Math Notes

Nicholas Lyu

Nov 27, 2023

Contents

1 Multilinear algebra

1.1 Module[1]

Recall that a ring R is a multiplicative monoid with identity 1 and an additive abelian group such that multiplication distributes on both sides over addition. A R-module A is an additive abelian group which admits ring action by R.

Definition 1.1 (*R-module*). A (left) R-module A is an additive abelian group with scalar multiplication $R \times A \to A$, written $(\kappa, a) \mapsto \kappa a$ such that

- $\forall \kappa \in R : A \xrightarrow{\kappa(\,\cdot\,)} A$ is a morphism of addition in A: $\kappa(a+b) = \kappa a + \kappa b$
- $\forall a \in A : R \xrightarrow{(\,\cdot\,)a} A$ is a morphism of addition: $(\kappa + \lambda)a = \kappa a + \lambda a$
- Scalar multiplication is monoid action: $(k\lambda)a = k(\lambda a), \quad 1a = a$

Definition 1.2 (*Free module*). A subset X of a R-module f with inclusion $\iota : X \to F$ freely generates X if to every function $f : X \to A$ there is exactly one linear map $t : F \to A$ such that the diagram below commutes. In other words, $\mathbf{Set}(X, U(A)) \cong \mathbf{Mod}(F, A)$.

$$\begin{tikzcd} X \arrow[r, "\iota"] \arrow[dr, "f"'] & F \arrow[d, dashed, "\exists!\, t"] \\ & A \end{tikzcd}$$

Definition 1.3 (*universal bilinear function*). Given K-modules A, B over a commutative ring K, a universal bilinear function $h_0 : A \times B \to D$ satisfies

$$\mathbf{Bilin}_K(A, B; C) \cong \mathbf{Mod}_K(D, C)$$

The universal bilinear function is the tensor product with $D = A \otimes B$. Equivalently,

$$\begin{tikzcd} A \times B \arrow[r, "h_0"] \arrow[dr, "f"'] & D \arrow[d, dashed, "\exists!\, g"] \\ & C \end{tikzcd}$$

Theorem 1.1 (*universality of scalar multiplication*). Let A be a K-module over a commutative ring K. Scalar multiplication $A \otimes K \xrightarrow{(a,\kappa)\mapsto a\kappa} A$ is the universal bilinear function.

Proof: consider a bilinear function $f : A \times K \to C$, where C is a K-module. Define $g : A \to B$ via $g(a) = f(a, 1)$. Then

$$h(a, \kappa) = h(a, 1)\kappa = g(a)\kappa = g(a\kappa) = (g \circ h_0)(a, \kappa)$$

[1] Maclane *Algebra*, V.1, IX.7

Corollary 1.1. Every K-module A is isomorphic to $A \otimes K$ by the following isomorphism.

$$\phi : A \otimes K \xrightarrow{(a,\kappa)\mapsto a\kappa} A$$

Proof: the universal bilinear map is unique up to isomorphism. Equivalently, every bilinear map from $A \times K \to C$ is equivalent to a linear map $A \to C$.

1.2 Tensor product via adjoints[2]

We construct the tensor product $A \otimes B$ as the codomain of the universal bilinear function on $A \times B$. We begin by constructing the free module F over $A \times B$.

1. Consider the set of functions $f : A \times B \to K$ with a finite number of nonzero values. These functions form a K-module under termwise addition and scalar multiplication.
2. Let $[a, b] \in F$ denote the special function that is 1 on (a, b) and 0 elsewhere. Then every function f is a finite linear combination of $\{[a \in A, b \in B]\}$.
3. F is free over $A \times B$ by definition 1.2: every set function $A \times B \xrightarrow{f} C$ induces a unique module morphism $\tilde{f}(u(a, b)) = f(a, b)$ for $u : A \times B \xrightarrow{[\cdot,\cdot]} F$.

Now u is not bilinear since $[a, b] + [a', b] \neq [a + a', b] \in F$. As of now, they are totally unrelated basis elements. We construct the tensor product by identifying these elements.

Definition 1.4 (*tensor product*). Let $S \subset F$ be the submodule spanned by

$$\begin{aligned}
&[a_1\kappa_1 + a_2\kappa_2, b] - [a_1, b]\kappa_1 - [a_2, b]\kappa_2 \\
&[a, b_1\kappa_1 + b_2\kappa_2] - [a, b_1]\kappa_1 - [a, b_2]\kappa_2
\end{aligned}$$

for every choice of κ_i and elements in A, B. Define the tensor product space and operation

$$A \otimes B = F/S, \quad a \otimes b = p([a, b])$$

The tensor product map is bilinear (linearity in the second argument follows similarly):

$$\begin{aligned}
(a_1\kappa_1 + a_2\kappa_2) \otimes b &= p\,[a_1\kappa_1 + a_2\kappa_2, b] \\
&= p\,(([a_1\kappa_1 + a_2\kappa_2, b] - [a_1, b]\kappa_1 - [a_2, b]\kappa_2) + [a_1, b]\kappa_1 + [a_2, b]\kappa_2) \\
&= p\,\,([a_1\kappa_1 + a_2\kappa_2, b] - [a_1, b]\kappa_1 - [a_2, b]\kappa_2) + p([a_1, b])\kappa_1 + p([a_2, b])\kappa_2 \\
&= (a_1 \otimes b)\kappa_1 + (a_2 \otimes b)\kappa_2
\end{aligned}$$

The following proposition is among the main reasons we study bilinear maps:

Proposition 1.2. $\mathbf{Bilin}(A, B; C)$ is a K-module under pointwise sum and scalar multiples.

$$\mathbf{Bilin}(A, B; C) \cong \mathbf{Hom}(A, \mathbf{Hom}(B, C))$$

[2] Maclane *Algebra*, IX.8

Proof: given bilinear $f : A \times B \to C$, the partial application map $f(\,\cdot\,, -)$ in the first argument is of type $A \to \mathbf{Hom}(B, C)$ by the linearity of f in the second argument. The assignment $f \mapsto f(\,\cdot\,, -)$ to partial application in the first argument is also linear since

$$(f + g\alpha)(\,\cdot\,, -) = f(\,\cdot\, -) + g(\,\cdot\,, -)\alpha$$

By partial application in the second argument we also have

$$\mathbf{Bilin}(A, B; C) \cong \mathbf{Hom}(B, \mathbf{Hom}(A, C))$$

The following universal characterization of tensor products is preferred, as it highlights the similarity to the Cartesian product construction in **Set**.

Theorem 1.3 (*tensor*-Hom *adjunction*). Fixing a module B,

$$(- \otimes B) \dashv \mathbf{Hom}(B, -) : \mathbf{Mod}_K \leftrightarrows \mathbf{Mod}_K$$

Proof: the proposition above, with the universal bilinear property of the tensor product, then the following bijection is natural in A and C:

$$\mathbf{Hom}(A \otimes B, C) \cong \mathbf{Hom}(A, \mathbf{Hom}(B, C))$$

1.3 Algebra[3]

An algebra has compatible ring and module structures.

Definition 1.5 (K-*algebra*). Let K be a commutative ring, a linear algebra A over the commutative ring K of scalars, abbreviated K-algebra A, is a right K-module which is also a ring over addition and, for every $a_1, a_2 \in A, \kappa \in K$:

$$(a_1 a_2)\kappa = (a_1 \kappa) a_2 = a_1(a_2 \kappa)$$

Scalar multiplication on A (as a module) is compatible with ring multiplication in R.

Example 1.1 (*examples of algebras*). Let K denote a commutative ring.

- *$n \times n$ matrices:* They form a K-module by scalar multiplication and element-wise addition and a ring by matrix multiplication.
- *Endomorphisms:* Fix a K-module C. The set $\mathrm{End}_K(C) = \mathbf{Mod}_K(C, C)$ of module endomorphisms is a module by pointwise addition and scalar multiplication as well as a ring by composition.
- *Commutative algebras:* The polynomial rings $K[x]$ is a module by K-scalar multiplication and addition and a ring under polynomial product.
- *Rings with center containing K:* The center K of a ring R is its largest commutative subring. Every R is a K-algebra with the algebra axiom giving commutativity.

[3]Maclane *Algebra*, IX.12

- $\mathbb{R}$-*algebra* The complex $\mathbb{C}$ and quaterions $\mathbb{Q}$ are algebras over $\mathbb{R}$. They are additionally division algebras (module + division ring).

Proposition 1.4. In a K-algebra A, ring multiplication is a K-bilinear map.

Proof: Distributive law for ring multiplication, and algebra definition.

$$a(a_1 + \kappa a_2) = aa_1 + a(\kappa a_2) = aa_1 + (aa_2)\kappa$$

By the universal property of tensor products, there is an unique linear map $A \otimes A \xrightarrow{\pi} A$ such that $a_1 a_2 = \pi(a_1 \otimes a_2)$. Then definition 1.5 can be expressed diagramatically.

Theorem 1.5 (*diagramatic characterization of K-algebra*). Let A be a K-module equipped with two K-linear maps

$$\pi : A \otimes A \to A, \quad u : K \to A$$

such that the diagrams below commute (recall the isomorphism ϕ in corollary 1.1)

$$\begin{tikzcd} A \otimes A \otimes A \arrow[r, "1\otimes\pi"] \arrow[d, "\pi\otimes 1"'] & A \otimes A \arrow[d, "\pi"] \\ A \otimes A \arrow[r, "\pi"] & A \end{tikzcd} \qquad \begin{tikzcd} K \otimes A \arrow[r, "\phi'"] \arrow[d, "u\otimes 1"'] & A \arrow[d, "1_A"] & A \otimes K \arrow[l, "\phi"'] \arrow[d, "1\otimes u"] \\ A \otimes A \arrow[r, "\pi"] & A & A \otimes A \arrow[l, "\pi"'] \end{tikzcd}$$

Then A is a K-algebra with product $a_1 a_2 = \pi(a_1 \otimes a_2)$ and multiplicative unit $u(1)$.

Proof: The diagram on the left characterizes associativity, making A into a monoid.

$$[\pi \circ (1 \otimes \pi)](a_1, a_2, a_3) = \pi(a_1, a_2 a_3) = a_1(a_2 a_3)$$
$$[\pi \circ (\pi \otimes 1)](a_1, a_2, a_3) = \pi(a_1 a_2, a_3) = (a_1 a_2)a_3$$

The K-linearity of π implies distributivity, making A into a ring. K-linearity of u implies that it is completely determined by $u(1)$ and $u(\kappa) = u(1)\kappa$. Commutative paths for the diagram on the right then reads

$$(1_A \circ \phi')(\kappa, a) = \kappa a$$
$$(1_A \circ \phi)(a, \kappa) = a\kappa$$
$$(\pi \circ (u \otimes 1))(\kappa, a) = \pi(u(1)\kappa, a) = (u(1)\kappa)a$$
$$(\pi \circ (1 \otimes u))(a, \kappa) = \pi(a, u(1)\kappa) = a(u(1)\kappa)$$

The second diagram captures exactly what it means for scalar and ring multiplication to be "compatible": scalar multiplication $a\kappa$ may be computed directly via $\phi(a, \kappa)$ via the module structure of A, or by embedding κ into a ring element $u(\kappa)$ then taking the product: $\pi(a \otimes u(\kappa))$ The second commutative diagram dictates that these two results are equal. This establishes the algebra axioms together with associativity.

Equivalently, A is an algebra if ring multiplication given by π satisfies the diagram on the right, i.e. scalar elements are"embedded" via u such that scalar multiplication and embedded multiplication are compatible.

1.4 Tensors[4]

We consider finite-dimensional vector space V over a field F. Tensors are elements in iterated products of V and V^*, regarded as transformations.

- Elements of V are contravariant tensors since $V \cong \mathbf{Hom}(V^*, F)$ transforms contravariantly. Similarly, $V^* \cong \mathbf{Hom}(V, F)$ are denoted covariant tensors.
- Covariance / contravariance describes the transformations, not elements. Distinguish between the types of $v \in V$ (covariant) and $v \in \mathbf{Hom}(V^*, F)$ (contravariant).
- Let $\mathbf{b} = \{b_i\}$ be a basis for V. Denote by $\{b^i\}$ the dual basis such that $b^i b_j = \delta^i_j$.
- Every $v \in V$ has coordinate representation $\{\xi^i = b^i\, v\}$ with repsect to $\mathbf{b}$ such that $v = \xi^i b_i = (b^i\, v) b_i$, where the sum is understood. Similarly, $\omega \in V^*$ has coordinate representation $\{\eta_i = \omega\, b_i\}$ such that $\omega = \eta_i b^i = (\omega\, b_i) b^i$
- Contravariant bases are written with subscripts (b_i as a basis in $V \cong V^{**}$).
- Contravariant components are written with superscripts ($\omega^i \in F$ as projection $b^i\omega$).
- Covariant bases are written with superscripts (b^i as a basis of V^*).
- Covariant components are written with subscripts ($v_i \in F$ as $b_i v$).

Consider the effect of a basis change $b \to c$, such a change is characterized by a $n \times n$ matrix with scalar entries $P^j_i = c^j b_i$. The contravariant basis transforms via

$$b_i = c_j (c^j b_i) = c_j P^j_i$$

Correspondingly, the contravariant components transform as

$$\omega = \omega^i_b b_i = \omega^i_b c_j P^j_i = (\omega^i_b P^j_i) c_j \implies \omega^i \mapsto \omega^i P^j_i$$

For a 2-contravariant tensor, the components transform as $\rho^{kl}_c = P^k_i P^l_j \rho^{ij}_b$.

$$r = (b_i \otimes b_j)\rho^{ij}_b = \left[(c_k P^k_i) \otimes (c_l P^l_j)\right] \rho^{ij}_b = (c_k \otimes c_l) \left(P^k_i P^l_j \rho^{ij}_b\right)$$

Theorem 1.6 (*contravariant component transformation rule*). Under a coordinate change $b \to c$ via $P^j_i = c^j b_i$, the components of a contravariant tensor $\omega \in V^{\otimes p}$ transform as

$$\omega_c^{j_1 \cdots j_p} = \omega_b^{i_1 \cdots i_p} \left(P^{j_1}_{i_1} \cdots P^{j_p}_{i_p}\right)$$

We next consider $b \to c$ effecting the dual (covariant) basis. Let $[Q^k_j] = [P^j_i]^{-1}$ so that

$$P^j_i Q^k_j = Q^k_j P^j_i = \delta^k_i$$

The covariant basis transforms as

$$b^i = b^j \delta^i_j = b^j P^k_j Q^i_k = c^k Q^i_k$$

Correspondingly, the covariant components transform as

$$v = (b_i\, v) b^i = (b_i\, v) c^k Q^i_k = c^k (v_i Q^i_k) \implies v_i \mapsto v_i Q^i_k$$

[4] Maclane *Algebra*, XVI.2

Theorem 1.7 (*component transformation rule*). Under a basis change $b \to c$ via $P_i^j = c^j b_i$ with inverse Q_j^k, the components $\alpha_{i_1 \cdots i_m}^{j_1 \cdots j_n}$ of a i-covariant, j-contravariant tensor transform as

$$\bar{\alpha}_{j_1 \cdots j_m}^{i_1 \cdots i_n} = \alpha_{h_1 \cdots h_m}^{l_1 \cdots l_n} \left(P_{l_1}^{i_1} \cdots P_{l_n}^{i_n}\right) \left(Q_{j_1}^{h_1} \cdots Q_{j_n}^{h_n}\right)$$

Proposition 1.8. $\mathbf{Hom}(A, A') \otimes \mathbf{Hom}(B, B') \cong \mathbf{Hom}(A \otimes B, A' \otimes B')$

Proof: δ_j^i forms a basis for $\mathbf{Hom}(A, A')$. A general element on the left hand side is

$$\eta_{ij}^{kl}(\delta_k^i \otimes \delta_l^j) \in \mathbf{Hom}(A, A') \otimes \mathbf{Hom}(B, B')$$

This map sends $a_i \to a'_k, b_j \to b'_l$. Define $A \otimes B \xrightarrow{F} A' \otimes B'$ by

$$F(a_i \otimes b_j) = \eta_{ij}^{kl}(a'_k \otimes b'_l)$$

1.5 Graded module[5]

Let K be a commutative ring and let module denote a K-module.

Definition 1.6 (*graded module*). A graded module G is sequence of modules indexed by $\mathbb{N}$.

$$G = (G_0, G_1, \cdots)$$

- A morphism $t : G \to G'$ of graded modules is a sequence of module morphism $G_p \to G'_p$.
- An element $g = G_p \in G$ is a module with a distinguished degree.
- Sums in G are only defined between elements of the same degree (homogenous).
- A graded submodule $S \subseteq G$ is a sequence of submodules $S_p \subseteq G_p$. Graded submodules of G are partially ordered by inclusion.
- Given a morphism $t : G \to G'$ of graded modules, its image is a graded submodule, as is its kernel. The quotient graded module is a sequence of quotient modules.

Definition 1.7 (*tensor product*). The graded tensor product of $G, H \in \mathbf{GrMod}$ is

$$(G \otimes H)_n = \bigoplus_{p+q=n} (G_p \otimes H_q) = (G_n \otimes H_0) \oplus \cdots \oplus (G_0 \otimes H_n)$$

The tensor product is associative, $(G \otimes H) \otimes M = G \otimes (H \otimes M)$. If G_n, H_n are all of dimension n, then $(G \otimes H)_n$ has dimension n^2.

Theorem 1.9 (*universality of graded tensor product*). Given graded modules G, H, M. For every family of bilinear maps $t_{p,q} : G_p \times H_q \to M_{p+q}$, there is exactly one morphism $t : G \otimes M \to M$ of graded modules with $t(g \otimes h) = t_{p,q}(g, h)$.

Proof: By the universality of tensor products, for each p, q there is unique module morphism $G_p \otimes H_q \xrightarrow{s_{p,q}} M_{p+q}$ with $s_{p,q}(g \otimes h) = t_{p,q}(g, h)$. Fixing n, there are n maps $\{s_{0,n} \cdots s_{n,0}\}$ which with codomain M_n. By the universality of insertions into biproducts $(G_p \otimes H_q) \xrightarrow{\iota} (G \otimes H)_{p+q}$, the factor map t_n is unique.

[5]Maclane *Algebra*, XVI.3

$$
\begin{tikzcd}
G_0 \otimes H_n \arrow[dr, "\iota"'] \arrow[drr, "s_{0,n}"] & & \\
\vdots & (G \otimes H)_n \arrow[r, dashed, "t_n"] & M_n \\
G_n \otimes H_0 \arrow[ur, "\iota"] \arrow[urr, "s_{n,0}"'] & &
\end{tikzcd}
$$

Definition 1.8 (*tensor product functor on graded modules*). The tensor product $\otimes$ is a bifunctor on **GrMod** by defining $G \otimes H$ via definition 1.7 and the tensor product of morphisms $(G \xrightarrow{u} G') \otimes (H \xrightarrow{v} H')$ the map $G \otimes H \to M$ defined via theorem 1.9 using $\{t_{p,q} = u_p \otimes v_q\}$.

- A graded module is "concentrated in degree 0" when $H_{n>0}$ is trivial. In other words, H is the sequence $(H_0, 0, \cdots)$; we identify the graded module with the module H_0.
- Let K denote the ring of scalars regarded as a graded module concentrated in degree 0. In particular, for every graded module G we have $G \otimes K \cong G$ by corollary 1.1.

Given G, consider all sequences $g = (g_0, \cdots)$ of elements with $g_n \in G_n$. Define the support of each sequence to be the set of all indices p with $f_p \neq 0$.

Definition 1.9 (*internally graded modules*). Each graded module G determines a module ΣG consisting of sequenes with finite support, with module operations defined pointwise.

$$
\begin{aligned}
(f + g)_p &= f_p + g_p \\
(f\kappa)_p &= f_p \kappa
\end{aligned}
$$

Every morphism of graded module $G \xrightarrow{t} G'$ induces a morphism of internally graded modules

$$
\Sigma G \xrightarrow{\Sigma t} \Sigma G', \quad (\Sigma t)\, f = \sum_p t\, f_p
$$

Moreover, $\Sigma : \mathbf{GrMod} \to \mathbf{Mod}$ is a functor.

In general, ΣG contains much more elements than G. As a set, G is like the disjoint union of $\{G_p\}$, while ΣG is its product (up to subtleties of finiteness).

Theorem 1.10 (*characterization of internally graded modules*). Call $S \in \mathbf{Mod}$ the sum of a sequence of its submodules $\{D_p\}_{p \in \mathbb{N}}$, when every $s \neq 0$ may be uniquely written as a finite sum with elements in D_p. A module M is an internally graded module if and only if there exists such a sequence of submodules which sum to M.

1.6 Graded algebra

Recall the definition 1.5 of algebra and its characterization 1.5.

Definition 1.10 (*graded algebra*). A graded algebra A is a graded module with

$$\pi : A \otimes A \to A$$
$$u : K \to A$$

both morphisms of graded modules, such that the algebra diagrams in 1.5 commute.

In particular, that u is a morphism of graded module implies $u(\kappa) \in A_0$ since morphisms preserve degree and $K \cong K_0$. Similarly, $a \otimes b$ is of degree $\deg a + \deg b$ by definition 1.7 of graded tensor product, so $\deg(ab) = \pi(a \otimes b) = \deg a + \deg b$.

Theorem 1.11 (*equivalent definition of graded algebra*). A graded module A is a graded algebra if and only if it is a ring with multiplication such that

$$\begin{aligned} \deg(ab) &= \deg a + \deg b \\ a(b_1 + b_2) &= ab_1 + ab_2, (b_1 + b_2)a = b_1 a + b_2 a, \qquad \text{if } \deg b_1 = \deg b_2 \\ (ab)\kappa &= a(b\kappa) = (a\kappa)b \end{aligned}$$

Proof: Given π, u, the product defined by $ab = \pi(a \otimes b)$ and scalar multiple $a\kappa = au(\kappa)$ satisfy these conditions. Conversely, these conditions state that $(a, b) \mapsto ab$, for each pair of degrees p, q, is a bilinear function $\pi_{p,q} : A_p \times A_q \to A_{p+q}$.

Example 1.2 (*Grassmann algebra*). The Grassmann algebra $G^{(1)}$ on one generator is the graded module

$$G^{(1)} = (K, Ke, 0, \cdots)$$

with unit 1 and product given by the product in K, bilinearity, and $1e = e1 = e, ee = 0$. The Grassmann algebra $G^{(2)}$ one two generators e, f is the graded module

$$G^{(2)} = (K, Ke \oplus Kf, Kef, 0, \cdots)$$

with unit 1 and multiplication given by the product in K, bilinearity, and

$$1e = e1 = e, \quad 1f = f1 = f, \quad e^2 = f^2 = 0, \quad fe = -ef$$

The simplest Grassmann algebra $G^{(1)}$ consists of linear combinations of $1 \in K = G^{(1)}_0$ and $\kappa e \in Ke = G^{(1)}_1$. The product of any two degree-1 elements vanish. Similarly, $G^{(2)}$ consists of linear combinations of 1 of degree 0, $\kappa_1 e + \kappa_2 f$ of degree 1, and $\kappa_3 ef$ of degree 2. Elements with degree > 2 vanishes, so the only interesting products are between degree-1 elements:

$$(\kappa_1 e + \kappa_2 f)(\kappa_3 e + \kappa_4 f) = \kappa_2\kappa_3 fe + \kappa_1\kappa_4 ef = (\kappa_1\kappa_4 - \kappa_2\kappa_3)ef$$

Example 1.3 (*graded polynomial algebra*). The graded polynomial algebra $P^{(1)}_K$ on 1 generator x of degree 2 is the graded module with $P_{2n+1} = 0$ and $P_{2n} = Kx_n$ for all n

$$P^{(1)}_K = (Kx_0, 0, Kx_1, 0, Kx_2, 0, \cdots)$$

with the product given by $x_n x_m = x_m x_n = x_{m+n}$. The graded polynomial algebra $P_k^{(2)}$ on 2 generators, each of degree 2, consists of the graded module with P_{2n+1} zero and each P_{2n} freely generated by the $n+1$ free generators $e_{n,0}, e_{n-1,1}, \cdots e_{0,n}$ denoting half-degree in each generator. The product $e_{p,q}e_{r,s} = e_{r,s}e_{p,q} = e_{p+r,q+s}$ defines the desired bilinear product. Let $x = e_{1,0}$ and $y = e_{0,1}$, we have $e_{p,q} = x^p y^q$ and every element of P_{2n} is of the form

$$\kappa_n x^n + \kappa_{n-1}x^{n-1}y + \kappa_{n-2}x^{n-2}y^2 + \cdots + \kappa_0 y^n$$

Note that elements of the polynomial algebra are monomials (one element of a specific degree). To describe the polynomials, we could have also chosen $P_K^{(1)} = (Kx_0, Kx_1, \cdots)$. The degree 2 of the generator is by convention for the following definition:

Definition 1.11 (*(skew) commutative graded algebra*). A graded algebra is commutative when for all elements $a_{\deg a}, b_{\deg p}$, we have

$$a_p b_q = (-1)^{pq} b_q a_p$$

Grassmann and graded polynomial algebra are both commutative algebra.

Definition 1.12 (*graded algebra of multilinear forms*). Fixing a module C over K, let C^p be the p-fold biproduct and each graded component $M_p(C)$ the set of all multilinear forms $C^p \to K$. It is a K-module under pointwise addition and scalar multiplication. Extend to $M_0(C) = (\mathbf{1} \to K) \cong K$. The graded sequence

$$(M_0(C), M_1(C), \cdots)$$

form a graded module with the product defined by

$$(h_p k_q)(c_1, \cdots, c_{p+q}) = h(c_1, \cdots, c_p)k(c_{p+1}, \cdots, c_{p+q})$$

This product is an element of $M_{p+q}(C)$, associative, bilinear in h, k, and makes $M(C)$ the graded algebra of multilinear forms on C.

Definition 1.13 (*morphism of graded algebra*). A morphism $t : A \to B$ of graded algebras is a morphism of graded modules which additionally respect the algebraic ring structure

$$t_{p+q}(ab) = (t_p a)(t_q b)$$
$$t_0(1_A) = 1_B$$

this is seen to be equivalent to

$$\begin{tikzcd} A \otimes A \arrow[r, "\pi"] \arrow[d, "t\otimes t"'] & A \arrow[d, "t"] \\ B \otimes B \arrow[r, "\pi"] & B \end{tikzcd} \qquad \begin{tikzcd} K \arrow[r, "u"] \arrow[d, "1"'] & A \arrow[d, "t"] \\ K \arrow[r, "u"] & B \end{tikzcd}$$

1.7 Universal graded constructions

Recall that an algebra is a ring. The kernel of a graded algebra morphism $A \xrightarrow{t} A'$ is a (two-sided) ring ideal in A. Since an graded algebra is also a graded module, the kernel is also a graded submodule.

Definition 1.14 (*ideal of a graded algebra*). An ideal D in a graded algebra A is a graded submodule of A (concretely, a graded sequence of submodules) such that

$$d \in D, a \in A \implies ad \in D, da \in D$$

Definition 1.15 (*graded quotient algebra*). Given an ideal D in a graded algebra A, the quotient graded algebra A/D inherits graded module and ring operations from A/D regarded as a graded module and ring, respectively. It is a graded algebra satisfying theorem 1.5. The projection $p : A \to A/D$ is universal for the functor $\mathbf{GrAlg}_{\sim(D)=0}(A, -)$ assigning to each graded algebra B the graded algebra morphisms from A with kernel containing D.

Every ring is a graded algebra over $\mathbb{Z}$ concentrated at degree 0, so the graded quotient algebra subsumes the quotient ring construction.

Example 1.4 (*degree truncation by quotients*). Fix a positive integer n, the sequence $D^{(n)} = (0, \cdots, 0, A_n, A_{n+1}, \cdots)$ is an ideal in the graded algebra A. The only nontrivial elements of A/D are of degree less than n.

Ideals in A are partially ordered by inclusion and closed under arbitrary intersection.

Definition 1.16 (*generated ideals*). Let X be any subset of A, the intersection of all ideals of A containing X is an ideal, and denoted the ideal generated by the subset X.

Theorem 1.12 (*characterization of generated ideals*). Let X be a set of elements of a graded algebra A, the p-th component D_p of the ideal D generated by X consists of the finite sums

$$a_1x_1b_1 + a_2x_2b_2 + \cdots + a_nx_nb_n$$

with $a_1, b_1 \in A$, each $x_i \in X$, and each $a_ix_ib_i$ of degree p. Note that "coefficients" a_i, b_i appear on both sides due to noncommutativity and in general contribute to the degree. Finally, also note that the finite sum suffices to make D an ideal, so the ideal consisting of possibly infinite sums may not be the smallest ideal.

Proof: the proposed set is clearly an ideal. This ideal contains $x \in X$, and any ideal containing X must contain all products axb and all their sums.

Theorem 1.13 (*universal property of generated ideals*). If $D(X) \subset A$ is the ideal generated by $X \subset A$, the projection $A \to A/D(X)$ is universal for morphisms $t : A \to B$ of graded algebras with $t(x) = 0$ for every $x \in X$. Equivalently,

$$\mathbf{GrAlg}_{\sim(X)=0}(A, -) \cong \mathbf{GrAlg}(A/D(X), -)$$

Proof: Since $D(X)$ is the smallest ideal containing X, and $\ker t$ is an ideal,

$$X \subset \ker t \implies D(X) \subset \ker t$$

then $\mathbf{GrAlg}_{\sim(X)=0}(A,-) \cong \mathbf{GrAlg}_{\sim(D(X))=0}(A,-) \cong \mathbf{GrAlg}(A/D(X),-)$.

Every graded algebra can be made into an (ordinary) algebra via the Σ functor from the category of graded algebras to the category of algebras.

Definition 1.17 (*internally graded algebra*). Given a graded algebra A, we can construct the internally graded module ΣA by definition 1.9. Element are sequences with finite support and module operations are defined poinwise. The module ΣA can be made into an algebra with the natural product sequence

$$(fg)_n = f_0 g_n + f_1 g_{n-1} + \cdots + f_n g_0$$

Definition 1.18 (*graded tensor algebra*). Given a module C over K, the graded tensor algebra $T(C)$ is the sequence

$$(T_0(C), T_1(C), T_2(C), T_3(C), \cdots) = (K, C, C^{\otimes 2}, C^{\otimes 3}, \cdots)$$

with module operations defined pointwise and (tensor) product defined by the isomorphism

$$T_p(C) \otimes T_q(C) \cong T_{p+q}(C)$$

The functor $T : \mathbf{Mod} \to \mathbf{GrAlg}$ from modules to graded algebras has $T(C \xrightarrow{f} D)$ via

$$T_p(C) = C^{\otimes p} \xrightarrow{f^{\otimes p}} D^{\otimes p} = T_p(D)$$

The graded tensor functor $T : \mathbf{Mod} \to \mathbf{GrAlg}$ is of interest as the "free" graded algebra generated by A_1, in the following sense.

Theorem 1.14 (*universal property of graded tensor functor*). Given $A \in \mathbf{GrAlg}$, to every morphism $s \in \mathbf{Mod}(C, A_1)$ there is a unique morphism $t \in \mathbf{GrAlg}(T(C), A)$ with $t_1 = s$. Equivalently, the functor $\mathbf{Mod}(C, (-)_1) : \mathbf{GrAlg} \to \mathbf{Set}$ is representable via

$$\mathbf{GrAlg}(T(C), A) \cong \mathbf{Mod}(C, A_1)$$

In terms of adjunctions: $T \dashv (-)_1 : \mathbf{Mod} \to \mathbf{GrAlg}$.

Proof: by $s \in \mathbf{Mod}(C, A_1)$, we are given $C \xrightarrow{s=t_1} A_1$ and $K \xrightarrow{u=t_0} A_0$ by the unit and linearity. By the graded algebra structure of $T(C)$, every morphism $T(C) \xrightarrow{t} A$ satisfies

$$T_p(C) \xrightarrow{t_p} A_p : \quad t_p(c_1 \otimes \cdots \otimes c_p) = (t_1\, c_1)(t_1\, c_2) \cdots (t_1\, c_p)$$

Given $s = t_1$, this formula uniquely determines t if it exists. Now t_p exists since

$$(c_1, \cdots, c_p) \mapsto (t_1\, c_1)(t_1\, c_2) \cdots (t_1\, c_p)$$

is multilinear and $\otimes^p$ is universal multilinear.

1.8 Exterior algebra

We consider graded algebras A over K with the property

$$a \in A_1 \implies a^2 = 0$$

Definition 1.19 (*exterior algebra*). Given a module C over K, let $D(C^2)$ be the ideal generated by $c^2, c \in C$. The exterior algebra is the graded quotient algebra

$$\Lambda(C) = T(C)/D(C^2)$$

Multiplication is written $(a, b) \mapsto a \wedge b$ and denoted the wedge (or exterior) product.

The exterior algebra is often defined in the following equivalent way which emphasizes mod-ing out symmetric elements.

Proposition 1.15. $D(C^2) = D(\{c_1 \otimes c_2 + c_2 \otimes c_1 \mid c_1, c_2 \in C\})$

Proof: $(c_1 + c_2)^2 - c_1^2 + c_2^2 = c_1 \otimes c_2 + c_2 \otimes c_1$. We can also take $c_1 = c_2$.

Proposition 1.16. Every commutative graded algebra A over a field $K = F$ of characteristic not 2 is an exterior algebra.

Proof: recall the definition of commutative algebra in definition 1.11. For $c \in T_1(C) = C$

$$c^2 = (-1)^{1 \cdot 1} c^2 \implies c^2 = 0$$

Note that when the field has characteristic 2, we have $1 \cdot 1 = 0$.

Proposition 1.17. For every K-module C, the graded exterior algebra $\Lambda(C)$ satisfies

- $\Lambda_0(C) = K, \Lambda_1(C) = C$
- generated by the set $\Lambda_1(C)$. In other words, every element in C_p can be written as a finite linear combination of wedge products of $c \in C_1$.
- commutative in the sense of definition 1.11
- the square $a \wedge a$ of every element a of odd degree is zero.

Proof:

- $\Lambda_0(C) = T_0(C) = K$ and $\Lambda_1(C) = T_1(C) = C$ since $D(C^2)$ does not contain any element of degree 0 or 1.
- In the tensor algebra, every element is a sum of products of elements of degree 1, so is the projection image $\Lambda(C)$.
- Since $c^2 = c \wedge c = 0 \implies 0 = (c_1 + c_2)^2 = c_1 \wedge c_2 + c_2 \wedge c_1$ Elements of degree 1 commute with the correct sign -1. Commutativity is established by induction.
- If a is a simple product $c_1 \wedge \cdots \wedge c_p$, then a^2 has a factor $c_1 \wedge c_1 = 0$. Otherwise,

$$(a_1 + \cdots + a_k)^2 = \sum a_i^2 + \sum_{i<j} (a_1 \wedge a_j + a_j \wedge a_i) = 0$$

Example 1.5 ($G^{(1)}$ *as exterior algebra*). Consider the free module on one generator b ober K. The exterior construction for $\Lambda(Kb)$ yields $b \wedge b = 0$. This effectively annihilates all higher degrees in $T(Kb)$, yielding

$$\Lambda(Kb) = (K, Kb, 0, \cdots)$$

Example 1.6 ($G^{(2)}$ *as exterior algebra*). Consider the free module $C = Kb_1 \oplus Kb_2$ on two generators b_1, b_2. We have $\Lambda_{p>2}(0)$ since any component of such an element contains a $b_1 \wedge b_1$ or $b_2 \wedge b_2$ factor by duplicity. The only interesting component $\Lambda_2(C)$ is spanned by $b_1 \wedge b_2$.

Note that $b_1 \wedge b_2$ is not zero: $D(Kb_1 \oplus Kb_2)$ is generated by $b_1^2, b_1b_2, b_2b_1, b_2^2$ or, equivalently, $b_1^2, b_1b_2 + b_2b_1, b_2^2, b_1b_2$. Every square $c^2 \in T_2(C)$ is a linear combination on the first free generators, so $b_1b_2 \notin D_2$ so $[b_1b_2] = b_1 \wedge b_2 \neq 0$. Then

$$\Lambda(Kb_1 \oplus Kb_2) = (K, Kb_1 \oplus Kb_2, K(b_1 \wedge b_2), 0, \cdots)$$

with product $b_1^2 = b_2^2 = 0$ and $b_1 \wedge b_2 = -b_2 \wedge b_1$.

Our construction of the exterior algebra is the most general algebra about C which factors out duplicate elements, in a sense made precise by this universal property:

Theorem 1.18 (*universal property of exterior algebra*). For $S \in \mathbf{GrAlg}$ with $s^2 = 0$ for every $s \in S_1$. For every $h \in \mathbf{Mod}(C, S_1)$ there is a unique $t \in \mathbf{GrAlg}(\Lambda(C), S)$ with $t_1 = h$.

Proof: Consider C as a graded module concentrated in degree 1.

$$\begin{tikzcd} T(C) \arrow[r, "\pi"] \arrow[dr, dashed, "k"] & \Lambda(C) \arrow[d, dashed, "t"] \\ C \arrow[u, "\iota"] \arrow[r, "h"] & S \end{tikzcd}$$

We are given h, by universality of $T(C)$ there exists a unique k such that the lower left triangle commutes. Since $\forall s \in S_1, s^2 = 0$, we must have $\forall c \in T_1(C), c^2 \in \ker k$. This implies $D(C^2) \in \ker k$. By the universality of the projection $T(C) \xrightarrow{\pi} \Lambda(C)$ there exists a unique t which makes the upper-right triangle commute.

There is an additional universal propery in terms of individual components of $\Lambda_p(C)$.

Definition 1.20 (*alternating maps*). $h \in \mathbf{Mod}(C^{\otimes p}; D)$ is alternating when

$$c_i = c_{j \neq i} \implies h(c_1, \cdots, c_p) = 0$$

Theorem 1.19 (*characterization of alternating maps*). h is alternating if and only if

$$i \neq j \implies h(c_1, \cdots, c_i, \cdots, c_j, \cdots) = -h(c_1, \cdots, c_j, \cdots, c_i, \cdots)$$

Proof: the given condition implies alternating since interchanging c_i, c_j introduces a negative sign but does not change value. Conversely, given alternating and ignoring other indices

$$\begin{aligned} 0 &= h(c_i + c_j, c_i + c_j) \\ &= h(c_i, c_i) + h(c_i, c_j) + h(c_j, c_i) + h(c_j, c_j) \\ &= h(c_i, c_j) + h(c_j, c_i) \end{aligned}$$

Theorem 1.20 (*universal property of the exterior space*).

$$\mathbf{Alt}(C^{\otimes p}, D) \cong \mathbf{Mod}(\Lambda_p(C), D)$$

Proof: It helps to describe Λ_p directly. By theorem 1.12, $D(C^2)$ consists of finite sums of ac^2b of degree p, with $a, b \in T(C)$. The kernel of $T_p \xrightarrow{\pi} \Lambda_p$ then consists of the submodule T_p spanned by all products $c_1 \otimes \cdots \otimes c_p$ with p factors with two successive factors are equal. Then $h \in \mathbf{Alt}(C^{\otimes p}, D)$ factors through π uniquely by annihilating repetitive elements.

1.9 Subspaces by exterior algebra

We begin by explicitly constructing a basis for $\Lambda_p(C)$ from a basis for C.

Proposition 1.21. If C is a free module with basis $b_1, \cdots, b_n$ (every vector space is free),

- $\Lambda_{p>n}(C) = 0$
- $\Lambda_{0\leq p\leq n}$ is free with $\binom{n}{p}$ basis $b^I = b_{i_1} \wedge \cdots b_{i_p}$, with I an monotonic multi-index, i.e.

$$j < k \implies i_j < i_k$$

Proof: The first part is trivial: duplicate tensor elements always wedge to 0. We first prove that $b_1 \wedge \cdots \wedge b_n \neq 0 \in \Lambda_n(C)$. The determinant is an alternating multilinear form $C^{\otimes n} \to K$ and $c_1 \wedge \cdots \wedge c_n$ is the universally alternating, so for some $\Lambda_n(C) \xrightarrow{t} K$,

$$\det(b_1, \cdots, b_n) = t(b_1 \wedge \cdots \wedge b_n) = 1 \implies b_1 \wedge \cdots \wedge b_n \neq 0 \in \Lambda_n(C)$$

Next, for any degree $p \leq n$, suppose for contradiction that $\{b_I\}$, for I spanning over all $\binom{n}{p}$ monotonic multi-indices, is linearly dependent, so that

$$\lambda^I b_I = 0$$

By linear dependence pick some multi-index J such that $\lambda^J \neq 0$. Form a monotonic multi-index J' with $q = n - p$ indices complementing J. Then

$$b_J \wedge b_{J'} = \pm b_1 \wedge \cdots \wedge b_n$$

If $I \neq J$, then $b_I, b_{J'}$ have at least one index in common so $b_I \wedge b_{J'} = 0$. Wedging with $b_{J'}$ "picks out" b_J and we have the following contradiction on $\lambda^J \neq 0$:

$$0 = b_{J'} \wedge 0 = b_{J'} \wedge \left(\lambda^I b_I\right) = (b_{J'} \wedge b_J)\, \lambda^J = (b_1 \wedge \cdots \wedge b_n)\, \lambda^J = 0$$

This proves that $\{b_I\}$ forms a basis with $\binom{n}{p}$ elements.

Exterior algebra provides a powerful, concise framework to describe linear subspaces. We consider a n-dimensional vector space V over a field F and its exterior algebra $\Lambda(V)$. Elements of $\Lambda_p(V)$ are called p-vectors, and $v_1 \wedge \cdots \wedge v_p$ decomposable p-vectors.

Theorem 1.22 (*linear independence*). A list of vectors $\{v_1, \cdots, v_p\}$ is linearly independent if and only if $\bigwedge v_j \neq 0 \in \Lambda_p(V)$.

Proof: If $\{v_j\}$ is independent, then it can be completed to a basis whose wedge product is a nonzero basis element of $\Lambda_p(V)$ by proposition 1.21 above. Conversely, suppose $\{v_j\}$ is dependent with

$$v_p = v_1\kappa_1 + \cdots + v_{p-1}\kappa_{p-1}$$

then every term of the expanded wedge product has a repeating factor, thus

$$v_1 \wedge \cdots \wedge v_p = v_1 \wedge \cdots \wedge v_{p-1} \wedge (v_1\kappa_1 \cdots + v_{p-1}\kappa_{p-1})$$

Proposition 1.23. Every p-vector $s \in \Lambda_p(V)$ defines a linear map $V \xrightarrow{(-)\wedge s} \Lambda_{p+1}(V)$. Let K denote the kernel of this map. Then

- $\dim K \leq p$
- each basis $\{b_1, \cdots, b_q\}$ of K yields a $(p-q)$-vector $u \in \Lambda_{p-q}(V)$ with $s = b_1 \wedge \cdots \wedge b_q \wedge u$.

Proof: Let $d = \dim K$. Given a basis $\{b_1, \cdots, b_d\}$ of K, we have

$$b_1 \wedge s = b_2 \wedge s = \cdots = b_q \wedge s = 0$$

Complete $\{b_1, \cdots, b_d\}$ to a basis of V. By proposition 1.21 this gives a basis of $\Lambda_p(V)$. Let

$$s = b_I\lambda^I \implies b_i \wedge s = (b_i \wedge b_I)\lambda^I$$

Every coefficients $\lambda^I \neq 0$ which contributes $b_I\lambda^I \neq 0$ to s must be annihilated by all of $b_1, \cdots, b_q$, so $\lambda^I \neq 0 \implies 1, \cdots, q \in I$. In particular, this means $q \leq p$ since I contains at most p elements. In the formula for s, factoring out the common wedge $b_1 \wedge \cdots \wedge b_q$ yields

$$s = b_I\lambda^I = (b_1 \wedge \cdots \wedge b_q)\, b_{I'}\lambda^I$$

Here I' denotes a multi-index I with possible occurences of $1, \cdots, q$ removed (in particular, occurences for all of them when $\lambda^I \neq 0$). Then the desired $(p-q)$-vector is

$$u = b_{I'}\lambda^I$$

Theorem 1.24 (*characterization of span by wedge product*). Two list of p linearly independent vectors $\{v_1, \cdots, v_p\}$ and $\{w_1, \cdots w_p\}$ span the same subspace of a finite-dimensional vector space V if and only if $\bigwedge v_j$ is a nonzero scalar multiple of $\bigwedge w_j$.

Proof: suppose they span the same subspace, then every w_j is a linear combination of v_k. By the multilinearity of $w_1 \wedge \cdots \wedge w_p$ we have the desired relation. Conversely, suppose $\bigwedge v_j = (\bigwedge w_j)\kappa$. Then $\ker((-) \wedge \bigwedge v_j) = \ker((-) \wedge \bigwedge w_j)$. This kernel contains $\{w_j\}$ and $\{v_j\}$ and has dimension at most p by proposition 1.23. By independence condition, $\{w_j\}, \{v_j\}$ are both bases for this kernel.

We denote two nonzero p-vectors of V equivalent if $s' = s\kappa$ for $\kappa \neq 0$.

Corollary 1.2. In a finite-dimensional vector space V, let $S(V)$ denote the space of all subspaces of V, the assignment

$$\Lambda_p(V) \xrightarrow{s \mapsto \ker(-\wedge s)} S(V)$$

induces a bijection between the equivalence classes of nonzero decomposable p-vectors to the set of all p-dimensional subspaces.

Proof: given decomposable $s = v_1 \wedge \cdots \wedge v_p \neq 0$, the subspace $\ker(- \wedge s) \subset V$ contains p linearly independent vectors $\{v_j\}$ and is of dimension p.

The natural follow-up is the characterization of decomposable p-vectors. Every 1-vector is decomposable, as well as n-vectors, for $n = \dim V$. However, p-vectors are not generally decomposable.

Proposition 1.25. In a n-dimensional vector space, every $(n-1)$-vector is decomposable.

Proof: Take a basis $\{b_j\}$ for V. This induces a basis for $\Lambda_{n-1}(V)$

$$s_i = b_1 \wedge \cdots \wedge b_{i-1} \wedge b_{i+1} \wedge \cdots \wedge b_n, \quad i = 1, \cdots, n$$

Every $(n-1)$-vector is a linear combination $s = \sum_{i=1}^n s_i \kappa^i$. Consider a vector $v = b_i v^i \in V$:

$$v \wedge s = (b_1 \wedge \cdots \wedge b_n) \sum (-1)^{i-1} \kappa_i v^i$$

The wedge factor is nonzero, and $(-1)^{i-1}\kappa_i v^i$ has solution of dimension 1. Therefore the kernel $\ker(- \wedge s)$ has dimension $n-1$ while $s \in \Lambda_{n-1}(V)$. Invoke the corollary above.

2 Affine and Projective Spaces

Materials from this section are from Maclane *Algebra* 3rd edition, appendix.

2.1 Affine line

Definition 2.1 (*One-dimensional real affine group, real affine line*). The group A_1 of one-dimensional real affine automorphism is the group $A_1 \cong \mathbb{R} \times (\mathbb{R} - \{0\})$ with multiplication

$$a(\mu', \kappa') \circ a(\mu, \kappa) = (\mu' + \kappa'\mu, \kappa'\kappa)$$

It acts on the real affine line L, a set with elements of $\mathbb{R}$ (without additional structure) by

$$a(\mu, \kappa)x = \kappa x + \mu$$

The projection $A_1 \xrightarrow{\pi_2} (\mathbb{R}^*, \cdot)$ is an epimorphism of groups given by

$$\pi\left[a(\mu', \kappa') \circ a(\mu, \kappa)\right] = \kappa'\kappa = \pi\left[a(\mu', \kappa')\right] \pi\left[a(\mu, \kappa)\right]$$

Its kernel consists of all translations $a(\mu, 1)$, which happens to be the image $(\mathbb{R}, +) \xrightarrow{\iota_1} A_1$. This gives rise to a short exact sequence

$$0 \to (\mathbb{R}, +) \xrightarrow{\mu \mapsto a(\mu,1)} A_1 \xrightarrow{a(\mu,\kappa) \mapsto \kappa} (\mathbb{R}, \cdot) \to 1$$

Definition 2.2 (*equivalent characterization of one-dimensional affine geometry*). The real affine linea L is the set $\mathbb{R}$ with translation group action $(\mathbb{R}, +) \times L \to L$ given by $\mu\, l = \mu + l$. For $l \in L, \mu \in \mathbb{R}, \omega_1 + \omega_2 = 1 \in \mathbb{R}$, the average of $l, \mu + l \in L$ by coefficients ω_1, ω_2 is by

$$(\mu\omega_2)\, l = l\omega_1 + (\mu + l)\omega_2$$

Averages may be defined in terms of appropriate translations. Affine transformations are functions $a : L \to L$ which preserve all averages.

Definition 2.3 (*affine property*). A property P of two or more points l, l' on a line L is an affine property when, for every affine transform $a \in A_1$.

$$P(l, l') \iff P(a(l), a(l'))$$

2.2 Affine spaces

From now on, we consider fields $F$of characteristic not 2 so $1 + 1 \neq 0$ and $\dfrac{1}{2} \in F$.

Definition 2.4 (*affine space*). An affine space P over F is a nonempty set equipped with a finite-dimensional vector space V over F and a regular abelian group action $V \times P \to P$, written $(v, p) \mapsto v+p$. Recall that regular group action is both transitive and free, that there exists one unique orbit and only the identity action has fixed points. In concrete terms:

$$\begin{aligned} \forall v, w \in V, p \in P: &\quad 0 + p = p, (v + w) + p = v + (w + p) \\ \forall p, q \in P, \exists! v \in V: &\quad p = v + q \end{aligned}$$

Given $p, q \in L$, the unique vector $v \in V$ such that $p = v + q$ is denoted "the vector from q to p" and written $v = p - q$.

The dimension of P is defined as $\dim P^\sharp$, and $P^\sharp$ is called the space of translations. There exists a natural isomorphism $V \cong P^\sharp$ by $v \leftrightarrow p \mapsto v + p$ and we use them interchangeably.

The one-dimensional affine line is an affine space with $L^\sharp = \mathbb{R}$. In general, given any vector space V over F with a finite-dimensional subspace $S \subset V$ and distinguished coset $P = S + u_0$, then P is an affine space over F with $P^\sharp = S$ under the identification

$$(t \in S)\ s + u_0 = t + s + u_0 \in P$$

Given any finite-dimensional vector space V, let $V \cong V^\flat \in \mathbf{Set}$. Then $V^\flat$ is a particular case of the example above with $S = V, P \cong V$. The corresponding translation space is $(V^\flat)^\sharp = V$.

Proposition 2.1. Every affine space P has a displacement operation $(-) : P \times P \to P^\sharp$ given by $p - q = v \in P^\sharp$ such that

$$\begin{aligned} (p - q) + p &= p \\ (p - q) + (q - r) &= p - r \\ p - q = 0 &\iff p = q \end{aligned}$$

Note that the difference of two points is a vector, while a sum of a vector plus point is a point. The difference $-_p$ is compatible with $+_{P^\sharp}, -_{P^\sharp}$ commutatively and associatively:

$$\begin{aligned} v + (p - q) &= (v + p) - q \\ (v + p) - (w + q) &= (v - w) + (p - q) \end{aligned}$$

Definition 2.5 (*weights and averages*). A list ω of n weights is a list of n scalars $\omega_i \in F$ with $\sum \omega_i = 1$. Let p denote a list of n points of P, then the average is defined as

$$\left(\sum (p_i - q)\omega_i\right) + p_0 \in P$$

The quantity above can be shown to be invariant towards different choices of q. The average of p with weights ω is denoted

$$\sum p_i \omega_i = \left(\sum (p_i - q)\omega_i\right) + p_0 \in P$$

Note that P is not equipped with a scalar multiplication, so the expression above only makes sense for $\sum \omega_i = 1$. Otherwise it is compatible in every way with the linear structure in $P^\sharp$.

The following lemma is useful in reducing all possible averages to the $n = 2$ case.

Lemma 2.2. Every $(n > 2)$-fold average can be written as a composite of 2-fold averages.

Proof: For $n > 2$ and $\omega_1 \neq 1$. Let $\lambda = (1 - \omega_1)^{-1}$, then

$$\sum_{j=1}^{n} p_j \omega_j = p_1 \omega_1 + \left[\sum_{j=2}^{n} p_j(\omega_j \lambda)\right] \lambda^{-1}$$

Similarly, for every $n > 3$-average with $\omega_1 + \omega_2 \neq 1$, for $\mu = (1 - \omega_1 - \omega_2)^{-1}$ we have

$$\sum_{j=1}^{n} p_j \omega_j = p_1 \omega_1 + p_2 \omega_2 + \left[\sum_{j=3}^{n} p_j(\omega_j \mu)\right] \mu^{-1}$$

For $n = 3$, $\omega_1 + \omega_2 + \omega_3 = 1$ implies that at least one ω_i is not 1 (since $2 \neq 0$), then very 3-fold average is a composition of 1-fold and 2-fold averages. For $n > 3$, at least one of ω_1, ω_2, or $\omega_1 + \omega_2$ is not 1, so every n-fold average reduces to one of the identities above. Apply induction on n.

Definition 2.6 (*affine space*). Given affine spaces P, P' over the same field F, an affine transformation $a : P \to P'$ is a set funcion $P \to P'$ with

$$a\left(\sum_{j=1}^{n} p_j \omega_j\right) = \sum_{j=1}^{n} a(p_j)\omega_j$$

for all n, all points $p_i \in P$ and all weights ω.

Lemma 2.3. Every translation is an affine transformation.

Proof: Let $a(p) = v + p$ for $v \in P^\sharp$, then

$$\begin{aligned} v + \sum_{j=1}^{n} p_j \omega_j &= v + \sum_{j=1}^{n} (p_j - p_0)\omega_j + p_0 \\ &= \sum_{j=1}^{n} (v + p_j - p_0)\omega_j + p_0 = \sum_{j=1}^{n} (v + p_j)\omega_j \end{aligned}$$

Lemma 2.4. A bijection $f : P \to P$ is a translation if and only if, for all $p, q \in P$:

$$f(q) = q1 + f(p)1 + p(-1)$$

Proof: Parallelogram law. The displacement is given by $v = f(p) - p$ for any $p \in P$.

Every affine space arises from some finite-dimensional vector space.

Theorem 2.5 (*characterization of affine spaces*). Every affine space P over F is affine-isomorphic to the affine space $(P^\sharp)^\flat$ for some finite-dimensional vector space $P^\sharp$ over F. Fixing a priori $p_0 \in P$, the assignment $\left(P^\sharp\right)^\flat \xrightarrow{v \mapsto v + p_0} P$ is an affine isomorphism.

Corollary 2.1. Given finite-dimensional vector spaces V, W over F. A set function $V \xrightarrow{f} W$ is a linear transformation if and only if $f(0) = 0$ and $f : V^\flat \to W^\flat$ is an affine transformation.

Proof: every linear transformation is affine since it preserves linear combinations. Conversely, given f affine, any linear combination $v_1\lambda_1 + v_2\lambda_2 \in V$ may be written as a 3-fold average in $V^\flat$ by

$$v_1\lambda_1 + v_2\lambda_2 = 0(1 - \lambda_1 - \lambda_2) + v_1\lambda_1 + v_2\lambda_2$$

The f preserves averages and 0, so it preserves linear combinations.

3 Differential Forms on $\mathbb{R}^n$

Reference for this section is Loring W. Tu, *an Introduction to manifolds, 2nd ed*, Chapter 3.

3.1 Multilinear algebra

Definition 3.1 (*permutation action*). The group action $S_k : \mathrm{Hom}(V^k, \mathbb{R}) \to \mathrm{Hom}(V^k, \mathbb{R})$ is defined by

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

This is a valid group action since $\tau\,(\sigma\, f) = (\tau\sigma)\, f$. Recall that the orbit of a k-linear function $f \in \mathrm{Hom}(V^k, \mathbb{R})$ is the set

$$S_k\, f = \{\sigma\, f \mid \sigma \in S_k\}$$

Definition 3.2 (*symmetric, alternating covectors*). A k-linear function $f : V^k \to \mathbb{R}$ is symmetric if, for every permutation $\sigma \in S_k$,

$$\sigma\, f = f$$

It is alternating if

$$\sigma\, f = (\operatorname{sgn}\sigma) f$$

The space of all such alternating covectors over V^k is denoted $A_k(V)$.

Definition 3.3 (*alternating and symmetrizing operator*). The symmetrizing operator S maps $\mathrm{Hom}(V^k, \mathbb{R})$ to the space of symmetric linear maps. The orbit $S_k\, f$, as a whole, is invariant under group action.

$$S\, f = \sum_{g \in (S_k\, f)} g = \sum_{\sigma \in S_k} (\sigma\, f)$$

The alternating operator A maps to the space of alternating maps.

$$A\, f = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma) \cdot (\sigma\, f)$$

Proposition 3.1. $\ker\, S = \langle \alpha - \sigma\,\alpha \mid \alpha \in \mathrm{Hom}(V^k, \mathbb{R}), \sigma \in S_k \rangle$

Proof: The symmetrizing operator is idempotent: $S^2 = S$. Now $\ker\, S \subseteq \mathrm{Im}(\mathrm{id} - S)$ since

$$\forall f \in \ker\, S, f = (\mathrm{id} - S)\, f$$

On the other hand, $\mathrm{Im}(\mathrm{id} - S) \subseteq \ker\, S$ since

$$S\ [\mathrm{Im}(\mathrm{id} - S)\, f] = (S^2 - S)\, f = 0$$

Note that the symmetrizing operator sends elements of the same orbit to the same result.

Proposition 3.2. $\ker A = \langle \alpha + \tau\alpha \mid \alpha \in \mathrm{Hom}(V^k, \mathbb{R}) \text{ and } \tau \text{ a transposition} \rangle$.

Proof: if $\beta = \tau\alpha$, then $A\,\beta = -(A\,\alpha)$. Let $\tau = (i, j)$, note that $\alpha + \tau\alpha$ is redundant in the i, j components.

Definition 3.4 (*wedge (exterior) product*). Consider the minimal alternating product of two alternating tensors $f \in A_k(V), g \in A_l(V)$. Since f, g are already alternating, without loss of generality, the coefficients of f, g can be respectively in ascending order. We denote by a permutation in S_{k+l} a (k, l)-shuffle if

$$\sigma\,1 < \cdots < \sigma\,k, \quad \sigma(k+1) < \cdots < \sigma(k+l)$$

Note that a (k, l)-shuffle contains $\binom{k+l}{k}$ terms, while S_{k+l} contains $(k+l)!$ terms. This is because for each $\tau \in S_{k+l}$, there are $k!l!$ permutations which only permute elements within the first k, l slots respectively without crossing. The wedge product of f, g is defined as

$$\begin{aligned}
(f \wedge g)(v_1, \cdots, v_{k+l}) &= \sum_{\sigma \in (k,l)\text{-shuffle}} f\left(v_{\sigma\,1}, \cdots, v_{\sigma\,k}\right) g\left(v_{\sigma\,k+1}, \cdots, v_{\sigma\,k+l}\right) \\
&= \frac{1}{k!\,l!} \sum_{\sigma \in S_{k+l}} f\left(v_{\sigma\,1}, \cdots, v_{\sigma\,k}\right) g\left(v_{\sigma\,k+1}, \cdots, v_{\sigma\,k+l}\right) \\
f \wedge g &= \sum_{\sigma \in S_{k+l}} \frac{1}{k!\,l!} A(f \otimes g)
\end{aligned}$$

Proposition 3.3. The wedge product is anticommutative

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

Proof: for each of the first k slots of f in $f \wedge g$, we need l transpositions to bubble it right to its corresponding position in $g \wedge f$.

Corollary 3.1. if f is an alternating tensor of odd degree on V, then $f \wedge f = 0$

Proof: $f \wedge f = (-1)^{(2n+1)^2} f \wedge f$

Proposition 3.4. $A(A(f) \otimes g) = k!A(f \otimes g)$

Proof: Adopting the extension that $\tau \in S_k$ is an element of S_{k+l} fixing the other elements

$$\begin{aligned}
A(A(f) \otimes g) &= \sum_{\sigma \in S_{k+l}} (\operatorname{sgn}\sigma) \cdot \sigma \left[\sum_{\tau \in S_k} (\operatorname{sgn}\tau)(\tau\, f) \otimes g \right] \\
&= \sum_{\tau \in S_k} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma\tau) \cdot (\sigma\tau)\,(f \otimes g) \\
&= \sum_{\tau \in S_k} \sum_{\mu \in S_{k+l}\tau} \operatorname{sgn}\mu \cdot \mu\,(f \otimes g) \\
&= k! \sum_{\mu \in S_{k+l}} (\operatorname{sgn}\mu) \cdot \mu(f \otimes g)
\end{aligned}$$

Proposition 3.5. The wedge product is associative.

Proof: Direct computation

$$\begin{aligned}(f \wedge g) \wedge h &= \frac{1}{(k+l)!m!} A((f \wedge g) \otimes h) \\ &= \frac{1}{(k+l)!m!k!l!} A(A(f \otimes g) \otimes h) \\ &= \frac{1}{m!k!l!} A(f \otimes g \otimes h)\end{aligned}$$

Let $[b^i_j]$ denote the matrix whose (i, j) entry is b^i_j.

Proposition 3.6. Given 1-covectors α^i over v and $v_1, \cdots, v_k \in V$

$$(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \cdots, v_k) = \det\left[\alpha^i\, v_j\right]$$

Proof: Direct computation

$$\begin{aligned}(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \cdots, v_k) &= A(\alpha^1 \otimes \cdots \otimes \alpha^k)(v_1, \cdots, v_k) \\ &= \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \prod_{j=1}^{k} \alpha^j\, v_{\sigma\, j}\end{aligned}$$

Proposition 3.7. Let $e_1, \cdots, e_n$ be a basis for V and $\alpha^1, \cdots, \alpha^n$ be its dual basis. For each $k \leq n$, the alternating k-linear tensors

$$\alpha^I = \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} = \bigwedge \alpha^{i_j}, \qquad i_1 < \cdots < i_k$$

form a basis for the space $A_k(V)$ of alternating k-linear tensors on V.

Proof: For linear independence, let J span over all $\binom{n}{k}$ strictly ascending k-indices.

$$\sum c_I \alpha^I = 0 \implies \forall J, \sum c_I (\alpha^I\, e_J) = \sum c_I \delta^I_J = c_J = 0$$

For spanning, consider any $f \in A_k(V)$, then f, g agree if they agree on all e_I for I ranging over all ascending k-indices.

$$\begin{aligned}g &= \sum_J (f\, \alpha^J) \cdot \alpha^J \\ g\, e_I &= \sum_J (f\, \alpha^J) \cdot (\alpha^J e_I) = f\, e_I\end{aligned}$$

Corollary 3.2. If $\dim V = n$, then $\dim A_k(V) = \binom{n}{k}$

Corollary 3.3. If $k > \dim V$, then $A_k(V)$ is trivial.

Proof: If $k > \dim V$, for each basis α^J, at least two indices must be the same j. For one-covectors, $\alpha^j \wedge \alpha^j = 0$.

Proposition 3.8. For two sets of covectors $\{\beta^1, \cdots, \beta^k\}$ and $\{\gamma^1, \cdots, \gamma^k\}$ over the same V

$$\beta = A\gamma \implies \bigwedge \beta = (\det A) \bigwedge \gamma$$

Proof: We use the scalar linearity of the wedge product and the big product formula on the second step, the redundant property of the wedge product on the third step, and the determinant definition on the last step.

$$\begin{aligned} \bigwedge \beta = \bigwedge_i \left(\sum_j a^i_j \gamma^j \right) &= \sum_{j_1=1}^{k} \cdots \sum_{j_k=1}^{k} \left(\prod_{l=1}^{k} a^l_{j_l} \cdot \bigwedge_{l=1}^{k} \gamma^{j_l} \right) \\ &= \sum_{\sigma \in S_k} \left(\prod_{l=1}^{k} a^l_{\sigma\, l} \cdot \bigwedge_{l=1}^{k} \gamma^{\sigma\, l} \right) = \sum_{\sigma \in S_k} \left(\operatorname{sgn} \sigma \cdot \prod_{l=1}^{k} a^l_{\sigma\, l} \right) \bigwedge_{l=1}^{k} \gamma^l = (\det A) \bigwedge \gamma \end{aligned}$$

3.2 Exterior derivative on $\mathbb{R}^n$

Note the similarity between the following rule and the Lebneiz rule for derivation on algebras.

Definition 3.5 (*antiderivation of a graded algebra*). Given a graded algebra $A = \bigoplus_{k=0}^{\infty} A^k$ over a field K, an antiderivation of A is a K-linear map $D : A \to A$ such that $\forall a \in A_k, b \in A_l$

$$D(ab) = (D\, a)b + (-1)^{\deg a}\, a(D\, b)$$

If there is an integer m such that $\forall k, D(A^k) \subseteq A^{k+m}$, then D is an antiderivation of degree m. For $A_{k<0} = 0$, an antiderivation can have negative degrees.

Definition 3.6 (*graded algebra of differential forms*). Let U be an open subset of $\mathbb{R}^n$, the set of differential forms $\Omega^*(U)$, with $\Omega^0(U) = C^\infty(U)$, is a module under pointwise addition and real scalar multiplication. It is also a compatible ring under pointwise addition and the wedge product, thus forming a graded algebra.

Definition 3.7 (*exterior derivative on* $\mathbb{R}^n$). The set of differential forms over $U \subset \mathbb{R}^n$ form a graded algebra $\Omega^*(U)$ under the wedge product, with $\Omega^0(U) = C^\infty(U)$. Define the zeroth order differential map $d : \Omega^n(U) \to \Omega^{n+1}(U)$ by the base case

$$df = (\partial_{x^i} f)\, dx_i$$

and inductively, for $\omega = a^I\, dx_I \in \Omega^k(U)$

$$d\omega = d\left(a^I\, dx_I\right) = da^I \wedge dx_I = \left(\partial_{x^j} a^I\right)\, dx_j \wedge dx_I \in \Omega^{k+1}(U)$$

Proposition 3.9. Exterior differentiation is an antiderivation of $\Omega^*(U)$ of degree 1.

$$d(\omega \wedge \tau) = (d\omega) \wedge \tau + (-1)^{\deg \omega} \omega \wedge d\tau$$

It satisfies $d^2 = 0$, and for $f \in \Omega^0(U), X \in \mathfrak{X}(U)$, $df\, X = X\, f$.

Proof: by linearity in a, b, let $\omega = f\, dx_I, \tau = g\, dx_J$, then

$$\begin{aligned} d(\omega \wedge \tau) &= d(fg\, dx_I \wedge dx_J) \\ &= \partial_{x^k}(fg)\, dx_k \wedge dx_I \wedge dx_J \\ &= (f\partial_{x^k} g + g\partial_{x^k} f)\ dx_k \wedge dx_I \wedge dx_J \end{aligned}$$

Compute the right hand side components

$$\begin{aligned} d\omega \wedge \tau &= ((\partial_{x^k} f)\ dx_k \wedge dx_I) \wedge (g\, dx_J) = (g\, \partial_{x^k} f)\ dx_k \wedge dx_I \wedge dx_J \\ \omega \wedge (d\tau) &= f\, dx_I \wedge ((\partial_{x^k} g)\ dx_k \wedge dx_J) = (f\, \partial_{x^k} g)\, (dx_I \wedge dx_k \wedge dx_J) \end{aligned}$$

Noting that $dx_I \wedge dx_k \wedge dx_J = (-1)^{\deg \omega}\, dx_k \wedge dx_I \wedge dx_J$, the desired equation is seen to hold.

For the second property, by linearity consider $\omega = f\, dx_I$, then

$$d^2\omega = d\left((\partial_{x^j} f)\, dx_j \wedge dx_I\right) = \left(\partial^2_{x^j, x^k} f\right)\ dx_k \wedge dx_j \wedge dx_I = 0$$

The sum is zero since when $k = j, dx_k \wedge dx_j = 0$, and when $j \neq k$, there is always a correspondingly pair $dx_k \wedge dx_j$ with the same coefficients (by the symmetry of second derivatives).

An induction argument establishes that the properties above uniquely characterize the exterior differentiation map.

Definition 3.8 (*closed and exact forms*). A k-form ω on $U \subseteq \mathbb{R}^n$ is closed if $d\omega = 0$ and exact if there exists a $(k-1)$-form τ such that $\omega = d\tau$. Every exact form is closed.

Definition 3.9 (*differential complex*). A collection of spaces $\{V_k\}_{k\in\mathbb{N}}$ equipped with $d_{k\in\mathbb{N}} : V_k \to V_{k+1}$ is a differential complex if $d_{k+1}d_k : V_k \to V_{k+2}$ vanishes for all k. The graded algebra $\Omega^*(U)$, with d, forms the de Rham complex of U

$$0 \xrightarrow{\iota} \Omega^0(U) \xrightarrow{d} \Omega^1(U) \xrightarrow{d} \Omega^2(U) \to \cdots$$

The closed forms are elements of $\ker d$, and the exact forms elements of $\operatorname{Im} d$.

3.3 Applications to $\mathbb{R}^3$ calculus

Let U be an open subset of $\mathbb{R}^3$, we can identify various forms with familiar quantities:

- A 1-form $P\, dx + Q\, dy + R\, dz \in \Omega^1(U)$ is identified a vector field $(P, Q, R) \in \mathfrak{X}(U)$.
- A 2-form $P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy \in \Omega^2(U)$ is identified with $(P, Q, R) \in \mathfrak{X}(U)$. Note the order $dz \wedge dx$ instead of $dx \wedge dz$.
- A 3-form $f\, dx \wedge dy \wedge dz \in \Omega^3(U)$ is identifed with a scalar field $f \in \mathfrak{X}(U)$.

Under this identification

- The gradient $\nabla : C^\infty(U) \to \mathfrak{X}(U)$ is of type

$$C^\infty(U) \cong \Omega^0(U) \xrightarrow{d} \Omega^1(U) \cong \mathfrak{X}(U)$$

- The curl $\nabla\times : \mathfrak{X}(U) \to \mathfrak{X}(U)$ is of type

$$\mathfrak{X}(U) \cong \Omega^1(U) \xrightarrow{d} \Omega^2(U) \cong \mathfrak{X}(U)$$

- The divergence $\nabla\cdot : \mathfrak{X}(U) \to C^\infty(U)$ is of type

$$\mathfrak{X}(U) \cong \Omega^2(U) \xrightarrow{d} \Omega^3(U) \cong C^\infty(U)$$

In terms of concrete coordinates, the curl reads,

$$\begin{aligned}
\nabla \times (p, q, r) \leftrightarrow\ & d\,[p\,dx + q\,dy + r\,dz] \\
=\ & -(\partial_y p)\ dx \wedge dy + (\partial_z p)\, dz \wedge dx \\
& + (\partial_x q)\ dx \wedge dy - (\partial_z q)\ dy \wedge dz \\
& - (\partial_x r)\ dz \wedge dx + (\partial_y r)\ dy \wedge dz \\
=\ & (\partial_x q - \partial_y p)\, dx \wedge dy + (\partial_z p - \partial_x r)\, dz \wedge dx + (\partial_y r - \partial_x q)\, dy \wedge dz
\end{aligned}$$

The divergence reads

$$\begin{aligned}
\nabla \cdot (p, q, r) \leftrightarrow\ & d\,[p\,dy \wedge dz + q\,dz \wedge dx + r\,dx \wedge dy] \\
=\ & (\partial_x p)\ dx \wedge dy \wedge dz + (\partial_y q)\ dy \wedge dz \wedge dx + (\partial_z r)\ dz \wedge dx \wedge dy \\
=\ & (\partial_x p + \partial_y q + \partial_z r)\, dx \wedge dy \wedge dz \\
\leftrightarrow\ & (\partial_x p + \partial_y q + \partial_z r)
\end{aligned}$$

This definition gives extremely concise proofs of several results from vector calculus.

Proposition 3.10. $\nabla \times (\nabla f) = 0, \quad \nabla \cdot (\nabla \times f) = 0$

Proof: under this definition, $\nabla \times (\nabla f) \leftrightarrow d^2 f = 0$. Similarly, $\nabla \cdot (\nabla \times X) \leftarrow d^2 \omega_X = 0$.

In $\mathbb{R}^3$, a vector field F is the gradient of some scalar f if and only if $\nabla \times F = 0$. In terms of forms, this says that every closed form is exact. This need not hold for general spaces, and we take the occasion to introduce relevant terminology:

Definition 3.10 (*de Rham cohomology*). the k-th de Rham cohomology of U is the quotient

$$H^k(U) = \frac{\ker\left[\Omega^k(U) \xrightarrow{d} \Omega^{k+1}(U)\right]}{\operatorname{Im}\left[\Omega^{k-1}(U) \xrightarrow{d} \Omega^k(U)\right]}$$

Theorem 3.11 (*Poincaré lemma*). every closed $k \geq 1$-form on $\mathbb{R}^n$ is exact. $H^{k\geq 1}(\mathbb{R}^n) = 0$.

4 Quaternion and Rotations

4.1 Quaternions

Definition 4.1 (*quaternions*). A quaternion $q \in \mathbb{H}$ is of the form, for $a, b, c, d \in \mathbb{R}$

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

Subject to associative multiplication and the rules

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$$

Remark 4.1. Just as the complex numbers make $\mathbb{R}^2$ a commutative division algebra (a field), the quaternions make $\mathbb{R}^4$ a noncommutative division algebra.

Recall the Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

They are involutary and multiply like the quaternions, up to the imaginary unit.

$$\sigma_x\sigma_y = i\sigma_z, \sigma_y\sigma_z = i\sigma_x, \sigma_z\sigma_x = i\sigma_y$$

This allows us to establish an isomorphism between the quaternion and Pauli algebras.

Theorem 4.1 (*Pauli representation of quaternions*). The quaternion and Pauli algebras are isomorphic under the following identification

$$(\mathbf{i}, \mathbf{j}, \mathbf{k}) \leftrightarrow (-i\sigma_x, -i\sigma_y, -i\sigma_z)$$

The consequent representation is

$$\mathbf{i} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \quad \mathbf{j} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \mathbf{k} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$$

Proof: Direct computation. For example, $\mathbf{ij} = (-i\sigma_x)(-i\sigma_y) = -i\sigma_z = \mathbf{k}$, as desired.

Definition 4.2 (*quaternion conjugate, norm*). The components $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are all imaginary units, and conjugating $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$ negates the imaginary components.

$$\bar{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$$

In the Pauli representation, $\bar{q} = q^\dagger$. Then conjugation is contravariant

$$\overline{q_1 q_2} = \bar{q}_2\bar{q}_1$$

The norm is defined as $q\bar{q} = |q|^2 = a^2 + b^2 + c^2 + d^2$. The inverse is seen to be

$$q^{-1} = \frac{1}{|q|^2}\bar{q}$$

In the Pauli representation, $|q|^2 = \det q$. Multiplicative property of the norm follows:

$$|qw| = \det(qw) = \det q \cdot \det w = |q| \cdot |w|$$

The quaternions with unit norm are called the unit quaternions. They constitute $\mathbf{S}^3 \subset \mathbb{R}^4$.

Proposition 4.2. Unit quaterions represent origin-preserving isometries (rotations) of $\mathbb{R}^4$.

Proof: Multiplication by q preserves the origin $u0 = 0$. It's an isometry since

$$|uv - uw| = |u(v - w)| = |u||v - w| = |v - w|$$

Definition 4.3 (*pure imaginary quaterions*). A pure imaginary quaternion is of form

$$u = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

We interchangeably regard them as vectors in $\mathbf{S}^2$. As such, they satisfy

$$uv = -u \cdot v + u \times v$$

In particular, every $u \in \mathbb{R}^3$ squares to -1 under quaternion multiplication

$$u^2 = -u \cdot u = -|u|^2 = -1$$

Every unit quaternion $t \in \mathbf{S}^3$ may be expressed in terms of $u \in S^2$ and $\theta \in [0, 2\pi)$ as

$$t = \cos\theta + u\sin\theta$$
$$\bar{t} = \cos\theta - u\sin\theta$$

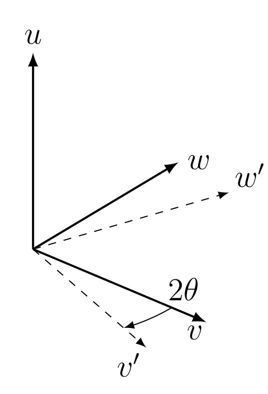
Theorem 4.3 ($SO(3) \cong \mathbb{RP}^3$). Every rotation of $\mathbf{S}^2$ uniquely corresponds to an antidopal pair of unit quaternions $\pm t \in \mathbf{S}^3$ via the conjugation map

$$u \mapsto t^{-1}ut$$

Proof: consider the projection representation of t as $u \in \mathbf{S}^2, \theta \in [0, 2\pi)$ and denote conjugation by $\varphi_u : \mathbf{S}^2 \to \mathbf{S}^3$. Choose $v \in \mathbf{S}^2$ orthogonal to u by the inner product in $\mathbb{R}^3$, so that $u \cdot v = 0$. This determines another orthogonal element

$$w = u \times v = uv - u \cdot v = uv \in \mathbf{S}^2$$

Note that $uv = u \times v = -v \times u = -vu$. Then $\{u, v, uv\}$ is a right-handed basis for $\mathbb{R}^3 \supset \mathbf{S}^2$.



We compute the action of φ_t on the basis u, v, w.

$$\begin{aligned}
t^{-1}ut &= (\cos\theta - u\sin\theta)u(\cos\theta + u\sin\theta) \\
&= (\cos\theta - u\sin\theta)(u\cos\theta - \sin\theta) \\
&= \sin\theta\cos\theta - \sin\theta\cos\theta + u(\sin^2\theta + \cos^2\theta) = u \\
t^{-1}vt &= (\cos\theta - u\sin\theta)v(\cos\theta + u\sin\theta) \\
&= (\cos\theta - u\sin\theta)(v\cos\theta + vu\sin\theta) \\
&= v\cos^2\theta - uvu\sin^2\theta - uv\sin\theta\cos\theta + vu\sin\theta\cos\theta \\
&= v\cos^2\theta + u^2v\sin^2\theta - 2uv\sin\theta\cos\theta \\
&= v\cos^2\theta - v\sin^2\theta - 2w\sin\theta\cos\theta \\
&= v\cos(2\theta) - w\sin(2\theta) \\
t^{-1}uvt &= (\cos\theta - u\sin\theta)uv(\cos\theta + u\sin\theta) \\
&= (\cos\theta - u\sin\theta)(uv\cos\theta + uvu\sin\theta) \\
&= uv\cos^2\theta - u^2vu\sin^2\theta - uuv\sin\theta\cos\theta + uvu\sin\theta\cos\theta \\
&= w\cos^2\theta - w\sin^2\theta + v\sin\theta\cos\theta + v\sin\theta\cos\theta \\
&= w\cos(2\theta) + v\sin(2\theta)
\end{aligned}$$

From u looking at the origin, conjugation by t effects clockwise rotation by 2θ. We also note two symmetries. First, rotating u by 2θ is equivalent to rotating $-u$ by -2θ. The quaternion which affects this rotation is our original quaternion

$$t = \cos(-\theta) + (-u)\sin(-2\theta)$$

Secondly, rotating u by 2θ is equivalent to rotating $-u$ by $2(\pi - \theta)$. The quaternion which affects this change is $-t$

$$\cos(\pi - \theta) + (-u)\sin(\pi - \theta) = -\cos\theta - u\sin\theta = -t$$

This equivalence may also be seen readily from the conjugation formula

$$\varphi_t(u) = t^{-1}ut = (-t^{-1})u(-t) = \varphi_{-t}(u)$$

Every rotation of $\mathbf{S}^2$ (and $\mathbb{R}^3$) is a clockwise rotation θ about an axis $u \in \mathbf{S}^2$. Rotations of $\mathbf{S}^2$ is then $\mathbf{S}^3$ up to antipodal equivalence $q \sim -q$, or, equivalently, $\mathbb{RP}^3$. Equivalently, every clockwise rotation of $\mathbf{S}^2$ around axis $u \in \mathbf{S}^2$ by θ is affected by conjugation via

$$t = \cos\frac{\theta}{2} + u\sin\frac{\theta}{2} \in \mathbf{S}^3$$

Remark 4.2. The antidopal 2-to-1 covering of rotations $SO(3)$ by unit quaternions $\mathbf{S}^3$ is responsible for the factor of 2.

Corollary 4.1. The effect of composing two rotations $(u, \theta), (v, \varphi)$ is another rotation.

Proof: This is a trivial geometric fact, but the formula we derive here is nontrivial. Consider the two unit quaterions affecting the given rotations

$$\begin{aligned} p &= \cos\frac{\theta}{2} + u\sin\frac{\theta}{2} \\ q &= \cos\frac{\varphi}{2} + v\sin\frac{\varphi}{2} \end{aligned}$$

Composing the rotations is then equivalent to applying

$$t \mapsto \left[(t \mapsto q^{-1}tq) \circ (t \mapsto p^{-1}tp)\right] p = (pq)^{-1}t(pq)$$

It suffices to calculate their product. To reduce clutter, let $\alpha = \theta/2, \beta = \varphi/2$

$$\begin{aligned} pq &= (\cos\alpha + u\sin\alpha)(\cos\beta + v\sin\beta) \\ &= \cos\alpha\cos\beta + u\sin\alpha\cos\beta + v\cos\alpha\sin\beta + (uv)\sin\alpha\sin\beta \\ &= \cos\alpha\cos\beta + u\sin\alpha\cos\beta + v\cos\alpha\sin\beta + (-u\cdot v + u\times v)\sin\alpha\sin\beta \\ &= [\cos\alpha\cos\beta - (u\cdot v)\sin\alpha\sin\beta] + u\sin\alpha\cos\beta + v\cos\alpha\sin\beta + (u\times v)\sin\alpha\sin\beta \end{aligned}$$

The real part of this quaternion consequently gives the rotational angle γ via

$$\cos\frac{\gamma}{2} = \cos\alpha\cos\beta - (u\cdot v)\sin\alpha\sin\beta$$

The axis is given by normalizing the imaginary part

$$u\sin\alpha\cos\beta + v\cos\alpha\sin\beta + (u\times v)\sin\alpha\sin\beta$$

5 Tangent Structure

5.1 Tangent map

Summary 1. The pointwise tangent map is equivalent defined by extending kinematic curves, transforming representations via $J_{x\to yf}$, and pulling back derivations. See remark 5.1.

Proposition 5.1. Given a smooth map $f : M \to N$, the tangent map $T_pf : T_pM \to T_qN$ is a linear map. It is an isomorphism if f is a diffeomorphism.

Proof: Taking the kinematic definition,

$$(T_pf)[c] = [fc]$$

Let x, y be charts on M, N, respectively. The vector space structure is defined on T_pM via

$$\begin{aligned}\alpha[d] + [c] &= [x^{-1}(x(p) + t(v_c + \alpha v_d))] \\ &= \left[x^{-1}\left(xp + t\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right)\right]\end{aligned}$$

Applying the tangent map

$$(T_pf)\left(\alpha[d] + [c]\right) = \left[fx^{-1}\left(xp + t\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right)\right]$$

Its representative vector in chart y is

$$\begin{aligned}&d_t\big|_0\left[yfx^{-1}\left(xp + t\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right)\right] \\ =&J_{yfx^{-1}}\big|_{xp}\left[\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right]\end{aligned}$$

On the other hand, the representative vector for $(T_pf)(\alpha[fd]) + (T_pf)([fc])$ in chart y is

$$\begin{aligned}&\alpha d_t\big|_0 yfd + d_t\big|_0 yfc \\ =&\alpha d_t\big|_0 yfx^{-1}xd + d_t\big|_0 yfx^{-1}xc \\ =&J_{yfx^{-1}}\big|_{xp}\left[\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right] \\ =&J_{x\to yf}\big|_p\left[\left(d_t\big|_0(xc) + \alpha d_t\big|_0(xd)\right)\right]\end{aligned}$$

In short, the tangent map transforms representative vectors as

$$T_pf\left(p, v, (U, x)\right) = \left(f(p), J_{x\to yf}\big|_p v, (V, y)\right)$$

Theorem 5.2 (*chain rule of tangent maps*). Given smooth maps $M \xrightarrow{f} N \xrightarrow{g} P$,

$$T_p(gf) = \left(T_{f(p)}g\right)T_pf$$

Proof: $T_p(gf)[c] = [g(fc)] = \left(T_{f(p)}g\right)T_pf$

Definition 5.1 (*pullback, pushforward of functions*). The pullback of $B \xrightarrow{g} C$ along $A \xrightarrow{\varphi} B$ is $A \xrightarrow{\varphi^* g} C = A \xrightarrow{\varphi} B \xrightarrow{g} C$. When φ is a bijection, the pushforward of $A \xrightarrow{f} C$ along φ is $B \xrightarrow{\varphi^{-1}} A \xrightarrow{f} C$. Note that the type transform is different from that of vector fields.

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[dr, dashed, "\varphi^* g"'] & B \arrow[d, "g"] \\ & C \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "\varphi"] \arrow[dr, "f"'] & B \arrow[d, dashed, "\varphi_* f"] \\ & C \end{tikzcd}$$

Definition 5.2 (*derivation definition of tangent maps*). Let v_p be a derivation at p, then $T_p f : C^\infty(M) \to C^\infty(N)$ is defined as

$$(T_p f\, v_p)\, g = v_p(gf) = v_p(f^*\, g)$$

Definition 5.3 (*(point) differential of a scalar*). The (point) differential of $f : C^\infty(M)$ at $p \in$M is $df(p) : T_p M \to \mathbb{R}$ via

$$df(p)\, v_p = v_p f$$

It takes a derivation and evaluates its derived value for f at point p.

For any vector space V and $p \in V$, there is a canonical map $T_p V \xrightarrow{\nu} V$. Note that, in the definition above,

$$T_p f \xrightarrow{df(p)} \mathbb{R} = T_p M \xrightarrow{T_p f} T_p \mathbb{R} \xrightarrow{\nu} \mathbb{R}$$

Definition 5.4 (*(point) differential of a vector*). Given a vector space V, the (point) differential of $M \to V$ at $p \in$M is $df(p) : T_p M \to V$ via

$$T_p f \xrightarrow{df(p)} V = T_p M \xrightarrow{T_p f} T_p V \xrightarrow{\nu} V$$

One may alternatively think of it as the component-aggregation of f_j.

Remark 5.1. Interpretations of the tangent map:

The key here is the chain of morphisms $I \xrightarrow{c} M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}$. A map $f : M \to N$ can

- induce a map $(I \to M) \to (I \to N)$: push a map into M to one into N.

  Extend $c : I \to M$ to $I \xrightarrow{c} M \xrightarrow{f} N$ via this map, then $T_p f\,[c] = \left[\mathbb{R} \xrightarrow{c} M \xrightarrow{f} N\right]$.
- induce a map $(N \to \mathbb{R}) \to (M \to \mathbb{R})$: pull a map out of N to one out of M.

  Extend $N \xrightarrow{g} \mathbb{R}$ to $M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}$. Then $(T_p f\, v_p)\left(N \xrightarrow{(\cdot)} \mathbb{R}\right) = v_p \left(M \xrightarrow{f} N \xrightarrow{(\cdot)} \mathbb{R}\right)$.
- more directly, transform a tangent vector as $(p, v_p, (x, U))$: $T_p f\,[c] = J_{x \to y f} v_p$.

The change of $N \xrightarrow{g} \mathbb{R}$ along $I \xrightarrow{c} M \xrightarrow{f} N$ is the same as that of $M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}$ along $I \xrightarrow{c} M$. Both are defined by the derivative of $I \xrightarrow{c} M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}$ at 0.

Definition 5.5 (*velocity of a curve*). Given $c : I \to M$, its velocity $\dot{c}(t_0) \in T_{c(t_0)}M$ is

$$\dot{c}(t_0) = (T_{t_0}c)\ \partial_u\Big|_{t_0}$$

As a derivation, this pulls on $M \xrightarrow{f} \mathbb{R}$, producing $I \xrightarrow{c} M \xrightarrow{f} \mathbb{R}$ and evaluating its rate of change along c at t_0.

$$\dot{c}(t_0)f = \partial_u\Big|_{t_0}(fc)$$

Proposition 5.3. Given $M \xrightarrow{f} N$. If $T_pf = 0$ for all $p \in M$, then f is locally constant.

Proof: choose an arbitrary $p \in M$ and (U, x) in M covering p. For $q = f\,p$ choose chart (V, y) in N covering q. Then $T_p(yfx^{-1}) = (T_qy)(T_pf)(T_{x^{-1}p}x^{-1}) = 0$. This is true for every $p' \in U$, so yfx^{-1} is constant on $x^{-1}U$. Since x^{-1}, y are both bijections, f is constant on U. This applies for every p and a chart domain U covering it.

Theorem 5.4 (*inverse mapping theorem*). Given a smooth map $M \xrightarrow{f} N$ such that T_pf is an isomorphism, there exists an open neighborhood of p such that $f(O)$ is open and $f_{|O}$ is a diffeomorphism.

Proof: $J_{x\to yf}\Big|_p : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism and apply the inverse mapping theorem: yfx^{-1} restricted on some open $O' \in \mathbb{R}^n$ is a diffeomorphism. Take $O = x^{-1}(O')$.

5.2 Tangent and Cotangent Bundles

Summary 2. Natural charts exist on the tangent and cotangent bundles. On principle components, tangent charts transform as $J_{x\to y}$. The cotangent charts are their contragredient and transform as $J^*_{y\to x}$. The contragredient is the natural covariant functor capturing duality, see lemma 5.9.

Definition 5.6 (*tangent lift*). Given $M \xrightarrow{f} N$, its tangent lift $TM \xrightarrow{Tf} TN$ is defined by

$$Tf(v_p) = T_pf(v_p)$$

Proposition 5.5. Given $M \xrightarrow{f} N \xrightarrow{g} K$

$$T(gf) = (Tg)(Tf)$$

Proof: $T(gf)\,v_p = T_p(gf)(v_p) = (T_{f\,p}\,g)(T_p\,f)v_p = (Tg)(Tf)\,p$

Definition 5.7 (*tangent functor*). The tangent functor $T : \mathbf{Man} \to \mathbf{Man}$ maps M to TM and $M \xrightarrow{f} N$ to $TM \xrightarrow{Tf} TN$.

Definition 5.8 (*tangent bundle projection map*). $TM \xrightarrow{\pi}_{TM} M$ is defined by

$$\pi_{TM}\, v_p = p$$

Given any chart $U \xrightarrow{x} V \subseteq \mathbb{R}^n$, we may define its tangent lift $TU \xrightarrow{Tx} V \times \mathbb{R}^n$.

Definition 5.9 (*natural chart*). A chart $U \xrightarrow{x} \mathbb{R}^n$ induces a natural chart $TU \xrightarrow{Tx} \mathbb{R}^{2n}$ with

$$Tx(v_i \partial_{x_i}\big|_p) = (x_1 p, \cdots, x_n p, v_1, \cdots v_n)$$

The mapping on the first n indices are apparent. Fix $i = 1$, note that Tx is linear in v_i.

$$\begin{aligned} Tx\left(\partial_{x_1}\big|_p\right) &= (x_1 p, \cdots, x_n p) \times \partial_{x_1}\big|_p x \\ &= (x_1 p, \cdots, x_n p) \times \partial_1\big|_p x x^{-1} \\ &= (x_1 p, \cdots, x_n p, 1, 0 \cdots) \end{aligned}$$

The tangent lift of each chart provides a chart on TM. It maps points to coordinates and tangent vectors to their representations.

Theorem 5.6 (*manifold structure of tangent bundle*). For a smooth n-manifold M, TM is a smooth $2n$-manifold in a natural way and $\pi_{TM} : TM \to M$ a smooth map.

Proof: Consider the coordinate transforms $(Ty)(Tx^{-1}) : (p, v) \to (q, w)$. Their components are smooth:

$$\begin{aligned} q &= \left(yx^{-1}\right) p \\ w &= J_{x \to y}\big|_p v \end{aligned}$$

Moreover, $Tx(TU \cap TV)$ is open since $J_{x \to y}$ is a homeomorphism. Projection is smooth since

$$(x\pi(Tx)^{-1})(p, v) = p$$

Definition 5.10 (*(local) trivialization*). A diffeomorphism $TM \xrightarrow{F} M \times V$ that's linear in its second argument and such that $\pi_1 F = \pi_{TM}$ is called a trivialization of TM, in which case TM is denoted trivial, i.e. reducible to the product with a vector space. For $U \subseteq M$, a trivialization of TU is a local trivialization of TM over U.

Proposition 5.7. In general, the tangent bundle cannot be globally trivialized, but it may be locally trivialized at every point over a chart domain.

Remark 5.2. In genral, here is no cotangent lift: the dual maps of pointed tangent lifts do not combine. This is because $T_pM \xrightarrow{T_pf} T_{fp}N$ may not be surjective.

Definition 5.11 (*natural (cotangent) chart*). Given a chart (U, x). Define (T^*U, T^*x) via

$$\begin{aligned} T^*U &= \bigcup_{p \in U} T_p^* M \\ T^*x\left(\omega_i\, \partial_{x_i}\big|_p\right) &= (x_1 p, \cdots, x_n p, \omega_1, \cdots \omega_n) \end{aligned}$$

Definition 5.12 (*dual of linear map*). Give a linear map $V \xrightarrow{T} W$, its dual is the map $W^* \xrightarrow{T^*} V^*$ such that for every $\omega \in W^*$, the diagram below commutes:

$$\begin{tikzcd}
& \mathbb{R} & \\
V \arrow[ur, "T^*\omega"] \arrow[rr, "T"'] & & W \arrow[ul, "\omega"']
\end{tikzcd}$$

In other words, $T^*\omega$ pulls ω back along T such that

$$\forall v \in V : \omega\,(Tv) = (T^*\omega)\,v$$

Proposition 5.8. $V^* \xrightarrow{(T^*)^{-1}} W^* = V^* \xrightarrow{(T^{-1})^*} W^*$

Proof: take any $T^*\omega \in V^*$, we wish to show

$$(T^{-1})^*\,(T^*\omega) = (T^*)^{-1}\,(T^*\omega) = \omega$$

To do this, take any $w \in W$ and invoke the duality relation in definition 5.12 twice

$$[(T^{-1})^*\,(T^*\omega)]\,w = (T^*\omega)\,(T^{-1}\,w) = \omega\,(TT^{-1}\,w) = \omega w$$

Consider the category of vector spaces with morphisms invertible map. The inverse and dual define two contravariant endofunctors since $(AB)^* = B^*A^*$ and $(AB)^{-1} = B^{-1}A^{-1}$. Proposition 5.8 shows that they commute.

Definition 5.13 (*contragredient*). The **contragredient** functor $[(\cdot)^{-1}]^*$ is a covariant functor resulting from composing the dual and inverse functors.

$$\left[(\cdot)^{-1}\right]^* V = V^*$$
$$\left[(\cdot)^{-1}\right]^* \left(V \xrightarrow{T} W\right) = V^* \xrightarrow{(T^{-1})^*} W^*$$

Dual and inverses are involutory and commute, so the contragredient is also involutory.

The lemma below characterizes the contragredient property.

Lemma 5.9. $V^* \xrightarrow{A} W^*$ is contragredient $V \xrightarrow{B} W$ when

$$\forall \nu \in V^*, v \in V : \nu\,v = (A\,\nu)\,(B\,v)$$

Proof: let $v = B^{-1}w$ for $w \in W$, then the condition is equivalent to

$$\forall \nu \in V^*, w \in W : \nu(B^{-1}\,w) = (A\,\nu)\,w$$

Invoke the definition of duality, this implies $A = (B^{-1})^*$.

Theorem 5.10 (*natural charts are contragredient*). Restricting our attention to principal components, $\forall p \in M$ and a chart x covering p, $T_pM \xrightarrow{T_p^* x} \mathbb{R}^n$ is the contragredient of $T_pM \xrightarrow{T_p x} \mathbb{R}^n$ when we invoke the natural identification $(R^n)^* \cong R^n$, having fixed a basis.

$$\left[T_p^* M \xrightarrow{T_p^* x} \mathbb{R}^n \cong (\mathbb{R}^n)^*\right] = \left[\left(T_pM \xrightarrow{T_p x} \mathbb{R}^n\right)^{-1}\right]^*$$

Proof: The chart representations of the tangent and cotangent vectors must be faithful to their relation in the original tangent and cotangent spaces, so by construction $\omega_p\, v_p = \omega\, v$. Invoke lemma 5.9.

Proposition 5.11. The cotangent chart's overlap map is $(T_p^* x)(T_p^* y)^{-1} = J^*_{x\to y}$. In other words, contangent vectors transform under $x \to y$ as $J^*_{y\to x}$.

Proof: the (inverse of the) cotangent map is contragredient to the (inverse of the) tangent map. The overlap map is the contragredient of the tangent overlap map $J_{x\to y}$. In detail:

$$\begin{aligned}
(T_p^* y)(T_p^* x)^{-1} &= \left[(T_p y)^{-1}\right]^* \left[\left[(T_p x)^{-1}\right]^*\right]^{-1} \\
&= \left[(T_p y)^{-1}\right]^* (T_p x)^* \\
&= \left[(T_p x)(T_p y)^{-1}\right]^* \\
&= J^*_{y\to x}
\end{aligned}$$

Remark 5.3. The contragredient transformation is the natural covariant functor capturing the duality between tangent and cotangent spaces.

5.3 Vector fields and Lie algebra

Summary 3. A smooth vector field is equivalent to a global derivation. Pointwise Lie derivative is equivalent to the pointwise application of the vector field. The Lie bracket defines an action of vector fields on themselves which also acts as a Lebneiz derivation when multiplication is taken to be composition. The pushforward and pullback of vector fields are naturally defined to respect the pushforward and pullback of scalar fields. The pushforward $\varphi_* X$ corresponds to the natural representation of X under the identification created by $\varphi_* X$.

Definition 5.14 (*vector field*). A smooth vector field on M is a smooth section $M \xrightarrow{X} TM$ such that $M \xrightarrow{X} TM \xrightarrow{\pi} M = \mathrm{id}$. The set of all smooth vector fields on M is denoted $\mathfrak{X}(M)$.

Definition 5.15 (*vector space structure of* $\mathfrak{X}(M)$). The following operations for $X, Y \in \mathfrak{X}(M), c \in \mathbb{R}$ makes $\mathfrak{X}(M)$ a vector space over the reals.

$$\begin{aligned}
(X + Y)\, p &= X\, p + Y\, p \\
(cX)\, p &= c(X\, p)
\end{aligned}$$

Definition 5.16 ($\mathfrak{X}(M)$ *is a* $C^\infty(M)$*-module*). For $f \in C^\infty(M)$, define the action of $f \in C^\infty(M)$ on $X \in \mathfrak{X}(M)$ as pointwise scaling.

$$(f X)\, p = (f\, p)(X\, p)$$

Recall that $C^\infty(M)$ is a ring with pointwise addition and multiplication, and $\mathfrak{X}(M)$ being an abelian group. The action above above is a ring action that's compatible with the inherent structure of $\mathfrak{X}(M)$ (i.e. commutes with the action above)

$$[(f + g \cdot h)X]\, p = (f X)\, p + [(gX)\, p] \cdot [(hX)\, p]$$

Definition 5.17 (*vector field along* $N \xrightarrow{f} M$). $N \xrightarrow{X} TM$ is a vector field along f if

$$\left(N \xrightarrow{X} TM \xrightarrow{\pi} M\right) = \left(N \xrightarrow{f} M\right)$$

The space of vector fields along f is denoted $\mathfrak{X}_f$. It is similarly a $C^\infty(N)$-module

Definition 5.18 (*(global) derivation*). A global derivation on $C^\infty(M)$ is a linear map $\mathcal{D} : C^\infty(M) \to C^\infty(M)$ such that

$$\mathcal{D}\,(f \cdot g) = (\mathcal{D}\, f) \cdot g + f \cdot (\mathcal{D}\, g)$$

We denote the set of such derivations by $\mathrm{Der}(C^\infty(M))$

Definition 5.19 (*Lie derivative*). A vector field $M \xrightarrow{X} TM$ induces a global derivation of scalar fields $\mathcal{L}_X \in \mathrm{Der}(C^\infty(M))$ defined via

$$(\mathcal{L}_X f)\, p = (X\, p)\, f$$

Given $M \xrightarrow{f} \mathbb{R}$, its Lie derivative $\mathcal{L}_X f : M \to \mathbb{R}$ evaluates the change of f along X pointwise. An equivalent definition is by noting the equation below, then $\mathcal{L}_X f = df \circ X$.

$$(X\, p)\, f = (TM \xrightarrow{df} \mathbb{R})\,(M \xrightarrow{X} TM)$$

Remark 5.4. We reemphasize the theme of delayed evaluation. Given a domain A with $a \in A$ and any map $\omega : A \to B$, we can construct a map $\mathrm{ev}_a : (A \to B) \to B$ such that

$$\mathrm{ev}_a\, \omega = \omega\, a$$

Note that ω is an argument on the left hand side and a map on the right hand side. One familiar example is the double-dual construction.

Similarly, consider a tangent vector $v_p : C^\infty(M) \to \mathbb{R} \cong T_pM$, it applies as a derivation on $f \in C^\infty(M)$. We may similarly introduce $\mathrm{ev}_f : T_p^*M \to \mathbb{R}$, or df, such that

$$\mathrm{ev}_f\, v_p = v_p\, f$$

Substitute $v_p = X\, p$ above to obtain $\mathcal{L}_X f = df \circ X$. Note that df is an applying map on the right hand side and an argument on the left hand side. The Lie derivative $\mathcal{L}_X$ is another example of delaying evaluation.

Proposition 5.12. Consider an open set $U \subset M$ and $X \in \mathfrak{X}(M)$. If $\mathcal{L}_X f = 0$ for all $f \in C^\infty(U)$, then $X|_U = 0$.

Proof: for any $p \in U$, choose a chart x and let f be the coordinate functions.

Definition 5.20 (*coordinate frame field*). Fixing a chart (U, x). the ordered set of fields $\{p \mapsto \partial_{x_i}|_p\}$ is called a coordinate (holonomic) field.

Every field $X \in \mathfrak{X}(M)$ apparantly defines a global derivation via $\mathcal{L}_X \in \mathrm{Der}(C^\infty(M))$. The converse is also true: we may think of smooth vector fields and derivations interchangeably.

Theorem 5.13 (*smooth vector fields biject with derivations*). Every $\mathcal{D} \in \mathrm{Der}(C^\infty(M))$ satisfies $\mathcal{D} = \mathcal{L}_X$ for a unique $X \in \mathfrak{X}(M)$.

Proof: Given $\mathcal{D}$, define a field X by $(X\,p)\,f = (\mathcal{D}\,f)\,p$. We need to show that $p \mapsto X\,p$ is smooth. It suffices to show that $p \mapsto (X\,p)\,f$ is smooth for every $f \in C^\infty(M)$ since we can take f to be chart components. Now $(X\,p)\,f = (\mathcal{D}\,f)\,p$ is smooth in p since $\mathcal{D}\,f \in C^\infty(M)$.

For uniqueness, suppose $\mathcal{L}_X = \mathcal{L}_{X'} = \mathcal{D}$, then $\mathcal{D} - \mathcal{D} = \mathcal{L}_{X-X'} = 0$ implies $X = X'$.

This theorem allows us to define the action of vector fields on real maps as derivations:

$$(X\,f)\,p := (X\,p)\,f$$

Remark 5.5. Vector fields account for all derivations only in the smooth case. (?)

Definition 5.21 (*Lie bracket*). Define the commutator of derivations

$$[\mathcal{D}_1, \mathcal{D}_2] = \mathcal{D}_1 \circ \mathcal{D}_2 - \mathcal{D}_2 \circ \mathcal{D}_1$$

Then $[\mathcal{D}_1, \mathcal{D}_2] \in \mathrm{Der}(C^\infty(M))$. The Lie bracket $[X, Y] \in \mathfrak{X}(M)$ is then defined naturally.

Proof: we derive equality by examining the derivation behavior of $[\mathcal{D}_1, \mathcal{D}_2]$ on $f \cdot g$.

$$\begin{aligned} \mathcal{D}_1\,(\mathcal{D}_2\,(f \cdot g)) &= \mathcal{D}_1\,\left[f \cdot (\mathcal{D}_2\,g) + g \cdot (\mathcal{D}_2\,f)\right] \\ &= f \cdot (\mathcal{D}_1 \mathcal{D}_2\,g) + (\mathcal{D}_1\,f) \cdot (\mathcal{D}_2\,g) + g \cdot (\mathcal{D}_1 \mathcal{D}_2\,f) + (\mathcal{D}_1\,g) \cdot (\mathcal{D}_2\,f) \end{aligned}$$

Subtract away the result for $\mathcal{D}_2 \mathcal{D}_1\,(fg)$ and we're left with

$$[\mathcal{D}_1, \mathcal{D}_2](fg) = f \cdot ([\mathcal{D}_1, \mathcal{D}_2]g) - g \cdot ([\mathcal{D}_1, \mathcal{D}_2]f)$$

Definition 5.22 (*Lie algebra*). A vector space $\mathfrak{a}$ is called a Lie algebra if it comes equipped with a antisymmetric bilinear multiplication map $[\,\cdot\,,\,\cdot\,] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$ such that $[x, \cdot\,] : \mathfrak{a} \to \mathfrak{a}$ is a derivation of $\mathfrak{a}$ for every $x \in \mathfrak{a}$. This is denoted the Jacobi identity.

$$[x, [y, z]] = [y, [x, z]] + [[x, y], z]$$

Remark 5.6. The bracket $[x, y]$ simultaneously denotes the vector object x acting on y and multiplication for the Lebneiz rule. The Lie axiom then reads that composition is derivation.

$$x(yz) = y(xz) + (xy)z$$

Theorem 5.14 (*commutators satisfy the Jacobi identity*). Whenever multiplication and subtraction is well-defined, the commutator $[a, b] = ab - ba$ satisfies the Jacobi identity.

Proof: Direct computation

$$\begin{aligned}
[x, [y, z]] &= xyz - xzy - yzx + zyx \\
[y, [x, z]] &= yxz - yzx - xzy + zxy \\
[[x, y], z] &= xyz - yxz - zxy + zyx
\end{aligned}$$

Commutators are apparantly also skew-symmetric. Define multiplication on $\mathrm{Der}(C^\infty(M))$ by composition. The Lie bracket is obviously bilinear.

Corollary 5.1. $\mathfrak{X}(M)$ is a Lie algebra.

Definition 5.23 (*abelian, subalgebra, and ideal*). A Lie algebra $\mathfrak{a}$ is abelian if the bracket always vanish. A subspace $\mathfrak{h}$ is a sub-algebra if it is closed under the bracket, and it is an ideal if if $[\mathfrak{a}, \mathfrak{h}] \subset \mathfrak{h}$.

Definition 5.24 (*pullback, pushforward of vector field*). Each diffeomorphism $M \xrightarrow{\varphi} N$ induces a pullback $\varphi^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$ and pushforward $\varphi_* : \mathfrak{X}(M) \to \mathfrak{X}(N)$ as follows:

$$\begin{tikzcd}
M \arrow[r, "\varphi"] \arrow[d, dashed, "\varphi^* Y"] & N \arrow[d, "Y"] \\
TM & TN \arrow[l, "T\varphi^{-1}"]
\end{tikzcd}
\qquad
\begin{tikzcd}
M \arrow[d, "X"'] & N \arrow[l, "\varphi^{-1}"'] \arrow[d, dashed, "\varphi_* X"] \\
TM \arrow[r, "T\varphi"'] & TN
\end{tikzcd}$$

Theorem 5.15 (*universal property of vector field pullback*). The pullback of a vector field is defined such that the Lie derivative is natural with respect to pushforwards and pullbacks of maps and fields along diffeomorphisms. Deriving a map pullback by a field pullback at an applied point is equivalent to deriving the map by the field at the original point.

$$\begin{aligned}
\varphi^*(Y\, g) &= (\varphi^* Y)\, (\varphi^* g) \\
\varphi_*(X\, f) &= (\varphi_* X)\, (\varphi_* f)
\end{aligned}$$

$$\begin{tikzcd}
TM & TN \\
M_{(p)} \arrow[u, dashed, "\varphi^* Y"] \arrow[r, "\varphi"] \arrow[d, dashed, "\varphi^* g"'] & N_{(q)} \arrow[u, "Y"'] \arrow[dl, "g"] \\
\mathbb{R} &
\end{tikzcd}
\qquad
\begin{tikzcd}
TM & TN \\
M \arrow[u, "X"] \arrow[r, "\varphi"] \arrow[d, "f"'] & N \arrow[u, dashed, "\varphi_* X"'] \arrow[dl, dashed, "\varphi_* f"] \\
\mathbb{R} &
\end{tikzcd}$$

Proof: Choose $p \in M$ and $q = \varphi\, p$. On the second step we use the tangent map definition.

The pushforward equation follows since $\varphi_* = (\varphi^{-1})^*$.

$$\begin{aligned}
[(\varphi^* Y)\,(\varphi^* g)]\,p &= [(T\varphi^{-1} \circ Y \circ \varphi)\,(g \circ \varphi)]\,p \\
&= [(T\varphi^{-1} \circ Y)\,q]\,(g\,\varphi) \\
&= [(T\varphi \circ T\varphi^{-1} \circ Y)\,q]\,g \\
&= (Y\,q)\,g = [(Y\,g) \circ \varphi]\,p \\
&= \varphi_*(Y\,g)\,p
\end{aligned}$$

Definition 5.25 (*f-related fields*). Given a not necessarily diffeomorphic map $M \xrightarrow{f} N$, the field $M \xrightarrow{X} TM$ is f-related to $N \xrightarrow{Y} TN$ if the following square commutes:

$$\begin{tikzcd}
TM \arrow[r, "Tf"] & TN \\
M \arrow[u, "X"] \arrow[r, "f"] & N \arrow[u, "Y"']
\end{tikzcd}$$

The following lemma similarly characterizes the universal property of f-relatedness.

Lemma 5.16. X, Y are f-related if and only if for all $N \xrightarrow{g} \mathbb{R}$,

$$X\,(gf) = (Y\,g)f \iff X\,(f^*\,g) = f^*\,(Y\,g)$$

Proof: Consider the diagram below

$$\begin{tikzcd}
TM \arrow[r, "Tf"] & TN \\
M_{(p)} \arrow[u, "X"] \arrow[r, "f"] \arrow[d, "gf"] & N \arrow[u, "Y"'] \arrow[dl, "g"] \\
\mathbb{R} &
\end{tikzcd}$$

The f-related condition $(Tf)X = Yf$ is of type $M \to TN$. Two objects of type $M \to TN$ are equal if they induce the same derivation in TN for all $p \in M$. That is, if

$$\begin{aligned}
[(Tf \circ X)\,p]\,g &= [(Y \circ f)\,p]\,g \\
[(Tf \circ X)\,g]\,p &= [Y\,(f\,p)]\,g \\
[X\,(g \circ f)]\,p &= (Y\,g)\,(f\,p) \\
[X\,(g \circ f)]\,p &= [(Y\,g) \circ f]\,p \\
X\,(gf) &= (Y\,g)f
\end{aligned}$$

Each step is an equivalent derivation.

Remark 5.7. The lemma above is a tautology considering remark 5.1. The pointwise tangent map (which extends to vector fields) is defined in terms of the pullback of derivations.

Proposition 5.17. Given a smooth $M \xrightarrow{f} N$ and $X_i \sim_f Y_i$ for $i = 1, 2$, then

$$[X_1, X_2] \sim_f [Y_1, Y_2]$$

In particular, if $M \xrightarrow{\varphi} M$ is a diffeomorphism, then $[\varphi_* X_1, \varphi_* X_2] = \varphi_*[X_1, X_2]$.

Proof: Invoke the lemma two times in reverse directions: at the beginning and the end

$$\begin{aligned}
[X_1, X_2]\,(gf) &= (X_1 X_2)\,(gf) - (X_2 X_1)\,(gf) \\
&= [(Y_1 Y_2 - Y_2 Y_1)\, g] \circ f \\
&= ([Y_1, Y_2] g) \circ f
\end{aligned}$$

5.4 Vector field and flow

Summary 4. Flow generates a vector field by spatially global time derivation. A vector field generates flow. A local flow box is smooth aggregation of pointwise centered integral curves. Local flow boxes always uniquely exist, and the maximal flow is always defined on an open spatial-temporal neighborhood.

Definition 5.26 (*integral curve*). A curve $I \xrightarrow{c} M$ is an integral curve for $M \xrightarrow{X} TM$ if

$$\forall t \in I : \dot{c} = X \circ c$$

Recall definition 5.5. Then

$$[X\, c(t)]\, f = (f \circ c)'(t)$$

Within a local chart (U, x) for which $X = X_i \partial_{x_i}$, this constitutes a system of ODEs

$$d_t x_i = X_i$$

Definition 5.27 (*(complete) flow*). A complete flow is a group homomorphism $\varphi : \mathbb{R} \to \mathrm{Diff}(M)$ from the additive group $\mathbb{R}$. A (local) flow is defined on some neighborhood of 0.

Recall that $(\mathbb{R} \to (M \to M)) \cong (\mathbb{R} \times M \to M)$. Fixing the second argument, differentiating along time yields $d_t\varphi : \mathbb{R} \to (M \to TM)$ below, where $t_0 = 0$ unless specified.

$$X^{\varphi}_{t_0}\, p = d_t\big|_{t_0} \varphi(x, p)$$

Recall that a Banach space is a complete, normed vector space in which every Cauchy sequence converges.

Theorem 5.18 (*existence and uniqueness theorem*). Given a Banach space $E, U \subset E$ and a smooth map $U \xrightarrow{F} E$. For any $x_0 \in V \subset U$, there is a unique locally smooth time-evolution $\varphi : (-\epsilon, \epsilon) \times V \to U$ which flows along X. In other words, $\forall x \in V, c(0) = x$:

$$c_x'(t) = F(c_x(t))$$

Lemma 5.19. If $c_1 : (-\epsilon_1, \epsilon_1) \to M$ and $c_2, (-\epsilon_2, \epsilon_2) \to M$ are integral curves of X with $c_1(0) = c_2(0)$, then $c_1 = c_2$ on the intersection of their domains.

Proof: Let $I = (-\epsilon_1, \epsilon_1) \cap (-\epsilon_2, \epsilon_2)$ and $K = \{t \in I : c_1(t) = c_2(t)\}$. First note that K is closed since M is Hausdorff, so $\mathbb{R} - K$ is open by continuity of c_1, c_2 . Now K is also open: for any $t_0 \in K$, choose a chart x covering $p = c_1(t_0) = c_2(t_0)$ and applying theorem 5.18 shows that K also contains a small interval $(t_0 - \epsilon, t_0 + \epsilon)$. By I connected, $K = I$.

Definition 5.28 (*flow box*). Given smooth vector field $M \xrightarrow{X} TM$, a flow box for X at p is a triple (U, a, φ^X) such that

1. $p \in U$, $\varphi^X : (-a, a) \to (U \to M)$ is smooth.
2. $\varphi^X(0, \cdot) : U \to M$ is the inclusion map.
3. $\forall p \in U, \varphi^X(\cdot, p) : (-a, a) \to M$ is an integral curve for x.

In short, a flow box is a smooth, local time-evolution map that is an inclusion when fixing $t = 0$, and an integral curve when fixing p.

Proposition 5.20. $\varphi^X(t, \varphi^X(s, p)) = \varphi^X(t + s, p)$

Proof: Fix s, p, then $c_1(0) = c_2(0) = \varphi^X(s, p)$. Invoke the uniqueness theorem.

$$I \xrightarrow{c_1} M = t \mapsto \varphi^X(t, \varphi^X(s, p))$$
$$I \xrightarrow{c_2} M = t \mapsto \varphi^X(t + s, p)$$

Example 5.1 (*complete fields do not form a vector space*). Over $\mathbb{R}^2$, the fields $y^2\partial_x$ and $x^2\partial_y$, corresponding to $\{\dot{x} = y^2, \dot{y} = 0\}$ etc. are complete. However, their sum is not complete:

$$y^2\partial_x + x^2\partial_y \iff \begin{cases} d_t x = y^2 \\ d_t y = x^2 \end{cases}$$

A vector field can be not complete because:

1. The curve approaches some point in a finite time where the vector field is not defined.
2. The curve goes to infinity in finite time.

Here, the tangent line to the field is $y = x$. Consider solving the equation for $x(t) = y(t)$ and $x(0) = 1$, the solution is $x(t) = 1/(1 - t)$.

Theorem 5.21 (*local existence of unique flow box*). Given a smooth field X on a n-manifold M. For every point $p_0 \in M$ there exists a flow box for X at p_0. Any two such flow boxes agree on their domain.

Proof: Work in a chart. This is acorollary of theorem 5.18.

The lemma below allows us to compose local flows.

Lemma 5.22. Given smooth fields $X_1, \cdots, X_k$ with local flow boxes $\{U_j, \varphi^j\}$. Given $p_0 \in M$ and a neighborhood O of p_0, there is an open set $U \subset \bigcap U_j$ and $\epsilon > 0$ such that

$$\varphi^k_{t_k} \circ \cdots \circ \varphi^1_{t_1}$$

is defined on U and maps into O whenever $t_j \in (-\epsilon, \epsilon)$.

Proof: By continuity we can choose V_k, ϵ_k such that $\varphi^k_{t_k \in (-\epsilon_k, \epsilon_k)}(V_k) \subseteq O$. Inductively choose V_j, ϵ_j. Let $U = \bigcap V_j, \epsilon = \min\{\epsilon_j\}$.

Given an integral curve $c_p : (a, b) \to M$ and $c_p(0) = p$. Consider the limit

$$p' \lim_{t \to b^-} c_p(t)$$

If $p_1 \in M$, concatenating with an integral curve beginning at p' yields a longer integral curve.

Definition 5.29 (*maximal integral curve*). Fixing $p \in M$ and a smooth vector field X. Consider the collection $\mathcal{J}_p$ of all pairs (J, α) such that the open interval $J \supset \{0\}$ and $\alpha : J \to M$ is an integral curve of X centered at p. The maximal interval is

$$J^X_p = (T^-_{p,X}, T^+_{p,X}) = \bigcup_{(J,\alpha) \in \mathcal{J}_p}$$

The maximal curve $c(t) = \alpha(t)$ whenever $t \in J$.

Definition 5.30 (*maximal flow (domain)*). The maximal flow domain is defined as

$$\mathcal{D}_X = \bigcup_p J^X_p \times \{p\}$$

The maximal flow $\varphi^X : \mathbb{R} \times \mathcal{D}_X \to M$ is naturally defined. Fixing t, the maximal domain of $\varphi^X_t = p \mapsto \varphi^X_t(p)$ is denoted

$$\mathcal{D}^t_X = \{p : t \in (T^-_{p,X}, T^+_{p,X})\} \subset M$$

The maximal domain $\mathcal{D}^t_X$ denotes the subset of M admitting an integral curve to time t.

Definition 5.31 (*complete vector field*). The vector field X is complete if all its integral curves are defined on $\mathbb{R}$. Equivalently, if it generates a complete maximal flow $\varphi : \mathbb{R} \to \mathrm{Diff}(M)$ such that

$$\forall t_0 \in \mathbb{R} : d_t\big|_{t_0} \varphi(t, p) = X\,p$$

The proposition below gives a feel for the implications of maximality.

Proposition 5.23. $X \in \mathfrak{X}(M)$ is complete if it permits a flow on $(-a, a) \times M$.

Proof: Every point p permits an integral curve, so integral curves may be extended to $\mathbb{R}$.

Theorem 5.24 (*characterization of maximal flow*). For any $X \in \mathfrak{X}(M)$, its maximal flow domain $\mathcal{D}_X$ is an open neighborhood of $\{0\} \times M \subset \mathbb{R} \times M$ and $\varphi^X : \mathcal{D}_X \to M$ is smooth. Whenever both sides are defined

$$\varphi^X(t+s,p) = \varphi^X(t, \varphi^X(s,p))$$

If the right side is defined, then so is the left. The converse is true if $t, s \geq 0$ or $t, s \leq 0$.

Proof: Fix p and let $q = \varphi^X(s,p)$. Suppose the right side is defined, then $s \in (T_p^-, T_p^+)$ and $t \in (T_q^-, T_q^+)$. Consider the time-shift of $\varphi^X(\cdot, p)$ by s. The maximal domain of $\psi : \tau \to \varphi^X(s+\tau, p)$ is $(T_p^- - s, T_p^+ - s)$ and ψ is an integral curve at $q = \psi^X(s,p)$. By maximality $(T_p^- - s, T_p^+ - s) = (T_q^-, T_q^+)$. Then $t \in (T_q^-, T_q^+) \implies t \in (T_p^- - s, T_p^+ - s)$, so the left side is defined. Conversely, let $t, s \geq 0$. Then $t+s \in (T_p^-, T_p^+) \supset [0, t+s]$ implies $t, s \in (T_p^-, T_p^+)$. By maximality the right hand side is defined.

To show open and smoothness, consider $\mathcal{S} \subset \mathcal{D}_X \subset \mathbb{R} \times M$ such that $(t,p) \in \mathcal{S}$ if there exists an interval $J \supset \{0, t\}$ and an open $U \subset M$ such that $\varphi^X_{|J \times U}$ is smooth. By construction $\mathcal{S}$ is open, and $\mathcal{S} = \mathcal{D}_X$. Suppose otherwise for contradiction and $(t', p') \in \mathcal{D}_X - \mathcal{S}$. Without loss of generality let $t' > 0$. Fix p' and consider $\tau = \sup \{t \mid (t, p') \in \mathcal{S}\}$. Now $\tau > 0$ since $(0, p_0)$ is contained in some flow box and $\tau \leq t'$. **todo**.

The corollaries below elucidate the structure of the maximal domain.

Corollary 5.2. Given $s, t > 0$, the domain of $\varphi_s^X \circ \varphi_t^X \subset \mathcal{D}_X^{s+t}$.

Proof: Recall that $\varphi_{t+s}^X = \varphi_s^X \circ \varphi_t^X$ and $\mathcal{D}_X^{s+t}$ is the domain of φ_{t+s}^X.

Corollary 5.3. For every t, $\mathcal{D}_X^t \subset M$ is open and $\varphi_t^X(\mathcal{D}_X^t) = \mathcal{D}_X^{-t}$.

Proof: Open is trivial. Now $\varphi_t^X(\mathcal{D}_X^t) \subset \mathcal{D}_X^{-t}$ since $\varphi_{-t}^X \left[\varphi_t^X(\mathcal{D}_X^t)\right] = \mathrm{Id}_{|\mathcal{D}_X^t}$ is well-defined. Any extraneous point in D_X^{-t} yields an integral curve contradicting the maximality of $\mathcal{D}_X^t$.

Definition 5.32 (*support of a vector field*). The support of a vector field $X \in \mathfrak{X}(M)$, denoted $\mathrm{supp}(X)$, is the closure of $\{p : X(p) \neq 0\}$.

Proposition 5.25. Every vector field with compact support is complete.

Proof: Points outside the support has a neighborhood of vanishing field, and invoking uniqueness on the trivial solution in the neighborhood shows that integral curves of the point vanish. Thus integral curves of points within the support stay in the support. For every $\epsilon > 0$, consider $U_\epsilon \subseteq M$ consisting of points such that $(-\epsilon, \epsilon) \subset (T_p^-, T_p^+)$, each of these sets is open. By local flow property $\{U_\epsilon\}_{\epsilon > 0}$ is a nested open cover for M. By compactness choose $\epsilon_0 : M \subseteq U_{\epsilon_0}$. Invoke proposition 5.23.

Theorem 5.26 (*straightening theorem*). Given $X \in \mathfrak{X}(M)$ and $X(p) \neq 0$ for some $p \in M$, then there is a chart (U, x) covering p such that the representation of X on U is simply ∂_{x_1}.

Proof: By locality let $M = \mathbb{R}^n, p = 0$ with standard coordinates u_i. Up to a linear transformation of the chart x, center $X(0) = \partial_{u_1}\big|_0$. Let φ denote a local flowbox about

$p = 0$. Define

$$\chi(a_1, \cdots, a_n) = \varphi_{a_1}(0, a_2, \cdots, a_n)$$

Note that $\chi(a_1, \cdots, a_n)$ is the result of flowing a point $(a_2, \cdots, a_n)$ on the hyperplane $u_1 = 0$ along for time a_1. This is the straightened inverse chart (up to shrinking) if $X\,p \neq 0$. We first show that $\partial_{u_1}\big|_a$ is indeed the representation of X at $\chi(a)$.

$$\begin{aligned}
\left[(T\chi)\,\partial_{u_1}\big|_a\right]\, f &= \partial_{u_1}\big|_a(f \circ \chi) \\
&= \lim_{t\to 0} \frac{1}{t}\left[f(\varphi_{a_1+t}(0, a_2, \cdots, a_n)) - f(\chi(a))\right] \\
&= \lim_{t\to 0} \frac{1}{t}\left[f(\varphi_t(\chi\, a)) - f(\chi\, a)\right] \\
&= (X\, f)(\chi\, a)
\end{aligned}$$

The Jacobian $T_0\chi$ is the identity. Invoke the inverse function theorem to make χ^{-1} a chart.

5.5 Lie derivative

Summary 5. The flow generated by a vector field creates a natural identification of points and pointed tangent spaces for an open temporal neighborhood about the origin. This allows us to define a consistent derivative.

One difficulty about defining global derivation on a manifold is the lack of a global structure on a manifold. The manifold counterpart of a global direction (e.g. ∂_x in $\mathbb{R}^n$) is a vector field $X \in \mathfrak{X}(M)$. This easily subsumes the real analysis definition of directional derivatives.

Another difficulty is about establishing a uniform definition of differentiation across the variety of the types we wish to differentiate. The directional derivative of a scalar field is easily defined by point-wise application of the vector fields, but it seems far from clear how we may extend this definition to vector fields, co-vector fields, or more complex objects.

We begin by reconsidering the operation $X\,f$. At each p, the existence of local flow boxes φ^X creates, for each t in a neighborhood $I_\delta = (-\delta, \delta)$ of 0, a natural identification between the one-point spaces $\{p\}$ and $\{\varphi_t^X(p)\}$. We define the derivation of f by X at p by differentiating the value of f along this continuum of spaces, indexed by t.

Definition 5.33 (*scalar Lie derivative by flow*).

$$(X\, f)\, p \equiv \lim_{t\to 0} \frac{f(\varphi^X(t,p)) - f(p)}{t} = \partial_t\big|_0[f \circ \varphi^X(t,p)] = \left[\partial_t\big|_0 \varphi^X(t,p)\right]\, f = (X\, p)\, f$$

This interpretation by flow is thus consistent with definition 5.19. We rewrite it more suggestively in terms of pushforwards.

$$(X\, f)\, p \equiv \partial_t\big|_0[f \circ \varphi^X(t,p)] = \left[\partial_t\big|_0 \left(\varphi_t^X\right)^* f\right]\, p$$

When X is not complete, for a fixed t, φ_t^X may not be well-defined for all M. However, the existence of local flow-boxes implies that $M = \bigcup_{t\neq 0} \mathcal{D}_X^t$. For any fixed p, there is still a neighborhood of $t=0$ on which $\varphi_t^X\, p$ and $(\varphi_t^{X*} f)\, p$ are well-defined. This allows us to define the derivative above. Similarly, we may well-define time-derivative of a pullback which satisfies universal property 5.15.

$$d_t\big|_0 \left(\varphi_t^{X*}\, Y\right)\, p = d_t\big|_0 \left(T\varphi_t^X\right)^{-1} Y \left(\varphi_t^X\, p\right)$$

We now consider differentiating a vector field $Y \in \mathfrak{X}(M)$ along the direction given by $X \in \mathfrak{X}(M)$. Vector fields take values in pointed tangent spaces. Similarly, a local automorphism $\varphi_t^X \in \operatorname{Diff}(M)$ creates, for each $p \in U$, a natural interval I_δ of tangent spaces $\{T_{\varphi_t^X(p)}M\}_{t\in I_\delta}$ via the tangent map which can be naturally "collapsed" together. Universal property 5.15 suggests that a natural representation of $Y_p \in T_pM$ in $T_{\varphi_t^X(p)}M$ is the pullback. The other is the pushforward, but they only differ up to a sign.

$$(\varphi_t^{X*} Y)_p = \left[(T\varphi_{-t}^X) \circ Y \circ \varphi_t^X\right]\, p$$

Definition 5.34 (*Lie derivative of vector fields*). Given $X, Y \in \mathfrak{X}(M)$, the Lie derivative of Y by X is defined as

$$\mathcal{L}_X Y = \partial_t\big|_0 \varphi_t^{X*} Y$$

The derivative measures how the representation of Y changes along the flow generated by X. Based on this definition, it is obvious that $\mathcal{L}_X Y \in \mathfrak{X}(M)$.

We now show that the Lie derivative above is the Lie bracket. This is among the highly nontrivial structural theorems in differential geometry. To do this, we need the following technical lemma to characterize the local behavior of $\varphi_t^{X*} f$.

Lemma 5.27. Given $X \in \mathfrak{X}(M)$ and $f \in C^\infty(U)$ for an open $U \supset \{p\}$, there is an open interval I_δ and a neighborhood V of p such that $\varphi^X(I_\delta \times V) \subset U$ and there exists $g \in C^\infty(I_\delta \times V)$ such that for all $(t, q) \in I_\delta \times V$,

$$f(\varphi^X(t,q)) = f(q) + t g(t,q), \quad g(0,q) = (X\, f)\, q$$

Suppress q and let $g_t : V \to M$ denote the partial application $q \mapsto g(t,q)$, we have

$$\varphi_t^{X*} f = f + t g_t, \quad g_0 = X\, f$$

Proof: $I_\delta \times V$ exists by a smooth local flow box. Note that $\tau, q \mapsto f(\varphi^X(\tau,q)) - f(q)$ is smooth on $I_\delta \times V$ and vanishes at $\tau = 0$. Consider

$$g(t,q) = \int_0^1 \partial_\tau\big|_{st} \left[f(\varphi^X(\tau,q)) - f(q)\right]\, ds$$

This is our desired quantity. Note that $\partial_\tau\big|_{st} = t\partial_s\big|_s$ under the substitution $\tau = st$.

$$\begin{aligned} tg(t,q) &= \int_0^1 (t\partial_\tau\big|_{st}) \left[f(\varphi^X(\tau,q)) - f(q)\right]\, ds \\ &= \int_0^1 \partial_s\big|_s \left[f(\varphi^X(st,q)) - f(q)\right]\, ds \\ &= f(\varphi^X(t,q)) - f(q) \end{aligned}$$

For the second part, note that g is smooth and take $t \to 0$.

Theorem 5.28 (*Lie derivative coincides with Lie bracket*). $\mathcal{L}_X Y = [X, Y]$.

Proof: Expand the Lie derivative definition by the pushforward definition 5.24.

$$\begin{aligned}
(\mathcal{L}_X Y)\, f &= \left(\partial_t\big|_0 \varphi_t^{X*} Y\right)\, f \\
&= \partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y \circ \varphi_t^X\right]\, f \\
&= \partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y\right] (f \circ \varphi_t^X)
\end{aligned}$$

Check types: $f \circ \varphi_t^X : C^\infty(M)$ and $\left[(T\varphi_{-t}^X) \circ Y\right] : \mathrm{Der}(C^\infty(M))$. Invoke lemma 5.27.

$$\begin{aligned}
\partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y\right] (f \circ \varphi_t^X) &= \partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y\right] [f + t g_t] \\
&= \partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y\right] f + \big|_{t=0} \left[(T\varphi_{-t}^X) \circ Y\right] g_t
\end{aligned}$$

In the limit $t \to 0$, $T\varphi_{-t}^X$ is trivial and the second term becomes

$$\big|_{t=0} \left[(T\varphi_{-t}^X) \circ Y\right] g_t = Y\, g_0 = Y\,(X\, f) = (Y \circ X)\, f$$

Simplify the first term by the definition of the pointed tangent map. Carefully note the action of the tangent map and recall the flow definition of scalar derivative.

$$\partial_t\big|_0 \left[(T\varphi_{-t}^X) \circ Y\right] f = \partial_t\big|_0 Y\,(f \circ \varphi_{-t}^X) = -\partial_t\big|_0 \varphi_t^{X*}(Y\, f) = -(X \circ Y)\, f$$

It is evident from the flow interpretation that the Lie derivative should be another vector field. At every point, the Lie derivative simply differentiates a vector-valued function on a neighborhood of identified tangent spaces. However, it is nontrivial that the Lie bracket should be a vector field, being the difference $XY - YX$ of two composed differentiation. The proposition below justify this: the higher-order terms cancel in local coordinates.

Proposition 5.29. Suppose in a local basis x that $X = \sum X_i \partial_{x_i}$ and $Y = \sum Y_i \partial_{x_i}$, then

$$[X, Y] = \sum_{ij} \left[X_j \left(\partial_{x_j} Y_i\right) - Y_j \left(\partial_{x_j} X_i\right)\right] \partial_{x_i}$$

Proof: Expand by definition.

$$\begin{aligned}
[X, Y]\, f &= X\,(Y_i \cdot \partial_{x_i} f) - Y\,(X_i \cdot \partial_{x_i} f) \\
&= X_j \partial_{x_j}(Y_i \cdot \partial_{x_i} f) - Y_j \partial_{x_j}(X_i \cdot \partial_{x_i} f) \\
&= X_j(\partial_{x_j} Y_i)(\partial_{x_i} f) + X_j Y_i(\partial^2_{x_i, x_j} f) - Y_j(\partial_{x_j} X_i)(\partial_{x_i} f) - Y_j X_i(\partial^2_{x_i, x_j} f) \\
&= \left[X_j(\partial_{x_j} Y_i) - Y_j(\partial_{x_j} X_i)\right] \partial_{x_i} f
\end{aligned}$$

Proposition 5.30. For $X, Y \in \mathfrak{X}(M)$, $[X, Y] = 0 \iff \varphi_t^X \varphi_s^Y = \varphi_s^Y \varphi_t^X$ whenever defined.

Proof: By the flow property, $\varphi_t^X \varphi_s^Y = \varphi_s^Y \varphi_t^X$ is defined and true only when

$$\varphi_s^Y = \varphi_{-t}^X \varphi_s^Y \varphi_t^X = \varphi_s^{\varphi_t^{X*} Y}$$

5.6 1-Forms

Definition 5.35 (*covector field*). A 1-form, or covector field, is a smooth section of the projection $\pi : T^*M \to M$. Such fields is a module over $C^r(M)$ and denoted $\mathfrak{X}^*(M)$.

Definition 5.36 (*differential*). A $C^{r\geq 1}$ function $f : M \to \mathbb{R}$ induces a 1-form $df \in \mathfrak{X}^*(M)$

$$(df\, X)\, p = df\big|_p (X\, p) = (X\, p)\, f = (X\, f)\, p$$

The differential map is of type $d : (M \to \mathbb{R}) \to \mathfrak{X}^*(M)$. A one-form is of type $M \to T^*M$ by definition, $TM \to \mathbb{R}$ by definition of the dual construction of the cotangent space, and by extension a linear map $\mathfrak{X}(M) \to C^\infty(M)$ by global application.

Proposition 5.31. Product rule for differentials

$$d(f \cdot g) = g \cdot df + f \cdot dg$$

Proof: $d(f \cdot g)\, X = X(f \cdot g) = g \cdot (X\, f) + f \cdot (X\, g) = g \cdot df + f \cdot dg$

Definition 5.37 (*coordinate (holonomic) coframe field*). Given a chart (U, x), the covector fields $\{dx_i\}$ is a basis for $T^*_{p\in U}M$. They constitute the coordinate coframe field over U. For every $X \in \mathfrak{X}(M)$:

$$X_{|U} = \sum (dx_i\, X) \cdot \partial_{x_i}$$

Definition 5.38 (*pullback of 1-form*). Given a C^∞ map $\varphi : M \to N$ (not necessarily diffeomorphism), the pullback of $\omega \in \mathfrak{X}^*(N)$ by φ is

$$\varphi^*\omega\, X = \omega\, (\varphi_* X)$$

In terms of point-wise application, for $v \in T_pM$

$$(\varphi^*\omega)_p\, v = \omega_{\varphi\, p}(T_p\varphi\, v)$$

The pushforward of vector field is contravariant, while the pullback of one-forms is contravariant. Both always exist regardless of whether they're induced by a diffeomorphism.

Theorem 5.32 (*local pullback formula*). Given $\varphi : M \to N$ and charts $(U, x), (V, y)$ for M, N. For $\omega \in \mathfrak{X}^*(N)$ with local coframe representation $\omega\, p = \vec{\omega}(p) \cdot dy$, its pullback is

$$\varphi^*\left(\vec{\omega} \cdot dy\right) = (\vec{\omega} \circ \varphi) \cdot (J_{x \to y\varphi} \cdot dx)$$
$$\varphi^*\left(\sum \omega_i\, dy_i\right) = \sum (\omega_i \circ \varphi)\left(\partial_{x_j}(y \circ \varphi)_i\right)\, dx_j$$

Proof: Consider a vector field $X\,p = X_i(p)\partial_{x_i}\big|_p = \vec{x}_p \cdot \partial_x\big|_p$.

$$\begin{aligned}
\omega(\varphi_* X)\,p &= \omega_{\varphi\,p}\ \left[\varphi_*(\vec{x}_p \cdot \partial_x\big|_p)\right] \\
&= \left[\vec{\omega}_{\varphi\,p} \cdot dy\big|_{\varphi(p)}\right]\ \left[\varphi_*(\vec{x}_p \cdot \partial_x\big|_p)\right] \\
&= \left[\vec{\omega}_{\varphi\,p} \cdot dy\big|_{\varphi(p)}\right]\ \left[\left(J_{x\to y\varphi}\big|_p\,\vec{x}_p\right) \cdot \partial_y\big|_{\varphi(p)}\right] \\
&= \vec{\omega}(\varphi\,p) \cdot (J_{x\to y\varphi}\big|_p \cdot \vec{x}(p)) \\
&= \left[(\vec{\omega} \circ \varphi) \cdot (J_{x\to y\varphi}\big|_p \cdot \vec{x})\right]\Bigg|_p
\end{aligned}$$

Definition 5.39 (*tautological 1-form*). The tangent bundle projection map $T^*M \xrightarrow{\pi} M$ has a tangent lift $TT^*M \xrightarrow{T\pi} TM$. Given a 1-form $\omega \in T_p^*M$ and a velocity vector $v_\omega \in T_\omega T^*M$, the tautological 1-form $\theta \in T^*T^*M$ is defined by

$$\theta\,v_\omega = \omega\ [(T\pi)\,v_\omega]$$

Consider typing with the diagram below. The action of θ is defined by the two anti-diagonals.

$$\begin{tikzcd}
TM_{[(T\pi)\,v_\omega]} & TT^*M_{(v_\omega)} \arrow[l, "T\pi"'] & \\
M \arrow[u, dashed, "T"] & T^*M_{(\omega)} \arrow[l, "\pi"] \arrow[u, dashed, "T"] \arrow[r, dashed, "T^*"] & T^*T^*M_{(\theta)}
\end{tikzcd}$$

Let (q, p) denote the natural chart for the point and principal components of T^*M, respectively. We consider the local representation of

$$\theta = (\theta\,\partial_{q_i})\,dq_i + (\theta\,\partial_{p_i})\,dp_i = p\,dq$$

Note that $\partial_{q_i}, \partial_{p_i}$ are both elements of TT^*M based at $\sum p_i\,dq_i$.

$$\begin{aligned}
\theta\,\partial_{q_i} &= (p\,dq)\ [(T\pi)\,\partial_{q_i}] = (p\,dq)\,\partial_{q_i} = p_i \\
\theta\,\partial_{q_i} &= (p\,dq)\ [(T\pi)\,\partial_{p_i}] = (p\,dq)\,0 = 0
\end{aligned}$$

5.7 A categorical perspective

This section explores the natural construction of induced pushforwards and pullbacks using the composition of morphisms. We begin with a simple map between two manifolds. The manifolds may coincide and the map needs not be a diffeomorphism.

$$M \xrightarrow{\varphi} N$$

Scalar fields are constructed with the contravariant functor $C^\infty(-, \mathbb{R}) : \mathbf{Man} \to \mathbf{Set}$. It acts on scalar fields as the pullback.

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[dr, dashed, "C^\infty(\varphi,\mathbb{R})\, f"'] & N \arrow[d, "f"] \\ & \mathbb{R} \end{tikzcd} \qquad C^\infty(M,\mathbb{R}) \xleftarrow{C^\infty(\varphi,\mathbb{R}) = \varphi^*} C^\infty(N,\mathbb{R})$$

Consider the contravariant functor $\mathrm{Der}(-,\mathbb{R}) : \mathbf{Set} \to \mathbf{Set}$ of pointed real derivations of scalar fields. The elements of its objects are tangent vectors. It acts on morphisms $\varphi^* : C^\infty(N,\mathbb{R}) \to C^\infty(M,\mathbb{R})$ as the pointed tangent map $T_p\varphi$.

$$\begin{tikzcd} C^\infty(N,\mathbb{R}) \arrow[r, "\varphi^*"] \arrow[dr, dashed, "T_p\varphi\, v_p"'] & C^\infty(M,\mathbb{R}) \arrow[d, "v_p"] \\ & \mathbb{R} \end{tikzcd} \qquad \mathrm{Der}(C^\infty(M,\mathbb{R}),\mathbb{R}) \xrightarrow{T_p\varphi} \mathrm{Der}(C^\infty(N,\mathbb{R}))$$

Composing the two contravariant functors yields the covariant pointed tangent functor. When φ is an injection, there will be no ambiguity in collecting is results along all points of the domain manifold, yielding the global tangent functor.

$$T_p(-) = \mathrm{Der}(C^\infty(-,\mathbb{R}),\mathbb{R}) : \mathbf{Man} \to \mathbf{Set}$$

The pointed cotangent functor is obtained by composing yet again with the contravariant dual map, this time yielding a contravariant functor.

$$T_p^*(-) = \mathrm{Hom}(\mathrm{Der}(C^\infty(-,\mathbb{R}),\mathbb{R}),\mathbb{R}) : \mathbf{Man} \to \mathbf{Set}$$

5.8 Line integrals and conservative fields

Definition 5.40 (*line integral*). The line integral of $\alpha \in \Omega^1(M)$ along smooth $[a,b] \xrightarrow{\gamma} M$ is

$$\int_\gamma \alpha = \int_{[a,b]} \gamma^* \alpha$$

This definition may be extended to accomodate (finitely) piecewise smooth curves.

Definition 5.41 (*conservative* 1*-form*). a 1-form $\alpha \in \Omega^1(M)$ is conservative if $\int_c \alpha = 0$ for all closed piecewise smooth curves c.

Lemma 5.33. Consider a set function $M \xrightarrow{f} \mathbb{R}$ and $\alpha \in \Omega^1(M)$. If for every $v_p \in T_pM$ and smooth curve c representing v_p i.e $\dot{c}(0) = v_p$, the derivative $d_t\big|_0 (fc)t$ exists and

$$d_t\big|_0 f(c(t)) = \alpha\, v_p$$

Then f is smooth and $df = \alpha$

Proof: Work in a chart (U,x), let $v_p = \partial_{x_i}$ and $c(t) = x^{-1}(x\,p + te_i)$. The hypothesis implies that all first-order partial derivatives of fx^{-1} exist and are continuous. Given this, $df_p\, v_p = \alpha_p\, v_p$ so $df = \alpha$. In particular, $\alpha\, \partial_{x_i} = \partial_{x_i} f$ are smooth, so f is smooth.

Proposition 5.34. A smooth 1-form $\alpha \in \Omega^1(M)$ is conservative if and only if it is exact.

Proof: If $\alpha = df$, its pullback along a closed curve corresponds to integration in a vanishing interval, so it is conservative. Conversely, fix $p_0 \in M$. Given conservative α we can unambiguously define

$$f(p) = \int_\gamma \alpha, \quad \gamma(0) = p_0, \gamma(1) = p$$

We use the previous lemma to show that $\alpha = df$: choose a curve c with $c(-1) = p_0, c(0) = p$ and $c'(0) = v_p$. This can be used for γ above. Then $c^*\alpha = g(t)\,dt$ for some g on the reals.

$$\begin{aligned} d_t\big|_0 f(c(t)) &= d_t\Big|_0 \int_{c|[-1,t]} \alpha = d_t\Big|_0 \int_{c|[0,t]} \alpha \\ &= d_t\Big|_0 \int_0^t c^*\alpha = d_t\Big|_0 \int_0^t g(t)\,dt = g(0) \end{aligned}$$

On the other hand, recall the definition of a pushback in terms of pushforwards

$$\alpha\, v_p = \alpha(c'(0)) = \alpha\left(T_0 c\,\partial_t\big|_0\right) \quad = c^*\alpha(\partial_t\big|_0) = g(0)\,dt\big|_0\left(\partial_t\big|_0\right) = g(0)$$

Definition 5.42 (*frame field*). Given an open subset $U \subset M$ with dimension n, a set $\{E_1, \cdots, E_n\}$ of smooth vector fields defined on U is a frame field over U if $\{E_1\,p, \cdots, E_n\,p\}$ forms a basis for T_pM for every $p \in M$. They constitute a global frame field if $U = M$ and are nonholomonic if there does not exist a chart in which $\{E_i\,p\} = \{\partial_{x_i}\big|_p\}$. A frame field with nonvanishing Lie bracket is seen to be nonholomoni.

Remark 5.8. *Second fundamental confusion of calculus (Penrose)*: Given a chart (U, x), the meaning of $\partial_{x_i} f$, fixing i, depends implicitly on the rest of the coordinate functions. Consider the following change of coordinates

$$\begin{aligned} x_1' &= x_1 + x_2 \\ x_2' &= x_1 - x_2 + x_3 \\ x_3' &= x_3 \end{aligned}$$

Here, $x_3' = x_3$ but $\partial_{x_3} = \partial_{x_2'} + \partial_{x_3'}$. This confusion arises frequently in thermodynamics: the functions P, V, T are not independent and may be viewed as functions on some 2-manifold. Any two of the three serves as a chart, thus $\partial_P f$ depends on whether we're using the chart (P, V) or (P, T).