

Lezioni di Geometria Differenziale

Marco Manetti, Domenico Monaco

4 gennaio 2025

Sommario

Note di circa 30 ore di lezioni del corso di Geometria Differenziale, a.a. 2008-09, Corso di Laurea in Matematica, Sapienza Università di Roma; le rimanenti 18 ore erano dedicate a curve e superfici in \mathbb{R}^3 .

Indice

1	Lezioni schizofreniche	2
1.1	Funzioni C^∞	2
1.2	Partizioni dell'unità C^∞ (su \mathbb{R}^n)	4
1.3	Derivazioni	8
1.4	Germi di funzioni C^∞ , spazio tangente, differenziale	11
1.5	Teorema di invertibilità locale	14
1.6	Esercizi svolti	15
2	Varietà differenziabili	18
2.1	Varietà topologiche	18
2.2	Varietà differenziabili	20
2.3	Sistemi di coordinate locali, sottovarietà	23
2.4	Germi di funzioni C^∞ , spazio tangente, differenziale (di nuovo!)	24
3	Il teorema di Sard	28
3.1	Prerequisiti di Analisi	28
3.2	Dimostrazione del teorema di Sard	29
3.3	Applicazioni del teorema di Sard	31
4	Trasversalità	33
5	Fibrato tangente	36
6	Campi di vettori	38
6.1	Campi di vettori su varietà	38
6.2	Algebre di Lie	39
6.3	Correlazione	41
6.4	Gruppi di Lie	42
7	Gruppi a un parametro	44
7.1	Flussi, curve integrali, generatori infinitesimali	44
7.2	Problema di Cauchy e sistemi dinamici	46
8	Algebra multilineare	49
8.1	Prodotti tensoriali	49
8.2	Prodotti esterni	52
8.3	Prodotto di contrazione	54

9	Forme differenziali	56
9.1	Fibrato cotangente	56
9.2	1-forme differenziali	57
9.3	k -forme differenziali	59

1 Lezioni schizofreniche

1.1 Funzioni C^∞

Ci rifacciamo alla definizione di *classe* C^∞ ben nota dai corsi di Analisi: dato un aperto $U \subseteq \mathbb{R}^n$, la funzione $f: U \rightarrow \mathbb{R}$ si dirà quindi di classe C^∞ se è continua, derivabile e le sue derivate parziali $\frac{\partial f}{\partial x_i}: U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, sono ancora C^∞ . Le funzioni prese in considerazione in Geometria si assumono tutte di classe C^∞ ; lo studio viene invece rivolto al loro dominio di definizione.

Se $U \subseteq \mathbb{R}^n$ è aperto, useremo spesso anche la notazione

$$C^\infty(U) := \{f: U \rightarrow \mathbb{R} : f \text{ è di classe } C^\infty\}.$$

L'insieme $C^\infty(U)$ è dotato naturalmente di una struttura di anello commutativo unitario nonché di quella di spazio vettoriale. Il problema che ci proponiamo di affrontare è come definire invece $C^\infty(A)$ per un generico sottoinsieme (non necessariamente aperto) $A \subseteq \mathbb{R}^n$.

Esempio 1.1. Consideriamo $f: S^1 \rightarrow \mathbb{R}$. Alcune idee per poter stabilire quando f sia di classe $C^\infty(S^1)$ possono essere le seguenti:

- esistono un aperto $U \supset S^1$ di \mathbb{R}^2 e una funzione $g \in C^\infty(U)$ tali che $f \equiv g|_{S^1}$;
- parametrizzando S^1 con una coordinata angolare ϑ possiamo pensare a f come a una funzione $g(\vartheta)$ di classe C^∞ e periodica di periodo 2π .

Non è difficile verificare che in realtà queste due definizioni risultano equivalenti (basta passare a coordinate polari).

Per definire in generale $C^\infty(A)$ ci conviene prima studiare le proprietà di $C^\infty(U)$ con U aperto di \mathbb{R}^n . Questo spazio gode infatti delle *proprietà di fascio*:

1. se U si può scrivere come $U = \bigcup_{i \in I} U_i$, dove gli $U_i \subseteq \mathbb{R}^n$ sono aperti, e f è una funzione di classe $C^\infty(U)$, allora

$$f \equiv 0 \iff f|_{U_i} \equiv 0 \forall i \in I \quad (\text{unicità}); \tag{1}$$

2. se U si può scrivere come $U = \bigcup_{i \in I} U_i$, dove gli $U_i \subseteq \mathbb{R}^n$ sono aperti, e f_i è una famiglia di funzioni di classe $C^\infty(U_i)$ per ogni $i \in I$, tali che

$$f_i|_{U_i \cap U_j} \equiv f_j|_{U_i \cap U_j} \quad \forall i, j \in I,$$

allora

$$\exists f \in C^\infty(U) : f|_{U_i} \equiv f_i \forall i \in I \quad (\text{incollamento}). \tag{2}$$

Osservazione 1.1. Dalla (1) e dalla (2) discende che la f come in (2) è unica: infatti, se \tilde{f} è una funzione che soddisfa entrambe le proprietà di fascio, allora

$$\tilde{f}|_{U_i} \equiv f_i \forall i \in I \implies (f - \tilde{f})|_{U_i} \equiv 0 \forall i \in I \stackrel{(1)}{\implies} f - \tilde{f} \equiv 0.$$

Le proprietà (1) e (2) sono quelle che vogliamo soddisfatte dagli spazi $C^\infty(A)$, che ci apprestiamo quindi a definire formalmente.

Definizione 1.1. Sia A un sottoinsieme qualsiasi di \mathbb{R}^n ; una funzione $f: A \rightarrow \mathbb{R}$ si dice *di classe C^∞* (e scriveremo, equivalentemente, $f \in C^\infty(A)$) se è *localmente* la restrizione di una funzione C^∞ definita su un aperto (e quindi per la quale vale la definizione finora utilizzata):

$$\forall x \in A \quad \exists U \subseteq \mathbb{R}^n \text{ aperto, } x \in U, \quad \exists g \in C^\infty(U) : g \Big|_{U \cap A} \equiv f \Big|_{U \cap A}.$$

Verifichiamo che effettivamente una funzione C^∞ corrispondente a questa definizione soddisfa le proprietà di fascio:

1. la (1) è verificata tautologicamente dalle funzioni a livello insiemistico;
2. per la (2), scriviamo $A = \bigcup_{i \in I} (U_i \cap A)$, dove gli U_i sono aperti di \mathbb{R}^n (e quindi gli $U_i \cap A$ sono aperti in A nella topologia di sottospazio). Prendiamo delle funzioni $f_i \in C^\infty(A \cap U_i)$ tali che

$$f_i \Big|_{A \cap U_i \cap U_j} \equiv f_j \Big|_{A \cap U_i \cap U_j} \quad \forall i, j \in I;$$

sappiamo che esiste una funzione $f: A \rightarrow \mathbb{R}$ tale che $f \Big|_{A \cap U_i} \equiv f_i$ (a livello di applicazione insiemistica). Ci resta da verificare che f soddisfa la definizione di C^∞ .

Prendiamo quindi $x \in A$: allora esiste un indice $i \in I$ tale che $x \in A \cap U_i$ (perché gli U_i ricoprono tutto A). Dato che $f_i \in C^\infty(A \cap U_i)$ esiste un aperto V contenente x e una funzione $g \in C^\infty(V)$ tale che $f_i \equiv g$ su $A \cap U_i \cap V$. Ma allora è evidente che $f \equiv g$ su $A \cap (U_i \cap V)$.

Definizione 1.2. Data un'applicazione $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, si definisce *supporto* di f la chiusura in U dell'insieme dei punti dove f non si annulla:¹

$$\text{supp } f := \{x \in U : \forall V \in \mathcal{J}(x) \exists y \in U \cap V : f(y) \neq 0\} = \overline{\{x \in U : f(x) \neq 0\}}^U.$$

Particolarmente interessanti sono le funzioni a *supporto compatto*: se $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è a supporto compatto e il dominio di definizione U è un aperto, questo vuol dire che f si annulla sul "bordo" di U (ovvero esiste una distanza minima positiva fra il supporto di f e la frontiera di U). Inoltre, nel caso in cui $U = \mathbb{R}^n$, per verificare che il supporto di f è compatto basta verificare che sia limitato (essendo già un chiuso).

Un tipo particolare di funzioni C^∞ a supporto compatto sono le cosiddette *funzioni a bernoccolo* (dall'inglese *bump function*) che risultano utili in alcune dimostrazioni. Vediamo di cosa si tratta.

Esempio 1.2 (Funzioni a bernoccolo). Consideriamo la funzione $\lambda: \mathbb{R} \rightarrow \mathbb{R}$ così definita:

$$\lambda(t) = \begin{cases} 0 & t \leq 0 \\ e^{-\frac{1}{t^2}} & t > 0 \end{cases}.$$

È un facile esercizio di calcolo differenziale verificare che $\lambda \in C^\infty(\mathbb{R})$ e che $\frac{d^n \lambda}{dt^n}(0) = 0$ per ogni $n \in \mathbb{N}$. Il supporto della funzione è $\text{supp } \lambda = [0, +\infty)$. L'andamento della funzione è illustrato nella Figura 1(a).

Fissato ora $\varepsilon > 0$, definiamo

$$\varphi_\varepsilon(t) := \frac{\lambda(t)}{\lambda(t) + \lambda(\varepsilon - t)}.$$

¹Denotiamo con $\mathcal{J}(x)$ l'insieme degli intorno di $x \in \mathbb{R}^n$.

Anche questa funzione è $C^\infty(\mathbb{R})$ in quanto il denominatore non si annulla mai; inoltre, vale 0 per $t \leq 0$, vale 1 per $t \geq \varepsilon$ e si raccorda in maniera C^∞ per $0 < t < \varepsilon$. Il suo supporto è ancora la semiretta $[0, +\infty)$, e il suo grafico è illustrato nella Figura 1(b).

Se fissiamo ora un terzo parametro $r > 0$, possiamo trovare una funzione C^∞ che vale 1 in una palla attorno all'origine di raggio r , e si annulla fuori dalla palla di centro l'origine e raggio $r + \varepsilon$: è la funzione

$$\varphi_{r,\varepsilon}(t) = 1 - \varphi_\varepsilon(|t| - r).$$

Anche se il modulo non è una funzione C^∞ vicino all'origine (non è neanche C^1 !), basta osservare che per $|t| < r$ la funzione $\varphi_\varepsilon(|t| - r)$ vale 0, ed è quindi infinitamente differenziabile. La funzione $\varphi_{r,\varepsilon}$ è la funzione a bernoccolo che cercavamo, ed è la prima funzione a supporto compatto che incontriamo, essendo $\text{supp } \varphi_{r,\varepsilon} = [-r - \varepsilon, r + \varepsilon]$. Il suo nome deriva dal grafico che traccia, illustrato nella Figura 1(c).

È ovvio che procedendo alla stessa maniera è possibile costruire funzioni a bernoccolo in \mathbb{R}^n , centrate in un punto qualsiasi (e non necessariamente nell'origine); in altre parole, comunque si scelgano $p \in \mathbb{R}^n$ e due costanti positive r e ε , esiste una funzione $\varphi: \mathbb{R}^n \rightarrow [0, 1]$ di classe C^∞ tale che $\varphi(x) = 1$ per ogni $x \in B(p, r)$ e $\varphi(x) = 0$ per ogni $x \in \mathbb{R}^n \setminus B(p, r + \varepsilon)$.

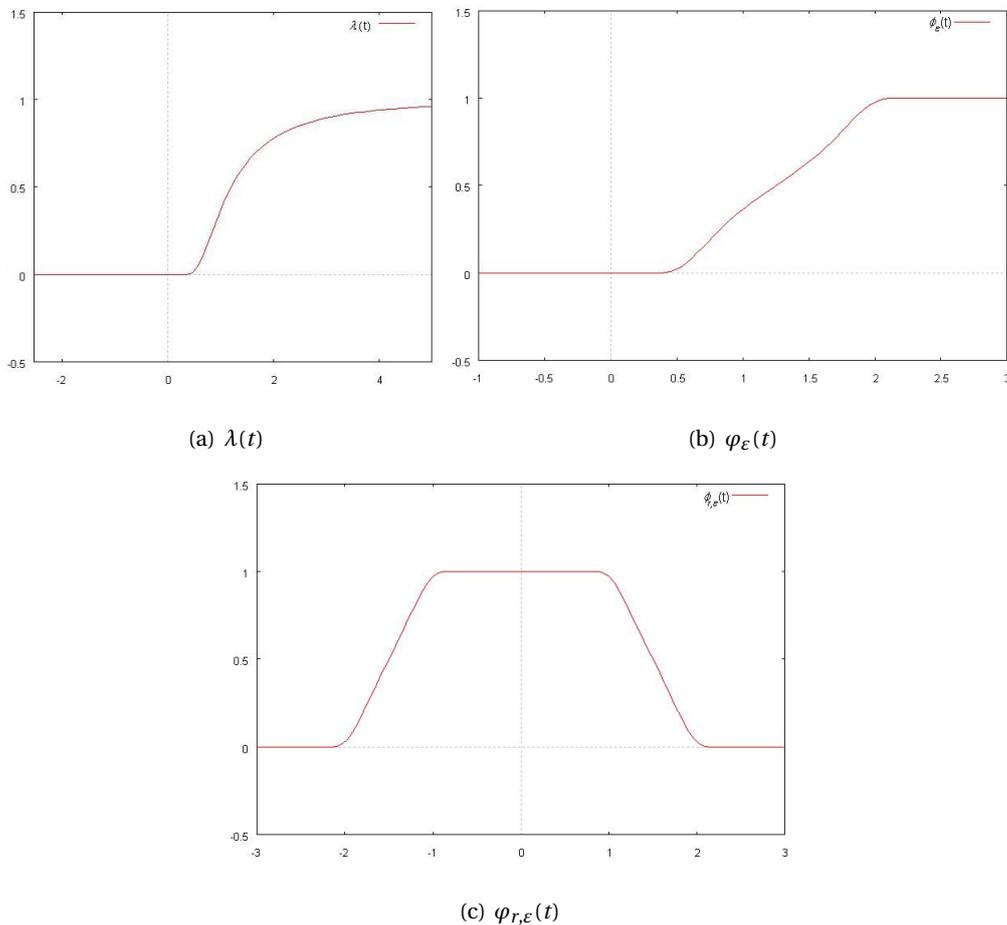


Figura 1: Una funzione a bernoccolo

1.2 Partizioni dell'unità C^∞ (su \mathbb{R}^n)

Ricordiamo che una famiglia di sottospazi $\{A_i\}_{i \in I}$ di uno spazio topologico X si dice *localmente finita* se per ogni $x \in X$ esiste un intorno U di x tale che $A_i \cap U \neq \emptyset$ al più per un numero finito di indici.

Definizione 1.3. Sia $\mathbb{R}^n = \bigcup_{i \in I} U_i$ un ricoprimento aperto di \mathbb{R}^n ; una *partizione dell'unità* C^∞ su \mathbb{R}^n subordinata al ricoprimento $\{U_i\}_{i \in I}$ è una famiglia di funzioni $\varphi_i: \mathbb{R}^n \rightarrow [0, 1]$, $i \in I$, di classe $C^\infty(\mathbb{R}^n)$ tali che:

1. $\text{supp } \varphi_i \subset U_i$ per ogni $i \in I$;
2. la famiglia $\{\text{supp } \varphi_i\}_{i \in I}$ è localmente finita;
3. $\sum_{i \in I} \varphi_i \equiv 1$ su \mathbb{R}^n .

Facciamo qualche osservazione su questa definizione. Innanzitutto, il nome “partizione” dell’unità deriva dal fatto che, per il punto 3., noi siamo capaci di scrivere 1 come somma di un certo numero di funzioni. Dato che la funzione φ_i si annulla “fuori” da U_i per il punto 1., la somma in 3. vale di fatto $\sum_{i \in I} \varphi_i \Big|_{U_i}$ (dato che il valore di φ_i fuori da U_i è nullo, questo non contribuisce alla somma). Da questo deduciamo che la somma è finita, in quanto, per il punto 2. della definizione, ogni aperto U_i interseca al più un numero finito di supporti (ovvero, al più un numero finito di funzioni non si annulla su U_i). Si può visualizzare una partizione dell’unità come una serie di bernoccoli (le φ_i) che ricoprono \mathbb{R}^n ma senza “accumularsi” troppo (per la locale finitezza) e tali che le loro altezze, se sommate, diano 1.

Quello che ci proponiamo di dimostrare ora è che ogni ricoprimento aperto di \mathbb{R}^n ammette partizioni dell’unità C^∞ : questo risultato, a sua volta, porterà ad altre interessanti applicazioni. Premettiamo tuttavia un lemma preliminare.

Lemma 1.1. Sia $\mathbb{R}^n = \bigcup_{i \in I} U_i$ un ricoprimento aperto di \mathbb{R}^n ; allora, esiste un ricoprimento localmente finito del tipo $\mathbb{R}^n = \bigcup_{h \in H} B(x_h, r_h)$ tale che per ogni indice $h \in H$ esiste un indice $i \in I$ con $\overline{B(x_h, r_h)} \subset U_i$.

Osservazione 1.2. Per l’assioma della scelta, esiste una *funzione di raffinamento* $\delta: H \rightarrow I$ tale che per ogni $h \in H$ valga $\overline{B(x_h, r_h)} \subset U_{\delta(h)}$.

Segue inoltre dal Lemma 1.1 che anche la famiglia $\{\overline{B(x_h, r_h)}\}$ è localmente finita: infatti, un aperto interseca la chiusura di una palla se e solo se interseca la palla stessa (come discende dalla definizione di chiusura).

Dimostrazione. Dal fatto che gli U_i formano un ricoprimento aperto deduciamo che per ogni $x \in \mathbb{R}^n$ esistono un raggio $r_x \in \left(0, \frac{1}{2}\right]$ ed un indice $i \in I$ tali che $\overline{B(x, r_x)} \subset U_i$. Infatti, essendo aperto, U_i contiene un intorno aperto, che possiamo scegliere sferico, di ogni suo punto: a meno di restringere il raggio di questo intorno, possiamo supporre che anche la chiusura della palla sia contenuta in U_i . Un ricoprimento di questo tipo, tuttavia, è “troppo grosso”: vediamo come possiamo fare per prendere un suo sottoricoprimento localmente finito.

Poniamo, per ogni $m \in \mathbb{N}$, $K_m := \{x \in \mathbb{R}^n : m \leq \|x\| \leq m+1\}$. Essendo compatto, l’insieme K_m è coperto da un numero finito di palle:

$$\forall m \in \mathbb{N} \exists J \text{ insieme finito di indici, } \exists x_j \in K_m, j \in J : K_m \subset \bigcup_{j \in J} B(x_j, r_{x_j}).$$

Dato che tutte le palle hanno raggio al più $\frac{1}{2}$, ogni palla centrata in un punto di K_m può intersecare al più i tre compatti K_{m-1} , K_m e K_{m+1} . Dato che per costruzione $\mathbb{R}^n = \bigcup_{m \in \mathbb{N}} K_m$, abbiamo che

$$\mathbb{R}^n = \bigcup_{m \in \mathbb{N}} \bigcup_{j \in J} B(x_j, r_{x_j})$$

è il ricoprimento localmente finito che cercavamo. □

Teorema 1.1. *Ogni ricoprimento aperto di \mathbb{R}^n ammette partizioni dell'unità C^∞ .*

Dimostrazione. Dato un ricoprimento aperto $\mathbb{R}^n = \bigcup_{i \in I} U_i$, per il Lemma 1.1 ne abbiamo anche uno localmente finito $\mathbb{R}^n = \bigcup_{h \in H} B(x_h, r_h)$ e una funzione di raffinamento $\delta: H \rightarrow I$ con $\overline{B(x_h, r_h)} \subset U_{\delta(h)}$ per ogni $h \in H$. Sia ora $\lambda_h: \mathbb{R}^n \rightarrow [0, 1]$ una funzione $C^\infty(\mathbb{R}^n)$ tale che $\lambda(x) > 0$ se e solo se $x \in B(x_h, r_h)$ (ad esempio, una funzione a bernoccolo): conseguentemente, avremo $\text{supp } \lambda_h = \overline{B(x_h, r_h)} \subset U_{\delta(h)}$.

Definiamo $g_i = \sum_{h: \delta(h)=i} \lambda_h$; tale definizione ha senso perché, quando vogliamo definire $g_i(x)$, ogni intorno del punto x può intersercare al più un numero finito di palle della forma $B(x_h, r_h)$, quindi le funzioni λ_h che sommiamo sono (localmente) in numero finito. Concludiamo allora che $g_i: \mathbb{R}^n \rightarrow [0, +\infty)$ è una funzione $C^\infty(\mathbb{R}^n)$, essendo la somma localmente finita e le λ_h a loro volta di classe C^∞ . Vale inoltre che $\text{supp } g_i \subset U_i$, essendo unione dei supporti delle (finite) λ_h sommate per formarla. Infatti si ha

$$\begin{aligned} x \in \text{supp } g_i &\iff \forall n \in \mathbb{N} \exists x_n \in B\left(x, \frac{1}{n}\right) : g_i(x_n) \neq 0 && \text{(per definizione di supporto)} \\ &\iff \forall n \in \mathbb{N} \exists x_n \in B\left(x, \frac{1}{n}\right) : g_i(x_n) > 0 && \text{(perché } g_i \text{ è non negativa)} \\ &\iff \forall n \in \mathbb{N} \exists x_n \in B\left(x, \frac{1}{n}\right) : \sum_{h: \delta(h)=i} \lambda_h(x_n) > 0 \\ &\iff \forall n \in \mathbb{N} \exists x_n \in B\left(x, \frac{1}{n}\right), \exists h : \delta(h) = i \text{ e } \lambda_h(x_n) > 0 \\ &\iff x \in \bigcup_{h: \delta(h)=i} \text{supp } \lambda_h \subset \bigcup_{h: \delta(h)=i} U_{\delta(h)} = U_i. \end{aligned}$$

Per avere ora delle funzioni che costituiscano una partizione dell'unità, è sufficiente “normalizzare” le g_i , ovvero definire $\varphi_i := \frac{g_i}{\sum_{i \in I} g_i} = \frac{g_i}{\sum_{h \in H} \lambda_h}$ (tale definizione ha senso perché la somma a denominatore è sempre positiva). Le funzioni φ_i soddisfano ora le tre proprietà di una partizione dell'unità: $\text{supp } \varphi_i = \text{supp } g_i \subset U_i$, la famiglia $\{\text{supp } \varphi_i\}$ è localmente finita perché lo è quella delle palle chiuse $\{\overline{B(x_h, r_h)}\}$, e infine

$$\sum_{i \in I} \varphi_i = \sum_{i \in I} \frac{g_i}{\sum_{i \in I} g_i} = \frac{\sum_{i \in I} g_i}{\sum_{i \in I} g_i} = 1.$$

□

Vediamo delle applicazioni di questo teorema. La prima ci dice che, per i chiusi, la definizione locale di $C^\infty(A)$ che abbiamo dato coincide con quella globale su \mathbb{R}^n .

Corollario 1.1. *Sia $A \subseteq \mathbb{R}^n$ un chiuso e $f: A \rightarrow \mathbb{R}$ una funzione di classe C^∞ ; allora, esiste una funzione $g \in C^\infty(\mathbb{R}^n)$ tale che $g|_A \equiv f$.*

Dimostrazione. Definiamo l'aperto $U_0 = \mathbb{R}^n \setminus A$. Per definizione di $C^\infty(A)$ esistono degli aperti U_i , $i \in I$ (supponiamo che I non contenga lo 0 che abbiamo già usato come indice) e delle funzioni $f_i \in C^\infty(U_i)$ tali che

$$A \subset \bigcup_{i \in I} U_i \quad \text{e} \quad f_i|_{A \cap U_i} \equiv f|_{A \cap U_i}.$$

Dato che $U_0 \cup \bigcup_{i \in I} U_i \supset (\mathbb{R}^n \setminus A) \cup A = \mathbb{R}^n$ è un ricoprimento aperto, ammette una partizione dell'unità $\varphi_i, i \in I \cup \{0\}$. Poniamo quindi

$$g_i(x) = \begin{cases} f_i(x)\varphi_i(x) & x \in U_i \\ 0 & x \notin U_i \end{cases} = \begin{cases} f_i(x)\varphi_i(x) & x \in U_i \\ 0 & x \in \mathbb{R}^n \setminus \text{supp } \varphi_i \end{cases};$$

dalla seconda scrittura (equivalente alla prima, in quanto se x è nel complementare del supporto di φ_i allora $\varphi_i(x) = 0$, ovvero le due definizioni sono coincidenti sull'intersezione) si deduce che le g_i sono $C^\infty(\mathbb{R}^n)$, essendo incollamento di due funzioni C^∞ su aperti.

Essendo la famiglia $\{\text{supp } g_i\}$ localmente finita, ha senso definire $g = \sum_{i \in I \cup \{0\}} g_i$; g risulta ancora $C^\infty(\mathbb{R}^n)$. Per vedere che g è la funzione richiesta, dobbiamo verificare che coincida con f su A . Sia quindi $x \in A$: si ha allora

$$\begin{aligned} g(x) &= \sum_{i \in I \cup \{0\}} f_i(x)\varphi_i(x) = \\ &= \sum_{i \in I \cup \{0\}} f(x)\varphi_i(x) = \quad (\text{su } U_i, \text{ dove } \varphi_i(x) \neq 0, \text{ si ha } f_i \equiv f) \\ &= f(x) \sum_{i \in I \cup \{0\}} \varphi_i(x) = f(x) \end{aligned}$$

dove nell'ultima uguaglianza abbiamo usato il fatto che, essendo φ_i una partizione dell'unità, si ha per definizione $\sum_{i \in I \cup \{0\}} \varphi_i \equiv 1$. \square

Lemma 1.2 (di Urysohn C^∞). *Siano $A, B \subseteq \mathbb{R}^n$ due chiusi disgiunti; allora, esiste una funzione $\varphi: \mathbb{R}^n \rightarrow [0, 1]$ di classe C^∞ , detta funzione di Urysohn, tale che*

$$\varphi \Big|_A \equiv 1, \quad \varphi \Big|_B \equiv 0.$$

Dimostrazione. Dal fatto che A e B sono chiusi disgiunti, segue che $\mathbb{R}^n = (\mathbb{R}^n \setminus A) \cup (\mathbb{R}^n \setminus B)$ è un ricoprimento aperto: consideriamo la partizione dell'unità associata $\varphi_A, \varphi_B: \mathbb{R}^n \rightarrow [0, 1]$ di classe C^∞ con $\text{supp } \varphi_A \subset \mathbb{R}^n \setminus A$, $\text{supp } \varphi_B \subset \mathbb{R}^n \setminus B$ e $\varphi_A + \varphi_B \equiv 1$. Poiché $\text{supp } \varphi_B \subset \mathbb{R}^n \setminus B$, abbiamo che $\varphi_B \Big|_B \equiv 0$; inoltre, dato che per lo stesso motivo $\varphi_A \Big|_A \equiv 0$, dal fatto che $\varphi_A + \varphi_B \equiv 1$ deduciamo che $\varphi_B \Big|_A \equiv 1$. È ora evidente che φ_B è la funzione che cercavamo. \square

Osservazione 1.3. Ovviamente, con trasformazioni affini possiamo trovare funzioni C^∞ che soddisfano $\varphi \Big|_A \equiv a, \varphi \Big|_B \equiv b$ per ogni $a, b \in \mathbb{R}$.

Il prossimo risultato ci dice che possiamo approssimare funzioni continue su compatti con funzioni C^∞ commettendo un errore piccolo a piacere.

Teorema 1.2 (di Stone-Weierstrass, caso particolare). *Sia $K \subset \mathbb{R}^n$ un compatto e $f: K \rightarrow \mathbb{R}$ continua; allora, per ogni $\varepsilon > 0$ esiste una funzione $g \in C^\infty(\mathbb{R}^n)$ tale che*

$$\max_{x \in K} |f(x) - g(x)| < \varepsilon.$$

Dimostrazione. Dato che l'immagine $f(K)$ è un compatto di \mathbb{R} per la continuità di f , possiamo scrivere $f: K \rightarrow [-M, M]$ per qualche $M \in [0, +\infty)$. Definiamo

$$A = \left\{ x \in K : f(x) \geq \frac{M}{3} \right\}, \quad B = \left\{ x \in K : f(x) \leq -\frac{M}{3} \right\};$$

A e B risultano allora essere chiusi disgiunti di K , e quindi di \mathbb{R}^n . Siamo dunque nelle ipotesi del lemma di Urysohn: esiste quindi una funzione $g_1: \mathbb{R}^n \rightarrow \left[-\frac{M}{3}, \frac{M}{3}\right]$ di classe C^∞ e tale che $g_1|_A = \frac{M}{3}$ e $g_1|_B = -\frac{M}{3}$. Di conseguenza $f - g_1$ manda K nell'intervallo $\left[-\frac{2}{3}M, \frac{2}{3}M\right]$: infatti, se $|f(x)| < \frac{M}{3}$ si aggiunge con g_1 al più $\frac{M}{3}$, mentre se $f(x) \geq \frac{M}{3}$ si toglie $\frac{M}{3}$ con g_1 o viceversa si aggiunge tale quantità se $f(x) \leq -\frac{M}{3}$.

Iterando questo ragionamento, possiamo trovare delle funzioni $g_i, i = 1, \dots, n$, di classe C^∞ e tali che $f - g_1 - \dots - g_n: K \rightarrow \left[-\left(\frac{2}{3}\right)^n M, \left(\frac{2}{3}\right)^n M\right]$ abbia immagine di ampiezza piccola a piacere. La funzione che cerchiamo è quindi $g = \sum_{i=1}^n g_i$. \square

1.3 Derivazioni

Sappiamo dall'Analisi cosa si intende per derivate parziali di una funzione $f: \mathbb{R}^n \rightarrow \mathbb{R}$, e le indichiamo con $\frac{\partial f}{\partial x_i}, i = 1, \dots, n$. Questa definizione presuppone tuttavia che si abbiano delle coordinate (canoniche) x_i rispetto alle quali derivare; definirle non è difficile in \mathbb{R}^n , ma negli spazi che tratteremo potrebbe non essere altrettanto ovvio. In questa sezione e nelle seguenti ci ripropriamo di dare una definizione più algebrica di derivata da poter utilizzare su spazi più generali del "semplice" \mathbb{R}^n , e di vedere come riadattare alcuni dei concetti ad essa collegati che già conosciamo dall'Analisi.

Definizione 1.4. Sia $(A, +, \cdot)$ un anello commutativo unitario che sia anche un \mathbb{R} -spazio vettoriale (in particolare $\mathbb{R} \subseteq A$). Un'applicazione $\delta: A \rightarrow A$ si dice *derivazione* se:

1. δ è \mathbb{R} -lineare;
2. vale la *regola di Leibniz*

$$\delta(a \cdot b) = a \cdot \delta(b) + b \cdot \delta(a) \quad \forall a, b \in A. \quad (3)$$

Indicheremo con $\text{Der}(A, A)$ l'insieme delle derivazioni $\delta: A \rightarrow A$.

Osservazione 1.4. Se $\lambda \in \mathbb{R} \subseteq A$ è una costante reale e $\delta: A \rightarrow A$ una derivazione, allora dalla \mathbb{R} -linearità deduciamo che $\delta(\lambda) = \lambda \cdot \delta(1)$. Del resto, sfruttando la regola di Leibniz, si ha che

$$\delta(1) = \delta(1 \cdot 1) = 1 \cdot \delta(1) + 1 \cdot \delta(1) = 2\delta(1) \quad \implies \quad \delta(1) = 0$$

e quindi $\delta(\lambda) = 0$ per ogni $\lambda \in \mathbb{R}$. Ritroviamo il noto fatto che "la derivata di una costante è nulla".

Esempio 1.3. Sia $U \subseteq \mathbb{R}^n$ un aperto: allora le derivate parziali usuali $\frac{\partial}{\partial x_i}: C^\infty(U) \rightarrow C^\infty(U), i = 1, \dots, n$, sono derivazioni. Più in generale, comunque si scelgano le funzioni $h_1, \dots, h_n \in C^\infty(U)$, l'applicazione

$$\delta: C^\infty(U) \rightarrow C^\infty(U), \quad \delta(f) := \sum_{i=1}^n h_i \cdot \frac{\partial f}{\partial x_i} \quad (4)$$

è una derivazione (questo è vero in generale per combinazioni lineari di derivazioni a coefficienti nell'anello di definizione). La verifica della \mathbb{R} -linearità è molto semplice ed è lasciata come esercizio al lettore; dimostriamo invece che δ soddisfa la regola di Leibniz. Si ha infatti

$$\delta(fg) = \sum_{i=1}^n h_i \frac{\partial(fg)}{\partial x_i} =$$

$$\begin{aligned}
&= \sum_{i=1}^n h_i \left(f \frac{\partial g}{\partial x_i} + g \frac{\partial f}{\partial x_i} \right) = \quad (\text{per la regola di Leibniz applicata a } \frac{\partial}{\partial x_i}) \\
&= \left(\sum_{i=1}^n h_i \frac{\partial g}{\partial x_i} \right) f + \left(\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i} \right) g = \\
&= f \cdot \delta(g) + g \cdot \delta(f).
\end{aligned}$$

Per le derivazioni in $C^\infty(U)$, quanto abbiamo dimostrato in questo esempio si inverte: tutte le derivazioni possono essere scritte in maniera univoca nella forma (4). Per dimostrare questo fatto, tuttavia, ci occorre esplorare più a fondo le proprietà di tali oggetti.

Lemma 1.3. *Le derivazioni $\delta: C^\infty(U) \rightarrow C^\infty(U)$ sono operatori locali, ovvero se $V \subseteq U$ è un aperto e $f \in C^\infty(U)$ è tale che $f|_V \equiv 0$, allora anche $\delta(f)|_V \equiv 0$.*

Dimostrazione. Nelle notazioni dell'enunciato, sia $p \in V$; vogliamo dimostrare che $\delta(f)(p) = 0$ sapendo che f si annulla identicamente in V . Prendiamo una funzione a bernoccolo $\varphi \in C^\infty(U)$ tale che $\text{supp } \varphi \subset V$ e $\varphi \equiv 1$ in un intorno di p (e quindi in particolare $\varphi(p) = 1$); la funzione $f\varphi$ è allora identicamente nulla su tutto \mathbb{R}^n . Infatti, dove f non si annulla (in $U \setminus V$) si annulla φ , e dove φ non si annulla (in V) si annulla f ; infine, $f\varphi$ può essere estesa in maniera C^∞ ponendola uguale a 0 in $\mathbb{R}^n \setminus U$. Conseguentemente, applicando la regola di Leibniz, otteniamo che su \mathbb{R}^n si ha

$$0 = \delta(0) = \delta(f\varphi) = f\delta(\varphi) + \varphi\delta(f)$$

e calcolando quest'uguaglianza in p abbiamo

$$0 = f(p)\delta(\varphi)(p) + \varphi(p)\delta(f)(p) = 0 \cdot \delta(\varphi)(p) + 1 \cdot \delta(f)(p) = \delta(f)(p)$$

come volevamo dimostrare. □

In virtù di quanto appena dimostrato, possiamo affermare che, se $f \in C^\infty(U)$ e δ è una derivazione di $C^\infty(U)$ in sé, il comportamento di $\delta(f)$ dipende, almeno localmente, dal comportamento di f . Infatti, se $g \in C^\infty(U)$ coincide con f in un aperto $V \subseteq U$, allora $(f - g)|_V \equiv 0$, e quindi anche $\delta(f - g)|_V \equiv 0$. Per linearità, questo vuol dire che $\delta(f)|_V \equiv \delta(g)|_V$, quindi anche le derivazioni di f e g coincidono su V .

Lemma 1.4 (formula di Taylor). *Sia $f \in C^\infty(U)$ e $p = (p_1, \dots, p_n) \in U$; allora, esistono un intorno aperto $V \subseteq U$ di p e delle funzioni $f_1, \dots, f_n \in C^\infty(U)$ tali che*

$$f(x) = f(p) + \sum_{i=1}^n (x_i - p_i) f_i(x) \quad \forall x = (x_1, \dots, x_n) \in V \quad e \quad f_i(p) = \frac{\partial f}{\partial x_i}(p) \quad \forall i = 1, \dots, n. \quad (5)$$

Dimostrazione. Non è restrittivo supporre che V sia una palla aperta centrata in p , $V = B(p, r)$. Per la convessità della palla, se x è in V allora tutto il segmento $p + t(x - p)$, $t \in [0, 1]$, è contenuto in V . Possiamo quindi calcolare

$$\begin{aligned}
f(x) &= f(p) + \int_0^1 \frac{df}{dt}(p + t(x - p)) dt \\
&= f(p) + \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p + t(x - p)) \frac{\partial}{\partial t}(p_i + t(x_i - p_i)) dt \\
&= f(p) + \sum_{i=1}^n (x_i - p_i) \int_0^1 \frac{\partial f}{\partial x_i}(p + t(x - p)) dt.
\end{aligned}$$

Ci basta a questo punto porre

$$f_i(x) := \int_0^1 \frac{\partial f}{\partial x_i}(p + t(x-p)) dt$$

per ottenere la tesi. Infatti, le f_i sono $C^\infty(V)$ per i ben noti teoremi di derivazione sotto il segno di integrale; a meno di moltiplicare ciascuna f_i per una funzione a bernoccolo che si annulla in $U \setminus V$, possiamo supporre che le f_i siano anche $C^\infty(U)$. Inoltre, per $x = p$ si ha

$$f_i(p) = \int_0^1 \frac{\partial f}{\partial x_i}(p) dt = \frac{\partial f}{\partial x_i}(p).$$

□

Teorema 1.3. Sia $\delta: C^\infty(U) \rightarrow C^\infty(U)$ una derivazione; allora, esistono e sono uniche delle funzioni $h_1, \dots, h_n \in C^\infty(U)$ tali che

$$\delta = \sum_{i=1}^n h_i \cdot \frac{\partial}{\partial x_i}.$$

Dimostrazione. Per l'unicità, basta osservare che $\delta(x_i) = h_i$ (la funzione $x_i: U \rightarrow \mathbb{R}$ è ovviamente C^∞), in quanto $\frac{\partial x_i}{\partial x_j} = \delta_i^j$.² Per quanto riguarda l'esistenza, sia $p \in U$ e $f \in C^\infty(U)$; per la formula di Taylor, abbiamo che in un intorno di p vale l'uguaglianza (5). Applicando δ a entrambi i membri (possiamo farlo perché le derivazioni sono operatori locali: la differenza fra i due membri è nulla, quindi è nulla la sua derivazione) e sfruttando le sue proprietà di \mathbb{R} -linearità e di Leibniz, otteniamo calcolando in p che

$$\begin{aligned} \delta(f)(p) &= \delta(f(p)) + \sum_{i=1}^n \delta(x_i - p_i) f_i \Big|_{x=p} + \sum_{i=1}^n \delta(f_i)(x_i - p_i) \Big|_{x=p} \\ &= \delta(f(p)) + \sum_{i=1}^n \delta(x_i - p_i)(p) f_i(p) + \sum_{i=1}^n \delta(f_i)(p)(p_i - p_i) \\ &= \sum_{i=1}^n \delta(x_i)(p) \frac{\partial f}{\partial x_i}(p). \end{aligned}$$

Infatti $\delta(f(p)) = 0$ perché derivazione di una costante, e $\delta(x_i - p_i) = \delta(x_i) - \delta(p_i) = \delta(x_i)$ per la \mathbb{R} -linearità e il fatto che, ancora una volta, $\delta(p_i) = 0$ perché le p_i sono costanti. Ci basta ora porre $h_i(x) := \delta(x_i)(x)$ per ottenere quanto richiesto. □

Vediamo ora di fornire anche per il concetto di *campo vettoriale* un'interpretazione algebrica. Innanzitutto, se $U \subseteq \mathbb{R}^n$ è un aperto, diremo che il campo vettoriale $\eta: U \rightarrow \mathbb{R}^n$ è di classe C^∞ se tutte le sue componenti $\eta_i: U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, sono C^∞ . Definiamo inoltre

$$\mathcal{X}(U) := \{\eta: U \rightarrow \mathbb{R}^n : \eta \in C^\infty(U)\}.$$

Osserviamo che se $V \subseteq U$ è aperto e $\eta \in \mathcal{X}(U)$, allora $\eta \Big|_V \in \mathcal{X}(V)$.

Ad un campo vettoriale $\eta \in \mathcal{X}(U)$ associamo una combinazione lineare formale

$$\eta = (\eta_1, \dots, \eta_n) \mapsto \sum_{i=1}^n \eta_i \frac{\partial}{\partial x_i};$$

in questa maniera possiamo pensare ciascun campo vettoriale come una derivazione su $C^\infty(U)$. Possiamo quindi riformulare il Teorema 1.3 affermando che lo spazio dei campi vettoriali di classe C^∞ su un aperto $U \subseteq \mathbb{R}^n$ è in biiezione con lo spazio delle derivazioni di $C^\infty(U)$ in se stesso:

$$\mathcal{X}(U) \simeq \text{Der}(C^\infty(U), C^\infty(U)).$$

²Il simbolo δ_i^j denota la delta di Kronecker

$$\delta_i^j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

1.4 Germi di funzioni C^∞ , spazio tangente, differenziale

Abbiamo visto che il comportamento di un campo vettoriale in un punto dipende dal comportamento dello stesso campo in un intorno arbitrariamente piccolo del punto stesso, come discende dal fatto che le derivazioni sono operatori locali. Questo ci porta naturalmente a pensare di introdurre una relazione d'equivalenza che ci permetta considerare "uguali" due funzioni che coincidono vicino a un punto, disinteressandoci di quanto facciano altrove.

Consideriamo dunque l'insieme

$$E(p) := \{(U, f) : U \subseteq \mathbb{R}^n \text{ aperto}, p \in U, f \in C^\infty(U)\}$$

e introduciamo su questo insieme la seguente relazione \sim :

$$(U, f) \sim (V, g) \quad \text{se} \quad \exists W \subseteq U \cap V \text{ intorno aperto di } p : f|_W \equiv g|_W.$$

Verifichiamo che la relazione \sim è d'equivalenza:

- i. per quanto riguarda la proprietà riflessiva, si ha che $(U, f) \sim (U, f)$ prendendo $W = U$ nella definizione di \sim ;
- ii. se $(U, f) \sim (V, g)$, allora vale anche $(V, g) \sim (U, f)$ perché banalmente se $f|_W \equiv g|_W$ allora è anche $g|_W \equiv f|_W$. Concludiamo che \sim è simmetrica;
- iii. supponiamo che $(U, f) \sim (V, g)$ e $(V, g) \sim (Z, h)$. Sappiamo quindi che esiste $W \subseteq U \cap V$ intorno aperto di p tale che $f|_W \equiv g|_W$ e che esiste $Y \subseteq V \cap Z$ intorno aperto di p tale che $g|_Y \equiv h|_Y$. Dato che W e Y sono aperti, lo è anche la loro intersezione $W \cap Y \subseteq U \cap V \cap Z$; inoltre per ipotesi abbiamo che $f|_{W \cap Y} \equiv g|_{W \cap Y}$ e $g|_{W \cap Y} \equiv h|_{W \cap Y}$. Ne deduciamo che vale anche $f|_{W \cap Y} \equiv h|_{W \cap Y}$, e quindi che $(U, f) \sim (Z, h)$: la relazione \sim è anche transitiva.

Definizione 1.5. Il quoziente $\mathcal{E}(p) := E(p)/\sim$ prende il nome di *anello dei germi di funzioni C^∞ in p* . Tale insieme è effettivamente un anello commutativo unitario se si definiscono le operazioni di *somma*

$$[U, f] + [V, g] := [U \cap V, f + g]$$

e *prodotto*

$$[U, f] \cdot [V, g] := [U \cap V, f \cdot g]$$

(la verifica che tali operazioni siano ben definite è lasciata al lettore).

Osserviamo che $\mathcal{E}(p)$ è anche uno spazio vettoriale su \mathbb{R} , visto come campo delle funzioni costanti. Denoteremo spesso $f = [U, f] \in \mathcal{E}(p)$ per alleggerire la notazione; del resto, il linguaggio dei germi serve a specificare l'essere locale di una data proprietà, più che *dove* tale proprietà valga. La terminologia riguardante le funzioni rimane invece pressoché invariata.

Se $f \in \mathcal{E}(p)$, è ovviamente ben definito $f(p)$ (tutte le funzioni nella classe di equivalenza di f coincidono con f su un qualche intorno di p , e quindi in particolare in p). Consideriamo quindi l'omomorfismo suriettivo di anelli di *valutazione in p*

$$\text{ev}_p : \mathcal{E}(p) \rightarrow \mathbb{R}, \quad f \mapsto f(p).$$

Essendo suriettivo, il suo nucleo $\mathcal{M}(p) := \ker \text{ev}_p = \{f \in \mathcal{E}(p) : f(p) = 0\}$ è un ideale massimale. Infatti, per il teorema di isomorfismo, l'anello quoziente $\mathcal{E}(p)/\mathcal{M}(p)$ è isomorfo all'immagine di ev_p ,

ovvero a tutto \mathbb{R} , risultando quindi un campo; è noto dai corsi di Algebra che un anello quoziente è un campo se e solo se l'ideale rispetto a cui si quozienta è massimale. In questo caso, però, $\mathcal{M}(p)$ è anche l'unico ideale massimale.

Lemma 1.5. 1. Il germe $f \in \mathcal{E}(p)$ è invertibile se, e solo se, $f(p) \neq 0$.

2. Ogni ideale proprio di $\mathcal{E}(p)$ è contenuto in $\mathcal{M}(p)$.

Dimostrazione 1. Se $f \in \mathcal{E}(p)$ è invertibile, esiste $g \in \mathcal{E}(p)$ tale che $f \cdot g \equiv 1$. Calcolando quest'uguaglianza in p si ottiene $f(p) \cdot g(p) = 1$, e quindi non può essere $f(p) = 0$.

Se, viceversa, $f(p) \neq 0$ e f è il rappresentante della coppia $(U, f) \in E(p)$, definiamo $V := \{x \in U : f(x) \neq 0\}$. Dato che f è C^∞ su U , in particolare è continua, quindi $V = f^{-1}(\mathbb{R} \setminus \{0\})$ è un aperto; inoltre, si ha banalmente che $(U, f) \sim \left(V, f \Big|_V \right)$. Dato che per definizione f non si annulla in V , ci basta definire

$$[U, f]^{-1} = \left[V, f \Big|_V \right]^{-1} = \left[V, \frac{1}{f} \Big|_V \right].$$

△

2. Sia $I \subset \mathcal{E}(p)$ un ideale proprio; allora non contiene elementi invertibili. Infatti, un ideale ha la proprietà che se a è un elemento dell'ideale e b in elemento qualsiasi dell'anello, allora ab è nell'ideale; se I contenesse un elemento invertibile f , allora conterrebbe $ff^{-1} = 1$ (vediamo f come elemento di I e f^{-1} come elemento dell'anello $\mathcal{E}(p)$) e quindi $1 \cdot g = g$ per ogni $g \in \mathcal{E}(p)$ (per lo stesso motivo), ottenendo l'assurdo $I = \mathcal{E}(p)$. Dato che per quanto appena dimostrato al punto 1. $\mathcal{M}(p)$ è l'insieme degli elementi non invertibili in $\mathcal{E}(p)$, si ha $I \subseteq \mathcal{M}(p)$. □

La prima applicazione della terminologia dei germi di funzioni è una definizione comoda dello spazio tangente in un punto ad un aperto di \mathbb{R}^n . Intuitivamente, lo spazio tangente all'aperto $U \subseteq \mathbb{R}^n$ nel punto $p \in U$ è l'insieme di vettori applicati in p tangenti a U : possiamo usare il formalismo delle derivazioni e dei germi per rendere più precisa quest'idea.

Definizione 1.6. Dato l'aperto $U \subseteq \mathbb{R}^n$, si definisce *spazio tangente a U in $p \in U$* lo spazio

$$T_p U := \text{Der}(\mathcal{E}(p), \mathbb{R}) = \{ \delta : \mathcal{E}(p) \rightarrow \mathbb{R} : \delta \text{ è } \mathbb{R}\text{-lineare, } \delta(fg) = f(p)\delta(g) + g(p)\delta(f) \forall f, g \in \mathcal{E}(p) \}.$$

In altre parole, a un vettore $v = (v_1, \dots, v_n)$ tangente a U in p facciamo corrispondere la derivazione $\sum_{i=1}^n v_i \frac{\partial}{\partial x_i}$; questa derivazione associa a $f \in \mathcal{E}(p)$ il numero reale $\sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(p)$. Calcoliamo questa "derivata direzionale" in p perché questo è l'unico punto dove è ben definito il valore delle (classi di equivalenza di) funzioni in $\mathcal{E}(p)$, e quindi delle loro derivate.

Osservazione 1.5. Se $p \in U \subseteq V$, dove U e V sono aperti di \mathbb{R}^n , allora $T_p U = T_p V$.

Lo spazio tangente $T_p U$ è ovviamente un \mathbb{R} -spazio vettoriale; dimostriamo ora che, in effetti, è una copia di \mathbb{R}^n .

Lemma 1.6. Sia $U \subseteq \mathbb{R}^n$ un aperto e $p \in U$; allora, l'applicazione

$$\varphi : \mathbb{R}^n \rightarrow T_p U, \quad v = (v_1, \dots, v_n) \mapsto \sum_{i=1}^n v_i \frac{\partial}{\partial x_i}$$

è un'isomorfismo di spazi vettoriali.

Dimostrazione. Ci basta verificare che φ sia una biiezione. Asseriamo che, se $\delta: \mathcal{E}(p) \rightarrow \mathbb{R}$ è una derivazione (ovvero un elemento di $T_p U$) e x_1, \dots, x_n sono coordinate in \mathbb{R}^n , l'applicazione

$$\psi: T_p U \rightarrow \mathbb{R}^n, \quad \delta \mapsto (\delta(x_1), \dots, \delta(x_n))$$

è l'inversa di φ . Sia infatti $f \in \mathcal{E}(p)$; per la formula di Taylor (che può essere applicata, perché il fatto che f sia un germe ci dà informazioni sulle sue proprietà locali), per certe $f_i \in C^\infty(U)$, $i = 1, \dots, n$, vale l'uguaglianza

$$f(x) = f(p) + \sum_{i=1}^n (x_i - p_i) f_i(x), \quad \text{con } f_i(p) = \frac{\partial f}{\partial x_i}(p).$$

Applicando δ ad entrambi i membri e calcolando in $x = p$, come abbiamo già fatto in precedenza (nella dimostrazione del Teorema 1.3), otteniamo

$$\delta(f) = \sum_{i=1}^n \delta(x_i) \frac{\partial f}{\partial x_i}(p). \quad (6)$$

Possiamo quindi calcolare

$$\begin{aligned} \varphi(\psi(\delta)) &= \varphi(\delta(x_1), \dots, \delta(x_n)) = \sum_{i=1}^n \delta(x_i) \frac{\partial}{\partial x_i} = \delta \quad (\text{per la (6)}), \\ \psi(\varphi(v_1, \dots, v_n)) &= \psi\left(\sum_{i=1}^n v_i \frac{\partial}{\partial x_i}\right) = \left(\sum_{i=1}^n v_i \frac{\partial x_1}{\partial x_i}, \dots, \sum_{i=1}^n v_i \frac{\partial x_n}{\partial x_i}\right) \\ &= (v_1, \dots, v_n) \quad (\text{perché } \frac{\partial x_j}{\partial x_i} = \delta_i^j) \end{aligned}$$

provando che φ e ψ sono l'una inversa dell'altra. \square

Consideriamo ora un'applicazione $F: U \subseteq \mathbb{R}^n \rightarrow V \subseteq \mathbb{R}^m$ di classe C^∞ , e sia $W \subseteq V$ un aperto. Definiamo l'omomorfismo di anelli

$$F^* := \cdot \circ F: C^\infty(W) \rightarrow C^\infty(F^{-1}(W)), \quad f \mapsto f \circ F.$$

Possiamo "restringere" F^* ai germi: se $p \in U$ e $q = F(p) \in V$ allora risulta ben definita

$$F^*: \mathcal{E}(q) \rightarrow \mathcal{E}(p), \quad [W, f] \mapsto [F^{-1}(W), f \circ F].$$

Quest'applicazione è un omomorfismo di anelli e un morfismo di \mathbb{R} -spazi vettoriali; in altre parole $F^*(1) = 1$ e $F^*(f)(p) = f(q)$ per ogni $f \in \mathcal{E}(q)$.

Se $\delta: \mathcal{E}(p) \rightarrow \mathbb{R}$ è una derivazione (ovvero un vettore tangente a U in p), allora anche $\delta \circ F^*: \mathcal{E}(q) \rightarrow \mathbb{R}$ è una derivazione (ovvero un vettore tangente a V in q). Verifichiamo che $\delta \circ F^*$ soddisfa la regola di Leibniz (la verifica della \mathbb{R} -linearità è ancora una volta lasciata per esercizio); se f e g sono germi in $\mathcal{E}(q)$, allora

$$\begin{aligned} \delta \circ F^*(fg) &= \delta(F^*(fg)) = \delta(F^*(f) \cdot F^*(g)) \quad (\text{perché } F^* \text{ è un omomorfismo di anelli}) \\ &= F^*(f)(p) \delta(F^*(g)) + F^*(g)(p) \delta(F^*(f)) \quad (\text{perché } \delta \text{ soddisfa la regola di Leibniz}) \\ &= f(q) \delta \circ F^*(g) + g(q) \delta \circ F^*(f). \end{aligned}$$

Definizione 1.7. Nelle notazioni precedenti, si definisce *differenziale di F in p* l'applicazione

$$DF_p: T_p U \simeq \text{Der}(\mathcal{E}(p), \mathbb{R}) \rightarrow T_q V \simeq \text{Der}(\mathcal{E}(q), \mathbb{R}), \quad \delta \mapsto \delta \circ F^*.$$

Verifichiamo che questa definizione coincide con quella “usuale”. Siano x_1, \dots, x_n delle coordinate su U e y_1, \dots, y_m coordinate su V in maniera tale che $y_j = F_j(x_1, \dots, x_n)$, $j = 1, \dots, m$. Consideriamo $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1, \dots, n}$ come base di $T_p U \simeq \text{Der}(\mathcal{E}(p), \mathbb{R})$, $p \in U$, e $\left\{ \frac{\partial}{\partial y_j} \right\}_{j=1, \dots, m}$ come base di $T_q V \simeq \text{Der}(\mathcal{E}(q), \mathbb{R})$, $q = F(p)$. Se $f \in \mathcal{E}(q)$, si ha

$$\begin{aligned} DF_p \left(\sum_{i=1}^n \frac{\partial}{\partial x_i} \right) (f) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} (F_1(p), \dots, F_m(p)) \\ &= \sum_{i=1}^n \sum_{j=1}^m \frac{\partial f}{\partial y_j} (F_1(p), \dots, F_m(p)) \frac{\partial F_j}{\partial x_i} (p), \end{aligned}$$

ovvero, volendo sottolineare la natura operatoriale di DF_p ,

$$DF_p \left(\sum_{i=1}^n \frac{\partial}{\partial x_i} \right) = \sum_{j=1}^m \left[\sum_{i=1}^n \frac{\partial F_j}{\partial x_i} (p) \right] \frac{\partial}{\partial y_j}.$$

Appare ora evidente che la matrice che rappresenta l'operatore DF_p in queste basi è

$$DF_p = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

ovvero la *matrice jacobiana* di F in p , $\mathcal{J}F_p = \left(\frac{\partial F_i}{\partial x_j} \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$.

1.5 Teorema di invertibilità locale

In questa sezione richiamiamo un risultato noto dai corsi di Analisi, riformulandolo nella terminologia che abbiamo appena introdotto, e ne esaminiamo alcune conseguenze. Ricordiamo che un *diffeomorfismo* è un'applicazione C^∞ che sia biunivoca e tale che l'inversa è ancora di classe C^∞ .

Teorema 1.4 (di invertibilità locale). *Siano U e V due aperti di \mathbb{R}^n , $F: U \rightarrow V$ una funzione C^∞ e $p \in U$; se $\det DF_p \neq 0$, allora esistono un intorno aperto $W \subseteq U$ di p e un intorno aperto $Z \subseteq V$ di $F(p)$ tali che $F|_W : W \rightarrow Z$ sia un diffeomorfismo.*

Osserviamo che un diffeomorfismo ha sempre jacobiana invertibile, dove definito. Infatti, dato che la jacobiana della composizione di due funzioni è la composizione (ovvero il prodotto, in termini matriciali) delle loro jacobiane, si ha che se F è un diffeomorfismo

$$\begin{aligned} F \circ F^{-1} = \text{id} &\implies \mathcal{J}(F \circ F^{-1}) = I = \mathcal{J}F \cdot \mathcal{J}F^{-1} \\ &\implies \mathcal{J}F_p \cdot \mathcal{J}F_p^{-1} = I \quad (\text{calcolando in } p) \\ &\implies \det \mathcal{J}F_p \cdot \det \mathcal{J}F_p^{-1} = 1 \\ &\implies \det \mathcal{J}F_p \neq 0. \end{aligned}$$

Corollario 1.2. *Siano U e V due aperti di \mathbb{R}^n ; allora, una funzione $F: U \rightarrow V$ di classe C^∞ è un diffeomorfismo globale se, e solo se, è biunivoca e il suo differenziale DF_p è invertibile in ogni punto $p \in U$.*

Dimostrazione. L'implicazione \implies è ovvia, quindi dimostriamo l'implicazione \impliedby . Per l'ipotesi di biunivocità di F , sappiamo che è ben definita $F^{-1}: V \rightarrow U$; quello che ci resta da verificare è che sia di classe C^∞ . Possiamo applicare in ogni punto $p \in U$ il Teorema 1.4, in quanto per ipotesi il differenziale ha sempre determinante non nullo: quindi per ogni $p \in U$ esistono un intorno aperto $W \subseteq U$ di p e un intorno aperto $Z \subseteq V$ di $F(p)$ tali che $F \Big|_W : W \rightarrow Z$ sia un diffeomorfismo, e quindi in particolare tali che $F^{-1} \Big|_Z : Z \rightarrow W$ sia di classe C^∞ . È sufficiente ora osservare che la proprietà di essere C^∞ è per definizione una proprietà locale, e che tutti i possibili Z che otteniamo ricoprono V . \square

Corollario 1.3 (Teorema del Dini o delle funzioni implicite). *Siano $U \subseteq \mathbb{R}^n$ e $V \subseteq \mathbb{R}^m$ aperti, $p \in U$ e $F: U \rightarrow V$ di classe C^∞ tale che $F(p) = 0$ e DF_p sia suriettivo (in particolare si ha $m \leq n$ e $\text{rank} DF_p = m$); allora, esistono un intorno aperto $W \subseteq U$ di p , un intorno aperto $Z \subseteq \mathbb{R}^n$ di 0 e un diffeomorfismo $h: W \rightarrow Z$ tali che $F \circ h^{-1}(z_1, \dots, z_n) = (z_1, \dots, z_m)$.*

In altre parole, a meno di un cambio di coordinate in partenza (ovvero della funzione h) possiamo considerare la funzione F come una proiezione sulle prime m coordinate.

Dimostrazione. Dall'ipotesi che $\text{rank} \left(\frac{\partial F_i}{\partial x_j}(p) \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = m$, a meno di ripermutare le colonne della matrice jacobiana (un'azione di cui possiamo eventualmente tenere conto nella definizione di h), possiamo supporre che il determinante del minore $\left(\frac{\partial F_i}{\partial x_j}(p) \right)_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$ sia non nullo. Definiamo quindi

$$h: U \rightarrow \mathbb{R}^n, \quad x = (x_1, \dots, x_n) \mapsto (F_1(x), \dots, F_m(x), x_{m+1}, \dots, x_n) = (z_1, \dots, z_n).$$

L'applicazione h è banalmente di classe C^∞ . Inoltre, la matrice jacobiana di h calcolata in p risulta essere la matrice a blocchi

$$Dh_p = \left(\begin{array}{c|c} \frac{\partial F_i}{\partial x_j}(p) & * \\ \hline 0 & I \end{array} \right) \begin{array}{l} \left. \vphantom{\frac{\partial F_i}{\partial x_j}(p)} \right\} m \\ \left. \vphantom{I} \right\} n-m \end{array}$$

ed è quindi invertibile. Quindi, a meno di restringersi a un intorno aperto più piccolo Z di 0 , possiamo affermare che $h: W \rightarrow Z$ è un diffeomorfismo. Per costruzione, inoltre, abbiamo che $F \circ h^{-1}$ è la proiezione sulle prime m coordinate, ottenendo quindi la tesi. \square

1.6 Esercizi svolti

Esercizio 1.1. Sia $\det: \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow \mathbb{R}$ il determinante usuale; sotto quali condizioni su $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ il differenziale $D\det_A: T_A \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow T_{\det A} \mathbb{R}$ è suriettivo?

Soluzione. Tale differenziale è suriettivo se e solo se A ha rango almeno $n-1$. Ad esempio, se A è invertibile allora le matrici che hanno determinante pari a $\det A$ sono, almeno localmente, un iperpiano.

Per dimostrare ciò, osserviamo che ogni vettore tangente $v \in T_A \mathfrak{M}_{n,n}(\mathbb{R}) \simeq \mathfrak{M}_{n,n}(\mathbb{R})$ si può scrivere nella forma

$$v = D\gamma \left(\frac{d}{dt} \right) \quad \text{dove } \gamma: (-\varepsilon, \varepsilon) \rightarrow \mathfrak{M}_{n,n}(\mathbb{R}), \quad 0 \mapsto A;$$

questo semplicemente perché possiamo interpretare sempre un vettore tangente come la “velocità” di una curva. Ad esempio, se $\gamma(t) = A + tA_1 + t^2A_2$, con $A_1, A_2 \in \mathfrak{M}_{n,n}(\mathbb{R})$ allora $D\gamma \left(\frac{d}{dt} \right) (0) = A_1$ (è la derivata formale rispetto a t dell'espressione di γ , calcolata per $t=0$; la tangente a γ in 0 è il “vettore” $v = A_1$).

Per studiare $D\det_A$ ci basta quindi studiare $D\det(\gamma): (-\delta, \delta) \rightarrow T_{\det A} \mathbb{R} \simeq \mathbb{R}$ al variare delle curve γ che mandano 0 in A ; ci siamo quindi ricondotti a studiare una “derivata in una variabile”. In particolare, dovendo studiare la suriettività del differenziale del determinante, siamo interessati a quali γ siano nel nucleo di $D\det$, ovvero per quali γ si abbia $D\det(\gamma) = 0$.

Sia B un generico vettore tangente a $\mathfrak{M}_{n,n}(\mathbb{R})$ in A (facciamo notare ancora una volta che di fatto $B = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ è una matrice!), e scriviamo $\gamma(t) = A + tB + o(t)$. Ricordando la definizione di determinante, otteniamo che

$$\det \gamma(t) = \det A + t \left(\sum_{i,j=1}^n \pm b_{ij} \det A^{ij} \right) + o(t).$$

Infatti, per calcolare il determinante della matrice $A + tB$ dobbiamo considerare le permutazioni di tutti gli elementi della matrice stessa, che non solo determinano il segno \pm nella somma ma anche il minore $(A + tB)^{ij}$ di cui dobbiamo calcolare, ricorsivamente, il determinante. Ne segue che i termini di grado 0 in t compaiono quando la permutazione non fa apparire termini della matrice tB (e quindi fa comparire solo elementi di A), mentre che i termini di grado 1 si ottengono quando compare un elemento della matrice tB e tutti gli altri $n - 1$ vengono presi dalla matrice A . Giustificata la scrittura che abbiamo usato, appare ora chiaro che se $\det A^{ij} = 0$ per ogni $i, j = 1, \dots, n$ (o, equivalentemente, se $\text{rank } A < n - 1$) allora $\det \gamma(t)$ è costante in t (valendo $\det A$), e $D\det_A = 0$. Viceversa, se il rango di A vale almeno $n - 1$, esiste un minore (che possiamo supporre essere quello “in alto a sinistra”, a meno di permutazioni) di determinante $s = \det(a_{ij})_{\substack{i=1,\dots,n-1 \\ j=1,\dots,n-1}} \neq 0$. Scegliendo come B la matrice che ha come unico elemento non nullo l’elemento $b_{nn} = 1$, otteniamo che $\det(A + tB) = \det A + ts$ e quindi $D\det(A + tB)_A = s \neq 0$. \diamond

Esercizio 1.2. Siano $S \subset \mathfrak{M}_{n,n}(\mathbb{R})$ il sottospazio delle matrici simmetriche e $F: \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow S$ l’applicazione C^∞ definita da $A \mapsto AA^T$. Dimostrare che se $E \in O(n)$ è una matrice ortogonale, allora DF_E è suriettivo.

Soluzione. Siano $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ e $\gamma_A: (-\varepsilon, \varepsilon) \rightarrow \mathfrak{M}_{n,n}(\mathbb{R})$ data da $t \mapsto (I + tA)E$. Si ha allora

$$\dot{\gamma}_A(0) := D\gamma_A \left(\frac{d}{dt} \right) (0) = AE$$

per cui, per l’invertibilità di E , al variare di $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ si possono così ottenere tutti i vettori tangenti di $T_E \mathfrak{M}_{n,n}(\mathbb{R}) \simeq \mathfrak{M}_{n,n}(\mathbb{R})$ (se vogliamo ottenere la matrice B , ci basta porre $A = BE^{-1}$).

Calcoliamo

$$F(\gamma_A(t)) = (I + tA)EE^T(I + tA^T) = I + t(A + A^T) + t^2 AA^T$$

dove abbiamo usato che $EE^T = I$ essendo E ortogonale. Da questo ricaviamo che

$$DF(\gamma_A) \left(\frac{d}{dt} \right) (0) = \frac{d}{dt} F \circ \gamma_A(0) = A + A^T;$$

dato che ogni matrice simmetrica è somma di una matrice e della sua trasposta (prendiamo come matrice la metà superiore della matrice simmetrica e dividiamo tutti gli elementi della diagonale per 2, mettendo tutti elementi nulli nella metà inferiore), abbiamo che DF_E è suriettivo. Osserviamo anche che $\ker DF_E = \{AE : A \text{ antisimmetrica}\}$: questo ci dice che lo spazio tangente a $O(n)$ è dato da quello delle matrici antisimmetriche. \diamond

Esercizio 1.3. Sia $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ di classe C^∞ e di coordinate $F = (F_1, \dots, F_m)$, e sia $Z = \{x \in \mathbb{R}^n : f(x) = 0\}$ il suo luogo di zeri. Se per ogni $p \in Z$ il differenziale $DF_p: T_p \mathbb{R}^n \rightarrow T_0 \mathbb{R}^m$ è suriettivo e $f \in C^\infty(\mathbb{R}^n)$ è tale che $f \Big|_Z \equiv 0$, allora esistono delle funzioni $f_j \in C^\infty(\mathbb{R}^n)$, $j = 1, \dots, m$, tali che

$$f = \sum_{j=1}^m F_j \cdot f_j.$$

Soluzione. Vediamo che un'uguaglianza del genere è vera localmente, poi estendiamo tale proprietà a tutto \mathbb{R}^n .

Sia dunque $p \in \mathbb{R}^n$. Se $F(p) \neq 0$, allora per la permanenza del segno esistono un indice $j \in \{1, \dots, m\}$ e un aperto U contenente p tali che $F_j(x) \neq 0$ per ogni $x \in U$. Ci basta quindi scrivere $f = F_j \cdot \frac{f}{F_j}$ per ottenere l'uguaglianza richiesta su U . Se, al contrario, $F(p) = 0$ (ovvero $p \in Z$), per il teorema delle funzioni implicite esistono un intorno aperto $V \subseteq \mathbb{R}^n$ di p , un intorno aperto $W \subseteq \mathbb{R}^m$ di 0 e un diffeomorfismo $h: V \rightarrow W$ tali che $h(p) = 0$ e $F \circ h^{-1}(z_1, \dots, z_n) = (z_1, \dots, z_m)$. Da questo segue che l'immagine secondo h di $Z \cap V$ è il sottospazio di equazioni $\{z_1 = \dots = z_m = 0\}$; interpretando quindi (F_1, \dots, F_m) come le prime m coordinate locali, ci basta vedere che esistono delle funzioni φ_j , $j = 1, \dots, m$, tali che in un'intorno di 0 contenuto in W si abbia

$$f \circ h^{-1} = z_1 \varphi_1 + \dots + z_m \varphi_m.$$

Questo discende da una generalizzazione della formula di Taylor, che enunciamo di seguito (la dimostrazione è lasciata per esercizio: basta riadattare opportunamente quella del Lemma 1.4). s

Lemma 1.7. *Siano $A \subseteq \mathbb{R}^m \times \mathbb{R}^s$ un aperto contenente $(0, 0)$, e $f \in C^\infty(A)$; allora esistono un intorno aperto $B \subseteq A$ di $(0, 0)$ e delle funzioni $f_j \in C^\infty(A)$, $j = 1, \dots, m$, tali che*

$$f(x, y) = f(0, y) + \sum_{j=1}^m x_j f_j(x, y) \quad \forall (x, y) \in B.$$

Applicando tale risultato nel nostro caso, otteniamo che

$$f \circ h^{-1}(z_1, \dots, z_m) = f \circ h^{-1}(0, \dots, 0, z_{m+1}, \dots, z_n) + \sum_{j=1}^m z_j \varphi_j(z_1, \dots, z_m) = \sum_{j=1}^m z_j \varphi_j(z_1, \dots, z_m)$$

poiché, essendo $(0, \dots, 0, z_{m+1}, \dots, z_n) \in Z \cap V$, è $f \circ h^{-1}(0, \dots, 0, z_{m+1}, \dots, z_n) = 0$.

Possiamo quindi concludere che esiste un ricoprimento aperto $\mathbb{R}^n = \bigcup_{h \in H} U_h$ tale che $f|_{U_h} = \sum_{j=1}^m F_j f_{jh}$, per ogni $h \in H$, per certe $f_{jh} \in C^\infty(U_h)$. Se $\{\varphi_h\}_{h \in H}$ è la partizione dell'unità subordinata a $\{U_h\}_{h \in H}$, allora su tutto \mathbb{R}^n vale

$$f = f \cdot 1 = f \sum_{h \in H} \varphi_h = \sum_{h \in H} f \varphi_h = \sum_{j=1}^m F_j \sum_{h \in H} f_{jh} \varphi_h.$$

Ci basta porre $f_j := \sum_{h \in H} f_{jh} \varphi_h$ per ottenere la tesi. ◇

2 Varietà differenziabili

In questa sezione introduciamo l'oggetto matematico che viene principalmente studiato dalla Geometria Differenziale: la varietà differenziabile. Occorre però partire dal concetto di varietà topologica.

2.1 Varietà topologiche

Definizione 2.1. Sia X uno spazio topologico; X si dice una *varietà topologica di dimensione n* se

1. X è di Hausdorff;
2. X è localmente omeomorfo ad aperti di \mathbb{R}^n , ovvero per ogni $p \in X$ esistono un intorno aperto U di p e un aperto V di \mathbb{R}^n tali che esista un omeomorfismo $\varphi: U \rightarrow V$; la tripla (U, V, φ) viene detta *carta locale*;
3. X è a base numerabile.

L'ipotesi 3. è talvolta sostituita in letteratura assumendo che X sia *paracompatto*, ovvero che ogni ricoprimento aperto di X ammetta un raffinamento localmente finito. Il fatto di essere Hausdorff, localmente omeomorfo ad aperti di \mathbb{R}^n e paracompatto equivale a dire che ogni componente connessa dello spazio è a base numerabile; in questo senso le due ipotesi di numerabilità e paracompattatezza coincidono se X è connesso o se ha al più un'infinità numerabile di componenti connesse (come accade nella quasi totalità dei casi).

Un qualsiasi aperto $U \subseteq \mathbb{R}^n$ è un esempio di varietà topologica di dimensione n ; è sufficiente la carta (U, U, id) .

Osservazione 2.1. Può sembrare che, nella definizione di varietà topologica, la seconda ipotesi renda la prima ridondante; essere Hausdorff è una proprietà stabile per omeomorfismo. In realtà non è così; consideriamo come controesempio la retta reale con "due zeri", ovvero il quoziente $X = \mathbb{R} \times \{0, 1\} / \sim$ dove \sim è la relazione d'equivalenza

$$(x, 0) \sim (y, 1) \quad \text{se} \quad x = y \neq 0.$$

Si ha che $(\mathbb{R} \times \{0\}, \mathbb{R}, \text{id})$ e $(\mathbb{R} \times \{1\}, \mathbb{R}, \text{id})$ sono due carte locali ($\mathbb{R} \times \{0\}$ e $\mathbb{R} \times \{1\}$ sono aperti nella topologia quoziente) ma i due zeri (che non vengono identificati dalla relazione \sim) non possono avere intorni disgiunti.

Definizione 2.2. Sia X una varietà topologica di dimensione n ; si definisce *atlante* per X una famiglia di carte $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ tali che gli U_i ricoprano X : $X = \bigcup_{i \in I} U_i$.

L'atlante \mathfrak{A} si dice *di classe C^k* , $k = 0, \dots, \infty$, se per ogni $i, j \in I$ tali che $U_i \cap U_j \neq \emptyset$ si ha che le *funzioni di transizione* (o di *cociclo*)

$$\begin{array}{ccc} & U_i \cap U_j & \\ \varphi_i \swarrow & & \searrow \varphi_j \\ \varphi_i(U_i \cap U_j) \subseteq V_i & \xrightarrow{\varphi_j \circ \varphi_i^{-1}} & \varphi_j(U_i \cap U_j) \subseteq V_j \end{array}$$

sono diffeomorfismi di classe C^k dove definite. Notiamo che l'insieme di definizione e il codominio di queste funzioni sono aperti di \mathbb{R}^n , quindi il concetto di C^k ha senso.

Esempio 2.1. La sfera $S^n = \left\{ x = (x_0, \dots, x_n) \in \mathbb{R}^{n+1} : \sum_{i=0}^n x_i^2 = 1 \right\}$ è una varietà topologica di dimensione n ; ne vogliamo dare un atlante C^∞ . Scriviamo $S^n = U_1 \cup U_2$, dove $U_1 = S^n \setminus \{(1, 0, \dots, 0)\}$ e $U_2 = S^n \setminus \{(-1, 0, \dots, 0)\}$ sono omeomorfi per proiezione stereografica a \mathbb{R}^n . Più precisamente, se

$$\begin{aligned} \varphi_1: U_1 &\rightarrow \mathbb{R}^n, & x = (x_0, \dots, x_n) &\mapsto \frac{1}{1-x_0}(x_1, \dots, x_n), \\ \varphi_2: U_2 &\rightarrow \mathbb{R}^n, & x = (x_0, \dots, x_n) &\mapsto \frac{1}{1+x_0}(x_1, \dots, x_n), \end{aligned}$$

allora $\mathfrak{A} = \{(U_1, \mathbb{R}^n, \varphi_1), (U_2, \mathbb{R}^n, \varphi_2)\}$ è un atlante per S^n . Quello che ci resta da verificare è che le funzioni di transizione (che in questo caso sono due, $\varphi_2 \circ \varphi_1^{-1}$ e la sua inversa $\varphi_1 \circ \varphi_2^{-1}$) sono di classe $C^\infty(\mathbb{R}^n \setminus \{0\})$. Dimosteremo solo che $\varphi_2 \circ \varphi_1^{-1}$ soddisfa questa proprietà, computando esplicitamente φ_1^{-1} ; in maniera del tutto analoga si dimostra che anche l'altra funzione di transizione è C^∞ .

Se $\varphi_1(x_0, \dots, x_n) = (y_1, \dots, y_n)$, allora per ogni $i = 1, \dots, n$ si ha

$$y_i = \frac{x_i}{1-x_0} \implies \sum_{i=1}^n y_i^2 = \frac{\sum_{i=1}^n x_i^2}{(1-x_0)^2} = \frac{1-x_0^2}{(1-x_0)^2} = \frac{1+x_0}{1-x_0}$$

e quindi

$$(1-x_0) \sum_{i=1}^n y_i^2 = 1+x_0 \implies x_0 = \frac{\sum_{i=1}^n y_i^2 - 1}{\sum_{i=1}^n y_i^2 + 1}.$$

Ne concludiamo che

$$\varphi_1^{-1}(y_1, \dots, y_n) = \left(\frac{\sum_{i=1}^n y_i^2 - 1}{\sum_{i=1}^n y_i^2 + 1}, \frac{2y_1}{\sum_{i=1}^n y_i^2 + 1}, \dots, \frac{2y_n}{\sum_{i=1}^n y_i^2 + 1} \right)$$

e quindi φ_1^{-1} è una funzione C^∞ dove definita. Dato che anche φ_2 è C^∞ , la funzione di transizione lo è in quanto composizione di funzioni C^∞ è C^∞ .

Esercizio 2.1. Dimostrare che lo spazio proiettivo $\mathbb{P}_{\mathbb{R}}^n$ è una varietà topologica di dimensione n che possiede un atlante C^∞ .

Suggerimento: scrivere

$$\mathbb{P}_{\mathbb{R}}^n = \bigcup_{i=0}^n U_i, \quad U_i = \{[x_0, \dots, x_n] \in \mathbb{P}_{\mathbb{R}}^n : x_i \neq 0\}$$

e dimostrare che ciascun U_i è omeomorfo a \mathbb{R}^n usando l'applicazione

$$\varphi_i: U_i \rightarrow \mathbb{R}^n, \quad [x_0, \dots, x_n] \mapsto \left(\frac{x_0}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_n}{x_i} \right).$$

Verificare inoltre che

$$\varphi_i^{-1}(y_1, \dots, y_n) = [y_1, \dots, y_{i-1}, 1, y_i, \dots, y_n].$$

Esercizio 2.2 (☛). Dimostrare che la Grassmanniana

$$G(k, n) := \{V \subseteq \mathbb{R}^n : \dim V = k\}$$

è una varietà topologica di dimensione $k(n-k)$ che possiede un atlante C^∞ .

2.2 Varietà differenziabili

D'ora in poi sottointenderemo sempre che gli atlanti siano atlanti C^∞ ; useremo anche la terminologia *atlanti differenziabili*.

Sia X una varietà topologica di dimensione n . Consideriamo la seguente relazione \sim fra atlanti di X :

$$\mathfrak{A}_1 \sim \mathfrak{A}_2 \quad \text{se} \quad \mathfrak{A}_1 \cup \mathfrak{A}_2 \text{ è ancora un atlante;}$$

in altre parole, se $\mathfrak{A}_1 = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ e $\mathfrak{A}_2 = \{(Z_j, W_j, \psi_j)\}_{j \in J}$, diremo che \mathfrak{A}_1 è equivalente ad \mathfrak{A}_2 se per ogni scelta di $i \in I$ e $j \in J$ tali che $U_i \cap Z_j \neq \emptyset$ la funzione di transizione $\psi_j \circ \varphi_i^{-1}: \varphi_i(U_i \cap Z_j) \subseteq V_i \rightarrow \psi_j(U_i \cap Z_j) \subseteq W_j$ è di classe C^∞ .

Verifichiamo la transitività della relazione \sim (riflessività e simmetria sono ovvie), in quanto faremo uso di un ragionamento che utilizzeremo anche in seguito. Siano dunque $\mathfrak{A}_1 = \{(U_i, V_i, \varphi_i)\}_{i \in I}$, $\mathfrak{A}_2 = \{(Z_j, W_j, \psi_j)\}_{j \in J}$ e $\mathfrak{A}_3 = \{(A_h, B_h, \eta_h)\}_{h \in H}$ tre atlanti su X , con $\mathfrak{A}_1 \sim \mathfrak{A}_2$ e $\mathfrak{A}_2 \sim \mathfrak{A}_3$. Quello che vogliamo mostrare è che per ogni $i \in I$ e $h \in H$ tali che $U_i \cap A_h \neq \emptyset$ la funzione di transizione $\eta_h \circ \varphi_i^{-1}: \varphi_i(U_i \cap A_h) \subseteq V_i \rightarrow \eta_h(U_i \cap A_h) \subseteq B_h$ è di classe C^∞ . Notiamo innanzitutto che, dato che $\{Z_j\}_{j \in J}$ è un ricoprimento di X , si ha

$$U_i \cap A_h = \bigcup_{j \in J} U_i \cap A_h \cap Z_j;$$

inoltre, poiché l'essere di classe C^∞ è una proprietà locale e $\{\varphi_i(U_i \cap A_h \cap Z_j)\}_{j \in J}$ è un ricoprimento aperto di $\varphi_i(U_i \cap A_h)$, ci basta verificare che $\eta_h \circ \varphi_i^{-1} \Big|_{\varphi_i(U_i \cap A_h \cap Z_j)}$ è di classe C^∞ . A tale scopo osserviamo che

$$\eta_h \circ \varphi_i^{-1} = (\eta_h \circ \psi_j^{-1}) \circ (\psi_j \circ \varphi_i^{-1})$$

è di classe C^∞ perché composizione di funzioni C^∞ (come discende dall'equivalenza di \mathfrak{A}_1 con \mathfrak{A}_2 e di \mathfrak{A}_2 con \mathfrak{A}_3).

$$\begin{array}{ccccc} & & U_i \cap A_h \cap Z_j & & \\ & \swarrow \varphi_i & \downarrow \psi_j & \searrow \eta_h & \\ \varphi_i(U_i \cap A_h \cap Z_j) \subseteq V_i & \xrightarrow{\psi_j \circ \varphi_i^{-1}} & \psi_j(U_i \cap A_h \cap Z_j) \subseteq W_j & \xrightarrow{\eta_h \circ \psi_j^{-1}} & \eta_h(U_i \cap A_h \cap Z_j) \subseteq B_h \\ & \searrow \eta_h \circ \varphi_i^{-1} & & & \end{array}$$

Definizione 2.3. Sia X una varietà topologica di dimensione n ; una *struttura differenziabile* su X è una classe di equivalenza di atlanti differenziabili rispetto alla relazione \sim appena introdotta. La varietà X dotata di una struttura differenziabile viene detta *varietà differenziabile*; la dimensione di una varietà differenziabile è la sua dimensione come varietà topologica.

La definizione di atlanti equivalenti ci è utile per aggiungere carte ad un atlante senza mutare la struttura differenziabile dello spazio. Infatti, se $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ è un atlante su X e $h: V_{i_0} \rightarrow W_{i_0}$ è un diffeomorfismo fra aperti di \mathbb{R}^n , allora $\mathfrak{A} \cup \{(U_{i_0}, W_{i_0}, h \circ \varphi_{i_0})\}$ è ancora un atlante, perché per ogni $i \in I$ la funzione di transizione $\varphi_i \circ (h \circ \varphi_{i_0})^{-1} = (\varphi_i \circ \varphi_{i_0}^{-1}) \circ h^{-1}$ è di classe C^∞ . Inoltre, se $Z_i \subseteq U_i$ è un aperto, allora $\mathfrak{A} \cup \left\{ \left(Z_i, \varphi_i(Z_i), \varphi_i \Big|_{Z_i} \right) \right\}$ è ancora un atlante (la verifica, molto semplice, è lasciata al lettore). Possiamo quindi modificare le carte nel nostro atlante mediante diffeomorfismi in arrivo oppure per restrizione e aggiungerle all'atlante stesso senza modificare la struttura differenziabile sulla varietà in esame.

Esempio 2.2. Sia $X = \mathbb{R}$, e consideriamo i due atlanti $\mathfrak{A}_1 = \{(\mathbb{R}, \mathbb{R}, \text{id})\}$ e $\mathfrak{A}_2 = \{(\mathbb{R}, \mathbb{R}, \varphi)\}$ dove $\varphi(t) = t^3$; allora i due atlanti \mathfrak{A}_1 e \mathfrak{A}_2 non sono equivalenti. Infatti si ha

$$\begin{array}{ccc} & \mathbb{R} & \\ \varphi \swarrow & & \searrow \text{id} \\ \mathbb{R} & \xrightarrow{\sqrt[3]{\cdot}} & \mathbb{R} \end{array}$$

e l'applicazione $t \mapsto \sqrt[3]{t}$ non è differenziabile nel punto $t = 0$ (e a maggior ragione non può essere C^∞).

Potevamo fare lo stesso ragionamento con $\varphi(t) = t^{2n+1} + a$ per ogni $n \in \mathbb{N}$ e $a \in \mathbb{R}$; questo potrebbe portarci a pensare che su \mathbb{R} sia possibile definire addirittura un'infinità continua di strutture differenziabili! Tuttavia, come vedremo in seguito, molte di queste sono canonicamente isomorfe.

Vogliamo ora generalizzare il concetto di funzione C^∞ da \mathbb{R}^n a varietà differenziabili qualsiasi; siamo motivati dal fatto che le varietà differenziabili sono localmente omeomorfe (e quindi, topologicamente parlando, "indistinguibili") ad aperti di \mathbb{R}^n , e che la definizione di C^∞ sia qualcosa che riguarda solo le proprietà locali delle funzioni.

Siano dunque X una varietà differenziabile, $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ un atlante che ne definisce la struttura differenziabile e U un aperto di X . Denotiamo (provvisoriamente) con $C_{\mathfrak{A}}^\infty(U)$ l'insieme delle funzioni che, "lette" sulle carte locali, sono C^∞ :

$$C_{\mathfrak{A}}^\infty(U) := \left\{ f: U \rightarrow \mathbb{R} : f \circ \varphi_i^{-1} \Big|_{\varphi_i(U \cap U_i)} \text{ è di classe } C^\infty \text{ per ogni } i \in I \right\}.$$

$$\begin{array}{ccc} U \cap U_i & \xrightarrow{i} & U & \xrightarrow{f} & \mathbb{R} \\ \varphi_i \downarrow & & \nearrow f \circ \varphi_i^{-1} & & \\ \varphi_i(U \cap U_i) & \subseteq & V_i & & \end{array}$$

Lemma 2.1. Siano \mathfrak{A} e \mathfrak{B} atlanti su X ; allora \mathfrak{A} è equivalente a \mathfrak{B} se, e solo se, per ogni aperto $U \subseteq X$ si ha $C_{\mathfrak{A}}^\infty(U) = C_{\mathfrak{B}}^\infty(U)$.

Il significato di questo **Lemma** è che le funzioni C^∞ dipendono solo dalla struttura differenziabile e la determinano univocamente.

Dimostrazione. Siano $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ e $\mathfrak{B} = \{(Z_j, W_j, \psi_j)\}_{j \in J}$ i due atlanti, e sia U un aperto di X . Supponiamo che $\mathfrak{A} \sim \mathfrak{B}$; per simmetria, basta dimostrare che $C_{\mathfrak{A}}^\infty(U) \subseteq C_{\mathfrak{B}}^\infty(U)$. Sia dunque $f: U \rightarrow \mathbb{R}$ di classe $C_{\mathfrak{A}}^\infty$; è sufficiente dimostrare che $f \circ \psi_j^{-1}$ è C^∞ su $\psi_j(U \cap Z_j)$ per ogni $j \in J$. Dato che $\{U \cap Z_j \cap U_i\}_{i \in I}$ è un ricoprimento aperto di $U \cap Z_j$, ci basta mostrare che $f \circ \psi_j^{-1}$ è C^∞ su $\psi_j(U \cap Z_j \cap U_i)$ per ogni $j \in J$ e per ogni $i \in I$; ma questo è vero perché

$$f \circ \psi_j^{-1} = (f \circ \varphi_i^{-1}) \circ (\varphi_i \circ \psi_j^{-1})$$

è composizione di applicazioni C^∞ (la prima perché $f \in C_{\mathfrak{A}}^\infty(U)$ e la seconda perché $\mathfrak{A} \sim \mathfrak{B}$). Viceversa, supponiamo che $\mathfrak{A} \not\sim \mathfrak{B}$: questo vuol dire che esistono due indici $i \in I$ e $j \in J$ tali che $h = \psi_j \circ \varphi_i^{-1}$ oppure h^{-1} non è C^∞ . Supponiamo per fissare le idee che h non sia di classe C^∞ ; questo vuol dire che una delle coordinate $\psi_{j,s}: U_i \cap Z_j \rightarrow \mathbb{R}$ di ψ_j non è in $C_{\mathfrak{A}}^\infty$. Del resto, $\psi_{j,s}$ è $C_{\mathfrak{B}}^\infty$ perché composizione di ψ_j con la proiezione sulla s -esima coordinata; quindi $C_{\mathfrak{A}}^\infty \neq C_{\mathfrak{B}}^\infty$. \square

Ora che sappiamo cos'è una funzione C^∞ da una varietà a \mathbb{R} , possiamo definire le applicazioni C^∞ tra varietà: dato che la differenziabilità è una proprietà locale, ci basterà "leggere" le applicazioni sulle carte locali, e utilizzare gli omeomorfismi con gli aperti di \mathbb{R}^n per ricondurci a funzioni che sappiamo già trattare. Chiameremo *carta* C^∞ per la varietà X la cui struttura differenziabile è definita dall'atlante \mathfrak{A} un omeomorfismo $\varphi: U \subseteq X \rightarrow V \subseteq \mathbb{R}^n$ tale che $\mathfrak{A} \cup (U, V, \varphi)$ sia ancora un atlante (e quindi un atlante equivalente a \mathfrak{A}).

Definizione 2.4. Siano X e Y due varietà differenziabili; diremo che la funzione $F: X \rightarrow Y$ è di classe C^∞ (o differenziabile) se è continua e per ogni coppia $(U \subseteq X, V \subseteq \mathbb{R}^n, \varphi), (Z \subseteq Y, W \subseteq \mathbb{R}^m, \psi)$ di carte C^∞ la funzione $\psi \circ F \circ \varphi^{-1}: \varphi(F^{-1}(Z) \cap U) \rightarrow W$ è di classe C^∞ .

$$\begin{array}{ccc} (F^{-1}(Z) \cap U) \subseteq X & \xrightarrow{F} & Z \subseteq Y \\ \varphi \downarrow & & \downarrow \psi \\ \varphi(F^{-1}(Z) \cap U) \subseteq V \subseteq \mathbb{R}^n & \xrightarrow{\psi \circ F \circ \varphi^{-1}} & W \subseteq \mathbb{R}^m \end{array}$$

Diremo inoltre che $F: X \rightarrow Y$ è un *diffeomorfismo* se è C^∞ , invertibile e con inversa C^∞ .

Osservazione 2.2. Composizione di applicazioni C^∞ è ancora C^∞ . Inoltre, varietà diffeomorfe hanno la stessa dimensione.

Esempio 2.3 (Immersione di Veronese). Consideriamo l'applicazione

$$F: \mathbb{P}_{\mathbb{R}}^1 \rightarrow \mathbb{P}_{\mathbb{R}}^2, \quad [x_0, x_1] \mapsto [x_0^2, x_0x_1, x_1^2].$$

Sicuramente l'applicazione $\hat{F}: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^3 \setminus \{0\}$, $\hat{F}(x_0, x_1) = (x_0^2, x_0x_1, x_1^2)$ è di classe C^∞ ; dalla definizione di topologia quoziente, possiamo dedurre che l'applicazione F è quantomeno continua. Verifichiamo ora che è anche C^∞ per una sola coppia di carte (facciamo riferimento alle notazioni dell'Esercizio 2.1): il lettore volenteroso può verificarlo per le altre coppie. Sia $U_1 = \{[x_0, x_1] \in \mathbb{P}_{\mathbb{R}}^1 : x_1 \neq 0\}$ una carta di $\mathbb{P}_{\mathbb{R}}^1$ e $V_0 = \{[y_0, y_1, y_2] \in \mathbb{P}_{\mathbb{R}}^2 : y_0 \neq 0\}$ una carta di $\mathbb{P}_{\mathbb{R}}^2$. Ricordiamo che gli omeomorfismi di queste carte sono

$$\varphi: U_1 \rightarrow \mathbb{R}, \quad [x_0, x_1] \mapsto \frac{x_0}{x_1} \quad \text{e} \quad \psi: V_0 \rightarrow \mathbb{R}^2, \quad [y_0, y_1, y_2] \mapsto \left(\frac{y_1}{y_0}, \frac{y_2}{y_0} \right).$$

Allora

$$F^{-1}(V_0) = \{[x_0, x_1] \in \mathbb{P}_{\mathbb{R}}^1 : x_0^2 \neq 0\}$$

e quindi

$$F^{-1}(V_0) \cap U_1 = \{[x_0, x_1] \in \mathbb{P}_{\mathbb{R}}^1 : x_0x_1 \neq 0\}.$$

L'applicazione $\psi \circ F \circ \varphi^{-1}: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}^2$ è quindi in questo caso l'applicazione

$$t \xrightarrow{\varphi^{-1}} [1, t] \xrightarrow{F} [1, t, t^2] \xrightarrow{\psi} (t, t^2)$$

che è ovviamente di classe C^∞ .

Più in generale, con ragionamenti del tutto analoghi è possibile dimostrare che se $p_0, \dots, p_N \in \mathbb{R}[x_0, \dots, x_n]$ sono polinomi omogenei di grado $r > 0$ e tali che l'intersezione di tutti i loro luoghi di zeri sia ridotta al solo $\{0\}$, allora la funzione

$$G: \mathbb{P}_{\mathbb{R}}^n \rightarrow \mathbb{P}_{\mathbb{R}}^N, \quad [x] = [x_0, \dots, x_n] \mapsto [p_0(x), \dots, p_N(x)]$$

è di classe C^∞ .

Esercizio 2.3. Dimostrare che la proiezione al quoziente $\pi: \mathbb{R}^{n+1} \setminus \{0\} \rightarrow \mathbb{P}_{\mathbb{R}}^n$ è di classe C^∞ .

Esempio 2.4. Riprendiamo l'Esempio 2.2 e dimostriamo che le due varietà $(\mathbb{R}, \mathcal{A}_1)$ e $(\mathbb{R}, \mathcal{A}_2)$ sono diffeomorfe. Consideriamo infatti l'applicazione $F: (\mathbb{R}, \mathcal{A}_1) \rightarrow (\mathbb{R}, \mathcal{A}_2)$, $F(t) = \sqrt[3]{t}$; allora l'applicazione $\varphi \circ F \circ \text{id}^{-1} = \text{id}$ (ricordiamo che $\varphi(t) = t^3$) è C^∞ e invertibile, e quindi F è un diffeomorfismo.

2.3 Sistemi di coordinate locali, sottovarietà

Abbiamo già notato in precedenza che uno dei numerosi “vantaggi” del lavorare su \mathbb{R}^n consiste nell’aver un sistema di coordinate molto semplice da definire (se si vuole, la base duale della base canonica). Dato che sappiamo che le varietà differenziabili sono localmente assimilabili ad aperti di \mathbb{R}^n , saremmo tentati di introdurre anche sulle varietà delle coordinate. Un modo formale per farlo è dato dalla seguente definizione.

Definizione 2.5. Sia X una varietà differenziabile di dimensione n e p un punto di X . Si definisce *sistema di coordinate locali in p* un insieme di n funzioni $x_1, \dots, x_n \in C^\infty(U)$, dove U è un intorno aperto di p , tali che l’applicazione $\varphi = (x_1, \dots, x_n) : U \rightarrow \mathbb{R}^n$ mandi p in 0 e determini una carta locale (ovvero $\varphi(U)$ è aperto in \mathbb{R}^n e $\varphi : U \rightarrow \varphi(U)$ è un diffeomorfismo).

Esempio 2.5. Sia $p = (p_0, \dots, p_n)$ un punto di S^n , e supponiamo per fissare le idee che $p_0 > 0$. Sia $U = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} : x_0 > 0\}$ e consideriamo l’applicazione

$$\varphi : U \rightarrow \mathbb{R}^{n+1}, \quad (x_0, \dots, x_n) \mapsto \left(\sum_{i=0}^n x_i^2 - 1, x_1, \dots, x_n \right) = (y_0, \dots, y_n).$$

Dato che $y_0 = x_0^2 + \sum_{i=1}^n x_i^2 - 1$, l’immagine

$$V = \varphi(U) = \left\{ (y_0, \dots, y_n) \in \mathbb{R}^{n+1} : y_0 > \sum_{i=1}^n y_i^2 - 1 \right\}$$

è un aperto in \mathbb{R}^{n+1} . Inoltre, φ è un diffeomorfismo, in quanto l’inversa è data da

$$\varphi^{-1} : V \rightarrow U, \quad (y_0, \dots, y_n) \mapsto \left(\sqrt{y_0 - \sum_{i=1}^n y_i^2 + 1}, y_1, \dots, y_n \right)$$

che è di classe C^∞ .

Visto che $\varphi(U \cap S^n) = V \cap \{y_0 = 0\}$, nelle coordinate y_0, \dots, y_n la sfera S^n viene identificata con l’iperpiano $\{y_0 = 0\}$; questo vuol dire che y_1, \dots, y_n sono coordinate locali in p per S^n . Dato che $(y_1, \dots, y_n) = (x_1, \dots, x_n)$, le coordinate locali si ottengono semplicemente proiettando S^n sull’iperpiano relativo a una coordinata non nulla del punto p in esame.

Definizione 2.6. Sia X una varietà differenziabile di dimensione n . Il sottoinsieme $Y \subseteq X$ si dice *sottovarietà di dimensione m* se per ogni punto $p \in Y$ esiste un sistema di coordinate locali x_1, \dots, x_n in U intorno aperto di p tali che $Y \cap U = \{x_{m+1} = \dots = x_n = 0\}$ (ovvero Y è localmente il luogo di zeri di $(n - m)$ coordinate locali).

Esempio 2.6. Una sottovarietà di un aperto di \mathbb{R}^n è un suo sottospazio vettoriale. Inoltre, l’Esempio 2.5 mostra che S^n è una sottovarietà di dimensione n di \mathbb{R}^{n+1} (corrispondente all’annullarsi della coordinata y_0).

Esercizio 2.4. Dimostrare che se Y è una sottovarietà di X , allora Y è localmente chiuso:

$$\forall p \in Y \exists U \text{ intorno aperto di } p : Y \cap U \text{ è chiuso in } U.$$

Suggerimento: dimostrare che un sottospazio topologico $Y \subseteq X$ è localmente chiuso se, e solo se, Y è aperto in \overline{Y} e Y è intersezione di un chiuso e di un aperto.

Esercizio 2.5. Una sottovarietà è una varietà differenziabile.

Concludiamo questa sottosezione con la definizione, molto naturale, di prodotto fra varietà differenziabili.

Definizione 2.7. Siano X una varietà differenziabile definita dall'atlante $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ di dimensione n e Y una varietà differenziabile definita dall'atlante $\mathfrak{B} = \{(Z_j, W_j, \psi_j)\}_{j \in J}$ di dimensione m . Il prodotto $X \times Y$ è la varietà differenziabile la cui struttura è definita dall'atlante

$$\mathfrak{A} \times \mathfrak{B} = \{(U_i \times Z_j, V_i \times W_j, (\varphi_i, \psi_j))\}_{i \in I, j \in J}.$$

Osserviamo che se la dimensione di X come varietà differenziabile è n e quella di Y è m , allora $X \times Y$ è una varietà differenziabile di dimensione $n + m$.

2.4 Germi di funzioni C^∞ , spazio tangente, differenziale (di nuovo!)

Abbiamo visto che la struttura differenziabile di una varietà X è univocamente determinata dal fascio strutturale

$$C^\infty : \{\text{aperti di } X\} \rightarrow \{\text{anelli commutativi unitari}\}, \quad U \mapsto C^\infty(U).$$

Anche la definizione che abbiamo dato di funzione C^∞ fra varietà è una definizione che riguarda le proprietà locali delle varietà stesse; quindi il comportamento di una funzione in un punto è determinato solo da ciò che accade in un intorno di quel punto. Siamo quindi nella stessa situazione delle funzioni C^∞ fra aperti di \mathbb{R}^n ; possiamo dunque considerare equivalenti due funzioni che coincidano su un aperto fissato e definire i germi di funzioni C^∞ fra varietà.

Definizione 2.8. Sia X una varietà differenziabile di dimensione n , e sia $p \in X$. Consideriamo l'insieme

$$E(p) := \{(U, f) : U \subseteq X \text{ aperto}, p \in U, f \in C^\infty(U)\}.$$

Quozientiamo $E(p)$ rispetto alla relazione d'equivalenza \sim definita nella maniera seguente:

$$(U, f) \sim (V, g) \quad \text{se} \quad \exists W \subseteq U \cap V \text{ intorno aperto di } p : f|_W \equiv g|_W$$

(la verifica che \sim sia una relazione d'equivalenza è identica a quella già fatta nel caso di $X = \mathbb{R}^n$). Definiamo *anello dei germi di funzioni C^∞ in p* il quoziente $\mathcal{E}(p) := E(p) / \sim$.

Dato che per definizione $\mathcal{E}(p)$ dipende solo dalla struttura differenziabile in un intorno di p , abbiamo che se x_1, \dots, x_n è un sistema di coordinate locali in p allora $\mathcal{E}(p)$ è isomorfo all'anello dei germi di funzioni C^∞ (secondo la vecchia definizione) nelle variabili x_1, \dots, x_n . Inoltre, il seguente diagramma di anelli commutativi unitari

$$\begin{array}{ccc} & C^\infty(U) & \\ & \swarrow \quad \searrow & \\ \mathcal{E}(p) & \xrightarrow{ev_p} & \mathbb{R} \end{array} \qquad \begin{array}{ccc} & f & \\ & \swarrow \quad \searrow & \\ [U, f] & \xrightarrow{\quad} & f(p) \end{array}$$

è commutativo.

Definizione 2.9. Sia X una varietà differenziabile di dimensione n , e sia $p \in X$. Si definisce *spazio tangente a X in p* lo spazio

$$T_p X := \text{Der}(\mathcal{E}(p), \mathbb{R}) = \{\delta : \mathcal{E}(p) \rightarrow \mathbb{R} : \delta \text{ è } \mathbb{R}\text{-lineare}, \delta(fg) = f(p)\delta(g) + g(p)\delta(f) \forall f, g \in \mathcal{E}(p)\}.$$

Sia ora $F : X \rightarrow Y$ una funzione C^∞ fra varietà; scegliamo un punto $p \in X$ e chiamiamo $q = F(p) \in Y$. Dato un aperto $V \subseteq Y$ contenente q , l'applicazione

$$F^* := \cdot \circ F : C^\infty(V) \rightarrow C^\infty(F^{-1}(V)), \quad f \mapsto f \circ F$$

è un morfismo di anelli; con un abuso di notazione, chiameremo F^* anche il morfismo di anelli

$$F^* : \mathcal{E}(q) \rightarrow \mathcal{E}(p), \quad [V, f] \mapsto [F^{-1}(V), f \circ F].$$

Esercizio 2.6. Verificare che $F^* : \mathcal{E}(q) \rightarrow \mathcal{E}(p)$ è ben definito e che soddisfa le seguenti proprietà:

1. $\text{id}_X^* = \text{id}_{\mathcal{E}(p)}$ (in questo caso $q = \text{id}_X(p) = p$);
2. $(F \circ G)^* = G^* \circ F^*$.

Dedurre che se F è un diffeomorfismo, allora F^* è un isomorfismo di anelli.

Definizione 2.10. Nelle notazioni precedenti, si definisce *differenziale di F in p* l'applicazione

$$DF_p := \cdot \circ F^* : T_p X \simeq \text{Der}(\mathcal{E}(p), \mathbb{R}) \rightarrow T_q Y \simeq \text{Der}(\mathcal{E}(q), \mathbb{R}), \quad \delta \mapsto \delta \circ F^*.$$

Vediamo a cosa corrispondono queste definizioni in coordinate locali. Siano quindi $p \in X$ e $x_1, \dots, x_n \in \mathcal{E}(p)$ un sistema di coordinate locali in un intorno U di p . Questo vuol dire, per definizione, che l'applicazione $\varphi : U \rightarrow \mathbb{R}^n$ di coordinate (x_1, \dots, x_n) è un diffeomorfismo fra U e un intorno di 0 in \mathbb{R}^n ; quindi il suo differenziale $D\varphi_p : T_p X \rightarrow T_0 \mathbb{R}^n$ è un isomorfismo. Sappiamo però che

$$T_0 \mathbb{R}^n \simeq \text{Der}(\mathcal{E}(0), \mathbb{R}) \simeq \left\{ \sum_{i=1}^n v_i \frac{\partial}{\partial x_i}, (v_1, \dots, v_n) \in \mathbb{R}^n \right\} \simeq \mathbb{R}^n;$$

perciò la scelta di un sistema di coordinate locali x_1, \dots, x_n in p determina la scelta di una *base* di $T_p X$ come \mathbb{R} -spazio vettoriale: questa base è proprio $\left\{ \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right\}$.

Inoltre, le funzioni x_1, \dots, x_n generano un ideale massimale in $\mathcal{E}(p)$; infatti, ci basta vedere la varietà localmente come un intorno di 0 in \mathbb{R}^n e utilizzare il Lemma 1.4 per ottenere che, se f è un germe di funzioni C^∞ in un intorno U di p , allora esistono delle funzioni $f_1, \dots, f_n \in \mathcal{E}(p)$ tali che

$$f(x) = f(p) + \sum_{i=1}^n x_i f_i(x) \quad \forall x \in U. \quad (7)$$

Possiamo definire la derivata di una funzione $g \in \mathcal{E}(p)$ rispetto alla coordinata i -esima sulla varietà come

$$\frac{\partial g}{\partial x_i}(p) := D\varphi_p^{-1} \left(\frac{\partial}{\partial x_i} \right) (g)(p) = \frac{\partial}{\partial x_i} (g \circ \varphi^{-1})(p) = \frac{\partial}{\partial x_i} (\varphi^{-1})^* (g)(p)$$

dove $\frac{\partial}{\partial x_i}$ è la derivata parziale usuale in \mathbb{R}^n in tutti i membri tranne nel primo. Si ha che se deriviamo le coordinate locali

$$\frac{\partial x_j}{\partial x_i} = \frac{\partial}{\partial x_i} (x_j \circ \varphi^{-1}) = \delta_i^j$$

perché $x_j \circ \varphi^{-1}$ è la proiezione (in \mathbb{R}^n) sulla j -esima coordinata.

Applicando la derivazione $\frac{\partial}{\partial x_i}$ (della varietà) alla (7), e ricordando che le coordinate locali mandano p in 0 , otteniamo

$$\frac{\partial f}{\partial x_i}(p) = 0 + \sum_{j=1}^n \frac{\partial x_j}{\partial x_i}(p) f_j(p) + \sum_{j=1}^n x_j(p) \frac{\partial f_j}{\partial x_i}(p) = f_i(p)$$

così come nel caso della formula di Taylor su \mathbb{R}^n .

Osservazione 2.3. Come cambia l'operatore di derivata parziale quando cambiamo sistema di coordinate? Supponiamo di avere delle nuove coordinate y_1, \dots, y_n in un intorno U di p . Dato che $\left\{ \frac{\partial}{\partial x_j} \right\}_{j=1, \dots, n}$ è una base per lo spazio tangente $T_p X$, esistono dei coefficienti a_{ij} tali che

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^n a_{ij} \frac{\partial}{\partial x_j}, \quad \forall i = 1, \dots, n.$$

Applicando queste derivazioni alle vecchie coordinate locali, otteniamo

$$\frac{\partial x_h}{\partial y_i} = \sum_{j=1}^n a_{ij} \frac{\partial x_h}{\partial x_j} = a_{ih} \quad \forall i = 1, \dots, n$$

e quindi otteniamo la formula

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^n \frac{\partial x_j}{\partial y_i} \frac{\partial}{\partial x_j}, \quad \forall i = 1, \dots, n \quad (8)$$

che ricorda la formula di derivazione di funzione composta (ma senza la funzione!).

La nozione di differenziale ci aiuta a riconoscere le sottovarietà. Difatti, abbiamo definito una sottovarietà come luogo di zeri di certe coordinate locali. Talvolta è però più facile identificarne una in una varietà X come controimmagine mediante una funzione $F: X \rightarrow Z$ di classe C^∞ di uno spazio che sappiamo già essere una sottovarietà di Z . Il caso più semplice è quando $Y = F^{-1}(q)$ è la controimmagine di un qualche punto $q \in Z$; vorremo che Y fosse una sottovarietà chiusa di X . Il prossimo risultato ci dice che questo è sicuramente vero se il differenziale di F è suriettivo.

Lemma 2.2. *Siano X una varietà differenziabile di dimensione n , Z una varietà differenziabile di dimensione m e $F: X \rightarrow Z$ una funzione C^∞ . Dato $q \in Z$, chiamiamo $Y = F^{-1}(q)$ la sua controimmagine. Se per ogni $p \in Y$ il differenziale $DF_p: T_p X \rightarrow T_q Z$ è suriettivo, allora Y è una sottovarietà chiusa di X di dimensione $n - m$.*

Osservazione 2.4. Se $n < m$ il differenziale di F non può mai essere suriettivo per questione di dimensioni.

Dimostrazione. Sia $p \in Y$ e scegliamo una carta locale (U, V, φ) contenente p ; analogamente, prendiamo una carta (S, T, ψ) intorno a q . A meno di diffeomorfismi in arrivo, possiamo supporre che V sia un intorno dell'origine in \mathbb{R}^n e T sia un intorno dell'origine in \mathbb{R}^m , e che inoltre si abbia $\varphi(p) = 0$, $\psi(q) = 0$. L'applicazione F induce un'applicazione $\hat{F} = \psi \circ F \circ \varphi^{-1}: V \subseteq \mathbb{R}^n \rightarrow T \subseteq \mathbb{R}^m$ di classe C^∞ ; ovviamente DF_p è suriettivo se e solo se $D\hat{F}_0$ è suriettivo. Per il teorema delle funzioni implicite, esiste un diffeomorfismo $h: V \rightarrow W \subseteq \mathbb{R}^n$ tale che $h \circ \varphi$ è ancora una carta locale e

$$\hat{F} \circ h^{-1}(z_1, \dots, z_n) = (z_1, \dots, z_m).$$

Quindi $\varphi(Y)$ è la controimmagine secondo $\hat{F} \circ h^{-1}$ di 0 , e quindi è un sottospazio vettoriale di \mathbb{R}^n ; conseguentemente Y è una sottovarietà. In più, se x_1, \dots, x_n sono le coordinate locali in p date dalle coordinate dell'applicazione $h \circ \varphi$, allora $Y = \{x_1 = \dots = x_m = 0\}$, e quindi ha dimensione $n - m$. \square

Esempio 2.7. Sia $F: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ data da $F(x_0, \dots, x_n) = \sum_{i=0}^n x_i^2$. La sua matrice jacobiana, che in questo caso si riduce a un vettore, è $\mathcal{J}F = (2x_0, \dots, 2x_n)$, e quindi il differenziale è suriettivo in ogni $x \neq 0$. Riotteniamo quindi il fatto che $S^n = F^{-1}(1)$ è una sottovarietà chiusa di \mathbb{R}^{n+1} di dimensione $(n+1) - 1 = n$.

Esempio 2.8. Sappiamo dall'Esercizio 1.1 che il differenziale dell'applicazione determinante

$$\det: \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow \mathbb{R}$$

è suriettivo quando è calcolato in una matrice di rango almeno $n-1$. Conseguentemente, $SL(n, \mathbb{R}) = \det^{-1}(1)$ è una sottovarietà chiusa di $\mathfrak{M}_{n,n}(\mathbb{R}) \simeq \mathbb{R}^{n^2}$ di dimensione $n^2 - 1$.

Esempio 2.9. Sia S lo spazio delle matrici simmetriche $n \times n$ e $F: \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow S$, $F(A) = AA^T$. Se $E \in O(n)$, abbiamo visto nell'Esercizio 1.2 che DF_E è suriettivo; quindi $O(n) = F^{-1}(I)$ è una sottovarietà chiusa di dimensione $n^2 - \binom{n}{2} = \frac{n(n-1)}{2}$.

Consideriamo ora una varietà X di dimensione n e una sua sottovarietà di dimensione m . Il morfismo di inclusione $i: Y \hookrightarrow X$ induce un morfismo di inclusione $Di_p: T_p Y \hookrightarrow T_p X$. Inoltre, se p è un punto di Y , per definizione di sottovarietà esistono delle coordinate locali x_1, \dots, x_n in p (visto però come punto di X) rispetto alle quali $Y = \{x_{m+1} = \dots = x_n = 0\}$; dunque x_1, \dots, x_m sono coordinate locali in p (visto stavolta come punto di Y). Ne deduciamo che $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1, \dots, m}$ è una base per $T_p Y$, che risulta dunque essere un sottospazio vettoriale di $T_p X$ (quello generato dai primi m vettori della base).

Esempio 2.10. Sia $F: X \rightarrow \mathbb{R}^{n-m}$ una funzione C^∞ e $Y = F^{-1}(0)$ (supponiamo ovviamente che 0 sia un valore effettivamente assunto da F !); come sappiamo Y è una sottovarietà di dimensione m se $DF_p: T_p X \rightarrow T_0 \mathbb{R}^{n-m} \simeq \mathbb{R}^{n-m}$ è suriettivo per ogni $p \in Y$. Del resto, dato che per definizione $F|_Y \equiv 0$, il suo differenziale è banale nei punti di Y , ovvero $DF_p|_{T_p Y} \equiv 0$. Dato che il nucleo del differenziale di F in p ha dimensione m così come $T_p Y$, si ha che

$$T_p Y = \ker DF_p.$$

Usiamo questa osservazione per individuare lo spazio tangente alla sfera n -dimensionale $T_p S^n$. Se definiamo

$$F: \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad (x_0, \dots, x_n) \mapsto \sum_{i=0}^n x_i^2 - 1$$

allora F è una funzione C^∞ e $S^n = F^{-1}(0)$. Se $p \in S^n$, sappiamo che un generico vettore tangente a p (visto come punto di \mathbb{R}^{n+1}) è della forma $\sum_{i=0}^n v_i \frac{\partial}{\partial x_i}$; allora

$$DF_p \left(\sum_{i=0}^n v_i \frac{\partial}{\partial x_i} \right) = \sum_{i=0}^n v_i \frac{\partial F}{\partial x_i}(p) = 2 \sum_{i=0}^n v_i p_i = 2 \langle v, p \rangle.$$

Ne deduciamo che

$$T_p S^n = \ker DF_p = \{v \in \mathbb{R}^{n+1} : v \perp p\}$$

che è esattamente l'idea intuitiva che abbiamo di spazio tangente alla sfera (l'iperpiano ortogonale al raggio vettore del punto). Più in generale, se $F: \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione di classe C^∞ che assume il valore 0 e tale che il differenziale di F è suriettivo in tutti i punti di $F^{-1}(0)$, allora

$$T_p F^{-1}(0) = (\nabla F)^\perp.$$

3 Il teorema di Sard

Abbiamo visto che la suriettività del differenziale di una funzione gioca spesso un ruolo importante per stabilire se determinati sottoinsiemi siano o meno sottovarietà. In questa sezione ci proponiamo di capire “in quanti punti” il differenziale può non essere suriettivo.

3.1 Prerequisiti di Analisi

Definizione 3.1. Siano X e Y due varietà differenziabili, e $F: X \rightarrow Y$ un'applicazione differenziabile. Un punto $p \in X$ si dice *punto critico* per F se $DF_p: T_p X \rightarrow T_{F(p)} Y$ non è suriettivo; altrimenti, p si dice *punto regolare*.

Un punto $q \in Y$ si dice *valore critico* per F se è immagine di un punto critico; altrimenti, q si dice *valore regolare*.

Osservazione 3.1. Se la dimensione di X è minore di quella di Y , allora tutti i punti di X sono critici.

Possiamo quindi rinunciare al Lemma 2.2 affermando che se $F: X \rightarrow Z$ è un'applicazione differenziabile fra varietà, q è un punto di Z e tutti i punti di $Y = F^{-1}(q)$ sono regolari, allora Y è una sottovarietà di X .

Quello che ci proponiamo di dimostrare in questa sezione è il seguente

Teorema 3.1 (di Sard). *Se $F: X \rightarrow Y$ è una funzione C^∞ fra varietà, allora l'immagine dell'insieme dei punti critici ha misura di Lebesgue nulla.*

Per farlo, abbiamo bisogno di alcuni “ingredienti” di Analisi, che ricordiamo qui di seguito.

Teorema 3.2 (del valor medio). *Sia $f: \overline{B(p, r)} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione di classe C^∞ , e supponiamo che tutte le derivate di ordine minore di m si annullino in p :*

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(p) = 0, \quad k = 1, \dots, m-1, \quad i_j \in \{1, \dots, n\}.$$

Allora

$$|f(x) - f(p)| \leq c \|x - p\|^m, \quad \text{dove} \quad c = \max_{x \in \overline{B(p, r)}} \left\| \frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}(x) \right\|.$$

Definizione 3.2. Un cubo di centro $p \in \mathbb{R}^n$ e lato $2a \in \mathbb{R}^+$ è un insieme della forma³

$$W = \{(x_1, \dots, x_n) \in \mathbb{R}^n : |x_i - p_i| \leq a, i = 1, \dots, n\}.$$

Si definisce *misura (di Lebesgue)* di un cubo $W \subseteq \mathbb{R}^n$ di lato $2a$ il valore $m(W) = (2a)^n$.

Un insieme $C \subseteq \mathbb{R}^n$ si dice avere *misura (di Lebesgue) nulla* ($m(C) = 0$) se per ogni $\varepsilon > 0$ esistono dei cubi $W_i, i \in \mathbb{N}$, tali che

$$C \subseteq \bigcup_{i \in \mathbb{N}} W_i \quad \text{e} \quad \sum_{i \in \mathbb{N}} m(W_i) < \varepsilon.$$

Nella definizione di insieme a misura nulla, non è restrittivo supporre che i cubi siano di lato arbitrariamente piccolo: infatti, ogni cubo di lato a è unione di k^n cubi di lato $\frac{a}{k}$. Inoltre, è immediato verificare che un sottoinsieme di un insieme a misura nulla ha ancora misura nulla, e che l'unione numerabile di insiemi a misura nulla è ancora a misura nulla.

Lemma 3.1. *Siano $U \subseteq \mathbb{R}^n$ un aperto e $F: U \rightarrow \mathbb{R}^n$ una funzione di classe C^∞ . Se $C \subseteq U$ ha misura nulla, allora anche la sua immagine $f(C)$ ha misura nulla.*

³Osserviamo che, per come è definito, un cubo è un sottoinsieme chiuso e limitato di \mathbb{R}^n , e quindi è compatto.

Dimostrazione. Possiamo ricoprire U con un'unione numerabile di cubi K_i , $i \in \mathbb{N}$; è sufficiente quindi provare che $m(F(C \cap K_i)) = 0$ per ogni $i \in \mathbb{N}$. Dato che K_i è compatto, per il teorema del valor medio esiste una costante c_i (che sarà il massimo del modulo delle derivate prime di f in K_i) tale che

$$\|F(x) - F(y)\| \leq c_i \|x - y\|.$$

Questo ci dice che se W è un cubo di centro p e lato a contenuto in K_i , allora $F(W)$ è contenuto in un cubo di centro $F(p)$ e lato $c'_i a$, dove $c'_i = c_i \sqrt{n}$. Sia ora $\varepsilon > 0$; per l'ipotesi che C ha misura nulla possiamo supporre che $C \subseteq K_i$ è contenuto nell'unione numerabile di cubi W_j , $j \in \mathbb{N}$, ognuno di lato a_j , e tali che $\sum_{j \in \mathbb{N}} a_j^n < \varepsilon$. Ma allora

$$F(C \cap K_i) \subseteq \bigcup_{j \in \mathbb{N}} F(W_j) \quad \text{e} \quad \sum_{j \in \mathbb{N}} m(F(W_j)) \leq \sum_{j \in \mathbb{N}} (c'_i)^n \sum_{j \in \mathbb{N}} a_j^n < (c'_i)^n \varepsilon$$

e quindi $F(C \cap K_i)$ ha misura nulla. □

Corollario 3.1. *Gli insiemi a misura nulla sono stabili per diffeomorfismo.*

Possiamo quindi dire quando un sottoinsieme di una varietà X è a misura nulla (finora abbiamo solo parlato di sottoinsiemi di \mathbb{R}^n).

Definizione 3.3. Un sottoinsieme C della varietà differenziabile X si dice avere *misura (di Lebesgue) nulla* se per ogni carta locale (U, V, φ) si ha che $\varphi(U \cap C)$ ha misura nulla.

Osserviamo che ci basta verificare questa proprietà per un atlante numerabile: infatti, dal fatto che le varietà differenziabili sono a base numerabile discende che ogni atlante ammette un sottoatlante numerabile.

L'ultimo risultato di Analisi che ci serve per dimostrare il teorema di Sard è il seguente.

Teorema 3.3 (di Fubini). *Sia $C \subseteq \mathbb{R}^n$ un chiuso. Se per ogni $\alpha \in \mathbb{R}$ l'insieme $C \cap \{x_1 = \alpha\}$ ha misura nulla (in \mathbb{R}^{n-1}) allora C ha misura nulla (in \mathbb{R}^n).*

3.2 Dimostrazione del teorema di Sard

Dimostrazione del Teorema 3.1. Osserviamo innanzitutto che se la dimensione di X è minore di quella di Y , allora $F(X)$ ha misura nulla e la tesi è banalmente verificata; è anzi sufficiente che F sia di classe C^1 .

Sia $Y = \bigcup_{i \in \mathbb{N}} U_i$ un ricoprimento numerabile di carte di Y , e $F^{-1}(U_i) = \bigcup_{j \in \mathbb{N}} U_{ij}$ un ricoprimento numerabile fatto con carte di X . Se D è l'insieme dei punti critici di F , allora $F(D) = \bigcup_{i,j} F(D \cap U_{ij})$.

Se inoltre $(U_{ij}, V_{ij}, \varphi_{ij})$ è una carta di X (con V_{ij} aperto in \mathbb{R}^n) corrispondente alla carta (U_i, V_i, φ_i) di Y (con V_i aperto in \mathbb{R}^m), allora, poiché U_{ij} è diffeomorfo a V_{ij} mediante φ_{ij} e U_i è diffeomorfo a V_i mediante φ_i , ci basta verificare che l'applicazione $\varphi_i \circ F \circ \varphi_{ij}^{-1}: V_{ij} \subseteq \mathbb{R}^n \rightarrow V_i \subseteq \mathbb{R}^m$ manda $\varphi_{ij}(D) \cap V_{ij}$ in un insieme di misura nulla. In altre parole, ci siamo ricondotti a una funzione da un aperto di \mathbb{R}^n a un aperto di \mathbb{R}^m ; vogliamo mostrare che per questa applicazione vale il teorema di Sard. Continueremo tuttavia a chiamare questa applicazione $F: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (il lettore dovrà tenere a mente che in realtà si tratta di $\varphi_i \circ F \circ \varphi_{ij}^{-1}$) e l'insieme dei suoi punti critici D (che di fatto è $\varphi(D) \cap V_{ij}$).

Osserviamo che se $m = 0$ la tesi è banalmente verificata, in quanto il differenziale di una funzione a valori in \mathbb{R}^0 (che è un punto!) è suriettivo in ogni $p \in U$. Supponiamo quindi $m > 0$ e dimostriamo il teorema per induzione su n . Se $n = 0$ allora U è un punto; quindi $F(U)$ è un punto in \mathbb{R}^m e in quanto tale ha misura nulla. Supponiamo ora il teorema vero per tutte le applicazioni $G: V \subseteq \mathbb{R}^{n-1} \rightarrow \mathbb{R}^m$ di classe C^∞ e dimostriamo che il teorema è vero anche per $F: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Siano (F_1, \dots, F_m) le coordinate di F e sia D_i , $i \in \mathbb{N}$, l'insieme dei punti dove tutte le derivate parziali i -esime delle F_j si annullano contemporaneamente; ovviamente si ha che $D_{i+1} \subseteq D_i$ e che i D_i sono tutti chiusi (in quanto luogo di zeri di applicazioni continue). Notiamo che l'insieme D dei punti critici contiene D_1 e che è anch'esso un chiuso, in quanto è l'insieme dei punti dove si annulla il determinante della matrice jacobiana di F . Ci proponiamo di dimostrare che:

1. $F(D \setminus D_1)$ ha misura nulla;
2. $F(D_i \setminus D_{i+1})$ ha misura nulla per ogni $i \in \mathbb{N}$;
3. per i sufficientemente grande (e vedremo anche *quanto* grande), $F(D_i)$ ha misura nulla.

Il terzo punto è utile per comprendere quale regolarità serve richiedere alla funzione F per applicare il teorema di Sard (ovvero è utile quando non consideriamo funzioni C^∞ ma solo di classe C^r , $r \in \mathbb{N}$). Di fatto, i primi due (nelle ipotesi che F sia di classe C^∞) sono sufficienti a dimostrare la nostra tesi originaria: infatti possiamo scrivere che $F(D)$ è l'unione di $F(D \setminus D_1)$ e di tutti i $F(D_i \setminus D_{i+1})$, al variare di $i \in \mathbb{N}$.

Dimostriamo quindi che $F(D \setminus D_1)$ ha misura nulla. Ci basta dimostrare che comunque si scelga $p \in D \setminus D_1$ esiste un intorno V di p tale che $F((D \setminus D_1) \cap V)$ ha misura nulla; considereremo poi un ricoprimento numerabile di $F(D \setminus D_1)$ fatto da intorni di questo tipo. Sia quindi $p \in D \setminus D_1$; a meno di permutazioni (ovvero di diffeomorfismi) possiamo supporre che il differenziale DF_p non sia suriettivo ma che $\frac{\partial F_1}{\partial x_1}(p) \neq 0$. Definiamo

$$h: U \rightarrow \mathbb{R}^n, \quad x = (x_1, \dots, x_n) \mapsto (F_1(x), x_2, \dots, x_n).$$

Dato che la matrice che rappresenta Dh_p è la matrice

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & * & \cdots & * \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

il differenziale di h in p è invertibile; per il teorema di invertibilità locale esiste un intorno V di p tale che $h: V \rightarrow h(V)$ sia un diffeomorfismo. Se $z = (z_1, \dots, z_n)$ è un punto in $h(V)$, allora è chiaro che

$$F \circ h^{-1}(z_1, \dots, z_n) = (z_1, G_2(z), \dots, G_m(z))$$

dove G_2, \dots, G_m sono le ultime $m-1$ coordinate di $F \circ h^{-1}$. Ma allora il differenziale di $F \circ h^{-1}$ nel punto z è rappresentato dalla matrice

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & \frac{\partial G_2}{\partial z_2}(z) & \cdots & \frac{\partial G_2}{\partial z_n}(z) \\ \vdots & \vdots & \ddots & \vdots \\ * & \frac{\partial G_m}{\partial z_2}(z) & \cdots & \frac{\partial G_m}{\partial z_n}(z) \end{pmatrix}.$$

Fissato quindi il valore della prima coordinata ($z_1 = \alpha$), i valori critici di $F \circ h^{-1}$ sono quelli dell'applicazione $z \mapsto G(\alpha, z_2, \dots, z_n) : (\{\alpha\} \times \mathbb{R}^{n-1}) \cap h(V) \rightarrow \mathbb{R}^{m-1}$; per l'ipotesi induttiva, questi formano un insieme di misura nulla. Sappiamo quindi che l'intersezione fra l'insieme dei punti critici di $F \circ h^{-1}$ e l'iperpiano $\{z_1 = \alpha\}$ ha misura nulla in \mathbb{R}^{n-1} per ogni $\alpha \in \mathbb{R}$; il teorema di Fubini ci permette di concludere che l'insieme dei punti critici di $F \circ h^{-1}$ ha misura nulla in \mathbb{R}^n . Poiché sia i

punti critici che gli insiemi di misura nulla sono stabili per diffeomorfismo, anche $F((D \setminus D_1) \cap V)$ ha misura nulla.

La dimostrazione del fatto che $F(D_i \setminus D_{i+1})$ ha misura nulla per ogni $i \in \mathbb{N}$ è del tutto analoga. Scegliamo un punto $p \in D_i \setminus D_{i+1}$ e dimostriamo che per qualche intorno V di p vale che $F((D_i \setminus D_{i+1}) \cap V)$ ha misura nulla. Sappiamo che, a meno di permutazioni, esiste una derivata parziale i -esima di una F_j che si annulla in p ma la cui derivata rispetto alla variabile x_1 è diversa da zero in p : denotiamo questa derivata parziale con w . Definiamo come prima

$$h: U \rightarrow \mathbb{R}^n, \quad x = (x_1, \dots, x_n) \mapsto (w(x), x_2, \dots, x_n).$$

Ancora una volta, il differenziale di h è invertibile in p e quindi esiste un intorno V di p tale che $h: V \rightarrow h(V)$ sia un diffeomorfismo. Dato che $w \Big|_{D_i} \equiv 0$, abbiamo che l'immagine mediante $F \circ h^{-1}$

di $h((D_i \setminus D_{i+1}) \cap V)$ è contenuta in $\{0\} \times \mathbb{R}^{n-1}$: definiamo dunque $G := F \circ h^{-1} \Big|_{\{z_1=0\}}$. Dato che in $(D_i \setminus D_{i+1}) \cap V$ si annullano le derivate di $F \circ h^{-1}$, a maggior ragione si annullano quelle di G ; quindi i valori critici di $F \circ h^{-1}$ sono contenuti nell'insieme dei valori critici di G che per l'ipotesi induttiva ha misura nulla. Concludiamo che anche l'insieme dei punti critici di F ha misura nulla.

Per quanto riguarda il terzo punto, sia p un punto in D_i , con $i \gg 0$. Dal teorema del valor medio sappiamo che esistono un raggio $r > 0$ e una costante $c > 0$ tali che per ogni $x \in \overline{B(p, r)} \cap D_i$ e per ogni $y \in \overline{B(p, r)}$ vale

$$\|F(x) - F(y)\| \leq c \|x - y\|^{i+1}.$$

Se $W \subset \overline{B(p, r)} \cap D_i$ è un cubo di lato a allora $F(W)$ è contenuto in un cubo di lato $c' a^{i+1}$, per una qualche costante positiva c' ; la misura di W è a^n e $F(W)$ è contenuto in un cubo di misura $(c')^m a^{(i+1)m}$. Quindi se D_i è contenuto in un'unione di cubi di misura più piccola di $\varepsilon > 0$, allora anche la sua immagine è contenuta in un'unione di cubi di misura piccola se $(i+1)m \geq n$, ovvero se $i \geq \frac{n}{m} - 1$: in queste ipotesi abbiamo quindi che $F(D_i)$ ha misura nulla. \square

Corollario 3.2 (Teorema di Brown). *Se $F: X \rightarrow Y$ è una funzione C^∞ fra varietà, allora l'insieme dei valori regolari di F è denso in Y .*

Dimostrazione. Dal teorema di Sard sappiamo che l'insieme dei valori critici ha misura di Lebesgue nulla, quindi in particolare non può contenere aperti. Il suo complementare (che è ovviamente l'insieme dei valori regolari) è dunque denso in Y . \square

Questi teoremi ci dicono quindi che il differenziale di una funzione C^∞ fra varietà, a patto che la dimensione dello spazio di arrivo sia minore o uguale di quella dello spazio di partenza, è "praticamente sempre" suriettivo!

3.3 Applicazioni del teorema di Sard

Esempio 3.1. Il teorema di Sard permette di verificare senza bisogno dell'Esercizio 1.1 che il gruppo speciale lineare $SL(n, \mathbb{R})$ è una sottovarietà dello spazio delle matrici $n \times n$. Consideriamo nuovamente l'applicazione $\det: \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow \mathbb{R}$; per il teorema di Sard esiste un valore regolare $s \in \mathbb{R} \setminus \{0\}$ (ogni aperto ne contiene almeno uno). Abbiamo quindi che $\det^{-1}(\{s\})$ è una sottovarietà chiusa di $\mathfrak{M}_{n,n}(\mathbb{R})$; del resto

$$\det^{-1}(\{s\}) = SL(n, \mathbb{R}) \cdot \begin{pmatrix} s & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Questa matrice è invertibile, e in quanto tale è un diffeomorfismo di $\mathfrak{M}_{n,n}(\mathbb{R})$ in sé; quindi $SL(n, \mathbb{R})$ è diffeomorfo a $\det^{-1}(\{s\})$.

Esercizio 3.1. Dare una dimostrazione alternativa del fatto che il gruppo delle matrici ortogonali $O(n)$ è una sottovarietà di $\mathfrak{M}_{n,n}(\mathbb{R})$.

Suggerimento: le matrici simmetriche definite positive formano un aperto di $\mathfrak{M}_{n,n}(\mathbb{R})$.

Esercizio 3.2. Sia $X \subseteq \mathbb{R}^N$ una sottovarietà di dimensione n . Dimostrare che se $N > 2n + 1$ allora esiste un iperpiano $H \subseteq \mathbb{R}^N$ tale che la proiezione ortogonale su H ristretta a X , $\pi: X \rightarrow \pi(X) \subseteq H$, è iniettiva.

Soluzione. La nostra strategia per individuare tale iperpiano sarà quella di trovare un vettore $\alpha \in \mathbb{R}^N$ tale che ogni retta che ha la sua stessa direzione interseca X in al più un punto: a quel punto sarà sufficiente porre $H = \alpha^\perp$.

Sappiamo che $X \times X$ è una varietà di dimensione $2n$. Sia⁴

$$Y = \{(x_1, x_2) \in X \times X : x_1 \neq x_2\} = (X \times X) \setminus \Delta_X.$$

Dato che le varietà differenziabili sono spazi di Hausdorff, la diagonale è chiusa nel prodotto, quindi Y è un aperto di $X \times X$; in quanto tale, risulta a sua volta una varietà di dimensione $2n$.

Consideriamo

$$F: Y \rightarrow S^{N-1}, \quad (x_1, x_2) \mapsto \frac{x_1 - x_2}{\|x_1 - x_2\|};$$

L'applicazione F è di classe C^∞ , e poiché $2n < N - 1$ il differenziale $DF_{(x_1, x_2)}$ non può essere suriettivo, per questioni di dimensione, in nessun punto di Y . Conseguentemente, tutti i punti di Y sono critici; per il teorema di Brown $S^{N-1} \setminus F(Y)$ è denso (e in particolare è non vuoto).

Prendiamo quindi $\alpha \in S^{N-1} \setminus F(Y)$ e poniamo $H = \alpha^\perp$; la proiezione ortogonale su H

$$\pi: X \rightarrow H, \quad \pi(x) = x - \langle x, \alpha \rangle \alpha$$

è allora iniettiva. Se così non fosse, esisterebbero infatti due punti distinti x_1 e x_2 in X che, proiettati su H , vanno a finire nello stesso punto, e quindi x_1 e x_2 si trovano sulla stessa retta ortogonale a H . Ma allora $x_1 - x_2$ è un multiplo di α : possiamo supporre, a meno di scambiarli fra di loro, che $x_1 - x_2 = \lambda \alpha$ con $\lambda \in \mathbb{R}^+$. Ma allora

$$\frac{x_1 - x_2}{\|x_1 - x_2\|} = \alpha$$

(perché $\alpha \in S^{N-1}$ ha norma unitaria, quindi deve essere $\lambda = \|x_1 - x_2\|$) che è assurdo, in quanto stiamo supponendo che $\alpha \notin F(Y)$. \diamond

⁴La *diagonale* di un prodotto $X \times X$ è $\Delta_X := \{(x, x) \in X \times X : x \in X\}$.

4 Trasversalità

Consideriamo le seguenti sottovarietà X e Y di dimensione 1 in \mathbb{R}^2 :

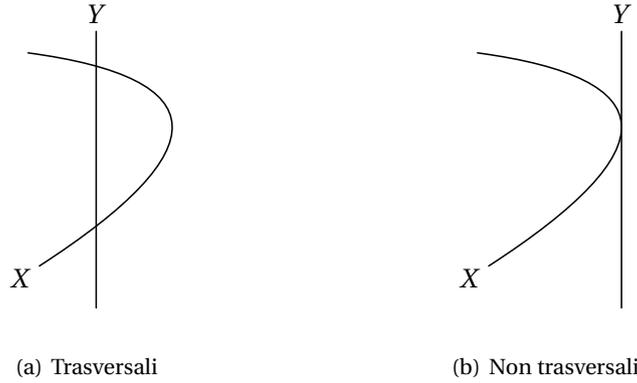


Figura 2: Trasversalità

Intuitivamente, nel primo caso (Figura 2(a)) i due sottospazi sono trasversali, nel senso che se approssimiamo le due sottovarietà nei punti di intersezione con le loro tangenti, esse non coincidono; così non è nel caso della Figura 2(b). È inoltre chiaro che se nella prima “perturbiamo” un po’ le posizioni di X e di Y , il loro essere trasversali non varia. Vediamo di rendere più formali queste idee intuitive.

Definizione 4.1. Siano $F: X \rightarrow Y$ una funzione di classe C^∞ fra varietà e $Z \subseteq Y$ una sottovarietà. Diremo che l’applicazione F è *trasversale* a Z se per ogni punto $p \in F^{-1}(Z)$ si ha che

$$DF_p(T_p X) + T_{F(p)} Z = T_{F(p)} Y$$

ovvero se l’immagine secondo il differenziale di F in p dello spazio tangente a X in p e lo spazio tangente a Z in $F(p)$ generano tutto lo spazio tangente a Y in $F(p)$. Scriveremo $F \pitchfork Z$ per indicare che F è trasversale a Z .

In Figura 2(a), possiamo immaginare che F sia una funzione che “incurva” un intervallo della retta reale e lo manda nel sottoinsieme X del piano; in questo caso F è trasversale a Y perché le rette tangenti a X e a Y sono sempre distinte (e quindi i vettori che individuano le loro direzioni sono linearmente indipendenti, generando così tutto il piano). Nel caso della Figura 2(b), invece, i vettori che rappresentano le tangenti a X e Y sono collineari, e quindi F non è trasversale a Y .

Osservazione 4.1. Possiamo dare due definizioni equivalenti di trasversalità (il lettore è invitato a verificare effettivamente l’equivalenza di queste definizioni con quella appena fornita). Consideriamo nuovamente una funzione $F: X \rightarrow Y$ di classe C^∞ fra varietà e $Z \subseteq Y$ una sottovarietà.

1. F è trasversale a Z se per ogni punto $p \in F^{-1}(Z)$ l’applicazione

$$T_p X \xrightarrow{DF_p} T_{F(p)} Y \xrightarrow{\pi} T_{F(p)} Y / T_{F(p)} Z \quad (9)$$

è suriettiva (l’ultimo spazio è il quoziente di spazi vettoriali fra $T_{F(p)} Y$ e $T_{F(p)} Z$, e π è la proiezione quoziente);

2. F è trasversale a Z se per ogni punto $p \in F^{-1}(Z)$ l’insieme

$$H_p := \{v \in T_p X : DF_p(v) \in T_{F(p)} Z\}$$

(che è il nucleo dell’applicazione $\pi \circ DF_p$ come in (9)) è un sottospazio vettoriale di $T_p X$ di dimensione $\dim X - (\dim Y - \dim Z)$.

La condizione di trasversalità ci consente di generalizzare il Lemma 2.2; corrisponde alla suriettività del differenziale come condizione sufficiente affinché $F^{-1}(Z)$ sia una sottovarietà, che abbiamo dimostrato nel caso in cui Z è un punto.

Lemma 4.1. *Sia $F: X \rightarrow Y$ una funzione di classe C^∞ fra due varietà X e Y di dimensione rispettivamente n e m , e sia $Z \subseteq Y$ una sottovarietà di dimensione k . Se F è trasversale a Z , allora $F^{-1}(Z)$ è una sottovarietà di X di dimensione $n - m + k$.*

Dimostrazione. Siano $p \in F^{-1}(Z)$ e $q = F(p) \in Z$. Dato che Z è una sottovarietà di Y , esistono delle coordinate locali y_1, \dots, y_m in un intorno aperto V di q rispetto alle quali $Z = \{y_{k+1} = \dots = y_m = 0\}$; sia $\varphi: V \rightarrow \mathbb{R}^{m-k}$ l'applicazione di coordinate (y_{k+1}, \dots, y_m) . Siano inoltre U un intorno aperto di p e $G: U \rightarrow \mathbb{R}^{m-k}$ l'applicazione $G = \varphi \circ F$. Essendo y_{k+1}, \dots, y_m coordinate locali, il differenziale $D\varphi_q: T_q Y \rightarrow \mathbb{R}^{m-k}$ coincide con la proiezione a quoziente $\pi: T_q Y \rightarrow T_q Y / T_q Z$; infatti,

$$D\varphi_q \left(\frac{\partial}{\partial y_j} \right) = 0 \quad \forall j = k+1, \dots, m.$$

Concludiamo dunque che $D\varphi_q$ è suriettivo e $\ker D\varphi_q = T_q Z$.

La condizione $F \pitchfork Z$ in p è quindi equivalente al fatto che DG_p è suriettivo (in virtù della prima definizione equivalente di trasversalità). Ci basta ora osservare che $F^{-1}(Z) = G^{-1}(0)$ e applicare il Lemma 2.2. \square

Vediamo ora che la condizione di trasversalità è invariante per generiche "perturbazioni".

Teorema 4.1 (di trasversalità). *Siano $F: X \times S \rightarrow Y$ una funzione di classe C^∞ fra varietà e $Z \subseteq Y$ una sottovarietà. Se F è trasversale a Z , allora per quasi tutti gli $s \in S$ l'applicazione*

$$F_s: X \rightarrow Y, \quad x \mapsto F(x, s)$$

è trasversale a Z .

Dimostrazione. Chiamiamo $H = F^{-1}(Z)$; dal lemma precedente sappiamo che H è una sottovarietà di $X \times S$ di dimensione $\dim H = \dim X + \dim S + \dim Z - \dim Y$. Fissiamo $s \in S$; in virtù della seconda definizione equivalente di trasversalità, per mostrare che $F_s \pitchfork Z$ occorre dimostrare che per ogni punto $p \in X$ tale che $(p, s) \in H$ lo spazio

$$H_p := \{v \in T_p X : DF_{s,p}(v) \in T_{F_s(p)} Z\}$$

è un sottospazio vettoriale di $T_p X$ di dimensione $\dim X + \dim Z - \dim Y$.

Notiamo che

$$\begin{aligned} DF_{s,p}(v) &= \sum_i v_i \frac{\partial}{\partial x_i} (f(F_s(x))) = \\ &= \sum_i v_i \sum_j \frac{\partial f}{\partial y_j} (F_s(x)) \frac{\partial F}{\partial x_i} (x, s) + \sum_h 0 \sum_j \frac{\partial f}{\partial y_j} (F_s(x)) \frac{\partial F}{\partial s_h} (x, s) = \\ &= DF_p(v, 0), \end{aligned}$$

in quanto $T_{(p,s)}(X \times S) \simeq T_p X \oplus T_s S$ (il lettore verifichi quest'ultima osservazione) e quindi $DF_{(p,s)} = DF_{s,p} + DF_{p,s}$, dove F_p è la funzione F come funzione di $s \in S$. Scegliamo dunque un valore regolare $s \in S$ della proiezione sul secondo fattore $\pi_2: H \rightarrow S$; l'esistenza ci è garantita dal teorema di Sard (e questo specifica il "quasi tutti" presente nell'enunciato). Essendo s regolare, il differenziale $D\pi_{2,s}$ è suriettivo; perciò

$$\dim \{(v, 0) \in T_{(p,s)} H\} = \dim \ker D\pi_{2,s} = \dim H - \dim S = \dim X + \dim Z - \dim Y.$$

Ma dato che per definizione $H = F^{-1}(Z)$ si ha

$$T_{(p,s)}H = \{(v, t) \in T_pX \oplus T_sS : DF_{(p,s)}(v, t) \in T_{(p,s)}Z\}$$

e quindi

$$\{(v, 0) \in T_{(p,s)}H\} \simeq \{v \in T_pX : DF_{s,p}(v) \in T_{(p,s)}Z\} = H_p$$

ovvero $F_s \pitchfork Z$. □

Esercizio 4.1. Sia $f: X \rightarrow \mathbb{R}^n$ una funzione C^∞ e $Z \subseteq \mathbb{R}^n$ una sottovarietà. Dimostrare che per ogni $\varepsilon > 0$ esiste un vettore $v \in \mathbb{R}^n$ tale che $\|v\| < \varepsilon$ e la funzione *traslata* di f

$$f_v: X \rightarrow \mathbb{R}^n, \quad x \mapsto f(x) + v$$

è trasversale a Z .

5 Fibrato tangente

Mediante la definizione di spazio tangente, abbiamo “attaccato” ad ogni punto di una varietà differenziabile uno spazio vettoriale. Prendendoli tutti, otteniamo una nuova varietà.

Definizione 5.1. Sia M una varietà differenziabile. Si definisce *fibrato tangente* alla varietà l'unione disgiunta di tutti gli spazi tangenti:

$$TM := \bigsqcup_{p \in M} T_p M.$$

Osservazione 5.1. Risulta ben definita l'applicazione $\pi: TM \rightarrow M$, $v \in T_p M \mapsto p$.

Se $M \subseteq \mathbb{R}^N$ è una sottovarietà di dimensione n , sappiamo che per ogni $p \in M$ lo spazio tangente $T_p M$ è un sottospazio vettoriale n -dimensionale. Supponiamo che in un intorno del punto p la varietà M sia l'intersezione dei luoghi di zeri delle applicazioni $F_1, \dots, F_{N-n} \in C^\infty(\mathbb{R}^N)$ (le coordinate locali), e che il rango della matrice jacobiana, ovvero del differenziale, della funzione $F = (F_1, \dots, F_{N-n})$ sia massimo (pari a $N - n$) in p ; per quanto abbiamo visto, lo spazio tangente $T_p M$ è il nucleo di questa matrice. In questo caso

$$TM = \{(p, v) \in \mathbb{R}^N \times \mathbb{R}^N : p \in M, v \in T_p M\}.$$

Dato che localmente si ha $TM = M \times T_p M$, ci aspettiamo che TM sia una varietà differenziabile di dimensione $2n$; verifichiamo che effettivamente è così. Siano quindi $p \in M$ e $v \in \mathbb{R}^n$ tali che $\mathcal{J}F_p(v) = 0$; in un intorno di p , M è il luogo di zeri di F , quindi in un intorno di (p, v) il fibrato tangente è l'insieme

$$TM = \{(x, w) \in \mathbb{R}^N \times \mathbb{R}^N : F_1(x) = \dots = F_{N-n}(x) = 0, \mathcal{J}F_p(w) = 0\}.$$

Ci appare quindi naturale definire l'applicazione

$$\vartheta: \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^{N-n} \times \mathbb{R}^{N-n}, \quad (x, w) \mapsto (F_1(x), \dots, F_{N-n}(x), \mathcal{J}F_p(w))$$

di cui TM è (localmente) il luogo di zeri; ci resta quindi da dimostrare che $D\vartheta_{(p,v)}$ ha rango massimo. Questo è molto semplice da verificare, dato che la matrice jacobiana di ϑ è la matrice a blocchi

$$\mathcal{J}\vartheta_{(p,v)} = \begin{pmatrix} \mathcal{J}F_p & O \\ * & \mathcal{J}F_p \end{pmatrix}$$

che è di rango massimo per ipotesi. Ne deduciamo che TM ha dimensione $(N+N) - (N-n+N-n) = 2n$.

Osservazione 5.2. Il blocco (*) della matrice $\mathcal{J}\vartheta_{(p,v)}$, che rappresenta le derivate delle funzioni F_1, \dots, F_{N-n} rispetto alle coordinate w_1, \dots, w_N , dà informazioni su quanto la varietà sia “curva”: se M è una varietà lineare, allora (*) è la matrice nulla, perché sono lineari anche le funzioni F_i . Lo studio di questo blocco è oggetto della Geometria Riemanniana.

Esempio 5.1. Il gruppo $SO(3)$ delle rotazioni in \mathbb{R}^3 è fatto dalle matrici che hanno per colonne i vettori $v_1, v_2, v_3 \in \mathbb{R}^3$ tali che:

- v_1 è un punto della sfera S^2 ;
- v_2 è un punto in S^2 ortogonale a v_1 , e quindi $v_2 \in T_{v_1} S^2$;
- v_3 è il prodotto vettoriale dei primi due vettori, $v_3 = v_1 \wedge v_2$.

Dato che v_3 è completamente determinato da $(v_1, v_2) \in S^2 \times T_{v_1} S^2$, si ha che $SO(3) \subset TS^2$. Non si ha l'uguaglianza perché v_2 deve avere norma unitaria in questa scelta, mentre i vettori in TS^2 possono avere norma arbitraria.

Esercizio 5.1. Si definisce *sezione nulla* del fibrato tangente alla varietà M l'insieme

$$\sigma_0(M) := \{(p, 0) \in TM : p \in M\}.$$

Computare il primo gruppo di omotopia $\pi_1(TS^2 \setminus \sigma_0(S^2))$.

Se M è una varietà di dimensione n , allora TM ha una struttura canonica di varietà di dimensione $2n$ anche se M non è semplicemente una sottovarietà di \mathbb{R}^N . Per definire tale struttura, quello che richiediamo è che:

1. la funzione $\pi: TM \rightarrow M$ definita nell'Osservazione 5.1 è di classe C^∞ ;
2. i sottoinsiemi $T_p M \subseteq TM$ sono sottovarietà diffeomorfe a \mathbb{R}^n ;
3. se $\mathfrak{A} = \{(U_i, V_i, \varphi_i)\}_{i \in I}$ è un atlante che definisce la struttura differenziabile di M , allora $\mathfrak{B} = \{(\pi^{-1}(U_i), V_i \times \mathbb{R}^n, \vartheta_i)\}_{i \in I}$ è un atlante che definisce la struttura differenziabile di TM .

Ci restano da definire le applicazioni ϑ_i . Sia (U, V, φ) una carta locale C^∞ su M ; consideriamo le coordinate di $\varphi = (x_1, \dots, x_n)$ come coordinate locali su U . Sappiamo che

$$D\varphi_p: T_p M \rightarrow T_{\varphi(p)} \mathbb{R}^n \simeq \mathbb{R}^n, \quad v = \sum_{i=1}^n v_i \frac{\partial}{\partial x_i} \mapsto (v_1, \dots, v_n)$$

è, per ogni $p \in U$, un isomorfismo fra $T_p M$ e \mathbb{R}^n (e anzi è l'identità, visto che stiamo considerando le coordinate di φ come coordinate locali x_1, \dots, x_n). A questo punto appare chiaro che possiamo definire

$$\vartheta: \bigsqcup_{p \in U} T_p \rightarrow V \times \mathbb{R}^n, \quad (p, v) \mapsto (\varphi(p), D\varphi_p(v)).$$

Questa applicazione è sicuramente una biezione fra i fibrati tangenti di U e V ; nel definire la struttura differenziabile di TM imponiamo che sia un diffeomorfismo. Questo è naturale in quanto, dato che U e V sono diffeomorfi, stiamo solo chiedendo che varietà diffeomorfe abbiano "lo stesso" fibrato tangente.

6 Campi di vettori

6.1 Campi di vettori su varietà

Come per gli aperti di \mathbb{R}^n , vogliamo definire i campi di vettori anche sulle varietà differenziabili generiche.

Definizione 6.1. Sia M una varietà differenziabile e $U \subseteq M$ un aperto. Si definisce *campo di vettori* su U un'applicazione

$$\eta : U \rightarrow \bigsqcup_{p \in U} T_p M, \quad p \mapsto \eta(p)$$

che sia una *sezione* (ovvero $\eta(p) \in T_p M$ per ogni $p \in U$) e che sia C^∞ . Denoteremo con $\mathcal{X}(U)$ l'insieme dei campi di vettori C^∞ su U .

Cerchiamo di capire meglio cosa vuol dire l'essere C^∞ per un campo di vettori: a questo scopo mettiamoci in coordinate locali. Siano quindi $p \in U$ e x_1, \dots, x_n coordinate locali in un intorno aperto $V \subseteq U$ di p ; come sappiamo, $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1, \dots, n}$ è una base per lo spazio tangente $T_q M$ per ogni $q \in V$. Sia quindi $\eta : U \rightarrow TM$ un'applicazione tale che $\eta(q) \in T_q M$ per ogni $q \in U$; questo vuol dire che in V vale una scrittura del tipo

$$\eta = \eta(x_1, \dots, x_n) = \sum_{i=1}^n \eta_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i}.$$

La condizione che η è un campo di vettori C^∞ su U sarà per noi equivalente a dire che $\eta_1, \dots, \eta_n : V \rightarrow \mathbb{R}$ sono C^∞ . Verifichiamo che questa è una buona definizione, ovvero che è indipendente dal sistema di coordinate locali scelto. Sia quindi y_1, \dots, y_n un secondo sistema di coordinate locali in p ; dalla formula del cambio di coordinate (8) ricaviamo che

$$\eta = \sum_{i=1}^n \eta_i \frac{\partial}{\partial x_i} = \sum_{i=1}^n \eta_i \sum_{j=1}^n \frac{\partial y_j}{\partial x_i} \frac{\partial}{\partial y_j} = \sum_{j=1}^n \left(\sum_{i=1}^n \eta_i \frac{\partial y_j}{\partial x_i} \right) \frac{\partial}{\partial y_j}$$

e le funzioni

$$\mu_j := \sum_{i=1}^n \eta_i \frac{\partial y_j}{\partial x_i}$$

sono C^∞ se e solo se le η_i sono C^∞ . D'altronde, le carte del fibrato tangente sono

$$\vartheta : \pi^{-1}(U) \rightarrow V \times \mathbb{R}^n, \quad \sum_{i=1}^n \eta_i \frac{\partial}{\partial x_i} \in T_p M \mapsto (p, \eta_1, \dots, \eta_n)$$

e quindi $\eta : U \rightarrow \pi^{-1}(U)$ è C^∞ se e solo se $\vartheta \circ \eta : U \rightarrow V \times \mathbb{R}^n$ è C^∞ , ovvero se e solo se η_1, \dots, η_n sono C^∞ (essendo la prima coordinata $(\vartheta \circ \eta)_1(p) = p$ evidentemente C^∞).

Osservazione 6.1. Se η è un campo di vettori C^∞ sulla varietà M e $f \in C^\infty(M)$, allora anche $f\eta$ è un campo di vettori C^∞ su M . Infatti, sia $p \in M$, e siano x_1, \dots, x_n coordinate locali in p ; in queste coordinate

$$\eta = \eta(x_1, \dots, x_n) = \sum_{i=1}^n \eta_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i}, \quad f = f(x_1, \dots, x_n).$$

Abbiamo allora che

$$f\eta(x) = f\eta(x_1, \dots, x_n) = \sum_{i=1}^n f(x)\eta_i(x) \frac{\partial}{\partial x_i}$$

perciò le coordinate di $f\eta$ sono ottenute moltiplicando il vettore $\eta(x)$ per lo scalare $f(x)$; tale operazione è ovviamente C^∞ .

Questo fatto si può anche sintetizzare affermando che $\mathcal{X}(M)$ è un $C^\infty(M)$ -modulo.⁵

⁵Ricordiamo che, se A è un anello commutativo unitario, un A -modulo è uno "spazio vettoriale" a coefficienti in A , ovvero un gruppo abeliano F dove è definito un prodotto $\cdot : A \times F \rightarrow F$ che soddisfa tutte le proprietà del prodotto per scalare di uno spazio vettoriale su un campo.

Osservazione 6.2. Dimostriamo che $\mathcal{X}(M) \subseteq \text{Der}(C^\infty(M), C^\infty(M))$. Infatti, sia x_1, \dots, x_n un sistema di coordinate locali in un punto $p \in M$. Se f è una funzione C^∞ su M , possiamo definire

$$\eta(f)(p) := \eta(p)(f) = \sum_{i=1}^n \eta_i(p) \frac{\partial f}{\partial x_i}(p) \quad (10)$$

che è ancora una funzione C^∞ su M .

Esercizio 6.1. Dimostrare che $\mathcal{X}(M) \simeq \text{Der}(C^\infty(M), C^\infty(M))$.

6.2 Algebre di Lie

In questa sottosezione apriamo una parentesi algebrica che ci sarà utile a dimostrare alcune proprietà dei campi di vettori su varietà.

Sia A un anello commutativo unitario che sia anche un \mathbb{R} -spazio vettoriale, e consideriamo l'insieme delle derivazioni

$$\text{Der}(A, A) = \{\delta : A \rightarrow A : \delta \text{ è } \mathbb{R}\text{-lineare, } \delta(ab) = a\delta(b) + b\delta(a)\}.$$

Lo spazio $\text{Der}(A, A)$ non è chiuso per composizione: se δ_1 e δ_2 sono derivazioni di A in sé, in generale $\delta_1 \circ \delta_2$ non è una derivazione. Introduciamo allora il *bracket* (o *commutatore*)

$$[\delta_1, \delta_2] := \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1;$$

allora $[\delta_1, \delta_2]$ è una derivazione. La \mathbb{R} -linearità discende naturalmente da quella di δ_1 e δ_2 ; inoltre, $[\delta_1, \delta_2]$ soddisfa la regola di Leibniz perché

$$\begin{aligned} [\delta_1, \delta_2](ab) &= (\delta_1 \circ \delta_2 - \delta_2 \circ \delta_1)(ab) \\ &= \delta_1(\delta_2(ab)) - \delta_2(\delta_1(ab)) \\ &= \delta_1(a\delta_2(b) + b\delta_2(a)) - \delta_2(a\delta_1(b) + b\delta_1(a)) \\ &= a(\delta_1(\delta_2(b))) + b(\delta_1(\delta_2(a))) - a(\delta_2(\delta_1(b))) - b(\delta_2(\delta_1(a))) \\ &= a(\delta_1 \circ \delta_2(b) - \delta_2 \circ \delta_1(b)) + b(\delta_1 \circ \delta_2(a) - \delta_2 \circ \delta_1(a)) \\ &= a[\delta_1, \delta_2](b) + b[\delta_1, \delta_2](a). \end{aligned}$$

Alcune proprietà dell'applicazione *bracket* $[\cdot, \cdot] : \text{Der}(A, A) \times \text{Der}(A, A) \rightarrow \text{Der}(A, A)$, di semplice verifica, sono le seguenti:

1. è \mathbb{R} -bilineare;
2. è antisimmetrica, ovvero per ogni $\delta_1, \delta_2 \in \text{Der}(A, A)$ si ha

$$[\delta_1, \delta_2] = -[\delta_2, \delta_1];$$

3. per ogni $\delta_1, \delta_2, \delta_3 \in \text{Der}(A, A)$ vale che

$$[\delta_1, [\delta_2, \delta_3]] + [\delta_2, [\delta_3, \delta_1]] + [\delta_3, [\delta_1, \delta_2]] = 0.$$

Definizione 6.2. Sia L uno spazio vettoriale sul campo \mathbb{K} (di caratteristica $\neq 2$), e sia $[\cdot, \cdot] : L \times L \rightarrow L$ un'applicazione, detta *bracket*, che soddisfa le seguenti proprietà:

1. è \mathbb{K} -bilineare;
2. è antisimmetrica, ovvero per ogni $v, w \in L$ si ha

$$[v, w] = -[w, v];$$

3. per ogni $v_1, v_2, v_3 \in L$ vale l'identità di Jacobi

$$[v_1, [v_2, v_3]] + [v_2, [v_3, v_1]] + [v_3, [v_1, v_2]] = 0. \quad (11)$$

Allora L si dice un'algebra di Lie.

Esempio 6.1. Per quanto appena visto, lo spazio delle derivazioni di un anello commutativo unitario in sé è un'algebra di Lie se dotato del commutatore. Inoltre, se L è un'algebra di Lie e $v \in L$, l'applicazione $\delta := [v, \cdot] : L \rightarrow L$ soddisfa la regola di Leibniz; infatti, dalle tre proprietà del *bracket* discende che

$$\begin{aligned} & [v, [w, z]] + [w, [z, v]] + [z, [v, w]] = 0 \\ \implies & [v, [w, z]] + [w, -[v, z]] - [[v, w], z] = 0 \\ \implies & [v, [w, z]] - [-[v, z], w] - [[v, w], z] = 0 \\ \implies & \delta([w, z]) = -[\delta(z), w] + [\delta(w), z] \\ \implies & \delta([w, z]) = [w, \delta(z)] + [\delta(w), z]. \end{aligned}$$

Altri esempi di algebre di Lie sono $L = \mathfrak{M}_{n,n}(\mathbb{R})$ con il *bracket* $[A, B] = AB - BA$, e le sue sottoalgebre

$$\mathfrak{sl}(n, \mathbb{R}) := \{A \in \mathfrak{M}_{n,n}(\mathbb{R}) : \text{tr}(A) = 0\} \quad \text{e} \quad \mathfrak{o}(n, \mathbb{R}) := \{A \in \mathfrak{M}_{n,n}(\mathbb{R}) : A^T = -A\}$$

rispettivamente delle matrici a traccia nulla e antisimmetriche.

Torniamo ora ai campi di vettori su una varietà M . Dato che ogni campo di vettori è in particolare una derivazione di $C^\infty(M)$ in sé, a cosa corrisponde il *bracket* di due campi di vettori? Per scoprirlo, siano $\eta, \mu \in \mathcal{X}(M)$ e sia $f \in C^\infty(M)$: verifichiamo che $[\eta, \mu]$, oltre a essere una derivazione, è anche un campo di vettori. Sia x_1, \dots, x_n un sistema di coordinate locali in un punto $p \in M$ e supponiamo che in queste coordinate valga una scrittura del tipo

$$\eta = \sum_{i=1}^n \eta_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i}, \quad \mu = \sum_{i=1}^n \mu_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i}.$$

Ricordando la (10) e il fatto che f , essendo di classe C^∞ , soddisfa il teorema di Schwartz sull'inversione dell'ordine di derivazione, calcoliamo dunque

$$\begin{aligned} [\eta, \mu](f) &= \eta(\mu(f)) - \mu(\eta(f)) \\ &= \eta\left(\sum_{i=1}^n \mu_i \frac{\partial f}{\partial x_i}\right) - \mu\left(\sum_{i=1}^n \eta_i \frac{\partial f}{\partial x_i}\right) \\ &= \sum_{j=1}^n \eta_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^n \mu_i \frac{\partial f}{\partial x_i}\right) - \sum_{j=1}^n \mu_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^n \eta_i \frac{\partial f}{\partial x_i}\right) \\ &= \sum_{i,j=1}^n \eta_j \frac{\partial \mu_i}{\partial x_j} \frac{\partial f}{\partial x_i} + \sum_{i,j=1}^n \eta_j \mu_i \frac{\partial^2 f}{\partial x_j \partial x_i} - \sum_{i,j=1}^n \mu_j \frac{\partial \eta_i}{\partial x_j} \frac{\partial f}{\partial x_i} - \sum_{i,j=1}^n \mu_j \eta_i \frac{\partial^2 f}{\partial x_j \partial x_i} \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n \eta_j \frac{\partial \mu_i}{\partial x_j} - \mu_j \frac{\partial \eta_i}{\partial x_j}\right) \frac{\partial f}{\partial x_i}. \end{aligned}$$

Otteniamo quindi che se η_1, \dots, η_n sono le coordinate di η e μ_1, \dots, μ_n sono quelle di μ , allora le coordinate di $[\eta, \mu]$ sono

$$\sum_{j=1}^n \eta_j \frac{\partial \mu_i}{\partial x_j} - \mu_j \frac{\partial \eta_i}{\partial x_j}, \quad i = 1, \dots, n,$$

e quindi anche $[\eta, \mu]$ è un campo di vettori C^∞ su M .

6.3 Correlazione

Data un'applicazione C^∞ tra varietà $F: M \rightarrow N$, grazie al differenziale $DF_p: T_p M \rightarrow T_{F(p)} N$ siamo capaci di "portare avanti" gli spazi tangenti. Questo non è più vero se vogliamo fare lo stesso con interi campi di vettori: infatti, dato $\eta \in \mathcal{X}(M)$, è possibile che F mandi il punto p_1 e il punto p_2 di M nello stesso punto q di N , e che non ci sia una maniera univoca di definire $\mu(q) = F_*(\eta(p_1)) = F_*(\eta(p_2))$. Dovremo quindi ricorrere a un concetto più debole.

Definizione 6.3. Sia $F: M \rightarrow N$ una funzione C^∞ tra varietà, e siano $\eta \in \mathcal{X}(M)$, $\mu \in \mathcal{X}(N)$. I due campi di vettori si dicono F -correlati se per ogni p in M si ha che

$$DF_p(\eta(p)) = \mu(F(p)).$$

Scriveremo che $\mu = F_*\eta$ per indicare che η e μ sono F -correlati.

Lemma 6.1. Due campi di vettori $\eta \in \mathcal{X}(M)$, $\mu \in \mathcal{X}(N)$ sono F -correlati se, e solo se, per ogni aperto $U \subseteq N$ e per ogni funzione $f \in C^\infty(U)$ si ha

$$\eta(f \circ F) = \mu(f) \circ F \quad \text{ovvero} \quad \eta(F^*(f)) = F^*(\mu(f)). \quad (12)$$

Dimostrazione. Supponiamo che $DF_p(\eta(p)) = \mu(F(p))$ per ogni $p \in M$. Siano $U \subseteq N$ un aperto e f una funzione $C^\infty(U)$. Allora

$$\begin{aligned} \eta(f \circ F)(p) &= \eta(p)(f \circ F) && \text{(per la (10))} \\ &= DF_p(\eta(p))(f) && \text{(per definizione di differenziale)} \\ &= \mu(F(p))(f) && \text{(per l'ipotesi)} \\ &= [\mu(f) \circ F](p) && \text{(ancora per la (10))} \end{aligned}$$

e quindi il "solo se" dell'enunciato è dimostrato.

Viceversa, supponiamo che per ogni aperto $U \subseteq N$ e per ogni funzione $f \in C^\infty(U)$ valga la (12). Scegliamo in particolare come U un aperto dotato di un sistema di coordinate y_1, \dots, y_m , che sono funzioni $C^\infty(U)$; quindi si ha

$$\eta(y_j \circ F) = \mu(y_j) \circ F, \quad j = 1, \dots, m.$$

Ma $\mu(y_j)$ sono proprio le coordinate di μ , nel senso che se $q \in U$ allora

$$\mu(q) = \sum_{j=1}^m \mu(y_j)(q) \frac{\partial}{\partial y_j} \in T_q N.$$

Dalla definizione di differenziale e dall'ipotesi ricaviamo che

$$DF_p(\eta(p))(y_j) = \eta(p)(y_j \circ F) = \mu(F(p))(y_j), \quad j = 1, \dots, m,$$

e quindi, dato che le y_j formano un sistema di coordinate,

$$DF_p(\eta(p)) = \mu(F(p))$$

che è quanto volevamo dimostrare. □

Corollario 6.1. Sia $F: M \rightarrow N$ una funzione C^∞ tra varietà, e siano $\eta_1, \eta_2 \in \mathcal{X}(M)$, $\mu_1, \mu_2 \in \mathcal{X}(N)$. Se η_i è F -correlato a μ_i , $i = 1, 2$, allora $[\mu_1, \mu_2]$ è F -correlato a $[\eta_1, \eta_2]$.

Dimostrazione. Utilizziamo la condizione necessaria e sufficiente espressa dalla (12). Siano $U \subseteq N$ un aperto e $f \in C^\infty(U)$. Si ha allora

$$\begin{aligned} [\eta_1, \eta_2](f) &= \eta_1(\eta_2(F^*(f))) - \eta_2(\eta_1(F^*(f))) \\ &= \eta_1(F^*(\mu_2(f))) - \eta_2(F^*(\mu_1(f))) \\ &= F^*(\mu_1(\mu_2(f))) - F^*(\mu_2(\mu_1(f))) \\ &= F^*([\mu_1, \mu_2](f)) \end{aligned}$$

e quindi $[\mu_1, \mu_2] = F_*[\eta_1, \eta_2]$, sempre per la (12). \square

Corollario 6.2. Sia $F: M \rightarrow N$ un diffeomorfismo tra varietà. Allora l'applicazione

$$F_*: \mathcal{X}(M) \rightarrow \mathcal{X}(N), \quad \eta \mapsto F_*\eta$$

è ben definita ed è un isomorfismo di algebre di Lie.

Dimostrazione. Dato un campo di vettori $\eta \in \mathcal{X}(M)$ e un punto $q \in N$, ci basta definire

$$F_*\eta(q) := DF_{F^{-1}(q)}(\eta(F^{-1}(q))). \quad (13)$$

La verifica che F_* è un isomorfismo è semplice ed è lasciata per esercizio al lettore. \square

6.4 Gruppi di Lie

Così come un gruppo topologico è uno spazio topologico che sia un gruppo nella categoria degli spazi topologici, esiste un concetto analogo per le varietà differenziabili.

Definizione 6.4. Sia (G, \cdot) un gruppo. Si dice che G è un *gruppo di Lie* se è anche una varietà differenziabile e se le applicazioni di prodotto e di inversione,

$$\cdot: G \times G \rightarrow G, \quad (g, h) \mapsto g \cdot h \quad \text{e} \quad {}^{-1}: G \rightarrow G, \quad g \mapsto g^{-1}$$

sono di classe C^∞ .

Esempio 6.2. I principali gruppi di matrici, ovvero $GL(n, \mathbb{R})$, $SL(n, \mathbb{R})$, $O(n)$, $SO(n)$ e così via, sono gruppi di Lie, dato che il prodotto righe per colonne e l'operazione di inverso di una matrice sono tutte espressioni polinomiali nei coefficienti della matrice stessa, e quindi C^∞ .

Osservazione 6.3. Se G è un gruppo di Lie e g è un elemento fissato di G , l'applicazione di moltiplicazione a sinistra

$$L_g: G \rightarrow G, \quad h \mapsto g \cdot h$$

è un diffeomorfismo di G in sé, con inverso $(L_g)^{-1} = L_{g^{-1}}$.

Definizione 6.5. Sia G un gruppo di Lie e $\eta \in \mathcal{X}(G)$. Il campo di vettori η si dice *invariante a sinistra* (in inglese *left invariant*) se per ogni $g \in G$ si ha che η è L_g -correlato a se stesso:

$$(L_g)_*\eta = \eta \quad \forall g \in G.$$

Denoteremo con \mathfrak{g} l'insieme dei campi di vettori C^∞ su G che siano invarianti a sinistra.

Dalla definizione (13) di $(L_g)_*$ segue che, se η è un campo di vettori C^∞ su G e η è invariante a sinistra, allora per ogni $g \in G$ si ha

$$\eta(g) = DL_{g^{-1}}(\eta(1)) \quad (14)$$

dove 1 denota l'elemento neutro di G : quindi η è univocamente determinato dal suo valore $\eta(1) \in T_1G$.

Osserviamo inoltre che se η e μ sono invarianti a sinistra, allora anche $[\eta, \mu]$ è invariante a sinistra. Infatti, per ogni $g \in G$ il campo di vettori η è L_g -correlato a se stesso e lo stesso vale per μ ; inoltre l'operazione di *bracket* commuta con la correlazione (6.1).

Lemma 6.2. *L'applicazione*

$$\text{ev}_1: \mathfrak{g} \rightarrow T_1G, \quad \eta \mapsto \eta(1)$$

è un isomorfismo di spazi vettoriali.

Dimostrazione. Dato un vettore $v \in T_1G$ e un elemento $g \in G$, vogliamo trovare un campo di vettori η che valutato in g valga v . Basta quindi osservare che dalla (14) l'inversa di ev_1 è data da

$$DL_{g_1}: T_1G \rightarrow T_gG, \quad v \mapsto DL_{g_1}(v).$$

Mostriamo che il campo di vettori η è C^∞ . Innanzitutto, η è C^∞ in un intorno V di g ; se $h \in V$, allora $\eta(h) = DL_{h_1}(v) = DL_{h_1}(DL_{g_1}^{-1}(\eta(g)))$ dipende in maniera C^∞ da $\eta(g)$. Se $a \in G$, allora $U = L_a(V)$ è un intorno di a diffeomorfo a V , e η è C^∞ su U . Essendo C^∞ su un intorno di ogni punto di G , η è C^∞ su G . \square

Ne deduciamo quindi che anche T_1G è un'algebra di Lie: dati due vettori $v, w \in T_1G$ esistono unici due campi di vettori invarianti a sinistra $\eta, \mu \in \mathfrak{g}$ tali che $\eta(1) = v$ e $\mu(1) = w$. Ci basta a questo punto porre

$$[v, w] := [\eta, \mu](1).$$

7 Gruppi a un parametro

In questa sezione ritorneremo a concetti più geometrici che algebrici.

7.1 Flussi, curve integrali, generatori infinitesimali

Definizione 7.1. Sia M una varietà differenziabile. Un *flusso* su M (o *gruppo a un parametro*, o *sistema dinamico*) è un'applicazione $\vartheta: \mathbb{R} \times M \rightarrow M$ che sia un'azione su M del gruppo abeliano $(\mathbb{R}, +)$, ovvero tale che:

1. per ogni $x \in M$ si ha $\vartheta(0, x) = x$;
2. per ogni $x \in M$ e per ogni $s, t \in \mathbb{R}$ si ha $\vartheta(s, \vartheta(t, x)) = \vartheta(s + t, x)$.

Osservazione 7.1. Fissato $t \in \mathbb{R}$, l'applicazione

$$\vartheta_t: M \rightarrow M, \quad x \mapsto \vartheta(t, x)$$

è un diffeomorfismo di M in sé, con inverso $(\vartheta_t)^{-1} = \vartheta_{-t}$.

Definizione 7.2. Fissato un punto $x \in M$, la *linea di flusso* (o *curva integrale*) è l'applicazione

$$\alpha_x: \mathbb{R} \rightarrow M, \quad t \mapsto \vartheta(t, x).$$

Possiamo immaginare a un gruppo a un parametro come all'azione di un campo di forze costante sulla varietà M , e alla curva integrale come alla traiettoria che il punto x segue mosso da tali forze. Ovviamente, per ogni $t \in \mathbb{R}$ e $x \in M$ si ha

$$\alpha_x(t) = \vartheta_t(x).$$

Definizione 7.3. Dato $p \in M$, il *generatore infinitesimale* del flusso ϑ in p è il vettore

$$X(p) := \dot{\alpha}_p(0) = \frac{d\alpha_p}{dt}(0) = D\alpha_{p_0} \left(\frac{d}{dt} \right) \in T_p M.$$

Esaminiamo i concetti fin qui esposti più da vicino. Sia p un punto di M e siano x_1, \dots, x_n coordinate locali in un intorno V di p ; il flusso $\vartheta: (-\varepsilon, \varepsilon) \times V \rightarrow M$ avrà coordinate $(\vartheta_1, \dots, \vartheta_n)$. Conseguentemente

$$\dot{\alpha}_p(x) = \left(\frac{\partial \vartheta_1}{\partial t}(0, x), \dots, \frac{\partial \vartheta_n}{\partial t}(0, x) \right) \quad \forall x \in V$$

ovvero

$$X(p) = \sum_{i=1}^n \frac{\partial \vartheta_i}{\partial t}(0, p) \frac{\partial}{\partial x_i};$$

X risulta quindi un campo di vettori C^∞ su M . Nell'interpretazione fisica dei flussi, questo campo rappresenta la velocità istantanea del punto nell'istante iniziale.

Esempio 7.1. Se il flusso ϑ è dato da

$$\vartheta: \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \vartheta(t, (x_1, x_2)) = (x_1 + t, x_2)$$

il suo generatore infinitesimale è $X = \frac{\partial}{\partial x_1}$.

Esempio 7.2 (Esponenziale di una matrice). Definiamo l'*esponenziale* di una matrice $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ come

$$e^A := \sum_{k=0}^{+\infty} \frac{A^k}{k!} \in \mathfrak{M}_{n,n}(\mathbb{R}).$$

Stiamo considerando $\mathfrak{M}_{n,n}(\mathbb{R})$ come uno spazio metrico completo; vi si possono infatti definire un prodotto scalare e una norma inducendoli da quelli di \mathbb{R}^{n^2} , ovvero

$$\langle A, B \rangle = \text{tr}(AB^T) = \sum_{i,j=1}^n a_{ij}b_{ij}, \quad \|A\|^2 = \langle A, A \rangle = \sum_{i,j=1}^n a_{ij}^2.$$

Valgono i seguenti fatti, che non dimostriamo:

1. $\|AB\| \leq \|A\|\|B\|$ e $\|A+B\| \leq \|A\| + \|B\|$ per ogni $A, B \in \mathfrak{M}_{n,n}(\mathbb{R})$;
2. $\|A^m\| \leq \|A\|^m$ per ogni $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ e per ogni $m \in \mathbb{N}$ (si ottiene facilmente per induzione usando il punto precedente);
3. se A e B sono due matrici tali che $AB = BA$, allora $e^{A+B} = e^A e^B$.

Usando il punto 2., si ha che

$$\left\| \sum_{k=0}^{+\infty} \frac{A^k}{k!} \right\| \leq e^{\|A\|} < +\infty \quad \forall A \in \mathfrak{M}_{n,n}(\mathbb{R})$$

e quindi la serie che definisce l'esponenziale di matrice è totalmente convergente: l'applicazione $A \mapsto e^A$ è dunque ben definita. Si può dimostrare inoltre che tale funzione è anche analitica⁶ e quindi C^∞ su $\mathfrak{M}_{n,n}(\mathbb{R})$.

Fissiamo ora una matrice $E \in \mathfrak{M}_{n,n}(\mathbb{R})$ e consideriamo il flusso

$$\vartheta: \mathbb{R} \times \mathfrak{M}_{n,n}(\mathbb{R}) \rightarrow \mathfrak{M}_{n,n}(\mathbb{R}), \quad (t, A) \mapsto Ae^{tE}.$$

Questa applicazione è effettivamente un flusso; infatti, dato che tutti i multipli scalari di E commutano fra loro,

$$\begin{aligned} \vartheta(0, A) &= Ae^{0E} = AI + A \sum_{k=1}^{+\infty} \frac{0^k}{k!} = A, \\ \vartheta(s+t, A) &= Ae^{(s+t)E} = Ae^{sE}e^{tE} = \vartheta(t, Ae^{sE}) = \vartheta(t, \vartheta(s, A)). \end{aligned}$$

Il suo generatore infinitesimale è

$$X(A) = \left. \frac{d}{dt} (Ae^{tE}) \right|_{t=0} = AEe^{tE} \Big|_{t=0} = AE \in T_A \mathfrak{M}_{n,n}(\mathbb{R}).$$

Ci chiediamo come le funzioni cambiano sotto l'azione di un flusso. Consideriamo quindi un gruppo a un parametro $\vartheta: \mathbb{R} \times M \rightarrow M$ e una funzione $f \in C^\infty(M)$. Se calcoliamo f lungo una curva integrale, allora abbiamo che per definizione di differenziale

$$\frac{df \circ \alpha_p}{dt}(t) = D\alpha_{p_t} \left(\frac{d}{dt} \right) (f)$$

e quindi

$$\lim_{t \rightarrow 0} \frac{f \circ \vartheta_t - f}{t} = X(f).$$

⁶Una funzione si dice *analitica* in un punto se è sviluppabile in serie di Taylor in un intorno di tale punto e coincide con essa in questo intorno.

Osservazione 7.2. Il generatore infinitesimale X di un flusso ϑ è ϑ_t -correlato a se stesso per ogni $t \in \mathbb{R}$. Infatti, se f è una funzione C^∞ su M e $p \in M$, si ha

$$\begin{aligned} X(f \circ \vartheta_t)(p) &= \lim_{h \rightarrow 0} \frac{f \circ \vartheta_t(\vartheta_h(p)) - f \circ \vartheta_t(p)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\vartheta_{h+t}(p)) - f(\vartheta_t(p))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f \circ \vartheta_h(\vartheta_t(p)) - f(\vartheta_t(p))}{h} \\ &= X(f(\vartheta_t(p))) \end{aligned}$$

e quindi $(\vartheta_t)_* X = X$ per la (12).

Lemma 7.1. Sia $\vartheta : \mathbb{R} \times M \rightarrow M$ un flusso a un parametro, e sia $X \in \mathcal{X}(M)$ il suo generatore infinitesimale. Allora $X(p) = 0$ se, e solo se, $\vartheta(t, p) = p$ per ogni $t \in \mathbb{R}$.

Dimostrazione. Dato che $X(p)$ è definito come $\dot{\alpha}_p(0)$, se α_p è costante allora ovviamente $X(p)$ è nullo. Viceversa, se $X(p) = 0$, allora $X(\vartheta_t(p)) = 0$ per ogni $t \in \mathbb{R}$, per l'osservazione precedente; del resto, $X(\vartheta_t(p)) = X(\alpha_p(t)) = \dot{\alpha}_p(t)$. Poiché $\alpha_p : \mathbb{R} \rightarrow M$, per la connessione di \mathbb{R} si ha che α_p è costante se e solo se è localmente costante. Scegliamo un intervallo $(a, a + \varepsilon) \subseteq \mathbb{R}$ sufficientemente piccolo in maniera che la sua immagine sia tutta contenuta in un aperto $U \subseteq M$, che sia il dominio di una carta (U, V, φ) . Allora $\varphi \circ \alpha_p : (a, a + \varepsilon) \rightarrow V \subseteq \mathbb{R}^n$ ha derivata nulla, perché $\dot{\alpha}_p = 0$; ma per il teorema di Fermat, $\varphi \circ \alpha_p$ è costante, e poiché φ è un diffeomorfismo anche α_p è costante. \square

Le curve integrali di un flusso hanno dunque la proprietà che se la loro "velocità" si annulla in un punto, allora è sempre nulla; quindi possono essere costanti (la loro immagine è un punto di M), periodiche (la loro immagine è una curva chiusa diffeomorfa a S^1), oppure iniettive (vedi Figura 7.1).

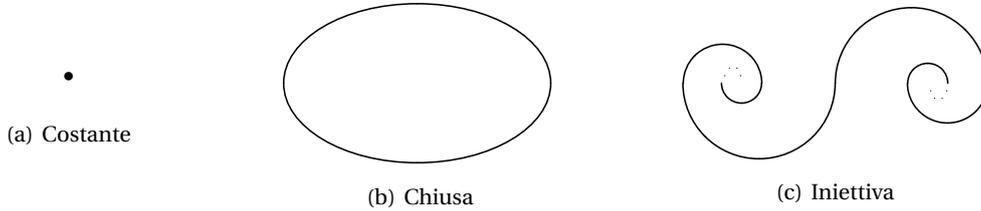


Figura 3: Curve integrali

7.2 Problema di Cauchy e sistemi dinamici

Per esplorare meglio le connessioni fra un flusso e il suo generatore infinitesimale, sfruttiamo un ben noto teorema di Equazioni Differenziali.

Teorema 7.1 (esistenza e unicità del problema di Cauchy). Sia V un intorno di $0 \in \mathbb{R} \times \mathbb{R}^n$ e siano $f_1, \dots, f_n \in C^\infty(V)$; allora, esistono $\varepsilon > 0$ e $r > 0$ tali che per ogni $a = (a_1, \dots, a_n) \in B(0, r) \subset \mathbb{R}^n$ esiste un'unica funzione $x_a = (x_1, \dots, x_n) : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ che risolve il problema di Cauchy

$$\begin{cases} \frac{dx_i}{dt}(t) = f_i(t, x_1(t), \dots, x_n(t)) & t \in (-\varepsilon, \varepsilon) \\ x_i(0) = a_i \end{cases}, \quad i = 1, \dots, n. \quad (15)$$

Inoltre, la funzione $(t, a) \mapsto x_a(t)$ è di classe C^∞ su $(-\varepsilon, \varepsilon) \times B(0, r)$.

Lemma 7.2. Sia $\vartheta: \mathbb{R} \times M \rightarrow M$ un flusso sulla varietà M e sia X il suo generatore infinitesimale. Se ϑ_1 è un altro flusso su M che ha per generatore infinitesimale X , allora ϑ_1 coincide con ϑ .

Dimostrazione. Fissiamo un punto $p \in M$ e consideriamo la curva integrale $\alpha_p(t) = \alpha(t) = \vartheta(t, p)$; per definizione di generatore infinitesimale $\dot{\alpha}(t) = X(\alpha(t))$. Supponiamo che $\beta_p = \beta$ sia la curva integrale in p del flusso ϑ_1 ; quello che vogliamo dimostrare è che $\alpha(t) = \beta(t)$ per ogni $t \in \mathbb{R}$.

Chiamiamo $A = \{t \in \mathbb{R} : \alpha(t) = \beta(t)\}$ (A è non vuoto perché $\alpha(0) = \beta(0) = p$) e dimostriamo che A è aperto e chiuso in \mathbb{R} ; la connessione di \mathbb{R} ci permetterà di concludere che $A = \mathbb{R}$. Dato che per definizione $A = (\alpha, \beta)^{-1}(\Delta_M)$ e M è uno spazio di Hausdorff, abbiamo subito che A è un chiuso perché le curve integrali sono funzioni continue. Sia ora $t_0 \in A$ e $q = \alpha(t_0) = \beta(t_0)$ la sua immagine in M secondo le curve integrali. Sia x_1, \dots, x_n un sistema di coordinate in un intorno aperto U di q ; in queste coordinate

$$X = \sum_{i=1}^n \eta_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i}, \quad \alpha = (\alpha_1, \dots, \alpha_n), \quad \beta = (\beta_1, \dots, \beta_n).$$

Dal fatto che $\dot{\alpha}(t) = X(\alpha(t))$, α risolve in U il problema di Cauchy

$$\begin{cases} \frac{d\alpha_i}{dt}(t) = \eta_i(\alpha_1(t), \dots, \alpha_n(t)) \\ \alpha_i(t_0) = q_i \end{cases}, \quad i = 1, \dots, n$$

che in un intorno I di t_0 ammette un'unica soluzione; del resto, anche β risolve lo stesso problema di Cauchy, e dunque $\alpha(t) = \beta(t)$ per ogni $t \in I$. Conseguentemente $I \subseteq A$, e A è anche aperto (perché intorno di ogni suo punto, dall'arbitrarietà di t_0). \square

Il risultato appena dimostrato, sotto un'ulteriore ipotesi di compattezza, si inverte.

Lemma 7.3. Sia X un campo di vettori su una varietà compatta M ; allora, esiste un unico flusso $\vartheta: \mathbb{R} \times M \rightarrow M$ di cui X è il generatore infinitesimale.

Idea di dimostrazione. Per ogni $p \in M$ esistono un numero positivo $\varepsilon_p > 0$, un intorno aperto V_p di p e una funzione $\vartheta_p: (-\varepsilon_p, \varepsilon_p) \times V_p \rightarrow M$ tale che

$$\frac{d\vartheta_p}{dt}(t, q) = X(\vartheta_p(t, q)), \quad \vartheta_p(0, p) = p; \quad (16)$$

ci basta risolvere il relativo problema di Cauchy, non prima di aver notato che V_p può essere scelto in maniera tale da essere tutto contenuto in una carta, e quindi diffeomorfo a un aperto di \mathbb{R}^n .

Chiamiamo

$$U = \bigcup_{p \in M} (-\varepsilon_p, \varepsilon_p) \times V_p \subseteq \mathbb{R} \times M;$$

ovviamente si ha che $\{0\} \times M \subseteq U$. Dall'unicità della soluzione del problema di Cauchy abbiamo che tutte le ϑ_p coincidono sull'intersezione dei loro domini; per cui possiamo costruire $\vartheta: U \rightarrow M$ per incollamento. Osserviamo ora che per ogni $q \in M$ l'insieme $U \cap (\mathbb{R} \times \{q\})$ è diffeomorfo a un intervallo aperto di \mathbb{R} contenente lo zero: è l'unione di tutti gli intervalli $(-\varepsilon_p, \varepsilon_p)$ estesa a tutti i p per i quali ϑ_q è definita su tale intervallo. Quello che ci resta da dimostrare è che $U = \mathbb{R} \times M$.

La famiglia \mathcal{F} delle coppie (U, ϑ) che soddisfano le proprietà appena dimostrate ($\{0\} \times M \subseteq U \subseteq \mathbb{R} \times M$, $\vartheta: U \rightarrow M$ soddisfa il problema (16), $U \cap (\mathbb{R} \times \{q\})$ è connesso per ogni $q \in M$) è non vuoto e può essere ordinato per estensione:

$$(U, \vartheta) \leq (V, \eta) \quad \text{se} \quad U \subseteq V, \quad \eta \Big|_U \equiv \vartheta.$$

Per il lemma di Zorn esiste una coppia (U, ϑ) massimale, con cui "sostituiamo" la coppia (U, ϑ) precedentemente costruita per incollamento.

Poiché M è compatto, esiste un numero finito di punti $p_1, \dots, p_m \in M$ tali che

$$M = \bigcup_{j=1}^m V_{p_j};$$

chiamando $\varepsilon = \min_{1 \leq j \leq m} \varepsilon_{p_j}$, si ha che $\varepsilon > 0$ e che $(-\varepsilon, \varepsilon) \times M \subseteq U$. Sia ora $t \in \mathbb{R}$; scegliamo un $N \in \mathbb{N}$

sufficientemente grande affinché $\left| \frac{t}{N} \right| < \varepsilon$. Ponendo ora

$$\vartheta(t, p) = \vartheta\left(\frac{t}{N}, \vartheta\left(\frac{t}{N}, \dots, \vartheta\left(\frac{t}{N}, \vartheta\left(\frac{t}{N}, p\right)\right) \dots\right)\right)$$

possiamo quindi estendere $\vartheta: \mathbb{R} \times M \rightarrow M$, e concludere così la dimostrazione. \square

Osservazione 7.3. L'ipotesi che M sia compatta può essere sostituita da quella che il supporto di X sia compatto (se la velocità si annulla in un punto interno alla varietà, la curva integrale rimane costantemente uguale a quel punto, quindi definita su tutto \mathbb{R}); la dimostrazione resta praticamente inalterata. Tuttavia essa non è un'ipotesi tecnica; è necessaria affinché la costante ε della dimostrazione sia uniforme su tutta la varietà (se così non fosse, è possibile che vicino al bordo ε debba essere definito come un estremo inferiore di certi ε_{p_j} , e quindi potrebbe essere nullo). Un controesempio è dato dal vettore $X = \frac{\partial}{\partial x_1}$ in $B(0, 1) \subset \mathbb{R}^2$; come abbiamo già visto, questo è il generatore del flusso di traslazione nella direzione dell'asse x_1 , e quindi tutte le curve integrali escono dalla palla unitaria.

8 Algebra multilineare

Questa sezione ci fornirà degli strumenti di Algebra lineare che ci ritorneranno utili in seguito.

8.1 Prodotti tensoriali

In questa sezione, I denoterà un insieme arbitrario di indici, e \mathbb{K} un campo (negli esempi che tratteremo in seguito molto spesso \mathbb{K} sarà il campo \mathbb{R} dei numeri reali).

Definizione 8.1. Sia $\{V_i\}_{i \in I}$ una famiglia di spazi vettoriali su \mathbb{K} . Si definisce:

- *prodotto diretto* dei V_i lo spazio vettoriale costituito dalle successioni di vettori di tali spazi:

$$\prod_{i \in I} V_i := \{(v_i) : v_i \in V_i \forall i \in I\};$$

- *somma diretta* dei V_i lo spazio vettoriale costituito dalle somme *finite* di vettori di tali spazi:

$$\bigoplus_{i \in I} V_i := \left\{ \sum_{i \in I} v_i : v_i \neq 0 \text{ per un numero finito di } i \in I \right\}.$$

Questi due insiemi sono effettivamente spazi vettoriali se dotati delle operazioni

$$\lambda(v_i) + \mu(w_i) = (\lambda v_i + \mu w_i), \quad \lambda \sum_{i \in I} v_i + \mu \sum_{i \in I} w_i = \sum_{i \in I} [\lambda v_i + \mu w_i], \quad \forall \lambda, \mu \in \mathbb{K};$$

osserviamo che nel secondo caso la combinazione lineare di somme finite è ancora una somma finita.

Osservazione 8.1. Nel caso in cui I sia finito, il prodotto diretto e la somma diretta dei V_i sono canonicamente isomorfi.

Il prodotto diretto dei V_i è il prodotto dei V_i nella categoria **Vec** degli spazi vettoriali, ovvero

1. esiste, per ogni $k \in I$, una proiezione

$$\pi_k : \prod_{i \in I} V_i \rightarrow V_k, \quad (\dots, v_k, \dots) \mapsto v_k$$

che è un'applicazione lineare;

2. per ogni spazio vettoriale W e per ogni famiglia di applicazioni lineari $f_k : W \rightarrow V_k$, $k \in I$, esiste un'unica applicazione lineare $f : W \rightarrow \prod_{i \in I} V_i$ tale che $f_k = \pi_k \circ f$, ovvero

$$\begin{array}{ccc} \prod_{i \in \{1,2\}} V_i & \xrightarrow{\pi_1} & V_1 \\ \downarrow \pi_2 & \nearrow \exists! f & \uparrow f_1 \\ V_2 & \xleftarrow{f_2} & W \end{array} \quad f(w) = (f_i(w)).$$

Analogamente, la somma diretta dei V_i è il coprodotto dei V_i (o, equivalentemente, il prodotto nella categoria opposta **Vec^o**), ovvero

1. esiste, per ogni $k \in I$, un'inclusione

$$i_k : V_k \hookrightarrow \bigoplus_{i \in I} V_i, \quad v_k \mapsto \sum_{i \in I} v_i, \text{ dove } v_i = 0 \text{ se } i \neq k$$

che è un'applicazione lineare;

2. per ogni spazio vettoriale W e per ogni famiglia di applicazioni lineari $f_k : V_k \rightarrow W$, $k \in I$, esiste un'unica applicazione lineare $f : \bigoplus_{i \in I} V_i \rightarrow W$ tale che $f_k = f \circ i_k$, ovvero

$$\begin{array}{ccc}
 \bigoplus_{i \in \{1,2\}} V_i & \xleftarrow{i_1} & V_1 \\
 \uparrow i_2 & \searrow \exists! f & \downarrow f_1 \\
 V_2 & \xleftarrow{f_2} & W
 \end{array}
 \quad
 f \left(\sum_{i \in I} v_i \right) = \sum_{i \in I} f_i(v_i);$$

osserviamo che quest'ultima è una somma finita (solo un numero finito di termini è diverso da zero) e quindi definisce un elemento di W .

Definizione 8.2. Sia S un insieme qualsiasi; si definisce *spazio vettoriale generato da S* lo spazio

$$\langle S \rangle := \bigoplus_{s \in S} \mathbb{K} \cdot s.$$

Esiste un'ovvia inclusione $\alpha : S \hookrightarrow \langle S \rangle$, $s \mapsto 1 \cdot s =: s$. Inoltre, S risulta essere per costruzione una base di $\langle S \rangle$.

Prendiamo ora in considerazione due spazi vettoriali V e W sul campo \mathbb{K} ; vogliamo rendere bilineare l'applicazione $\alpha : V \times W \rightarrow \langle V \times W \rangle$. A tale scopo, chiamiamo \mathcal{T} la famiglia dei sottospazi $H \subseteq \langle V \times W \rangle$ tali che la fattorizzazione a quoziente $\pi \circ \alpha : V \times W \rightarrow \langle V \times W \rangle / H$ sia bilineare (in questo caso π denota la proiezione a quoziente $\langle V \times W \rangle \rightarrow \langle V \times W \rangle / H$). La famiglia \mathcal{T} è non vuota perché contiene l'intero spazio $\langle V \times W \rangle$: se “mandiamo tutto a zero” ogni applicazione diventa bilineare! Dato che l'intersezione di due sottospazi vettoriali è ancora un sottospazio, possiamo scegliere l'elemento “minimale” fra quelli di \mathcal{T} , ovvero

$$H_0 = \bigcap_{H \in \mathcal{T}} H.$$

Lemma 8.1. Lo spazio H_0 è lo spazio generato dalle relazioni di bilinearità

- (a) $\{(v_1 + v_2, w) - (v_1, w) - (v_2, w) : v_1, v_2 \in V, w \in W\}$;
- (b) $\{(v, w_1 + w_2) - (v, w_1) - (v, w_2) : v \in V, w_1, w_2 \in W\}$;
- (c) $\{(\lambda v, w) - \lambda(v, w) : v \in V, w \in W, \lambda \in \mathbb{K}\}$;
- (d) $\{(v, \lambda w) - \lambda(v, w) : v \in V, w \in W, \lambda \in \mathbb{K}\}$.

Dimostrazione. Denotiamo con K l'insieme generato da tali relazioni. È molto semplice verificare che $K \subseteq H_0$; infatti, se H è in \mathcal{T} allora per definizione

$$\alpha(v_1 + v_2, w) \equiv (v_1, w) + (v_2, w) \pmod{H} \quad \forall v_1, v_2 \in V, \forall w \in W$$

e analoghe equivalenze valgono rispetto alle altre relazioni di bilinearità. Se tali equivalenze valgono per tutti gli $H \in \mathcal{T}$, devono valere anche per la loro intersezione H_0 . Viceversa, essendo K definito proprio dalle relazioni di bilinearità, l'applicazione $\pi \circ \alpha : V \times W \rightarrow \langle V \times W \rangle / K$ è sicuramente bilineare, per cui K è in \mathcal{T} ; essendo H_0 l'intersezione di tutti i sottospazi contenuti in \mathcal{T} , troviamo anche l'altra inclusione $H_0 \subseteq K$. Possiamo concludere che $H_0 = K$, come volevamo dimostrare. \square

Definizione 8.3. Nelle notazioni precedenti, si definisce *prodotto tensoriale* degli spazi vettoriali V e W lo spazio vettoriale

$$V \otimes W := \langle V \times W \rangle / H_0.$$

Si denota con $v \otimes w$ il *prodotto tensoriale* di v con w , ovvero l'immagine della coppia $(v, w) \in V \times W$ secondo la mappa $\iota := \pi \circ \alpha : V \times W \rightarrow V \otimes W$.

Osserviamo che, come si evince dalla dimostrazione del Lemma 8.1, ι è per costruzione un'applicazione bilineare, e che inoltre l'insieme dei prodotti tensoriali $v \otimes w$, al variare di $v \in V$ e $w \in W$, generano l'intero spazio $V \otimes W$ (nel senso che lo spazio da essi generato non è contenuto in nessun sottospazio proprio di $V \otimes W$).

Teorema 8.1 (proprietà universale del prodotto tensoriale). *Comunque si prendano uno spazio vettoriale Z e un'applicazione bilineare $\beta: V \times W \rightarrow Z$, esiste un'unica applicazione lineare $f: V \otimes W \rightarrow Z$ tale che $\beta = f \circ \iota$.*

In altre parole, la proprietà universale del prodotto tensoriale consiste nel trasformare applicazioni bilineari (dal prodotto cartesiano di due spazi vettoriali) in applicazioni lineari (dal loro prodotto tensoriale): per ogni spazio vettoriale Z si ha che

$$\begin{array}{ccc} V \times W & \xrightarrow{\beta} & Z \\ \downarrow \iota & \searrow \exists! f & \\ V \otimes W & & \end{array} \quad \text{Lin}(V \otimes W, Z) \xrightarrow{\cdot \circ \iota} \text{Bil}(V \times W, Z) \text{ è un isomorfismo.}$$

Dimostrazione. L'applicazione bilineare β induce un'applicazione lineare $\tilde{f}: \langle V \times W \rangle \rightarrow Z$, che manda $1 \cdot (v, w)$ in $\beta(v, w)$. Del resto, dato che β è bilineare e $H_0 \subset \langle V \times W \rangle$ è generato dalle relazioni di bilinearità, si ha $\tilde{f}(H_0) = \{0\}$; è quindi possibile fattorizzare \tilde{f} a $f: \langle V \times W \rangle / H_0 \rightarrow Z$. Ci basta ora osservare che per definizione $\langle V \times W \rangle / H_0 = V \otimes W$.

L'unicità di tale f segue dal fatto che, per costruzione, $f(v \otimes w) = \beta(v, w)$, e i prodotti tensoriali $v \otimes w$ generano tutto $V \otimes W$. \square

Esercizio 8.1. Siano V e W spazi vettoriali di dimensione finita, con $\dim V = n$ e $\dim W = m$. Dimostrare che se $\{e_i\}_{i=1, \dots, n}$ è una base di V e $\{\varepsilon_j\}_{j=1, \dots, m}$ è una base di W , allora le coppie $\{e_i \otimes \varepsilon_j\}_{\substack{i=1, \dots, n \\ j=1, \dots, m}}$ formano una base di $V \otimes W$. Dedurne che $\dim V \otimes W = nm$.

Esempio 8.1. Se V e W sono spazi di dimensione finita, allora $V \otimes W^* \simeq \text{Hom}(W, V)$. Definiamo infatti

$$\beta: V \times W^* \rightarrow \text{Hom}(W, V), \quad \beta(v, \varphi)(w) := \varphi(w)v;$$

l'immagine di β è data da tutte le applicazioni lineari da W in V di rango minore o uguale a 1. Essendo ogni $\varphi \in W^*$ un'applicazione lineare $\varphi: W \rightarrow \mathbb{K}$, β è un'applicazione bilineare. In base alla proprietà universale del prodotto tensoriale, è possibile fattorizzarla ad un'applicazione $f: V \otimes W^* \rightarrow \text{Hom}(W, V)$: verifichiamo che f è un isomorfismo. Siano dunque n e m le dimensioni rispettivamente di V e di W ; scegliamo una base $\{e_i\}_{i=1, \dots, n}$ di V e una base $\{\varepsilon_j\}_{j=1, \dots, m}$ di W . Se $\{\varphi_j\}_{j=1, \dots, m}$ è la base duale di quella scelta per W , allora

$$f(e_i \otimes \varphi_j)(\varepsilon_k) = \beta(e_i, \varphi_j)(\varepsilon_k) = \varphi_j(\varepsilon_k) e_i = \delta_j^k e_i$$

e quindi $f(e_i \otimes \varphi_j)$ è l'applicazione lineare da W in V rappresentata dalla matrice che ha come unico elemento non nullo quello alla riga i -esima e colonna j -esima, che vale 1. Al variare di i da 1 a n e j da 1 a m , tali matrici generano ovviamente tutto lo spazio $\text{Hom}(W, V)$.

Il prodotto tensoriale di due spazi vettoriali V e W è a sua volta uno spazio vettoriale, quindi possiamo considerarne il prodotto tensoriale con un terzo spazio Z . Si verifica che tale prodotto è associativo:

$$V \otimes W \otimes Z = (V \otimes W) \otimes Z = V \otimes (W \otimes Z).$$

In particolare, denoteremo con

$$\bigotimes^n V := \underbrace{V \otimes V \otimes \dots \otimes V}_{n \text{ volte}}$$

il prodotto tensoriale di uno spazio vettoriale V con se stesso n volte, ovvero quello che si ottiene quozientando lo spazio $\bigoplus_{i=1}^n V$ per il sottospazio generato dalle relazioni di multilinearità (ovvero di linearità componente per componente). Chiameremo inoltre *tensori semplici* gli elementi di $\otimes^n V$, ovvero quelli della forma $v_1 \otimes v_2 \otimes \dots \otimes v_n$.

Osservazione 8.2. La proprietà universale del prodotto tensoriale (Teorema 8.1) si generalizza a prodotti tensoriali di un numero arbitrario di spazi considerando le applicazioni multilineari.

8.2 Prodotti esterni

Definizione 8.4. Sia A il sottospazio di $\otimes^n V$ generato dalle relazioni alternanti:

$$A = \langle \{v_1 \otimes \dots \otimes v_n : v_i = v_j \text{ per qualche } i, j\} \rangle.$$

Si definisce *prodotto esterno n -esimo* di V lo spazio vettoriale

$$\bigwedge^n V := (\otimes^n V) / A.$$

Si pone per convenzione $\bigwedge^0 V = \mathbb{K}$.

Ad esempio, dato uno spazio vettoriale V , il suo secondo prodotto esterno è

$$\bigwedge^2 V := (V \times V) / \langle \{v \otimes v : v \in V\} \rangle.$$

Indichiamo con $\lambda = \pi \circ \iota : V \times V \rightarrow \bigwedge^2 V$ l'inclusione del prodotto cartesiano di V con se stesso nel suo secondo prodotto esterno (π è la proiezione a quoziente $V \otimes V \rightarrow \bigwedge^2 V$). Ovviamente il prodotto esterno è generato dalle immagini delle coppie $(v_1, v_2) \in V \times V$ secondo λ , che denoteremo con il simbolo $v_1 \wedge v_2$.

Teorema 8.2 (proprietà universale del prodotto esterno). *Comunque si prendano uno spazio vettoriale Z e un'applicazione bilineare alternante $\beta : V \times V \rightarrow Z$, esiste un'unica applicazione lineare $f : \bigwedge^2 V \rightarrow Z$ tale che $\beta = f \circ \lambda$.*

Dimostrazione. Per la proprietà universale del prodotto tensoriale, essendo β bilineare si può fattorizzare a $\tilde{f} : V \otimes V \rightarrow Z$. Si ha per costruzione che $\tilde{f}(v \otimes v) = \beta(v, v) = 0$ per ogni $v \in V$, dato che β è anche alternante; conseguentemente, \tilde{f} fattorizza ulteriormente a $f : \bigwedge^2 V \rightarrow Z$. L'unicità segue ancora dal fatto che i prodotti $v_1 \wedge v_2$ generano tutto il prodotto esterno. \square

Possiamo affermare che la proprietà universale del prodotto esterno trasforma quindi le applicazioni bilineari alternanti (dal prodotto cartesiano di uno spazio vettoriale con se stesso) in applicazioni lineari (dal secondo prodotto esterno): per ogni spazio vettoriale Z

$$\begin{array}{ccc} V \times V & \xrightarrow{\beta} & Z \\ \downarrow \lambda & \nearrow \exists! f & \\ \bigwedge^2 V & & \end{array} \quad \text{Lin}(\bigwedge^2 V, Z) \xrightarrow{\circ \lambda} \text{BilAlt}(V \times V, Z) \text{ è un isomorfismo.}$$

Anche in questo caso, se consideriamo i prodotti $\bigwedge^n V$ possiamo generalizzare tale proprietà ad applicazioni multilineari alternanti.

Osservazione 8.3. Sia V uno spazio vettoriale. Per le proprietà di multilinearità e alternanza si ha che

$$\begin{aligned} 0 &= (v_1 + v_2) \wedge (v_1 + v_2) \wedge v_3 \wedge \dots \wedge v_n = \\ &= v_1 \wedge v_1 \wedge v_3 \wedge \dots \wedge v_n + v_2 \wedge v_1 \wedge v_3 \wedge \dots \wedge v_n + v_1 \wedge v_2 \wedge v_3 \wedge \dots \wedge v_n + \\ &\quad + v_2 \wedge v_2 \wedge v_3 \wedge \dots \wedge v_n = \\ &= v_2 \wedge v_1 \wedge v_3 \wedge \dots \wedge v_n + v_1 \wedge v_2 \wedge v_3 \wedge \dots \wedge v_n \end{aligned}$$

e quindi

$$v_2 \wedge v_1 \wedge v_3 \wedge \cdots \wedge v_n = -v_1 \wedge v_2 \wedge v_3 \wedge \cdots \wedge v_n.$$

Ovviamente, lo stesso si può dire se si scambiano i vettori corrispondenti ad una qualsiasi coppia di indici distinti i e j :

$$v_1 \wedge \cdots \wedge v_j \wedge \cdots \wedge v_i \wedge \cdots \wedge v_n = -v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_j \wedge \cdots \wedge v_n.$$

Inoltre, se σ è una qualsiasi permutazione degli indici $\{1, 2, \dots, n\}$ e $(-1)^\sigma$ è la sua segnatura, allora è facile verificare che

$$v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \cdots \wedge v_{\sigma(n)} = (-1)^\sigma v_1 \wedge v_2 \wedge \cdots \wedge v_n. \quad (17)$$

Consideriamo ora uno spazio vettoriale V di dimensione finita m ; ci proponiamo di trovare una base per $\wedge^n V$. Sappiamo già che se e_1, \dots, e_m è una base per V , allora i tensori semplici $\{e_{i_1} \otimes \cdots \otimes e_{i_n}\}_{i_1, \dots, i_n \in \{1, \dots, m\}}$ formano una base per $\otimes^n V$. Il prodotto esterno è ottenuto da quest'ultimo quotizzando rispetto alle relazioni alternanti; quindi per ottenere una base di $\wedge^n V$ dobbiamo scegliere le immagini secondo λ dei tensori semplici ottenuti scegliendo elementi distinti della base di V . Inoltre, due scelte che comprendono gli stessi vettori ma in un'ordine diverso danno una relazione di dipendenza lineare, come segue dalla (17): possiamo dunque scrivere che $\wedge^n V$ è generato dall'insieme dei prodotti $e_{i_1} \wedge \cdots \wedge e_{i_n}$, $i_1, \dots, i_n \in \{1, \dots, m\}$, con $i_1 < i_2 < \dots < i_n$.

Ci resta dunque da verificare che questi siano anche linearmente indipendenti. Identifichiamo a tale scopo il prodotto cartesiano $V \times \cdots \times V$, n volte, con lo spazio delle matrici $\mathfrak{M}_{n,m}(\mathbb{K})$; ogni fattore del prodotto cartesiano determina un vettore colonna di una matrice. Fissiamo una scelta di indici $i_1 < i_2 < \dots < i_n$ e chiamiamo $\beta: \mathfrak{M}_{n,m}(\mathbb{K}) \rightarrow \mathbb{K}$ l'applicazione che alla matrice A associa il determinante del minore ottenuto scegliendo da A le righe i_1, \dots, i_n . Essendo un determinante, β è un'applicazione multilineare alternante: per la proprietà universale del prodotto esterno, possiamo fattorizzarla ad un'applicazione lineare $f: \wedge^n V \rightarrow \mathbb{K}$. È ora chiaro che

$$f(e_{j_1} \wedge e_{j_2} \wedge \cdots \wedge e_{j_n}) = \beta(e_{j_1}, e_{j_2}, \dots, e_{j_n}) = \begin{cases} 1 & \text{se } j_1 = i_1, \dots, j_n = i_n \\ 0 & \text{altrimenti;} \end{cases}$$

infatti, nel primo caso il minore che stiamo considerando nel calcolo di β è l'identità, altrimenti almeno una delle sue colonne è nulla.

Otteniamo in particolare che

$$\dim \wedge^n V = \binom{m}{n}.$$

Osserviamo che, per $n > m$, $\wedge^n V = \{0\}$: non possiamo scegliere n vettori distinti da una base di m elementi! Analogamente, $\wedge^m V = \mathbb{K}$: da questo segue che l'applicazione determinante $\det: \mathfrak{M}_{m,m}(\mathbb{K}) \rightarrow \mathbb{K}$ è univocamente determinata dall'essere multilineare, alternante e tale che $\det I = 1$. Qualsiasi altra applicazione multilineare alternante $g: \mathfrak{M}_{m,m}(\mathbb{K}) \rightarrow \mathbb{K}$ deve pertanto essere un multiplo di \det .

Osservazione 8.4. Data un'applicazione lineare $f: V \rightarrow W$ tra due spazi vettoriali, questa induce un'applicazione lineare tra i relativi prodotti esterni

$$\wedge^n f: \wedge^n V \rightarrow \wedge^n W, \quad v_1 \wedge \cdots \wedge v_n \mapsto f(v_1) \wedge \cdots \wedge f(v_n).$$

Essa è ben definita: l'applicazione $f: V \rightarrow W$ si può estendere a un'applicazione $(f, \dots, f): V \times \cdots \times V \rightarrow W \times \cdots \times W$, n volte, che fattorizza ad un'applicazione multilineare e alternante dal prodotto cartesiano di V con se stesso n volte al prodotto esterno n -esimo di W . L'applicazione $\wedge^n f$ si costruisce allora sfruttando la proprietà universale del prodotto esterno.

Definizione 8.5. Si definisce *algebra esterna* di V lo spazio vettoriale

$$\bigwedge^* V := \bigoplus_{n=0}^{\dim V} \bigwedge^n V.$$

Il prodotto di quest'algebra è quello che si ottiene fattorizzando a quoziente il prodotto bilineare alternante

$$\bigotimes^a V \otimes \bigotimes^b V \rightarrow \bigotimes^{a+b} V, \quad (v_1 \otimes \cdots \otimes v_a) \otimes (w_1 \otimes \cdots \otimes w_b) \mapsto v_1 \otimes \cdots \otimes v_a \otimes w_1 \otimes \cdots \otimes w_b,$$

ovvero

$$\bigwedge^a V \wedge \bigwedge^b V \rightarrow \bigwedge^{a+b} V, \quad (v_1 \wedge \cdots \wedge v_a) \wedge (w_1 \wedge \cdots \wedge w_b) \mapsto v_1 \wedge \cdots \wedge v_a \wedge w_1 \wedge \cdots \wedge w_b.$$

Osserviamo che $\dim \bigwedge^* V = 2^{\dim V}$.

8.3 Prodotto di contrazione

Sia V un \mathbb{K} -spazio vettoriale di dimensione finita m . Vogliamo dimostrare che il prodotto esterno n -esimo del suo duale, $\bigwedge^n V^*$, è canonicamente isomorfo al duale del suo prodotto esterno n -esimo, $(\bigwedge^n V)^*$.

A tale scopo, fissiamo un vettore $v \in V$ e definiamo l'operazione di *prodotto di contrazione* come segue:

$$v \lrcorner : \bigwedge^n V^* \rightarrow \bigwedge^{n-1} V^*, \quad v \lrcorner (\varphi_1 \wedge \cdots \wedge \varphi_n) = \sum_{i=1}^n (-1)^{i+1} \varphi_i(v) \varphi_1 \wedge \cdots \wedge \varphi_{i-1} \wedge \varphi_{i+1} \wedge \cdots \wedge \varphi_n. \quad (18)$$

Si vede subito che questa applicazione è lineare (anche in v). Per verificare che è anche ben definita, consideriamo l'applicazione multilineare

$$(\varphi_1, \dots, \varphi_n) \in V^* \times \cdots \times V^* \mapsto \sum_{i=1}^n (-1)^{i+1} \varphi_i(v) \varphi_1 \wedge \cdots \wedge \varphi_{i-1} \wedge \varphi_{i+1} \wedge \cdots \wedge \varphi_n \in \bigwedge^{n-1} V^*;$$

il segno $(-1)^{i+1}$ posto nella sommatoria rende tale applicazione anche alternante. Infatti, se ad esempio $\varphi_1 = \varphi_2$, allora la sommatoria vale

$$\varphi_1(v) \varphi_3 \wedge \cdots \wedge \varphi_n - \varphi_1(v) \varphi_3 \wedge \cdots \wedge \varphi_n + \sum_{i=3}^n (-1)^{i+1} \varphi_i(v) \varphi_1 \wedge \varphi_1 \wedge \cdots \wedge \varphi_{i-1} \wedge \varphi_{i+1} \wedge \cdots \wedge \varphi_n$$

e questo valore è nullo in quanto tutti i termini dell'ultima sommatoria contengono prodotti esterni di elementi uguali. Questa applicazione multilineare alternante fattorizza allora al prodotto di contrazione definito in (18).

Abbiamo quindi un'applicazione bilineare $\lrcorner : V \times \bigwedge^n V^* \rightarrow \bigwedge^{n-1} V^*$; possiamo estenderla definendo ricorsivamente

$$\lrcorner : \bigwedge^k V \times \bigwedge^n V^* \rightarrow \bigwedge^{n-k} V^*, \quad (v_1 \wedge \cdots \wedge v_k) \lrcorner (\varphi_1 \wedge \cdots \wedge \varphi_n) = (v_1 \wedge \cdots \wedge v_{k-1}) \lrcorner [v_k \lrcorner (\varphi_1 \wedge \cdots \wedge \varphi_n)].$$

In particolare, per $k = n$, si verifica che

$$\lrcorner : \bigwedge^n V \times \bigwedge^n V^* \rightarrow \mathbb{K}, \quad (v_1 \wedge \cdots \wedge v_n) \lrcorner (\varphi_1 \wedge \cdots \wedge \varphi_n) = \sum_{\sigma \in \Sigma_n} (-1)^\sigma \varphi_{\sigma(1)}(v_1) \cdots \varphi_{\sigma(n)}(v_n)$$

ovvero che il prodotto di contrazione corrisponde al determinante della matrice $(\varphi_j(v_i))_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$. In particolare, \lrcorner è non degenere, ovvero

$$(v_1 \wedge \cdots \wedge v_n) \lrcorner (\varphi_1 \wedge \cdots \wedge \varphi_n) = 0 \quad \forall \varphi_1 \wedge \cdots \wedge \varphi_n \in \bigwedge^n V^* \quad \implies \quad v_1 \wedge \cdots \wedge v_n = 0;$$

infatti, basta scegliere come v_1, \dots, v_n un insieme di vettori di V linearmente indipendenti, estenderli a una base e poi scegliere come $\varphi_1, \dots, \varphi_n$ i corrispondenti vettori della base duale per ottenere un prodotto di contrazione non banale.

A questo punto, appare chiaro che

$$\bigwedge^n V^* \rightarrow (\bigwedge^n V)^*, \quad \varphi_1 \wedge \dots \wedge \varphi_n \mapsto \cdot \lrcorner (\varphi_1 \wedge \dots \wedge \varphi_n)$$

è un isomorfismo canonico, non dipendendo dalla scelta di basi negli spazi che coinvolge.

9 Forme differenziali

Torniamo ora allo studio delle varietà differenziabili: faremo uso delle nozioni apprese nella precedente sezione per “dualizzare” molti dei concetti esposti nel corso di queste note. Va detto che gli oggetti che stiamo per introdurre possono in realtà essere definiti in maniera più intrinseca rispetto alle varietà stesse (ad esempio senza l’uso di coordinate locali ma con argomenti puramente algebrici) e che tali definizioni si rivelano poi adatte anche a spazi più generali. Abbiamo preferito seguire la strada inversa perché permette di “maneggiare” per primi oggetti con cui lo studente ha già avuto un primo approccio nei corsi di Calcolo precedenti.

9.1 Fibrato cotangente

Consideriamo una varietà differenziabile X di dimensione n e un suo punto p . Abbiamo già visto più volte che a questo punto possiamo associare lo spazio vettoriale $T_p X$, che abbiamo chiamato *spazio tangente a X in p* , e che abbiamo definito come lo spazio delle derivazioni dall’anello dei germi di funzioni C^∞ in p a valori reali, $T_p X := \text{Der}(\mathcal{E}(p), \mathbb{R})$; una volta introdotto un sistema di coordinate locali in p , diciamo x_1, \dots, x_n , una base dello spazio tangente è data dalle derivazioni $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$. Abbiamo poi osservato che $\frac{\partial x_j}{\partial x_i} = \delta_i^j$, come potevamo aspettarci.

Definizione 9.1. Si definisce *spazio cotangente a X in p* il duale dello spazio tangente a X in p :

$$T_p^* X := \text{Hom}(T_p X, \mathbb{R}).$$

Gli elementi dello spazio cotangente vengono detti *covettori*.

Per quanto appena osservato, le coordinate x_j (che sono funzioni C^∞ in un intorno di p) si comportano in un certo qual modo come la base duale di $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1, \dots, n}$. Denotiamo quindi con $\{dx_i\}_{i=1, \dots, n}$ tale base duale (che risulterà quindi essere una base per $T_p^* X$):

$$dx_j \left(\frac{\partial}{\partial x_i} \right) = \delta_i^j.$$

Possiamo definire lo spazio cotangente anche in una maniera più algebrica, come già accennato. Consideriamo infatti l’anello $\mathcal{E}(p)$ dei germi di funzioni C^∞ in un intorno di p ; abbiamo già visto (Lemma 1.5) che esso contiene un solo ideale massimale

$$\mathcal{M}(p) = \{f \in \mathcal{E}(p) : f(p) = 0\}.$$

Lemma 9.1. Esiste un isomorfismo naturale fra $T_p^* X$ e l’anello quoziente $\mathcal{M}(p)/\mathcal{M}^2(p)$.

Ricordiamo che se I è un ideale di un anello A , allora I^2 denota l’ideale generato dall’insieme dei prodotti di elementi di I . In generale, infatti, questo insieme non è a sua volta un ideale.

Dimostrazione. Consideriamo un sistema di coordinate locali x_1, \dots, x_n in p : non è restrittivo supporre che $x_i(p) = 0$ per ogni $i = 1, \dots, n$. Dalla formula di Taylor (5) segue che ogni $f \in \mathcal{M}(p)$ si può scrivere come

$$f(x) = \sum_{i=1}^n x_i f_i(x) \text{ per certe } f_i \in \mathcal{E}(p) \text{ tali che } f_i(p) = \frac{\partial f}{\partial x_i}(p).$$

Ne segue dunque che $\mathcal{M}(p)$ è generato da x_1, \dots, x_n , e conseguentemente che le coppie $x_i x_j$, al variare di i e j da 1 a n , generano $\mathcal{M}^2(p)$.

Applicando poi la formula di Taylor alle f_i , troviamo che per ogni $i = 1, \dots, n$ si ha

$$f_i(x) = \frac{\partial f_i}{\partial x_i}(p) + \sum_{j=1}^n x_j g_{ij}(x) \text{ per certe } g_{ij} \in \mathcal{E}(p) \text{ tali che } g_{ij}(p) = \frac{\partial f_i}{\partial x_j}(p)$$

ovvero

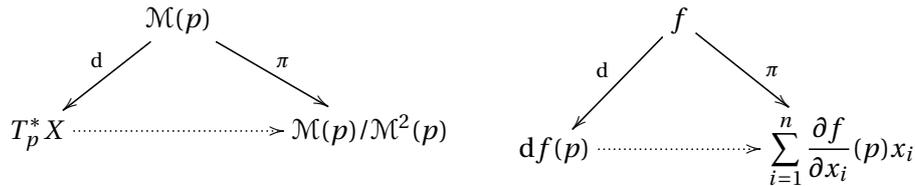
$$f(x) = \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(p) + \xi(x) \text{ dove } \xi(x) = \sum_{i,j=1}^n x_i x_j g_{ij}(x) \in \mathcal{M}^2(p).$$

Osserviamo che tale scrittura è unica: le f_i e le g_{ij} sono univocamente determinate, come si evince dalla dimostrazione del Lemma 1.4.

Se definiamo il *differenziale di de Rham* di f in p (che, a differenza del differenziale definito precedentemente, è un covettore) ponendo

$$df(p) := \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) dx_i,$$

quello che vogliamo fare è rendere commutativo il diagramma



cosa che possiamo fare perché se $f \in \mathcal{M}^2(p)$ allora $\frac{\partial f}{\partial x_i}(p) = 0$ per ogni $i = 1, \dots, n$, e quindi $df(p)$ è nullo. Ci basta associare alla funzione x_i il covettore dx_i . □

Definizione 9.2. Si definisce *fibrato cotangente* alla varietà X l'unione disgiunta di tutti gli spazi cotangenti:

$$T^* X := \bigsqcup_{p \in X} T_p^* X.$$

Anche in questo caso, esiste un'ovvia proiezione $\pi : T^* X \rightarrow X$, $\varphi \in T_p^* X \mapsto p$. Nel definire una struttura differenziabile sul fibrato cotangente, richiediamo che tale proiezione sia C^∞ .

9.2 1-forme differenziali

Vogliamo in questa sezione definire l'analogo di un campo di vettori, ma partendo dal fibrato cotangente: otterremo un campo di covettori.

Definizione 9.3. Siano X una varietà differenziabile di dimensione n e U un aperto di X . Una *1-forma differenziale* è un'applicazione

$$\omega : U \rightarrow \bigsqcup_{p \in U} T_p^* X, \quad p \mapsto \omega(p)$$

che sia una sezione (ovvero $\omega(p) \in T_p^* X$ per ogni $p \in U$) e che sia C^∞ . Denoteremo con $\mathcal{A}^1(U)$ l'insieme delle 1-forme differenziali su U .

Dobbiamo specificare in che senso una forma differenziale possa essere di classe C^∞ . Consideriamo quindi un punto $p \in U$ e un sistema di coordinate locali x_1, \dots, x_n in un intorno $V \subseteq U$ di p . Dato che dx_1, \dots, dx_n è una base di $T_p^* X$, allora esistono delle funzioni scalari $\omega_i : V \rightarrow \mathbb{R}$, $i = 1, \dots, n$, tali che

$$\omega(x) = \sum_{i=1}^n \omega_i(x) dx_i;$$

diremo che ω è C^∞ se lo sono le funzioni ω_i . Da questa definizione segue facilmente che $\mathcal{A}^1(U)$ è un $C^\infty(U)$ -modulo: il prodotto di una forma differenziale per una funzione C^∞ è ancora una forma differenziale.

Una maniera alternativa per definire l'essere C^∞ per una 1-forma differenziale ω è quella di considerarla come un'applicazione dall'insieme dei campi di vettori su U a valori nelle funzioni su U :

$$\omega: \mathcal{X}(U) \rightarrow \{f: U \rightarrow \mathbb{R}\}, \quad (\omega, \eta)(p) = (\omega(p), \eta(p))$$

dove $(\cdot, \cdot): T_p^*X \times T_pX \rightarrow \mathbb{R}$ denota l'accoppiamento di dualità.⁷ Diremo perciò che ω è C^∞ se per ogni $\eta \in \mathcal{X}(U)$ la funzione (ω, η) è C^∞ .

Esercizio 9.1. Dimostrare che le due definizioni di 1-forma differenziale C^∞ sono equivalenti.

Se denotiamo con $\mathcal{A}^0(U)$ l'insieme delle funzioni $C^\infty(U)$, allora per ogni aperto U di X esiste un'applicazione

$$d: \mathcal{A}^0(U) \rightarrow \mathcal{A}^1(U), \quad (df, \eta) := \eta(f) \quad \forall \eta \in \mathcal{X}(U)$$

detta *differenziale di de Rham* di f (la funzione $\eta(f)$ è stata definita nella (10)). Dalla definizione segue immediatamente che df è una 1-forma differenziale C^∞ : $df(p) \in T_p^*X$ per ogni $p \in U$ e per ogni $\eta \in \mathcal{X}(U)$ si ha che $(df(p), \eta(p))$ è una funzione $C^\infty(U)$.

Fissiamo ora un sistema di coordinate x_1, \dots, x_n in un intorno di p ; per definizione di differenziale si ha che

$$\left(df, \frac{\partial}{\partial x_i} \right) = \frac{\partial f}{\partial x_i}.$$

Del resto

$$\left(\sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j, \frac{\partial}{\partial x_i} \right) = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \left(dx_j, \frac{\partial}{\partial x_i} \right) = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \delta_i^j = \frac{\partial f}{\partial x_i}.$$

Dato che $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1, \dots, n}$ è una base per T_pX si ha che

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j$$

che è la definizione di differenziale di de Rham usata per dimostrare il Lemma 9.1. In particolare, se $f = x_i$ è una delle coordinate locali, allora

$$dx_i = \sum_{j=1}^n \frac{\partial x_i}{\partial x_j} dx_j = \sum_{j=1}^n \delta_i^j dx_j = dx_i (!)$$

ovvero il differenziale di de Rham della i -esima coordinata locale corrisponde all' i -esimo elemento della base duale di $\left\{ \frac{\partial}{\partial x_j} \right\}_{j=1, \dots, n}$.

Il lettore è invitato a verificare che il differenziale di de Rham è una derivazione su $\mathcal{A}^0(U)$: è \mathbb{R} -lineare e soddisfa la regola di Leibniz

$$d(fg) = f dg + g df.$$

Consideriamo ora un'applicazione differenziabile fra varietà $F: X \rightarrow Y$, dove X ha dimensione n e Y ha dimensione m . Come sappiamo, se V è un aperto di Y è ben definito il morfismo di anelli $F^*: \mathcal{A}^0(V) \rightarrow \mathcal{A}^0(F^{-1}(V))$; vogliamo estendere questa applicazione alle 1-forme, in maniera da rendere commutativo il diagramma

$$\begin{array}{ccc} \mathcal{A}^0(V) & \xrightarrow{F^*} & \mathcal{A}^0(F^{-1}(V)) \\ \downarrow d & & \downarrow d \\ \mathcal{A}^1(V) & \xrightarrow{F^*} & \mathcal{A}^1(F^{-1}(V)) \end{array} \quad (19)$$

⁷Dato un \mathbb{K} -spazio vettoriale V , l'accoppiamento di dualità è l'applicazione $(\cdot, \cdot): V^* \times V \rightarrow \mathbb{K}$, $(\varphi, v) := \varphi(v)$.

Al solito, definiremo questa applicazione $F^* : \mathcal{A}^1(V) \rightarrow \mathcal{A}^1(F^{-1}(V))$ punto per punto. Sia dunque p un punto di $F^{-1}(V)$ e $q = F(p) \in V$. L'applicazione lineare $DF_p : T_p X \rightarrow T_q Y$ porta vettori tangenti in vettori tangenti; passando ai duali, l'applicazione $DF_p^T : T_q^* Y \rightarrow T_p^* X$ porta covettori in covettori. Dato quindi un covettore $\omega \in T_q^* Y$, definiamo $F^* \omega := DF_p^T(\omega) = \omega \circ DF_p$. Estendiamo questa definizione alle 1-forme: data $\omega \in \mathcal{A}^1(V)$, per ogni campo di vettori $\eta \in \mathcal{X}(F^{-1}(V))$ definiamo

$$(F^* \omega, \eta)(p) := (\omega(q), F_* \eta(q)) = (\omega(q), DF_p(\eta)(q)).$$

Verifichiamo che la forma differenziale sul fibrato cotangente di X così definita è effettivamente C^∞ , e che commuta con i differenziali di de Rham. A questo scopo, prendiamo un sistema di coordinate locali y_1, \dots, y_m in un intorno di q ; allora, la forma $\omega \in \mathcal{A}^1(V)$ si può scrivere come

$$\omega = \sum_{j=1}^m \omega_j dy_j, \quad \omega_j \in \mathcal{A}^0(V).$$

Applicando F^* a questa uguaglianza e sfruttando la linearità, si ha

$$F^* \omega(p) = \sum_{j=1}^m F^*(\omega_j)(q) F^* dy_j, \quad F^*(\omega_j) \in \mathcal{A}^0(F^{-1}(V))$$

dove il primo dei due F^* nella sommatoria è quello che opera sulle funzioni C^∞ (ovvero sulle 0-forme). Abbiamo allora che $F^* \omega$ è una combinazione lineare a coefficienti C^∞ di $F^* dy_j$; se possiamo vedere queste ultime come differenziali di funzioni C^∞ , allora usando queste come coordinate locali in p avremmo verificato che $F^* \omega$ è una 1-forma differenziale C^∞ .

Più in generale, dimostriamo che per ogni funzione $f \in C^\infty(V)$ si ha che

$$F^* df = d(F^*(f)),$$

ovvero che il diagramma (19) commuta; avremo la nostra tesi applicando questa formula a $f = y_j$. Si ha dunque, per ogni $\eta \in \mathcal{X}(F^{-1}(V))$,

$$\begin{aligned} (F^* df, \eta) &= (DF_p^T(df), \eta) \\ &= (df, DF_p(\eta)) \quad (\text{per definizione di } DF_p^T) \\ &= DF_p(\eta)(f) \quad (\text{per definizione di differenziale di de Rham } df) \\ &= \eta(F^*(f)) \quad (\text{per definizione di differenziale } DF_p) \\ &= (d(F^*(f)), \eta), \end{aligned}$$

come volevamo dimostrare.

L'applicazione $F^* : \mathcal{A}^1(V) \rightarrow \mathcal{A}^1(F^{-1}(V))$ prende il nome di *pullback di forme*, perché “porta indietro” le forme differenziali dal fibrato cotangente di Y a quello di X . Osserviamo che, mentre l'applicazione di correlazione fra campi di vettori $F_* : \mathcal{X}(U) \rightarrow \mathcal{X}(F(U))$ non è sempre ben definita, il *pullback* di forme (ovvero il suo analogo F^* con i covettori) lo è.

9.3 k -forme differenziali

Sia X una varietà differenziabile di dimensione n , e sia p un punto di X . Lo spazio cotangente $T_p^* X$ è uno spazio vettoriale; perciò possiamo considerarne il prodotto esterno k -esimo $\wedge^k T_p^* X$. In particolare, se x_1, \dots, x_n sono coordinate locali in p , allora

$$\{dx_{i_1}, \dots, dx_{i_k}\}_{i_1 < \dots < i_k}$$

è una base per $\wedge^k T_p^* X$ (ognuno degli i_j può essere scelto da 1 a n).

Definizione 9.4. Sia U un aperto di X . Una k -forma differenziale è un'applicazione

$$\omega : U \rightarrow \bigsqcup_{p \in U} \bigwedge^k T_p^* X, \quad p \mapsto \omega(p)$$

tale che $\omega(p) \in \bigwedge^k T_p^* X$ per ogni $p \in U$ e che sia C^∞ . Chiameremo k il *grado* di ω . Denoteremo con $\mathcal{A}^k(U)$ l'insieme delle k -forme differenziali su U .

In coordinate, abbiamo che una k -forma differenziale ha una scrittura del tipo

$$\omega = \sum_{i_1 < \dots < i_k} \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}, \quad \omega_{i_1, \dots, i_k} \in \mathcal{A}^0(U)$$

ovvero è una combinazione lineare a coefficienti in $C^\infty(U)$ degli elementi della base citata sopra. Per utilizzare una notazione più compatta, nel seguito useremo il multiindice I per indicare l'insieme dei pedici $i_1 < \dots < i_k$, $i_j \in \{1, \dots, n\}$, e scriveremo inoltre

$$\omega_{i_1, \dots, i_k} = \omega_I, \quad dx_{i_1} \wedge \dots \wedge dx_{i_k} = dx_I.$$

Osservazione 9.1. Questa definizione di k -forma differenziale coincide con quella già data nei casi $k = 0$ e $k = 1$, dato che

$$\bigwedge^0 T_p^* X = \mathbb{R}, \quad \bigwedge^1 T_p^* X = T_p^* X.$$

Il prodotto dell'algebra esterna $\bigwedge^a T_p^* X \wedge \bigwedge^b T_p^* X \rightarrow \bigwedge^{a+b} T_p^* X$ si può estendere puntualmente anche alle k -forme:

$$\mathcal{A}^r(U) \wedge \mathcal{A}^s(U) \rightarrow \mathcal{A}^{r+s}(U), \quad (\omega \wedge \psi)(x) := \omega(x) \wedge \psi(x).$$

Dalla (17) discende facilmente che

$$\omega \wedge \psi = (-1)^{rs} \psi \wedge \omega \quad \forall \omega \in \mathcal{A}^r(U), \psi \in \mathcal{A}^s(U).$$

È anche possibile estendere il differenziale di de Rham $d: \mathcal{A}^k(U) \rightarrow \mathcal{A}^{k+1}(U)$, ponendo

$$d\omega = d\left(\sum_I \omega_I dx_I\right) := \sum_I d\omega_I \wedge dx_I \quad (20)$$

dove $d\omega_I$ è il differenziale di de Rham di $\omega_I \in \mathcal{A}^0(U)$ precedentemente definito. L'applicazione $d: \mathcal{A}^k(U) \rightarrow \mathcal{A}^{k+1}(U)$ così definita risulta essere \mathbb{R} -lineare e soddisfa la *regola di Leibniz graduata*

$$d(\omega \wedge \psi) = (d\omega) \wedge \psi + (-1)^r \omega \wedge (d\psi), \quad \omega \in \mathcal{A}^r(U), \psi \in \mathcal{A}^s(U). \quad (21)$$

È sufficiente (per la linearità del differenziale) verificare questa identità sulle forme del tipo $\omega = f dx_I$, $\psi = g dx_J$, dove f e g sono funzioni $C^\infty(U)$, I è un multiindice di grado r e J un multiindice di grado s . Si ha infatti in questo caso

$$\begin{aligned} d(\omega \wedge \psi) &= d(fg) \wedge dx_I \wedge dx_J \\ &= g df \wedge dx_I \wedge dx_J + f dg \wedge dx_I \wedge dx_J, \\ (d\omega) \wedge \psi &= df \wedge dx_I \wedge (g dx_J) \\ &= g df \wedge dx_I \wedge dx_J, \\ \omega \wedge (d\psi) &= f dx_I \wedge dg \wedge dx_J \\ &= (-1)^k f dg \wedge dx_I \wedge dx_J, \end{aligned}$$

dove abbiamo sfruttato più volte il fatto che $d^2 = 0$ e ancora la (17) nell'ultima uguaglianza.

Il fatto che $d^2 = d \circ d : \mathcal{A}^k(U) \rightarrow \mathcal{A}^{k+2}(U)$ è identicamente nullo è molto semplice da verificare. Se infatti ω è una k -forma del tipo $\omega = f dx_I$, allora

$$d\omega = df \wedge dx_I = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i \wedge dx_I$$

e quindi

$$\begin{aligned} d(d\omega) &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i \wedge dx_I \\ &= \sum_{i < j} \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i \wedge dx_I + \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j \wedge dx_I \\ &= \sum_{i < j} \left(\frac{\partial^2 f}{\partial x_j \partial x_i} - \frac{\partial^2 f}{\partial x_i \partial x_j} \right) dx_j \wedge dx_i \wedge dx_I = 0. \end{aligned}$$

Infatti, per $i = j$ si ha $dx_i \wedge dx_i \wedge dx_I = 0$, e quindi possiamo sommare solo sulle coppie $i < j$; inoltre, scambiando l'ordine di dx_i e dx_j nel prodotto esterno, tale prodotto cambia segno. L'ultima uguaglianza segue poi dal noto teorema di Schwarz sull'inversione dell'ordine di derivazione.

Teorema 9.1. *L'operatore $d : \mathcal{A}^k(U) \rightarrow \mathcal{A}^{k+1}(U)$ definito in (20) è l'unico operatore tale che*

- i. d è \mathbb{R} -lineare;
- ii. d soddisfa la regola di Leibniz graduata (21);
- iii. $d^2 = 0$;
- iv. d estende il differenziale di de Rham $d : \mathcal{A}^0(U) \rightarrow \mathcal{A}^1(U)$ precedentemente definito.

Come conseguenza di questo teorema, otteniamo che la definizione (20) di differenziale di de Rham è in realtà indipendente dal sistema di coordinate scelto.

Dimostrazione. L'esistenza è garantita dal fatto che il differenziale di de Rham soddisfa queste proprietà. Per dimostrare l'unicità, sia $\tilde{d} : \mathcal{A}^k(U) \rightarrow \mathcal{A}^{k+1}(U)$ un altro operatore che le soddisfa. Supponiamo che $\omega \in \mathcal{A}^k(U)$ si scriva in coordinate locali come

$$\omega = \sum_{i_1 < \dots < i_k} \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} = \sum_I \omega_I dx_I, \quad \omega_{i_1, \dots, i_k} \in \mathcal{A}^0(U).$$

Allora

$$\begin{aligned} \tilde{d}\omega &= \sum_{i_1 < \dots < i_k} \tilde{d}(\omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}) \quad (\text{per la } \mathbb{R}\text{-linearità}) \\ &= \sum_I \tilde{d}\omega_I \wedge dx_I + \sum_{j=1}^k \sum_{i_1 < \dots < i_k} \pm \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge \tilde{d}(dx_{i_j}) \wedge \dots \wedge dx_{i_k} \quad (\text{per Leibniz}) \\ &= \sum_I d\omega_I \wedge dx_I + \sum_{j=1}^k \sum_{i_1 < \dots < i_k} \pm \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge \tilde{d}(dx_{i_j}) \wedge \dots \wedge dx_{i_k} \\ &\hspace{15em} (\text{perché } \tilde{d} \text{ coincide con } d \text{ sulle 0-forme}) \\ &= \sum_I d\omega_I \wedge dx_I \quad (\text{perché } \tilde{d}^2 = 0) \end{aligned}$$

e quindi l'operatore \tilde{d} coincide con il differenziale di de Rham già definito. \square

Data un'applicazione differenziabile $F: X \rightarrow Y$, è possibile definire il *pullback* di k -forme in maniera del tutto analoga a quanto abbiamo già fatto per le 1-forme. Se y_1, \dots, y_m sono coordinate locali nell'aperto V di Y , intorno del punto q (ovviamente intendiamo che m è la dimensione di Y come varietà), allora una k -forma su V è della forma

$$\omega = \sum_{j_1 < \dots < j_k} \omega_{j_1, \dots, j_k} dy_{j_1} \wedge \dots \wedge dy_{j_k}, \quad \omega_{j_1, \dots, j_k} \in \mathcal{A}^0(V);$$

definiremo dunque la forma $F^* \omega \in \mathcal{A}^k(F^{-1}(V))$ come

$$\begin{aligned} F^* \omega &= \sum_{j_1 < \dots < j_k} F^*(\omega_{j_1, \dots, j_k}) F^*(dy_{j_1} \wedge \dots \wedge dy_{j_k}) \\ &= \sum_{j_1 < \dots < j_k} F^*(\omega_{j_1, \dots, j_k}) F^* dy_{j_1} \wedge \dots \wedge F^* dy_{j_k} \\ &= \sum_{j_1 < \dots < j_k} F^*(\omega_{j_1, \dots, j_k}) d(F^*(y_{j_1})) \wedge \dots \wedge d(F^*(y_{j_k})), \end{aligned}$$

dove nell'ultima uguaglianza abbiamo sfruttato implicitamente l'Osservazione 8.4.

Possiamo lavorare anche in maniera più intrinseca (senza far ricorso alle coordinate). Sia dunque p un punto di $F^{-1}(V)$ tale che $F(p) = q$. Quello che vogliamo è che per ogni k -upla di vettori η_1, \dots, η_k in $T_p X$ si abbia

$$\eta_1 \lrcorner \dots \lrcorner \eta_k \lrcorner F^* \omega = DF_p(\eta_1) \lrcorner \dots \lrcorner DF_p(\eta_k) \lrcorner \omega;$$

perciò F^* è il trasposto dell'applicazione $\wedge^k DF_p: \wedge^k T_p X \rightarrow \wedge^k T_q Y$.