)} = \int_{\mathbb{R}^n} \hat{f}(\zeta)\overline{\hat{g}(\zeta)} d\mu_s \tag{117}$$

For completeness, we compute the norm on H_s in eq. (118).

$$\|f\|_{(s)} = \|\hat{f}\|_{L^2(\mu_s)} = \sqrt{\int_{\mathbb{R}^n} |\hat{f}(\zeta)|^2 d\mu_s} \tag{118}$$

■

Lemma 2.3

The space of rapidly decreasing smooth functions, $\mathcal{S}$ is toplinearly embedded in L^p for usual p and $p = +\infty$. Furthermore, $(L^p)^*$ embeds into $\mathcal{S}'$ toplinearly — where $\mathcal{S}'$ and $(L^p)^*$ are endowed with the $\sigma(\mathcal{S}', \mathcal{S})$ and $\sigma((L^p)^*, L^p)$ topologies.

Proof. We prove the continuity of the inclusion map $\iota : \mathcal{S} \to L^p$ for usual p and $p = +\infty$. If $p = +\infty$, then $\|f\|_\infty = \|f\|_{(0,0)}$ and continuity follows. On the other hand, if p is usual, we let $N = n + 1$ where n is the dimension of the domain. We see that

$$(1 + |x|)^0 \leq (1 + |x|)^{N(p-1)} \quad \text{implies} \quad (1 + |x|)^{-Np} \leq (1 + |x|)^{-N}$$

Hence,

$$\begin{aligned} \|f\|_p^p = \int_{\mathbb{R}^n} |f|^p (1 + |x|)^{Np} (1 + |x|)^{-Np} &\leq \int_{\mathbb{R}^n} \Big(|f| (1 + |x|)^N \Big)^p (1 + |x|)^{-Np} dx \\ &\leq \int_{\mathbb{R}^n} \Big(|f| (1 + |x|)^N \Big)^p (1 + |x|)^{-N} dx \leq \|f\|_{(N,0)}^p \int_{\mathbb{R}^n} (1 + |x|)^{-N} dx \\ &\leq \|f\|_{(N,0)}^p \Big\| (1 + |x|)^{-N} \Big\|_1 \end{aligned}$$

Taking the p-th root, we see that $\|f\|_p \leq \|f\|_{(N,0)} \|(1 + |x|)^{-N}\|_1^{1/p}$, and the inclusion map is a toplinear embedding.

Let $\iota : \mathcal{S} \to L^p$ be the toplinear embedding, and ι^* be the adjoint map of $(L^p)^*$ by precomposing any functional $f \in (L^p)^*$ by ι. Such that for every $\phi \in \mathcal{S}$,

$$(\iota^* f)(\phi) = f(\iota \phi) \quad \forall f \in (L^p)^*$$

The composition of continuous maps is again continuous, so ι^* maps into the tempered distributions. To show continuity, fix a convergent sequence $f_n \to f$ in $\sigma((L^p)^*, L^p)$ and $\phi \in \mathcal{S}$,

$$|\iota^* f_n(\phi) - \iota^* f(\phi)| = |f_n(\iota\phi) - f(\iota\phi)| \to 0$$

therefore $\iota^* f_n \to \iota^* f$ in $\sigma(\mathcal{S}', \mathcal{S})$. So ι^* is a toplinear embedding. ■

Proposition 2.1: Key results of H_s over $\mathbb{C}$

If $s \in \mathbb{R}$, the following are true.

- The space of Schwartz functions $\mathcal{S}$ is dense in H_s
- If $t < s$, then H_s is densely embedded in H_t. In particular, *the spaces decrease, and the norms increase.* That is,
$$t < s \implies H_s \subseteq H_t \quad \text{and} \quad \|\cdot\|_{(t)} \leq \|\cdot\|_{(s)}$$
- Λ_t is a unitary isomorphism from H_s to H_{s-t}, here we view Λ_t as a map between Sobolev spaces; whose topology differs from the weak-$*$ topology of $\mathcal{S}'$,
- $H_0 = L^2$, and $\|\cdot\|_{(0)} = \|\cdot\|_{L^2}$

- $\partial^\alpha : \mathcal{S}' \to \mathcal{S}'$ is a bounded linear map from H_s to $H_{s-|\alpha|}$, for multi-indices α.

Proof of first claim. It suffices to show $\mathcal{S}$ is dense in $L^2(\mu_s)$. Let $f \in L^2(\mu_s)$, the computation below shows that $(1+|\zeta|^2)^{s/2} f \in L^2(d\zeta)$

$$\int_{\mathbb{R}^n} |f|^2(1+|\zeta|^2)^s d\zeta = \|f\|^2_{L^2(\mu_s)} = \left\|(1+|\zeta|^2)^{s/2} f\right\|_{L^2(d\zeta)} < +\infty$$

Furthermore, if $g \in \mathcal{S}$ lem. 1.1 tells us that $h = (1+|\zeta|^2)^{-s/2} g$ is in $\mathcal{S}$ as well. Using the density of $\mathcal{S}$ in $L^2(d\zeta)$, if $\varepsilon > 0$ we obtain a $g \in \mathcal{S}$ with $\int |(1+|\zeta|^2)^{s/2} f - g|^2 d\zeta < \varepsilon^2$

$$\int |(1+|\zeta|^2)^{s/2} f - g|^2 d\zeta = \int (1+|\zeta|^2)^s |f-h|^2 d\zeta = \|f-h\|^{\,2}_{L^2(\mu_s)} < \varepsilon^2$$

This proves the first claim. ■

Proof of the second claim. Given $t < s$ then $(1+|\zeta|^2)^t \leq (1+|\zeta|^2)^s$ pointwise, and eq. (118) gives

$$\|f\|_{(t)} = \|f\|_{L^2(\mu_t)} \leq \|f\|_{L^2(\mu_s)} = \|f\|_{(s)} \tag{119}$$

Hence H_s is a subspace of H_t. But $\mathcal{S}$ is dense in both H_s and H_t, so that H_s is dense in H_t. Finally, eq. (119) tell us that the inclusion map $\iota : H_s \to H_t$ is a toplinear embedding. ■

Proof of the rest. Next, we show the multiplication map on the space of tempered distributions, $\Lambda_t : \mathcal{S}' \to \mathcal{S}'$, with $f \mapsto \Lambda_t f \in \mathcal{S}'$ is a unitary isomorphism from $H_s \to H_{s-t}$, and $\Lambda_t^{-1} = \Lambda_{-t}$. Fix f, g in H_s, noting that $\Lambda_{s-t}\Lambda_t f = \Lambda_s f$ (verify), using the Hilbert space structure on H_s and H_t instead of the weak-$*$ topology inherited from $\mathcal{S}'$, we compute the inner product

$$\langle f, g\rangle_{(s)} = \langle \Lambda_s f, \Lambda_s g\rangle_{L^2(dx)} = \langle \Lambda_{s-t}\Lambda_t f, \Lambda_{s-t}\Lambda_t g\rangle_{L^2(dx)} = \langle \Lambda_t f, \Lambda_t g\rangle_{(s-t)}$$

This establishes the fourth and fifth claims. Finally, to show $\partial^\alpha : H_s \to H_{s-|\alpha|}$ is toplinear. To this, we use the properties of the Fourier Transform for tempered distributions.

- $(\partial^\alpha F)^\wedge = (2\pi i\zeta)^\alpha \hat{F}$, where $\hat{F}$ denotes the distributional Fourier Transform, and the factor of $(2\pi i\zeta)^\alpha$ is a slowly increasing (smooth) function.
- Computing the Fourier Transform of $(\Lambda_{s-|\alpha|}\partial^\alpha f)$, we get

$$(\Lambda_{s-|\alpha|}\partial^\alpha f)^\wedge = (1+|\zeta|^2)^{s/2-|\alpha|/2}(\partial^\alpha f)^\wedge = (2\pi i)^{|\alpha|}(1+|\zeta|^2)^{(s-|\alpha|)/2}\zeta^\alpha \hat{f}$$

So $(\Lambda_{s-|\alpha|}\partial^\alpha f^\wedge) = C_\alpha(1+|\zeta|^2)^{(s-|\alpha|)/2}\zeta^\alpha \hat{f}$, where $C_\alpha = (2\pi i)^{|\alpha|}$. We know that

$$\Lambda_s f \in L^2 \iff (\Lambda_s f)^\wedge = (1+|\zeta|^2)^{s/2}\hat{f} \in L^2(d\mu_s)$$

So the expression that *defines* $(\Lambda_{s-|\alpha|}\partial^\alpha f)^\wedge$ as a pointwise function. We can compute its L^2 norm,

$$\begin{aligned}
\|\partial^\alpha f\|_{(s-|\alpha|)} &= \|(\Lambda_{s-|\alpha|}\partial^\alpha f)^\wedge\|_{\,L^2} = \|(1+|\zeta|^2)^{s/2-|\alpha|/2}(\partial^\alpha f)^\wedge\|_{\,L^2} \\
&= \left\|(1+|\zeta|^2)^{(s-|\alpha|)/2} C_\alpha \zeta^\alpha \hat{f}\right\|_{L^2} \\
&\lesssim_\alpha \left\|(1+|\zeta|^2)^{s/2}\hat{f}\right\|_{L^2} \lesssim_\alpha \|f\|_{(s)}
\end{aligned}$$

the first equality follows from Plancherel, and the second last estimate is justified in lem. 1.3. ■

Before stating the Sobolev Embedding Theorem, we note that if $s > 0$, then H_s embeds continuously into $H_0 = L^2$. Identifying $f \in H_s$ with its L^2 representative, it makes sense to evaluate $f(x) \in \mathbb{C}$ up to a null set.

If the L^2 representative of f coincides a.e with a continuous function g we can *identify* f again with this continuous function. If g is a member of any of the *continuous function spaces* we have discussed (e.g: C_0^k, $\mathcal{S}$) then we *say* $f \in C_0^k$ or $f \in \mathcal{S}$.

If every member of H_s belongs some continuous function space, for example C_0^k, then we write $H_s \subseteq C_0^k$. The obvious question becomes, given H_s for $s > 0$,

- When can be identify each H_s as a subset of C_0^k?
- When is the inclusion map $j : H_s \to C_0^k$ toplinear?
- When is the inclusion map compact?

Clearly the first question can be answered using regularity properties of the Fourier Transform, and the second depends on finding an estimate for the norm $f \in H_s$

$$\sum_{|\alpha| \leq k} \|\partial^\alpha f\|_u = \|f\|_{C_0^k} \lesssim_k \|f\|_{(s)}$$

Proposition 2.2: Sobolev Embedding Theorem

Let $s > k + n2^{-1}$, where $k = 0, 1, \ldots$.

- For every $f \in H_s$ and multi-index α with $|\alpha| \leq k$: the Fourier Transform of the α-distributional derivative of f is in L^1.

  $$(\partial^\alpha f)^\wedge \in L^1 \quad \text{and} \quad \|\mathcal{F}(\partial^\alpha f)\|_{L^1} \leq C\|f\|_{(s)}$$

  where C is given by eq. (120)

  $$C = (2\pi)^k \sqrt{\int (1 + |\zeta|^2)^{k-s} d\zeta} \tag{120}$$

  is independent of f and depends only on $k - s$.
- H_s can be identified as a subspace of C_0^k, and the inclusion map $j : H_s \to C_0^k$ is continuous.

Remark 2.1: Sharper constant C

We can obtain a sharper estimate by replacing the factor $(2\pi)^k$ with $(2\pi)^{k-s}$ in eq. (120).

Proof. We start with some semantics. Since $s > k + n2^{-1} > 0$, $H_s \subseteq L^2$. The Fourier Transform of $f \in H_s$ is a pointwise function in L^2. The first bullet point is a *statement* about the integrability of the pointwise function $\mathcal{F}(\partial^\alpha f)$, where $\mathcal{F}$ and ∂^α should be interpreted in the distributional sense, but it so happens that it produces a pointwise function, which we will *identify* with its tempered distribution.

$$\mathcal{F}(\partial^\alpha f) \in L^1 \subseteq \mathcal{S}'$$

Computing the Fourier Transform of $\partial^\alpha f$, where $\partial^\alpha : \mathcal{S}' \to \mathcal{S}'$ is the distributional derivative on the space of tempered distributions, the prop. 2.2

$$\partial^\alpha : H_s \to H_{s-|\alpha|} \quad \text{is continuous}$$

But $|\alpha| \leq k \implies s - |\alpha| \geq s - k > n2^{-1} > 0$. So the $\partial^\alpha f$ is also L^2, and

$$\mathcal{F}(\partial^\alpha f) = (2\pi i)^{|\alpha|}\zeta^\alpha \hat{f}(\zeta)$$

Computing a pointwise estimate, using lem. 1.4 and

$$|\mathcal{F}(\partial^\alpha f)(\zeta)| \leq (2\pi)^k |\zeta^\alpha||\hat{f}| \lesssim_k (1 + |\zeta|^2)^{|\alpha|/2}|\hat{f}| \lesssim_k (1 + |\zeta|^2)^{k/2}|\hat{f}|$$

We integrate over $\mathbb{R}^n$, still denoting $(1 + |\zeta|^2)^{t/2}$ by A_t, and by Cauchy-Schwartz:

$$\begin{aligned}\|(\partial^\alpha f)^\wedge(\zeta)\|_{L^1} &\lesssim_k \|A_k \hat{f}\|_{L^1} \lesssim_k \|A_{k-s}\|_{L^2}\|A_s \hat{f}\|_{L^2} \\ &\lesssim_k \|A_{k-s}^2\|_{L^1}^{1/2}\|f\|_{(s)} = C\|f\|_{(s)}\end{aligned}$$

It suffices to show the integral defining $C = (2\pi)^k \|A_{k-s}^2\|_{L^1}^{1/2}$ converges. This is summarized in note 2.1.

Note 2.1: Integrability of C

A measurable function $f \in \mathbb{B}_{\mathbb{R}^n} = \mathbb{B}$ is *radially symmetric* whenever there exists a $g \in \mathbb{B}$ such that $f(x) = g(|x|)$. A moment's thought will show that this is independent of the choice of a.e representative.

Lemma 2.4: Radial Symmetry Lemma

If f is radially symmetric, with $f(x) = g(|x|)$ then

$$\int_{\mathbb{R}^n} f(x)dx = \sigma S^{n-1} \int_0^\infty g(r) r^{n-1} dr \tag{121}$$

where $\sigma(S^{n-1})$ is the surface measure of the $n-1$ sphere $\{x \in \mathbb{R}^n,\ |x| = 1\}$.

Proof. Postponed. ■

Let $B = [|x| < c]$ for some $c > 0$.

- The limit in eq. (122) exists whenever $n - a > 0$.

$$\int_0^c r^{(n-a)-1} dr = \left(\frac{1}{n-a}\right) r^{n-a}\Big|_0^c \tag{122}$$

- The limit in eq. (123) exists whenever $n - a < 0$.

$$\int_c^\infty r^{(n-a)-1} dr = \left(\frac{1}{n-a}\right) r^{n-a}\Big|_c^\infty \tag{123}$$

Lemma 2.5: Integrability of Radially Symmetric Functions

Let $f(x) = g(|x|)$ be radially symmetric, sufficient conditions for f to be integrable

- about the origin: suppose $(n-a) > 0$, if $|f(x)|\chi_B \lesssim |x|^a \chi_B$, then $f\chi_B \in L^1$.
- away from origin: suppose $(n-a) < 0$, if $|f(x)|\chi_{B^c} \lesssim |x|^a \chi_{B^c}$, then $f\chi_{B^c} \in L^1$.

Proof. Use eqs. (122) and (123). ∎

To prove $|A^2_{k-s}|$ is in L^1, we rearrange the equation $s > k + n2^{-1}$ to obtain $n - 2(s-k) < 0$ and $(k-s) < 0$. Since A^2_{k-s} is continuous, it suffices to prove $A^2_{k-s}\chi_{B^c} \in L^1$. The Sobolev factor $(1+|\zeta|^2)^{t/2}$ is bounded above by $|\zeta|^t$ for $t < 0$. Substituting $t = 2(k-s)$ reads

$$A^2_{k-s} = (1+|\zeta|^2)^{(k-s)} \leq |\zeta|^{-2(s-k)}$$

The right member has *negative exponent* $a = 2(s-k) > n$, therefore A^2_{k-s} is integrable away from the origin.

This proves the first bullet point.

The Fourier inversion integral converges for $\mathcal{F}(\partial^\alpha f) \in L^1$, and by Riemann-Lebesgue we see that

$$\partial^\alpha f = \mathcal{F}^{-1}\mathcal{F}(\partial^\alpha f) \in C_0$$

where $\mathcal{F}^{-1}\mathcal{F} = \mathcal{F}\mathcal{F}^{-1} = \mathrm{id}_{\mathcal{S}'}$, and the inversion formula $\mathcal{F}^{-1}(g(\zeta)) = \mathcal{F}(g(-\zeta))$ for $g \in L^1$ implies

$$\|\check{g}\|_u = \|(\hat{g}(-\zeta))\|_u \leq \|g(-\zeta)\|_{L^1} = \|g\|_{L^1}$$

let $g = \mathcal{F}(\partial^\alpha f)$ and we see that

$$\|\partial^\alpha f\|_u \leq \|(\partial^\alpha f)^\wedge\|_{L^1} \lesssim_k \|f\|_{(s)} \quad \forall |\alpha| \leq k$$

Summing over all such α, we obtain an estimate for the C^k_0 norm of f.

$$\|f\|_{C^k_0} = \sum_{|\alpha|\leq k} \|\partial^\alpha f\|_u \lesssim_{k,n} \|f\|_{(s)}$$

and this proves the second bullet point. ∎

Corollary 2.1

If $f \in H_s$ for all $s > 0$ as a tempered distribution, then it can be identified pointwisely as a function in $C^\infty \cap C_0$.

Chapter 2: Periodic Distributions

Periodic Fourier Transform

If f is periodic, we define the Fourier Transform of f to be

$$\mathcal{F}f(k) = \hat{f}(k) = \int_{\mathbb{T}^n} f(x)E_{-k}(x)dx \tag{124}$$

where $E_{-k} = e^{-2\pi i\langle k,x\rangle}$.

Some properties of the integral defined in eq. (124).

- If $f \in L^p(\mathbb{T}^n)$, then eq. (124) converges at every $k \in \mathbb{Z}^n$.
- Hausdorff-Young: If $f \in L^p(\mathbb{T}^n)$, for $1 \leq p \leq 2$, and q is conjugate to p, then
$$\|\hat{f}\|_{l^q} \leq \|f\|_{L^p(\mathbb{T}^n)}$$
- Regularity transforms into integrability (or decay): if $f \in C^k(\mathbb{T}^n)$, then
$$(\partial^\alpha f)^\wedge(k) = (2\pi ik)^\alpha \hat{f}(k) \tag{125}$$
and
$$\|(2\pi ik)^\alpha \hat{f}(k)\|_{l^\infty} \leq \|\partial^\alpha f\|_{L^1(\mathbb{T}^n)} \tag{126}$$
for $|\alpha| \leq k$.

The last claim is summarized in the following Lemma.

Lemma 1.1: Regularity transforms into decay

If $f \in C^p(\mathbb{T}^n)$, then eqs. (125) and (126) holds for all $|\alpha| \leq p$. Furthermore,

$$|\hat{f}(k)| \lesssim_f (1+|k|)^{-p} \quad \forall k \in \mathbb{Z}^n \tag{127}$$

Proof. We want to prove that, differentiation transforms into multiplication by coordinate functions. We will use induction on $|\alpha|$. Let $f \in C^1(\mathbb{T}^n)$. The partial derivative with respect to the jth coordinate,

$$\partial^{e_j} f(x) = Df(x)(e_j)$$

Then Fourier Transform of $\partial^{e_j} f$ then becomes,

$$\langle \partial^{e_j} f, E_k\rangle = \int_{\mathbb{T}^n} m(Df(x)(e_j), E_{-k}(x))dx$$

where m denotes the multiplication map. By the product rule,

$$m(Df(x)(e_j), \phi(x)) = Dm(f(x), \phi(x))(e_j) - m(f(x), D\phi(x)(e_j))$$

because $m(f(x), \phi(x))(e_j)$ is periodic, the boundary terms disappear when integrated against x_j. Since $DE_{-k}(x)(e_j) = (-2\pi ik^j)E_{-k}(x)$, we have

$$\begin{aligned}
&\int \cdots \int_{\prod_{\underline{n}} \mathbb{T}} m(Df(x)(e_j), E_{-k}(x)) \prod d\mu(x_{\underline{n}}) = \\
&\quad \int \cdots \int_{\prod_{\underline{n-1}} \mathbb{T}} \left(m(f(x+te_j), \phi(x+te_j))\Big|_{\partial\mathbb{T}} - \int_{\mathbb{T}} m(f(x), DE_{-k}(x)(e_j))d\mu(x_j) \right) \prod d\mu(x_{i\neq j}) \\
&\qquad = (2\pi ik^j) \int_{\mathbb{T}^n} m(f(x), E_{-k}(x))d\mu(x)
\end{aligned} \tag{128}$$

therefore $\langle \partial^{e_j} f, E_k \rangle = (2\pi i k^j)\hat{f}(k)$. The general case follows from the fact that $\partial^\alpha f = \partial^{e_j}\partial^{\alpha - e_j} f$. This proves eq. (125), and eq. (126) follows by Holder's inequality with respect to the counting measure.

Remark 1.1

The argument above relies on the fact that the boundary terms vanish. This is why we require $\partial^\alpha f \in C_0(\mathbb{R}^n)$ in order to diagonalise differentiation.

Next, given $f \in C^p(\mathbb{T}^n)$ we see that if $|\beta| \leq p$, $|k^\beta||\hat{f}(k)| \leq C_\beta$. We can take the maximum of all such $|\beta| \leq p$, and relabel C. To proceed any further we will need the following very useful estimate, which allows us to bound the polynomial factor by powers of multi-indices with degree less than N.

$$(1 + |x|)^N \lesssim_N \sum_{|\beta| \leq N} |x^\beta| \tag{129}$$

Note 1.1: Proof of the estimate

We offer a sketch. Let $N \geq 1$ be fixed, prove that $\sum |x_{\underline{n}}|^n \geq \delta > 0$ on S^{n-1} by a continuity argument. Then, use a functional analytic 'homogeneity' argument and $\sum |x_{\underline{n}}|^N \geq \delta |x|$ for all $x \in \mathbb{R}^n$. Using the binomial theorem to break apart $(1 + |x|)^N$ into smaller pieces, each of which can be bounded by $(1 + |x|^N)$, we obtain the result.

Using eq. (129), we see that

$$(1 + |k|)^p |\hat{f}(k)| \lesssim_p \sum_{|\beta| \leq p} C_\beta \lesssim_{p,f} C$$

Multiplying across proves eq. (127). ■

Exercise 9.22-9.24

Few more things, if $\phi \in C_c^\infty(\mathbb{R}^n)$, we define the *periodization map* P by the sum in eq. (130).

$$P\phi(x) = \sum_{k \in \mathbb{Z}^n} \tau_k \phi(x) = \sum_{k \in \mathbb{Z}^n} \phi(x - k) \tag{130}$$

We define the *distributional periodization map* P' to be a map

$$P : \mathbf{D}'(\mathbb{T}^n) \to \mathbf{D}'(\mathbb{R}^n)$$

that precomposes $F \in \mathbb{D}'(\mathbb{T}^n)$ with P. That means

$$\langle P'F, \phi \rangle_{\mathcal{D}(\mathbb{R}^n)} = \langle F, P\phi \rangle_{(\mathcal{D}(\mathbb{T}^n))}$$

Proposition 2.1: Exercise 9.22

The periodization map $P : C_c^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{T}^n)$ and its adjoint, defined in eq. (130) are toplinear.

Proof. Let $\phi_j \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$. Since $\mathcal{D}(\mathbb{R}^n)$ is topologized in a manner that allows us to pass to some compact $K \subseteq \mathbb{R}^n$, this is equivalent to $\phi_j \to \phi$ in $\mathcal{D}(K)$. Where $\mathcal{D}(K)$ is a Frechet space with the norms $\|\partial^\alpha \phi|_K\|_u$.

Suppose $x \in \mathbb{T}^n$, because $\tau_k K$ is locally finite, there exists a neighbourhood U and M (dependent on x) where

$$P\phi_j|_U = \sum_{|k| \leq M} \tau_k \phi_j$$

similarly for $P\phi|_U$ as well. This means we can interchange P and ∂^α by linearity, and

$$\|\partial^\alpha[P\phi_j - P\phi]\|_{u,U} = \|\sum_{k \leq M} P\partial^\alpha(\phi_j - \phi)\|_{u,U} \lesssim_M \|\partial^\alpha(\phi_j - \phi)\|_{u,K}$$

Taking the maximum over all such U and M (because $\mathbb{T}^n$ is compact) proves the continuity of P and of its adjoint. ■

Because $\tau_k \circ P = P$ for all $k \in \mathbb{Z}^n$, and the adjoint of P precomposes the distribution by P itself, if $F \in \mathcal{D}'(\mathbb{T}^n)$

$$\langle \tau_k P' F \phi \rangle_{\mathcal{D}(\mathbb{R}^n)} = \langle F, P\tau_k \phi \rangle_{\mathcal{D}(\mathbb{T}^n)} = \langle P'F, \phi \rangle_{\mathcal{D}(\mathbb{R}^n)}$$

With this, we define the range of P' — the space of *shift invariant distributions.*

$$\mathcal{D}'_{per}(\mathbb{R}^n) = \left\{ F \in \mathcal{D}'(\mathbb{R}^n),\ \tau_k F = F\ \forall k \in \mathbb{Z}^n \right\} \tag{131}$$

Our goal in this section is to show that the mapping P' is a bijection.

Lemma 2.1

It can also be shown that $\mathcal{D}'_{per}(\mathbb{R}^n) \subseteq \mathcal{S}'$, which states: *every* shift-invariant distribution is tempered. Moreoever, P maps $\mathcal{D}'(\mathbb{T}^n)$ into $\mathcal{D}'_{per}(\mathbb{R}^n)$. This is a bijection.

Proof. ■

Lemma 2.2

If $g \in C_s^\infty(\mathbb{Z}^n)$, meaning $|g(k)| \leq C(1+|k|)^N$ for some C, N. Then, the Fourier Series $\sum_{k \in \mathbb{Z}^n} g(k) E_k(x)$ converges in $\mathcal{D}'(\mathbb{T}^n)$ to a distribution F such that $\hat{F} = g$.

Moreoever, it converges to a tempered distribution $G \in \mathcal{S}'(\mathbb{R}^n)$ such that $G = P'F$, and G is *shift invariant* — that is, $\tau_k G = G$ for all k.

Proof. Let $g : \mathbb{Z}^n \to \mathbb{C}$ that is polynomially bounded, this means $|g(k)| \leq C(1 + |k|)^N$ for some C and N. Then, for every $\phi \in C^\infty(\mathbb{T}^n)$,

$$\left\langle \sum g(k) E_k,\ \phi \right\rangle_{\mathcal{D}} \quad \text{converges absolutely}$$

and that $\sum g(k)E_k \in \mathcal{D}'$. For any finite sum,

$$\sum^{\wedge} |g(k)\hat{\phi}(k)| \leq \sum^{\wedge} (1+|k|)^N |\hat{\phi}(k)| \leq \sum_{k\in\mathbb{Z}^n} (1+|k|)^{N+2} |\hat{\phi}(k)| (1+|k|)^{-2}$$
$$\leq \left\|(1+|k|)^{N+2}\hat{\phi}\right\|_{l^\infty} \cdot \left\|(1+|k|)^{-2}\right\|_{l^1} \leq_g \|\partial^\alpha \phi\|_{L^1(\mathbb{T}^n)}$$

If $\phi_n \to \phi$ in $\mathcal{D}$, its Fourier coefficients converge pointwise as well. Applying Lebesgue's theorem with respect to the counting measure tells us $\sum g(k)E_k$ is in $\mathcal{D}$. ■

Periodic Sobolev Spaces H_s over $\mathbb{C}$

We give the first *definition* of the *periodic Sobolev spaces* H_s over the complex plane by defining H_s as a subspace of $\mathcal{D}'(\mathbb{T}^n)$ by imposing integrability condition on its Fourier Transform. Let Λ_s be the map

$$\Lambda_s : \mathcal{D}'(\mathbb{T}^n) \to \mathcal{D}'(\mathbb{T}^n) \quad \Lambda_s F = \mathcal{F}^{-1}((1+|k|^2)^{s/2}\hat{F}(k))$$

where $\mathcal{F}$ and its inverse should be viewed from $\mathcal{D}'(\mathbb{T}^n)$ to $C_s(\mathbb{Z}^n)$.

Remark 3.1

Just like $C_s^\infty(\mathbb{R}^n)$, the space of slowly increasing sequences $C_s(\mathbb{Z}^n)$ is closed under pointwise multiplication, and $B_s(k) = (1+|k|^2)^{s/2}$ is in $C_s(\mathbb{Z}^n)$ for every $s \in \mathbb{R}$.

Periodic Distributions

Let $\mathbb{T}^n = (S^1)^n = (\mathbb{R}/\mathbb{Z})^n$ is a compact Hausdorff space.

- A function $f : \mathbb{R}^n \to \mathbb{C}$ is *periodic* (or 1-periodic) if $f(x+k) = f(x)$ for every $k \in \mathbb{Z}^n$.
- We write $Q = [0,1)^n$

Definition 4.1: Periodic function

A function $f : \mathbb{R}^n \to \mathbb{C}$ is *periodic* (or 1-periodic) if $f(x+k) = f(x)$ for every $k \in \mathbb{Z}^n$. f can be uniquely identified within the space of 1-periodic functions by its values in $Q = [0,1)^n$. The space of smooth periodic functions on $\mathbb{R}^n$ is denoted by $C^\infty(\mathbb{T}^n)$.

$$C^\infty(\mathbb{T}^n) = \left\{ f \in C^\infty(\mathbb{R}^n),\ f \text{ is periodic.} \right\}$$

Because the n-torus is compact, we can endow $C^\infty(\mathbb{T}^n)$ with the Frechet topology of uniform convergence, and if $f \in C^\infty(\mathbb{T}^n)$, its *restriction* onto the Q is smooth. Since Q represents $\mathbb{T}^n$ as a quotient space, there exists an injection from smooth functions $C^\infty(Q) \to C^\infty(\mathbb{T}^n)$.

If $g \in C^\infty(Q)$ the values of g are completely determine by points in Q. We can extend g to a pointwise function on $\mathbb{R}^n$ by the *extension map* $W : C^\infty(\mathbb{T}^n) \to C^\infty(\mathbb{R}^n)$. We note that this sum in eq. (132) converges to a smooth function in $C^\infty(\mathbb{R}^n)$, and $W : C^\infty(\mathbb{T}^n) \to C^\infty(\mathbb{R}^n)$ can be shown to be toplinear.

$$Wg(t) = W_g(t) = g(x) \quad \text{where } t - x \in \mathbb{Z}^n \tag{132}$$

This coincides with the definition of $C^\infty(\mathbb{T}^n)$, where each $f \in C^\infty(\mathbb{T}^n)$ is uniquely determined by its values in Q, meaning

$$Wf|_Q = f$$

We are now in the position to discuss the Fourier Transform of distributions on $\mathbb{T}^n$. Since $C^\infty = C_c^\infty$ on $\mathbb{T}^n$, their dual spaces satisfy $\mathcal{E}' = \mathcal{D}' = \mathcal{S}'$.

Recall, the Fourier Transform of a $F \in \mathcal{S}'(\mathbb{R}^n)$ is *defined* by precomposing F with the Fourier Transform on $\mathcal{S}$ — which is a linear isomorphism. With this, we are ready to define $\mathcal{F}$ on $\mathcal{D}'(\mathbb{T}^n)$.

Definition 4.2: Fourier Transform on Periodic Distributions

The *Fourier Transform on* $\mathcal{D}'(\mathbb{T}^n)$ is a linear mapping

$$\mathcal{F} : \mathcal{D}'(\mathbb{T}^n) \to C_s(\mathbb{Z}^n) \quad \text{and} \quad \mathcal{F}F(k) = \hat{F}(k) = \langle F, E_{-k} \rangle_{\mathcal{D}}$$

where $E_{-k}(x) = e^{-2\pi i \langle k, x \rangle} \in C^\infty(\mathbb{T}^n)$.

Remark 4.1: $\mathcal{F}$ maps into slowly increasing sequences

The definition of $\mathcal{F}$ on $\mathcal{D}'(\mathbb{T}^n)$ maps into the space of complex-valued sequences $\mathbb{C}^{\mathbb{Z}^n}$. By definition of continuity between TVS, $F : C^\infty(\mathbb{T}^n) \to \mathbb{C}$. Because there exists a constant C and N, such that

$$|\langle F, \phi \rangle| \le C \sum_{|\alpha| \le N} \|\partial^\alpha \phi\|_u \quad \forall \phi \in C^\infty(\mathbb{T}^n)$$

for some N that is dependent on F. Applying this to $\hat{F}(k) = \langle F, E_{-k} \rangle$, and using lem. 1.3 for $|k^\alpha| \le |k|^{|\alpha|}$, we see that

$$|\hat{F}(k)| \lesssim_F \sum_{|\alpha| \le N} \|\partial^\alpha E_{-k}\|_u \lesssim_F \sum_{|\alpha| \le N} (2\pi k)^\alpha \|E_{-k}\|_u \lesssim_F (1 + |k|)^N$$

We list some facts of $\mathcal{F}$ on $\mathcal{D}'(\mathbb{T}^n)$

- The Fourier Transform is actually a linear isomorphism from $\mathcal{D}'(\mathbb{T}^n)$ to $C_s(\mathbb{Z}^n)$.
- Furthermore, the *Fourier Series* defined by taking linear combinations of $\hat{F}(k)E_k(x) \in C^\infty(\mathbb{T}^n)$ *converges in* $\mathcal{D}'(\mathbb{T}^n)$ (in weak-$*$) to F itself.
- A surprising but non-trivial result is that if we view linear combinations of $\hat{F}(k)E_k(x)$ as elements in $\mathcal{S}'(\mathbb{R}^n)$, then $\sum \hat{F}(k)E_k$ *converges in* $\mathcal{S}'(\mathbb{R}^n)$ (in weak-$*$) to $P'F$.
- Finally, the continuity of $\mathcal{F}$ on $\mathcal{S}'$ gives us the following result: the Fourier Transform of $\sum \hat{F}(k)E_k$ for $F \in \mathcal{D}'(\mathbb{T}^n)$ must *converge to the Fourier Transform of* $P'F$.

$$(P'F)^\wedge = \mathcal{F}(\sum_{k \in \mathbb{Z}^n} \hat{F}(k)E_k) = \sum \hat{F}(k)\tau_k \delta$$

Definition 4.3: Periodic Sobolev Spaces

If $s \in \mathbb{R}$, the *periodic Sobolev space* H_s is a subspace of $\mathcal{D}'(\mathbb{T}^n)$ where each element $f \in H_s$ satisfies

$$\mathcal{F}(\Lambda_s f) \in l^2(\mathbb{Z}^n) \quad \text{or} \quad \sum_{k \in \mathbb{Z}^n} (1+|k|^2)^s |\hat{f}(k)|^2 < +\infty$$

As in the case for $H_s(\mathbb{R}^n)$, we define the inner product by pulling back the inner product on l^2. This makes H_s a Hilbert space, and the Fourier Transform is a unitary isomorphism from H_s into $l^2(\mathbb{Z}^n, B_s^2(k)dk)$, where dk is the counting measure on $\mathbb{Z}^n$, $B_s(k) = (1+|k|^2)^{s/2}$ and the σ-algebra on $\mathbb{Z}^n$ is assumed to be maximal.

For all $f, g \in H_s$, the inner product on H_s is then given by

$$\langle f, g \rangle_{(s)} = \langle \Lambda_s f, \Lambda_s g \rangle_{L^2} = \langle (\Lambda_s f)^\wedge, (\Lambda_s g)^\wedge \rangle_{L^2} = \sum_{k \in \mathbb{Z}^n} (1+|k|^2)^s \hat{f}(k) \overline{\hat{g}(k)} \tag{133}$$

with the usual norm induced by the inner product $\|f\|_{(s)} = \langle f, f \rangle_{(s)}^{1/2}$

$$\|f\|_{(s)} = \|\Lambda_s f\|_{l^2} = \sqrt{\sum_{k \in \mathbb{Z}^n} (1+|k|^2)^s |\hat{f}(k)|^2} \tag{134}$$

If $s \geq 0$ it is fruitful to consider another choice of Λ_s that induces the same norm (hence topology), but with a different inner product. Let us write

$$A_s(k) = \delta_0 + 2\pi|k|^s \quad \text{and} \quad \|f\|_A = \|A_s \hat{f}\|_{l^2} \tag{135}$$

and

$$B_s(k) = (1+|k|^2)^{s/2} \quad \text{and} \quad \|f\|_B = \|B_s \hat{f}\|_{l^2} \tag{136}$$

We wish to show the norms induced by A_s and B_s are equal. The proof of this is summarized in the following note.

Note 4.1: Equivalent norms whenever $s \geq 0$

Lemma 4.1

Let $s \geq 0$ and $C = (2\pi)$, then $|A_s| \leq C|B_s|$ pointwise for all k.

Proof. If $k \neq 0$, then $(2\pi)^{2/s}|k|^2 \leq (2\pi)^{2/s}(1+|k|^2)$, and taking the $s/2 \geq 0$ power reads

$$A_s(k) = 2\pi|k|^s \leq (2\pi)(1+|k|^2)^{s/2} = 2\pi B_s(k)$$

Also, $(2\pi)B_s(1) \geq 1 = A_s(1)$, therefore $A_s(k) \leq (2\pi)B_s(k)$. ∎

Lemma 4.2

Let $s \geq 0$ and $C = \max(2^{s/2}(2\pi)^{-1}, 1)$, then $|B_s| \leq C|A_s|$ pointwise for all k.

Proof. Notice that if $k \neq 0$, then $|k| = \sqrt{\sum_{i=\underline{n}} |k_i|^2} \geq 1$. Hence,

$$2^{-1/2} \leq |k|(1+|k|^2)^{-1/2} \implies (1+|k|^2)^{s/2} \leq 2^{s/2}|k|^s$$

Since $k \neq 0$, it follows that

$$B_s(k) = (1+|k|^2)^{s/2} \leq |k|^s 2^{s/2} \leq C(\delta_0 + 2\pi|k|^s) = CA_s(k)$$

If $k = 0$, then $CA_s(0) = C \geq 1 = B_s(0)$, and the proof is complete. ■

Remark 4.2: Redefining Λ_s

From now on Λ_s will refer to the map

$$\Lambda_s : \mathcal{D}'(\mathbb{T}^n) \to \mathcal{D}'(\mathbb{T}^n) \quad \Lambda_s f = (A_s \hat{f})^\vee = \mathcal{F}^{-1}((1+2\pi|k|^s)\hat{f}(k))$$

where A_s is found in eq. (135). In terms of Fourier coefficients, this corresponds to

- $(\Lambda_s F)^\wedge(0) = \hat{F}(0)$, while
- $(\Lambda_s F)^\wedge(k) = 2\pi|k|^s \hat{F}(k)$ for $k \neq 0$

Definition 4.4: Periodic Sobolev Spaces (redefined)

If $s \geq 0$, we define the *periodic Sobolev space* H_s to be the subspace of distributions on $\mathbb{T}^n$ that satisfies

$$H_s = \left\{ f \in \mathcal{D}'(\mathbb{T}^n),\ (\Lambda_s f)^\wedge \in l^2(\mathbb{Z}^n, dk) \right\}$$

Alternatively, we can absorb the factor of Λ_s into the measure, by writing

$$H_s = \left\{ f \in \mathcal{D}'(\mathbb{T}^n),\ \hat{f} \in l^2(\mathbb{Z}^n, \mu_s) \right\}$$

where $\mu_s(A) = \sum_{j \in A}(\delta_0(j) + |j|^{2s})$ which is simply the integral of the additional 'factor' with respect to the counting measure dk.

Remark 4.3: Identifying $H_s \subseteq L^2$

We can simplify things further if we identify $H_s \subseteq L^2$ (because $s \geq 0$), and

$$H_s = \left\{ f \in L^2(\mathbb{T}^n, dx),\ \hat{f} \in l^2(\mathbb{Z}^n, \mu_s) \right\}$$

But the Fourier Transform is a unitary isomorphism between $L^2(\mathbb{T}^n, dx)$ and $l^2(\mathbb{Z}^n, dk)$, combining the first and last characterization, we write

$$H_s = \left\{ f \in L^2(\mathbb{T}^n, dx),\ \Lambda_s f \in L^2(\mathbb{T}^n, dx) \right\}$$

similar to Definition 8.1, the claim $\Lambda_s f \in L^2(\mathbb{T}^n)$ should be interpreted with respect to $\mathcal{D}'$. This means *there exists $g \in L^2(\mathbb{T}^n)$ that realizes the duality pairing*

$$\langle \Lambda_s f, \phi \rangle_{\mathcal{D}} = \langle g, \iota\phi \rangle_{L^2}$$

where $\iota : C^\infty(\mathbb{T}^n) \to L^2(\mathbb{T}^n)$ is the toplinear embedding.

The inner product and the norm on H_s is now given by

$$\begin{aligned} \langle f, g \rangle_{(s)} &= \langle \Lambda_s f, \Lambda_s g \rangle_{L^2} = \left\langle (\Lambda_s f)^\wedge,\ (\Lambda_s g)^\wedge \right\rangle_{L^2} \\ &= \sum_{k \in \mathbb{Z}^n} (\delta_0 + 2\pi |k|^s) \hat{f}(k) \overline{\hat{g}(k)} \end{aligned} \tag{137}$$

We define $A_s(j) = \delta_0(j) + \sqrt{2\pi}|j|^s$ for $j \in \mathbb{Z}^n$, so that

$$\langle f, g \rangle_{(s)} = \sum_{k \in \mathbb{Z}^n} |A_s|^2 \left\langle \hat{f}(k),\ \hat{g}(k) \right\rangle_{\mathbb{C}} = \left\langle \hat{f}(0),\ \hat{g}(0) \right\rangle_{\mathbb{C}} + 2\pi \sum_{\substack{k \in \mathbb{Z}^n \\ k \neq 0}} |k|^{2s} \left\langle \hat{f}(k),\ \hat{g}(k) \right\rangle_{\mathbb{C}} \tag{138}$$

Vector-valued H_s loops over $\mathbb{C}$

We will now consider the case where the domain is $\mathbb{R}^1 = \mathbb{R}$, and measurable which are vector valued, i.e $f : \mathbb{R} \to \mathbb{C}^n$, where $n \geq 1$.

If $f = (f_1, \ldots, f_n)$ where each f_i is $(\mathbb{R}, \mathbb{C})$ measurable. We say f is L^p if each $f_i \in L^p$. Continuity and smoothness properties of f should be interpreted in a geometric setting. If $f \in C^p$, then it is a *morphsim of class C^p*.
See 'vector-valued-lp-spaces' post for a summary. For each $f \in L^2(\mathbb{T}, \mathbb{C}^n)$,

$$\hat{f} : \mathbb{Z} \to \mathbb{C}^n \quad \hat{f}(k) = (\hat{f}_1(k),\ \ldots,\ \hat{f}_n(k))$$

The $L^2(\mathbb{T}, \mathbb{C}^n) = L^2$ inner product of f, g is defined similarly,

$$\langle f, g \rangle_{L^2} = \sum_k \langle \hat{f}(k), \hat{g}(k) \rangle_{\mathbb{C}^n} = \sum_k \sum_i \langle \hat{f}_i(k), \hat{g}_i(k) \rangle_{\mathbb{C}} = \sum_{i \leq n} \langle f_i, g_i \rangle_{L^2}$$

Proposition 5.1

Prop 3: If $t > s \geq 0$, the Sobolev spaces decrease, while the norms increase.

$$H_t \subseteq H_s \quad \text{and} \quad \|\cdot\|_{(s)} \leq \|\cdot\|_{(t)}$$

Moreoever, the inclusion $I : H_t \to H_s$ is a continuous compact map.

Proof. The first two claims follow immediately from the definition of vector-valued H_s, and from Theorem 9.1, 9.2.

To show compactness, we approximate ι with finite-rank operators (the symmetric partial sums S_m in this case).

$$S_m f = \sum_{|k|\leq N} E_k \hat{f}(k)$$

The idea is to use the fact that the norms on H_s are defined through the pullback

$$\Lambda_s : f \mapsto \mathcal{F}^{-1}(A_s(k)\hat{f}(k))$$

with $A_s = \delta_0 + \sqrt{2\pi}|k|^s$. We approximate the inclusion map $I : H_t \to H_s$

$$\begin{aligned}
\|S_n f - If\|_{H_s}^2 &= \Big\| \sum_{|k|>N} \hat{f}(k) E_k \Big\|_{H_s}^2 = 2\pi \sum_{|k|>N} |\hat{f}|^2 |A_s|^2 \\
&= 2\pi \sum_{|k|>N} |\hat{f}|^2 |k|^{2s} = 2\pi \sum_k |k|^{2(s-t)} |k|^{2t} |\hat{f}|^2 \\
&\leq 2\pi |N|^{2(s-t)} \sum_k |k|^{2t} |\hat{f}|^2 \lesssim N^{-2a} \|f\|_{H_t}^2
\end{aligned}$$

for some $a = t - s > 0$. Taking square roots gives $\|S_N f - If\|_{H_s} \lesssim N^{-a} \|f\|_{H_t}$. This holds for an arbitrary $f \in H_t$, which yields

$$\|S_N - I\|_{\mathcal{L}(H_t, H_s)} \lesssim N^{-a} \quad \text{and} \quad \forall M > N,\ \|S_M - I\| \lesssim N^{-a} \to 0.$$

Therefore I is compact. ∎

Proposition 5.2

Prop 4: If $s > k + 2^{-1}$, then $H_s(S^1) \subseteq C^k(S^1, \mathbb{C}^n)$. Essentially the periodic analogue of the Sobolev Embedding Theorem, moreover

$$\|\mathcal{F}(\partial f)\|_{l^1} \lesssim_{k,k-s} \|f\|_{H_s} \quad \text{and} \quad \|\partial f\|_u \lesssim_{k,k-s} \|f\|_{H_s}$$

for all multi-indices $|\alpha| \leq k$.

Proof. We first compute the first estimate for the l^1 norm of the weak-α derivative of f. The following holds pointwise for $j \in \mathbb{Z}$.

$$|\mathcal{F}(\partial^\alpha f)| = |2\pi|^{|\alpha|} |j^\alpha| |\hat{f}|$$

Because the domain is 1-dimensional, the α is a scalar, so $|j^\alpha| = |j|^\alpha$.

$$\|\mathcal{F}(\partial^\alpha f)\|_{l^1} \lesssim_k \Big\| |j|^k |\hat{f}| \Big\|_{l^1} \lesssim \Big\| |j|^s |\hat{f}| \Big\|_{l^2} \Big\| |j|^{k-s} \Big\|_{l^2} \lesssim_{k,k-s} \|f\|_{H_s}$$

The last estimate is justified by

- $|j|^s \leq A_s(j)$ pointwise for $j \in \mathbb{Z}$, and

- $\sum_j |j|^{2(k-s)}$ has exponent $2(k-s) < -1$, so it converges to *something* finite.

Now, use the Weierstrass M-test to show the series:

$$\sum_{k\in\mathbb{Z}} \hat{f}(k)E_k \quad \text{converges absolutely, uniformly to some } g \in C(S^1)$$

so f (viewed as an a.e class of functions) admits a continuous representative. Furthermore, all the weak-derivatives of f exist (up to order k) and are continuous, by the previous section - there exists a unique C^k representative of f, whose ordinary derivatives represent the corresponding weak derivatives of f.

The M-test also gives us the estimate:

$$\|\partial^\alpha f\|_u \leq \|\mathcal{F}(\partial^\alpha f)\|_{l^1} \lesssim_{k,k-s} \|f\|_{H_s}$$

if we equip C^k with the standard norm $\|f\|_{C^k} = \sum_{|\alpha|\leq k} \|\partial^\alpha f\|_u$, then $\|f\|_{C^k} \lesssim_s \|f\|_{H_s}$ as well. ∎

Corollary 5.1

If $f_n \to f$ in H_s, and k be a non-negative integer, such that $s > k + 2^{-1}$, then each f_n (resp. f) admits unique C^k representatives, whose ordinary derivatives represent the weak derivatives of f_n (resp. f) up to order k. And $f_n \to f$ in C^k.

Adjoint map $j : H_{1/2} \to L^2$

Proposition 6.1

Let $j : H_{1/2} \to L^2$ be the inclusion map. It is a compact continuous linear map, and so is the adjoint map $j^* : L^2 \to H_{1/2}$ defined by

$$\forall x \in H_{1/2},\ y \in L^2 \quad \langle j(x), y\rangle_{L^2} = \langle x, j^*y\rangle_{H_{1/2}}$$

If $y \in L^2$, then

$$j^*y = \hat{y}(0) + \sum_{k\neq 0} (2\pi|k|)^{-1}\hat{y}(k)E_k$$

The adjoint/pullback map also embeds L^2 into H_1, with

$$\|j^*y\|_{H_{1/2}} \leq \|j^*y\|_{H_1} \leq \|y\|_{L^2}$$

Proof. From the definition of j^*, fix $x \in H_{1/2}$ and $y \in L^2$. The left hand side becomes

$$\langle j(x), y\rangle_{L^2} = \langle x, j^*y\rangle_{H_{1/2}} = \langle \mathcal{F}(jx), \mathcal{F}y\rangle_{l^2} = \sum \langle \hat{x}(k), \hat{y}(k)\rangle_{\mathbb{C}^n}$$

And RHS:

$$\langle \hat{x}(0),\ (j^*y)^\wedge(0)\rangle_{\mathbb{C}^n} + 2\pi \sum_{k\neq 0} |k| \langle \hat{x}(k),\ (j^*y)^\wedge(k)\rangle_{\mathbb{C}^n}$$

We equate both sides using a technique we will reuse in later sections, setting x to an orthonormal basis vector with Fourier representation

$$x = E_k e_i \quad k \in \mathbb{Z}^n,\ 1 \leq i \leq n$$

(recall each $\hat{x}(k)$ is an element in $\mathbb{C}^n$). The "i" in the exponent refers to the imaginary unit, while the "i" in the lower index is a dummy variable, and $e_i = (0, \ldots, 1, \ldots, 0)$ is a standard basis vector in $\mathbb{C}^n$. We see that $\hat{y}(0) = (j^*y)^\wedge(0)$, and $\hat{y}(k) = 2\pi|k|(j^*y)^\wedge(k)$. Computing the H_1 norm of j^*y, we see

$$\begin{aligned}\|j^*y\|_{H_1}^2 &= |\hat{y}(0)|^2 + 2\pi \sum_{k\neq 0} |k|^2 |\underbrace{2\pi|k|^{-1}\hat{y}(k)}_{\mathcal{F}(j^*y)(k)}|^2 \\ &= |\hat{y}(0)|^2 + (2\pi)^{-1} \sum_{k\neq 0} |\hat{y}(k)|^2\end{aligned}$$

which is clearly less than $\|\hat{y}\|_{l^2}^2 = \|y\|_{L^2}^2$, and $\|j^*y\|_{H_{1/2}} \leq \|j^*\|_{H_1}$ follows because norms increase. ■

Chapter 1: Multilinear maps

Bilinear maps

Definition 1.1: Bilinear map

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

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

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 $m : E_1 \times E_2 \to F$ is continuous if and only if there exists a $C \geq 0$, where

$$|m(x, y)| \leq C|x||y| \tag{139}$$

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

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

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 eithe 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 $(x, y) \in E_1 \times E_2$, $\omega(x, y) = \varphi(x)(y)$.

Proof. Let $\varphi_\omega : E_1 \to L(E_2, F)$ 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 $\varphi : L(E_1, E_2; F) \to L(E_1, L(E_2, F))$ 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 $\omega : E_1 \times E_2 \to F$ 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 (139), 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 $i = \underline{k}$, $v_i \in E_i$,

$$\varphi(\underline{\cdot^{i-1}}, v_i, \underline{\cdot^{k-i}}) : \Pi(E_{\underline{i-1}}, E_{i+\underline{k-i}}) \to F \quad \text{is } (k-1)\text{-linear}$$

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 $\theta \in S_{\underline{k}}$,

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

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 $\varphi : \prod E_{\underline{k}} \to F$ 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:

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

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

$$(x_n^{\underline{k}}) \to (x^{\underline{k}}) \quad \text{as } n \to +\infty \tag{140}$$

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

$$\varphi(b^{\underline{k}}) - \varphi(a^{\underline{k}}) = \sum_{i=\underline{k}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{k-i}}) \tag{141}$$

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. (141) follows by setting $m = k$ in

$$\varphi(a^{\underline{k}}) = \varphi(b^{\underline{m}}, a^{m+\underline{k-m}}) - \sum_{i=\underline{m}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{k-i}}) \tag{142}$$

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

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

Induction hypothesis: suppose eq. (142) holds for a fixed m. Since $a^{m+1} = b^{m+1} - \Delta_{m+1}$,

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

and this proves eq. (141)

We substitute $a^i = x^i$, and $b^i = x_n^i$ for $i = \underline{k}$, and eq. (141) becomes eq. (143)

$$\varphi(x_n^{\underline{k}}) - \varphi(x^{\underline{k}}) = \sum_{i=\underline{k}} \varphi(x_n^{\underline{i-1}}, x_n^i - x^i, x^{i+\underline{k-i}}) \tag{143}$$

Then the triangle inequality reads

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

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

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 $L(E_{\underline{k}}; F) = L(E^k, F) = L^k(E, F)$. *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 $U \overset{\circ}{\subseteq} X$. If U is a *neighbourhood* of a point $p \in X$, we write $p \overset{\circ}{\in} U$.

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 $\lim_{t\to 0} o(t)/t = 0$. A map $\psi : U \to F$ where $U \overset{\circ}{\subseteq} E$ contains 0 in E, is said to be $o(h)$ if $|\psi(h)|/|h| \to 0$ as $h \to 0$ in E.

Definition 1.3: Differentiability

Let $f : E \to F$ 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

$$f(x+h) = f(x) + \lambda h + o(h) \quad \text{for sufficiently small } h \tag{144}$$

The role $o(h)$ plays here is a map from $U \to F$, 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. (144) 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 $\varphi(h) \to 0$ as $h \to 0$. Using the operator norm, we see that

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

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

$$\begin{aligned} f(x+h) + g(x+h) &= (f(x) + g(x)) + (\lambda_f(h) + \lambda_g(h)) + o(h) + o(h) \\ (f+g)(x+h) &= (f+g)(x) + (\lambda_f + \lambda_g)(h) + o(h) \end{aligned} \tag{145}$$

since eq. (145) satisfies eq. (144), 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 $x \in E$,

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x) \tag{146}$$

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

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 $m : \prod F_{\underline{k}} \to G$ be a continuous 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

$$m(f_{\underline{k}})(x) = m(f_{\underline{k}}(x)) \tag{147}$$

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

$$Dm(f_{\underline{k}})(x)(y) = \sum_{i=\underline{k}} m(f_{\underline{i-1}}(x), Df_i(x)(y), f_{i+\underline{k-i}}(x)) \tag{148}$$

Proof. Let x be fixed. Equation (148) is proven if we show eq. (149)

$$m(f_{\underline{k}})(x+h) = m(f_{\underline{k}})(x) + \left(\sum_{i=\underline{k}} m(f_{\underline{i-1}}(x), Df_i(x)(h), f_{i+\underline{k-i}}(x)) \right) + o(h) \tag{149}$$

and for sufficiently small h we have

$$f_i(x+h) - f_i(x) = Df_i(x)(h) + o(h^i) \tag{150}$$

We will use the difference formula in eq. (142), with the following substitutions

$$f_i(x+h) = b^i \qquad f_i(x) = a^i \tag{151}$$
$$Df_i(x)(h) = c^i \qquad o(h^i) = \varepsilon^i \tag{152}$$
$$f_i(x+h) - f_i(x) = c^i + \varepsilon^i \qquad \Delta^i = o(h^i) + c^i \tag{153}$$

With these substitutions, the equation we want to prove (eq. (148)) becomes eq. (154)

$$m(b^{\underline{k}}) - m(a^{\underline{k}}) = \left(\sum_{i=\underline{k}} m(a^{\underline{i-1}}, c^i, a^{i+\underline{k-i}}) \right) + o(h) \tag{154}$$

Starting from eq. (142),

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

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

$$m(b^{\underline{i-1}}, \Delta^i, a^{i+\underline{k-i}}) = m(b^{\underline{i-1}}, c^i, a^{i+\underline{k-i}}) + m(b^{\underline{i-1}}, o(h^i), a^{i+\underline{k-i}}) \tag{155}$$

Let us study the first term in eq. (155), and with i held fixed, define

$$m_i(z^{\underline{i-1}}) = m(z^{\underline{i-1}}, c_i, a^{i+\underline{k-i}}) \tag{156}$$

Expanding the first term within eq. (155), and because m_i as defined in eq. (156) is $i-1$-linear (because it is a k-linear map with $k-(i-1)$ variables held constant); we use eq. (142) again.

$$m_i(b^{\underline{i-1}}) = \left(\sum_{i=\underline{k}} m_i(b^{\underline{j}}, \Delta^j, a^{j+\underline{(i-1)-j}}) \right) + m_i(a^{\underline{i-1}}) \tag{157}$$

Unboxing the last term in eq. (157) using the definition of m_i reads

$$m(b^{\underline{i-1}}, \Delta^i, a^{i+\underline{k-i}}) = m(a^{\underline{i-1}}, c^i, a^{i+\underline{k-i}}) + \sum_{j=\underline{i-1}} m_i(b^{\underline{j}}, \Delta^j, a^{j+\underline{(i-1)-j}}) \tag{158}$$

We wish to remove all of the b^is. Since $\Delta^i = c^i + \varepsilon^i$ (eq. (153)), we have

$$\begin{aligned}
m(b^{\underline{k}}) - m(a^{\underline{k}}) &= \sum_{i=\underline{k}} m(b^{\underline{i-1}}, c^i, a^{i+\underline{k-i}}) + m(b^{\underline{i-1}}, \varepsilon^i, a^{i+\underline{k-i}}) \\
&= \left(\sum_{i=\underline{k}} m_i(b^{\underline{i-1}}) \right) + \sum_{i=\underline{k}} m(b^{\underline{i-1}}, \varepsilon^i, a^{i+\underline{k-i}}) \\
&= \left(\sum_{i=\underline{k}} m_i(a^{\underline{i-1}}) + \sum_{j=\underline{i-1}} m_i(b^{\underline{j-1}}, \Delta^j, a^{j+\underline{(i-1)-j}}) \right) + \sum_{i=\underline{k}} m(b^{\underline{i-1}}, \varepsilon^i, a^{i+\underline{k-i}}) \\
&= \left(\sum_{i=\underline{k}} m_i(a^{\underline{i-1}}) \right) + \sum_{\substack{i=\underline{k} \\ j=\underline{i-1}}} m_i(b^{\underline{j-1}}, \Delta^j, a^{j+\underline{(i-1)-j}}) + \sum_{i=\underline{k}} m(b^{\underline{i-1}}, \varepsilon^i, a^{i+\underline{k-i}})
\end{aligned} \tag{159}$$

The last term within eq. (159) is $o(h)$, since it is a linear combination of $o(h^i)$s.

$$|\sum_{i=\underline{k}} m(b^{\underline{i-1}}, \varepsilon^i, a^{i+\underline{k-i}})| \lesssim_{m,a,b} |o(h)| \tag{160}$$

Each summand in the second last term in eq. (159) is $o(h)$ as well, as

$$|m_i(\underline{b^{j-1}}, \Delta^j, a^{j+\underline{(i-1)-j}})| \le |m_i| \, \Pi \, (\underline{b^{j-1}}, \Delta^j, a^{j+\underline{(i-1)-j}}) \tag{161}$$

$$\le |m| \cdot \left(\Pi(c^i, a^{i+\underline{k-i}})\right)\left(\Pi(\underline{b^{j-1}}, \Delta^j, a^{j+\underline{(i-1)-j}})\right) \lesssim_{m,a,b} \sup_{\substack{i=\underline{k} \\ j=\underline{i-1}}} |c^i| \cdot |\Delta^j|$$

$$\lesssim_{m,a,b} \sup_{\substack{i=\underline{k} \\ j=\underline{i-1}}} |Df_i(x)(h)| \cdot |f_j(x+h) - f_j(x)|$$

$$\lesssim_{m,a,b} |Df_i(x)||h| \sup_{\substack{i=\underline{k} \\ j=\underline{i-1}}} |\Delta^j| \lesssim_{m,a,b} |o(h)| \tag{162}$$

for the second last estimate we used $\Delta^j \to 0$. Therefore the second term in eq. (159) is $o(h)$, and eq. (149) is proven. Therefore $m(f_{\underline{k}})$ is differentiable at x. Continuity of $Dm(f_{\underline{k}})$ follows from the fact that

$$Dm(f_{\underline{k}})(x) = \sum_{i=\underline{k}} m(f_{\underline{i-1}}(x), Df_i(x)(\cdot), f_{i+\underline{k-i}}(x)) \tag{163}$$

and each of the summands eq. (163) can be broken down as the product of the compositions shown in eqs. (164) and (165)

$$x \mapsto (f_{\underline{i-1}}(x), f_{i+\underline{k-i}}(x)) \mapsto m(f_{\underline{i-1}}(x), \cdot, f_{i+\underline{k-i}}(x)) \tag{164}$$

$$x \mapsto Df_i(x)(\cdot) \tag{165}$$

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

Corollary 1.1: Higher order product rule

Let $m : \prod F_{\underline{k}} \to G$ be a continuous k-linear map between $F_{\underline{k}}$ and G. Suppose $f_i \in C^p(E, F_i)$ for $i = \underline{k}$. Then $m(f_{\underline{k}})$ is in $C^p(E, G)$ as defined in eq. (147), and

$$D^p m(f_{\underline{k}})(x) = \sum_{|\alpha|=p} \left(\frac{p!}{\alpha!}\right) m(D^{\alpha_{i=\underline{k}}} f_i(x)) \tag{166}$$

where $D^0 f_i(x) = f_i(x)$, α is a k multi-index with entries $\alpha_i \ge 0$ and

$$\alpha! = \prod_{i=\underline{k}} \alpha_i!$$

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 $p_0 \leq p_1 \leq \cdots \leq p_N$.

$$P = (a = p_0, p_1, \ldots, p_N = b) \quad \text{or} \quad P = (p_0, p_{\underline{N}}) \tag{167}$$

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. (167), and $Q = (q_0, q_{\underline{N}})$ 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 $P = (p_0, p_{\underline{N}})$, 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 ith 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 $\text{cell}(P) = \Big(\text{cell}(p_{\underline{N}})\Big)$.
- 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,

$$\Sigma = \Big\{ f : [a, b] \to E, \text{ there exists a } N\text{-partition } P \in I_p,\ \{v_{\underline{N}}\} \subseteq E \text{ such that } f|_{(p_{i-1}, p_i)} = v_i \ \forall i = \underline{N} \Big\} \tag{168}$$

I $f \in \Sigma$, 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. (168), we define the integral of f by

$$\int_a^b f = \sum_{i=\underline{N}} (p_i - p_{i-1}) v_i \tag{169}$$

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. (169) 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
$$\int \hat{\sum} f_{\underline{k}} = \hat{\sum} \int f_{\underline{k}} \tag{170}$$
- The integral over $[b, a]$ is *defined* to be the negative of eq. (169):
$$\int_a^b f = -\int_b^a f \tag{171}$$
- The integral is domain-additive, if $b = c$, then
$$\int_a^b f + \int_c^d g = \int_a^d (f + g) \tag{172}$$
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. 16.1

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

There are two ways of defining the space of step-mappings from I into E eqs. (173) and (174). Using a combinatorical argument with common refinements, it is not hard to see the two are subsets of each other.

$$\Sigma_E^1 = \left\{ f : I \to E,\ \mathrm{proj}_i\, f \in \Sigma_i\ \forall i = \underline{k} \right\} \tag{173}$$

$$\Sigma_E^2 = \left\{ f : I \to E,\ f \text{ is a step mapping.} \right\} \tag{174}$$

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

$$\int_a^b f = \left(\int_a^b \mathrm{proj}_{\underline{k}} f\right) = \left(\int_a^b \mathrm{proj}_1 f, \ldots, \int_a^b \mathrm{proj}_k f\right) \tag{175}$$

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 $f : I \to E$ 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 $P = (p_0, p_{\underline{n}})$ is a partition.

We construct a step mapping by sampling values of f. Set $g|_{\mathrm{cell}(p_i)} = f(p_i)$, $g(a) = f(a)$, $g(p_i) = f(p_i)$. Defining the endpoints is necessary, and g still remains a member of Σ_I by eq. (168). Each $x \in I \setminus P$ belongs in some 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 $f : I \to E$ 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 $f : I \to E$,

$$|\int_a^b f| \leq \int_a^b |f| \leq (b-a)\|f\|_u \tag{176}$$

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. (176) holds. If f is induced by some n-partition P,

$$\int_a^b f = \sum_{i=\underline{n}} |\mathrm{cell}(p_i)| f(p_i) \leq \sum_{i=\underline{n}} |\mathrm{cell}(p_i)||f(p_i)| = \int_a^b |f| \tag{177}$$

The integral in eq. (177) should be interpreted as a Riemann integral on $\mathbb{R}$, and eq. (178) is immediate:

$$\int_a^b |f| \leq |b-a|\|f\|_u \tag{178}$$

Next, let $\{f_n\}_{n\geq 1}$ be a sequence of step mappings in I which converges uniformly to $f \in \overline{\Sigma}$. Equation (178) tells us the sequence of integrals is uniformly Cauchy, as

$$|\int_a^b f_m - \int_a^b f_n| \leq |b-a| \|f_m - f_n\|_u \tag{179}$$

Hence $\int_a^b f$ is well defined, eq. (176) 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 $f \in \overline{\Sigma_E}$, then $\{\lambda f_n\} \to \lambda f$ uniformly in $\overline{\Sigma_F}$. Moreover,

$$\lambda\left(\int_a^b f\right) = \int_a^b \lambda f \tag{180}$$

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. (180) for step mappings, and the general case follows from continuity. ■

Fundamental Theorem of Calculus

Proposition 5.1

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

$$\varphi(t) = \int_a^t f \tag{181}$$

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

$$(D\varphi)(t_0) = f(t_0) \tag{182}$$

Remark 5.1: Identifications

The left hand side in eq. (182) 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 $\varepsilon(h) \to 0$ as $h \to 0$. Applying the well-known technique of estimating the integrand $f(t) = [f(t) - f(t_0)] + f(t_0)$, we have

$$\begin{aligned} \varphi(t_0 + h) - \varphi(t_0) &= \int_{t_0}^{t_0+h} f(t)dt \\ &= f(t_0) \cdot h + \int_{t_0}^{t_0+h} [f(t) - f(t_0)]dt \end{aligned} \tag{183}$$

The last term within eq. (183) 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 $L(E, F) \times E \to F$ is bilinear and continuous. For any regulated mapping $\alpha : I \to L(E, F)$,

$$\int_a^b \alpha(t) \cdot x dt = \left(\int_a^b \alpha(t)dt \right) \cdot x \tag{184}$$

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. 16.1 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

$$f(b) - f(a) = \int_a^b f'(t)dt \tag{185}$$

where the integrand in eq. (185) is — rigourously speaking — a map $\mathbb{R} \to L(\mathbb{R}, E)$, but we treat $f'(t) \in E$.

Proof. Throughout this proof, we will treat $f' : \mathbb{R} \to E$. Because f' is continuous everywhere, it is regulated. Define $\varphi(t) = \int_a^t f'(t)dt$, by eq. (182):

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

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. (185) 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 $U \overset{\circ}{\subseteq} E$ 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. (186) holds.

$$f(x+y) = f(x) + \int_0^1 Df(x+ty)ydt = \left(\int_0^1 Df(x+ty)dt\right) \cdot y \tag{186}$$

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, $g(1) - g(0) = \int_0^1 Df(x+ty) \cdot ydt$. 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

$$X = L(E, L(E, L(E, \cdots F))p \text{ times} \overset{\mathcal{L}}{\rightleftharpoons} L(E^p, F)$$

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 $\omega : E_1 \times E_2 \to F$ 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 $D^2A(x) = D^3\omega = 0$.

■

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

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

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

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

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

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
$$\Delta = g(x+w) - g(x) \tag{187}$$
- Again, for y sufficiently close to x, define $h(y) = f(y+w) - f(y)$, and
$$\Delta = h(x+v) - h(x) \tag{188}$$
- 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
$$Dg(y) = D(\tau_{-w} f)(y) - Df(y) = Df(y+w) - Df(y) \tag{189}$$

Using MVT twice, once on Equation (187) (the line segment $x + tw$, $0 \le t \le 1$ is contained in the domain of g), and another time on Equation (189) (with $y = x + tw$ in the integrand). We obtain:

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

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

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

Note 2.2

The error term δ_1 in the integrand is given by

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

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

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

$$D^2 f(x) \cdot (v, w) - D^2 f(x) \cdot (w, v) = \int_0^1 \int_0^1 \delta_1(tw, sv) ds\, dt - \int_0^1 \int_0^1 \delta_2(tw, sv) ds\, dt \tag{190}$$

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 (190) 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

$$|\int_0^1 \int_0^1 \delta(tw, sv) ds\, dt| \leq \int_0^1 \int_0^1 |\delta(tw, sv)| ds\, dt \tag{191}$$

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

$$|\delta_j| \leq |D^2 f(x + tw + sv) - D^2 f(x)||v||w| \tag{192}$$

- 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 $D^2 f$ means

$$R(v, w, \delta) = \sup_{y \in A} |D^2 f(x + y) - D^2 f(x)| \quad \text{is finite,} \quad \text{and} \quad \lim_{(v,w) \to 0} R(v, w, \delta) = 0 \tag{193}$$

  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 (191) to (193), we obtain the following bound on Equation (190)

$$\begin{aligned} |D^2 f(x) \cdot (v,w) - D^2 f(x) \cdot (w,v)| &\leq |\iint \delta_1(tw, sv) ds\, dt - \iint \delta_2(tw, sv) ds\, dt| \\ &\leq \iint |\delta_1| ds\, dt + \iint |\delta_2| ds\, dt \\ &\leq |v||w| R(v,w) \end{aligned} \tag{194}$$

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 $\theta : E^k \to F$ (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 $\lim_{(u_{\underline{k}}) \to 0} \theta(u_{\underline{k}}) = 0$, 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

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

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

■

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 $f, Df, D^2 f, \ldots, D^p f$ 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 $(v_2, \ldots, v_p) = (v_{1+\underline{p-1}})$,

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

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

- the transposition $(1, 2, \ldots) \mapsto (2, 1, \ldots)$ 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 $\lambda : L(E^{p-2}, F) \to F$ be the evaluation map at $(v_3, \ldots, v_p) = (v_{2+\underline{p-2}})$. Using the base case on $D^{p-2} f = g \in C^2(E, L(E^{p-2}, F))$,

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

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

$$(D^2 g)(x)(v_1, v_2) \cdot (v_{2+\underline{p-2}}) = (D^2 g)(x)(v_2, v_1) \cdot (v_{2+\underline{p-2}}) \tag{195}$$

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

$$\lambda(D^2(D^{p-2} f)) = D(\lambda(D(D^{p-2} f)) = D(D\lambda \circ D^{p-2} f) = D^2(\lambda \circ D^{p-2} f) \tag{196}$$

Substituting Equation (195) for the rightmost hand side of Equation (196) 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 $(D^2 g)(x) \in L(E^2, L(E^{p-2}, F))$. 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

$$\begin{aligned} \big(\,\theta(\underline{p})\,\big) &= \big(1, \theta(1+\underline{p-1})\big) \\ &= \big(1, 1+\beta(\underline{p-1})\big) \end{aligned} \tag{197}$$

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:

$$\begin{aligned} (D^p f)(x)(v_{\underline{p}}) &= D\big(\lambda D^{p-1} f\big)(x)(v_1) && \text{Equation (196)} \\ &= D\big(\lambda' D^{p-1} f\big)(x)(v_1) && \text{ind. hyp.} \\ &= (D^p f)(x)(v_{\theta(\underline{p})}) && \text{Equation (196)} \end{aligned}$$

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 post-composing any multi-linear map $A \in L(E^p, F)$ by λ. Denoting this map by λ_*,

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

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

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

Now, recall that for $p = 1$

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

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

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

Dropping the x from the expression, for $p \geq 2$ *assuming a similar formula holds*, then we write $\big[D^p(\lambda \circ f)\big] = \lambda_* \circ D^p f$. 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: CLMs commute past 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 in contained in U, then

$$f(x+y) = f(x) + \left(\sum_{i=\underline{p-1}} \frac{D^i f(x) \cdot (y^{(i)})}{(p-1)!} \right) + R_p \tag{198}$$

where $\cdot(y^{(i)})$ denotes the consecutive application of y for i times. The remainder R_p is given by eq. (199)

$$R_p = \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p f(x+ty) dt \cdot (y^{(p)}) \tag{199}$$

Furthermore, we include the pth term in the series using eq. (200)

$$f(x+y) = f(x) + \sum_{i=\underline{p}} \frac{D^i f(x) \cdot (y^{(i)})}{i!} + \theta(y) \tag{200}$$

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

$$|\theta(y)| \leq \sup_{0 \leq t \leq 1} \frac{|D^p f(x+ty) - D^p f(x)|}{p!} |y|^p \tag{201}$$

Proof. Postponed. ■

Bibliography

[1] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer Science & Business Media, 1 edition, nov 10 2010.

[2] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, jun 11 2013.

[3] Serge Lang. *Fundamentals of Differential Geometry*. Springer Science & Business Media, dec 6 2012.

[4] John M. Lee. *Introduction to Smooth Manifolds*. Springer Science & Business Media, 2 edition, mar 9 2013.

[5] John M. Lee. *Introduction to Riemannian Manifolds*. Springer, 2 edition, jan 2 2019.

[6] Steven Roman. *Advanced Linear Algebra*. Springer Science & Business Media, 3 edition, dec 31 2007.