

# Una pessima guida all'Analisi II

Giorgio Mangioni, with special thanks to Pietro Majer

2 gennaio 2023

## Introduzione

Oh sh\*t, here we go again.

---

*CJ di GTA San Andreas.*

Eh sì, anche quest'anno ho deciso di mettermi a scrivere le dispense del corso di Analisi II del professor Pietro Majer, dell'università di Pisa (e modestamente ex Normalista). Sarà un lavoro lungo e faticoso, pertanto ogni aiuto è calorosamente accolto. Per questioni di vitale importanza potete contattarmi a [g.mangioni@unipi.it](mailto:g.mangioni@unipi.it).

In alcuni punti integrerò quanto visto a lezione con un po' del mio, ad esempio nella risoluzione di alcuni esercizi dati in classe o nel provare teoremi di cui possiedo una dimostrazione mirabile (o almeno, più di quella proposta). In ogni caso, più o meno tutto quello che si è fatto in classe c'è, eccetto alcune oscenità irrilevanti su cui ho voluto glissare.

L'anno scorso (vedasi [Mangioni, 2019]) avevo pubblicato un testo impreciso e parziale, privo di immagini e di citazioni sufficientemente inadeguate. Cercherò di fare ammenda, nei limiti del possibile, in modo da alleviare (almeno in parte) il tedio tipico dello studio della materia di cui in oggetto. Spero vi piaccia, e se non vi piace scrivetevele voi le dispense.

Firmato: *Uno che fa l'eroe per hobby.*

# Indice

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Strumenti utili (?) e riassunto delle puntate precedenti</b>  | <b>7</b>  |
| 1.1      | Induzione transfinita . . . . .                                  | 7         |
| 1.2      | Semicontinuità . . . . .   | 8         |
| 1.3      | Proprietà di Urysohn . . . . .                                   | 9         |
| 1.4      | Compattezza Tuttingusti+1 . . . . .                              | 9         |
| 1.4.1    | Estensione del teorema di Heine-Cantor . . . . .                 | 10        |
| 1.4.2    | Lipschitzianità su un compatto . . . . .                         | 11        |
| 1.5      | Differenziale del determinante . . . . .                         | 11        |
| <b>2</b> | <b>Spazi vettoriali normati</b>                                  | <b>13</b> |
| 2.1      | Norma e topologia indotta . . . . .                              | 13        |
| 2.2      | Spazi normati e prodotti scalari . . . . .                       | 14        |
| 2.3      | Spazi $l^p$ . . . . .  | 14        |
| 2.3.1    | Disuguaglianza di Young . . . . .                                | 15        |
| 2.3.2    | Disuguaglianza di Hölder . . . . .                               | 15        |
| 2.3.3    | Definizione alternativa per la norma . . . . .                   | 16        |
| 2.3.4    | Disuguaglianza di Minkowski . . . . .                            | 16        |
| 2.4      | Equivalenza di norme . . . . .                                   | 17        |
| 2.4.1    | Le norme su $\mathbb{R}^n$ sono equivalenti . . . . .            | 17        |
| 2.5      | Spazi vettoriali complessi . . . . .                             | 18        |
| 2.5.1    | Relazione con gli $\mathbb{R}$ -spazi vettoriali . . . . .       | 18        |
| 2.5.2    | Norma complessa . . . . .  | 18        |
| 2.5.3    | Prodotto hermitiano . . . . .                                    | 19        |
| <b>3</b> | <b>Completezza</b>   | <b>21</b> |
| 3.1      | Spazi completi . . . . .   | 21        |
| 3.1.1    | Completezza e chiusura . . . . .                                 | 21        |
| 3.2      | Estensione di mappe UC a valori in spazi completi . . . . .      | 21        |
| 3.3      | Spazi di Banach . . . . .  | 22        |
| 3.3.1    | Connessione degli aperti . . . . .                               | 22        |
| 3.3.2    | Quozienti di spazi di Banach . . . . .                           | 23        |
| 3.3.3    | Spazi di funzioni completi . . . . .                             | 24        |
| 3.3.4    | Ogni spazio metrico si immerge nelle funzioni continue . . . . . | 25        |
| 3.4      | Completamento di uno spazio metrico . . . . .                    | 25        |
| 3.5      | Teorema delle contrazioni . . . . .                              | 26        |
| 3.5.1    | Variante per spazi compatti . . . . .                            | 26        |
| 3.6      | Teorema di Baire . . . . .                                       | 26        |
| 3.6.1    | Insieme di continuità di una funzione . . . . .                  | 27        |
| <b>4</b> | <b>Operatori lineari</b>   | <b>29</b> |
| 4.1      | Continuità delle applicazioni lineari . . . . .                  | 29        |
| 4.1.1    | Duale topologico . . . . .                                       | 30        |
| 4.1.2    | Applicazioni bilineari . . . . .                                 | 30        |
| 4.2      | Completezza delle funzioni lineari continue . . . . .            | 30        |
| 4.3      | Serie di Neumann . . . . .                                       | 31        |

|          |   |           |
|----------|---|-----------|
| 4.4      | Perturbazioni lipschitziane dell'identità . . . . .           | 31        |
| 4.5      | Gli isomorfismi lineari sono un aperto . . . . .              | 32        |
| 4.6      | $GL(E)$ come gruppo di Lie . . . . .                          | 32        |
| 4.7      | Isometrie affini . . . . .                                    | 33        |
| <b>5</b> | <b>Topologia e convergenza in spazi di funzioni</b>           | <b>35</b> |
| 5.1      | Criteri di convergenza uniforme . . . . .                     | 35        |
| 5.1.1    | Teorema di Dini . . . . .                                     | 35        |
| 5.2      | Approssimazioni polinomiali . . . . .                         | 35        |
| 5.2.1    | Strani operatori lineari . . . . .                            | 35        |
| 5.2.2    | Polinomi di Bernstein . . . . .                               | 36        |
| 5.2.3    | Proprietà dei polinomi di Bernstein . . . . .                 | 37        |
| 5.2.4    | I polinomi sono densi nelle funzioni continue . . . . .       | 37        |
| 5.2.5    | Caso vettoriale . . . . .                                     | 38        |
| 5.3      | Stoned . . . . .  | 40        |
| 5.3.1    | Teorema di Stone . . . . .                                    | 40        |
| 5.3.2    | Classificazione delle algebre chiuse . . . . .                | 42        |
| 5.3.3    | Insight algebrico . . . . .                                   | 42        |
| 5.4      | Teorema di Ascoli-Arzelà . . . . .                            | 42        |
| 5.5      | Come non arrivare in ritardo . . . . .                        | 43        |
| 5.6      | Totale limitatezza di spazi funzionali . . . . .              | 44        |
| 5.7      | Estensione di mappe UC con modulo concavo . . . . .           | 44        |
| <b>6</b> | <b>Calcolo differenziale in spazi di Banach</b>               | <b>47</b> |
| 6.1      | Definizione di "o piccolo" . . . . .                          | 47        |
| 6.2      | Differenziale di Fréchet . . . . .                            | 47        |
| 6.2.1    | Differenziali particolari . . . . .                           | 47        |
| 6.2.2    | Derivata direzionale . . . . .                                | 48        |
| 6.2.3    | Differenziabilità in spazi complessi . . . . .                | 48        |
| 6.3      | Differenziale secondo . . . . .                               | 49        |
| 6.3.1    | Matrice Hessiana . . . . .                                    | 50        |
| 6.4      | Polinomio di Taylor . . . . .                                 | 50        |
| 6.4.1    | Derivata rispetto al centro dello sviluppo . . . . .          | 51        |
| 6.4.2    | Resto di Peano . . . . .                                      | 51        |
| 6.4.3    | Resto di Lagrange . . . . .                                   | 52        |
| 6.5      | Massimi e minimi liberi . . . . .                             | 53        |
| 6.6      | Massimi e minimi vincolati . . . . .                          | 54        |
| 6.7      | Teorema del differenziale totale . . . . .                    | 55        |
| 6.8      | Derivate parziali miste . . . . .                             | 57        |
| 6.9      | Lipschitz in più variabili, ma anche no . . . . .             | 58        |
| 6.10     | Invertibilità locale . . . . .                                | 58        |
| 6.11     | Teorema della funzione implicita . . . . .                    | 59        |
| 6.12     | Passaggio al limite sotto il segno di differenziale . . . . . | 59        |
| 6.13     | Derivare l'integrando . . . . .                               | 60        |
| 6.14     | Funzioni di classe $C^k$ . . . . .                            | 61        |
| 6.15     | Funzioni omogenee di Eulero . . . . .                         | 62        |
| <b>7</b> | <b>1-forme</b>  | <b>63</b> |
| 7.1      | Forme chiuse e esatte . . . . .                               | 64        |
| 7.1.1    | Costruzione di una primitiva per aperti stellati . . . . .    | 67        |
| 7.2      | Pullback di forme differenziali . . . . .                     | 67        |

|           |   |            |
|-----------|---|------------|
| <b>8</b>  | <b>Teoria della misura</b>  | <b>69</b>  |
| 8.1       | Algebre e anelli . . . . .  | 69         |
| 8.2       | Misura . . . . .  | 69         |
| 8.3       | Misura esterna ed estensione di misure . . . . .                            | 70         |
| 8.3.1     | Completamento di misure . . . . .   | 73         |
| 8.4       | Misura prodotto . . . . .   | 75         |
| 8.5       | Sistemi di Dynkin e $\pi$ -sistemi . . . . .                                | 75         |
| 8.5.1     | Coincidenza di misure . . . . .   | 76         |
| 8.6       | Costruzione della misura di Lebesgue . . . . .                              | 77         |
| 8.6.1     | Proprietà della misura di Lebesgue . . . . .                                | 78         |
| 8.7       | Misure boreliane su $[0, 1]$ . . . . .                                      | 80         |
| 8.7.1     | Bene. . . . .   | 80         |
| 8.7.2     | . . . ma non benissimo . . . . .  | 81         |
| <b>9</b>  | <b>Teoria dell'integrazione</b>   | <b>83</b>  |
| 9.1       | Funzioni misurabili . . . . .   | 83         |
| 9.1.1     | Funzioni misurabili rispetto al completamento . . . . .                     | 83         |
| 9.1.2     | Trasporto di misure . . . . .   | 84         |
| 9.2       | Costruzione dell'integrale . . . . .  | 84         |
| 9.2.1     | Lebesgue contro Riemann . . . . .   | 88         |
| 9.3       | Convergenza di integrali . . . . .  | 88         |
| 9.3.1     | Integrazione per serie . . . . .  | 90         |
| 9.4       | Integrali doppi . . . . .   | 91         |
| 9.4.1     | Insiemi trascurabili nel prodotto . . . . .                                 | 93         |
| 9.4.2     | Affettare gli integrali . . . . .   | 93         |
| 9.5       | Densità in $L^1$ . . . . .  | 93         |
| 9.6       | Continuità uniforme in $L^1$ . . . . .                                      | 94         |
| 9.6.1     | Approssimazioni continue . . . . .  | 95         |
| 9.7       | Compattezza in $L^1$ (incompleto) . . . . .                                 | 97         |
| 9.8       | Cambio di variabili . . . . .   | 98         |
| 9.8.1     | Immagine continua di insiemi misurabili . . . . .                           | 98         |
| 9.8.2     | Comportamento della misura di Lebesgue sotto applicazioni lineari . . . . . | 99         |
| 9.8.3     | Mettiamo assieme i pezzi . . . . .  | 99         |
| 9.8.4     | Radon-Nikodym . . . . .   | 102        |
| 9.8.5     | Integrazione in coordinate polari . . . . .                                 | 103        |
| 9.9       | Continuità di operatori di sostituzione . . . . .                           | 103        |
| 9.9.1     | Call of DUTIS . . . . .   | 104        |
| 9.10      | L'integrale come funzionale continuo (e viceversa) . . . . .                | 105        |
| 9.10.1    | Problema dei momenti di Hausdorff . . . . .                                 | 107        |
| <b>10</b> | <b>(Gran) Varietà</b>   | <b>109</b> |
| 10.1      | Varietà in $\mathbb{R}^n$ . . . . .   | 109        |
| 10.2      | Generalizzazione ai Banach . . . . .  | 109        |
| 10.3      | Come spammare sottovarietà . . . . .  | 110        |
| 10.3.1    | Grafici di funzioni . . . . .   | 110        |
| 10.3.2    | Controimmagini di valori regolari . . . . .                                 | 110        |
| 10.3.3    | Differenziabilità . . . . .   | 111        |
| 10.4      | Punti estremanti su sottovarietà . . . . .                                  | 112        |
| 10.5      | Misura su sottovarietà . . . . .  | 112        |
| 10.5.1    | Lemma di Sard per varietà . . . . .   | 113        |
| 10.5.2    | Cambio di variabile . . . . .   | 114        |
| 10.5.3    | Misura su grafici . . . . .   | 114        |
| 10.5.4    | Insight geometrico, senza le dovute prove . . . . .                         | 115        |
| 10.6      | Teorema della divergenza . . . . .  | 115        |

|           |   |            |
|-----------|---|------------|
| <b>11</b> | <b>Equazioni differenziali ordinarie lineari (EDOL)</b>                 | <b>119</b> |
| 11.1      | Esponenziale di matrici . . . . .                                       | 119        |
| 11.1.1    | Proprietà dell'esponenziale . . . . .                                   | 120        |
| 11.1.2    | Immagine della mappa esponenziale . . . . .                             | 121        |
| 11.2      | Sistemi di EDO a coefficienti costanti . . . . .                        | 121        |
| 11.3      | Stabilità delle soluzioni omogenee . . . . .                            | 122        |
| 11.4      | Sistema lineare generico . . . . .                                      | 123        |
| 11.4.1    | Unicità . . . . .   | 124        |
| 11.4.2    | Matrice di Wronski . . . . .  | 124        |
| 11.4.3    | Proprietà della matrice Wronskiana . . . . .                            | 125        |
| 11.4.4    | Soluzione dell'equazione generale . . . . .                             | 127        |
| 11.5      | Equazioni lineari periodiche (teoria di Floquet) . . . . .              | 127        |
| 11.6      | Stabilità dei sistemi periodici . . . . .                               | 128        |
| <b>12</b> | <b>Equazioni differenziali non lineari</b>                              | <b>131</b> |
| 12.1      | Teorema di Cauchy-Lipschitz(-Picard-Lindelöf-compagnia bella) . . . . . | 131        |
| 12.2      | Equazioni autonome . . . . .  | 133        |
| 12.3      | Esistenza locale per funzioni continue . . . . .                        | 133        |
| 12.4      | Estensione di soluzioni . . . . .                                       | 134        |
| 12.5      | Criterio di esistenza globale . . . . .                                 | 136        |
| 12.6      | Dipendenza continua dai dati iniziali . . . . .                         | 136        |
| <b>13</b> | <b>Metodi analitici in topologia</b>                                    | <b>141</b> |
| 13.1      | Teorema di non retrazione e teorema di Brouwer . . . . .                | 141        |
| 13.1.1    | Teorema di Perron-Frobenius . . . . .                                   | 143        |
| <b>14</b> | <b>Delirio e voli pindarici</b>   | <b>145</b> |
| 14.1      | Proiettili vaganti . . . . .  | 145        |
| 14.1.1    | Generalizzazione del caso precedente . . . . .                          | 146        |
| 14.2      | Coniugio di contrazioni . . . . .                                       | 146        |
| 14.3      | Derivate del Dini . . . . .   | 147        |
| 14.4      | Luogo di zeri di una derivata . . . . .                                 | 147        |
| 14.5      | Derivate di Pompeiu . . . . .   | 148        |
| 14.5.1    | Ma non ne bastava una sola? . . . . .                                   | 149        |
| 14.6      | Formula di Stirling, rivisitata e corretta . . . . .                    | 149        |
| 14.7      | Quoziente di Rayleigh . . . . .   | 150        |
| 14.8      | Equazioni ai commutatori . . . . .                                      | 151        |
| 14.8.1    | Proiettori e commutatori . . . . .                                      | 151        |

# Capitolo 1

## Strumenti utili (?) e riassunto delle puntate precedenti

### 1.1 Induzione transfinita

Ricordiamo la definizione di insieme ordinato.

**Definizione 1** (Insieme ordinato). *Sia  $X$  un insieme. Una relazione  $<$  su  $X$  si dice d'ordine (parziale) se è*

- *Antiriflessiva, cioè  $x \not< x$ ;*
- *Antisimmetrica, cioè se  $x < y$ ,  $y \not< x$ ;*
- *Transitiva, cioè  $x < y < z \implies x < z$ .*

**Definizione 2** (Ordine totale). *Due punti  $x, y \in X$  si dicono confrontabili se  $x < y$  oppure  $y < x$ . Un ordine si dice totale se ogni coppia di punti distinti è confrontabile.*

**Definizione 3** (Buon ordine). *Un insieme ordinato  $(X, <)$  è ben ordinato se ogni sottoinsieme non vuoto ammette minimo.*

Si osservi che un buon ordine è totale, perché comunque si scelgano  $x \neq y$  esiste il minimo di  $\{x, y\}$  e dunque i due punti sono confrontabili.

Gli elementi di un buon ordine si possono classificare in tre categorie, a seconda dell'insieme di coloro che li precedono. Dato  $x \in X$ , esso deve essere uno solo dei seguenti:

- il minimo di  $X$ ;
- il successore di  $y$  se quest'ultimo è il massimo degli elementi  $\{z \in X \text{ t.c. } z < x\}$ ; in tal caso indicheremo  $x$  con  $y + 1$ ;
- un elemento limite se l'insieme degli elementi più piccoli  $\{z \in X \text{ t.c. } z < x\}$  non ha un massimo; diremo allora che  $x = \lim_{z < x} z$ .

Per i buoni ordini vale un potenziamento del teorema di induzione dei numeri naturali:

**Teorema 1.1.1** (Induzione transfinita). *Sia  $(X, <)$  un insieme ben ordinato e sia  $P(x)$  un predicato su  $X$ , ossia un'affermazione che dipende da  $x \in X$ . Se vale che*

1.  *$P(x_0)$  è vera, dove  $x_0 = \min X$ ;*
2. *Se  $P(y)$  è vera, allora anche  $P(y + 1)$  è vera;*
3. *Se  $x = \lim_{z < x} z$  e  $P(z)$  è vera per ogni  $z < x$  allora  $P(x)$  è vera;*

*allora  $P(x)$  è vera per ogni  $x \in X$ .*

*Dimostrazione.* Sia  $S = \{x \in X \text{ t.c. } P(x) \text{ è falsa}\}$ . Osserviamo che  $S$  non ha minimo. Infatti, per ipotesi il suo minimo non può essere  $x_0$ . Inoltre non può essere un successore, perché se lo fosse la proprietà  $P$  varrebbe per il suo precedente e ciò violerebbe l'ipotesi 2., né può essere un limite, dato che  $P$  sarebbe vera per ogni  $z$  più piccolo e ciò contraddirebbe l'ipotesi 3.. Poiché  $S$  non ha minimo deve essere vuoto, poiché  $X$  è ben ordinato.  $\square$

## 1.2 Semicontinuità

**Definizione 4** (Semicontinuità inferiore). *Una funzione  $f : (X, \tau) \rightarrow \mathbb{R}$  si dice semicontinua inferiormente nel punto  $x_0$  se  $\liminf_{x \rightarrow x_0} f(x) \geq f(x_0)$ , ossia se*

$$\forall \varepsilon > 0 \exists U \text{ intorno di } x_0 \text{ t.c. } x \in U \implies f(x) > f(x_0) - \varepsilon$$

**Lemma 1.2.1.** *Sono equivalenti:*

1.  $f$  è semicontinua inferiormente in ogni punto;
2. Per ogni  $c \in \mathbb{R}$ , l'insieme  $f^{-1}(]c, +\infty[)$  è aperto. In altri termini,  $f$  è continua da  $X$  nella retta estesa  $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$  con la topologia inferiore  $\tau_{inf} = \{]c, +\infty[, \emptyset, \overline{\mathbb{R}}\}$
3. Il sottografico  $\{(x, c) \in X \times \mathbb{R} \text{ t.c. } f(x) > c\}$  è aperto;
4. Se  $X$  è primo numerabile, per ogni  $x_0$  e per ogni  $x_k \rightarrow x_0$  vale  $\liminf_{k \rightarrow +\infty} f(x_k) \geq f(x_0)$ .

*Dimostrazione.* Mostriamo le varie implicazioni.

- (1.  $\implies$  4.) Ovvio.
- (4.  $\implies$  1.) Procedendo per assurdo si costruisce una successione che tende a  $x_0$  ma il cui  $\liminf$  è minore di  $f(x_0)$ .
- (1.  $\implies$  2.) Se  $f(x_0) > c$  esiste un  $\varepsilon$  per cui  $f(x_0) - c > \varepsilon$ . Per semicontinuità inferiore esiste un  $U$  aperto tale che, se  $x \in U$ , allora  $f(x) > f(x_0) - \varepsilon > c$ . Allora  $U \subseteq f^{-1}(]c, +\infty[)$ .
- (2.  $\implies$  3.) Mostriamo che il sottografico è intorno di ogni suo punto. Sia  $(x_0, c)$  con  $f(x_0) > c$ . Allora fissato  $\varepsilon$  tale che  $f(x_0) > c + \varepsilon$ , l'insieme  $f^{-1}(]c + \varepsilon, +\infty[) \times (c - \varepsilon, c + \varepsilon)$  è aperto, contiene  $(x_0, c)$  ed è totalmente contenuto nel sottografico.
- (3.  $\implies$  1.) Per ogni  $\varepsilon > 0$ , il punto  $(x_0, f(x_0) - \varepsilon)$  appartiene al sottografico, perciò in particolare esiste un intorno aperto della forma  $U \times (f(x_0) - \varepsilon - \delta, f(x_0) - \varepsilon + \delta)$  contenuto nel sottografico. Dunque per ogni  $x \in U$  si ha che  $f(x) > f(x_0) - \varepsilon + \delta > f(x_0) - \varepsilon$ .

$\square$

La semicontinuità inferiore è una proprietà che si conserva quando si prende l'estremo superiore puntuale di una famiglia di funzioni. Vale infatti che:

**Lemma 1.2.2.** *L'involuppo superiore di una famiglia di funzioni reali semicontinue inferiormente è ancora semicontinuo inferiormente.*

*Dimostrazione.* Sia  $(f_i)_{i \in I}$  una famiglia di funzioni continue  $X \rightarrow (\overline{\mathbb{R}}, \tau_{inf})$ . Per mostrare che anche  $g(x) = \sup_{i \in I} f_i(x)$  è semicontinua basta osservare che per ogni numero reale  $c$  l'insieme

$$\{g(x) > c\} = \bigcup_{i \in I} \{f_i(x) > c\}$$

è aperto perché unione di aperti.  $\square$

Per le funzioni semicontinue vale inoltre il seguente indebolimento del teorema di Weierstrass:

**Teorema 1.2.3.** *Una funzione semicontinua inferiormente su un compatto ammette minimo, e simmetricamente una funzione semicontinua superiormente ammette massimo.*

*Dimostrazione.* Trattiamo solo il caso del minimo. Poiché  $f$  è continua nella topologia inferiore,  $f(X) \subseteq \overline{\mathbb{R}}$  è compatto. Osserviamo ora che i compatti della topologia inferiore sono tutti e soli i sottoinsiemi dotati di minimo. Infatti, se  $Y \subseteq \overline{\mathbb{R}}$  ha minimo, da ogni ricoprimento aperto si può scegliere una semiretta della forma  $]c, +\infty]$  contenente il minimo, che dunque ricopre tutto  $Y$ . Viceversa, se  $Y$  è privo di minimo, sia  $y_n \in Y$  una successione decrescente all'estremo inferiore. Allora dal ricoprimento aperto  $]y_n, +\infty]$  non si può estrarre un sottoricoprimento finito.  $\square$

### 1.3 Proprietà di Urysohn

**Teorema 1.3.1** (Urysohn). *Sia  $(X, \tau)$  uno spazio topologico. Una successione  $x_n$  di elementi di  $X$  converge a un punto  $x$  se e solo se da ogni sottosuccessione si estrae una sotto-sottosuccessione convergente a  $x$ .*

*Dimostrazione.* Una freccia è banale: se  $x_n \rightarrow x$ , ogni sottosuccessione converge a  $x$ . Viceversa, se  $x \not\rightarrow x$ , esiste una sottosuccessione che non sta mai in un intorno di  $x$ , dalla quale non si può tirar fuori nulla di utile.  $\square$

Con lo stesso spirito si ottiene che

**Teorema 1.3.2.** *Una funzione  $f$  tra spazi topologici è continua per successioni in un punto  $x$  se e solo se, per ogni successione  $x_n \rightarrow x$ , esiste una sottosuccessione tale che  $f(x_{n(k)}) \rightarrow f(x)$ .*

### 1.4 Compattezza Tuttigusti+1

Elenchiamo un sacco di proprietà le cui relazioni sono quasi evidenti. Questa lezione è stata una gran noia, ergo mi scuso in anticipo per eventuali imprecisioni o falsità.

**Definizione 5.** *Un punto  $x \in X$  si dice di  $\omega$ -accumulazione per  $Y \subseteq X$  se ogni suo intorno contiene infiniti punti di  $Y$ .*

**Definizione 6.** *Uno spazio topologico è compatto secondo Bolzano-Weierstrass se ogni sottoinsieme infinito ha un punto di  $\omega$ -accumulazione.*

**Definizione 7.** *Uno spazio topologico  $X$  è numerabilmente compatto se ogni ricoprimento aperto numerabile ammette un sottoricoprimento finito.*

**Definizione 8.** *Uno spazio topologico  $X$  ha la proprietà di Lindelöf se ogni ricoprimento aperto ammette un sottoricoprimento numerabile.*

**Definizione 9.** *Una famiglia di chiusi ha la proprietà dell'intersezione finita se ogni sottofamiglia finita ha intersezione non vuota.*

**Lemma 1.4.1.**  *$X$  è compatto se e solo se ogni famiglia di chiusi con intersezione vuota ammette una sottofamiglia finita a intersezione vuota, o equivalentemente se ogni famiglia di chiusi con la proprietà dell'intersezione finita ha intersezione non vuota. Una proprietà analoga vale per la numerabile compattezza.*

**Teorema 1.4.2.** *Vale che:*

1. *Uno spazio topologico è compatto se e solo se è numerabilmente compatto e ha la proprietà di Lindelöf.*
2. *Uno spazio topologico è numerabilmente compatto se e solo se è compatto secondo Bolzano-Weierstrass.*
3. *La sequenziale compattezza implica la compattezza secondo Bolzano-Weierstrass. Le due proprietà sono equivalenti se  $X$  è primo-numerabile.*
4. *Se  $X$  è secondo-numerabile ha la proprietà di Lindelöf. Il viceversa è vero per spazi metrici.*
5. *se  $X$  è metrico, la compattezza sequenziale implica la secondo-numerabilità.*

6.  $(X, d)$  è totalmente limitato se e solo se il suo completamento è compatto.

*Dimostrazione.* Allora:

1. Ovvio.
2. Da un lato, sia  $A$  un insieme infinito e sia  $x_n$  una successione di elementi di  $A$ . Si osserva facilmente che l'insieme dei punti limite della successione (che in particolare sono punti di  $\omega$ -accumulazione di  $A$ ) è l'intersezione delle chiusure delle code, cioè

$$L(x_n) = \bigcap_n \overline{\{x_i\}_{i \geq n}}$$

Allora per numerabile compattezza esiste almeno un punto limite, dato che le chiusure delle code hanno la proprietà dell'intersezione finita.

Viceversa, sia data una famiglia numerabile di chiusi  $F_i$  con la proprietà dell'intersezione finita, e sia  $x_n \in \bigcap_{i=1}^n F_i$ . Se gli  $x_n$  sono finiti c'è un  $x_n$  che si ripete, cioè che frequentemente appartiene a  $F_i$ , e dunque appartiene all'intersezione degli  $F_i$ . Altrimenti, sia  $x$  un punto di  $\omega$ -accumulazione per gli  $x_n$ . Allora  $x$  è di accumulazione per ogni  $F_i$ , perché contiene infiniti elementi  $x_n$ , e quindi sta in tutti gli  $F_i$ .

3. L'implicazione è banale. Viceversa, basta osservare che, per 2), la proprietà di Bolzano-Weierstrass è equivalente alla numerabile compattezza. Ricordando che negli spazi primo-numerabile un insieme è chiuso se e solo se contiene i limiti di tutte le successioni convergenti composta da suoi punti, segue la tesi.

4. Dato che ogni aperto è unione di elementi della base, è ovvio che si riesca a estrarre un sottoricoprimento numerabile da un qualsiasi ricoprimento aperto (basta prendere un aperto che estende ogni elemento della base).

Viceversa, se  $X$  è uno spazio metrico basta mostrare che la proprietà di Lindelöf implica la separabilità, cioè l'esistenza di un sottoinsieme numerabile denso. Per ogni  $n$  consideriamo il ricoprimento di  $X$  ottenuto con le palle di raggio  $1/n$ . Per proprietà di Lindelöf ne possiamo estrarre una quantità numerabile. Allora l'insieme dei centri delle palle estratte dista al più  $1/n$  da ogni punto di  $X$ , per costruzione. Facendo l'unione su  $n$  di questi insiemi di centri otteniamo un insieme denso (giacché si accumula su ogni punto di  $X$  per costruzione) e numerabile (dato che l'unione numerabile di insiemi numerabili è numerabile).

5. Vista a Geometria 2.<sup>1</sup>

6. Basta mostrare il seguente:

**Lemma 1.4.3.**  $Y \subseteq X$  è totalmente limitato se e solo se lo è la sua chiusura.

Questo è vero perché, se  $Y$  è contenuto nell'unione di un numero finito di palle di raggio  $\varepsilon$ , la sua chiusura è contenuta nella chiusura di questa unione di palle, che è contenuta nelle palle (aperte) di raggio  $2\varepsilon$ . Il viceversa è ovvio, e in generale si può mostrare che un sottoinsieme di un totalmente limitato è anch'esso totalmente limitato. Ora,  $X$  è totalmente limitato se e solo se la sua chiusura nel completamento (cioè il completamento stesso) è totalmente limitato. Allora il completamento è compatto perché ovviamente è completo.

□

### 1.4.1 Estensione del teorema di Heine-Cantor

**Lemma 1.4.4** (Heine-Cantor V2). *Sia  $f : X \rightarrow Y$  una funzione continua tra spazi metrici e sia  $K \subseteq X$  compatto. Allora la continuità di  $f$  su  $K$  è uniforme, nel senso che*

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ t.c. } \forall y \in K, \forall x \in X, d(x, y) \leq \delta \implies d(f(x), f(y)) \leq \varepsilon$$

<sup>1</sup>Colgo l'occasione per fare pubblicità progresso consigliando [Vergamini-Marchetti-Traino, 2020].

Il lemma ci dice che  $f$  è uniformemente continua sulle coppie di punti di cui almeno uno sta nel compatto. Equivalentemente, è possibile definire una sorta di modulo di continuità  $\eta$  tale che  $\forall y \in K$  e  $\forall x \in X$   $d(f(x), f(y)) \leq \eta(d(x, y))$ .

*Dimostrazione.* Il ragionamento ricalca quello del teorema principale. Per assurdo esistano un  $\varepsilon > 0$  e una coppia di successioni  $\{x_n\}_{n \in \mathbb{N}} \subseteq X$  e  $\{y_n\}_{n \in \mathbb{N}} \subseteq K$  tali che  $d(x_n, y_n) \leq 1/n$  ma  $d(f(x_n), f(y_n)) \geq \varepsilon$ . A meno di estrarre una sottosuccessione,  $y_n \rightarrow y$ , e perciò anche  $x_n \rightarrow y$  dato che la distanza tra i termini delle successioni è infinitesima. Tuttavia per continuità dovremmo avere che  $d(f(y), f(y)) \geq \varepsilon$ , il che è ovviamente assurdo.  $\square$

### 1.4.2 Lipschitzianità su un compatto

**Teorema 1.4.5.** *Una funzione  $f : K \rightarrow \mathbb{R}$  localmente lipschitziana su uno spazio metrico compatto  $K$  è lipschitziana.*

*Dimostrazione.* Consideriamo per ogni punto di  $K$  un intorno aperto su cui la funzione è lipschitziana ed estraiamo un sottoricoprimento finito. Sia  $L$  la massima costante di Lipschitz tra quelle degli aperti estratti, e sia  $\varepsilon$  un numero di Lebesgue per il ricoprimento. Allora  $f$  è lipschitziana di costante  $\max\left\{\frac{2\|f\|_{\infty, K}}{\varepsilon}, L\right\}$ . Infatti, se due punti  $x, y \in K$  distano meno di  $\varepsilon$  devono essere contenuti in uno degli aperti del ricoprimento estratto, dove la funzione è lipschitziana di costante al più  $f$ ; se invece i due punti sono più lontani si ha che  $|f(x) - f(y)| \leq 2\|f\|_{\infty, K} \frac{d(x, y)}{\varepsilon}$ .  $\square$

## 1.5 Differenziale del determinante

**Teorema 1.5.1.**  $\det(A + H) = \det(A) + \text{tr}(A^*H) + o(H)$ , dove  $A^*$  è la matrice aggiunta classica. Se  $A$  è invertibile, la relazione diventa  $\det(A + H) = \det(A)(1 + \text{tr}(A^{-1}H)) + o(H)$ .

*Dimostrazione.* Il determinante è una funzione continua a valori reali definita sulle matrici  $M(n, \mathbb{R})$ , che possiamo identificare con  $\mathbb{R}^{n^2}$  munito del prodotto scalare standard. Quest'ultimo è la traccia del prodotto di una matrice e della trasposta dell'altra, in quanto

$$\text{tr}(A^T B) = \sum_{i,j} a_{i,j} b_{i,j}$$

Dunque la tesi è che il gradiente del determinante (rispetto a questo prodotto scalare) è la matrice aggiunta  $A_{ij}^* = (-1)^{i+j} \det C_{ji}$ , dove  $C_{ji}$  è il cofattore ottenuto togliendo la riga  $j$  e la colonna  $i$ . Ora, sviluppiamo il determinante di  $A + H$  con lo sviluppo di Leibnitz e consideriamo solo i termini che dipendono linearmente dagli elementi di  $H$ .

$$\begin{aligned} \det(A + H) &= \sum_{\sigma \in S_n} (-1)^{\text{sgn}(\sigma)} (a_{1,\sigma(1)} + h_{1,\sigma(1)}) \cdots (a_{n,\sigma(n)} + h_{n,\sigma(n)}) = \\ &= \det(A) + \sum_{i,j} h_{i,j} \sum_{\sigma \in S_n, \sigma(i)=j} \frac{(-1)^{\text{sgn}(\sigma)}}{a_{i,\sigma(i)}} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} + o(H) = \\ &= \det(A) + \sum_{i,j} h_{i,j} \sum_{\sigma \in S_{n-1}} (-1)^{\text{sgn}(\sigma)} \prod_{k \neq i} a_{k,\sigma(k)} + o(H) = \det(A) + \text{tr}(A^*H) + o(H) \end{aligned}$$

$\square$



# Capitolo 2

## Spazi vettoriali normati

### 2.1 Norma e topologia indotta

Ricordiamo che, dato uno spazio vettoriale reale  $V$ , una norma è una funzione  $\|\cdot\| : V \rightarrow \mathbb{R}$  con le seguenti proprietà:

- Non degenerazione:  $\|v\| = 0 \iff v = 0$ ;
- 1-omogeneità:  $\|\lambda v\| = |\lambda| \|v\| \forall \lambda \in \mathbb{R}$ ;
- Subadditività:  $\|v + w\| \leq \|v\| + \|w\|$

Ogni norma induce una distanza  $d(a, b) = \|a - b\|$ , la quale a sua volta induce una topologia. Mostriamo che le funzioni

1. Somma  $+$  :  $V \times V \rightarrow V$ ;
2. Prodotto per scalare  $\cdot$  :  $\mathbb{R} \times V \rightarrow V$ ;
3. Norma  $\|\cdot\|$  :  $V \rightarrow \mathbb{R}$ ;

sono continue nella topologia indotta (assumiamo che la topologia sul prodotto di due spazi sia quella indotta dalla distanza pari alla somma delle distanze tra le componenti).

*Dimostrazione.* 1. La somma è 1-Lipschitziana, in quanto

$$\|(x + y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\|$$

2. Fissato un vettore  $x_0$  e uno scalare  $\lambda_0$  abbiamo che

$$\|\lambda x - \lambda_0 x_0\| = \|\lambda(x - x_0) + x_0(\lambda - \lambda_0)\| \leq |\lambda| \|x - x_0\| + \|x_0\| |\lambda - \lambda_0|$$

L'espressione trovata è  $\mathcal{O}(1)$  per  $x \rightarrow x_0$  e  $\lambda \rightarrow \lambda_0$ .

3. Per banale applicazione della disuguaglianza triangolare

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|$$

Dunque anche la norma è 1-Lipschitziana.

□

## 2.2 Spazi normati e prodotti scalari

Sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare definito positivo. Esso induce la norma  $\|v\| = \sqrt{\langle v, v \rangle}$ . Per mostrare che la funzione definita è effettivamente una norma, e in particolare per ottenere la disuguaglianza triangolare, serve dimostrare preliminarmente il teorema seguente.

**Teorema 2.2.1** (Disuguaglianza di Cauchy-Schwarz).

$$\langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2$$

L'uguaglianza vale se e solo se i vettori sono linearmente dipendenti.

*Dimostrazione.* Consideriamo la funzione di variabile reale  $f(t) = \langle v + tw, v + tw \rangle$ . Osserviamo che tale funzione è un polinomio di grado 2 in  $t$ , in quanto

$$\langle v + tw, v + tw \rangle = t^2 \|w\|^2 + 2t \langle v, w \rangle + \|v\|^2$$

Poiché il polinomio assume sempre valori positivi o nulli, il suo discriminante è negativo o nullo. Allora

$$\langle v, w \rangle^2 - \|v\|^2 \|w\|^2 \leq 0$$

Da cui la tesi. In particolare, l'uguaglianza vale se e solo se esiste un  $t$  per cui  $f$  si annulla, ossia  $v + tw = 0$  e dunque i vettori sono linearmente dipendenti.  $\square$

Dal risultato ottenuto segue banalmente che la funzione di cui sopra è una norma. I dettagli sono lasciati al lettore.

## 2.3 Spazi $l^p$

Per fare Minkowski ci vuole  
Hölder, per fare Hölder ci vuole  
Young...

---

*Arrangiamento di una canzone popolare*

Consideriamo l'insieme  $\mathbb{R}^{\mathbb{N}}$  delle successioni a valori reali. Dato un numero reale  $p \in [1, +\infty[$ , definiamo la norma  $p$ -esima di una successione  $x_n$  come

$$\|x_n\|_p = \left( \sum_{k=0}^{+\infty} |x_k|^p \right)^{1/p}$$

Similmente, si definisce la norma "infinito" di una successione come

$$\|x_n\|_{\infty} = \sup_{k \in \mathbb{N}} |x_k|$$

L'insieme delle successioni la cui norma  $p$ -esima è limitata è uno spazio vettoriale<sup>1</sup> e che  $\|x_n\|_p$  è effettivamente una norma. Mentre le proprietà di non degenerazione e omogeneità sono banali, per giungere alla disuguaglianza triangolare serve una serie di lemmi concatenati.

---

<sup>1</sup>Per dimostrarlo, l'unica cosa non banale da fare è far vedere che la somma termine a termine di successioni la cui norma è finita ha norma finita. A tal fine si mostra prima che  $\|x_n\|_p \leq \|x_n\|_1$  per ogni  $p \in [1, +\infty[$ , mediante una generalizzazione della disuguaglianza delle medie, e poi che la somma di due serie convergenti è convergente.

### 2.3.1 Disuguaglianza di Young

**Teorema 2.3.1.** *Dati due numeri reali non negativi  $a, b$  e due numeri  $p, q \in [1, +\infty[$  tali che  $\frac{1}{p} + \frac{1}{q} = 1$ , vale che*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

*Inoltre vale l'uguaglianza se e solo se  $a^p = b^q$ .*

*Dimostrazione.* La relazione tra  $p$  e  $q$  si può riscrivere come  $\frac{1}{p-1} = q-1$  smantettando un po' con l'algebra. Ora, il lato destro della disuguaglianza può essere visto come un integrale:

$$\frac{a^p}{p} + \frac{b^q}{q} = \int_0^a x^{p-1} dx + \int_0^b x^{q-1} dx = \int_0^a x^{p-1} dx + \int_0^b x^{1/(p-1)} dx$$

Osserviamo che i due integrandi sono l'uno la funzione inversa dell'altro, perciò l'integrale di una è l'area tra il grafico dell'altra e l'asse  $y$ . La tesi segue allora osservando la Figura 2.1 e osservando che  $a^p = b^q$  si riscrive come  $b = a^{p-1}$ .  $\square$

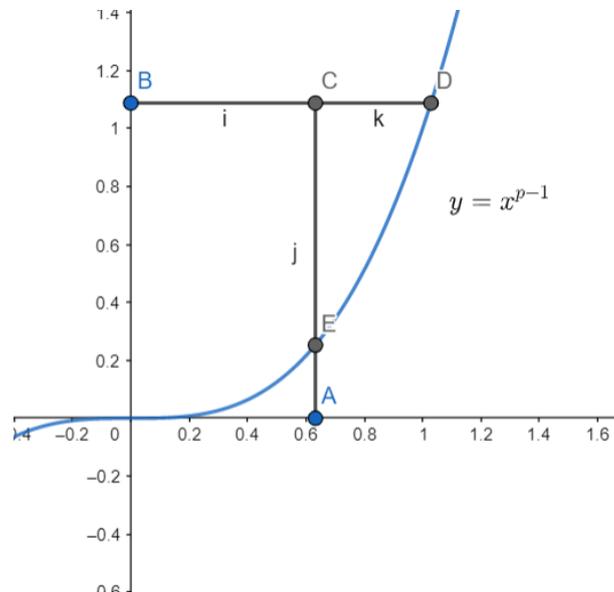


Figura 2.1: Dimostrazione grafica della disuguaglianza di Young. Il prodotto corrisponde all'area del rettangolo, minore della somma degli integrali.

### 2.3.2 Disuguaglianza di Hölder

**Teorema 2.3.2.** *Date due successioni  $x_n, y_n$  e due numeri  $p, q \in [1, +\infty]$  tali che  $\frac{1}{p} + \frac{1}{q} = 1$ ,<sup>2</sup> vale che*

$$\langle x_n, y_n \rangle \leq \|x_n\|_p \|y_n\|_q$$

*Dove il prodotto scalare considerato generalizza quello canonico in  $\mathbb{R}^n$ :*

$$\langle x_n, y_n \rangle = \sum_{k=0}^{+\infty} x_k y_k$$

*Inoltre, l'uguaglianza vale se e solo se, al variare di  $i \leq n$ , i prodotti  $x_i y_i$  sono tutti positivi e esiste una costante positiva  $\lambda$  per cui  $\lambda |x_i|^p = |y_i|^q$ .*

<sup>2</sup>Con la convenzione che  $1/\infty = 0$ .

*Dimostrazione.* Escludiamo innanzitutto il caso banale in cui  $p = +\infty$ . Negli altri casi la relazione da dimostrare è omogenea, vale a dire che è invariante per moltiplicazione per scalari. In altri termini, si può dimostrare il teorema supponendo che  $\|x_n\|_p = \|y_n\|_q = 1$  e poi generalizzare moltiplicando per uno scalare. Allora mostriamo che, restringendosi al caso detto,  $\langle x_n, y_n \rangle \leq 1$ . Applicando la disuguaglianza di Young termine a termine otteniamo

$$\langle x_n, y_n \rangle = \sum_{k=0}^{+\infty} x_k y_k \leq \sum_{k=0}^{+\infty} |x_k y_k| \leq \frac{1}{p} \sum_{k=0}^{+\infty} |x_k|^p + \frac{1}{q} \sum_{k=0}^{+\infty} |y_k|^q = \frac{1}{p} \|x_n\|_p^p + \frac{1}{q} \|y_n\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

Perché si ottenga un'uguaglianza si deve avere innanzitutto che  $x_k y_k = |x_k y_k|$ , da cui la prima condizione. Inoltre, deve valere l'uguaglianza quando si applica Young, ossia si deve avere che, per ogni  $i$ ,  $|x_i|^p = |y_i|^q$ . Tuttavia, queste non sono le vere componenti dei vettori, bensì la loro versione rinormalizzata che ci serviva per semplificare la dimostrazione. Allora, dette  $X_i, Y_i$  le componenti originali, la condizione è equivalente a

$$\frac{|X_i|^p}{\|x\|_p^p} = \frac{|Y_i|^q}{\|y\|_q^q}$$

Da cui la tesi. □

Osserviamo che, nel caso in cui  $p = q = 2$ , si ottiene la disuguaglianza di Cauchy-Schwarz.

### 2.3.3 Definizione alternativa per la norma

L'obiettivo di questa sezione, che divaga leggermente dal resto<sup>3</sup>, è ottenere una definizione diversa della norma  $p$ -esima di un vettore, nota la definizione della norma  $q$ -esima dove  $q$  è l'esponente coniugato di  $p$ . Mostriamo innanzitutto che, dato un vettore  $x \in \mathbb{R}^N$ , esiste un vettore  $y$  per cui la disuguaglianza di Hölder è un'uguaglianza, ossia

$$\langle x_n, y_n \rangle = \|x_n\|_p \|y_n\|_q$$

Infatti, imponiamo l'esistenza di un vettore  $y_n$  la cui norma  $q$ -esima è  $\|x_n\|_p^{p-1}$ .

$$\|x_n\|_p^{p-1} = \left( \sum_{k=0}^n |x_k|^p \right)^{(p-1)/p} = \left( \sum_{k=0}^n (|x_k|^{p-1})^q \right)^{1/q} = \|y_n\|_q$$

Dove la componente  $k$ -esima di  $y$  è quella tra parentesi. Ora, considerando il loro prodotto scalare, si ottiene la tesi, in quanto

$$\langle x_n, y_n \rangle = \sum_{k=0}^n |x_k|^p = \|x_n\|_p^p = \|x_n\|_p \|y_n\|_q$$

Osserviamo altresì che è possibile rinormalizzare  $y$  in modo che sia di norma 1, così che  $\langle x_n, y_n \rangle = \|x_n\|_p$ . Dalla disuguaglianza di Hölder otteniamo dunque la seguente definizione alternativa per la norma di  $x$ :

$$\|x_n\|_p = \max_{\|y_n\|_q=1} \langle x_n, y_n \rangle$$

\*

### 2.3.4 Disuguaglianza di Minkowski

**Teorema 2.3.3.** *Vale la disuguaglianza triangolare per le norme  $p$ -esime, ossia*

$$\|x_n + y_n\|_p \leq \|x_n\|_p + \|y_n\|_p$$

*Dimostrazione.* Se  $p = +\infty$  la tesi è banale. In caso contrario osserviamo che, fissato  $k \in \mathbb{N}$ , si ha

$$|x_k + y_k|^p = |x_k + y_k|^{p-1} |x_k + y_k| \leq |x_k + y_k|^{p-1} (|x_k| + |y_k|)$$

<sup>3</sup>Come peraltro ogni sezione di questo pdf.

Sommando a destra e a sinistra si ottiene

$$\|x_n + y_n\|_p^p \leq \sum_{k=0}^{+\infty} |x_k + y_k|^{p-1} |x_k| + \sum_{k=0}^{+\infty} |x_k + y_k|^{p-1} |y_k| = \langle |x_n + y_n|^{p-1}, |x_n| \rangle + \langle |x_n + y_n|^{p-1}, |y_n| \rangle$$

Applicando a entrambi i prodotti scalari la disuguaglianza di Hölder, con esponenti (nell'ordine)  $q = \frac{p}{p-1}$  e  $p$ , si ottiene

$$\begin{aligned} \|x_n + y_n\|_p^p &\leq \left\| (|x_n + y_n|)^{p-1} \right\|_{p/(p-1)} (\|x_n\|_p + \|y_n\|_p) \\ \|x_n + y_n\|_p^p &\leq \left( \sum_{k=0}^{+\infty} |x_k + y_k|^p \right)^{(p-1)/p} (\|x_n\|_p + \|y_n\|_p) = \|x_n + y_n\|_p^{p-1} (\|x_n\|_p + \|y_n\|_p) \end{aligned}$$

Da cui

$$\|x_n + y_n\|_p \leq \|x_n\|_p + \|y_n\|_p$$

□

## 2.4 Equivalenza di norme

Si dice che una norma  $\|\cdot\|_1$  sullo spazio  $V$  è più fine della norma  $\|\cdot\|_2$ , sempre su  $V$ , se la topologia indotta dalla prima contiene quella indotta dalla seconda. In altri termini, la prima topologia “distingue più cose”, in quanto ha più aperti.

Similmente, si dice che due norme  $\|\cdot\|_1, \|\cdot\|_2$  sono equivalenti (topologicamente) se le corrispondenti distanze inducono la stessa topologia, ossia se sono l'una più fine dell'altra e viceversa. Si osserva che è equivalente chiedere che le norme siano “bi-lipschitzianamente equivalenti”, ossia che

$$\exists C \in \mathbb{R}^+ \text{ t.c. } \forall x \in V \quad \frac{1}{C} \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1$$

*Dimostrazione.* Se vale che  $\|x\|_2 \leq C \|x\|_1$ , allora  $B_1(0, r) \subseteq B_2(0, Cr)$ , dove  $B_1(0, r)$  è la palla nella topologia 1 di centro l'origine e raggio  $r$ . In tal modo abbiamo ottenuto che ogni palla della topologia 2 contiene una palla della topologia 1, e ciò equivale ad affermare che la 1 è più fine.

Viceversa, supponiamo che  $B_1(0, 1) \subseteq B_2(0, C)$  per un certo  $C$ . Riscaldando, otteniamo che per ogni  $r$  vale che  $B_1(0, r) \subseteq B_2(0, Cr)$ . Allora, se per assurdo esistesse un  $x$  per cui  $\|x\|_2 > C \|x\|_1$ , supponendo che  $\|x\|_2 = C(\|x\|_1 + \delta)$  per un certo  $\delta > 0$ , avremmo che  $B_1(0, \|x\|_1 + \delta) \not\subseteq B_2(0, C(\|x\|_1 + \delta))$ , in quanto la prima contiene  $x$  ma non la seconda. □

### 2.4.1 Le norme su $\mathbb{R}^n$ sono equivalenti

**Teorema 2.4.1.** *Dato uno spazio vettoriale reale di dimensione finita, tutte le norme su di esso sono bi-Lipschitzianamente equivalenti.*

*Dimostrazione.* Sia data una norma  $\|\cdot\|$ . Mostriamo che essa è equivalente alla norma  $\|\cdot\|_1$ , cosa che ci permette di concludere per evidenti questioni di transitività. Innanzitutto, sia  $\{e_1, \dots, e_n\}$  la base canonica. Dato un qualsiasi vettore  $x = \sum_{i=1}^n a_i e_i$ , la sua norma si stima come

$$\|x\| \leq \sum_{i=1}^n |a_i| \|e_i\| \leq \max_{1 \leq i \leq n} \|e_i\| \sum_{i=1}^n |a_i| = \max_{1 \leq i \leq n} \|e_i\| \|x\|_1$$

Da cui segue che la topologia 1 è più fine della topologia della norma considerata. Dunque, poiché la norma  $\|\cdot\|$  è una funzione continua su  $V$  con la topologia che induce, lo è anche su  $V$  con la topologia indotta dalla norma 1.

Ora, consideriamo la sfera  $S_1(0, 1)$  nella norma 1. Essa è un insieme chiuso e limitato di  $\mathbb{R}^n$ , dunque è compatta<sup>4</sup>. Ciò significa che la funzione  $\|\cdot\| : S_1 \rightarrow \mathbb{R}$  ammette minimo, sia esso  $m$ . Allora  $\|x\| \geq m$  per gli  $x \in S_1$ . Rinormalizzando si ottiene che  $\forall x \in V$  vale che  $\|x\| \geq m \|x\|_1$ , da cui la tesi. □

<sup>4</sup>Si può generalizzare e mostrare che uno spazio vettoriale dotato di una distanza è di dimensione finita se e solo se la sfera unitaria è compatta. Prima o poi fornirò una dimostrazione.

Il teorema può ulteriormente essere perfezionato. Si può mostrare che esiste una sola topologia su  $\mathbb{R}^n$  che rende continue le funzioni somma e prodotto per scalare.

## 2.5 Spazi vettoriali complessi

### 2.5.1 Relazione con gli $\mathbb{R}$ -spazi vettoriali

**Teorema 2.5.1** (Caratterizzazione dei  $\mathbb{C}$ -spazi vettoriali). *Un insieme  $X$  è uno spazio vettoriale su  $\mathbb{C}$  se e solo se è uno spazio vettoriale reale con un endomorfismo  $\mathbb{R}$  lineare  $J$  tale che  $J^2 = -I$ .*

*Dimostrazione.* Innanzitutto, uno spazio vettoriale complesso è uno spazio vettoriale reale per restrizione del prodotto per scalari a un campo più piccolo. Inoltre, sia  $J$  il prodotto per lo scalare  $i$ . Tale applicazione è lineare su  $\mathbb{R}$  e ovviamente soddisfa la proprietà richiesta.

Viceversa, dato uno spazio vettoriale reale munito di una  $J$  come sopra, estendiamo a uno spazio vettoriale complesso allargando il prodotto per scalari a  $\mathbb{C}$  ponendo, per ogni  $a + ib \in \mathbb{C}$  e per ogni  $x \in V$

$$(a + ib)x = ax + bJx$$

Si verifica banalmente che tale operazione è effettivamente un prodotto per scalare.<sup>5</sup> □

In questa ottica, dato uno spazio vettoriale reale  $V$  è possibile considerarne il complessificato,<sup>6</sup> ossia lo spazio vettoriale su  $V \times V$  in cui l'applicazione  $J$  è

$$J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$$

Ora, un'applicazione lineare  $L$  su uno spazio vettoriale complesso è  $\mathbb{C}$ -lineare se e solo se è  $\mathbb{R}$ -lineare ed è tale che  $Lix = iLx$ , ossia se commuta con la  $J$  di sopra. Coi conti espliciti otteniamo che  $L$  deve essere della forma

$$L = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

### 2.5.2 Norma complessa

**Definizione 10.** *Una norma  $\|\cdot\|$  su uno spazio vettoriale complesso  $V$  è una norma complessa se  $\forall \lambda \in \mathbb{C}$  e  $\forall x \in V$  vale  $\|\lambda x\| = |\lambda| \|x\|$ .*

Il seguente lemma mostra che, sotto certe condizioni, è possibile estendere una norma reale a una norma complessa che induce la stessa topologia. Peraltro, ciò significa anche che la topologia di uno spazio vettoriale complesso, sotto certe condizioni, dipende solo dalla sua struttura reale (e in particolare che  $\mathbb{C}^n$  è omeomorfo a  $\mathbb{R}^{2n}$ ).

**Lemma 2.5.2.** *Se  $V$  è uno spazio vettoriale complesso munito di una norma reale  $\|\cdot\|$  e se l'applicazione  $J$  è continua, allora esiste una norma complessa topologicamente equivalente alla norma reale.*

*Dimostrazione.* Consideriamo la norma complessa

$$\|x\|_{\mathbb{C}} = \max_{|\lambda|=1} \|\lambda x\|$$

Il massimo esiste perché, fissato  $x \in V$ , stiamo facendo il sup sul compatto  $\mathbb{S}^1$ . Verifichiamo che questa è effettivamente una norma.

- La non degenerazione è ovvia.

<sup>5</sup>La costruzione fatta con  $J$  si può generalizzare a una generica applicazione lineare  $A$ . In tal senso si ottiene che uno spazio vettoriale reale  $X$  è un  $K[A]$ -modulo, cioè uno spazio vettoriale sull'anello dei polinomi in  $A$ . Nel nostro caso,  $\mathbb{C}$  è lo spazio vettoriale sull'anello dei polinomi in  $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ .

<sup>6</sup>Per i più algebristi e i meno analisti, il complessificato è il prodotto tensore tra  $\mathbb{C}$  e  $V$ .

- Moltiplicando per uno scalare complesso  $\nu$  otteniamo

$$\|\nu x\|_{\mathbb{C}} = \max_{|\lambda|=1} \|\lambda \nu x\| = \max_{|\lambda|=1} \|\lambda|\nu|x\| = |\nu| \|x\|_{\mathbb{C}}$$

in quanto, dati due numeri complessi con lo stesso modulo, il loro rapporto appartiene a  $\mathbb{S}^1$ .

- Per quanto riguarda la triangolare, per le proprietà del max abbiamo

$$\|x + y\|_{\mathbb{C}} = \max_{|\lambda|=1} \|\lambda(x + y)\| \leq \max_{|\lambda|=1} \|\lambda x\| + \|\lambda y\| \leq \max_{|\lambda|=1} \|\lambda x\| + \max_{|\lambda|=1} \|\lambda y\| = \|x\|_{\mathbb{C}} + \|y\|_{\mathbb{C}}$$

Per quanto riguarda l'equivalenza, osserviamo che sicuramente  $\|x\| \leq \|x\|_{\mathbb{C}}$  per come è definita. Per l'altra disuguaglianza, scriviamo  $\lambda = a + ib$  e interpretiamo la moltiplicazione per  $i$  come l'applicazione  $J$ .

$$\|x\|_{\mathbb{C}} = \max_{a^2+b^2=1} \|ax + bJx\| \leq \max_{\theta \in [-1,1]} |\cos \theta| \|x\| + |\sin \theta| \|Jx\|$$

Poiché  $J$  è continua per ipotesi, e dunque è lipschitziana, esiste un  $M > 0$  tale che

$$\|x\|_{\mathbb{C}} \leq \max_{\theta \in [-1,1]} |\cos \theta| \|x\| + |\sin \theta| M \|x\| = C \|x\|$$

per una certa costante  $C$ . Allora  $\|x\| \leq \|x\|_{\mathbb{C}} \leq \|C\|x\|$ , perciò le norme sono topologicamente equivalenti.  $\square$

### 2.5.3 Prodotto hermitiano

**Definizione 11.** Un'applicazione  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  su uno spazio vettoriale complesso si dice prodotto hermitiano (definito positivo) se

- è lineare nella prima componente e antilineare nella seconda (si dice in tal caso che è sesquilineare):

$$\langle ax + y, cw + z \rangle = a \langle x, cw + z \rangle + \langle y, cw + z \rangle = a \bar{c} \langle x, w \rangle + \bar{c} \langle y, w \rangle + a \langle x, z \rangle + \langle y, z \rangle$$

- $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- $\langle x, x \rangle \geq 0$  e l'uguaglianza vale se e solo se  $x = 0$ .

Si osserva facilmente che un prodotto hermitiano definito positivo induce la norma complessa

$$\|x\| = \sqrt{\langle x, x \rangle}$$

**Teorema 2.5.3** (Disuguaglianza di Cauchy-Schwarz). Per ogni  $x, y \in V$  vale che

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

*Dimostrazione.* Osserviamo innanzitutto che, indicando con  $\Re$  la parte reale,  $\Re(\langle x, y \rangle)$  è un prodotto scalare reale sulla soggiacente struttura di  $\mathbb{R}$ -spazio vettoriale. Infatti è simmetrico, in quanto sappiamo che  $\Re(\langle x, y \rangle) = \Re(\langle y, x \rangle)$ , ed è bilineare, dato che se  $a, c$  sono reali vale che

$$\Re(\langle ax + y, cw + z \rangle) = \Re(ac \langle x, w \rangle + c \langle y, w \rangle + a \langle x, z \rangle + \langle y, z \rangle) = ac \Re(\langle x, w \rangle) + c \Re(\langle y, w \rangle) + a \Re(\langle x, z \rangle) + \Re(\langle y, z \rangle)$$

Essendo inoltre definito positivo vale la disuguaglianza di Cauchy-Schwarz reale:

$$\Re(\langle x, y \rangle) \leq \sqrt{\Re(\langle x, x \rangle)} \sqrt{\Re(\langle y, y \rangle)}$$

Ora, poiché il prodotto hermitiano era definito positivo (e quindi in particolare la norma di un vettore è reale), la parte reale alla RHS è inutile, perciò otteniamo

$$\Re(\langle x, y \rangle) \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} = \|x\| \|y\|$$

Per quanto riguarda la LHS, osserviamo che esiste un numero complesso  $\lambda$  di modulo 1 tale che  $\lambda = \frac{\langle x, y \rangle}{|\langle x, y \rangle|}$ . Tale numero si indica con  $\text{sgn}(\langle x, y \rangle)$ . Invertendo la relazione possiamo scrivere

$$|\langle x, y \rangle| = \bar{\lambda} \langle x, y \rangle = \langle x, \lambda y \rangle$$

Poiché la quantità alla LHS è reale lo è anche quella alla RHS. Dunque

$$|\langle x, y \rangle| = \Re(\langle x, \lambda y \rangle) \leq \sqrt{\Re(\langle x, x \rangle)} \sqrt{\lambda \bar{\lambda} \Re(\langle y, y \rangle)} = \|x\| \|y\|$$

Da cui la tesi. □

# Capitolo 3

## Completezza

### 3.1 Spazi completi

**Definizione 12.** Una successione  $x_n \in (X, d)$  si dice di Cauchy se

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \text{ t.c. } \forall n, m \geq n_0 d(x_n, x_m) < \varepsilon$$

In altri termini, una successione di Cauchy si addensa progressivamente. Osserviamo che una successione convergente è di Cauchy, ma non sempre vale il viceversa. Ad esempio, in  $\mathbb{Q}$  una successione che tende a  $\sqrt{2}$  non ha limite nell'insieme considerato. Uno spazio metrico in cui ogni successione di Cauchy converge si dice completo.

#### 3.1.1 Completezza e chiusura

**Lemma 3.1.1.** Un sottoinsieme  $Y$  chiuso di uno spazio metrico  $X$  completo è a sua volta completo.

*Dimostrazione.* Innanzitutto, osserviamo che in uno spazio metrico<sup>1</sup> la chiusura è equivalente alla chiusura per successioni, ossia se una successione  $y_n$  a valori in  $Y$  converge a un certo  $x \in X$ , allora  $x \in Y$ . Ora, consideriamo una successione di Cauchy in  $Y$ . Essa converge in  $X$ , e per chiusura sequenziale il limite appartiene a  $Y$ . In altri termini, la successione di Cauchy converge in  $Y$ .  $\square$

Una sorta di inverso del lemma appena dimostrato è il seguente.

**Lemma 3.1.2.** Un sottoinsieme completo  $Y \subseteq X$  è chiuso.

*Dimostrazione.* Supponiamo che una successione  $y_n$  converga a un certo  $x$ . Allora la successione è di Cauchy, e perciò converge all'interno dell'insieme  $Y$ . In altri termini  $Y$  contiene tutti i suoi punti di accumulazione, ed è dunque chiuso.  $\square$

### 3.2 Estensione di mappe UC a valori in spazi completi

**Teorema 3.2.1.** Sia  $E \subseteq (X, d)$ , sia  $(Y, D)$  completo e sia  $f : (E, d) \rightarrow (Y, D)$  una mappa uniformemente continua, di modulo di continuità  $\omega(t)$ . Allora  $f$  si estende in modo unico a una funzione  $F : (\bar{E}, d) \rightarrow (Y, D)$  definita sulla chiusura del dominio e con lo stesso modulo di continuità.

*Dimostrazione.* Sia  $x \in \bar{E}$  e sia  $x_n$  una successione che vi tende. In particolare, tale successione è di Cauchy. Allora anche la successione delle immagini lo è, dato che per uniforme continuità

$$D(f(x_n), f(x_m)) \leq \omega(d(x_n, x_m))$$

e il modulo di continuità è infinitesimo in zero. Poiché lo spazio di arrivo  $Y$  è completo, la successione delle immagini converge a un certo  $y$ , che poniamo uguale a  $F(x)$ .

<sup>1</sup>In realtà, basta che lo spazio sia primo-numerabile. Si osservi che gli spazi metrici sono primo-numerabili perché, dato un punto  $x$ , le palle di centro  $x$  e raggio razionale descrivono un sistema fondamentale di intorni numerabile.

Osserviamo che la definizione è ben posta perché è indipendente dalla scelta della successione. Se, infatti, prendiamo una diversa successione  $x'_n \rightarrow x$ , si ha che  $d(x_n, x'_n) \rightarrow 0$ . Per uniforme continuità anche le immagini hanno distanza infinitesima e devono perciò convergere allo stesso elemento  $y$ .

Abbiamo dunque definito una funzione  $F$  su  $\bar{E}$ . Mostriamo che essa ha lo stesso modulo di continuità di  $f$ . Siano allora  $x, y \in \bar{E}$  a distanza  $d(x, y) = \varepsilon - \delta < \varepsilon$  e  $x_n \rightarrow x, y_n \rightarrow y$  due successioni che vi tendono rispettivamente. Sia  $N$  tale che, per ogni  $n > N$ , le due successioni stanno entrambe a meno di  $\delta/2$  dai rispettivi limiti. Allora la distanza tra i termini generali è stimata da

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n) < \varepsilon$$

Per uniforme continuità di  $f$ ,  $D(f(x_n), f(y_n)) \leq \omega(\varepsilon)$ , e tale relazione passa al limite per  $n \rightarrow +\infty$ . Abbiamo allora che, se  $d(x, y) < \varepsilon$ ,  $D(F(x), F(y)) \leq \omega(\varepsilon)$ .  $\square$

### 3.3 Spazi di Banach

Uno spazio vettoriale normato, completo rispetto alla distanza indotta dalla norma, si dice di Banach. Come vedremo, più o meno tutta l'analisi si può ripetere per filo e per segno considerando un qualsiasi spazio di Banach al posto dei numeri reali, a costo di sacrificare la nozione, peraltro superflua, di ordinamento totale.

Il seguente teorema fornisce un metodo operativo per determinare la completezza di uno spazio normato.

**Teorema 3.3.1** (Criterio di convergenza totale). *Uno spazio vettoriale è di Banach se e solo se, comunque si scelga una successione  $x_n$  la cui serie delle norme converge, essa converge come successione delle somme parziali. In simboli,  $(V, \|\cdot\|)$  è di Banach sse*

$$\sum_{n \in \mathbb{N}} \|x_n\| < +\infty \implies \sum_{n=0}^{\infty} x_n \in V$$

*Dimostrazione.* Supponiamo che lo spazio sia di Banach e che la serie delle norme di  $x_n$  converga (si dice in tal caso che la serie converge totalmente). Allora la successione delle somme parziali  $s_N = \sum_{n=0}^N x_n$  è di Cauchy, in quanto la distanza tra due somme parziali si stima con la coda della serie delle norme, infinitesima perché la serie converge. Essendo lo spazio completo, la serie converge.

Viceversa, supponiamo che valga l'implicazione della tesi e mostriamo che ogni successione di Cauchy  $x_n$  converge. Per la proprietà di Cauchy si può estrarre una sottosuccessione  $x_{n_k}$  tale che  $\|x_{n_k} - x_{n_{k-1}}\| < 2^{-k}$ . Allora definiamo  $v_0 = x_{n_0}$  e  $v_k = x_{n_k} - x_{n_{k-1}}$ . Tale successione converge totalmente, perché la serie delle sue norme è maggiorata dalla serie geometrica. Allora essa converge come successione delle somme parziali. Osserviamo infine che la somma parziale di tale serie è  $\sum_{k=0}^n v_k = x_{n_0} + \sum_{k=1}^n x_{n_k} - x_{n_{k-1}} = x_{n_n}$ . Abbiamo allora trovato una sottosuccessione convergente di una successione di Cauchy, e dunque tutta la successione converge. Il lettore è invitato a verificare la validità dei due lemmi che ho usato ma non ho avuto voglia di dimostrare.  $\square$

#### 3.3.1 Connessione degli aperti

**Teorema 3.3.2.** *Un aperto  $U$  di uno spazio di Banach è connesso se e soltanto se è connesso per archi.*

*Dimostrazione.* Sia  $x \in U$  e sia  $A$  l'insieme dei punti di  $U$  connessi per archi ad  $x$ .  $A$  è aperto, perché per ogni  $y \in A$  esiste una palla convessa  $B(y, r) \subseteq U$  che è connessa per archi lineari a  $y$ , e dunque a  $x$ . D'altro canto, anche il complementare di  $A$  è aperto, perché se  $z$  non è connettibile ad  $x$  deve esistere un suo intorno di punti non connettibili a  $x$ . Dunque, dato che  $U$  è aperto, chiuso e non vuoto deve coincidere con  $U$  per connessione.  $\square$

**Teorema 3.3.3.** *Sia  $U$  un aperto di uno spazio di Banach e sia  $D$  un insieme denso in  $U$ . Se due punti  $x, y \in U$  sono connessi da un arco  $\gamma \subseteq U$  esiste una spezzata  $\delta \subseteq U$  che connette  $x$  a  $y$  i cui vertici intermedi sono contenuti in  $D$ .*

*Dimostrazione.* Innanzitutto, consideriamo un intorno della curva della forma

$$V = \bigcup_{t \in [0,1]} B(\gamma(t), \varepsilon) \subseteq U$$

Se per assurdo un tale intorno non esistesse, potremmo trovare due successioni  $x_n = \gamma(t_n)$  e  $y_n \notin U$  tali che  $\|x_n, y_n\| \leq \frac{1}{n}$  per ogni  $n$ . Per compattezza potremmo estrarre una successione  $x_{n(k)} \rightarrow x \in \gamma([0, 1])$ . Ma allora  $x \notin U$ , perché sarebbe anche il limite della successione  $y_{n(k)}$  a valori nel chiuso  $U^C$ .

Ora, sia  $\omega$  il modulo di continuità di  $\gamma$  e sia  $n \in \mathbb{N}$  per cui  $\omega\left(\frac{1}{n}\right) < \frac{\varepsilon}{10}$ . Suddividiamo  $[0, 1]$  in intervalli della forma  $\left[\frac{k}{n}, \frac{k+1}{n}\right]$  e costruiamo induttivamente la spezzata  $\delta$  ponendo

$$\begin{cases} \delta(0) = \gamma(0) \\ \delta(1) = \gamma(1) \\ \delta\left(\frac{k}{n}\right) = z_k \text{ t.c. } z_k \in D, \|z_k - \gamma\left(\frac{k}{n}\right)\| \leq \frac{\varepsilon}{10} \\ \delta\left((1-t)\frac{k}{n} + t\frac{k+1}{n}\right) = (1-t)z_k + tz_{k+1} \end{cases}$$

Osserviamo che

$$\|z_{n+1} - z_n\| \leq \frac{\varepsilon}{5} + \omega\left(\frac{1}{n}\right) \leq \frac{3\varepsilon}{10}$$

Se dimostriamo che  $\delta([0, 1]) \subseteq V$ , ossia che  $\|\gamma - \delta\|_\infty < \varepsilon$ , abbiamo concluso. Dobbiamo vedere cosa succede nei punti diversi dai punti di giunzione. Posto allora  $h < \frac{1}{n}$  abbiamo

$$\begin{aligned} \left\| \delta\left(\frac{k}{n} + h\right) - \gamma\left(\frac{k}{n} + h\right) \right\| &\leq \left\| \delta\left(\frac{k}{n} + h\right) - \delta\left(\frac{k}{n}\right) \right\| + \left\| \delta\left(\frac{k}{n}\right) - \gamma\left(\frac{k}{n}\right) \right\| + \left\| \gamma\left(\frac{k}{n}\right) - \gamma\left(\frac{k}{n} + h\right) \right\| \\ &\leq \|z_{n+1} - z_n\| + \frac{\varepsilon}{10} + \omega\left(\frac{1}{n}\right) \end{aligned}$$

Da cui  $\left\| \delta\left(\frac{k}{n} + h\right) - \gamma\left(\frac{k}{n} + h\right) \right\| \leq \frac{\varepsilon}{2}$ . □

Percorrendo le spezzate con velocità tale da annullare tutte le derivate nei punti di giunzione (cosa che si può fare con abili magheggi utilizzando la funzione  $\exp(-1/x^2)$ , che ha tutte le derivate nulle nell'origine) otteniamo il seguente risultato:

**Corollario 3.3.3.1.** *Un aperto di uno spazio di Banach è connesso se e solo se è connesso per archi  $C^\infty$*

Infine, possiamo dire che

**Corollario 3.3.3.2.** *Le curve  $C^\infty$  sono uniformemente dense nelle curve continue a valori in uno spazio di Banach.*

Basta infatti osservare che, se  $\gamma$  è una curva, scegliendo come aperto un intorno “tubolare” di spessore  $\varepsilon$  attorno all'immagine di  $\gamma$  e applicando i due teoremi precedenti si ottiene una spezzata  $C^\infty$  a distanza uniforme minore di  $\varepsilon$  da  $\gamma$ .

### 3.3.2 Quozienti di spazi di Banach

Sia  $E$  uno spazio vettoriale e sia  $\|\cdot\|$  una seminorma su  $E$ , cioè una norma degenera per la quale esistono vettori non nulli  $x$  di norma nulla. Sia  $F$  un sottospazio di  $E$  e consideriamo l'insieme quoziente  $E/F = \{x + F, x \in E\}$  di tutti i traslati di  $F$ . Muniamo tale spazio della seguente norma

$$\|x + F\| = \inf_{\pi(y)=x+F} \|y\| = d(x, F)$$

dove  $\pi$  è la proiezione canonica l'ultimo termine è la distanza tra il sottospazio  $F$  e il punto  $y$ . Si noti che se  $\pi(y) = x + F$  allora  $x - y \in F$ , cioè

$$d(x, F) = \inf_{z \in F} \|x - z\| = \inf_{z \in F} \|y - (y - x + z)\| = \inf_{z' \in F} \|y - z'\| = d(y, F)$$

La funzione  $\|\cdot\|$  indotta è banalmente subadditiva e 1-omogenea, e dunque è una seminorma. Inoltre, i traslati la cui norma è nulla sono quelli per cui  $d(y, F) = 0$  per ogni  $y \in x + F$ . Ciò significa che  $y \in \bar{F}$ , cioè il traslato sta nella chiusura di  $F$ . Allora se  $F$  è chiuso la seminorma quoziente è una norma, perché l'unico traslato a distanza nulla è  $F$  stesso.

Mettiamoci ora nel caso in cui  $E$  è di Banach. Vale il seguente teorema che relaziona la completezza di  $E$  alla completezza dei suoi sottospazi e dei quozienti.

**Teorema 3.3.4.** *Sia  $E$  uno spazio normato e  $F$  un suo sottospazio chiuso.  $E$  è di Banach se e solo se lo sono  $F$  e  $E/F$ , quest'ultimo con la norma indotta dal quoziente.*

*Dimostrazione.* Supponiamo che  $E$  sia completo. Allora  $F$  è completo perché è chiuso in un completo. Per la completezza di  $E/F$  usiamo il criterio di convergenza totale. Sia allora  $\sum_{k=0}^{\infty} \pi(V_n)$  una sommatoria normalmente convergente in  $E/F$ , cioè tale che  $\sum_{k=0}^{\infty} \|\pi(V_n)\| = \sum_{k=0}^{\infty} \inf_{\pi(x_n)=V_n} \|x_n\|$ . Se scegliamo una successione  $x_n \in X$  tale che  $\pi(x_n) = V_n$  che sia normalmente convergente in  $X$  abbiamo finito, perché allora la sommatoria delle proiezioni è la proiezione della sommatoria per continuità di  $\pi$ . Ma ciò è banale, perché per definizione di inf esiste un vettore  $x_n$  la cui norma è minore del doppio della norma di  $V_n$ , e dunque gli  $x_n$  convergono totalmente.

Supponiamo ora che  $F$  e  $E/F$  siano di Banach e mostriamo che lo è  $E$ . Consideriamo allora una successione  $x_n \in E$  di Cauchy. Per definizione della norma quoziente, anche  $\pi(x_n) \in E/F$  lo è, e dunque per completezza converge a un certo  $\pi(x)$ . Ciò significa in particolare che la seguente quantità è infinitesima per  $n \rightarrow \infty$ :

$$\|\pi(x_n) - \pi(x)\| = \|\pi(x_n - x)\| = \inf_{y \in F} \|x_n - x - y\|$$

Il secondo passaggio segue dalla linearità della proiezione. Ora, sia  $y_n \in F$  che si avvicina all'inf a meno di  $1/n$ , ossia tale che

$$\|x_n - x - y_n\| \leq \|\pi(x_n - x)\| + 1/n$$

Mostriamo che anche la successione  $y_n$  è di Cauchy. Infatti

$$\|y_n - y_m\| \leq \|x_n - x - y_n\| + \|x_m - x - y_m\| \leq \|\pi(x_n - x)\| + 1/n + \|\pi(x_m - x)\| + 1/m$$

Che è infinitesimo. Allora per completezza  $y_n \rightarrow y$ . Mostriamo che  $x_n \rightarrow x + y$ , cosa che conclude la dimostrazione. Infatti

$$\|x_n - x - y\| \leq \|x_n - x - y_n\| + \|y_n - y\|$$

Ed entrambe le quantità sono infinitesime. □

### 3.3.3 Spazi di funzioni completi

Consideriamo l'insieme delle funzioni  $f : S \rightarrow (E, \|\cdot\|)$ , dove  $S$  è un generico insieme e  $E$  è uno spazio di Banach. Tale insieme è uno spazio di Banach a sua volta, con la norma del sup:

$$\|f\|_{\infty} = \sup_{x \in S} \|f(x)\|$$

Si osserva che la norma del sup metrizza la convergenza uniforme di funzioni.

Consideriamo ora il sottospazio vettoriale  $B(S, E)$  delle funzioni limitate, ossia la cui norma è finita.

**Teorema 3.3.5.**  *$B(S, E)$  è completo.*

*Dimostrazione.* Sia  $f_n$  una successione di funzioni che sia di Cauchy nella norma del sup, ossia tale che, definitivamente,  $\|f_n - f_m\| < \varepsilon$ . Per come è definita la norma, la relazione vale anche puntualmente, ossia fissato  $x$  la successione  $f_n(x)$  è di Cauchy in  $E$ . Dunque, poiché  $E$  è completo, la successione converge a un certo  $f(x)$ . In tal modo abbiamo identificato naturalmente una funzione che è limite puntuale della successione di funzioni considerata.

Ora, mostriamo che la convergenza è uniforme. Per farlo, fissiamo un punto e osserviamo che definitivamente  $\|f_n(x) - f_m(x)\| < \varepsilon$  in modo indipendente da  $x$ . Allora la tesi segue mandando  $m \rightarrow +\infty$  e ottenendo che  $\forall x, \|f_n(x) - f(x)\| < \varepsilon$ , da cui  $\|f_n - f\|_{\infty} < \varepsilon$  definitivamente. □

La dimostrazione precedente mostra, in particolare, che l'insieme  $l^{\infty}$  definito in 2.3 è completo, in quanto è l'insieme delle funzioni da  $\mathbb{N}$  a  $\mathbb{R}$  che assumono valori equilimitati.

Supponiamo ora che  $S$  sia uno spazio topologico e consideriamo il sottospazio  $BC(S, E)$  delle funzioni continue e limitate.

**Teorema 3.3.6.**  *$BC(S, E)$  è completo.*

*Dimostrazione.* Basta mostrare che  $BC(S, E)$  è chiuso in  $B(S, E)$  e concludere per quanto detto in 3.1.1. Ciò è vero perché, come già visto nel corso di Analisi 1, il limite uniforme di una successione di funzioni continue è una funzione continua.  $\square$

Riallacciandoci all'insieme  $l^p$ , l'equivalente delle funzioni continue e limitate è l'insieme delle successioni infinitesime  $c_0$ . Infatti, muniamo  $\mathbb{N} \cup \{+\infty\}$  della topologia generata da  $\mathcal{P}(\mathbb{N})$  e dalle "semirette", ossia dagli insiemi della forma  $[n, +\infty]$ . Allora  $c_0$  è l'insieme delle funzioni nulle all'infinito e continue in tale punto, dato che le semirette sono gli intorni di  $+\infty$ . Inoltre, poiché sui naturali veri e propri la topologia considerata è quella discreta, ogni successione è continua in  $n$  per ogni  $n$  naturale.

### 3.3.4 Ogni spazio metrico si immerge nelle funzioni continue

**Teorema 3.3.7** (Fréchet-Kuratowski). *Dato uno spazio metrico  $(X, d)$ , esso si immerge isometricamente in uno spazio di Banach, e più precisamente in  $(BC(X, \mathbb{R}), \|\cdot\|_\infty)$ .*

*Dimostrazione.* Escludiamo ovviamente il caso banale in cui  $X$  è vuoto. Fissiamo un punto  $x_0 \in X$  e definiamo l'operatore che associa a ogni  $x \in X$  la funzione  $\phi_x$  così definita

$$\phi_x(y) = d(x, y) - d(x_0, y)$$

Tale funzione è ovviamente continua. Inoltre è limitata, in quanto

$$\|\phi\|_\infty = \sup_{y \in X} |d(x, y) - d(x_0, y)| \leq d(x, x_0)$$

Allora  $\phi_x \in BC(X, \mathbb{R})$ . Infine, la funzione  $x \rightarrow \phi_x$  è un'isometria, dato che

$$\|\phi_x - \phi_y\|_\infty = \sup_{z \in X} |d(x, z) - d(y, z)| = d(x, y)$$

$\square$

## 3.4 Completamento di uno spazio metrico

**Definizione 13.** *Sia  $(X, d)$  uno spazio metrico. Un suo completamento è uno spazio metrico completo  $(X', d')$  in cui  $X$  si immerge isometricamente, e tale che l'immagine dell'immersione è densa in  $X'$ . Più precisamente,  $(X', d')$  è un completamento di  $(X, d)$  se esiste un'isometria  $i : (X, d) \rightarrow (X', d')$  tale che  $i(\overline{X}) = X'$ .*

Di fatto, ciò ci permette di identificare  $X$  con un sottoinsieme di  $X'$ .

**Teorema 3.4.1.** *Ogni spazio metrico ammette un completamento, che è unico a meno di isometrie.*

*Dimostrazione.* Per l'esistenza ci sono due vie. Una, quella canonica, prevede di immergere  $X$  nello spazio delle successioni di Cauchy a valori in  $X$ , ma è una noia mortale e soprattutto fa fare un sacco di confusione perché ci sono troppi indici che ballano. Preferiamo allora osservare che il Teorema 3.3.7 garantisce un'immersione isometrica  $i : X \rightarrow BC(X, \mathbb{R})$ . Allora, poiché lo spazio di arrivo è completo, la chiusura dell'immagine  $\overline{i(X)}$  è un insieme completo (e anzi, è uno spazio di Banach).

Per quanto riguarda l'unicità, consideriamo due completamenti  $Y$  e  $Z$  e siano  $i : X \rightarrow Y$  e  $j : X \rightarrow Z$  le immersioni isometriche. Per teorema di estensione di funzioni uniformemente continue<sup>2</sup> alla chiusura di  $X$ , che si può identificare isometricamente sia con  $Y$  sia con  $Z$ , le isometrie si estendono in modo unico a due isometrie  $\bar{i} : Z \rightarrow Y$  e  $\bar{j} : Y \rightarrow Z$  che coincidono con  $i$  e  $j$  su  $X$ . Mostriamo allora che le estensioni sono l'una l'inversa dell'altra. Infatti,  $\bar{i}\bar{j}$  è un'isometria da  $Z$  in  $Z$  che estende l'identità su  $j(X)$ , e sempre per il teorema di estensione deve essere unica. Allora  $\bar{i}\bar{j} = id_Z$ . Per simmetria segue la tesi.  $\square$

<sup>2</sup>Vedasi 3.2.

### 3.5 Teorema delle contrazioni

**Teorema 3.5.1** (Banach-Cacciopoli). *Sia  $(X, d)$  completo non vuoto e sia  $T : X \rightarrow X$  una contrazione di costante  $k$ . Allora  $T$  ammette un unico punto fisso  $\bar{x}$ . Inoltre, dato un punto  $x$ , la successione  $T^n(x)$  converge a  $\bar{x}$ , e vale la seguente stima per la velocità di convergenza:*

$$d(\bar{x}, T^n(x)) \leq \frac{k^n}{1-k} d(x, T(x))$$

*Dimostrazione.* Preso un punto  $x$ , la successione delle iterate di  $T$  applicate a  $x$  è di Cauchy, in quanto, scelti comunque  $m > n$  naturali,

$$d(T^n(x), T^m(x)) \leq k^n d(x, T^{m-n}(x)) \leq k^n (d(x, T(x)) + \dots + d(T^{m-n-1}(x), T^{m-n}(x)))$$

$$d(T^n(x), T^m(x)) \leq k^n d(x, T(x)) \sum_{j=0}^{m-n-1} k^j \leq \frac{k^n}{1-k} d(x, T(x))$$

Che è infinitesima per  $n \rightarrow +\infty$ . Allora per completezza di  $X$  la successione ammette limite  $\bar{x}$ , che è un punto fisso perché, per continuità di  $T$ ,

$$T(\bar{x}) = T\left(\lim_{n \rightarrow +\infty} T^n(x)\right) = \lim_{n \rightarrow +\infty} T^{n+1}(x) = \bar{x}$$

Inoltre, mandando al limite in  $m$  la relazione della seconda riga si ottiene la stima cercata.

Per quanto riguarda l'unicità, se per assurdo ci fossero due punti fissi distinti  $x \neq y \in X$ , si avrebbe che  $d(x, y) = d(T(x), T(y)) = kd(x, y)$ , il che è impossibile essendo  $k < 1$ .  $\square$

Osserviamo che è possibile indebolire ulteriormente le ipotesi, come mostrato nel seguente corollario.

**Corollario 3.5.1.1.** *Sia  $(X, d)$  uno spazio metrico completo non vuoto e sia  $T : X \rightarrow X$  una funzione tale che esiste un  $N \in \mathbb{N}$  per cui  $T^N$  è una contrazione di costante  $k$ . Allora  $T$  ammette un unico punto fisso  $\bar{x}$ . Inoltre, dato un punto  $x$ , la successione  $T^n(x)$  converge a  $\bar{x}$*

*Dimostrazione.* Per teorema delle contrazioni,  $T^N$  ammette un unico punto fisso  $\bar{x}$ , il quale è dunque un punto periodico di periodo  $N$  per  $T$ . Ma allora anche  $T(\bar{x})$  è un punto fisso per  $T^N$ , e dunque deve coincidere con  $\bar{x}$  per unicità. Allora  $\bar{x}$  è un punto fisso per  $T$ .

Sia ora  $x \in X$  e consideriamo la successione  $T^n(x)$ . Essa converge a  $\bar{x}$  perché è l'unione di  $N$  sottosuccessioni che vi convergono, in quanto per le iterate di  $T^N$  vale il teorema delle contrazioni.  $\square$

#### 3.5.1 Variante per spazi compatti

**Teorema 3.5.2.** *Sia  $(X, d)$  uno spazio metrico compatto. Allora una funzione  $T : X \rightarrow X$  che riduce le distanze, cioè tale che  $\forall x, y \in X$   $d(T(x), T(y)) < d(x, y)$ , ammette un unico punto fisso.*

*Dimostrazione.* Come visto nel teorema delle contrazioni, se il punto fisso esiste è unico. Per quanto riguarda l'esistenza, consideriamo la funzione che associa a  $x \in X$  la sua distanza da  $T(x)$ . Questa funzione è continua e su un compatto ammette minimo. Se tale minimo non è zero, sia  $\bar{x}$  che realizza il minimo. Allora  $d(\bar{x}, T(\bar{x})) > d(T(\bar{x}), T(T(\bar{x})))$ , e ciò contraddice l'ipotesi di minimalità.  $\square$

### 3.6 Teorema di Baire

Diciamo che:

- Un sottoinsieme è di prima categoria se si può ricoprire con un'unione numerabile di chiusi di parte interna vuota.
- Un sottoinsieme non di prima categoria si dice di seconda categoria.
- Un'intersezione numerabile di aperti si dice  $G_\delta$ .

- Un sottoinsieme il cui complementare è di prima categoria si dice residuale. In altri termini, un insieme residuale estende un  $G_\delta$  denso.
- Uno spazio topologico in cui ogni aperto è di seconda categoria si dice di Baire. Equivalentemente, se un'unione numerabile di chiusi  $\bigcup_{n \in \mathbb{N}} C_n$  ha parte interna non vuota  $A$ , allora esiste un  $C_i$  avente parte interna non vuota, perché  $A$  è di prima categoria.
- Passando al complementare nel punto precedente, uno spazio è di Baire se e solo se l'intersezione numerabile di aperti densi  $\bigcap_{n \in \mathbb{N}} A_n$  è densa.

**Teorema 3.6.1** (della categoria di Baire). *Uno spazio metrico  $(X, d)$  completo è uno spazio di Baire.*

*Dimostrazione.* Mostriamo che, dati  $A_n$  densi,  $\bigcap_{n \in \mathbb{N}} A_n$  è denso. Ciò significa mostrare che interseca  $B(x_0, r_0)$  per ogni coppia  $x_0, r_0$ . Definiamo la seguente successione di palle innestate:

$$\begin{cases} B_0 = B(x_0, r_0) \\ B_n = B(x_n, r_n) \\ \overline{B_n} \subseteq B_{n-1} \cap A_n \\ r_n \leq 1/n \end{cases}$$

Tale successione esiste perché  $B_{n-1} \cap A_n \neq \emptyset$ , per densità. Inoltre, i centri  $x_n$  definiscono una successione di Cauchy, dato che definitivamente tutti i centri stanno in  $B_n$  il cui raggio è infinitesimo. Allora  $x_n \rightarrow x$ . Il limite è contenuto in ogni  $\overline{B_n}$  e dunque in ogni  $A_n$ . Segue la tesi.  $\square$

Poiché uno spazio metrico è aperto nella sua topologia, vale il seguente corollario:

**Corollario 3.6.1.1.** *Uno spazio metrico completo non è unione numerabile di chiusi con parte interna vuota.*

### 3.6.1 Insieme di continuità di una funzione

**Teorema 3.6.2.** *Data una funzione a valore in uno spazio metrico  $f : (X, \tau) \rightarrow (Y, D)$ , l'insieme dei punti su cui  $f$  è continua è un  $G_\delta$ . In particolare, se è denso è residuale.*

*Dimostrazione.* Osserviamo che l'insieme considerato contiene tutti e soli i punti la cui oscillazione è nulla. In particolare, è l'intersezione numerabile degli insiemi  $A_n = \{x \in X \text{ t.c. } osc(x) \leq 2^{-n}\}$ , che sono aperti in quanto se un punto ha oscillazione inferiore a  $2^{-n}$  esiste un suo intorno in cui la funzione ha oscillazione minore, e dunque tutti i punti dell'intorno appartengono a  $A_n$ . Da cui la tesi.  $\square$

**Corollario 3.6.2.1.** *Poiché  $\mathbb{Q}$  è di prima categoria, non esistono funzioni reali di variabile reale continue solo sui razionali.*

Invece, è possibile costruire una funzione continua solo sugli irrazionali. Ad esempio, è possibile numerare i razionali, ottenendo una numerazione  $q_n$ , e definire

$$f(x) = \sum_{q_n < x} 2^{-n}$$

che è continua sugli irrazionali e discontinua (a destra) su  $\mathbb{Q}$ .



# Capitolo 4

## Operatori lineari

### 4.1 Continuità delle applicazioni lineari

Sia  $L : (E, \|\cdot\|_E) \rightarrow (F, \|\cdot\|_F)$  una mappa lineare. In dimensione finita, sappiamo che essa è continua (e peraltro differenziabile, ma questa è un'altra storia). In generale ciò non è vero in dimensione superiore. Tuttavia, sono davvero pochi i requisiti per la continuità, come mostra il seguente risultato.

**Teorema 4.1.1** (CNES per la Lipschitzianità). *Sono equivalenti:*

1.  $L$  è localmente limitata, ossia esiste un intorno  $U$  dell'origine e una costante positiva  $C$  tali che  $\forall x \in U \ \|Lx\|_F < C$ ;
2.  $L$  è continua;
3.  $L$  è sublineare, ossia esiste una costante  $K$  per cui  $\|Lx\|_F \leq K\|x\|_E$ ;
4.  $L$  è Lipschitziana.

*Dimostrazione.* Osserviamo innanzitutto che le proprietà 3 e 4 sono ovviamente equivalenti (basta porre  $x = v - w$ ) e che ciascuna proprietà discende dalla successiva. Ci basta allora mostrare che  $1 \implies 3$ . WLOG, possiamo assumere che l'intorno  $U$  sia una palla  $B(0, \varepsilon)$ . Allora, comunque si prenda un vettore  $x \in E$ , possiamo rinormalizzarlo perché entri nella palla, per esempio considerando il vettore  $y = \frac{x\varepsilon}{2\|x\|_E}$ . Allora

$$\left\| L \frac{x\varepsilon}{2\|x\|_E} \right\|_F \leq C \quad \longrightarrow \quad \|Lx\|_F \leq K\|x\|_E$$

□

La migliore costante di lipschitzianità di  $L$  si dice norma dell'operatore ed è data da

$$\|L\|^* = \sup_{x \in E \setminus \{0\}} \frac{\|Lx\|_F}{\|x\|_E}$$

Poiché le norme sono 1-omogenee, è possibile dividere sopra e sotto per la norma  $E$  di  $x$  e restringere la definizione agli  $x$  della sfera unitaria:

$$\|L\| = \sup_{\|x\|_E=1} \|Lx\|_F$$

Si può dimostrare che essa è effettivamente una norma sullo spazio  $L(E, F)$  delle funzioni lineari continue. Inoltre, la norma è compatibile con la composizione, nel senso che  $\|LM\| \leq \|L\|\|M\|$ , in quanto la composizione di funzioni lipschitziane è lipschitziana con costante il prodotto delle costanti.

### 4.1.1 Duale topologico

Se  $E$  è uno spazio vettoriale normato reale, l'insieme  $E^* = L(E, \mathbb{R})$  dei funzionali continui prende il nome di spazio duale topologico, ed è munito della norma duale  $\|\cdot\|^*$ . Come vedremo in 4.2, tale norma lo rende uno spazio di Banach.

Possiamo osservare facilmente che, se  $E$  ha dimensione finita, il duale topologico coincide con lo spazio duale studiato nel corso di Geometria 1 ed è dunque isomorfo (e anzi, isometrico con la norma  $\|\cdot\|_\infty$ ) allo stesso  $E$ . Vedremo facilmente che in generale ciò non è vero se si considerano spazi di dimensione infinita.

### 4.1.2 Applicazioni bilineari

Siano  $E, F, G$  spazi di Banach. Consideriamo lo spazio vettoriale delle applicazioni bilineari continue  $L^2(E \times F, G)$ . Si mostra facilmente che tale spazio è isomorfo a  $L(E, L(F, G))$  tramite l'applicazione che manda  $b$  bilineare in  $f(b) = \bar{b}$  tale che  $b(e, f) = (\bar{b}(e))(f)$ . Ciò ricalca la bigezione insiemistica tra  $G^{F \times E}$  e  $(G^F)^E$ .

La funzione  $f$  induce una norma su  $L^2(E \times F, G)$  in modo da diventare un'isometria:

$$\|b\| = \|\bar{b}\| = \sup_{x \in E, \|x\|=1} \|\bar{b}(x)\| = \sup_{x \in E, \|x\|=1} \sup_{y \in F, \|y\|=1} \|\bar{b}(x)(y)\| = \sup_{\|x\|=1, \|y\|=1} \|b(x, y)\|$$

## 4.2 Completezza delle funzioni lineari continue

Siano  $(E, \|\cdot\|_E)$  e  $(F, \|\cdot\|_F)$  due spazi vettoriali normati, di cui il secondo di Banach. Sia  $S_E$  la sfera unitaria nella norma di  $E$ . Consideriamo l'applicazione che a ogni applicazione lineare continua  $L$  associa la sua restrizione alla palla unitaria  $L|_{S_E}$ .

$$R : (L(E, F), \|\cdot\|^*) \longrightarrow (B(S_E, F), \|\cdot\|_\infty)$$

Per come è definita la norma operatoriale, abbiamo che

$$\|L|_{S_E}\|_\infty = \sup_{\|x\|_E=1} \|Lx\|_F = \|L\|^*$$

Perciò l'applicazione che si è definita è un'isometria, dunque  $L(E, F)$  si immerge isometricamente in  $B(S_E, F)$ . Inoltre, l'immagine dell'applicazione  $R$  è chiusa, in quanto il limite uniforme<sup>1</sup> di una successione di funzioni lineari è lineare. Perciò  $L(E, F)$ , che si identifica con la sua immagine, è uno spazio di Banach in quanto è un chiuso nell'insieme completo  $B(S_E, F)$ .

Peraltro, quanto detto consente di costruire in modo naturale un completamento di uno spazio normato. Indicato con  $E^* = L(E, \mathbb{R})$  il duale topologico e con  $E^{**}$  il biduale, entrambi completi in quanto lo è  $\mathbb{R}$ , la seguente funzione è un embedding isometrico

$$\begin{aligned} i : E &\longrightarrow E^{**} \\ x &\longrightarrow \text{val}_x : E^* \longrightarrow \mathbb{R} \\ &f \longrightarrow f(x) \end{aligned}$$

Infatti  $\|\text{val}_x\| = \sup_{\|f\|=1} \|f(x)\| \leq \|x\|$ . Inoltre, vale l'uguaglianza se si considera il funzionale che vale  $k$  su  $k \frac{x}{\|x\|}$  e zero altrove. Basta dunque considerare la chiusura di  $i(E)$  per ottenere un completamento di  $E$ .

<sup>1</sup>In realtà basta il limite puntuale.

## 4.3 Serie di Neumann

Se la lavagna non sapesse che stiamo parlando di operatori, vedrebbe una serie geometrica e penserebbe di essere ancora ad Analisi 1.

---

Pietro Majer

**Teorema 4.3.1.** *Sia  $E$  uno spazio di Banach e sia  $H \in L(E)$  di norma strettamente minore di 1. Allora l'operatore  $I - H$  è invertibile con inversa in  $L(E)$ , e la serie geometrica  $\sum_{k=0}^{\infty} H^k$  converge totalmente a  $(I - H)^{-1}$ .*

*Dimostrazione.* Osserviamo che la serie  $\sum_{k=0}^{\infty} H^k$  converge totalmente, in quanto  $\|H^k\| \leq \|H\|^k$  poiché la norma è compatibile con la composizione. Allora sia  $L = \sum_{k=0}^{\infty} H^k$ .

Osserviamo ora che l'operatore di composizione a sinistra con  $H$ , che associa a ogni  $A \in L(E)$  il prodotto  $HA$ , è lineare e localmente limitato, in quanto se  $\|A\| \leq 1$ , allora  $\|HA\| \leq \|H\|$ . Per quanto visto in 4.1.1 ciò implica che è continuo, dunque passa ai limiti. Nel nostro caso abbiamo che

$$HL = H \sum_{k=0}^{\infty} H^k = \sum_{k=0}^{\infty} H^{k+1} = L - I$$

Da cui  $(I - H)L = I$ . Allora  $I - H$  è invertibile e l'inversa è  $L$ . □

Osserviamo che, in realtà, basterebbe avere che  $\|H^n\| < 1$  per un qualche  $n \in \mathbb{N}$  per concludere che  $I - H$  ha come inversa la serie di Neumann. Si può infatti mostrare la convergenza totale frammentando la somma in  $n$  somme di potenze di  $\|H^n\|$ , e ripetere il ragionamento dell'operatore di moltiplicazione.

## 4.4 Perturbazioni lipschitziane dell'identità

Sia  $(E, \|\cdot\|)$  di Banach. Nello spirito di quanto visto in 4.3, mostriamo che nell'insieme delle funzioni limitate  $B(E)$  esiste una palla di funzioni invertibili attorno all'identità.

**Teorema 4.4.1.** *Sia  $\Omega \subseteq E$  aperto e sia  $g : \Omega \rightarrow E$  una funzione  $k$ -Lipschitz, con  $k < 1$ .<sup>2</sup> Sia  $f = id_E + g$ . Allora vale che*

1. *Detta  $D(a, r)$  la palla chiusa di centro  $a$  e raggio  $r$ , se  $D(a, r) \subseteq \Omega$  allora vale la relazione  $D(f(a), (1 - k)r) \subseteq f(D(a, r))$ ;*
2. *La stessa relazione vale tra palle aperte;*
3.  *$f(\Omega)$  è aperto;*
4.  *$f : \Omega \rightarrow f(\Omega)$  è un omeomorfismo bilipschitziano, di costante  $1 + k$  e la cui inversa ha costante  $\frac{1}{1 - k}$ .*

*Dimostrazione.* Sia  $y \in D(f(a), (1 - k)r)$ . Mostriamo che è immagine di un certo  $x \in D(a, r)$ , ossia che esiste un  $x$  tale che  $f(x) = x + g(x) = y$ . Possiamo riscrivere la relazione come  $x = y - g(x)$  e osservare che stiamo cercando il punto fisso della funzione  $h(x) = y - g(x)$ .

Verifichiamo allora che quest'ultima è una contrazione da  $D(a, r)$  in sé. Intanto  $h$  è  $k$ -lipschitziana perché lo è  $g$ . Inoltre, dato  $z \in D(a, r)$ , abbiamo che

$$\|h(z) - a\| = \|y - g(z) - a\| = \|y - g(z) - f(a) + g(a)\| \leq \|y - f(a)\| + \|g(z) - g(a)\| \leq (1 - k)r + k\|z - a\| \leq r$$

Allora esiste un unico punto fisso di  $h$ , cosa che dimostra il primo punto del teorema.

Per mostrare che la stessa relazione vale per palle aperte, basta prendere tutte le palle chiuse dentro una

---

<sup>2</sup>Non la chiamo contrazione, termine preferito per le funzioni che mandano il dominio in sé.

palla aperta data, applicare il risultato precedente e passare all'estremo superiore sui raggi. Risulta allora ovvio che  $f(\Omega)$  è aperto, in quanto contiene una palla (aperta, volendo) di ogni suo punto.

Ora, osserviamo che  $f$  è iniettiva. Infatti se per due punti  $x, y$  vale che  $x + g(x) = y + g(y)$ , passando alle norme otteniamo  $\|x - y\| = \|g(x) - g(y)\| \leq k\|x - y\|$ , da cui  $x = y$ . Inoltre è una funzione lipschitziana perché somma di lipschitziane, e la sua costante è la somma delle costanti. Infine, vale la relazione

$$\|f(x) - f(y)\| = \|x - y + g(x) - g(y)\| \geq \|x - y\| - \|g(x) - g(y)\| \geq \|x - y\|(1 - k)$$

Da cui, chiamando  $x' = f(x)$  e analogamente  $y' = f(y)$ , otteniamo

$$\|f^{-1}(x') - f^{-1}(y')\| \leq \frac{1}{1 - k} \|x' - y'\|$$

ossia anche l'inversa di  $f$  è lipschitziana. □

**Corollario 4.4.1.1.** *Se nelle ipotesi del teorema precedente  $\Omega = E$ , allora anche  $f(\Omega) = E$ .*

*Dimostrazione.* Basta osservare che  $E$  contiene palle centrate nei suoi punti di dimensione arbitraria, e dunque ogni punto di  $E$  viene prima o poi raggiunto aumentando  $r$  nella dimostrazione precedente. □

Sempre nello spirito di quanto visto per gli operatori lineari, possiamo determinare una classe di funzioni in  $B(\Omega, E)$  che formano un aperto.

**Corollario 4.4.1.2.** *Sia  $\Omega \subseteq (E, \|\cdot\|)$  e sia  $f : \Omega \rightarrow f(\Omega)$  un omeomorfismo con inversa lipschitziana di costante  $k$ . Data una funzione  $h$  su  $\Omega$  di costante di Lipschitz minore di  $\frac{1}{k}$ , allora anche  $f + h$  è un omeomorfismo su  $\Omega$  con inversa lipschitziana. In altri termini, gli omeomorfismi con inversa lipschitziana sono un aperto delle funzioni  $B(\Omega, E)$ .*

*Dimostrazione.* Basta osservare che  $f + h = (id + h \circ f^{-1}) \circ f$  e che per il primo membro vale il teorema precedente. □

La lezione è finita, andate in pace.

Papa Pietro IV

## 4.5 Gli isomorfismi lineari sono un aperto

**Teorema 4.5.1.**  *$GL(E)$  è aperto in  $L(E)$ . In altri termini, perturbando leggermente un'applicazione invertibile essa resta tale.*

*Dimostrazione.* Sia  $A \in GL(E)$  e mostriamo che la palla  $B\left(A, \frac{1}{\|A^{-1}\|}\right)$  nella norma operatoriale è totalmente contenuta in  $GL(E)$ . Consideriamo l'operatore di moltiplicazione a sinistra per  $A^{-1}$ , che è un automorfismo di  $GL(E)$  e manda la palla considerata in  $B(I, 1)$ . Infatti, se  $H \in B\left(A, \frac{1}{\|A^{-1}\|}\right)$ , vale che  $\|A^{-1}H - I\| < \frac{1}{\|A^{-1}\|}$  e dunque

$$\|I - A^{-1}H\| \leq \|A^{-1}\| \|A^{-1}H - I\| < 1$$

Per quanto visto in 4.3,  $B(I, 1)$  contiene solo funzioni invertibili, perché se  $\|I - K\| < 1$ , allora anche  $I - (I - K) = K$  è invertibile. Perciò anche la controimmagine  $B\left(A, \frac{1}{\|A^{-1}\|}\right)$  è composta solo da applicazioni invertibili. □

## 4.6 $GL(E)$ come gruppo di Lie

**Teorema 4.6.1.** *Le operazioni di composizione e di inversione in  $GL(E)$  sono continue.*

*Dimostrazione.* Consideriamo innanzitutto la composizione

$$\begin{aligned} \circ : G \times G &\longrightarrow G \\ (L, H) &\longrightarrow LH \end{aligned}$$

Fissata una coppia di applicazioni lineari  $(L_0, H_0)$ , mostriamo che la norma di  $LH - L_0H_0$  è infinitesima per  $(L, H) \rightarrow (L_0, H_0)$ . Sommando e sottraendo come Dio comanda otteniamo

$$\|LH - L_0H_0\| \leq \|L\|\|H - H_0\| + \|L - L_0\|\|H_0\|$$

Fissato un  $\varepsilon$ , scegliamo un intorno di  $L_0$  dove  $\|L - L_0\|\|H_0\| \leq \varepsilon/2$ . Dopodiché, dato che la norma di  $L$  si stima con  $\|L_0\| + \varepsilon/\|H_0\|$ , scegliamo un intorno di  $H_0$  dove  $(\|L_0\| + \varepsilon/\|H_0\|)\|H - H_0\| \leq \varepsilon/2$ . In tal modo abbiamo ottenuto la continuità del prodotto.

Passiamo ora all'inversione

$$\begin{aligned} \circ : G &\longrightarrow G \\ L &\longrightarrow L^{-1} \end{aligned}$$

Per quanto osservato in 4.5, per ogni  $H$  di norma sufficientemente piccola vale

$$(L + H)^{-1} = L^{-1} \sum_{k=0}^{+\infty} (-1)^k (L^{-1}H)^k = L^{-1} - L^{-1}HL^{-1} + (L^{-1}H)^2L^{-1} \sum_{k=0}^{+\infty} (-1)^k (L^{-1}H)^k$$

Se mostriamo che l'ultimo termine è un  $o(\|H\|)$  in norma ricaviamo che l'inversione è continua, anzi differenziabile con differenziale dato da  $-L^{-1}HL^{-1}$ .

$$\|(L^{-1}H)^2L^{-1} \sum_{k=0}^{+\infty} (-1)^k (L^{-1}H)^k\| \leq \|L^{-1}\|^3 \|H\|^2 \frac{1}{1 - \|L^{-1}\|\|H\|} = o(\|H\|)$$

Come volevasi dimostrare. □

Nel caso particolare in cui  $E = \mathbb{R}^n$ ,  $GL(E)$  è uno spazio di matrici, che si identifica con un aperto di  $\mathbb{R}^{n \times n}$ . In altri termini,  $GL(\mathbb{R}^n)$  è sia una varietà differenziabile<sup>3</sup> sia un gruppo le cui operazioni sono continue. Un gruppo che rispetta queste proprietà si dice gruppo di Lie.

## 4.7 Isometrie affini

In generale, una funzione che conserva la distanza non è per forza affine. Ad esempio, data una funzione 1-lipschitziana  $f : \mathbb{R} \rightarrow \mathbb{R}$ , l'isometria  $g : (\mathbb{R}, d_\infty) \rightarrow (\mathbb{R}^2, d_\infty)$  che manda  $x$  in  $(x, f(x))$  può fare schifo a piacere. Invece, si può mostrare quanto segue.

**Teorema 4.7.1** (Mazur-Ulam). *Un'isometria surgettiva tra spazi vettoriali normati è affine.*

*Dimostrazione.* Sia  $f : (E, \|\cdot\|_E) \rightarrow (F, \|\cdot\|_F)$  l'isometria considerata. Dire che è affine equivale a dire che conserva le combinazioni affini. In particolare, basta mostrare che conserva il punto medio, ossia che  $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$ . In tal modo, si riesce a ottenere ogni combinazione affine i cui coefficienti sono potenze di 2, e si conclude sfruttando la continuità delle isometrie.

Fissiamo allora due punti  $x, y$  e definiamo la "deficienza" della funzione come

$$\delta(f, x, y) = \left\| f\left(\frac{x+y}{2}\right) - \frac{f(x)+f(y)}{2} \right\|_F$$

Tale valore misura quanto la funzione "non è affine", e il nostro obiettivo è mostrare che è costantemente nulla. Mostriamo innanzitutto che, comunque si scelga  $f$ , la sua deficienza si stima dall'alto con  $\frac{1}{2}\|x-y\|_E$ . Infatti

$$\delta(f, x, y) = \left\| f\left(\frac{x+y}{2}\right) - \frac{f(x)+f(y)}{2} \right\|_F \leq \frac{1}{2} \left\| f\left(\frac{x+y}{2}\right) - f(x) \right\|_F + \frac{1}{2} \left\| f\left(\frac{x+y}{2}\right) - f(y) \right\|_F$$

<sup>3</sup>Vedasi 10.

Essendo  $f$  un'isometria abbiamo che

$$\delta(f, x, y) \leq \frac{1}{2} \left\| \frac{x+y}{2} - x \right\|_E + \frac{1}{2} \left\| \frac{x+y}{2} - y \right\|_E = \frac{1}{2} \|x - y\|_E$$

Ora, consideriamo la funzione  $f^*(v) = f^{-1}(f(x) + f(y) - f(v)) = f^{-1} \circ A \circ f(v)$ , dove  $Aw = f(x) + f(y) - w$  è un'isometria affine. Osserviamo che  $f^*$  è ben definita, perché la surgettività di  $f$  ci permette di scrivere  $f^{-1}$ , e che è un'isometria, in quanto composizione di isometrie. La sua deficienza vale

$$\begin{aligned} \delta(f^*, x, y) &= \left\| f^* \left( \frac{x+y}{2} \right) - \frac{f^*(x) + f^*(y)}{2} \right\|_E = \\ &= \left\| f^{-1} \left( f(x) + f(y) - f \left( \frac{x+y}{2} \right) \right) - \frac{f^{-1}(f(x) + f(y) - f(x)) + f^{-1}(f(x) + f(y) - f(y))}{2} \right\|_E = \\ &= \left\| f^{-1} \left( f(x) + f(y) - f \left( \frac{x+y}{2} \right) \right) - \frac{x+y}{2} \right\|_E = \left\| f(x) + f(y) - f \left( \frac{x+y}{2} \right) - f \left( \frac{x+y}{2} \right) \right\|_F \end{aligned}$$

L'ultimo passaggio segue dal fatto che  $f^{-1}$  è un'isometria. Osserviamo allora che

$$\delta(f^*, x, y) = 2\delta(f, x, y)$$

Procedendo induttivamente, se  $f$  non fosse affine, potremmo costruire un'isometria la cui deficienza è grande a piacere. Ma ciò contraddirebbe il fatto che la deficienza è limitata.  $\square$

## Capitolo 5

# Topologia e convergenza in spazi di funzioni

### 5.1 Criteri di convergenza uniforme

#### 5.1.1 Teorema di Dini

**Teorema 5.1.1** (Dini). *Sia  $X$  uno spazio topologico compatto e sia  $f_n : X \rightarrow \mathbb{R}$  una successione di funzioni continue che converge puntualmente a una funzione continua  $f$  in modo monotono. Allora la convergenza è uniforme.*

*Dimostrazione.* WLOG possiamo supporre che le  $f_n$  convergano decrescendo a  $f \equiv 0$ . Per assurdo la convergenza non sia uniforme, ossia esiste un  $\varepsilon$  tale che per ogni  $n$  l'insieme  $A_n = \{f_n(x) \geq \varepsilon\}$  sia non vuoto. Poiché gli  $A_n$  sono una successione di chiusi innestati con la proprietà dell'intersezione finita, la loro intersezione è non vuota, ossia esiste un  $x$  per cui  $f_n(x) \geq \varepsilon$  per ogni  $n$ . Ciò viola l'ipotesi di convergenza puntuale a zero.  $\square$

### 5.2 Approssimazioni polinomiali

Scopo della seguente sezione è arrivare, dopo lunghi e penosi calcoli, a dimostrare il teorema di Weierstrass, secondo cui i polinomi su  $[0, 1]$  sono densi nelle funzioni continue reali. La prossima sottosezione sarà una serie di inutili complicazioni che vi consiglio di skippare allegramente, andandovi solo a guardare la definizione dei polinomi di Bernstein.

#### 5.2.1 Strani operatori lineari

**Definizione 14** (Derivata approssimata). *L'operatore lineare  $D_n : C^0([0, 1], \mathbb{R}) \rightarrow C^0([0, 1], \mathbb{R})$  definito come*

$$D_n(u)(x) = n \left( u \left( \left( 1 - \frac{1}{n} \right) x + \frac{1}{n} \right) - u \left( \left( 1 - \frac{1}{n} \right) x \right) \right)$$

*si dice derivata approssimata  $n$ -esima.*

Si osservi che, se  $u$  è derivabile, per teorema di Lagrange  $D_n(u)(x) = u'(\zeta)$ , con

$$|\zeta - x| \leq \frac{x}{n} + \left| \zeta - \frac{n-1}{n}x \right| \leq \frac{2}{n}$$

In particolare, se  $u \in C^1$   $\|D_n(u) - u'\| \leq \omega_{u'}\left(\frac{2}{n}\right)$ , e dunque le derivate approssimate convergono uniformemente a  $u'$ .

Ora, dato  $a > 0$  definiamo gli operatori lineari di traslazione e di dilatazione come segue

$$\tau_a(f)(x) = f(x + a)$$

$$\delta_a(f)(x) = f(ax)$$

Osserviamo che vale la seguente regola di commutazione

$$\tau_b \delta_a = \delta_a \tau_{ab}$$

Infatti

$$\tau_b \delta_a(f)(x) = \tau_b(\delta_a(f))(x) = \tau_b(f)(ax) = \tau_b(f(a \cdot))(x) = f(a \cdot)(x+b) = f(a(x+b))$$

$$\delta_a \tau_{ab}(f)(x) = \delta_a(f)(x+ab) = f(ax+ab)$$

Possiamo dunque riscrivere la derivata approssimata come l'operatore

$$D_n = n \delta_{\frac{n-1}{n}} \left( \tau_{\frac{1}{n}} - id \right) = n \left( \tau_{\frac{1}{n-1}} - id \right) \delta_{\frac{n-1}{n}}$$

## 5.2.2 Polinomi di Bernstein

Costruiamo ora una successione di operatori lineari  $B_n : C^0([0,1], \mathbb{R}) \rightarrow (\mathbb{R}[x], [0,1])$  a valori nei polinomi che soddisfino la relazione

$$\begin{cases} B_0(f)(x) = f(0) \text{ per ogni } x \in [0,1] \\ B_n(f)(0) = f(0) \text{ per ogni } n \in \mathbb{N} \\ DB_n = B_{n-1} D_n \end{cases}$$

dove l'operatore  $D$  è la derivata.

Innanzitutto mostriamo che, se esistono, assumono effettivamente valori nei polinomi. Procediamo induttivamente su  $n$ . Per costruzione  $B_0(f) \in \mathbb{R}[x]$ . Basta poi osservare che  $B_n(f)(x) = \int_0^x B_{n-1} D_n(f)(t) dt$ , e per ipotesi induttiva  $B_{n-1}$  ha valori nei polinomi. Si osservi che la dimostrazione mostra anche che  $B_n(f)$  ha grado massimo  $n$ .

Essendo polinomi, i  $B_n(f)$  coincidono con il loro sviluppo di Taylor in zero. Dunque, detto  $va_x$  l'operatore lineare di valutazione nel punto  $x$  e analogamente  $va_0$  la valutazione in zero, abbiamo che

$$B_n(f)(x) = va_x(B_n(f)) = \sum_{k=0}^n \frac{x^k}{k!} va_0 D^k B_n(f)$$

Attenzione:  $va_0$ , non  $va_x$ . Infatti stiamo calcolando le derivate nel centro dello sviluppo, cioè in zero, mentre  $x$  funge da coefficiente nella sommatoria.

Occupiamoci ora di capire chi siano i termini della sommatoria. Applicando ripetutamente la definizione dei  $B_n$  otteniamo

$$D^k B_n = B_{n-k} D_{n-k+1} \dots D_n = B_{n-k} (n-k+1) \delta_{\frac{n-k}{n-k+1}} \left( \tau_{\frac{1}{n-k+1}} - id \right) \dots n \delta_{\frac{n-1}{n}} \left( \tau_{\frac{1}{n}} - id \right)$$

Raccogliendo i coefficienti numerici e applicando le regole di commutazione di  $\delta$  e  $\tau$  risulta

$$D^k B_n = B_{n-k} \frac{n!}{(n-k)!} \delta_{\frac{n-k}{n-k+1}} \dots \delta_{\frac{n-1}{n}} \left( \tau_{\frac{1}{n}} - id \right)^k$$

Infatti  $\delta_{\frac{n-1}{n}} \left( \tau_{\frac{1}{n}} - id \right) = \left( \tau_{\frac{1}{n-1}} - id \right) \delta_{\frac{n-1}{n}}$ , e dunque è possibile portare tutti i fattori  $\left( \tau_{\frac{1}{n-1}} - id \right)$  a destra. Ora, componendo tra loro i  $\delta$  otteniamo

$$D^k B_n = B_{n-k} \frac{n!}{(n-k)!} \delta_{\frac{n-k}{n}} \left( \tau_{\frac{1}{n}} - id \right)^k$$

Ora sostituiamo nella serie di Taylor, ottenendo

$$B_n(f)(x) = \sum_{k=0}^n \binom{n}{k} x^k va_0 B_{n-k} \delta_{\frac{n-k}{n}} \left( \tau_{\frac{1}{n}} - id \right)^k (f) = \sum_{k=0}^n \binom{n}{k} x^k va_0 \left( \tau_{\frac{1}{n}} - id \right)^k (f)$$

L'ultimo passaggio segue dal fatto che  $B_n(f)(0) = f(0)$  e che  $\delta(f)(0) = f(0)$ . Possiamo dunque scrivere la somma come un binomio di Newton di operatori lineari

$$B_n(f)(x) = va_0 \sum_{k=0}^n \binom{n}{k} x^k (\tau_{\frac{1}{n}} - id)^k (f) = va_0 \left( id + x (\tau_{\frac{1}{n}} - id) \right)^n (f)$$

$$B_n(f)(x) = va_0 \left( id(1-x) + x\tau_{\frac{1}{n}} \right)^n (f) = va_0 \sum_{k=0}^n \binom{n}{k} (1-x)^{n-k} x^k \tau_{\frac{1}{n}}^k (f)$$

$$B_n(f)(x) = \sum_{k=0}^n \binom{n}{k} (1-x)^{n-k} x^k f\left(\frac{k}{n}\right)$$

Chi ha già dato EPS riconoscerà il valore atteso della variabile binomiale di parametro  $x$ .

### 5.2.3 Proprietà dei polinomi di Bernstein

1.  $B_n(1) = 1$ , dove 1 è la funzione costante.

Questo è banalmente vero perché  $B_n(1)(x) = \sum_{k=0}^n \binom{n}{k} (1-x)^{n-k} x^k = 1$ .

2. Per  $n > 0$ ,  $B_n(x) = x$ , dove  $x$  è la funzione identità.

Infatti  $B_1(x) = x$ . Inoltre, se  $B_{n-1}(x) = x$ , allora

$$B_n(x) = \int_0^x B_{n-1} D_n x(t) dt = \int_0^x B_{n-1} n \left( \left(1 - \frac{1}{n}\right)t + \frac{1}{n} - \left(1 - \frac{1}{n}\right)t \right) dt$$

$$B_n(x) = \int_0^x B_{n-1}(1) dt = \int_0^x dt = x$$

Dalle prime due condizioni otteniamo che  $B_n(ax + b) = ax + b$  per linearità.

3.  $B_n(x^2)$  converge uniformemente a  $x^2$ .

Infatti, esplicitando  $D_n$  otteniamo

$$B_n(x) = \int_0^x B_{n-1} D_n x(t) dt = \int_0^x B_{n-1} n \left( \left( \left(1 - \frac{1}{n}\right)t + \frac{1}{n} \right)^2 - \left(1 - \frac{1}{n}\right)^2 t^2 \right) dt$$

$$B_n(x) = \int_0^x B_{n-1} \left( \frac{1}{n} + 2 \left(1 - \frac{1}{n}\right)t \right) dt = \int_0^x \frac{1}{n} + 2 \left(1 - \frac{1}{n}\right)t dt = \frac{x}{n} + \left(1 - \frac{1}{n}\right)x^2 \rightarrow x^2$$

4. I  $B_n$  sono positivi, ossia mandano funzioni positive in funzioni positive. Dunque i  $B_n$  sono anche monotoni, ossia se  $f - g \geq 0$  allora anche  $B_n(f) \geq B_n(g)$ .

La proprietà segue banalmente dalla formula esplicita.

5. I  $B_n$  sono continui. In generale, ogni operatore lineare positivo è continuo.

Basta mostrare che  $\|B_n\| = \sup_{\|f\|_\infty=1} \|B_n(f)\|_\infty$  è finito. Ma se  $\|f\|_\infty = 1$  allora  $-1 \leq f \leq 1$ , e per monotonia  $-B_n(1) \leq B_n(f) \leq B_n(1)$ , da cui la tesi.

6. Dalla proprietà precedente segue che i  $B_n$  sono equilipschitziani di costante  $1 = \|B_n(1)\|$ .

### 5.2.4 I polinomi sono densi nelle funzioni continue

**Teorema 5.2.1** (Weierstrass). *I polinomi sono densi in  $C^0([0, 1], \mathbb{R})$ . Più precisamente, data una funzione continua  $f$ , la successione dei polinomi di Bernstein di  $f$  vi converge uniformemente.*

Si potrebbero fare una marea di conti coi polinomi di Bernstein, però direi che ci siamo già un po' rotti di maneggiarli. Allora possiamo ricorrere al seguente trucchetto:

**Teorema 5.2.2** (Korovkin). *Sia  $T_n : (C^0[0, 1], \mathbb{R}) \rightarrow (C^0[0, 1], \mathbb{R})$  una successione di operatori lineari positivi tali che, se  $p \in \mathbb{R}_2[x]$  è un polinomio di grado al più 2, i  $T_n(p)$  convergono uniformemente a  $p$ . Allora per ogni  $f \in C^0$  i  $T_n(f)$  convergono uniformemente a  $f$ .*

Il teorema di Weierstrass segue osservando che, nella sottosezione precedente, abbiamo mostrato esattamente le proprietà richieste per i polinomi di Bernstein.

*Dimostrazione.* Mostriamo dapprima il seguente:

**Lemma 5.2.3.** *Data  $f \in (C^0[0, 1], \mathbb{R})$  e  $\varepsilon > 0$ , esiste un numero finito di polinomi  $p_i \in \mathbb{R}_2[x]$  tali che  $f \leq \min_{i=1, \dots, m} p_i \leq f + \varepsilon$ .*

Infatti, fissato un punto  $a \in [0, 1]$ , sia  $p_a(x) = f(a) + \varepsilon/2 + M(x - a)^2$ . Scegliamo  $M$  in modo che  $f(x) \leq p_a(x)$  su tutto  $[0, 1]$ . Ad esempio, possiamo scegliere  $M = \max_{x \in [0, 1]} \frac{f(x) - f(a) - \varepsilon/2}{(x - a)^2}$ , che è ben definito perché il rapporto è continuo all'infuori di  $x = a$ , dove diverge a  $-\infty$ .

Per ogni  $a$  esiste un intorno  $I_a$  in cui  $p_a(x) \leq f(x) + \varepsilon$  per continuità. Dunque possiamo estrarre dagli  $I_a$  un sottoricoprimento finito e considerare il minimo dei  $p_a$  estratti.

Ora, data  $f \in (C^0[0, 1], \mathbb{R})$  fissiamo  $\varepsilon$  e consideriamo i polinomi  $p_1, \dots, p_m$  descritti nel lemma. Per positività degli operatori abbiamo che, per ogni  $k = 1, \dots, m$  e per ogni  $n \in \mathbb{N}$ ,  $T_n(f) \leq T_n(p_k)$ . Inoltre definitivamente  $T_n(p_k) \leq p_k + \varepsilon$  per ogni  $k$ , dato che  $T_n(p_k)$  converge a  $p_k$ . Dunque definitivamente

$$T_n(f) \leq \min_{k=1, \dots, m} p_k + \varepsilon \leq f + 2\varepsilon$$

Per linearità, ripetendo il ragionamento con  $-f$ , otteniamo che, definitivamente,

$$T_n(f) \geq f + 2\varepsilon$$

Da cui la tesi. □

Se la funzione considerata è sufficientemente regolare, la convergenza dei polinomi di Bernstein è molto più forte:

**Teorema 5.2.4.** *Sia  $f \in C^1([0, 1], \mathbb{R})$ . Allora la successione  $B_n(f)$  converge a  $f$  in norma  $C^1$ , ossia anche le derivate convergono uniformemente a  $f'$ .*

*Dimostrazione.* Innanzitutto, possiamo scrivere

$$DB_n(f) = B_{n-1}D_n(f)$$

Ricordiamo che, data una successione di funzioni equicontinue tra spazi metrici  $h_n$  convergenti puntualmente a  $h$  e una successione  $x_n \rightarrow x$ , vale che  $h_n(x_n) \rightarrow h(x)$ . Infatti, detto  $\omega$  il loro modulo di continuità,

$$d(h_n(x_n), h(x)) \leq d(h_n(x_n), h_n(x)) + d(h_n(x), h(x)) \leq \omega(d(x_n, x)) + d(h_n(x), h(x)) = o(1)$$

Dunque è sufficiente notare che, come operatori sulle funzioni continue, i  $B_n$  convergono puntualmente all'identità (per teorema di Weierstrass) e  $D_n(f)$  converge uniformemente a  $f'$  per quanto osservato in 5.2.1. Inoltre, i  $B_n$  sono equilipschitziani di costante 1. Allora  $B_{n-1}D_n(f) \rightarrow D(f)$  uniformemente, come volevasi dimostrare. □

Con un po' più di conti si riesce a mostrare anche che, se  $f \in C^k$ , la convergenza dei polinomi di Bernstein è in norma  $C^k$ .

### 5.2.5 Caso vettoriale

Si può estendere la definizione dei polinomi di Bernstein per funzioni  $f : [0, 1] \rightarrow (E, \|\cdot\|)$  a valori vettoriali. Definendo  $B_n, D_n, D$  esattamente allo stesso modo possiamo replicare in toto gli argomenti visti finora, eccetto le considerazioni sulla positività (che hanno poco senso in spazi non ordinati) e il teorema di Korovkin (nel quale si fa riferimento alla funzione  $x^2$ , non ben definita se  $x$  assume valori in un generico spazio di Banach).

Il primo problema si argina osservando che

$$\|B_n(f)(x)\| = \left\| \sum_{k=0}^n \binom{n}{k} (1-x)^{n-k} x^k f\left(\frac{k}{n}\right) \right\| \leq \sum_{k=0}^n \binom{n}{k} (1-x)^{n-k} x^k \left\| f\left(\frac{k}{n}\right) \right\| = B_n(\|f\|_\infty)(x)$$

Dunque anche in questo caso i  $B_n$  sono equilipschitziani di costante 1, e il teorema di convergenza  $C^1$  è ancora valido.<sup>1</sup>

Per quanto riguarda la seconda questione, serve una dimostrazione del teorema di Weierstrass nuova di pacca.

*Dimostrazione.* Sia  $\omega$  un modulo di continuità per  $f$  e sia  $\delta > 0$ . Mostriamo che

$$\|f(x) - B_n(f)(x)\| \leq \omega(\delta) + 2 \frac{\|f\|_\infty}{\delta} \sqrt{\frac{x(1-x)}{n}} \leq \omega(\delta) + \frac{\|f\|_\infty}{\delta\sqrt{n}}$$

Da cui, definitivamente,  $\|f - B_n(f)\|_\infty \leq \omega(\delta) + \delta$  per ogni  $\delta > 0$ . Infatti, sfruttando lo sviluppo del binomio di Newton abbiamo che

$$\begin{aligned} \|f(x) - B_n(f)(x)\| &= \left\| \sum_{k=0}^n f(x) \binom{n}{k} x^k (1-x)^{n-k} - \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \right\| \\ \|f(x) - B_n(f)(x)\| &\leq \sum_{k=0}^n \left\| f(x) - f\left(\frac{k}{n}\right) \right\| \binom{n}{k} x^k (1-x)^{n-k} \end{aligned}$$

Spezziamo la somma in due parti:

$$\begin{aligned} \|f(x) - B_n(f)(x)\| &\leq \sum_{|k/n-x| < \delta} \left\| f(x) - f\left(\frac{k}{n}\right) \right\| \binom{n}{k} x^k (1-x)^{n-k} + \sum_{|k/n-x| \geq \delta} \left\| f(x) - f\left(\frac{k}{n}\right) \right\| \binom{n}{k} x^k (1-x)^{n-k} \\ \|f(x) - B_n(f)(x)\| &\leq \omega(\delta) + 2\|f\|_\infty \sum_{|k/n-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k} \end{aligned}$$

Basta allora stimare uniformemente  $\sum_{|k/n-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k}$ . Osserviamo che

$$\sum_{|k/n-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{\delta} \sum_{0 \leq k \leq n} \left| \frac{k}{n} - x \right| \binom{n}{k} x^k (1-x)^{n-k}$$

Applicando la disuguaglianza di Cauchy-Schwarz al prodotto scalare dei vettori  $u_k = \left| \frac{k}{n} - x \right| \sqrt{\binom{n}{k} x^k (1-x)^{n-k}}$  e  $v_k = \sqrt{\binom{n}{k} x^k (1-x)^{n-k}}$  si ha che

$$\sum_{|k/n-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{\delta} \sqrt{\sum_{0 \leq k \leq n} \left( \frac{k}{n} - x \right)^2 \binom{n}{k} x^k (1-x)^{n-k}} \sqrt{\sum_{0 \leq k \leq n} \binom{n}{k} x^k (1-x)^{n-k}}$$

Esplicitando i termini del quadrato ritroviamo i polinomi di Bernstein scalari delle funzioni 1,  $x$ ,  $x^2$ :

$$\sum_{|k/n-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{\delta} \sqrt{B_n(x^2) - 2xB_n(x) + x^2B_n(1)} = \frac{1}{\delta} \sqrt{\left(1 - \frac{1}{n}\right)x^2 + \frac{x}{n} - x^2} = \frac{1}{\delta} \sqrt{\frac{x(1-x)}{n}}$$

Come volevasi dimostrare.  $\square$

Notiamo che l'estensione dei risultati del caso scalare alle curve a valori vettoriali ci regala una discreta soddisfazione, che completa le considerazioni della sezione 3.3.1:

**Corollario 5.2.4.1.** *Le curve  $C^\infty$  sono dense nelle curve  $C^k$  con la norma  $C^k$ .*

<sup>1</sup>In realtà ciò non è verissimo, perché avevamo usato che  $D_n(f)(x) = f'(\zeta)$  e l'esistenza di tale elemento era garantita dal teorema di Lagrange per curve a valori reali. La generalizzazione di quest'ultimo al caso vettoriale, che vedremo nella sezione 6.4.3, permette solo di ottenere stime sul modulo di  $D_n$ , che non mi sembrano fare al caso nostro.

## 5.3 Stoned

Cerchiamo ora di generalizzare la trattazione del teorema di Weierstrass da un punto di vista più algebrico. D'ora innanzi  $C(X)$  saranno le funzioni continue da  $X$  a valori reali.

**Definizione 15** (Algebra). *Uno spazio  $\mathcal{A}$  munito di tre operazioni  $(+, *, \cdot)$ , di cui le prime due sono interne e la terza è un prodotto per uno scalare reale, si dice  $\mathbb{R}$ -algebra se*

- $(A, +, \cdot)$  è uno spazio vettoriale reale;
- $(A, +, *)$  è un anello commutativo.

Si osservi che  $(C(X), +, *, \cdot)$ , dove  $*$  rappresenta il prodotto valore per valore, è un'algebra reale.

**Definizione 16** (Reticolo). *Una famiglia di funzioni  $\mathcal{R}$  è un reticolo se per ogni  $f, g \in \mathcal{R}$  anche  $\min\{f(x), g(x)\}, \max\{f(x), g(x)\} \in \mathcal{R}$ .*

Si osservi che se  $\mathcal{R}$  è anche un'algebra è sufficiente richiedere che, se  $f \in \mathcal{R}$ , allora anche  $|f| \in \mathcal{R}$ . Infatti, date  $f, g \in \mathcal{R}$ , si ha che  $|f - g| \in \mathcal{R}$ . Dopodiché basta osservare che massimo e minimo si ottengono dal seguente sistema:

$$\begin{cases} \max\{f(x), g(x)\} + \min\{f(x), g(x)\} = f + g \\ \max\{f(x), g(x)\} - \min\{f(x), g(x)\} = |f - g| \end{cases}$$

**Lemma 5.3.1.** *Una  $\mathbb{R}$ -algebra  $\mathcal{A}$  di funzioni limitate che sia chiusa per convergenza uniforme è un reticolo.*

*Dimostrazione.* Sia  $f \in \mathcal{A}$ . A meno di dividere per uno scalare, posso supporre che  $\|f\|_\infty \leq 1$ , da cui anche  $0 \leq f^2 \leq 1$ . Ora, detti  $B_n$  i polinomi di Bernstein, abbiamo che

$$|f(x)| = \sqrt{f^2(x)} = \lim_{n \rightarrow +\infty} B_n(\sqrt{t})(f^2(x))$$

Basta allora osservare che, poiché  $f^2(x) \in [0, 1]$ , i polinomi di Bernstein convergono uniformemente (su  $[0, 1]$ ) a  $\sqrt{t}$ , e dunque calcolando il tutto in  $f^2(x)$  otteniamo la convergenza uniforme di  $B_n(\sqrt{t})(f^2(x))$  a  $|f(x)|$ . Poiché l'algebra è chiusa per convergenza uniforme si ha la tesi.  $\square$

### 5.3.1 Teorema di Stone

**Lemma 5.3.2.** *Sia  $X$  compatto. Sia  $\mathcal{R} \subseteq C(X, \mathbb{R})$  un reticolo tale che, per ogni due punti  $x, y \in X$  e ogni due valori reali  $a, b$  esiste una funzione  $f \in \mathcal{R}$  tale che  $f(x) = a$ ,  $f(y) = b$  (cioè esiste sempre un'interpolazione reale a due nodi). Allora  $\mathcal{R}$  è denso in  $C(X, \mathbb{R})$ .*

*Dimostrazione.* La dimostrazione è molto simile a quella del teorema di Korovkin. Sia  $f \in C(X)$  e sia  $\varepsilon > 0$ . Per ipotesi, per ogni coppia di punti  $x, y \in X$  e per ogni scelta di due valori di interpolazione esiste  $g_{x,y} \in \mathcal{R}$  che assume quei valori. Scegliamo delle funzioni  $g_{x,y}$  tali che  $f(z) < g_{x,y}(z) < f(z) + \varepsilon$  in un intorno aperto  $U_{x,y}$  di  $x, y$ . Fissato  $x$ , il ricoprimento  $\{U_{x,y}\}_{y \in X}$  ammette un sottoricoprimento finito per compattezza. Siano allora  $y_1, \dots, y_n$  i punti corrispondenti agli intorni estratti e sia  $g_x = \max_{i=1, \dots, n} g_{x,y_i}$ , che appartiene a  $\mathcal{R}$  per proprietà del reticolo. Si osservi che, per costruzione,  $g_x \geq f$  su tutto  $X$  e che in un intorno di  $x$  dato dall'intersezione degli  $U_{x,y_i}$  vale  $g_x \leq f + \varepsilon$ . Si conclude allora come nel teorema 5.2.2.  $\square$

**Definizione 17** (Algebra separante). *Un'algebra  $\mathcal{A} \subseteq Y^X$  si dice separante se per ogni  $x \neq y \in X$  esiste una funzione  $f \in \mathcal{A}$  tale che  $f(x) \neq f(y)$ .*

Osserviamo che se  $\mathcal{A} \subseteq C(X)$  è separante allora  $X$  è  $T_2$ , dato che se esiste una funzione continua che assume valori diversi su  $x$  e su  $y$  esistono due intorni aperti disgiunti di  $x$  e di  $y$ , ad esempio  $\left\{f(z) > \frac{f(x)+f(y)}{2}\right\}$  e  $\left\{f(z) < \frac{f(x)+f(y)}{2}\right\}$ .

**Definizione 18** (Algebra unitaria). *Un'algebra  $\mathcal{A}$  è unitaria se contiene la costante 1. In particolare,  $\mathcal{A}$  contiene tutte le funzioni costanti, grazie alla chiusura per il prodotto per scalare.*

**Teorema 5.3.3** (Stone, caso reale). *Sia  $X$  compatto e sia  $\mathcal{A} \subseteq C(X)$  un'algebra chiusa uniformemente e separante. Allora  $\mathcal{A} = C(X)$  oppure esiste un  $x_0 \in X$  per il quale  $\mathcal{A} = M_{x_0} = \{f \in C(X) \text{ t.c. } f(x_0) = 0\}$ . In altri termini, un'algebra unitaria che separa i punti è densa nelle funzioni continue.*

*Dimostrazione.* Dobbiamo distinguere due casi.

- Supponiamo che per ogni  $x \in X$  esista una funzione  $f_x \in \mathcal{A}$  che non si annulla su  $x$ . Per mostrare che  $\mathcal{A} = C(X)$  basta far vedere che è densa, e per fare ciò possiamo ricorrere al lemma 5.3.2. Sappiamo già che un'algebra chiusa di funzioni limitate è un reticolo (la limitatezza è garantita dalla compattezza di  $X$ ), quindi dobbiamo mostrare che  $\mathcal{A}$  risolve ogni problema di interpolazione a due punti. Ora, fissiamo due punti  $x, y$  e sia  $g_{x,y} \in \mathcal{A}$  una funzione che assume valori distinti su  $x$  e su  $y$ . Possiamo assumere che entrambi i valori assunti siano non nulli, in quanto se per caso  $g_{x,y}(x) = 0$  possiamo sostituire  $g_{x,y}$  con la funzione  $g_{x,y} + \delta f_x$ , la quale continua a separare  $x$  e  $y$  per  $\delta$  sufficientemente piccolo. Ora, l'insieme  $\{(u(x), u(y)) \in \mathbb{R}^2 \text{ t.c. } u \in \mathcal{A}\}$  è un sottospazio vettoriale, perché l'algebra è chiusa per somma e prodotto per scalare. Vogliamo mostrare che coincide con  $\mathbb{R}^2$ , ossia che per ogni coppia di valori  $(a, b) \in \mathbb{R}^2$  esiste una  $u \in \mathcal{A}$  che realizza tale coppia su  $x, y$ . Per farlo basta mostrare che esistono due vettori linearmente indipendenti, e a tal fine possiamo considerare  $(g_{x,y}(x), g_{x,y}(y))$  e  $(g_{x,y}^2(x), g_{x,y}^2(y))$ , i quali sono linearmente indipendenti perché

$$\begin{vmatrix} g_{x,y}(x) & g_{x,y}(y) \\ g_{x,y}^2(x) & g_{x,y}^2(y) \end{vmatrix} = g_{x,y}(x)g_{x,y}(y)(g_{x,y}(y) - g_{x,y}(x)) \neq 0$$

- Supponiamo ora che esista un  $x_0 \in X$  tale che  $\mathcal{A} \subseteq M_{x_0}$ . Costruiamo il morfismo lineare

$$\begin{aligned} L : \mathbb{R} \times \mathcal{A} &\longrightarrow \{\lambda : X \rightarrow \mathbb{R} \text{ costanti}\} + \mathcal{A} \\ (\lambda, f) &\longrightarrow \lambda + f \end{aligned}$$

$L$  è un'applicazione lineare continua (in norma uniforme) e l'applicazione  $N(g) = (g(x_0), g - g(x_0))$  è una sua inversa continua. Dunque è un omeomorfismo lineare, e perciò è bilipschitziano per quanto sappiamo da 4.1.1. Ora, una funzione bilipschitziana manda successioni di Cauchy in successioni di Cauchy, e dunque preserva la completezza. Poiché  $\mathcal{A}$  è chiusa nel completo  $C(X)$  lo spazio  $\mathbb{R} \times \mathcal{A}$  è completo, e tramite  $L$  anche  $\{\lambda : X \rightarrow \mathbb{R} \text{ costanti}\} + \mathcal{A}$  lo è. In particolare è un'algebra chiusa contenente le costanti, e sapendo che è anche separante deve coincidere con  $C(X)$  per il punto precedente. Rimuovendo le costanti otteniamo l'uguaglianza  $\mathcal{A} = M_{x_0}$ , come voluto. □

**Teorema 5.3.4** (Stone, caso complesso). *Sia  $X$  uno spazio compatto e sia  $\mathcal{A} \subseteq C^0(X, \mathbb{C})$  un'algebra uniformemente chiusa, separante e chiusa per coniugio. Allora  $\mathcal{A} = C^0(X, \mathbb{C})$  oppure  $\mathcal{A} = M_{x_0}$  per un certo  $x_0 \in X$ .*

*Dimostrazione.* Sia  $\mathcal{B} = \left\{ \Re(f) = \frac{f + \bar{f}}{2}, f \in \mathcal{A} \right\}$ , che è una sottoalgebra reale di  $\mathcal{A}$  grazie alla chiusura per coniugio. Osserviamo che  $\mathcal{B}$  coincide anche con la sottoalgebra delle parti immaginarie, dato che  $\mathcal{A}$  è chiusa per moltiplicazione per lo scalare  $i$ . Allora possiamo scrivere  $\mathcal{A} = \mathcal{B} + i\mathcal{B}$  e studiare solo le proprietà di  $\mathcal{B}$ , che è uniformemente chiusa e separante (perché se una funzione di  $\mathcal{A}$  assume due valori diversi su  $x$  e su  $y$  devono essere distinte le parti reali o le parti immaginarie). Dunque per teorema di Stone reale  $\mathcal{B} = C(X)$ , caso in cui  $\mathcal{A} = C^0(X, \mathbb{C})$ , oppure  $\mathcal{B} = M_{x_0}^{\mathbb{R}}$ , e dunque anche  $\mathcal{A} = M_{x_0}^{\mathbb{C}}$ . □

Dal teorema di Stone segue banalmente il teorema di Weierstrass, dato che i polinomi su  $[0, 1]$  sono ovviamente separanti e contengono le costanti. Più in generale, se  $X \subseteq \mathbb{R}^n$  è un compatto, le funzioni polinomiali a valori reali sono un'algebra unitaria e separante (se due punti sono diversi hanno almeno una coordinata diversa, e le proiezioni sulle coordinate sono polinomi), e dunque sono dense nelle funzioni continue.

### 5.3.2 Classificazione delle algebre chiuse

Sia  $\mathcal{A} \subseteq C(X)$  un'algebra chiusa sul compatto  $X$ . Consideriamo la relazione di equivalenza

$$x \sim y \iff \forall f \in \mathcal{A} f(x) = f(y)$$

In pratica stiamo identificando i punti che  $\mathcal{A}$  non sa distinguere, in modo da renderla separante. Infatti, detta  $\pi$  la proiezione al quoziente, è ben definita l'algebra

$$\mathcal{A}_\sim = \{f_\sim \mid \exists f \in \mathcal{A} \text{ t.c. } f = f_\sim \circ \pi\} \subseteq C(X_\sim)$$

Per costruzione  $\mathcal{A}_\sim$  è separante. Inoltre è chiusa, perché è l'immagine del completo  $\mathcal{A}$  mediante l'isometria (nella norma uniforme)  $f \rightarrow f_\sim$  e un'isometria manda successioni di Cauchy in successioni di Cauchy. Allora  $\mathcal{A}_\sim$  coincide con  $C(X_\sim)$  o con un certo  $M_{x_\sim}$ .

Osserviamo inoltre che le fibre  $\{F_i\}_{i \in I}$  della proiezione al quoziente sono chiuse. Infatti  $X_\sim$  è  $T_2$  perché ammette un'algebra separante, perciò i suoi singoletti sono chiusi e dunque lo sono anche le fibre per continuità della proiezione.

Possiamo allora dire che

**Corollario 5.3.4.1.** *Data un'algebra chiusa  $\mathcal{A} \subseteq C(X)$  sul compatto  $X$ , esiste una partizione in chiusi  $\{F_i\}_{i \in I}$  tali che, ristretti a ogni  $F_i$ , gli elementi di  $\mathcal{A}$  sono costanti. Inoltre, o  $\mathcal{A}$  è unitaria o esiste un  $F_0$  su cui le restrizioni delle  $f$  sono identicamente nulle.*

### 5.3.3 Insight algebrico

Gli insiemi  $M_x$  descritti nel teorema di Stone sono tutti e soli gli ideali massimali non banali dell'anello  $C(X)$ . Infatti, da un lato  $M_x = \ker(va_x)$ , e dunque è un'ideale, ed è massimale perché il quoziente  $C(X)/M_x \cong \mathbb{R}$  è un campo. Dall'altro, dato un ideale massimale  $I$ , o è contenuto in un certo  $M_x$ , e dunque coincide con esso per massimalità, oppure per ogni  $x$  esiste una funzione  $f_x$  che non si annulla in un intorno aperto di  $x$ . Estraendo un sottoricoprimento finito otteniamo una famiglia finita di funzioni  $f_1, \dots, f_n$  tali che  $G(x) = \sum_{k=1}^n f_k^2(x) > 0$ . Dunque esiste un'inversa moltiplicativa di  $G(X)$ , e cioè la costante 1 appartiene all'ideale. Allora  $I = C(X)$ .

## 5.4 Teorema di Ascoli-Arzelà

Ricordiamo che una famiglia di funzioni è equicontinua se ammettono tutte lo stesso modulo di continuità  $\omega$ . Equivalentemente, si può dire che la famiglia  $\Lambda$  di funzioni su  $X$  è equicontinua se la mappa di valutazione  $EVA: \Lambda \times X \rightarrow X$  tale che  $EVA(f, x) = f(x)^2$  è uniformemente continua, dove su  $\Lambda \times X$  si considera la distanza prodotto tra la distanza infinito e la distanza su  $X$ .

*Dimostrazione.* Mostriamo solo l'esistenza del modulo di continuità per  $EVA$ , dato che l'altra implicazione è ovvia. Se  $\Lambda$  è equicontinua allora

$$d(EVA(f, a), EVA(g, b)) \leq d(f(a), f(b)) + d(f(b), g(b)) \leq \omega(d(a, b)) + d_\infty(f, g)$$

La RHS è una funzione della distanza  $d((f, a), (g, b))$  che è nulla e continua in zero. Dunque per definizione è un modulo di continuità per la funzione  $EVA$ .  $\square$

**Teorema 5.4.1** (Ascoli-Arzelà). *Sia  $X$  uno spazio metrico compatto ed  $E$  uno spazio di Banach. Sia  $\Lambda \subseteq (C^0(X, E), d_\infty)$ . Allora  $\Lambda$  è compatta se e solo se è chiusa, equicontinua e per ogni  $x \in X$  l'insieme delle valutazioni  $\Lambda(x) = \{f(x), f \in \Lambda\}$  è relativamente compatto, ossia la sua chiusura è compatta.<sup>3</sup>*

*Dimostrazione.* .

( $\implies$ ) Sia  $\Lambda$  compatto, e dunque ovviamente chiuso. La mappa  $EVA$  è uniformemente continua perché è continua sul compatto  $\Lambda \times X$ , dunque  $\Lambda$  è equicontinua. Infine,  $\Lambda(x)$  è l'immagine tramite la funzione  $EVA$  del compatto  $\Lambda \times \{x\}$ , e perciò è compatto.

( $\impliedby$ ) Ci serviranno un paio di risultati preliminari sulle successioni di funzioni equicontinue.

<sup>2</sup>Direi che questo corso è traumatico come il First Impact.

<sup>3</sup>In realtà, come si evincerà dalla dimostrazione, è possibile sostituire la terza richiesta con la seguente proprietà: per ogni  $x \in S$ , con  $S \subseteq X$  un sottoinsieme denso, l'insieme  $\Lambda(x)$  è relativamente compatto.

**Lemma 5.4.2.** *Sia  $f_n$  una successione equicontinua di funzioni da uno spazio metrico a uno spazio di Banach. Allora*

1. *L'insieme  $C$  dei punti  $x \in X$  su cui  $f_n(x)$  converge è chiuso.*
2. *Se  $f_n \rightarrow f$  puntualmente, la convergenza è uniforme sui compatti e  $f$  ha lo stesso modulo di continuità delle  $f_n$ .*

Mostriamo le due implicazioni:

1. Sia  $x \in \overline{C}$  e sia  $x_k \rightarrow x$ , con  $x_k \in C$  per ogni  $k$ . Basta provare che  $f_n(x)$  è di Cauchy per completezza di  $E$ . Vale che

$$\|f_n(x) - f_m(x)\| \leq \|f_n(x) - f_n(x_k)\| + \|f_n(x_k) - f_m(x_k)\| + \|f_m(x_k) - f_m(x)\|$$

$$\|f_n(x) - f_m(x)\| \leq 2\omega(d(x, x_k)) + \|f_n(x_k) - f_m(x_k)\|$$

Dunque fissato  $\varepsilon > 0$  Basta scegliere un  $k$  tale che  $2\omega(d(x, x_k)) < \varepsilon/2$  e poi, dato che  $f_n(x_k)$  è di Cauchy, scegliere  $m, n$  sufficientemente grandi da rendere  $\|f_n(x_k) - f_m(x_k)\| < \varepsilon/2$ . Dunque la successione  $f_n(x)$  è di Cauchy.

2. Per quanto riguarda il modulo di continuità basta osservare che la relazione puntuale che definisce la continuità uniforme, ossia  $d(f_n(x), f_n(y)) \leq \omega(d(x, y))$  passa al limite in  $n$ . Fissiamo ora un compatto  $K \subseteq X$  e supponiamo WLOG che  $f = 0$ . Allora per compattezza per ogni  $n \in \mathbb{N}$  esiste un  $x_n \in K$  tale che  $\|f_n\|_{\infty, K} = f_n(x_n)$ . Sia  $x_{n(k)}$  una sottosuccessione convergente a un certo  $x \in K$ . Allora  $\|f_{n(k)}\|_{\infty, K} = f_{n(k)}(x_{n(k)}) \rightarrow 0$  perché le  $f_n$  convergono a zero puntualmente.<sup>4</sup> Ora, lo stesso ragionamento si può applicare a ogni sottosuccessione delle  $f_n$ , ossia da qualsiasi sottosuccessione si può estrarre una sotto-sottosuccessione convergente uniformemente alla funzione identicamente nulla. Per proprietà di Uryshon<sup>5</sup> questo significa che  $f_n \rightarrow 0$  uniformemente.

Torniamo alle nostre funzioni. Poiché uno spazio metrico compatto è separabile, consideriamo un sottoinsieme denso numerabile  $Y = \{y_k\}_{k \in \mathbb{N}} \subseteq X$ . Data una successione  $\{f_n\}_{n \in \mathbb{N}} \subseteq \Lambda$  vogliamo esibire una sottosuccessione convergente in  $\Lambda$ . Per farlo sfrutteremo l'argomento diagonale di Cantor. Sia  $f_{n,1}$  una sottosuccessione convergente in  $y_1$ , che esiste per compattezza di  $\overline{\Lambda}(y_1)$ . Da questa estraiamo una sotto-sottosuccessione  $f_{n,2}$  che converge in  $y_2$ , e reiteriamo induttivamente il ragionamento dimodoché la successione  $f_{n,k}$  converga sui primi  $k$  elementi di  $Y$ . Consideriamo infine la successione  $f_{n,n}$ , che per costruzione converge puntualmente su  $Y$ . Poiché per il lemma l'insieme di convergenza è chiuso,  $f_{n,n}$  deve convergere puntualmente su tutto  $X = \overline{Y}$ . Inoltre sempre per il lemma la convergenza è uniforme. Dunque poiché  $\Lambda$  è chiuso per successioni la funzione limite  $f$  appartiene alla famiglia  $\Lambda$ , che è dunque sequenzialmente compatta.  $\square$

## 5.5 Come non arrivare in ritardo

**Teorema 5.5.1** (Hopf-Rinow per spazi metrici). *Sia  $(X, d)$  compatto e siano  $x_0, x_1 \in X$ . Se esiste un arco di lunghezza finita che collega  $x_0$  a  $x_1$ , allora ne esiste uno di lunghezza minima.*

*Dimostrazione.* Il funzionale lunghezza  $L : C^0([0, 1], (X, d)) \rightarrow \mathbb{R} \cup \{+\infty\}$  è semicontinuo inferiormente, perché per definizione è l'involuppo superiore sulle partizioni finite  $\pi \subseteq [0, 1]$  dei funzionali continui  $L_\pi(\gamma) = \sum_{t_i \in \pi} d(\gamma(t_{i-1}), \gamma(t_i))$ . Dunque ammette minimo su un compatto di  $C^0([0, 1], (X, d))$ ,<sup>6</sup> che andremo ora a delineare.

Osserviamo che è possibile restringere la ricerca del minimo alle curve di lunghezza finita (ovviamente) di una certa costante di Lipschitz. Infatti, sappiamo dal corso di Analisi 1<sup>7</sup> che la lunghezza di una curva è invariante per riparametrazioni del dominio, cioè se componiamo una curva con un cambio di parametro  $\sigma : [0, 1] \rightarrow [0, c]$  strettamente monotono e continuo otteniamo una curva della stessa

<sup>4</sup>Abbiamo usato che, se  $f_n$  è una famiglia di funzioni equicontinue tra spazi metrici che converge puntualmente a  $f$  e se  $x_n \rightarrow x$ , allora  $f_n(x_n) \rightarrow f(x)$ . Questo fatto viene mostrato all'interno della dimostrazione del teorema 5.2.4.

<sup>5</sup>Vedasi [Mangioni, 2019].

<sup>6</sup>Vedasi 1.2.

<sup>7</sup>Sì, sul Mangionario old but gold c'è, vedasi [Mangioni, 2019].

lunghezza. Sappiamo inoltre che, se una curva è continua e ha lunghezza finita, la funzione  $t \rightarrow L(\gamma, [0, t])$  è continua e debolmente crescente, e dunque per ogni curva possiamo considerare la riparametrizzazione  $\sigma(t) = t + L(\gamma, [0, t])$ . Osserviamo che la curva così parametrizzata è definita su  $[0, 1 + L(\gamma)]$  ed è 1-lipschitziana. Infatti

$$d(\gamma(\sigma^{-1}(r_1)), \gamma(\sigma^{-1}(r_2))) \leq L(\gamma \circ \sigma^{-1}, [r_1, r_2])$$

Ora, se  $r_{1,2} = \sigma(t_{1,2})$  la lunghezza di questa curva è per costruzione  $L(\gamma, [t_1, t_2])$ . Dunque

$$d(\gamma(\sigma^{-1}(r_1)), \gamma(\sigma^{-1}(r_2))) \leq L(\gamma, [t_1, t_2]) = L(\gamma, [0, t_2]) - L(\gamma, [0, t_1])$$

$$d(\gamma(\sigma^{-1}(r_1)), \gamma(\sigma^{-1}(r_2))) \leq t_2 + L(\gamma, [0, t_2]) - t_1 - L(\gamma, [0, t_1]) = r_2 - r_1$$

Ora, dobbiamo ritornare alle curve definite su  $[0, 1]$  per poter applicare il funzionale lunghezza, quindi componiamo con la riparametrizzazione lineare  $r \rightarrow \frac{r}{1+L(\gamma)}$ , ottenendo in tal modo una curva di costante di Lipschitz  $1 + L(\gamma)$ . Possiamo allora restringerci a considerare le curve di costante di Lipschitz al più  $1 + L_0$ , dove  $L_0$  è la lunghezza della curva la cui esistenza è garantita dalle ipotesi. Tali funzioni sono una famiglia equicontinua, uniformemente chiusa e tali che, per ogni punto  $t \in [0, 1]$ , l'insieme delle valutazioni delle curve in  $t$  è relativamente compatto, in quanto le curve hanno valori nel compatto  $X$ . La tesi segue dunque per teorema di Ascoli-Arzelà.  $\square$

## 5.6 Totale limitatezza di spazi funzionali

**Definizione 19.** Sia  $(X; d)$  uno spazio metrico. Un sottoinsieme  $Y \subseteq X$  è una  $\varepsilon$ -rete, o è  $\varepsilon$ -denso, se ogni punto di  $X$  dista da  $Y$  al più  $\varepsilon$ . Equivalentemente,  $Y$  è una  $\varepsilon$ -rete se l'unione delle palle di raggio  $\varepsilon$  e centro in  $Y$  ricopre  $X$ .

Si può dire che la totale limitatezza equivale all'esistenza di una  $\varepsilon$ -rete finita per ogni  $\varepsilon > 0$ .

**Teorema 5.6.1.** Siano  $X, Y$  spazi metrici. Sia  $\{x_1, \dots, x_m\}$  una  $\delta$ -rete finita di  $X$  e sia  $\{y_1, \dots, y_n\}$  una  $\varepsilon$ -rete finita di  $Y$ . Sia  $C_\omega \subseteq (C^0(X, Y), d_\infty)$  il sottospazio delle funzioni con modulo di continuità  $\omega$ . Allora  $C_\omega$  ammette una  $2(\varepsilon + \omega(\delta))$ -rete finita di al più  $n^m$  punti.

**Corollario 5.6.1.1.** Se  $X, Y$  sono totalmente limitati lo è anche  $C_\omega$ .

*Dimostrazione.* Per ogni  $\alpha : [m] \rightarrow [n]$  definiamo

$$S_\alpha = \{f \in C^0(X, Y) \text{ t.c. } \forall j \leq m \, d(f(x_j), y_{\alpha(j)}) < \varepsilon\}$$

Osserviamo che ovviamente  $S_\alpha$  è aperto. Poiché inoltre  $\{y_i\}_{i \leq n}$  è una  $\varepsilon$ -rete gli insiemi  $S_\alpha$  ricoprono  $C^0(X, Y)$ , e dunque anche  $C_\omega$ . Consideriamo ora due funzioni  $f, g \in S_\alpha \cap C_\omega$  e cerchiamo di stimarne la distanza uniforme. Per farlo, osserviamo che per ogni  $x \in X$  esiste un  $x_j$  della  $\delta$ -rete a distanza al più  $\delta$  da  $x$ . Allora, poiché  $\|f - g\|_\infty = \sup_{x \in X} d(f(x), g(x))$ , abbiamo

$$\|f - g\|_\infty \leq \sup_{x \in X} d(f(x), f(x_j)) + d(f(x_j), y_{\alpha(j)}) + d(y_{\alpha(j)}, g(x_j)) + d(g(x_j), g(x)) \leq 2(\omega(\delta) + \varepsilon)$$

Questo conclude la dimostrazione, perché scegliendo una funzione  $f_\alpha \in S_\alpha \cap C_\omega$  per ognuno dei  $n^m$  valori di  $\alpha$  possibili otteniamo una  $2(\omega(\delta) + \varepsilon)$ -rete di  $C_\omega$ .  $\square$

## 5.7 Estensione di mappe UC con modulo concavo

**Teorema 5.7.1** (Tietze). Sia  $X \subseteq (E, \|\cdot\|)$ . Sia  $f : X \rightarrow \mathbb{R}$  una funzione con modulo di continuità subadditivo. Allora  $f$  si estende a tutto  $E$  con lo stesso modulo di continuità.

*Dimostrazione.* Dato  $y \in X$  poniamo  $\phi_y(x) = f(y) + \omega(\|x - y\|)$ , che per costruzione è maggiore di  $f(x)$ . Osserviamo che  $\phi_y$  si estende naturalmente a una funzione da  $E$  in  $\mathbb{R}$  con modulo di continuità  $\omega$ , in quanto per subadditività di  $\omega$  abbiamo

$$\phi_y(x) - \phi_y(z) = \omega(\|x - y\|) - \omega(\|z - y\|) \leq \omega(\|\|x - y\| - \|z - y\|\|) \leq \omega(\|x - z\|)$$

Poniamo ora  $\bar{f}(x) = \inf_{y \in X} \phi_y(x)$ . Osserviamo che, se  $x \in X$ , allora  $\bar{f}(x) \leq f(x)$  perché lo sono le  $\phi(x)$ , e d'altra parte  $\bar{f}(x) \geq \phi_x(x) = f(x)$ . Dunque  $\bar{f}$  estende  $f$  a tutto  $E$ .

Bisogna ora mostrare che anche  $\bar{f}$  ha modulo di continuità  $\omega$ . In generale è vero che l'involuppo inferiore di una famiglia equicontinua è equicontinuo. Infatti, dati due punti  $x, z \in E$ , sia  $\phi_n(x)$  una successione che converge a  $\inf_y \phi_y(x)$ . Consideriamo la corrispondente successione  $\phi_n(z)$  e osserviamo che, per uniforme continuità,  $\phi_n(z) \leq \phi_n(x) + \omega(\|x - z\|)$ . Dunque possiamo dire che  $\inf_y \phi_y(z) \leq \inf_y \phi_y(x) + \omega(\|x - z\|)$ . Inoltre per simmetria  $\inf_y \phi_y(x) \leq \inf_y \phi_y(z) + \omega(\|x - z\|)$ , da cui  $|\inf_y \phi_y(x) - \inf_y \phi_y(z)| \leq \omega(\|x - z\|)$ .  $\square$

Alla luce di quanto appena visto è rilevante trovare dei criteri per determinare se una funzione ammetta un modulo di continuità subadditivo. Osserviamo innanzitutto che è equivalente richiedere la concavità di  $\omega$ , come mostra il seguente lemma.

**Lemma 5.7.2.** *Sia  $\omega : [0, +\infty[ \rightarrow [0, +\infty[$  nulla e continua in zero. Allora valgono le seguenti implicazioni (dall'alto verso il basso):*

1.  $\omega$  è concava;
2.  $\omega$  è subadditiva, cioè  $\omega(a + b) \leq \omega(a) + \omega(b)$ ;
3.  $\omega$  è sublineare, ossia esistono  $a, b$  tali che  $\omega(t) \leq at + b$ ;
4.  $\omega$  è dominata da un modulo di continuità concavo.

Il lemma ci dice in particolare che una funzione  $f$  uniformemente continua ammette un modulo di continuità subadditivo se e solo se è possibile scegliere un altro modulo di continuità che sia anche concavo.

*Dimostrazione.* Vale che:

- 1)  $\implies$  2) Se  $\omega$  è concava i suoi rapporti incrementali sono decrescenti. Allora

$$\omega(a + b) - \omega(a) \leq \omega(b) + \omega(0) = \omega(b)$$

- 2)  $\implies$  3) Per subadditività abbiamo che  $\omega(nt) \leq n\omega(t)$  per  $n \in \mathbb{N}$ . Osserviamo ora che

$$\omega(t) = \omega\left(\left(\lfloor t \rfloor + 1\right) \frac{t}{\lfloor t \rfloor + 1}\right) \leq (\lfloor t \rfloor + 1)\omega\left(\frac{t}{\lfloor t \rfloor + 1}\right) \leq (t + 1) \sup_{s \in [0, 1]} \omega(s)$$

Da cui la tesi con  $a = b = \sup_{s \in [0, 1]} \omega(s)$ .

- 3)  $\implies$  4) Sia  $\bar{\omega}(t) = \inf\{ct + d \text{ t.c. } cs + d \geq \omega(s) \forall s\}$ . Per costruzione  $\bar{\omega}$  maggiore  $\omega$ , quindi affinché sia un modulo di continuità bisogna mostrare che è nullo e continuo in zero. Fissato  $\delta$  tale che, per ogni  $x < \delta$ ,  $\omega(x) < \omega(\delta)$ , che esiste per continuità, cerchiamo una retta  $r(s) = cs + d$  tale che  $r \geq \omega$  e  $r(0) = \omega(\delta)$ . Possiamo scegliere  $r(s) = \omega(\delta) + \frac{(a\delta + b - \omega(\delta))s}{\delta}$ , dove  $ax + b$  è la retta della sublinearità. Le due rette si incontrano in  $x = \delta$  per costruzione. Dopo  $\delta$   $r(x) > ax + b \geq \omega(x)$ , mentre prima  $r(x) \geq \omega(\delta) \geq \omega(x)$  per le ipotesi su  $\delta$ . Dunque  $r(x)$  è tra le rette di cui  $\bar{\omega}$  è l'involuppo inferiore, e dunque  $\bar{\omega}$  è nullo in zero. Per la continuità credo basti osservare che, per ogni  $\varepsilon > 0$  esiste una retta del tipo  $r(x)$  e un intorno dell'origine dove  $\bar{\omega} \leq r(x) \leq \varepsilon$ .

$\square$

Diamo ora due teoremi che permettono di caratterizzare le funzioni con modulo di continuità concavo.

**Teorema 5.7.3.** *Una funzione uniformemente continua  $f : X \subseteq (E, \|\cdot\|) \rightarrow (Y, d)$  con  $X$  convesso ammette modulo di continuità concavo.*

*Dimostrazione.* Consideriamo il modulo di continuità per  $f$  dato da  $\omega(t) = \sup_{\|x - y\| \leq t} d(f(x), f(y))$  e mostriamo che è concavo. Osserviamo che, dati  $x, y \in X$ ,

$$d(f(x), f(y)) \leq d(f(x), f(\lambda(x) + (1 - \lambda)y)) + d(f(y), f(\lambda(x) + (1 - \lambda)y))$$

$$d(f(x), f(y)) \leq \omega((1 - \lambda)\|x - y\|) + \omega(\lambda\|x - y\|)$$

Passando al sup sulle coppie di punti per cui  $\|x - y\| \leq t$  otteniamo

$$\omega(t) \leq \omega((1 - \lambda)t) + \omega(\lambda t)$$

Cioè  $\omega$  è subadditivo. Per il lemma 5.7.2  $f$  ammette un modulo concavo.  $\square$

Ricordando che ogni spazio metrico si immerge isometricamente in un Banach, come visto in 3.3.7, dal teorema precedente otteniamo il seguente risultato:

**Corollario 5.7.3.1.** *Una funzione  $f : (X, d) \rightarrow \mathbb{R}$  ammette un modulo di continuità subadditivo se e solo se si estende a una funzione uniformemente continua su uno spazio normato (che è ovviamente convesso) in cui si immerge isometricamente*

**Teorema 5.7.4.** *Una funzione uniformemente continua  $f : (X, d) \rightarrow \mathbb{R}$  ammette un modulo di continuità concavo se e solo se è limite uniforme di funzioni lipschitziane. In altri termini, la chiusura uniforme delle funzioni lipschitziane sono le funzioni con modulo di continuità concavo.*

*Dimostrazione.* ( $\Leftarrow$ ) Sia  $(g_n)_{n \in \mathbb{N}}$  una successione di funzioni lipschitziane, di costanti  $L_n$ , convergente uniformemente a  $f$ . Allora per ogni  $n \in \mathbb{N}$  si ha che

$$\|f(x) - f(y)\| \leq \|f(x) - g_n(x)\| + \|g_n(x) - g_n(y)\| + \|g_n(y) - f(y)\| \leq 2\|f - g_n\|_\infty + L_n(d(x, y))$$

A meno di considerare una sottosuccessione possiamo assumere che  $\|f - g_n\| \leq 1/n$ . Dunque possiamo porre  $\omega(t) = \inf_{n \in \mathbb{N}} \frac{2}{n} + L_n(t)$ , che è una funzione nulla e continua in zero. Inoltre  $\omega$  è concava perché inviluppo inferiore di funzioni concave  $h_n$ . Infatti

$$t \inf_n h_n(x) + (1 - t) \inf_m h_m(y) \leq \inf_n (th_n(x) - (1 - t)h_n(y)) \leq \inf_n h_n(tx + (1 - t)y)$$

( $\Rightarrow$ ) Sia  $f$  una funzione che ammette modulo di continuità concavo. Osserviamo innanzitutto che ammette come modulo anche

$$\omega(t) = \inf\{at + b \mid a, b > 0, \forall x, x' \in X \mid f(x) - f(x') \leq ad(x, x') + b\}$$

In pratica, questo modulo è l'inviluppo inferiore di tutte le rette che maggiorano almeno un modulo di continuità per  $f$ . Al fine di mostrare che  $\omega$  è un modulo di continuità l'unica cosa non banale da osservare è la continuità nell'origine, la quale si ottiene con un ragionamento molto simile a quello visto nell'implicazione 3)  $\Rightarrow$  4) del lemma 5.7.2.

Esibiamo ora una successione  $g_n \rightarrow f$  di funzioni lipschitziane. Innanzitutto consideriamo una successione di rette della forma  $a_n t + b_n$  tali che  $a, b > 0$ ,  $b_n \rightarrow 0$  e  $\forall x, x' \in X \mid f(x) - f(x') \leq a_n d(x, x') + b_n$ . In pratica, stiamo prendendo una successione di rette dall'insieme il cui inviluppo inferiore è  $\omega$ . Poniamo  $g_n(x) = \sup_{y \in X} \{f(y) - a_n d(x, y)\}$ . Ovviamente  $g_n(x) \geq f(x)$ , e inoltre

$$|g_n(x) - f(x)| = g_n(x) - f(x) = \sup_{y \in X} \{f(y) - f(x) - a_n d(x, y)\} \leq b_n \rightarrow 0$$

Mostriamo dunque che  $g_n$  è  $a_n$ -lipschitziana.

$$|g_n(x) - g_n(z)| = \left| \sup_{y \in X} \{f(y) - a_n d(x, y)\} - \sup_{y \in X} \{f(y) - a_n d(z, y)\} \right|$$

$$|g_n(x) - g_n(z)| \leq \left| \sup_{y \in X} \{a_n(d(z, y) - d(x, y))\} \right| \leq a_n d(x, z)$$

$\square$

## Capitolo 6

# Calcolo differenziale in spazi di Banach

Vedremo che non cambia davvero niente rispetto a  $\mathbb{R}$ , ma che si può esprimere tutto in modo indipendente dalle coordinate (cosa molto comoda in spazi bruttini dove le basi magari esistono ma chissà quali sono).

### 6.1 Definizione di “o piccolo”

Data una funzione  $f : E \rightarrow F$  tra spazi di Banach, si dice che  $f(x) = o(1)$  per  $x \rightarrow x_0$  se e solo se  $\|f(x)\| = o(1)$ . Vale una definizione analoga per il concetto di “O grande”. Generalizzando, si dice che  $f(x) = o(g(x))$  per  $x \rightarrow x_0$  se  $\|f(x)\| = o(\|g(x)\|)$ .

Osserviamo in particolare che  $L \in L(E, F)$  è tale che  $Lx = o(x)$  per  $x \rightarrow x_0$  se e solo se  $L = 0$ . In altri termini, una funzione lineare è strettamente sublineare se e solo se è identicamente nulla.<sup>1</sup> Infatti, se per ogni  $\varepsilon$  esiste un intorno di zero in cui  $\|Lx\| \leq \varepsilon\|x\|$ , per omogeneità la relazione vale per ogni  $x$ , da cui  $\|Lx\| = 0$  identicamente.

### 6.2 Differenziale di Fréchet

Dato un aperto  $U \subseteq E$  di uno spazio di Banach, un punto  $x_0 \in U$  e una funzione  $f : E \rightarrow F$  tra spazi di Banach, si dice che  $f$  è differenziabile in  $x_0$  se esiste un'applicazione lineare continua  $df_{x_0}$ , detta differenziale (secondo Fréchet) tale che

$$f(x) = f(x_0) + df_{x_0}(x - x_0) + o(x - x_0)$$

Osserviamo innanzitutto che, se il differenziale esiste, è unico. Infatti, siano  $L^1$  e  $L_2$  due differenziali. Allora  $(L^1 - L_2)(x - x_0) = o(x - x_0)$ , cosa che, per quanto osservato in 6.1 ci assicura che  $L^1 \equiv L_2$ .

Come ci si può aspettare, valgono un sacco di teoremi analoghi a quelli sulle derivate, che non dimostrerò perché hanno lo stesso livello di sfida della settimana enigmistica. Mi limito a sottolineare i due risultati seguenti:

- $d(f \circ g)_x = df_{g(x)} \circ dg_x$ , cioè i differenziali compongono. Questo è il prototipo di una serie di “chain rules” che invocheremo all’occorrenza.
- Se  $f : U \subseteq E \rightarrow V \subseteq F$  è un omeomorfismo tra aperti e il suo differenziale in  $x_0 \in U$  è un isomorfismo lineare, allora  $f^{-1}$  è differenziabile in  $f(x_0)$  con differenziale  $df_{x_0}^{-1}$ .

#### 6.2.1 Differenziali particolari

Se  $\gamma : \mathbb{R} \rightarrow F$  è una curva  $C^1$ , il suo differenziale appartiene a  $L(\mathbb{R}, F)$ , che è isomorfo isometricamente a  $F$  mediante la mappa di valutazione in 1. Allora  $d\gamma_{x_0}(1)$  è un qualche vettore, che per costruzione

---

<sup>1</sup>Ma dai, non mi dire.

coincide con la derivata di  $\gamma$ . Allora

$$d\gamma_{x_0}(1) = \gamma'(x_0)$$

Se invece  $f : U \subseteq E \rightarrow \mathbb{R}$ , dove  $E$  è uno spazio di Hilbert (cioè è munito di un prodotto scalare definito positivo che lo rende uno spazio di Banach rispetto alla norma indotta), allora  $df_{x_0} \in L(E, \mathbb{R})$ . Ora, per un teorema di rappresentazione di Riesz che ricalca quello visto a Geometria 1 per  $\mathbb{R}^n$ , ogni funzionale lineare continuo è il prodotto scalare per un certo vettore. Definiamo allora il gradiente di  $f$  in  $x_0$ , e lo indichiamo con  $\vec{\nabla} f(x_0)$ , l'unico vettore tale che

$$df_{x_0}(v) = \langle \vec{\nabla} f(x_0), v \rangle$$

In particolare, se  $E = \mathbb{R}^n$  con la base canonica  $e_i$ , il gradiente assume la forma

$$\vec{\nabla} f = \left( \langle \vec{\nabla} f, e_i \rangle \right)_{i \leq n} = (\partial_i f)_{i \leq n}$$

in quanto la restrizione della funzione  $f$  alla retta  $x_0 + \text{Span}(e_i)$  è una curva la cui derivata è la derivata parziale di  $f$ .

Similmente, se  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $f(x_0)$  si decompone in coordinate come  $(f_i(x_0))_{i \leq n}$ . Allora il suo differenziale tra gli spazi muniti delle basi canoniche è la seguente matrice, che prende il nome di matrice Jacobiana:

$$Jf_{x_0} = (e_i \cdot df_{x_0}(e_j))_{i,j \leq n}$$

Ora, il prodotto scalare per  $e_i$  è la proiezione sulla coordinata  $i$ -esima, che ha come differenziale se stessa perché è lineare. Allora per composizione dei differenziali

$$Jf_{x_0} = (d(e_i \cdot f)_{x_0}(e_j))_{i,j \leq n} = (\partial_j(f_i(x_0)))_{i,j \leq n}$$

In altri termini, la matrice Jacobiana è la matrice delle derivate parziali. Ogni riga contiene nell'ordine le derivate parziali della componente  $i$ -esima.

### 6.2.2 Derivata direzionale

Sia  $f$  una funzione da  $E$  in  $F$ . Dato un vettore  $v \in E$  e un punto  $x_0 \in E$ , consideriamo la curva affine  $\gamma(t) = x_0 + vt$ , a valori in un intorno di  $x_0$ . Si definisce derivata direzionale lungo  $v$  di  $f$  il seguente limite, se esiste:

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

Nel caso  $v = e_i$  ritroviamo la derivata parziale. Inoltre se  $f$  è differenziabile ammette tutte le derivate direzionali, in quanto la derivata parziale risulta la derivata della curva  $f \circ \gamma$ :

$$\partial_v f(x_0) = df_{\gamma(0)}(\gamma'(0)) = df_{x_0}(v)$$

### 6.2.3 Differenziabilità in spazi complessi

**Definizione 20.** Una funzione  $f : \Omega \subseteq E \rightarrow F$  tra  $\mathbb{C}$ -spazi vettoriali muniti di una norma complessa si dice olomorfa se è differenziabile e il differenziale in ogni punto è  $\mathbb{C}$ -lineare.

Per la caratterizzazione vista in 2.5 è equivalente dire che il differenziale  $\mathbb{R}$ -lineare commuta con l'applicazione  $J$  che rappresenta il prodotto per  $i$ . In questo modo si riescono a tradurre praticamente tutti i teoremi visti finora nel caso reale verificando solo che qualche matrice commuti con  $J$ .

Per il seguente teorema ci serve una definizione:

**Definizione 21.** Una funzione  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  si dice armonica se il suo laplaciano è identicamente nullo, ossia se

$$\nabla^2 f = \sum_{i=1}^n \partial_i^2 f = 0$$

**Teorema 6.2.1.** *Se una funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  di classe  $C^2$  è olomorfa<sup>2</sup> allora le sue componenti  $f_x, f_y$  sono funzioni armoniche.*

*Dimostrazione.* Supponiamo che  $f$  sia olomorfa. Per la caratterizzazione equivalente possiamo considerare  $\mathbb{C}$  come  $\mathbb{R}^2$  e imporre che la matrice Jacobiana commuti con  $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ , e cioè che  $f = (f_x, f_y)$  soddisfi le cosiddette equazioni di Cauchy-Riemann:

$$\begin{cases} \partial_x f_x = \partial_y f_y \\ \partial_x f_y = -\partial_y f_x \end{cases}$$

Derivando le due equazioni rispettivamente in  $x$  e in  $y$  otteniamo

$$\begin{cases} \partial_x^2 f_x = \partial_x \partial_y f_y \\ \partial_y \partial_x f_y = -\partial_y^2 f_x \end{cases}$$

Sommando membro a membro otteniamo

$$\nabla^2 f_x = \partial_x^2 f_x + \partial_y^2 f_x = \partial_x \partial_y f_y - \partial_y \partial_x f_y = 0$$

L'ultimo passaggio segue dal teorema di Schwarz applicato a  $f_y$ , che è  $C^2$  in quanto lo è  $f$ . Procedendo in maniera analoga si ottiene che anche  $f_y$  è armonica.  $\square$

Si può mostrare anche una sorta di viceversa:

**Teorema 6.2.2.** *Sia  $u \in C^2(D, \mathbb{R})$  armonica su un dominio  $D$  semplicemente connesso. Allora  $u$  è la parte reale (o immaginaria) di una funzione olomorfa.*

*Dimostrazione.* Stiamo cercando una funzione  $v \in C^2(D, \mathbb{C})$  che soddisfi con  $u$  le condizioni di Cauchy-Riemann, così da poter scrivere  $f = u + iv$ . In altri termini vogliamo che la forma  $\omega = -\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy$  sia esatta, così da poter scegliere come  $v$  la primitiva di  $\omega$ . Ma essendo  $D$  semplicemente connesso basta che  $\omega$  sia chiusa, ossia che  $\frac{\partial \omega}{\partial y} - \frac{\partial \omega}{\partial x} = 0$ . Questo è vero perché tale differenza è esattamente il laplaciano di  $u$  che è nullo per ipotesi.  $\square$

## 6.3 Differenziale secondo

Dato che  $df$  è a sua volta una funzione tra spazi di Banach, è possibile considerarne il differenziale  $d(df) : \Omega \rightarrow L(E, L(E, F))$ . Sarà comodo identificare lo spazio di arrivo con le applicazioni bilineari da  $E$  in  $F$ , che indicheremo con  $L^2(E \times E, F)$ , sia perché il concetto di differenziale  $n$ -esimo si generalizza molto più facilmente in questa forma, sia perché risparmio tempo se scrivo meno parentesi. Poniamo allora

$$(d(df)_x(u))(v) = d^2 f_x(u, v)$$

Dove il primo elemento determina la funzione lineare che viene calcolata nel secondo.

Osserviamo che vi è una stretta relazione tra il differenziale e le “differenze seconde” della funzione, ossia le scritture del tipo “ $f(x + u + v) - f(x + u) - f(x + v) + f(x)$ ”. Mostriamo, in particolare, che vale la seguente relazione:

$$\|f(x + u + v) - f(x + u) - f(x + v) + f(x) - d^2 f_x(u, v)\| = o(\|u\| + \|v\|)^2)$$

*Dimostrazione.* Facciamo un po' di conti.<sup>3</sup> Con il solito trucco della curva calcolata agli estremi, applichiamo il teorema del valor medio alla funzione  $f(x + u + tv) - f(x + tv) - td^2 f_x(u, v)$  e otteniamo

$$\begin{aligned} \|f(x + u + v) - f(x + u) - f(x + v) + f(x) - d^2 f_x(u, v)\| &= \|f(x + u + tv) - f(x + tv) - td^2 f_x(u, v)\|_{t=0}^{t=1} \leq \\ &\leq \sup_{0 \leq t \leq 1} \|df_{x+u+tv}(v) - df_{x+tv}(v) - d^2 f_x(u, v)\| \leq \sup_{0 \leq t \leq 1} \|df_{x+u+tv} - df_{x+tv} - d^2 f_x(u)\| \|v\| \end{aligned}$$

<sup>2</sup>In futuro sapremo che le funzioni olomorfe sono analitiche, perciò l'ipotesi che sia  $C^2$  sarà superflua.

<sup>3</sup>Che novità.

Dove l'ultima norma è quella operatoriale, che funge da costante di Lipschitz. Ora, per come sono definiti i differenziali, possiamo scrivere l'espressione precedente come

$$\begin{aligned} & \sup_{0 \leq t \leq 1} \|df_x + d^2 f_x(u + tv) + o(\|u + tv\|) - df_x - d^2 f_x(tv) + o(\|tv\|) - d^2 f_x(u)\| \|v\| = \\ & = \sup_{0 \leq t \leq 1} \|o(\|u + tv\|) + o(\|tv\|)\| \|v\| \end{aligned}$$

Ora, poiché  $t \in [0, 1]$ , la quantità tra parentesi è  $o(\|u\| + \|v\|)$  uniformemente in  $t$ . Dunque il prodotto è  $o((\|u\| + \|v\|)^2)$ .  $\square$

Un'altra cosa che si può mostrare è che

**Lemma 6.3.1.**

$$d^2 f_x(u, v) = \partial_t(\partial_s f(x + tu + sv)|_{s=0})|_{t=0}$$

*Dimostrazione.* Per definizione di derivata parziale,

$$\partial_s f(x + tu + sv)|_{s=0} = df_{x+tu}(v)$$

Allora

$$\partial_t(\partial_s f(x + tu + sv)|_{s=0})|_{t=0} = \partial_t(df_{x+tu}(v))|_{t=0}$$

Adesso, la derivata del termine tra parentesi, che si può scrivere come  $Val_v(df_{x+tu})$  dove  $Val_v$  è l'operatore di valutazione, è data dalla composizione dei differenziali, che coincide con il differenziale secondo di  $f$  valutato in  $v$  dato che la valutazione è lineare. Dunque, come voluto, si ha

$$\partial_t(\partial_s f(x + tu + sv)|_{s=0})|_{t=0} = d^2 f_x(u, v)$$

$\square$

### 6.3.1 Matrice Hessiana

Consideriamo una funzione  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  differenziabile due volte. Il differenziale secondo è una forma bilineare a valori in  $\mathbb{R}$ , dunque si può rappresentare con una matrice, detta Hessiana di  $f$ . Tale matrice ha elementi

$$Hf_{ij}(x) = d^2 f_x(e_i, e_j) = \partial_i \partial_j f(x)$$

Volendo, si può osservare che l'Hessiana è la matrice Jacobiana del gradiente di  $f$ , rappresentato dal vettore delle derivate parziali.

## 6.4 Polinomio di Taylor

Non è l'eroe che ci meritiamo, ma è l'eroe di cui abbiamo bisogno.

---

*Gordon, capo del GCPD*

**Definizione 22.** Data una funzione  $f : E \rightarrow F$  differenziabile  $n$  volte (con  $n \geq 1$ ), definiamo il suo polinomio di Taylor di ordine  $n$  e centrato in  $x_0$  come

$$T_n(x, x_0) = \sum_{k=0}^n \frac{d^k f_{x_0}(x - x_0)^k}{k!}$$

Dove  $d^k f_{x_0}(x - x_0)^k = d^k f_{x_0}(x - x_0, x - x_0, \dots, x - x_0)$ .

### 6.4.1 Derivata rispetto al centro dello sviluppo

In linea con quanto visto ad Analisi 1, possiamo vedere cosa succede differenziando il polinomio di Taylor rispetto al centro dello sviluppo.

**Teorema 6.4.1.** *Se  $g$  è differenziabile  $n + 1$  volte nel punto  $y$ , vale che*

$$d_y T_n(x, y) = \frac{d^{n+1} g_y((x-y)^n, \cdot)}{n!}$$

*Dimostrazione.* Iniziamo a osservare cosa succede a un termine dello sviluppo quando viene differenziato rispetto a  $y$ . Innanzitutto scriviamo  $d^k g_y(x-y)^k = Val_k(d^k g_y, (x-y)^k)$ , dove indichiamo con  $Val_k(f, (x_1, \dots, x_k))$  l'applicazione che calcola  $f$  nel punto  $(x_1, \dots, x_k)$ . Se  $f$  è  $k$ -lineare, come nel nostro caso,  $Val_k$  è  $(k+1)$ -lineare, ossia lineare sia nelle entrate che nella funzione. Ciò significa in particolare che, se si fissano  $k$  entrate, la funzione  $Val_k$  diventa lineare, e il suo differenziale parziale in quella entrata è lei stessa calcolata nel differenziale dell'entrata. Scriviamo allora il suo differenziale rispetto a  $y$  come somma dei  $k+1$  differenziali parziali:

$$\begin{aligned} d_y Val_k(f, (x_1, \dots, x_k)) &= (d_{1,y} + \dots + d_{k+1,y}) Val_k(f, (x_1, \dots, x_k)) = \\ &= Val_k(d_y f(\cdot), (x_1, \dots, x_k)) + Val_k(f, (d_y(x_1)(\cdot), \dots, x_k)) + \dots + Val_k(f, (x_1, \dots, d_y(x_k)(\cdot))) \end{aligned}$$

Ha senso che in ogni termine vi sia un'entrata libera, indicata con  $(\cdot)$ , in quanto un differenziale mangia vettori. Ora, applichiamo quanto detto finora al nostro caso, nel quale le cose si semplificano parecchio perché le  $k$  entrate in fondo sono tutte uguali a  $x-y$ , il cui differenziale è palesemente  $d_y(x-y)(\cdot) = -(\cdot)$ . Allora per linearità otteniamo

$$d_y(d^k g_y(x-y)^k)(\cdot) = d_y Val_k(d^k g_y, (x-y)^k) = Val_k(d^{k+1} g_y, (x-y)^k) - k Val_k(d^k g_y, (x-y)^{k-1}, \cdot)$$

Possiamo osservare che il secondo pezzo è una valutazione in  $k$  entrate di cui una libera, perciò di fatto è una valutazione in  $k-1$  entrate che mangia un vettore  $(\cdot)$ . Allora otteniamo

$$d_y(d^k g_y(x-y)^k)(\cdot) = d^{k+1} g_y((x-y)^k, \cdot) - k d^k g_y((x-y)^{k-1}, \cdot)$$

La tesi segue moltiplicando per  $1/k!$  e sommando su  $k = 1, \dots, n$ , cosa che genera una serie telescopica in cui sopravvive solo l'ultimo termine.  $\square$

### 6.4.2 Resto di Peano

**Teorema 6.4.2.**  $\|f(x) - T_n(x, x_0)\| = o(\|x - x_0\|^n)$ .

*Dimostrazione.* Innanzitutto scriviamo  $x = x_0 + u$  per un certo vettore  $u$ . Allora possiamo restringere  $f$  lungo la retta  $x_0 + tu$  e ridurci al caso in cui  $E$  è un intervallo di  $\mathbb{R}$ . In altri termini, consideriamo la curva  $\gamma(t) = f(x_0 + tu)$ . La sua derivata è  $\gamma'(t) = df_{x_0+tu}(u)$ , e induttivamente possiamo calcolare la derivata  $k$ -esima  $\gamma^{(k)}(t) = d^k f_{x_0+tu}(u)^k$ . Infatti

$$\gamma^{(k)}(t) = f(x_0 + tu)^{(k)} = (f(x_0 + tu)^{(k-1)})' = (d^{k-1} f_{x_0+tu}(u)^{(k-1)})' = d^k f_{x_0+tu}(u)^k$$

Ora, consideriamo la funzione  $g(t) = \gamma(t) - T_n(t, 0)$ . Essa ha tutte le derivate fino alla  $n$ -esima nulle per costruzione, dunque ci basta mostrare che  $\|g(t)\| = o(t)^n$ . Procediamo per induzione su  $n$ .

- Se  $n = 1$ ,  $\|g(t)\| = \|g(t) - g(0) - tg'(0)\| = o(t)$  per definizione di derivabilità.
- Supponiamo la tesi vera per  $n = k$  e dimostriamola per  $n = k + 1$ . La derivata  $g'(t)$  è derivabile  $k$  volte con tutte le derivate nulle, quindi per induzione  $\|g'(t)\| = o(t)^k$ . Allora per teorema del valore medio

$$\|g(t)\| = \|g(t) - g(0)\| \leq \sup_{s \in [0, t]} \|g'(s)\| t = o(t^{k+1})$$

Osserviamo che la dimostrazione fatta finora si estende bene al caso in cui  $E$  è generico, perché l'o piccolo determinato dipende solo da  $t$  (e dunque dal modulo di  $u$ ) e cioè è uniforme sulla palla di centro  $x_0$  e raggio  $\|u\|$ .  $\square$

### 6.4.3 Resto di Lagrange

**Teorema 6.4.3.** *Sia  $f$  una funzione tra Banach differenziabile in  $x_0$   $n + 1$  volte. Allora*

$$\|f(x_0 + u) - T_n(x_0 + u, x_0)\| \leq \sup_{t \in (0,1)} \|d^{n+1} f_{x_0+tu}\| \frac{\|u\|^{n+1}}{(n+1)!}$$

*Dimostrazione.* Come prima, ci restringiamo al caso in cui  $f$  è una curva  $\gamma(t) = f(x + tu)$ , di cui ci interesserà il resto per  $t = 1$ . Inoltre, esiste un trucco (che usa un cannone che non siamo in grado di dimostrare) per ridursi a funzioni a valori in  $\mathbb{R}$ . Infatti, sia  $F$  lo spazio di arrivo e sia  $y = \gamma(1) - T_n(1, 0, \gamma)$ . Se  $y = 0$  ovviamente risulta

$$0 = \|\gamma(1) - T_n(1, 0, \gamma)\| \leq \frac{\sup_{t \in (0,1)} \|\gamma^{(n+1)}(t)\|}{(n+1)!}$$

In caso contrario consideriamo un funzionale lineare  $\phi$  lungo la retta  $\text{Span}\{y\}$  definito nel seguente modo:

$$\phi(ty) = t\|y\|$$

ed estendiamo a un funzionale lineare continuo su  $F$ . Per fare ciò ci serve il seguente:<sup>4</sup>

**Teorema 6.4.4** (di Hahn-Banach). *Un funzionale lineare continuo definito su un sottospazio si estende in modo lineare e continuo a un funzionale su tutto lo spazio, il quale ne conserva la norma.*

Ad esempio, per  $F = \mathbb{R}^n$  possiamo scegliere una funzione che sulla retta  $\text{Span}\{y\}$  coincide con  $\phi$  e che sull'ortogonale è nulla, cioè la distanza dal piano ortogonale presa con segno.

Ora, data questa funzione (che continueremo a chiamare  $\phi$ ), consideriamo la funzione reale di variabile reale  $u(t) = \phi(\gamma(t))$ . Osserviamo che  $u^{(k)} = \phi(\gamma^{(k)})$ , in quanto ogni volta che si deriva  $d\phi_x = \phi$ . Allora  $u$  è derivabile  $n + 1$  volte e dunque esiste il suo polinomio di Taylor, che non è altro che l'immagine secondo  $\phi$  del polinomio di  $\gamma$ . Ma  $u$  è reale di variabile reale, perciò per la formula del resto di Lagrange, che sappiamo valere in  $\mathbb{R}$ , si ha

$$\begin{aligned} u(1) - T_n(1, 0, u) &\leq \frac{\sup_{t \in (0,1)} \|u^{(n+1)}(t)\|}{(n+1)!} \\ \phi(\gamma(1) - T_n(1, 0, \gamma)) &\leq \frac{\sup_{t \in (0,1)} \|\phi(\gamma^{(n+1)}(t))\|}{(n+1)!} \end{aligned}$$

Ora, nella disuguaglianza soprastante a sinistra abbiamo  $\phi(\gamma(1) - T_n(1, 0, \gamma)) = \|\gamma(1) - T_n(1, 0, \gamma)\|$ , per come avevamo definito  $\phi$ . A destra possiamo maggiorare la norma con  $\|\gamma^{(n+1)}(t)\|$ , in quanto  $\phi$  ha norma operatoriale 1 come la sua restrizione alla retta lungo  $y$ . Otteniamo allora

$$\|\gamma(1) - T_n(1, 0, \gamma)\| \leq \frac{\sup_{t \in (0,1)} \|\gamma^{(n+1)}(t)\|}{(n+1)!}$$

□

**Corollario 6.4.4.1.** *Se  $f$  è di classe  $C^n$ , il resto si stima come*

$$\|f(x_0 + u) - T_n(x_0 + u, x_0, f)\| \leq \sup_{t \in (0,1)} \|d^n f_{x_0+tu} - d^n f_{x_0}\| \frac{\|u\|^n}{n!}$$

*Dimostrazione.* Innanzitutto spezziamo il polinomio  $n$ -esimo nella somma di un termine e del polinomio  $n - 1$ -esimo.

$$\|f(x_0 + u) - T_n(x_0 + u, x_0, f)\| = \left\| f(x_0 + u) - T_{n-1}(x_0 + u, x_0, f) - \frac{d^n f_{x_0}(x - x_0)^n}{n!} \right\|$$

<sup>4</sup>Il problema non è tanto conservare la linearità, che si ottiene definendo la funzione su una base (la quale esiste per una conseguenza del Lemma di Zorn), quanto la continuità, che come sappiamo non è automatica. Per i curiosi rimando a [https://it.wikipedia.org/wiki/Teorema\\_di\\_Hahn-Banach](https://it.wikipedia.org/wiki/Teorema_di_Hahn-Banach).

Osserviamo che, essendo  $\frac{d^n f_{x_0}(x-x_0)^n}{n!}$   $n$ -lineare, il suo sviluppo di Taylor all'ordine  $n-1$  è identicamente nullo, in quanto nel calcolo vengono fuori solo applicazioni lineari da calcolare in  $x_0 - x_0 = 0$ . Allora, posto  $g(x_0 + u) = f(x_0 + u) - \frac{d^n f_{x_0}(x-x_0)^n}{n!}$  possiamo scrivere

$$\|f(x_0 + u) - T_n(x_0 + u, x_0, f)\| = \left\| f(x_0 + u) - \frac{d^n f_{x_0}(x-x_0)^n}{n!} - T_{n-1}(x_0 + u, x_0, g) \right\|$$

Allora, applicando la formula di Lagrange all'ordine  $n-1$ , otteniamo

$$\|f(x_0 + u) - T_n(x_0 + u, x_0, f)\| \leq \frac{\|u\|^n}{n!} \sup_{t \in (0,1)} \left\| d^n f_{x_0+tu} - \frac{d_x^n (d^n f_{x_0}(x-x_0)^n)_{x=x_0+tu}}{n!} \right\|$$

Con un po' di ragionamenti molto simili a quelli fatti in 6.4.1, il differenziale rispetto a  $x$  di  $d^n f_{x_0}(x-x_0)^n$  è  $n! d^n f_{x_0}$ , che non ha entrate  $x$  e quindi è costante al variare di  $t$ . Otteniamo allora la tesi:

$$\|f(x_0 + u) - T_n(x_0 + u, x_0, f)\| \leq \sup_{t \in (0,1)} \|d^n f_{x_0+tu} - d^n f_{x_0}\| \frac{\|u\|^n}{n!}$$

□

Abbiamo ottenuto una forma carina per il resto di una funzione  $C^n$ . Vi ricorda qualcosa? Purtroppo, quanto visto finora è, in nuce, la generalizzazione di una nostra vecchia conoscenza, che enuncio semplicemente perché evoca Vietnam flashbacks.

**Teorema 6.4.5** (Classificazione delle funzioni  $C^n$ ). *Una funzione tra spazi di Banach è di classe  $C^n$  se e solo se ammette uno sviluppo polinomiale di ordine  $n$  con coefficienti continui (rispetto a tutte le variabili) e resto stimato dall'alto da un'espressione del tipo  $\sigma(x, y)\|x-y\|^n$ , dove  $\sigma$  è continua e infinitesima sulla diagonale.*

No, not that again!

---

*Living Colours, "Love Rears Up  
Its Ugly Head"*

## 6.5 Massimi e minimi liberi

**Definizione 23.** *Una funzione  $b : E^2 \rightarrow \mathbb{R}$  bilineare simmetrica si dice positiva se  $b(x, x) \geq 0$  per ogni  $x \in E$ . Nelle stesse ipotesi,  $b$  si dice definita positiva se esiste un  $a > 0$  tale che  $b(x, x) \geq a\|x\|^2$  per ogni  $x \in E$ .*

Osserviamo che la seconda condizione è molto più potente di quella vista a Geometria 1 (ossia che  $b(x, x) > 0$ ). Possiamo notare che in  $E = \mathbb{R}^n$  la nuova definizione è equivalente a quella nota, perché basta prendere come  $a$  il minimo di  $b(x, x)$  sulla palla unitaria (che esiste per compattezza).

Questa definizione ci servirà per legare l'esistenza dei differenziali primo e secondo ai punti di massimo e minimo. Ci limiteremo al caso in cui i punti estremanti sono "liberi", ossia in cui non vi sono vincoli sul dominio.

**Teorema 6.5.1** (CNES per punti estremanti liberi). *Sia  $f : E \rightarrow \mathbb{R}$  una funzione differenziabile due volte nel punto  $x_0$ . Se questo è un punto di minimo locale (cioè in un suo intorno) allora  $\vec{\nabla} f_{x_0} = 0$  e  $d^2 f_{x_0}$  è positivo. Il viceversa vale nel caso in cui  $df_{x_0} = 0$  e  $d^2 f_{x_0}$  è definito positivo. Il teorema vale anche scambiando "minimo" con "massimo" e "positivo" con "negativo".*

*Dimostrazione.* Restringiamoci al caso del minimo. Se  $x_0$  è un punto di minimo, allora lo è lungo ogni direzione  $v$ . Allora la derivata direzionale lungo  $v$  è nulla per ogni  $v$ , e dunque, poiché  $\vec{\nabla} f_{x_0}(v) = \partial_v f(x_0)$ , otteniamo che  $\vec{\nabla} f_{x_0} = 0$ . Per lo stesso motivo, la derivata direzionale seconda è maggiore o uguale a zero

lungo ogni direzione, perciò il differenziale secondo è positivo perché  $d^2 f_{x_0}(v)^2 = \partial_v^2 f(x_0)$ . Viceversa, sviluppando  $f$  con il suo polinomio di Taylor al secondo ordine otteniamo

$$f(x_0 + u) - f(x_0) = \frac{d^2 f_{x_0}(u)^2}{2} + o(\|u\|^2) \geq \frac{\alpha \|u\|^2}{2} + o(\|u\|^2)$$

L'espressione di sopra è positiva per  $\|u\|$  sufficientemente piccolo, da cui la tesi. Si noti che abbiamo avuto bisogno della nuova definizione di "definito positivo", per evitare che il differenziale secondo venisse pappato dall' $o(\|u\|^2)$ .  $\square$

## 6.6 Massimi e minimi vincolati

Supponiamo di avere una famiglia finita di funzioni  $C^1$   $f_1, \dots, f_n : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , che chiameremo vincoli, e una funzione  $C^1$   $f_0 : \Omega \rightarrow \mathbb{R}$ , detta obiettivo. Consideriamo l'intersezione degli insiemi di livello dei vincoli, ossia  $\Sigma = \{x \in \mathbb{R}^n \text{ t.c. } f_i(x) = 0 \text{ per } i = 1, \dots, n\}$ , e supponiamo di voler minimizzare  $f$  su tale insieme. In parole povere, stiamo cercando un punto di minimo di  $f$  imponendo alcune condizioni sulla soluzione. Ci viene in aiuto, ancora una volta, il solito Lagrange.

**Teorema 6.6.1** (Principio dei moltiplicatori di Lagrange). *Nelle ipotesi dette sopra, se  $\bar{x} \in \Sigma$  è un minimo locale (cioè in un intorno della forma  $\overline{B(\bar{x}, r)} \cap \Sigma$ ) per  $f|_{\Sigma}$ , i gradienti  $\{\vec{\nabla} f_i(\bar{x})\}_{i=0, \dots, n}$  sono linearmente dipendenti.*

*Dimostrazione.* Seguiremo la dimostrazione di De Giorgi, grande matematico vissuto a Pisa nel secolo scorso.<sup>5</sup> Supporremo che  $\bar{x}$  sia un punto di minimo locale.

Consideriamo, al variare di  $k \in \mathbb{N}$ , la funzione

$$g_k(x) = \|x - \bar{x}\|^2 + f_0(x) + k \sum_{i=1}^n (f_i(x))^2$$

Per continuità sul compatto  $\overline{B(\bar{x}, r)}$ ,  $g_k$  ammette minimo  $x_k \in \overline{B(\bar{x}, r)}$ . L'obiettivo è mostrare che  $x_k \rightarrow \bar{x}$  (a meno di considerare una sottosuccessione) e che dunque la successione sta definitivamente nella palla aperta  $B$ , dove i minimi sono liberi. Procediamo con ordine.

Innanzitutto, osserviamo che la successione  $g_k(x_k)$  è limitata in  $\mathbb{R}$ . Infatti

$$\min_{x \in \overline{B(\bar{x}, r)}} f_0(x) \leq \|x_k - \bar{x}\|^2 + f_0(x_k) + k \sum_{i=1}^n (f_i(x_k))^2 = \min_{x \in \overline{B(\bar{x}, r)}} g_k(x) \leq g_k(\bar{x}) = f_0(\bar{x})$$

Dove l'ultima uguaglianza segue dal fatto che i vincoli si annullano in  $\bar{x}$ . Allora esiste una sottosuccessione tale che  $x_{k(j)} \rightarrow \zeta \in \overline{B(\bar{x}, r)}$  e  $g_{k(j)}(x_{k(j)}) \rightarrow c \in \mathbb{R}$  per  $j \rightarrow +\infty$ .

Ora,  $\zeta$  appartiene a  $\Sigma$ , ossia annulla tutti i vincoli. Infatti consideriamo la successione

$$k(j) \sum_{i=1}^n f_i(x_{k(j)})^2 = g_{k(j)}(x_k) - \|x_{k(j)} - \bar{x}\|^2 - f_0(x_{k(j)}) \rightarrow c - \|\zeta - \bar{x}\|^2 - f_0(\zeta)$$

Poiché tale successione ha limite finito per  $k(j) \rightarrow +\infty$  ed è il prodotto di  $\sum_{i=1}^n f_i(x_{k(j)})^2$  per una successione divergente, si deve avere che  $\lim_{j \rightarrow +\infty} \sum_{i=1}^n f_i(x_{k(j)})^2 = \sum_{i=1}^n f_i(\zeta)^2 = 0$ .

Poiché  $\zeta \in \overline{B(\bar{x}, r)} \cap \Sigma$ ,  $f_0(\bar{x}) \leq f_0(\zeta)$ . Dunque, riprendendo una disuguaglianza della prima stima, abbiamo che

$$\begin{aligned} \|x_{k(j)} - \bar{x}\|^2 + f_0(x_k) &\leq f_0(\bar{x}) \leq f_0(\zeta) \\ \|x_{k(j)} - \bar{x}\|^2 &\leq f_0(\zeta) - f_0(x_k) = o(1) \end{aligned}$$

Dunque il limite della sottosuccessione è  $\bar{x}$ . Questo significa che, definitivamente, la successione appartiene alla palla  $B(\bar{x}, r)$  aperta. Dunque, poiché  $x_{k(j)}$  è un minimo su un aperto, cioè un minimo libero, il gradiente di  $g_k$  in tale punto è nullo. Allora, calcolando un po' di gradienti,

$$\vec{\nabla} g_{k(j)}(x_{k(j)}) = 2(x_{k(j)} - \bar{x}) + \vec{\nabla} f_0(x_{k(j)}) + k(j) \sum_{i=1}^n 2f_i(x_{k(j)}) \vec{\nabla} f_i(x_{k(j)}) = 0$$

<sup>5</sup>Spero che citarlo in queste dispense non lo faccia rivoltare nella tomba.

Ora come ora non possiamo ancora passare al limite, perché i  $k(j)$  potrebbero scappare.<sup>6</sup> Riscriviamo tutto come

$$\sum_{i=0}^n c_{i,j} \vec{\nabla} f_i(x_{k(j)}) = 2(\bar{x} - x_{k(j)})$$

per certi coefficienti  $c_{i,j}$ . Definiamo il vettore  $c_j = (c_{0,j}, \dots, c_{n,j})$ , che ha norma non nulla perché  $c_{0,j} = 1$  per costruzione. Possiamo allora dividere ambo i membri dell'uguaglianza per  $\|c_j\|$ .

$$\sum_{i=0}^n \lambda_{i,j} \vec{\nabla} f_i(x_{k(j)}) = 2 \frac{\bar{x} - x_{k(j)}}{\|c_j\|}$$

Dove si è posto  $\lambda_j = c_j/\|c_j\|$ . Ora, ciascun  $\lambda_{i,j}$  al variare di  $i = 0, \dots, n$  è limitato in modulo da 1, in quanto  $\|(\lambda_{i,j})_{i=1, \dots, n}\| = 1$  per costruzione. Allora, a meno di estrarre delle sottosuccessioni, possiamo supporre che  $\lambda_{i,j} \rightarrow \lambda_i$  per  $j \rightarrow +\infty$ . Si osservi che, dato che il vettore  $(\lambda_{i,j})_{i=1, \dots, n}$  ha norma 1 anche il limite sta nella sfera unitaria, e dunque in particolare definisce una combinazione lineare con coefficienti non tutti nulli. Otteniamo allora

$$\sum_{i=0}^n \lambda_i \vec{\nabla} f_i(x_{k(j)}) = \lim_{j \rightarrow +\infty} 2 \frac{\bar{x} - x_{k(j)}}{\|c_j\|} = 0$$

in quanto  $\|c_j\| \geq 1$ . Ma allora abbiamo appena mostrato che i gradienti sono linearmente indipendenti.  $\square$

## 6.7 Teorema del differenziale totale

Ammettere tutte le derivate parziali non implica la differenziabilità. Ad esempio, consideriamo la funzione

$$f(x, y) = \begin{cases} \phi\left(\frac{y}{x^2}\right) & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

Dove

$$\phi(t) = \begin{cases} t - 1 & \text{se } 1 \leq t \leq 2 \\ 3 - t & \text{se } 2 \leq t \leq 3 \\ 0 & \text{altrove} \end{cases}$$

La funzione solleva la parabola  $y = 2x^2$  ad altezza 1 in modo continuo, ma è nulla nell'origine, perciò non è ivi continua. Ciò nonostante, essa ammette tutte le derivate direzionali in zero, che sono nulle perché ogni retta è definitivamente al di sopra di ogni parabola.

Per giungere a un teorema che assicuri la differenziabilità introduciamo il concetto di differenziale parziale. Sia  $f : \Omega \subseteq E_1 \times E_2 \rightarrow F$  una funzione tra spazi di Banach e sia  $(x, y) \in \Omega$ . Consideriamo la funzione  $j_y : E_1 \rightarrow E_1 \times \{y\}$  tale che  $j_y(x') = (x', y)$ , banalmente continua, e restringiamola all'insieme  $\Omega_y = \{x' \in E_1 \text{ t.c. } (x', y) \in \Omega\}$ , cioè agli  $x'$  la cui "ordinata" è  $y$  e che con essa formano una coppia in  $\Omega$ . Tale insieme è ancora aperto, perché controimmagine di  $\Omega$  secondo la proiezione su  $x$  e poi secondo  $j_y$ . Si dice differenziale parziale nella prima coordinata il differenziale, se esiste, di  $f \circ j_y : \Omega_y \rightarrow F$ . In tal caso si pone

$$d_2 f(x, y) = d(f \circ j_y)x$$

Si dà un'analogia definizione del differenziale parziale nella seconda coordinata.

Il concetto di differenziale parziale estende quello di derivata parziale. Sotto questa luce è banale osservare che se  $f$  è differenziabile ammette tutti i differenziali parziali, e che il differenziale "totale" è la somma dei differenziali parziali applicati alle componenti. Il seguente teorema mostra che, sotto certe condizioni, l'implicazione si può invertire.

<sup>6</sup>De Giorgi, perdonami perché non so quello che faccio.

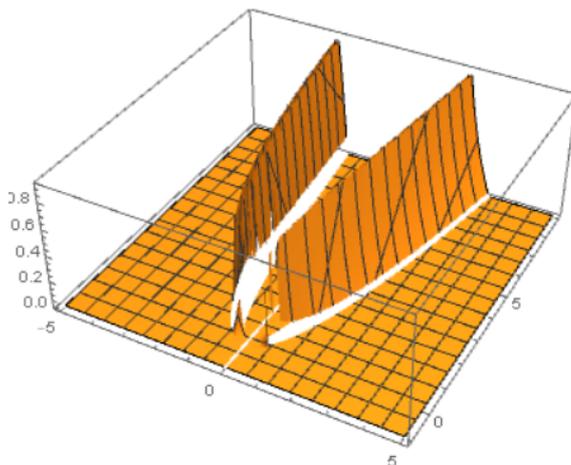


Figura 6.1: Rappresentazione della funzione definita in 6.7. La qualità è quella che è, cercate di vederlo più con la mente che con gli occhi.

**Teorema 6.7.1** (Differenziale totale). *Sia  $f : E_1 \times E_2 \times \dots \times E_n \rightarrow F$  una funzione definita in un intorno di  $x = (x_1, \dots, x_n)$ . Se in  $x$  esistono tutti i differenziali parziali e  $n - 1$  di questi sono continui in un intorno di  $x$  (come funzioni delle restanti variabili), allora  $f$  è differenziabile in  $x$ .*

*Dimostrazione.* WLOG ci riduciamo al caso in cui esistono solo due differenziali parziali  $d_1$  e  $d_2$ , di cui il secondo continuo. Mostriamo che  $f$  ammette differenziale pari alla somma dei differenziali. In altri termini scriviamo

$$f(x + u, y + v) - f(x, y) - d_1 f_{(x,y)}(u) - d_2 f_{(x,y)}(v)$$

E osserviamo che è un  $o(\|u\| + \|v\|)$ . Sommiamo e sottraiamo per "completare" il primo differenziale parziale:

$$\begin{aligned} f(x + u, y + v) - f(x + u, y) + (f(x + u, y) - f(x, y) - d_1 f_{(x,y)}(u)) - d_2 f_{(x,y)}(v) = \\ = f(x + u, y + v) - f(x + u, y) - d_2 f_{(x,y)}(v) + o(\|u\|) \end{aligned}$$

Ora, l'espressione rimanente è la differenza tra i valori assunti dalla curva  $f(x + u, y + tv) - td_2 f_{(x,y)}(v)$  in  $t = 1$  e  $t = 0$ . Allora per teorema del valore medio per curve

$$[f(x + u, y + tv) - td_2 f_{(x,y)}(v)]_0^1 \leq (1 - 0) \sup_{0 \leq t \leq 1} \|d_2 f_{(x+u, y+tv)}(v) - d_2 f_{(x,y)}(v)\|$$

E la quantità interna è  $o(\|v\|)$  uniformemente in  $u$  per continuità di  $d_2 f$  in entrambe le variabili.  $\square$

Notare che si poteva dire subito che

$$f(x + u, y + v) - f(x + u, y) - d_2 f_{(x,y)}(v) = d_2 f_{(x+u, y)}(v) - d_2 f_{(x,y)}(v) + o(\|v\|)$$

Tuttavia, non si poteva concludere per la continuità di  $d_2 f$ . Infatti, otterremmo solo che, per ogni  $u$  fissato, la quantità sopra è  $o(\|v\|)$ . Questo non ci assicura la convergenza, perché la relazione dipende dall' $u$  scelto. È come quando abbiamo osservato che la continuità di tutte le restrizioni sulle rette non equivale alla continuità globale.

Diremo che una funzione  $f : E \rightarrow F$  tra spazi di Banach è di classe  $C^1$  se ammette differenziale in ogni punto e la funzione  $df : E \rightarrow L(E, F)$  che manda ogni punto nel suo differenziale è continua. Rileggendo la dimostrazione, possiamo enunciare il seguente corollario:

**Corollario 6.7.1.1.** *Una funzione è di classe  $C^1$  se e solo se è parzialmente derivabile con continuità in ogni restrizione.*

## 6.8 Derivate parziali miste

Le derivate parziali commutano per definizione.

---

*Gigi Rolandi*

In questa sezione vedremo un teorema che permette di scambiare le derivate parziali miste sotto certe condizioni (abbastanza lasche, devo dire).

**Teorema 6.8.1** (di Schwarz). *Siano  $I, J$  due intervalli di  $\mathbb{R}$ , sia  $E$  un Banach e sia  $f : I \times J \rightarrow E$ . Sia inoltre  $(x_0, y_0)$  un punto interno del dominio. Se valgono le seguenti condizioni:*

- *Esiste la derivata parziale nella prima variabile  $\partial_x f$  in un intorno di  $(x_0, y_0)$ ;*
- *Esiste la derivata mista  $\partial_y \partial_x f$  in tale intorno, ed è continua nel punto  $(x_0, y_0)$ ;*
- *Esiste la derivata parziale nella seconda variabile  $\partial_y f$  lungo la retta  $(x, y_0)$*

*Allora esiste l'altra derivata mista  $\partial_x \partial_y f$  nel punto  $(x_0, y_0)$ , e le due derivate miste coincidono.*

*Dimostrazione.* Consideriamo la quantità

$$\|f(x_0 + u, y_0 + v) - f(x_0 + u, y_0) - f(x_0, y_0 + v) + f(x_0, y_0) - \partial_y \partial_x f(x_0, y_0)uv\|$$

Vogliamo mostrare che tale quantità è  $o(uv)$ . Se questo è vero, allora dividendo per  $|v|$  otteniamo

$$\left\| \frac{f(x_0 + u, y_0 + v) - f(x_0 + u, y_0)}{v} - \frac{f(x_0, y_0 + v) - f(x_0, y_0)}{v} - \partial_y \partial_x f(x_0, y_0)u \right\| = o(u)$$

Dove l' $o$  piccolo è uniforme in  $v$ . Allora possiamo mandare  $v \rightarrow 0$  e ottenere la derivata parziale nella seconda componente, che per ipotesi esiste lungo la retta a  $y_0$  fissato

$$\|\partial_y f(x_0 + u, y_0) - \partial_y f(x_0, y_0) - \partial_y \partial_x f(x_0, y_0)u\| = o(u)$$

Da cui, per definizione, la derivata mista esiste e vale che  $\partial_x \partial_y f(x_0, y_0) = \partial_y \partial_x f(x_0, y_0)$ .

Ora, per ottenere quanto voluto, invociamo il solito teorema del valor medio.

$$\begin{aligned} & \|f(x_0 + u, y_0 + v) - f(x_0 + u, y_0) - f(x_0, y_0 + v) + f(x_0, y_0) - \partial_y \partial_x f(x_0, y_0)uv\| \leq \\ & \leq \sup_{t \in [0,1]} \|\partial_x f(x_0 + tu, y_0 + v) - \partial_x f(x_0 + tu, y_0) - \partial_y \partial_x f(x_0, y_0)v\| |u| \end{aligned}$$

Poiché la derivata mista  $\partial_y \partial_x$  esiste in un intorno di  $(x_0, y_0)$ , possiamo reiterare il ragionamento e stimare la quantità di sopra con

$$\sup_{t,s \in [0,1]} \|\partial_y \partial_x f(x_0 + tu, y_0 + sv) - \partial_y \partial_x f(x_0, y_0)\| |uv| = o(uv)$$

Dove l'ultimo passaggio segue dalla continuità della derivata mista. □

**Corollario 6.8.1.1.** *Sia  $f : E \rightarrow F$  una funzione differenziabile tra spazi di Banach. Se il differenziale secondo di  $f$  è continuo in un intorno  $U$  di  $x$  allora è simmetrico negli argomenti nel punto  $x$ .*

*Dimostrazione.* Poniamo  $g(t, s) = f(x + tu + sv)$  e osserviamo che, a meno di rinominare le variabili, per quanto osservato in 6.3.1 vale l'uguaglianza

$$\partial_t \partial_s g(t, s) = \partial_{t'} \partial_{s'} f(x + t'u + s'v)|_{s'=s|t'=t} = d^2 f_{x+tu+sv}(u, v)$$

Poiché il differenziale è continuo in un intorno di  $x$ ,  $\partial_t \partial_s g(t, s)$  è continua, e dunque

$$d^2 f_{x+tu+sv}(u, v) = \partial_t \partial_s g(t, s) = \partial_s \partial_t g(t, s) = d^2 f_{x+tu+sv}(v, u)$$

Si conclude per continuità mandando  $t$  e  $s$  a zero. □

## 6.9 Lipschitz in più variabili, ma anche no

**Teorema 6.9.1.** *Sia  $\Omega \subseteq E$  aperto convesso e sia  $f : \Omega \rightarrow F$  differenziabile con differenziale limitato da  $K$  (ossia i differenziali dei singoli punti hanno norma operatoriale al più  $K$ ). Allora  $f$  è lipschitziana.*

*Dimostrazione.* Dati due punti  $x, y$  consideriamo la curva affine  $\gamma(t) = tx + (1-t)y$ , totalmente contenuta in  $\Omega$  per convessità. Allora per teorema del valore medio

$$\|f(x) - f(y)\| \leq \sup_{t \in [0,1]} \|df_{\gamma(t)}(\gamma'(t))\| = \sup_{t \in [0,1]} \|df_{tx+(1-t)y}(x-y)\| \leq \sup_{t \in [0,1]} \|df_{tx+(1-t)y}\| \|x-y\| \leq K \|x-y\|$$

□

Il seguente esempio mostra che è necessaria la convessità (che non è proprio la cosa più facile da ottenere). Consideriamo la funzione  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definita su  $\Omega = B(O, 1) \setminus \{(x, 0), x \geq 0\}$  che è la palla unitaria a cui si toglie il raggio per  $(1, 0)$ . Definiamo  $f$  in coordinate polari come  $f(r, \theta) = r\theta$ . Tale funzione ha differenziale uniformemente limitato, ma non è lipschitziana perché ha un salto quando si passa da una parte all'altra del raggio critico.

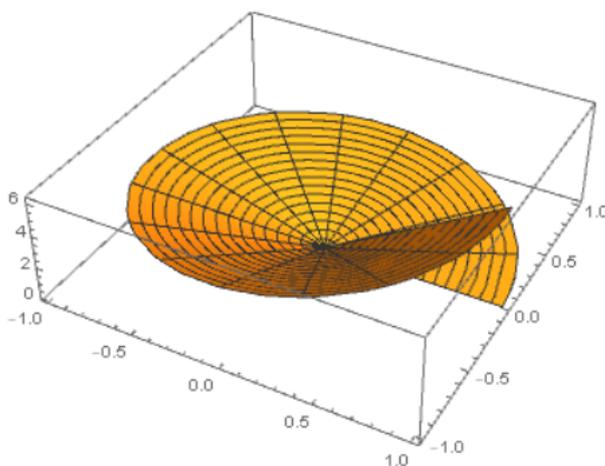


Figura 6.2: Rappresentazione della funzione di cui in 6.9.

## 6.10 Invertibilità locale

**Teorema 6.10.1.** *Sia  $f : E \supseteq \Omega \rightarrow F$  una funzione  $C^1$  tra spazi di Banach. Supponiamo che in  $x_0 \in \Omega$  il differenziale  $df_{x_0}$  sia invertibile. Allora esiste un intorno  $U$  di  $x_0$  tale che  $f|_U : U \rightarrow f(U)$  è un diffeomorfismo  $C^1$  tra aperti.*

*Dimostrazione.* Possiamo restringerci al caso semplice in cui  $df_{x_0}$  è l'identità. Infatti possiamo svolgere la dimostrazione per la funzione  $g = (df_{x_0})^{-1} \circ f : \Omega \rightarrow E$ , che è un diffeomorfismo se e solo se lo è  $f$  in quanto composizione di diffeomorfismi. Ora, se  $dg_{x_0} = I$ , per continuità del differenziale esiste un intorno sferico  $U = B(x_0, r)$  di  $x_0$  in cui  $\|dg_x - I\| \leq 1/2$ . In tale intorno i differenziali  $dg_x$  sono ancora invertibili, per quanto visto in 4.3. In particolare la funzione  $g - I : B(x_0, r) \rightarrow E$  ha differenziale limitato da  $1/2$ , e dunque è  $1/2$ -Lipschitz perché la palla è convessa, come visto in 6.9. Adesso scriviamo  $g = I + (g - I)$  e osserviamo che tale scrittura è una perturbazione lipschitziana dell'identità di costante minore di 1. Dunque ha immagine aperta ed è un omeomorfismo con la sua immagine, come mostrato in 4.4. Inoltre, nella sezione 6.2 (non) abbiamo mostrato che la funzione inversa è differenziabile e il suo differenziale è  $dg_y^{-1} = (dg_{g^{-1}(y)})^{-1}$ , continuo in  $y$  in quanto composizione di tre applicazioni continue<sup>7</sup>:

$$F \xrightarrow{g^{-1}} U \xrightarrow{dg} L(E) \xrightarrow{-1} L(E)$$

□

<sup>7</sup>Per la continuità dell'inversione di applicazioni lineari vedasi 4.6.

## 6.11 Teorema della funzione implicita

**Teorema 6.11.1** (di Dini). *Siano  $E, F, G$  spazi di Banach,  $\Omega \subseteq E \times F$  aperto e  $f : \Omega \rightarrow G$  di classe  $C^1$ . Sia  $v \in G$  e sia  $\Gamma = \{(x, y) \in \Omega \text{ t.c. } f(x, y) = v\}$  l'insieme di livello della funzione. Sia  $(x_0, y_0) \in \Gamma$  tale che  $d_2f_{(x_0, y_0)}$  invertibile (con inversa lineare e continua). Allora esiste un intorno aperto di  $(x_0, y_0)$  della forma  $U \times V$  ristretto al quale  $\Gamma$  è un grafico. Più precisamente, esiste una funzione  $g : U \rightarrow V$  tale che*

$$\Gamma \cap (U \times V) = \{(x, g(x)), x \in U\}$$

*Inoltre  $g$  è di classe  $C^1$  e ha differenziale  $dg_x = -(d_2f_{(x, g(x))})^{-1} \circ d_1f_{(x, g(x))}$ .*

*Dimostrazione.* A meno di comporre con una traslazione, assumiamo che  $v = 0$ . Per dimostrare la prima parte del teorema ci ricondurremo al teorema di invertibilità locale, considerando la funzione  $\Phi : \Omega \rightarrow E \times G$  definita come  $\Phi(x, y) = (x, f(x, y))$ . Infatti tale funzione è di classe  $C^1$  e il suo differenziale<sup>8</sup> si può esprimere nella seguente forma matriciale a blocchi

$$d\Phi_{(x, y)} = \begin{bmatrix} I & 0 \\ d_1f_{(x, y)} & d_2f_{(x, y)} \end{bmatrix}$$

Dove  $I$  è l'identità e  $0$  è l'applicazione identicamente nulla. Si osserva facilmente che il differenziale di  $\Phi$  è invertibile se e solo se lo è  $d_2f_{(x, y)}$ . Dunque per teorema della funzione implicita  $\Phi$  è un diffeomorfismo con immagine aperta in un intorno di  $(x_0, y_0)$ , che possiamo scegliere della forma  $W \times V$ . Ora, restringiamo l'intorno considerando l'insieme  $U \subseteq W$  degli  $x$  per cui esiste almeno un  $y \in V$  tale che  $(x, y) \in \Gamma$ , o in altri termini che  $\Phi(x, y) = (x, 0)$ .  $U$  è aperto in quanto è la controimmagine continua di  $\Phi(U \times V) \cap (E \times \{0\})$ , che è aperto nella topologia di  $E \times \{0\}$ .

Ora, restringendo  $\Phi$  a  $\Gamma \cap (U \times V)$  otteniamo un omeomorfismo che associa a ogni coppia  $(x, y)$  la coppia  $(x, 0)$ . Possiamo allora definire  $y = g(x) = \pi_y \circ \Phi^{-1}(x, 0)$ , dove  $\pi_y$  è la proiezione sulla seconda componente, e osservare che questa è esattamente la funzione cercata.

Per quanto riguarda la differenziabilità di  $g$ , dalla relazione  $f(x, g(x)) = 0$  otteniamo

$$df_{(x, g(x))} \circ d(x, g(x))_x = \begin{bmatrix} d_1f_{(x, g(x))} & d_2f_{(x, g(x))} \end{bmatrix} \circ \begin{bmatrix} I \\ dg_x \end{bmatrix} = d_1f_{(x, g(x))} + d_2f_{(x, g(x))} \circ dg_x = 0$$

Da cui

$$dg_x = -(d_2f_{(x, g(x))})^{-1} \circ d_1f_{(x, g(x))}$$

□

## 6.12 Passaggio al limite sotto il segno di differenziale

Chiediamoci ora quando una successione di funzioni differenziabili conservi la differenziabilità in un passaggio al limite. Osserviamo subito che la differenziabilità non si preserva passando al limite uniforme.

Ad esempio, la successione di funzioni  $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$  tende uniformemente a  $y = |x|$ , non derivabile nell'origine. Come vedremo in un teorema il cui acronimo è tutto un programma, la condizione giusta da chiedere è la convergenza uniforme dei differenziali.

**Teorema 6.12.1** (LSD). *Siano  $E, F$  due Banach e sia  $\Omega \subseteq E$  un aperto connesso. Sia  $f_n : \Omega \rightarrow F$  una successione di funzioni differenziabili tali che*

- *esiste un  $x_0 \in \Omega$  sul quale la successione  $f_n(x_0)$  converge puntualmente a un certo  $v$ ;*
- *la successione dei differenziali  $df_n : \Omega \rightarrow L(E, F)$  converge localmente uniformemente a una funzione  $g$ , ossia per ogni  $a \in \Omega$  esiste un intorno sferico  $B(a, r_a)$  in cui la successione  $df_n$  converge uniformemente a  $g$ .*

*Allora  $f_n$  converge uniformemente localmente (sugli stessi intorni sferici) a una funzione  $f$ , il cui differenziale coincide con  $g$  e che vale  $v$  in  $x_0$ .*

<sup>8</sup>In realtà, questo è un leggero abuso di notazione, però rende perfettamente l'idea.

*Dimostrazione.* Per ogni  $a \in \Omega$ , la successione  $f_n(x) - f_n(a)$  è di Cauchy uniforme per ogni  $x \in B(a, r_a)$ . Infatti, per teorema del valor medio

$$\|f_n(x) - f_n(a) - f_m(x) + f_m(a)\| \leq \|d(f_n - f_m)\| \|x - a\| \leq \|df_n - df_m\| r_a$$

Poiché  $F$  è di Banach,  $f_n(x) - f_n(a)$  converge, perciò  $\forall x \in B(a, r_a)$   $f_n(x)$  converge se e solo se  $f_n(a)$  converge. Allora l'insieme  $C = \{x \in \Omega \text{ t.c. } f_n(x) \text{ converge}\}$  è aperto, perché se contiene  $a$  contiene anche tutta la palla  $B(a, r_a)$ . Inoltre  $C$  è chiuso, perché per lo stesso ragionamento il suo complementare è aperto. Poiché  $C$  è non vuoto in quanto contiene  $x_0$  si deve avere che  $C = \Omega$ , per connessione del dominio. Osserviamo inoltre che la successione  $f_n(x)$  è di Cauchy uniforme sulla palla, perché

$$\|f_n(x) - f_m(x)\| \leq \|f_n(x) - f_n(a) - f_m(x) + f_m(a)\| + \|f_n(a) - f_m(a)\| \leq \|df_n - df_m\| r_a + \|f_n(a) - f_m(a)\|$$

e la seconda quantità dipende solo dalla velocità di convergenza al centro. Dunque la convergenza è localmente uniforme.

Ora, mostriamo che la funzione limite ha  $g$  come differenziale. Per farlo mostriamo che, fissato  $a \in \Omega$ , per ogni  $x \in B(a, r_a)$  la quantità  $\|f(x) - f(a) - g_a(x - a)\|$  è un  $o(\|x - a\|)$ . Indicheremo con  $g_a(v)$  la funzione  $g$  nel punto  $a$  e calcolata nel vettore  $v$ .

$$\begin{aligned} & \|f(x) - f(a) - g_a(x - a)\| \leq \\ & \leq \|f(x) - f(a) - (f_n(x) - f_n(a))\| + \|f_n(x) - f_n(a) - df_{n_a}(x - a)\| + \|df_{n_a}(x - a) - g_a(x - a)\| \leq \\ & \leq \lim_{m \rightarrow +\infty} \|f_m(x) - f_m(a) - (f_n(x) - f_n(a))\| + o(\|x - a\|) + \|df_n - g\|_{\infty, \overline{B(a, r_a)}} \|x - a\| \leq \\ & \leq \lim_{m \rightarrow +\infty} \|df_m - df_n\|_{\infty, \overline{B(a, r_a)}} \|x - a\| + \|df_n - g\|_{\infty, \overline{B(a, r_a)}} \|x - a\| + o(\|x - a\|) = \\ & = 2\|df_n - g\| \|x - a\| + o(\|x - a\|) = o(\|x - a\|) \end{aligned}$$

□

Osserviamo che il teorema non implica la convergenza uniforme delle  $f_n$ , nemmeno supponendo la convergenza uniforme sul dominio delle derivate. Ad esempio, la successione  $f_n(x) = x/n$  converge puntualmente, ma non uniformemente, alla funzione identicamente nulla, anche se le derivate convergono uniformemente su tutto  $\mathbb{R}$ .

### 6.13 Derivare l'integrando

Mostriamo ora un teorema di “Derivating Under The Integral Sign”, più noto tra i fisici con l'acronimo “DUTIS” dato che per loro far commutare le derivate e gli integrali è un dovere morale.<sup>9</sup>

**Teorema 6.13.1** (DUTIS). *Sia  $K \subseteq (E, \|\cdot\|)$  un sottoinsieme compatto di uno spazio di Banach e sia  $f(x, t) : K \times [a, b] \rightarrow \mathbb{R}$  una funzione continua che ammette differenziale parziale rispetto a  $x$  continuo (in entrambe le variabili). Allora la funzione*

$$g(x) = \int_a^b f(x, t) dt$$

è di classe  $C^1$  e il suo differenziale è l'integrale del differenziale parziale dell'argomento, cioè

$$dg_x[h] = \int_a^b d_x f_{(x,t)}[h] dt$$

*Dimostrazione.* Osserviamo che

$$\left\| g(x+h) - g(x) - \int_a^b d_x f_{(x,t)}[h] dt \right\| = \left\| \int_a^b (f(x+h, t) - f(x, t) - d_x f_{(x,t)}[h]) dt \right\|$$

<sup>9</sup>Se chiedete a un fisico perché sia possibile farlo vi risponderà qualcosa del tipo: “Le ipotesi sono sempre verificate. Quali? Boh.”

$$\left\| g(x+h) - g(x) - \int_a^b d_x f_{(x,t)}[h] dt \right\| = \left\| \int_a^b [f(x+sh, t) - s d_x f_{(x,t)}[h]]_{s=0}^{s=1} dt \right\|$$

Detto  $\omega$  il modulo di continuità di  $d_x f$  su  $K \times [a, b]$  possiamo scrivere

$$\left\| g(x+h) - g(x) - \int_a^b d_x f_{(x,t)}[h] dt \right\| \leq (b-a) \sup_{t \in [a,b], s \in [0,1]} \|d_x f_{(x+sh,t)}[h] - d_x f_{(x,t)}[h]\|$$

$$\left\| g(x+h) - g(x) - \int_a^b d_x f_{(x,t)}[h] dt \right\| \leq (b-a) \|h\| \omega(\|h\|) = o(\|h\|)$$

Dunque otteniamo la tesi per definizione di differenziale. Adesso è banale verificare che  $dg_x$  è continua, sfruttando sempre il modulo di continuità di  $d_x f_{(x,t)}$ .  $\square$

## 6.14 Funzioni di classe $C^k$

In questa sezione intendo fare un patchwork di tutto quanto detto finora in modo da ottenere dei risultati per le funzioni tra  $\mathbb{R}^n$  e  $\mathbb{R}$ . Premettiamo innanzitutto la seguente notazione.

**Definizione 24.** Dato un multi-indice  $\alpha = (\alpha_1, \dots, \alpha_n) \subseteq \mathbb{N}^n$ , definiamo il suo peso come

$$|\alpha| = \alpha_1 + \dots + \alpha_n$$

**Definizione 25.** Data una funzione  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e un multi-indice  $\alpha$ , indichiamo con  $\partial_\alpha f$  la derivata mista (se esiste)

$$\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}$$

Con la convenzione che  $\partial_{x_i}^0 f = f$ . Diciamo inoltre che tale derivata ha ordine  $|\alpha|$ .

Osserviamo che se  $f$  è di classe  $C^k$ , con  $k \geq |\alpha|$ , per il teorema di Schwarz l'ordine delle derivate parziali è indifferente.

Abbiamo allora gli strumenti per enunciare il seguente lemma:

**Lemma 6.14.1.** Una funzione  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  è di classe  $C^k$  se e solo se tutte le derivate parziali miste fino all'ordine  $k$  esistono e sono continue.

La dimostrazione è banale, perché in un verso segue dalla definizione di differenziabilità e nell'altro è un'applicazione del teorema del differenziale totale.

**Definizione 26.** L'insieme delle funzioni continue e limitate fino all'ordine  $k$  è l'insieme

$$BC^k(\mathbb{R}^n) = \{f : \mathbb{R}^n \rightarrow \mathbb{R} \text{ t.c. } \partial_\alpha f \text{ è limitata } \forall \alpha \in \mathbb{N}^n \text{ con } |\alpha| \leq k\}$$

**Teorema 6.14.2.**  $BC^k$  è uno spazio di Banach con la norma

$$\|f\|_{C^k} = \max_{|\alpha| \leq k} \|\partial_\alpha f\|_\infty$$

*Dimostrazione.* Per dimostrare la completezza osserviamo che, detto  $N = \#\{\alpha \text{ t.c. } |\alpha| \leq k\}$ ,<sup>10</sup> l'insieme considerato si immerge isometricamente nel prodotto  $BC(\mathbb{R}^n)^N$  munito della norma infinito sulle componenti, mediante l'applicazione che manda  $f$  nella  $N$ -upla delle sue derivate. L'insieme di arrivo è chiuso, perché se una successione di funzioni  $C^k$  converge uniformemente per ogni ordine di derivate il limite sarà anch'esso di classe  $C^k$ , per teorema dell'LSA. Allora  $BC^k$  si identifica con un chiuso in un prodotto finito di spazi metrici completi, che è banalmente completo.  $\square$

<sup>10</sup>Potete divertirvi a calcolare l'effettivo valore di  $N$ ; io passo tranquillamente.

## 6.15 Funzioni omogenee di Eulero

**Definizione 27.** Sia  $\Omega \subseteq E$  aperto tale che  $0 \notin \Omega$  e che, per ogni  $x \in \Omega$ , il segmento  $\left\{ tx + (1-t)\frac{x}{\|x\|} \right\}_{t \in [0,1]}$  è totalmente contenuto in  $\Omega$ . Una funzione  $f : \Omega \rightarrow \mathbb{R}$  si dice  $p$ -omogenea se  $f(tx) = t^p f(x)$  per ogni  $x$  nel dominio e per ogni  $t \neq 0$ .

**Teorema 6.15.1** (Eulero). Una funzione differenziabile è  $p$ -omogenea se e solo se  $df_y[y] = pf(y)$  per ogni  $y \in \Omega$ .

*Dimostrazione.* Supponiamo che  $f$  sia  $p$ -omogenea. Derivando la formula rispetto a  $t$  otteniamo

$$df_{tx}(x) = df_{tx}[x] = pt^{p-1}f(x)$$

Da cui la tesi con  $t = 1$ .

Viceversa, supponiamo che  $df_y[y] = pf(y)$  e mostriamo che  $f(tx) = t^p f(x)$ . Per farlo, possiamo mostrare che la quantità  $f(tx)t^{-p}$  è costante rispetto a  $t$ , da cui segue che è uguale a  $f(x)$  ponendo  $t = 1$ . Allora deriviamo tale formula rispetto a  $t$  e mostriamo che la derivata è nulla.

$$(f(tx)t^{-p})' = f_{tx}[x]t^{-p} - pf(tx)t^{-p-1} = (df_{tx}[tx] - pf(tx))t^{-p-1}$$

Ora, per ipotesi  $df_{tx}[tx] = pf(tx)$ , dove abbiamo posto  $y = tx$ . Dunque  $(f(tx)t^{-p})' = 0$ , come voluto.  $\square$

Si osservi che, se  $f$  è definita su  $\mathbb{R}^n \setminus \{0\}$ , abbiamo appena determinato tutte e sole le soluzioni dell'equazione differenziale alle derivate parziali

$$\vec{\nabla} f_y \cdot y = \sum_{i=1}^n \partial_i f(y) y_i = pf(y)$$

Per le 1-forme chiuse  $p$ -omogenee è possibile determinare esplicitamente una primitiva.

**Teorema 6.15.2.** Sia  $\Omega \subseteq E$  aperto tale che, per ogni  $x \neq 0 \in \Omega$ , il segmento  $\left\{ tx + (1-t)\frac{x}{\|x\|} \right\}_{t \in [0,1]}$  è totalmente contenuto in  $\Omega$ . Sia  $\omega$  una forma chiusa e omogenea secondo Eulero di grado  $p \neq 1$ . Allora  $\omega$  è esatta e una sua primitiva è

$$f(x) = \frac{1}{p+1} \omega(x)[x]$$

*Dimostrazione.* Per costruzione

$$df_x[u] = \frac{1}{p+1} d(\omega(x)[x])[u] = \frac{1}{p+1} (d\omega(x)[u, x] + \omega(x)[u])$$

Per chiusura di  $\omega$  vale che

$$df_x[u] = \frac{1}{p+1} (d\omega(x)[x] + \omega(x)[u])$$

Dal momento che, essendo  $\omega$   $p$ -omogenea,  $d\omega(x)[x] = p\omega(x)$ , si ha che  $df_x[u] = \omega(x)[u]$ , come voluto.  $\square$

# Capitolo 7

## 1-forme

D'ora in avanti, sia  $\Omega \subseteq (E, \|\cdot\|)$  un aperto di uno spazio di Banach. Ricordiamo che possiamo identificare il differenziale di una funzione continua  $f \in C(\Omega, \mathbb{R})$  con la funzione che associa a ogni punto  $x_0 \in \Omega$  il funzionale lineare  $df_{x_0} \in L(E, \mathbb{R}) = E^*$ . Vogliamo provare a rispondere alla domanda opposta: quando è vero che una funzione  $\omega : \Omega \rightarrow E^*$  è il differenziale di una funzione?

Premettiamo alcune definizioni.

**Definizione 28** (1-forma). Una 1-forma<sup>1</sup> è una funzione continua  $\omega : \Omega \rightarrow E^*$ .

Possiamo pensare a una 1-forma differenziale come un campo di “covettori”, cioè di elementi del duale. Se  $\Omega \subseteq \mathbb{R}^n$  un covettore si scrive come combinazione lineare degli elementi della base duale associata alla base canonica  $e_i^*$ , che si possono indicare con  $dx_i$ . Dunque possiamo scrivere

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

Dove  $\alpha_i \in C^0(\Omega, \mathbb{R})$  sono coefficienti continui. Inoltre, per teorema di Riesz a ogni funzionale continuo  $\phi$  corrisponde un vettore  $v$  dello spazio tale che  $\phi(w) = \langle v, w \rangle$ . Dunque possiamo associare a  $\omega$  un campo vettoriale  $X : \Omega \rightarrow \mathbb{R}^n$ , dato da

$$X(x) = \sum_{i=1}^n \alpha_i(x) e_i$$

**Definizione 29** (Integrale di una 1-forma). Sia  $\gamma \in C^1([a, b], \Omega)$  una curva. Definiamo l'integrale di linea di una 1-forma  $\omega$  lungo  $\gamma$  come

$$\int_{\gamma} \omega = \int_a^b \omega(\gamma(t))[\gamma'(t)] dt$$

La definizione si estende naturalmente alle curve  $C^1$  a tratti, sommando sulle componenti differenziabili.

Per costruzione l'integrale di una 1-forma gode di tutte le proprietà dell'integrale di Riemann. Inoltre, è invariante per riparametrazioni (strettamente crescenti) della curva, in quanto se  $\gamma = \delta \circ \sigma$  abbiamo, per teorema di cambio di variabile:

$$\begin{aligned} \int_{\gamma} \omega &= \int_a^b \omega(\gamma(t))[\gamma'(t)] dt = \int_a^b \omega(\delta(\sigma(t)))[(\delta \circ \sigma)'(t)] dt \\ \int_{\gamma} \omega &= \int_a^b \omega(\delta(\sigma(t)))[\delta'(\sigma(t))]\sigma'(t) dt = \int_{\sigma(a)}^{\sigma(b)} \omega(\delta(s))[\delta'(s)] ds = \int_{\delta} \omega \end{aligned}$$

Si osservi che se invece la riparametrazione invertisse l'ordine degli estremi di integrazione, ossia se  $\sigma$  fosse strettamente decrescente, l'integrale cambierebbe di segno.

---

<sup>1</sup>Molti, tra cui Majer, parlano di “forme differenziali”, ma ho preferito riservare l'aggettivo “differenziale” alle forme su varietà, che forse vedremo. In ogni caso, mi scuso in anticipo se dovessi venir meno a questo formalismo, cosa che succederà sicuramente.

Come conseguenza della possibilità di riscalarare il dominio e dell'ovvia additività dell'integrale di Riemann sugli estremi di integrazione si ha che, detta  $\gamma * \delta$  la giunzione di due curve  $C^1$  a tratti,

$$\int_{\gamma * \delta} \omega = \int_{\gamma} \omega + \int_{\delta} \omega$$

ossia l'integrale è additivo sulla curva.

Infine, il seguente fatto è degno di nota:

**Teorema 7.0.1** (Continuità nella curva). *L'integrale di  $\omega$ , come funzione su  $C^1([0, 1], \Omega)$ , è continuo nella distanza  $C^1$ .*

*Dimostrazione.* Dobbiamo mostrare che, se una successione  $\{\gamma_n\}_{n \in \mathbb{N}}$  converge uniformemente a  $\gamma$  e lo stesso vale per le derivate, il limite degli integrali è l'integrale del limite. Per fare ciò, è sufficiente mostrare che la successione delle funzioni integrande  $\omega(\gamma_n(t))[\gamma'_n(t)]$  converge uniformemente a  $\omega(\gamma(t))[\gamma'(t)]$  e sfruttare il teorema di limite sotto il segno di integrale visto ad Analisi 1. Osserviamo dunque che, per un certo  $t \in [0, 1]$  fissato,

$$\|\omega(\gamma_n(t))[\gamma'_n(t)] - \omega(\gamma(t))[\gamma'(t)]\| \leq \|\omega(\gamma_n(t)) - \omega(\gamma(t))\| \|\gamma'_n(t)\| + \|\omega(\gamma(t))[\gamma'_n(t) - \gamma'(t)]\|$$

Fissiamo ora un  $\varepsilon > 0$  e un  $n$  per cui sia le curve che le loro derivate distano uniformemente dai rispettivi limiti al più  $\varepsilon$ . Poiché  $\|\gamma'(t)\|$  è una funzione continua su un compatto è possibile definirne un massimo  $M$ . Dunque per convergenza uniforme possiamo stimare definitivamente  $\|\gamma'_n(t)\| \leq M + \varepsilon$ . Inoltre,  $\|\omega(\gamma(t))\|$  ammette un massimo, sia esso  $L$ , perché è continua sul compatto  $[0, 1]$ . Dunque il secondo addendo è minore di  $L\varepsilon$ .

Infine,  $\|\omega(\gamma_n(t)) - \omega(\gamma(t))\|$  è infinitesimo uniformemente in  $t$ , dato che  $\gamma([0, 1])$  è compatto e dunque si può applicare l'estensione del teorema di Heine-Cantor vista in 1.4.4. Stimiamo allora tale fattore con un certo  $\eta(\varepsilon)$  infinitesimo, ottenendo dunque

$$\sup_{t \in [0, 1]} \|\omega(\gamma_n(t))[\gamma'_n(t)] - \omega(\gamma(t))[\gamma'(t)]\| \leq \eta(\varepsilon)(M + \varepsilon) + L\varepsilon$$

Perciò gli integrandi convergono, come voluto. □

## 7.1 Forme chiuse e esatte

**Definizione 30** (1-forma esatta). *Una 1-forma  $\omega$  si dice esatta se è il differenziale di una funzione, ossia se esiste  $f : \Omega \rightarrow \mathbb{R}$  tale che  $\omega = df$ . Una 1-forma si dice localmente esatta se per ogni punto  $x \in \Omega$  esiste un intorno su cui  $\omega$  è il differenziale di una funzione.*

In  $\mathbb{R}^n$ , ciò equivale a dire che il campo vettoriale  $X$  associato è conservativo, cioè che esiste un potenziale  $f$  di cui è il gradiente.

**Teorema 7.1.1** (CNES per l'esattezza). *Sia  $\omega$  una 1-forma differenziale su  $\Omega$ . I seguenti fatti sono equivalenti:*

1.  $\omega$  è esatta;
2. l'integrale di  $\omega$  lungo una qualsiasi curva chiusa di classe  $C^1$  a tratti è nullo.
3. l'integrale di  $\omega$  lungo una curva  $\gamma$  di classe  $C^1$  a tratti dipende solo dal punto iniziale e dal punto finale;

*Dimostrazione.* 1)  $\implies$  2) Sia  $\omega = df$  e sia  $\gamma$  un laccio  $C^1$  a tratti. Allora per teorema fondamentale del calcolo integrale

$$\int_{\gamma} \omega = \int_a^b df(\gamma(t))[\gamma'(t)]dt = \int_a^b (f \circ \gamma)'(t)dt = f(\gamma(b)) - f(\gamma(a)) = 0$$

2)  $\implies$  3) Basta osservare che, se  $\gamma$  e  $\delta$  hanno gli stessi estremi, detta  $\gamma * \bar{\delta}$  la concatenazione di  $\gamma$  con la curva inversa di  $\delta$  si ha, per additività dell'integrale di Riemann, che

$$0 = \int_{\gamma * \bar{\delta}} \omega = \int_{\gamma} \omega + \int_{\bar{\delta}} \omega = \int_{\gamma} \omega - \int_{\delta} \omega$$

L'ultimo passaggio viene dal fatto che  $\int_{\bar{\delta}} \omega = \int_b^a \omega(\delta(t))[\delta'(t)]dt$ .

3)  $\implies$  1) Supponiamo che  $\Omega$  sia connesso, e dunque connesso per archi  $C^1$  essendo un aperto di uno spazio di Banach. Questa assunzione non è restrittiva dato che possiamo restringerci a ciascuna componente connessa. Sia  $x_0 \in \Omega$  e per ogni  $x \in \Omega$  sia  $\gamma_x$  una curva  $C^1$  che collega  $x_0$  a  $x$ . Poniamo  $f(x) = \int_{\gamma_x} \omega$ , che è ben definita dato che il cammino non dipende dalla curva scelta. Mostriamo ora che  $df = \omega$ .

Per farlo, sia  $B(x, r) \subseteq \Omega$  una palla centrata in  $x$  e sia  $\|h\| \leq r$ . Abbiamo che

$$f(x+h) - f(x) - \omega(x)[h] = \int_{\gamma_{x+h}} \omega - \int_{\gamma_x} \omega - \omega(x)[h]$$

A meno di cambiare curva, posso supporre che  $\gamma_{x+h}$  si ottenga dalla concatenazione di  $\gamma_x$  con il segmento  $\delta(t) = x+th$  che collega  $x$  a  $x+h$ . Dunque possiamo semplificare la differenza per additività dell'integrale e ottenere

$$f(x+h) - f(x) - \omega(x)[h] = \int_{\delta} \omega - \omega(x)[h] = \int_0^1 (\omega(x+th)[h] - \omega(x)[h])dt$$

Passando ai moduli, per continuità di  $\omega$  otteniamo

$$|f(x+h) - f(x) - \omega(x)[h]| \leq \sup_{t \in [0,1]} \|\omega(x+th) - \omega(x)\| \|h\| = o(1)\|h\|$$

Da cui la tesi.  $\square$

**Definizione 31** (Forma chiusa). Una forma  $\omega \in C^1(\Omega, E^*)$  si dice chiusa se il suo differenziale, come applicazione bilineare  $d\omega \in L(E, E^*) \cong E^{**}$ , è simmetrico.

Sempre nell'analogia tra forme e campi su  $\mathbb{R}^n$ , una forma è chiusa se il campo vettoriale associato  $X$  è irrotazionale, ossia per ogni  $i, j \leq n$  vale che

$$\frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}$$

Infatti, possiamo scrivere  $\omega$  come una combinazione lineare dei  $dx_i$  e osservare che

$$(d\omega_x[u])[v] = \sum_{i=1}^n (d\alpha_i(x)[u])dx_i[v] = \sum_{i,j=1}^n \partial_j \alpha_i(x) u_j dx_i[v] = \sum_{i=1}^n \partial_j \alpha_i(x) (dx_j, dx_i)[u, v]$$

Abbiamo ottenuto una forma bilineare<sup>2</sup> che è simmetrica se e solo se vale la condizione di irrotazionalità. Si osservi che in dimensione 3 ciò significa che il rotore di  $X$  è nullo.

Vogliamo studiare la relazione tra forme chiuse ed esatte. Iniziamo con il seguente.

**Teorema 7.1.2.** Una forma  $C^1$  esatta è chiusa.

*Dimostrazione.* Se in un qualche intorno di  $x$   $\omega = df$ , dato che  $\omega$  è di classe  $C^1$  il suo differenziale è continuo e coincide con il differenziale secondo di  $f$ , che è simmetrico per quanto visto in 6.3.  $\square$

Il viceversa non è necessariamente vero. Per esempio, consideriamo su  $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$  la forma definita da  $\omega(x, y)[v] = \langle (-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}), v \rangle$ .  $\omega$  è chiusa, perché

$$\frac{\partial X_x}{\partial y} = -\frac{x^2+y^2-2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{\partial X_y}{\partial x}$$

Tuttavia non è esatta, perché il suo integrale sulla curva chiusa  $\gamma(t) = (\cos t, \sin t)$  non è nullo. Infatti

$$\int_{\gamma} \omega = \int_0^{2\pi} \langle (-\sin t, \cos t), (-\sin t, \cos t) \rangle dt = \int_0^{2\pi} 1 dt = 2\pi$$

Il problema sta nel dominio  $\Omega$ , che ha un "buco" attorno a cui si può girare. Infatti, si ha che

<sup>2</sup>Come forse vedremo, non è una 2-forma.

**Teorema 7.1.3.** *Una forma chiusa su un dominio semplicemente connesso è esatta. In particolare, una forma  $C^1$  è chiusa se e solo se è localmente esatta, perché ogni punto  $x \in \Omega$  ammette un intorno semplicemente connesso (ad esempio, una palla centrata in  $x$  e contenuta in  $\Omega$  è sempre convessa).*

Mostreremo il seguente teorema più generale:<sup>3</sup>

**Teorema 7.1.4.** *Se  $\omega$  è una forma chiusa e  $\gamma, \delta$  sono curve  $C^1$  omotope a estremi fissi, si ha che*

$$\int_{\gamma} \omega = \int_{\delta} \omega$$

Da questa proposizione, in virtù dell'equivalenza 7.1.1, segue il teorema 7.1.3 osservando che in un dominio semplicemente connesso ogni curva chiusa è omotopa a una mappa costante, la quale ha ovviamente integrale nullo.

*Dimostrazione.* Sia  $H(t, s)$  l'omotopia tra  $\gamma = H(\cdot, 0)$  e  $\delta = H(\cdot, 1)$ . Ci restringiamo dapprima considerare curve di classe  $C^2$ . In tal caso non è restrittivo supporre che anche l'omotopia sia di classe  $C^2$ . Questo perché un'omotopia è di fatto un arco in  $C^0([0, 1], C_{\gamma(0), \gamma(1)}^0([0, 1], \Omega))$  che connette  $\gamma$  a  $\delta$  all'interno del sottospazio chiuso delle curve con estremi  $\gamma(0)$  e  $\gamma(1)$ . Dunque, per quanto osservato nel teorema 3.3.3 è possibile connettere  $\gamma$  e  $\delta$  con una spezzata i cui vertici sono curve  $C^2$ , in quanto queste sono dense negli archi continui. Ora, per linearità se i vertici di questa spezzata (compresi il punto iniziale e quello finale) sono curve  $C^2$  allora ogni punto della spezzata lo è, cioè la nuova omotopia è, a ogni tempo fissato, una curva  $C^2$ . Infine, percorrendo la spezzata in modo da annullare derivata prima e seconda in corrispondenza delle giunture è possibile ottenere un'omotopia che sia  $C^2$  anche come mappa sul quadrato.

Sotto le nuove ipotesi, poniamo

$$I(s) = \int_{H(\cdot, s)} \omega = \int_0^1 \omega(H(t, s))[\partial_t H(t, s)] dt$$

Il nostro obiettivo è far vedere che  $I(s)$  è costante. Osserviamo che, per 6.13.1,  $I$  è una funzione  $C^1$  la cui derivata è

$$I'(s) = \int_0^1 \partial_s (\omega(H(t, s))[\partial_t H(t, s)]) dt = \int_0^1 (d\omega(H(t, s))[\partial_s H(t, s), \partial_t H(t, s)] + \omega(H(t, s))[\partial_s \partial_t H(t, s)]) dt$$

Poiché  $\omega$  è chiusa è possibile invertire gli argomenti del suo differenziale. Inoltre, dato che  $H(t, s)$  è di classe  $C^2$  le sue derivate parziali commutano, per teorema di Schwarz. Dunque possiamo ricomporre l'integrando come

$$I'(s) = \int_0^1 (d\omega(H(t, s))[\partial_t H(t, s), \partial_s H(t, s)] + \omega(H(t, s))[\partial_t \partial_s H(t, s)]) dt$$

$$I'(s) = \int_0^1 \partial_t (\omega(H(t, s))[\partial_s H(t, s)]) dt = \omega(H(1, s))[\partial_s H(1, s)] - \omega(H(0, s))[\partial_s H(0, s)]$$

Ma la RHS è nulla perché  $H(0, s) = \text{cost}$  e  $H(1, s) = \text{cost}$ , essendo  $\gamma$  e  $\delta$  omotope a estremi fissi mediante  $H$ . Dunque  $I' \equiv 0$ , perciò i due integrali sono uguali.

Resta ora da mostrare come estendere il risultato a curve che non sono di classe  $C^2$ . Sia  $\gamma$  una curva  $C^1$  a tratti e sia  $\varepsilon > 0$  tale che l'insieme  $B = \{x \in \Omega \text{ t.c. } d(x, \gamma([0, 1])) < \varepsilon\}$  sia totalmente contenuto in  $\Omega$ , che esiste per quanto visto nella dimostrazione di 3.3.3. L'insieme delle curve a valori in  $B$  e di estremi  $\gamma(0)$  e  $\gamma(1)$  è dunque un insieme stellato da  $\gamma$ , e perciò tutti i suoi elementi sono omotopi a estremi fissi. Sia allora  $\gamma_n$  una successione di curve in  $C_{\gamma(0), \gamma(1)}^2([0, 1], B)$  che converge in norma  $C^1$  a  $\gamma$ . Una tale successione esiste per il teorema 5.2.4.1. Inoltre essendo tutte omotope a estremi fissi per quanto visto prima hanno tutte lo stesso integrale, il quale coincide con l'integrale su  $\gamma$  per il teorema di convergenza 7.0.1. Basta ora osservare che, ripetendo le considerazioni per  $\delta$  e scegliendo una qualsiasi coppia  $\gamma_n, \delta_n$ , l'omotopia tra  $\gamma$  e  $\delta$  si estende per concatenazione a un'omotopia tra  $\gamma_n$  e  $\delta_n$ , che dunque hanno lo stesso integrale essendo  $C^2$ .  $\square$

<sup>3</sup>L'intera trattazione è stata affrontata in maniera più semplice nel corso di Geometria 2, sebbene in quel frangente ci si sia ristretti a considerare curve a valori in  $\mathbb{R}^2$ . Rimando a [Vergamini-Marchetti-Traino, 2020].

Dal procedimento seguito emerge la possibilità di estendere la definizione di integrale di una 1-forma chiusa a un qualsiasi cammino continuo  $\gamma$ , ponendo

$$\int_{\gamma} \omega = \int_{\gamma_n} \omega$$

dove  $\gamma_n$  è una qualsiasi curva  $C^2$  a valori nell'intorno  $B$  definito nella dimostrazione. Si può allora concludere che

**Corollario 7.1.4.1.** *L'integrale (generalizzato) di una 1-forma chiusa sui cammini continui è invariante per omotopie a estremi fissi.*

### 7.1.1 Costruzione di una primitiva per aperti stellati

**Teorema 7.1.5.** *Sia  $\Omega \subseteq E$  un aperto stellato rispetto al punto  $x_0$ . Se  $\omega$  è chiusa una sua primitiva è la funzione*

$$f(x) = \int_0^1 \omega(tx + (1-t)x_0)[x] dt$$

cioè è l'integrale di  $\omega$  lungo il cammino rettilineo da  $x_0$  a  $x$ .

*Dimostrazione.* WLOG assumiamo che  $x_0 = 0$  e poniamo  $g(x, t) = \omega(tx)[x]$ . Osserviamo che, per chain rule applicata alla funzione di valutazione  $g(x, t) = \text{eva}(\omega(tx), x)$ , si ha

$$d_1 g_{(x,t)}[u] = t d\omega(tx)[u, x] + \omega(tx)[u]$$

Poiché la forma è chiusa, posso invertire gli argomenti del differenziale e ottenere

$$d_1 g_{(x,t)}[u] = t d\omega(tx)[x, u] + \omega(tx)[u] = \partial_t(t\omega(tx)[u])$$

Adesso, per il teorema di differenziale sotto il segno di integrale, che abbiamo visto in 6.13.1, possiamo scrivere

$$df_x[u] = \int_0^1 d_1 g_{(x,t)}[u] dt = \int_0^1 \partial_t(t\omega(tx)[u]) dt = \omega(x)[u]$$

da cui la tesi. □

## 7.2 Pullback di forme differenziali

**Definizione 32.** *Sia  $f : \Omega_0 \subseteq E_0 \rightarrow \Omega_1 \subseteq E_1$  una funzione  $C^1$  tra aperti di spazi di Banach e sia  $\omega$  una 1-forma su  $\Omega_1$ . Il suo pullback mediante  $f$  è la 1-forma su  $\Omega_0$  definita come*

$$f^* \omega(x)[v] = \omega(f(x))[df_x(v)]$$

Il pullback gode delle seguenti proprietà:

1. La funzione  $f^* : C^0(\Omega_1, E_1^*) \rightarrow C^0(\Omega_0, E_0^*)$  è lineare in modo ovvio.
2. Se  $f \in C^1(\Omega_0, \Omega_1)$  e  $g \in C^1(\Omega_1, \Omega_2)$ , allora  $(g \circ f)^* = f^* \circ g^*$ . Infatti

$$(g \circ f)^* \omega(x)[v] = \omega(g(f(x)))[d(g \circ f)_x(v)] = \omega(g(f(x)))[dg_{f(x)}(df_x(v))]$$

$$(g \circ f)^* \omega(x)[v] = g^* \omega(f(x))[df_x(v)] = f^* \circ g^* \omega(x)[v]$$

3. Se  $\omega$  è esatta lo è anche  $f^* \omega$ . Infatti

$$f^* dh_x[v] = dh(f(x))[df_x(v)] = d(h \circ f)_x[v]$$

4. Se  $\omega$  è chiusa lo è anche  $f^* \omega$ . Questo perché una forma  $C^1$  è chiusa se e solo se è localmente esatta, e per il punto precedente la locale esattezza si trasporta tramite pullback.



# Capitolo 8

## Teoria della misura

Ci apprestiamo ora allo studio di una branca della matematica che si prefigge di sistemare i fondamenti di una vasta gamma di concetti fisici, come il volume e la massa, e che ha strettissimi legami con la teoria assiomatica della probabilità descritta da Kolmogorov.

Attenzione: in tutti i libri inglesi si fa riferimento ai risultati di teoria della misura con l'aggettivo *metric* (MEasure TheoRetIC) per il pessimo gusto di Lebesgue e amici, e ciò porta (più spesso di quanto si possa credere) a non sempre banali ambiguità tra spazi metrici e *metric spaces* (spazi di misura). Eviterò questa notazione perché aborro la stupidità.

### 8.1 Algebre e anelli

**Definizione 33** (Algebra). *Un sottoinsieme  $\mathcal{A}$  delle parti di un insieme  $X$  è un'algebra se*

- $\emptyset \in \mathcal{A}$ ;
- se  $A \in \mathcal{A}$ , anche  $A^C = X \setminus A \in \mathcal{A}$ ;
- se una famiglia finita  $\{A_k\}_{k=0, \dots, n} \subseteq \mathcal{A}$ , allora  $\bigcup_{k=0, \dots, n} A_k \in \mathcal{A}$ .

**Definizione 34** ( $\sigma$ -algebra). *Un'algebra  $\mathcal{A}$  di parti di  $X$  è una  $\sigma$ -algebra se è chiusa per unioni numerabili.*

Si mostra facilmente che dalle proprietà di  $\mathcal{A}$  segue la chiusura per differenza insiemistica  $A \setminus B$  e per intersezioni numerabili, sfruttando le formule di de Morgan  $((\bigcup A_n)^C = \bigcap A_n^C)$ .

**Definizione 35** (Spazio misurabile). *Un insieme  $X$  munito di una  $\sigma$ -algebra si dice spazio misurabile.*

Si osservi che l'intersezione di  $\sigma$ -algebre è ancora una  $\sigma$ -algebra, perché ovviamente contiene  $X$  e se degli insiemi  $A_n$  stanno in ogni  $\sigma$ -algebra anche i complementari e l'unione sono contenuti in ogni  $\sigma$ -algebra. Inoltre, l'insieme delle parti è una  $\sigma$ -algebra, per ovvi motivi, e dunque dato un sottoinsieme qualsiasi  $S \subseteq \mathcal{P}(X)$  esiste una  $\sigma$ -algebra che lo contiene. Da questi due fatti segue che la seguente definizione è ben posta:

**Definizione 36.** *Dato un sottoinsieme delle parti  $S \subseteq \mathcal{P}(X)$ , la  $\sigma$ -algebra generata da  $S$  è*

$$\sigma(S) = \bigcap_{S \subseteq \mathcal{A} \subseteq \mathcal{P}(X) \text{ } \sigma\text{-algebra}} \mathcal{A}$$

Dunque, se  $(X, \tau)$  è uno spazio topologico è possibile munirlo della  $\sigma$ -algebra  $\mathbb{B}$  generata dagli aperti, detta  $\sigma$ -algebra di Borel o dei boreliani.

### 8.2 Misura

**Definizione 37** (Misura). *Sia  $X$  un insieme e sia  $\mathcal{A} \subseteq \mathcal{P}(X)$  una  $\sigma$ -algebra. Una funzione di sottoinsiemi  $\mu : \mathcal{A} \rightarrow [0, +\infty]^1$  si dice misura se  $\mu(\emptyset) = 0$  ed è  $\sigma$ -additiva, ossia, data una famiglia numerabile di*

<sup>1</sup>Sebbene sia possibile definire misure a valori in tutta la retta estesa, a meno di specificarlo considereremo sempre misure a valori in  $[0, +\infty]$ , anche e soprattutto perché non vogliamo preoccuparci delle condizioni di convergenza di sommatorie a segni casuali.

insiemi disgiunti  $A_n \in \mathcal{A}$ , la loro unione  $\sqcup_{n \in \mathbb{N}} A_n$  appartiene a  $\mathcal{A}$  e vale che

$$\mu(\sqcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$$

**Definizione 38** (Probabilità). *Una misura non negativa che associa 1 a tutto lo spazio è detta probabilità.*

Si osservi che la definizione di misura estende il concetto di sommatoria generalizzata visto al primo anno. Infatti, dato un insieme  $X$  e una famiglia di numeri reali non negativi  $\{c_x\}_{x \in X}$ , la funzione  $\mu(A) = \sum_{x \in A} c_x$  definita su  $\mathcal{P}(X)$  è una misura, perché la  $\sigma$ -additività è garantita dalla possibilità di riordinare le sommatorie. In particolare, se tutti i  $c_x$  valgono 1, la misura “conterà i punti”, cioè assocerà a un insieme finito la sua cardinalità e a un insieme infinito il valore  $+\infty$ .

Un altro possibile esempio è la  $\delta$  di Dirac del punto  $x \in X$ , cioè la misura

$$\delta_x(A) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

**Definizione 39** (Spazio misurabile). *Un insieme  $X$  munito di una  $\sigma$ -algebra  $\mathcal{A}$  e di una misura  $\mu$  su  $\mathcal{A}$  si dice spazio di misura. Se  $\mu$  è di probabilità,  $X$  si dice misura di probabilità.*

Faremo spesso uso di questa notazione.

**Definizione 40** (Insieme trascurabile). *Un insieme  $A \in \mathcal{A}$  si dice  $\mu$ -trascurabile se ha misura nulla. Si dice che il complementare di un insieme trascurabile ha misura piena.*

**Definizione 41.** *Una proprietà su uno spazio di misura  $(X, \mathcal{A}, \mu)$  è valida  $\mu$ -quasi ovunque (indicato con  $\mu$ -q.o.) se è vera per tutti gli  $x$  in un insieme di misura piena.*

Una misura gode delle seguenti proprietà di regolarità nel passaggio al limite:

- Se  $A_n \subseteq A_{n+1}$ , allora  $\mu(A_n) \uparrow \mu(\bigcup_{n \in \mathbb{N}} A_n)$   
Basta considerare la successione  $B_0 = A_0$  e  $B_n = A_n \setminus \bigcup_{k < n} A_k$  e utilizzare la  $\sigma$ -additività.
- Se  $A_{n+1} \subseteq A_n$  e definitivamente  $\mu(A_n) < \infty$ , allora  $\mu(A_n) \downarrow \mu(\bigcap_{n \in \mathbb{N}} A_n)$   
Infatti, supponendo che  $\mu(A_0) < \infty$  a meno di riscalarare gli indici, abbiamo che

$$\mu(A_n) = \mu(A_0) - \mu(A_0 \setminus A_n) \downarrow \mu(A_0) - \mu\left(A_0 \setminus \bigcap_{n \in \mathbb{N}} A_n\right)$$

E la tesi segue semplificando. Questa cosa non è legale se  $\mu(A_0) = \infty$ , e infatti esistono esempi di insiemi a misura infinita la cui intersezione ha misura nulla (ad esempio, le semirette  $[n, +\infty[$  hanno lunghezza infinita ma la loro intersezione è vuota).

### 8.3 Misura esterna ed estensione di misure

**Definizione 42** (Misura esterna). *Sia  $X$  un insieme e sia  $S \subseteq \mathcal{P}(X)$  una famiglia di parti. Una funzione di sottoinsiemi  $\phi : S \rightarrow [0, +\infty]$  si dice  $\sigma$ -subadditiva se, per ogni  $A \in S$  e per ogni famiglia numerabile  $(F_n)_{n \in \mathbb{N}} \subseteq S$  che ricopre  $A$  vale*

$$\phi(A) \leq \sum_{n \in \mathbb{N}} \phi(F_n)$$

Una funzione  $\sigma$ -subadditiva su  $\mathcal{P}(X)$  si dice misura esterna.

Data ora una famiglia di parti  $S \subseteq \mathcal{P}(X)$  e una funzione  $\phi : S \rightarrow [0, +\infty]$ , definiamo la misura esterna indotta  $\phi^*$  come la massima misura esterna su  $\mathcal{P}(X)$  che sia minore di  $\phi$  sugli insiemi di  $S$ .

Mostriamo che la definizione è ben posta, e cioè che il massimo è effettivamente realizzato. Osserviamo che, se  $\psi$  è una misura esterna con la proprietà richiesta, per ogni insieme  $A \subseteq X$  e per ogni ricoprimento  $A$  con una famiglia numerabile  $(F_n)_{n \in \mathbb{N}} \subseteq S$  si ha

$$\psi(A) \leq \sum_{n \in \mathbb{N}} \psi(F_n) \leq \sum_{n \in \mathbb{N}} \phi(F_n)$$

Dunque

$$\psi(A) \leq \inf \left\{ \sum_{n \in \mathbb{N}} \phi(F_n) \mid F_n \in S \text{ ricoprimento numerabile di } A \right\}$$

Vogliamo far vedere che  $\phi^*(A)$  è il membro di destra della disuguaglianza precedente, nella quale si intende che l'estremo inferiore di un insieme vuoto vale  $+\infty$ . Per farlo dobbiamo mostrare la  $\sigma$ -subadditività, ossia che per ogni  $A \subseteq X$  e per ogni ricoprimento numerabile di  $A$   $(B_n)_{n \in \mathbb{N}} \subseteq \mathcal{P}(X)$  vale che  $\phi^*(A) \leq \sup_{n \in \mathbb{N}} \phi^*(B_n)$ . Se  $\phi^*(B_n) = +\infty$  per un certo  $n$  la tesi è banale. Altrimenti, fissato un  $\varepsilon > 0$ , per ogni  $B_n$  esiste un suo ricoprimento numerabile  $\mathcal{F}_n \subseteq S$  tale che  $\sum_{F \in \mathcal{F}_n} \phi(F) \leq \phi^*(B_n) + \varepsilon/2^n$ . Unendo gli  $\mathcal{F}_n$  otteniamo un ricoprimento di  $A$  con insiemi di  $S$ , per cui

$$\phi^*(A) \leq \sum_{n \in \mathbb{N}} \sum_{F \in \mathcal{F}_n} \phi(F) \leq \varepsilon + \sum_{n \in \mathbb{N}} \phi^*(B_n)$$

La tesi segue per arbitrarietà di  $\varepsilon$ .

Possiamo allora dare la seguente definizione:

**Definizione 43** (Misura esterna indotta). *Sia  $\phi$  una funzione di insieme positiva definita su una famiglia di parti  $S$ . La misura esterna generata da  $\phi$  è la funzione*

$$\phi^*(A) = \inf \left\{ \sum_{n \in \mathbb{N}} \phi(F_n) \mid F_n \in S \text{ ricoprimento numerabile di } A \right\}$$

Si osservi che la misura esterna coincide con  $\phi$  su  $S$  se e solo se  $\phi$  è già  $\sigma$ -subadditiva su  $S$ . Infatti, un insieme  $F \in S$  è ovviamente un ricoprimento di se stesso, e per ogni altro ricoprimento numerabile  $(F_n)_{n \in \mathbb{N}} \subseteq S$  vale  $\phi(F) \leq \sum_{n \in \mathbb{N}} \phi(F_n)$ .

Vedremo ora che una misura esterna si riesce sempre a restringere a una misura su una qualche  $\sigma$ -algebra. Premettiamo qualche definizione.

**Definizione 44.** *Sia  $\mu^*$  una misura esterna su  $\mathcal{P}(X)$ . Un insieme  $A \subseteq X$  si dice misurabile secondo Carathéodory se per ogni  $E \subseteq X$  vale*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^C)$$

Equivalentemente,  $A$  è misurabile se vale la disuguaglianza

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^C)$$

dato che l'altra disuguaglianza è sempre vera per subadditività. La condizione si può ulteriormente indebolire se  $\mu^*$  è la misura indotta da una funzione subadditiva.

**Lemma 8.3.1.** *Sia  $\phi : S \subseteq \mathcal{P}(X) \rightarrow [0, +\infty]$  una funzione  $\sigma$ -subadditiva e sia  $\phi^*$  la misura esterna indotta. Un insieme  $A \subseteq X$  è  $\phi^*$ -misurabile se e solo se per ogni  $F \in S$  vale*

$$\phi^*(F) \geq \phi^*(F \cap A) + \phi^*(F \cap A^C)$$

*Dimostrazione.* Consideriamo solo la freccia non ovvia. Sia  $E \subseteq X$ . Se  $\phi^*(E) = +\infty$  non c'è nulla di cui discutere, altrimenti per ogni  $n \in \mathbb{N}$  scegliamo un ricoprimento  $(F_k^n)_{k \in \mathbb{N}} \subseteq S$  di  $E$  tale che  $\sum_{k \in \mathbb{N}} \phi(F_k^n) \leq \phi^*(E) + 1/n$ . Per ipotesi possiamo scomporre i termini della somma intersecando con  $A$  e  $A^C$ . Dunque per subadditività

$$\phi^*(E) + 1/n \geq \sum_{k \in \mathbb{N}} \phi(F_k^n) = \sum_{k \in \mathbb{N}} \phi(F_k^n \cap A) + \sum_{k \in \mathbb{N}} \phi(F_k^n \cap A^C) \geq \phi^* \left( \bigcup_{k \in \mathbb{N}} F_k^n \cap A \right) + \phi^* \left( \bigcup_{k \in \mathbb{N}} F_k^n \cap A^C \right)$$

$$\phi^*(E) + 1/n \geq \phi^*(E \cap A) + \phi^*(E \cap A^C)$$

Da cui la tesi passando al limite su  $n$ . □

**Definizione 45** (Misura completa). *Una misura su una  $\sigma$ -algebra  $\mathcal{A}$  si dice completa se ogni sottoinsieme di un insieme trascurabile è trascurabile.*

La definizione è meno banale di quanto appaia: il vero problema non è tanto il valore nullo della misura, il quale segue facilmente dalla subadditività, quanto il fatto che il sottoinsieme potrebbe non appartenere alla  $\sigma$ -algebra.

**Teorema 8.3.2** (Carathéodory). *Sia  $\mu^*$  una misura esterna nulla sull'insieme vuoto. La famiglia  $\overline{\mathcal{A}}$  degli insiemi misurabili secondo Carathéodory è una  $\sigma$ -algebra, ristretta alla quale  $\mu^*$  è una misura completa.*

*Dimostrazione.*  $\overline{\mathcal{A}}$  è ovviamente chiusa per complementare; inoltre per ipotesi l'insieme vuoto vi appartiene. Perché sia una  $\sigma$ -algebra bisogna far vedere che contiene tutte le unioni numerabili di suoi elementi. Mostriamo innanzitutto la chiusura per unione finita. Per ogni  $E \in \mathcal{P}(X)$  e per ogni coppia  $A, B \in \overline{\mathcal{A}}$ , per misurabilità di  $A$  abbiamo che

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap (A \cup B) \cap A) + \mu^*(E \cap (A \cup B) \cap A^C) = \mu^*(E \cap A) + \mu^*(E \cap A^C \cap B)$$

Inoltre

$$\mu^*(E \cap (A \cup B)^C) = \mu^*(E \cap A^C \cap B^C)$$

Sommando e applicando progressivamente la misurabilità di  $B$  e di  $A$  risulta

$$\mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^C) = \mu^*(E \cap A) + \mu^*(E \cap A^C) = \mu^*(E)$$

Basta ora far vedere che  $\mathcal{A}$  contiene le unioni numerabili disgiunte. Infatti, l'unione di una qualsiasi famiglia numerabile  $(A_n)_{n \in \mathbb{N}}$  si riconduce all'unione disgiunta degli insiemi  $A_n \setminus \bigcup_{k < n} A_k$ , che si trovano in  $\overline{\mathcal{A}}$  in quanto questa è chiusa per unioni finite e complementari (e dunque anche per differenza).

Sia allora  $(A_n)_{n \in \mathbb{N}}$  una famiglia numerabile disgiunta di insiemi di  $\overline{\mathcal{A}}$ . Per ogni  $E \subseteq X$  e per ogni  $n \in \mathbb{N}$  vale

$$\mu^*(E) = \mu^*\left(E \cap \bigcup_{k=0}^n A_k\right) + \mu^*\left(E \cap \left(\bigcup_{k=0}^n A_k\right)^C\right) \geq \mu^*\left(E \cap \bigcup_{k=0}^n A_k\right) + \mu^*\left(E \cap \left(\bigcup_{k=0}^{\infty} A_k\right)^C\right)$$

Per misurabilità di  $A_n$  il primo addendo si può spezzare come

$$\mu^*\left(E \cap \bigcup_{k=0}^n A_k\right) = \mu^*\left(E \cap \bigcup_{k=0}^n A_k \cap A_n\right) + \mu^*\left(E \cap \bigcup_{k=0}^n A_k \cap A_n^C\right) = \mu^*(E \cap A_n) + \mu^*\left(E \cap \bigcup_{k=0}^{n-1} A_k\right)$$

Procedendo induttivamente otteniamo allora

$$\mu^*\left(E \cap \bigcup_{k=0}^n A_k\right) = \sum_{k=0}^n \mu^*(E \cap A_k)$$

Sostituendo nella disuguaglianza originaria e passando all'estremo superiore su  $n$  otteniamo quanto cercato:

$$\mu^*(E) \geq \sum_{k=0}^{\infty} \mu^*(E \cap A_k) + \mu^*\left(E \cap \left(\bigcup_{k=0}^{\infty} A_k\right)^C\right) \geq \mu^*\left(E \cap \bigcup_{k=0}^{\infty} A_k\right) + \mu^*\left(E \cap \left(\bigcup_{k=0}^{\infty} A_k\right)^C\right)$$

Ora,  $\mu^*$  è una misura su  $\overline{\mathcal{A}}$ . Infatti, data una famiglia numerabile disgiunta  $(A_n)_{n \in \mathbb{N}}$ , basta porre  $E = \bigcup_{n \in \mathbb{N}} A_n$  e osservare che, per quanto si ricava dal punto precedente

$$\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{k \in \mathbb{N}} \mu^*\left(\bigcup_{n \in \mathbb{N}} A_n \cap A_k\right) + \mu^*\left(\bigcup_{n \in \mathbb{N}} A_n \cap \left(\bigcup_{n \in \mathbb{N}} A_n\right)^C\right) = \sum_{k \in \mathbb{N}} \mu^*(A_k)$$

Per mostrare che  $\mu^*$  è completa bisogna far vedere che se  $N \in \overline{\mathcal{A}}$  è trascurabile ogni suo sottoinsieme  $F$  sta in  $\overline{\mathcal{A}}$  e ha misura nulla. Ora, per subadditività la misura esterna di  $F$  è zero. Da ciò segue anche che  $F$  è misurabile secondo Carathéodory, perché per ogni  $E \subseteq X$  abbiamo

$$\mu^*(E) \geq \mu^*(E \cap F^C) = \mu^*(E \cap F^C) + \mu^*(E \cap F)$$

Entrambi i passaggi sono conseguenze della  $\sigma$ -subadditività. □

Mettendo insieme i risultati visti finora, abbiamo mostrato il seguente teorema:

**Teorema 8.3.3.** *Sia  $\phi : S \subseteq \mathcal{P}(X) \rightarrow [0, +\infty]$  una funzione  $\sigma$ -subadditiva e sia  $\phi^*$  la misura esterna indotta. Se per ogni  $A, E \in S$  vale  $\phi(E) = \phi^*(E \cap A) + \phi^*(E \cap A^C)$ , allora  $\phi^*$  estende  $\phi$  ed esiste una  $\sigma$ -algebra  $\overline{A} \supseteq \sigma(S)$  su cui  $\phi^*$  è una misura completa. In particolare,  $\phi$  è  $\sigma$ -additiva, perché si estende a una misura.*

La condizione  $\phi(E) = \phi^*(E \cap A) + \phi^*(E \cap A^C)$  serve ad assicurarci che la  $\sigma$ -algebra generata nel processo estenda l'insieme di definizione di  $\phi$ . Mostriamo ora che, quando  $\phi$  è  $\sigma$ -additiva, per particolari classi di insiemi la richiesta è sempre soddisfatta.

**Definizione 46** (Quasi-anello). *Una famiglia di parti  $S$  si dice quasi-anello se  $\forall A, B \in S$ ,  $A \cap B$  e  $A \setminus B$  si possono scrivere come unione numerabile disgiunta di insiemi di  $S$ .*

**Definizione 47** (Anello). *Una famiglia di parti  $S$  si dice anello se  $\forall A, B \in S$ ,  $A \cap B$  e  $A \setminus B$  stanno in  $S$ .*

**Teorema 8.3.4.** *Una funzione di insiemi  $\phi$  non negativa e  $\sigma$ -additiva su un quasi-anello  $S$  si estende a una misura su  $\sigma(S)$ .*

*Dimostrazione.* Ai fini della dimostrazione è sufficiente mostrare che le ipotesi del teorema 8.3.3 sono soddisfatte. Innanzitutto verifichiamo che una funzione non negativa e  $\sigma$ -additiva su un quasi-anello è anche  $\sigma$ -subadditiva. Sia  $A \subseteq \bigcup_{n \in \mathbb{N}} F_n$ , con  $A, F_n$  appartenenti a  $S$ . Vogliamo mostrare che  $\phi(A)$  è maggiorata dalla somma dei  $\phi(F_n)$ . Possiamo supporre che gli  $F_n$  siano disgiunti. Infatti, possiamo sostituirli con

$$\begin{cases} G_0 = F_0 \\ G_n = F_n \setminus \bigcup_{k < n} F_k = F_n \setminus F_1 \setminus \dots \setminus F_{n-1} \end{cases}$$

A loro volta i  $G_n$  si possono sostituire con delle unioni numerabili disgiunte di elementi di  $S$ , per le proprietà di quasi-anello. Ora, per ogni  $n$  esistono degli  $H_{n,k} \in S$  per cui  $F_n \cap A = \bigsqcup_{k \in \mathbb{N}} H_{n,k}$ . Inoltre esistono dei  $J_{n,k} \in S$  per cui  $F_n \setminus A = \bigsqcup_{k \in \mathbb{N}} J_{n,k}$ . Dunque per  $\sigma$ -additività

$$\mu(A) = \mu(\bigsqcup_{n \in \mathbb{N}} F_n \cap A) = \sum_{n,k \in \mathbb{N}} \mu(H_{n,k}) \leq \sum_{n,k \in \mathbb{N}} \mu(H_{n,k}) + \sum_{n,k \in \mathbb{N}} \mu(J_{n,k}) \leq \sum_{n \in \mathbb{N}} \mu(F_n)$$

Vediamo ora che ogni elemento di  $S$  verifica le ipotesi del lemma 8.3.1. Dati dunque due insiemi  $A, B \in S$ , possiamo scrivere  $B \cup A$  e  $B \setminus A$  come unioni numerabili disgiunte di elementi di  $S$ , che indicheremo rispettivamente con  $F_n$  e  $G_n$ . Allora, sfruttando la  $\sigma$ -subadditività di  $\phi^*$ , otteniamo

$$\phi^*(B \cap A) + \phi^*(B \cap A^C) = \phi^*(B \cap A) + \phi^*(B \setminus A) \leq \sum_{n \in \mathbb{N}} \phi(F_n) + \sum_{n \in \mathbb{N}} \phi(G_n)$$

Poiché gli addendi sono tutti disgiunti tra loro e  $\phi$  è  $\sigma$ -additiva possiamo scrivere

$$\phi^*(B \cap A) + \phi^*(B \cap A^C) \leq \sum_{n \in \mathbb{N}} \phi(F_n) + \sum_{n \in \mathbb{N}} \phi(G_n) = \phi\left(\bigcup_{n \in \mathbb{N}} F_n \cup \bigcup_{n \in \mathbb{N}} G_n\right) = \phi(B)$$

□

Poiché una  $\sigma$ -algebra è in particolare un quasi-anello, la dimostrazione precedente ci dice anche che

**Corollario 8.3.4.1.** *Una misura è  $\sigma$ -subadditiva.*

### 8.3.1 Completamento di misure

Il prossimo obiettivo è mostrare che una misura soddisfa le ipotesi del teorema 8.3.3. Premettiamo una definizione.

**Definizione 48.** *Sia  $\mu$  una misura. Un insieme  $E \subseteq X$  è finito rispetto a  $\mu$  se la sua misura è finita, ed è  $\sigma$ -finito rispetto a  $\mu$  se è unione numerabile di insiemi finiti.  $\mu$  si dice ( $\sigma$ -)finita se lo è lo spazio  $X$ .*

**Lemma 8.3.5.** *Se  $\mu$  è una misura sulla  $\sigma$ -algebra  $\mathcal{A}$ , allora*

$$\mu^*(A) = \inf \{ \mu(F) \mid A \subseteq F \in \mathcal{A} \}$$

*ed esiste un "coperchio" misurabile  $F = A^*$  che realizza l'estremo inferiore. Inoltre, se  $\mu$  è  $\sigma$ -finita e  $A$  è misurabile secondo Carathéodory è possibile scegliere il coperchio in modo che  $\mu^*(A^* \setminus A) = 0$ .*

*Dimostrazione.* Dato un qualsiasi ricoprimento numerabile  $\mathcal{F} \subseteq \mathcal{A}$  di  $A$ , l'unione  $\bigcup_{F \in \mathcal{F}} F$  appartiene all'algebra, ricopre  $A$  e ha misura minore o uguale alla somma delle misure degli  $F$ . Dunque otteniamo  $\mu^*(A) \geq \inf \{ \mu(F) \mid A \subseteq F \in \mathcal{A} \}$ . Il viceversa è ovvio.

Per mostrare l'esistenza di  $A^*$ , consideriamo una successione  $F_n \supseteq A$  tale che  $\mu(F_n) \downarrow \mu^*(A)$ . A meno di intersecarli assumiamo che gli  $F_n$  siano innestati. Poniamo  $A^* = \bigcap_{n \in \mathbb{N}} F_n$  e osserviamo che, essendo questo un sottoinsieme di tutti gli  $F_n$ ,

$$\mu(A^*) \leq \inf_{n \in \mathbb{N}} \mu(F_n) = \mu^*(A)$$

Dato che  $A \subseteq A^* \in \mathcal{A}$  vale anche l'altra disuguaglianza, da cui la tesi.

Ora, dato un insieme  $A \subseteq X$ , se  $\mu^*(A) < +\infty$  ovviamente  $\mu^*(A^* \setminus A) = 0$ , in quanto essendo  $A$  misurabile abbiamo

$$\mu(A^*) = \mu^*(A) + \mu^*(A^* \setminus A)$$

In caso contrario sia  $(Y_n)_{n \in \mathbb{N}}$  una partizione di  $X$  composta da insiemi di misura finita. Possiamo scrivere  $A = \sqcup (Y_n \cap A)$  e per ogni elemento della somma disgiunta possiamo trovare un coperchio di misura finita  $A_n^* \subseteq Y_n$  tale che  $\mu(A_n^* \setminus (Y_n \cap A)) = 0$ , all'interno della  $\sigma$ -algebra ristretta  $\mathcal{A}_n = \{Y_n \cap B, B \in \mathcal{A}\}$ . Definiamo dunque  $A^* = \sqcup S_n^*$ , che per costruzione sta in  $\mathcal{A}$ , e osserviamo che

$$\mu^*(A^* \setminus A) \leq \sum_{n \in \mathbb{N}} \mu(A_n^* \setminus (Y_n \cap A)) = 0$$

□

**Teorema 8.3.6.** *Data una misura  $\mu$  su una  $\sigma$ -algebra  $\mathcal{A}$ , esiste una  $\sigma$ -algebra  $\overline{\mathcal{A}} \supseteq \mathcal{A}$  e una misura completa  $\overline{\mu}$  su  $\overline{\mathcal{A}}$  che estende  $\mu$ .*

*Dimostrazione.* Ci basta far vedere che gli elementi di  $\mathcal{A}$  sono  $\mu^*$ -misurabili secondo Carathéodory. Siano  $A \in \mathcal{A}$  e  $E \subseteq X$  e sia  $E \subseteq E^* \in \mathcal{A}$  che realizza la misura esterna. Allora, sfruttando l'additività di  $\mu$  e la coincidenza di questa con  $\mu^*$  su  $\mathcal{A}$ , otteniamo

$$\mu^*(E) = \mu(E^*) = \mu(E^* \cap A) + \mu(E^* \cap A^C) = \mu^*(E^* \cap A) + \mu^*(E^* \cap A^C) \geq \mu^*(E \cap A) + \mu^*(E \cap A^C)$$

□

Per certe misure è possibile descrivere più esplicitamente questo completamento (che tuttavia potrebbe non essere l'unico possibile).

**Teorema 8.3.7.** *Se  $\mu$  è una misura  $\sigma$ -finita, la  $\sigma$ -algebra  $\overline{\mathcal{A}}$  coincide con la famiglia*

$$\mathcal{B} = \{A \setminus N \text{ t.c. } A \in \mathcal{A}, N \subseteq P \in \mathcal{A}, \mu(P) = 0\}$$

*E vale che  $\overline{\mu}(A \setminus N) = \mu(A)$ .*

*Dimostrazione.* Ovviamente  $\mathcal{B} \subseteq \overline{\mathcal{A}}$ , perché  $\overline{\mu}$  è completa su  $\overline{\mathcal{A}}$  e dunque quest'ultima deve contenere tutti i sottoinsiemi degli insiemi  $\mu$ -trascurabili. Viceversa, sia  $E \in \overline{\mathcal{A}}$ . Possiamo considerare un coperchio  $E^* \in \mathcal{A}$  per cui  $\mu^*(E^* \setminus E) = 0$ . Dunque  $E^* \setminus E$  è un sottoinsieme di un certo coperchio  $F \in \mathcal{A}$  avente misura nulla, dunque  $E = E^* \setminus (E^* \setminus E) \in \mathcal{B}$ .

Ora,  $\overline{\mu}(A \setminus N) = \mu^*(A \setminus N) \leq \mu(A)$  perché  $A$  lo ricopre e sta in  $\mathcal{A}$ . D'altra parte per subadditività  $\mu(A) \leq \mu^*(A \setminus N) + \mu^*(A \cap N) = \mu^*(A \setminus N)$ , dato che  $N$  ha un coperchio di misura nulla. Segue allora l'uguaglianza cercata. □

In generale, si potrebbe dimostrare separatamente che, data una qualsiasi misura, la famiglia  $\mathcal{B}$  costruita nel teorema è una  $\sigma$ -algebra completa contenente tutti gli insiemi  $\sigma$ -finiti, che non per forza coincide con il completamento costruito con il teorema di Carathéodory.

## 8.4 Misura prodotto

Siano  $(X, \mathcal{A}, \mu)$  e  $(Y, \mathcal{B}, \nu)$  spazi di misura. Sul prodotto  $X \times Y$  consideriamo la  $\sigma$ -algebra  $\mathcal{A} \otimes \mathcal{B}$  generata dalla famiglia  $S = \{A \times B \text{ t.c. } A \in \mathcal{A}, B \in \mathcal{B}\}$ .<sup>2</sup> Definiamo su  $S$  la misura prodotto

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$$

**Teorema 8.4.1.** *La misura prodotto si estende a  $\mathcal{A} \otimes \mathcal{B}$ .*

*Dimostrazione.* Vogliamo applicare il teorema 8.3.4. Mostriamo innanzitutto che  $S$  è un quasi-anello. Dati  $A \times B$  e  $A' \times B'$  in  $S$ , la loro intersezione è  $(A \cap A') \times (B \cap B')$ , che appartiene a  $S$ , mentre la loro differenza è data da

$$A \times B \setminus A' \times B' = \{(a, b) \in A \times B \text{ t.c. } a \notin A' \vee b \notin B'\} = (A \setminus A') \times B \sqcup (A \cap A') \times (B \setminus B')$$

Che è un'unione disgiunta di insiemi di  $S$ .

Resta da dimostrare che  $\mu \otimes \nu$  è  $\sigma$ -additiva su  $S$ . Per ogni  $x \in X$  possiamo considerare l'inclusione  $j_x : Y \rightarrow X \times Y$  tale che  $j_x(y) = (x, y)$ . Tale immersione è misurabile. Per verificarlo basta mostrare la misurabilità delle controimmagini di un insieme di generatori di  $\mathcal{A} \otimes \mathcal{B}$ , ad esempio  $S$ . Sia allora  $A \times B \in S$ . Se  $x \in A$   $j_x^{-1}(A \times B) = B \in \mathcal{B}$ , mentre se  $x \notin A$   $j_x^{-1}(A \times B) = \emptyset$ .

Tornando al teorema, siano  $(A_n \times B_n)_{n \in \mathbb{N}}$  disgiunti e supponiamo che la loro unione sia della forma  $A \times B \in S$ . Applicando  $j_x^{-1}$  con  $x \in A$  otteniamo

$$B = j_x^{-1}(A \times B) = j_x^{-1}(\sqcup_n A_n \times B_n) = \sqcup_n j_x^{-1}(A_n \times B_n) = \sqcup_{x \in A_n} B_n$$

Dunque, prendendo la misura

$$\nu(B) = \sum_{x \in A_n} \nu(B_n) = \sum_n \nu(B_n) \chi_{A_n}(x)$$

D'altra parte, se  $x \notin A$

$$\emptyset = j_x^{-1}(A \times B) = \sqcup_n j_x^{-1}(A_n \times B_n) = \emptyset$$

Per quanto l'identità sia banale, prendendo le misure possiamo formulare l'uguaglianza tra le funzioni

$$\nu(B) \chi_A(x) = \sum_n \nu(B_n) \chi_{A_n}(x)$$

Applicando Beppo Levi alla successione delle somme parziali, crescenti in quanto ogni termine è positivo, otteniamo

$$\nu(B) \mu(A) = \sum_n \nu(B_n) \mu(A_n)$$

Da cui la tesi. □

Con un ragionamento induttivo la dimostrazione si estende al caso di un prodotto di  $n$  misure  $\mu_1, \dots, \mu_n$ .

## 8.5 Sistemi di Dynkin e $\pi$ -sistemi

**Definizione 49.** *Una famiglia di parti  $\pi \subseteq \mathcal{P}(X)$  si dice  $\pi$ -sistema se è non vuoto e chiuso per intersezioni finite.*

**Definizione 50** (Sistema di Dynkin). *Una famiglia di parti  $\lambda \subseteq \mathcal{P}(X)$  si dice sistema di Dynkin, o  $\lambda$ -sistema, se*

- $\emptyset \in \lambda$ ;
- Se  $A \in \lambda$ , allora  $A^C \in \lambda$ ;

<sup>2</sup>In virtù della definizione 51, che vedremo a breve, questa  $\sigma$ -algebra è la più piccola a rendere misurabili le proiezioni  $X \times Y \implies X$  e  $X \times Y \implies Y$ .

- Se  $(A_n)_{n \in \mathbb{N}} \subseteq \lambda$  è una famiglia disgiunta allora  $\sqcup_{n \in \mathbb{N}} A_n \in \lambda$ .

Si osservi che una  $\sigma$ -algebra è un sistema di Dynkin, ma che non vale il viceversa. Ad esempio, sull'insieme  $X = \{1, 2, 3, 4\}$ , la famiglia degli insiemi di cardinalità pari è un  $\lambda$ -sistema ma non una  $\sigma$ -algebra, ad esempio perché non è chiusa per l'unione degli insiemi  $\{1, 2\}$  e  $\{1, 3\}$ .

Quello che manca a un sistema di Dynkin per essere una  $\sigma$ -algebra è la chiusura per intersezioni finite.

**Lemma 8.5.1.** *Una famiglia di parti è una  $\sigma$ -algebra se e solo se è contemporaneamente un  $\pi$ -sistema e un sistema di Dynkin.*

*Dimostrazione.* La freccia ( $\implies$ ) è ovvia. D'altra parte, per verificare che una famiglia di parti è una  $\sigma$ -algebra bisogna far vedere che contiene il vuoto, è chiusa per complementare e per unioni numerabili. I primi due requisiti sono garantiti dalle proprietà del  $\lambda$ -sistema. Per quanto riguarda la terza, sia  $(A_n)_{n \in \mathbb{N}}$  una famiglia numerabile. Definiamo la successione

$$\begin{cases} B_0 = A_0 \\ B_{n+1} = A_{n+1} \setminus \bigcup_{k=0}^n A_k = A_{n+1} \cap \bigcap_{k=0}^n A_k^C \end{cases}$$

Per ogni  $n$   $B_n$  appartiene alla famiglia perché lo abbiamo costruito usando solo complementari e intersezioni finite. Inoltre i  $B_n$  sono disgiunti e la loro unione coincide con quella degli  $A_n$ . Perciò  $\bigcup_{n \in \mathbb{N}} A_n$  appartiene alla famiglia per le proprietà del sistema di Dynkin.  $\square$

Con gli stessi ragionamenti visti nella sezione 8.1 si mostra che, data una famiglia di parti  $S$ , è ben definito il più piccolo sistema di Dynkin che la estende, il quale prende ovviamente il nome di  $\lambda$ -sistema generato da  $S$  e viene indicato con  $\lambda(S)$ . Questo concetto compare nel seguente lemma, la cui potenza va ben oltre quanto si possa immaginare.

**Lemma 8.5.2** (Dynkin, o della classe monotona). *Se  $\pi$  è un  $\pi$ -sistema,  $\lambda(\pi)$  è una  $\sigma$ -algebra. In altri termini, se un  $\pi$ -sistema è contenuto in un sistema di Dynkin  $\lambda$ , anche la  $\sigma$ -algebra generata da  $\pi$  sta dentro  $\lambda$ .*

*Dimostrazione.* Per quanto detto nel lemma 8.5.1 basta mostrare che  $\lambda(\pi)$  è un  $\pi$ -sistema. Dato  $A \in \lambda(\pi)$ , sia

$$K(A) = \{B \in \lambda(\pi) \text{ t.c. } A \cap B \in \lambda(\pi)\}$$

La tesi equivale a dire che per ogni  $A \in \lambda(\pi)$   $K(A) \supseteq \lambda(\pi)$ , dato che l'altra inclusione è ovvia. Osserviamo innanzitutto che  $K(A)$  è un sistema di Dynkin. Infatti:

- Ovviamente  $\emptyset \in K(A)$ ;
- Se  $A \cap B \in \lambda(\pi)$  allora per proprietà del  $\lambda$ -sistema

$$A \cap B^C = (A \cap B)^C \setminus A^C = (A \cap B)^C \cap A = ((A \cap B) \sqcup A^C)^C \in \lambda(\pi)$$

- Se  $(B_n)_{n \in \mathbb{N}}$  è una famiglia numerabile disgiunta tale che  $A \cap B_n \in \lambda(\pi)$  per ogni  $n$ , allora

$$A \cap \bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} A \cap B_n \in \lambda(\pi)$$

Ora, se  $A \in \pi$  per ipotesi  $\pi \subseteq K(A)$ . Dunque essendo  $K(A)$  un sistema di Dynkin  $K(A) \supseteq \lambda(\pi)$ , cioè  $\forall B \in \lambda(\pi)$  e per ogni  $A \in \pi$  abbiamo che  $B \in K(A)$ . D'altra parte  $B \in K(A)$  se e solo se  $A \in K(B)$ , perché entrambe le condizioni equivalgono a richiedere che  $A \cap B \in \lambda(\pi)$ . Quindi, dato che la condizione vale per ogni  $A \in \pi$ , abbiamo che per ogni  $B \in \lambda(\pi)$   $\pi \subseteq K(B)$ , da cui la tesi.  $\square$

### 8.5.1 Coincidenza di misure

Coincidenze? Io non credo.

---

Alan Kadmond

**Teorema 8.5.3** (Coincidenza). *Siano  $\mu, \nu$  due misure  $\sigma$ -finite su  $\mathcal{A} \subseteq \mathcal{P}(X)$  che coincidono sul  $\pi$ -sistema  $\pi$ . Se quest'ultimo contiene una successione di insiemi  $A_n \uparrow X$  di misura finita e se  $\sigma(\pi) = \mathcal{A}$ , allora le due misure coincidono su tutta  $\mathcal{A}$ .*

*Dimostrazione.* Assumiamo dapprima che  $\mu(X)$  sia finita. Per il lemma di Dynkin, è sufficiente mostrare che la famiglia degli insiemi su cui  $\mu$  e  $\nu$  coincidono è un sistema di Dynkin.

- Poiché le due misure coincidono sulla successione  $A_n \uparrow X$ , abbiamo che

$$\mu(X) = \lim_{n \rightarrow +\infty} \mu(A_n) = \lim_{n \rightarrow +\infty} \nu(A_n) = \nu(X)$$

- Se  $\mu(B) = \nu(B)$ , allora

$$\mu(B^C) = \mu(X) - \mu(B) = \nu(X) - \nu(B) = \nu(B^C)$$

- Se gli elementi  $B_n$  sono disgiunti a due a due e se per ogni  $n \in \mathbb{N}$   $\mu(B_n) = \nu(B_n)$ , allora

$$\mu\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \sum_{n \in \mathbb{N}} \mu(B_n) = \sum_{n \in \mathbb{N}} \nu(B_n) = \nu\left(\bigcup_{n \in \mathbb{N}} B_n\right)$$

Supponiamo ora che  $\mu(X) = \nu(X) = \infty$ . Fissato  $n \in \mathbb{N}$  definiamo

$$\mathcal{A}_n = \{B \subseteq A_n \text{ t.c. } B \in \mathcal{A}\}$$

Dato che  $\mu$  e  $\nu$  coincidono sul  $\pi$ -sistema

$$\pi_n = \{B \subseteq A_n \text{ t.c. } B \in \pi\}$$

per quanto visto nel caso precedente le misure coincidono anche su  $\sigma(\pi_n)$ . Mostriamo che questa famiglia contiene  $\mathcal{A}_n$ . Infatti la famiglia

$$\{B \subseteq X \text{ t.c. } B \cap A_n \in \sigma(\pi_n)\}$$

è una  $\sigma$ -algebra e contiene  $\pi$ , perché  $A_n \in \pi$  e  $\pi$  è chiuso per intersezioni. Allora deve contenere  $\mathcal{A}$  per il lemma di Dynkin. Dunque ogni elemento di  $\mathcal{A}$  contenuto in  $A_n$  appartiene a  $\sigma(\pi_n)$ , e perciò per ogni  $B \in \mathcal{A}$  abbiamo che

$$\mu(B \cap A_n) = \nu(B \cap A_n)$$

Passando al limite per  $n \rightarrow +\infty$  abbiamo la tesi. □

Ricollegandoci a quanto visto nel teorema 8.3.4, possiamo dire che

**Teorema 8.5.4.** *Se  $\phi$  è  $\sigma$ -additiva su un quasi-anello  $S$  si estende a una misura su  $\sigma(S)$ , in modo unico se  $\phi$  è  $\sigma$ -finita.*

## 8.6 Costruzione della misura di Lebesgue

Ricordiamo che, per secondo-numerabilità della topologia euclidea, si ha che

**Lemma 8.6.1.** *Un aperto  $A \subseteq \mathbb{R}$  è unione numerabile di intervalli aperti.*

Questo ci mostra in particolare che

**Corollario 8.6.1.1.** *I boreliani  $\mathbb{B}$  di  $\mathbb{R}$  sono generati dagli intervalli aperti.*

Infatti, se una  $\sigma$ -algebra contiene gli intervalli aperti deve contenere tutti gli aperti di  $\mathbb{R}$ , essendo chiusa per unione numerabile, e questi ultimi generano i boreliani per definizione. Dunque per definire una misura su  $\mathbb{R}$  ci basterà lavorare con gli intervalli.

**Teorema 8.6.2.** *La funzione  $\lambda$  definita sugli intervalli di  $\mathbb{R}$  da  $\lambda(I) = |I| = \sup(I) - \inf(I)$  è  $\sigma$ -additiva.*

*Dimostrazione.* Sia  $I = \sqcup_{n \in \mathbb{N}} I_n$  un intervallo dato dall'unione numerabile disgiunta di intervalli. Se l'unione è finita la tesi è banale, perché si tratta di scrivere esplicitamente una somma telescopica. In caso contrario dobbiamo distinguere due casi.

- Supponiamo che  $I$  sia limitato. Fissato  $\varepsilon > 0$  esiste un sottointervallo compatto  $K \subseteq I$  tale che  $\lambda(I) - \lambda(K) \leq \varepsilon$ . Inoltre, per ogni  $I_n$  è possibile determinare un intervallo aperto  $J_n \supseteq I_n$  tale che  $\lambda(J_n) - \lambda(I_n) \leq \varepsilon/2^n$ . Per compattezza di  $K$  dal ricoprimento aperto  $\{J_n\}_{n \in \mathbb{N}}$  si può estrarre un sottoricoprimento finito, sia esso  $\{J_1, \dots, J_k\}$ . Dunque

$$\lambda(I) - \varepsilon \leq \lambda(K) \leq \sum_{n=0}^k \lambda(J_n) \leq \sum_{n \in \mathbb{N}} \lambda(I_n) + \varepsilon/2^n = \varepsilon + \sum_{n \in \mathbb{N}} \lambda(I_n)$$

Dunque, passando al limite per  $\varepsilon \rightarrow 0$ , otteniamo che  $\lambda(I) \leq \sum_{n \in \mathbb{N}} \lambda(I_n)$ . Per quanto riguarda la disuguaglianza opposta basta banalmente osservare che  $\lambda(I) \geq \sum_{n=0}^k \lambda(I_n)$  e passare al limite.

- Se  $I$  è illimitato, per ogni  $k \in \mathbb{N}$  possiamo considerare un sottointervallo  $J_k \subseteq I$  di lunghezza  $k$  e osservare che, per quanto mostrato nel punto precedente,

$$k = \lambda(J_k) = \sum_{n \in \mathbb{N}} \lambda(J_k \cap I_n) \leq \sum_{n \in \mathbb{N}} \lambda(I_n)$$

Dunque la somma  $\sum_{n \in \mathbb{N}} \lambda(I_n)$  è maggiore di ogni naturale  $k$ , e cioè tende a  $\lambda(I) = +\infty$ , come richiesto. □

Ora, dati due intervalli  $I$  e  $J$ , ovviamente  $\lambda(I) = \lambda(I \cap J) + \lambda^*(I \cap J^C)$ . Dunque, in virtù del teorema 8.3.3 abbiamo ottenuto

**Teorema 8.6.3.** *La lunghezza degli intervalli si estende a una misura  $\lambda$ , detta misura di Lebesgue, su una  $\sigma$ -algebra  $\mathcal{L}$  che contiene i boreliani.*

Negli integrali la misura di Lebesgue si indica generalmente con  $dx$ .

Per quanto detto in 8.4 sui boreliani<sup>3</sup> di  $\mathbb{R}^n$  si può definire la misura  $\lambda^n$  data dal prodotto della misura di Lebesgue con se stessa  $n$  volte.

### 8.6.1 Proprietà della misura di Lebesgue

- Un punto, considerato come un prodotto di intervalli di lunghezza nulla, ha misura di Lebesgue nulla. Dunque, per  $\sigma$ -additività della misura, gli insiemi numerabili sono tutti trascurabili. Questo in particolare ci dice che

**Corollario 8.6.3.1.** *In  $\mathbb{R}$  i numeri razionali, e in generale i numeri reali algebrici, hanno misura di Lebesgue nulla.*

- La misura di Lebesgue è  $\sigma$ -finita, dato che  $\mathbb{R}^n = \bigcup_{k \in \mathbb{N}} (-k, k)^n$ . Dunque, per il teorema 8.3.7, otteniamo che<sup>4</sup>

**Corollario 8.6.3.2.** *I Lebesgue-misurabili in dimensione  $n$  sono tutti e soli gli insiemi della forma*

$$\{B \setminus N \text{ t.c. } B \in \mathbb{B}(\mathbb{R}^n), N \subseteq P \in \mathbb{B}(\mathbb{R}^n), \lambda(P) = 0\}$$

<sup>3</sup>In realtà, la misura costruita in 8.4 si estende all'algebra prodotto ottenuta dall'algebra di Lebesgue. Tuttavia non è chiaro se quest'ultima coincida effettivamente con il completamento dei boreliani di  $\mathbb{R}^n$ , ossia se il prodotto dei completamenti sia il completamento dell'algebra del prodotto (cosa che in generale dovrebbe essere falsa). Quindi meglio evitare di incasinarci.

<sup>4</sup>Sicuramente il teorema vale in dimensione 1. Dovrebbe essere vero in generale, ma bisogna prendere tutto con le pinze per i motivi detti nella nota precedente.

- La  $\sigma$ -algebra di Lebesgue è il completamento secondo Carathéodory dei boreliani, ossia della  $\sigma$ -algebra generata dal quasi-anello degli intervalli (aperti). Ripercorrendo la definizione della misura esterna indotta e la dimostrazione del lemma 8.3.5, possiamo mostrare che

**Lemma 8.6.4** (Regolarità della misura di Lebesgue). *Per ogni Lebesgue-misurabile  $B \in \mathcal{L}(\mathbb{R})$  esiste una successione di aperti  $A_n \supseteq A_{n+1} \supseteq B$  la cui intersezione è un coperchio di  $B$ , ossia tali che*

$$\lambda(A_n) \downarrow \lambda\left(\bigcap_n A_n\right) = \lambda(B)$$

*In altre parole, un insieme misurabile secondo Lebesgue coincide con un  $G_\delta$  cui si sottrae un insieme di misura nulla.*

- Prendendo il complementare nella relazione precedente, otteniamo che

**Corollario 8.6.4.1.** *Per ogni insieme misurabile  $B \in \mathcal{L}(\mathbb{R})$  esiste una successione crescente di chiusi  $C_n \subseteq C_{n+1} \subseteq B$  la cui unione coincide con  $B$  a meno di un insieme di misura nulla. A meno di intersecare i  $C_n$  con le palle  $\overline{B}(0, n)$  è possibile scegliere i  $C_n$  compatti.*

- Si mostra che

**Lemma 8.6.5.** *La misura di Lebesgue  $n$ -dimensionale è invariante per traslazioni.*

*Dimostrazione.* Sia  $A \in \mathcal{L}(\mathbb{R}^n)$  e sia  $h \in \mathbb{R}^n$ . Osserviamo dapprima che anche  $A + h \in \mathcal{L}$ , perché un Lebesgue misurabile si ottiene da un  $G_\delta$ , che traslato rimane un  $G_\delta$  dato che gli aperti rimangono aperti, cui si sottrae un insieme di misura nulla, che traslato rimane trascurabile (Infatti un insieme è trascurabile se si ricopre con una famiglia numerabile di plurirettangoli la cui somma delle misure è arbitrariamente piccola, e traslando un plurirettangolo se ne ottiene un altro della stessa misura). Mostriamo ora che gli insiemi  $A$  e  $A + h$  hanno la stessa misura. Per farlo, osserviamo che  $\mu(A) = \lambda(A + h)$  è una misura che coincide con  $\lambda$  sul  $\pi$ -sistema dei rettangoli della forma  $\prod [a_i, b_i]$ . Dunque le due misure coincidono sui boreliani, e per completamento sui Lebesgue-misurabili.  $\square$

- Più in generale

**Teorema 8.6.6.** *La misura di Lebesgue  $\lambda^n$  è l'unica misura su  $\mathbb{B}(\mathbb{R}^n)$  invariante per traslazioni e normalizzata a 1 su  $[0, 1]^n$ .*

*Dimostrazione.* Sia  $\mu$  un'altra misura con le caratteristiche dell'ipotesi. Ci basta mostrare che coincide con  $\lambda$  sui rettangoli di intervalli e concludere come nel lemma 8.6.5.

Osserviamo innanzitutto che i prodotti della forma  $\{x\} \times [0, 1]^{n-1}$  hanno misura zero. Infatti, l'insieme  $A = (\{0\} \cup \{1/n, n \in \mathbb{N} \setminus \{0\}\}) \times [0, 1]^{n-1}$  è boreliano, in quanto unione numerabile di chiusi, ed ha misura minore o uguale a 1 essendo contenuto in  $[0, 1]^n$ . Ma dato che tutti i chiusi considerati devono avere la stessa misura per invarianza per traslazioni si deve verificare che  $\mu(A) = \sum_{n \in \mathbb{N}} \mu(\{x\} \times [0, 1]^{n-1}) = 0$ , da cui  $\mu(\{x\} \times [0, 1]^{n-1}) = 0$ .

Dunque possiamo suddividere  $[0, 1]^n$  in tre parti:

$$[0, 1] = ([0, 1/2) \cup [1/2, 1) \cup \{1\}) \times [0, 1]^{n-1}$$

Di queste, il punto è trascurabile, mentre le altre due parti hanno misura uguale per invarianza per traslazioni. Otteniamo allora che  $\mu([0, 1/2) \times [0, 1]^{n-1}) = 1/2$ .

Ripetendo il ragionamento, possiamo mostrare che  $\mu([0, 2^{-n}) \times [0, 1]^{n-1}) = 2^{-n}$ . Infine, dato un numero reale positivo  $c$ , possiamo scriverlo in base 2 come  $c = \sum_{n \in \mathbb{N}} a_n 2^{-n}$ . Allora, suddividendo  $[0, c) \times [0, 1]^{n-1}$  in segmenti semiaperti di lunghezza  $2^{-n}$  secondo la sua scrittura binaria otteniamo

$$\mu([0, c) \times [0, 1]^{n-1}) = \sum_{a_n=1} 2^{-n} = c$$

Dunque  $\mu([a, b) \times [0, 1]^{n-1}) = b - a = \lambda([a, b) \times [0, 1]^{n-1})$  per invarianza sotto traslazioni.

Ora possiamo reiterare i passaggi con la seconda coordinata, tenendo fissa la prima, ottenendo che  $\mu([a, b) \times [c, d) \times [0, 1]^{n-2}) = (b - a)(d - c)$ . Ragionando induttivamente e applicando il lemma di Dynkin si ottiene la tesi.  $\square$

## 8.7 Misure boreliane su $[0, 1]$

Sia  $\mu$  una misura boreliana finita su  $[0, 1]$ . Consideriamo la funzione, detta funzione di ripartizione di  $\mu$ :

$$\begin{aligned}\phi : [0, 1] &\longrightarrow \mathbb{R}_0^+ \\ x &\longrightarrow \mu([0, x])\end{aligned}$$

Tale funzione è nulla in zero e crescente per evidenti motivi. Inoltre è continua a sinistra, perché

$$\lim_{x \rightarrow x_0^+} \phi(x) = \sup_{x < x_0} \mu([0, x]) = \mu([0, x_0]) = \phi(x_0)$$

Dimostriamo che una funzione come  $\phi$  determina una misura boreliana.

**Teorema 8.7.1.** *Sia  $\phi : [0, 1] \rightarrow \mathbb{R}_0^+$  nulla in zero, crescente e continua a sinistra, e sia  $c \in \mathbb{R}_0^+$ . Esiste un'unica misura boreliana tale che  $\mu([0, x]) = \phi(x)$  e  $\mu(\{1\}) = c$ .*

*Dimostrazione.* Poniamo

$$\begin{aligned}\mu([a, b]) &= \phi(b) - \phi(a) \\ \mu(\{1\}) &= c\end{aligned}$$

Ragionando come nel teorema 8.6.2 si mostra che  $\mu$  è  $\sigma$ -additiva sull'anello  $\mathcal{A}$  degli intervalli semiaperti di  $[0, 1]$  a cui uniamo  $\{1\}$ , dunque per il teorema 8.3.4  $\mu$  si estende a una misura sui boreliani di  $[0, 1]$ . Tale misura è unica per il teorema 8.5.3, in quanto è univocamente definita sul  $\pi$ -sistema  $\mathcal{A}$ , dal quale si estrae un ricoprimento di  $[0, 1]$  con insiemi di misura finita.  $\square$

### 8.7.1 Bene...

**Lemma 8.7.2.** *I boreliani della retta reale sono strettamente contenuti nei Lebesgue-misurabili.*

Dimostriamo innanzitutto che

**Lemma 8.7.3.** *I boreliani hanno la cardinalità del continuo.*

La dimostrazione risulta un po' tecnica per chi non ha già visto un po' di nozioni di teoria degli insiemi, perciò consiglio di passare avanti a chi non avesse gli strumenti.

*Dimostrazione.* Osserviamo innanzitutto che è possibile costruire per induzione transfinita la  $\sigma$ -algebra di Borel<sup>5</sup>. Poniamo  $B_0 = \tau$ , dove  $\tau$  è la topologia euclidea. Dopodiché definiamo, per induzione transfinita sugli ordinali  $\alpha$

$$B_{\alpha+1} = B_\alpha \cup \left\{ A^C, \bigcup A_n \text{ t.c. } A, A_n \in B_\alpha \right\}$$

E per gli ordinali limite

$$B_\alpha = \bigcup_{\beta < \alpha} B_\beta$$

Ora, detto  $\omega_1$  il primo ordinale non numerabile,  $B_{\omega_1}$  è una  $\sigma$ -algebra. Infatti

- $X \in \tau \subseteq B_{\omega_1}$ .
- Se  $A \in B_{\omega_1}$ , esiste un ordinale  $\alpha < \omega_1$  tale che  $A \in B_\alpha$ . Allora  $A^C \in B_{\alpha+1} \subseteq B_{\omega_1}$ .
- Data una successione  $(A_n)_{n \in \mathbb{N}} \subseteq B_{\omega_1}$ , essendo  $\aleph_1$  un cardinale regolare deve esistere un  $\alpha$  tale che  $(A_n)_{n \in \mathbb{N}} \subseteq B_\alpha$ , o si potrebbe costruire una successione numerabile cofinale in  $\omega_1$ . Dunque  $\bigcup_{n \in \mathbb{N}} (A_n) \in B_{\alpha+1} \subseteq B_{\omega_1}$ .

<sup>5</sup>Si osservi che lo stesso ragionamento fornisce una caratterizzazione "esplicita" (...?) della  $\sigma$ -algebra generata da una qualsiasi famiglia di parti  $S$ . Basta infatti porre  $B_0 = S$  e procedere analogamente.

$B_{\omega_1}$  risulta contenuta nei boreliani, in quanto abbiamo solo eseguito operazioni di complementazione e unione numerabile a partire dalla topologia euclidea. Dunque  $B_{\omega_1}$  deve coincidere con l'algebra di Borel  $\mathbb{B}$ , che è la più piccola  $\sigma$ -algebra che estende  $\tau$ .

Inoltre,  $B_{\omega_1}$  ha la cardinalità del continuo  $c$ . Questo perché

- $|\tau| = c$  (volendo, perché ogni aperto è unione di una sottofamiglia di una base numerabile);
- Se  $|B_\alpha| = c$ , allora

$$|B_{\alpha+1}| \leq |B_\alpha| + |\{A^C \text{ t.c. } A \in B_\alpha\}| + \left| \left\{ \bigcup A_n \text{ t.c. } A_n \in B_\alpha \right\} \right| \leq c + c + c^{\aleph_0} = c$$

- Se  $\alpha$  è un limite,

$$|B_\alpha| \leq \sum_{\beta < \alpha} |B_\beta| \leq |\alpha| \cdot c$$

E dato che  $|\alpha| \leq |\omega_1| = \aleph_1 \leq c$  anche  $|B_\alpha| = c$

□

Invece, i Lebesgue-misurabili hanno cardinalità  $2^c$ . Infatti, l'insieme di Cantor è Borel-misurabile, dato che è chiuso, ed ha misura nulla perché si ottiene come l'intersezione degli insiemi della successione

$$\begin{cases} I_0 = [0, 1] \\ I_{n+1} = \frac{1}{3}I_n \cup \left(\frac{2}{3} + \frac{1}{3}I_n\right) \end{cases}$$

La cui misura tende a zero. Dunque  $\mathcal{L}$  deve contenere tutti i sottoinsiemi dell'insieme di Cantor, che ha la cardinalità del continuo dato che coincide con gli elementi di  $[0, 1]$  la cui scrittura in base 3 non contiene 1. Perciò  $|\mathcal{L}| \geq 2^c$ , e deve valere l'uguaglianza dato che  $\mathcal{L} \subseteq \mathcal{P}(\mathbb{R})$ .

## 8.7.2 ... ma non benissimo

**Teorema 8.7.4.** *Esiste un sottoinsieme di  $\mathbb{R}$  che non è Lebesgue-misurabile.*

Nella dimostrazione faremo uso del teorema di Steinhaus, di cui parleremo in 9.6.2.

*Dimostrazione.* Consideriamo  $\mathbb{R}$  come spazio vettoriale su  $\mathbb{Q}$ . Per un corollario dell'assioma della scelta<sup>6</sup> esiste una base  $B$ , e ciò determina una decomposizione della forma  $\mathbb{R} = \mathbb{Q} \oplus V$ , dove  $V$  è lo spazio generato dagli elementi irrazionali della base.  $V$  prende il nome di insieme di Vitali, e vogliamo mostrare che non è misurabile. Infatti non può avere misura nulla, perché altrimenti

$$\lambda(\mathbb{R}) = \lambda(\bigsqcup_{q \in \mathbb{Q}} (V + q)) = \sum_{q \in \mathbb{Q}} \lambda(V) = 0$$

Abbiamo usato che i  $V + q$  hanno tutti la stessa misura per invarianza traslazionale. D'altra parte se  $V$  avesse misura positiva  $V - V$  dovrebbe essere un intorno dell'origine, ma dato che  $V$  è uno spazio vettoriale  $V - V = V$  e questo non può avere parte interna o conterrebbe dei razionali non nulli. □

La situazione è ancora meno rosea:

**Teorema 8.7.5.** *Ogni insieme misurabile secondo Lebesgue in  $\mathbb{R}^n$  di misura positiva contiene un insieme non misurabile.*

*Dimostrazione.* Facciamo il caso  $n = 1$ , dato che il caso generale è analogo. Sia  $A \subseteq \mathbb{R}$  misurabile e di misura positiva. Allora, detto  $V$  l'insieme di Vitali,

$$A = \mathbb{R} \cap A = (\bigsqcup_{q \in \mathbb{Q}} (V + q)) \cap A = (\bigsqcup_{q \in \mathbb{Q}} (V + q) \cap A)$$

Ora, poiché  $A$  ha misura positiva, uno dei  $(V + q) \cap A$  dovrebbe avere misura non nulla. D'altra parte  $((V + q) \cap A) - ((V + q) \cap A) \subseteq V$ , e si conclude come nel teorema precedente. □

<sup>6</sup>Non dovrebbe essere possibile aggirare questo passaggio, dato che, sotto certe condizioni, potrebbero esistere modelli di  $ZF \setminus C$  in cui non vale l'assioma della scelta e nei quali ogni sottoinsieme di  $\mathbb{R}$  è Lebesgue-misurabile. Vedasi a riguardo [https://en.wikipedia.org/wiki/Solovay\\_model](https://en.wikipedia.org/wiki/Solovay_model). D'altra parte, se non accettate l'assioma della scelta non siete i benvenuti nelle mie dispense.



# Capitolo 9

## Teoria dell'integrazione

### 9.1 Funzioni misurabili

**Definizione 51** (Funzione misurabile). *Una funzione tra spazi di misura  $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$  si dice misurabile se*

$$\forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{A}$$

Si noti che, affinché  $f$  sia misurabile, è sufficiente richiedere che la controimmagine di un insieme di generatori di  $\mathcal{B}$  sia contenuta in  $\mathcal{A}$ . Si osservi inoltre che la composizione di funzioni misurabili è misurabile, esattamente come succede per le funzioni continue. Inoltre, le funzioni continue tra spazi topologici sono misurabili rispetto ai Boreliani in partenza e in arrivo.

Nel caso in cui  $(Y, \mathcal{B}) = (\mathbb{R}, \mathbb{B})$  abbiamo

**Definizione 52** (Funzione Borel-misurabile). *Una funzione tra spazi di misura  $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathbb{B})$  si dice Borel-misurabile se la controimmagine di un boreliano è misurabile. Equivalentemente si può richiedere soltanto che gli insiemi  $\{f < c\}$  siano misurabili, perché le semirette generano il boreliano.*

A differenza del caso generale, non è sempre detto che la composizione di funzioni su  $(\mathbb{R}, \mathcal{L})$  ma non sui boreliani. Invece è vero che

**Lemma 9.1.1.** *Sia  $f_n : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathbb{B})$  una famiglia di funzioni Borel-misurabili. La somma finita, il prodotto finito e gli involucri  $\sup_n f_n$  e  $\inf_n f_n$  (e perciò anche il  $\limsup$  e il  $\liminf$  della successione) sono Borel-misurabili.*

*Dimostrazione.* Per quanto riguarda somma e prodotto la tesi è ovvia, sfruttando la caratterizzazione delle funzioni Borel-misurabili con le preimmagini delle semirette. Mostriamo che l'involuppo inferiore è misurabile.

$$\{\inf_n f_n < c\} = \bigcup_n \{f_n < c\} \in \mathcal{A}$$

Per quanto riguarda l'involuppo superiore basta osservare che  $\sup_n f_n = -\inf_n(-f_n)$ . □

Poiché gli involucri superiori ereditano la misurabilità, abbiamo anche che

**Lemma 9.1.2.** *Se  $f$  è misurabile, lo sono anche  $f_+ = \max(f, 0)$  e  $f_- = -\min(f, 0)$ .*

#### 9.1.1 Funzioni misurabili rispetto al completamento

**Teorema 9.1.3.** *Sia  $(X, \mathcal{A}, \mu)$  uno spazio di misura  $\sigma$ -finito. Se  $(X, \overline{\mathcal{A}}, \overline{\mu})$  è il suo completamento, una funzione  $f : X \rightarrow (\mathbb{R}, \mathbb{B})$  è  $\overline{\mathcal{A}}$ -misurabile se e solo se esiste una funzione  $g$   $\mathcal{A}$ -misurabile che coincide con  $f$   $\overline{\mu}$ -quasi ovunque.*

*Dimostrazione.* Mostriamo le due implicazioni.

- ( $\implies$ ) Innanzitutto la tesi è vera per le funzioni caratteristiche di insiemi di  $\overline{\mathcal{A}}$ . Infatti, dato  $A \in \overline{\mathcal{A}}$ , per quanto risulta da 8.3.7 possiamo scrivere  $A = A^* \setminus N$  dove  $A^* \in \mathcal{A}$  e  $N \in \mathcal{A}$  è  $\bar{\mu}$ -trascurabile. Ovviamente possiamo supporre  $N \subseteq A^*$ . Dunque  $\chi_A = \chi_{A^*} - \chi_N$ , e cioè  $\chi_A = \chi_{A^*}$  quasi ovunque. Da quanto appena detto segue facilmente la tesi per una funzione semplice, dato che basta rimpiazzare ogni indicatrice come nel caso precedente. Ora, se  $f = \lim_n \phi_n$  e  $\phi_n$  coincide quasi ovunque con una combinazione lineare  $\psi_n$  di caratteristiche di insiemi in  $\mathcal{A}$ , la famiglia  $\psi_n$  converge in un insieme di misura piena (precisamente, nell'intersezione degli insiemi su cui  $\phi_n = \psi_n$ ) a una certa funzione  $g$ , che è  $\mathcal{A}$ -misurabile perché limite di funzioni  $\mathcal{A}$ -misurabili e che coincide con  $f$  quasi ovunque.
- ( $\impliedby$ ) Se  $f, g$  sono come nelle ipotesi e  $g = f$  quasi ovunque, possiamo scrivere  $f = g + (f - g)$ . Dunque, dato che la somma di funzioni misurabili è misurabile, è sufficiente far vedere che  $f - g$  è  $\overline{\mathcal{A}}$ -misurabile, e a tal fine basta verificare che per ogni  $c \in \mathbb{R}$  l'insieme  $A = \{f - g < c\}$  è misurabile in quanto le semirette  $\{x < c\}$  generano i boreliani. Questo è vero perché se  $c \leq 0$   $A$  è un sottoinsieme di un trascurabile, e in caso contrario coincide con  $X$  meno un sottoinsieme di un insieme di misura nulla. Dunque per completezza della  $\sigma$ -algebra  $\overline{\mathcal{A}}$  abbiamo che  $A \in \overline{\mathcal{A}}$ , da cui la tesi. □

### 9.1.2 Trasporto di misure

**Definizione 53** (Push-forward di una misura). Sia  $F : (X, \mathcal{A}, \mu) \longrightarrow (Y, \mathcal{B})$  una funzione misurabile. La funzione di insieme su  $\mathcal{B}$  definita da

$$F_{\#}\mu(B) = \mu(F^{-1}(B))$$

è una misura, detta "push-forward" di  $\mu$  tramite  $F$ .

Le verifiche sono immediate.

## 9.2 Costruzione dell'integrale

**Definizione 54** (Funzione semplice). Una funzione misurabile  $\phi : (X, \mathcal{A}) \longrightarrow (\mathbb{R}, \mathbb{B})$  si dice semplice se è una combinazione lineare reale di funzioni caratteristiche di insiemi misurabili. In simboli, esistono  $c_1, \dots, c_k \in \mathbb{R}$  e  $A_1, \dots, A_k \in \mathcal{A}$  tali che

$$\phi = \sum_{i=1}^k c_i \chi_{A_i}$$

Si osservi che una funzione semplice assume un numero finito di valori reali. Ha dunque senso definire la forma canonica di una funzione semplice come

$$\phi = \sum_{c \in \mathbb{R}} c \chi_{\phi=c}$$

**Definizione 55** (Integrale di una funzione semplice). Sia  $\phi : (X, \mathcal{A}, \mu) \longrightarrow (\mathbb{R}, \mathbb{B})$  una funzione semplice non negativa. Definiamo il suo integrale (rispetto alla misura  $\mu$ ) come

$$\int_X \phi d\mu = \sum_{i=1}^k c_i \mu(A_i)$$

**Lemma 9.2.1.** La definizione non dipende dalla combinazione lineare che definisce  $\phi$ .

*Dimostrazione.* Supponiamo che  $\phi = \sum_{i=1}^k c_i \chi_{A_i} = \sum_{c \in \mathbb{R}} c \chi_{\phi=c}$ . Allora, dato che gli insiemi della seconda suddivisione partizionano  $X$ , possiamo scrivere

$$\sum_{i=1}^k c_i \mu(A_i) = \sum_{c \in \mathbb{R}} \sum_{i=1}^k c_i \mu(A_i \cap \{\phi = c\})$$

Basta allora mostrare che, per ogni  $c \in \mathbb{R}$ ,

$$\sum_{i=1}^k c_i \mu(A_i \cap \{\phi = c\}) = c \mu(\{\phi = c\})$$

Per farlo, per ogni sottoinsieme di indici  $I \subseteq \{1, \dots, k\}$  consideriamo l'insieme

$$A_I = \{x \in \{\phi = c\} : x \in A_i \iff i \in I\}$$

Gli  $A_I$  definiscono una partizione misurabile di  $\{\phi = c\}$ . Osserviamo inoltre che, se  $i \in I$ ,  $A_I \subseteq A_i$ , altrimenti  $A_I \cap A_i = \emptyset$ . Abbiamo allora

$$\sum_{i=1}^k c_i \mu(A_i \cap \{\phi = c\}) = \sum_{i=1}^k \sum_{I \subseteq \{1, \dots, k\}} c_i \mu(A_i \cap A_I) = \sum_{I \subseteq \{1, \dots, k\}} \sum_{i \in I} c_i \mu(A_I)$$

Ora, se  $A_I \neq \emptyset$ ,  $\sum_{i \in I} c_i = c$ , perché se  $\phi(x) = c$  allora

$$c = \sum_{i=1}^k c_i \chi_{A_i}(x) = \sum_{i=1}^k c_i \chi_{A_i \cap \{\phi=c\}}(x) = \sum_{I \subseteq \{1, \dots, k\}} \sum_{i \in I} c_i \chi_{A_I}(x)$$

E dunque, scegliendo  $x \in A_I$ , otteniamo  $\sum_{i \in I} c_i = c$ . Allora

$$\sum_{i=1}^k c_i \mu(A_i \cap \{\phi = c\}) = \sum_{I \subseteq \{1, \dots, k\}} c \mu(A_I) = c \mu(\{\phi = c\})$$

□

**Lemma 9.2.2.** *Sia  $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathbb{B})$  misurabile e non negativa. Esiste una successione crescente di funzioni semplici che converge a  $f$  puntualmente.*

*Dimostrazione.* Basta scegliere

$$\phi_n = n \chi_{f \geq n} + \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \chi_{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}}$$

Le verifiche sono banali. □

**Definizione 56** (Integrale di funzioni non negative). *Sia  $f : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathbb{B})$  misurabile e non negativa. Definiamo l'integrale di  $f$  come*

$$\int_X f d\mu = \sup_{0 \leq \phi \leq f \text{ semplice}} \int_X \phi d\mu$$

*Una funzione positiva si dice integrabile se ha integrale finito.*

Nella presente definizione stiamo includendo anche le funzioni che assumono valore infinito. Ovviamente se una funzione siffatta è integrabile deve essere finita quasi ovunque. Spesso ometteremo la misura rispetto a cui si integra, anche solo per banale dimenticanza.

Dalla definizione segue facilmente che l'integrale è monotono nell'integrando. In altri termini, se  $f \leq g$  allora  $\int_X f \leq \int_X g$ .

**Definizione 57** (Integrale su sottoinsiemi). *Sia  $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathbb{B})$  misurabile e non negativa. Definiamo l'integrale di  $f$  su  $A \subseteq X$  come*

$$\int_A f d\mu = \int_X f \chi_A d\mu$$

*O equivalentemente come*

$$\int_A f d\mu = \int_A f|_A d\mu_A$$

La seconda definizione è ben posta perché essendo  $A$  misurabile la restrizione di  $f$  ad  $A$  è misurabile rispetto alla  $\sigma$ -algebra indotta  $\{A \cap B, B \in \mathcal{A}\}$ . Si osservi che l'integrale è additivo sul dominio, cioè che

$$\int_{A \sqcup B} f = \int_A f + \int_B f$$

Questo perché da ogni funzione semplice che approssima  $f$  dal basso su  $A \sqcup B$  se ne ricavano una su  $A$  e una su  $B$  moltiplicando per le rispettive caratteristiche; viceversa date due funzioni semplici, una su  $A$  e una su  $B$ , la loro somma è una funzione semplice su  $A \sqcup B$ .

**Teorema 9.2.3** (Beppo Levi). *Sia  $(f_n)_{n \in \mathbb{N}}$  una successione di funzioni misurabili non negative che converge crescendo a  $f$ . Allora*

$$\int_X f_n \uparrow \int_X f$$

*Dimostrazione.* Osserviamo innanzitutto che la tesi ha senso, perché per quanto si ricava dal lemma 9.1.1 il limite di una successione di funzioni misurabili è misurabile. Inoltre, poiché  $f_n \leq f$ , per monotonia segue facilmente che  $\lim_{n \rightarrow +\infty} \int_X f_n \leq \int_X f$ . Per l'altra disuguaglianza basta mostrare che, per ogni funzione semplice  $0 \leq \phi \leq f$  vale che  $\lim_{n \rightarrow +\infty} \int_X f_n \geq \int_X \phi$ . Se scriviamo  $\phi$  in forma canonica, posto  $A_c = \{\phi = c\}$ , basta mostrare che per ogni  $c \in \mathbb{R}$  esiste un  $n$  per cui  $\int_{A_c} f_n \geq c\mu(A_c)$ . Supponiamo allora WLOG che  $\phi = 1 = \text{cost}$  su tutto  $X$ , a meno di spezzare l'integrale, e sia  $\lambda < 1$ . Allora, essendo  $f_n$  positiva per ogni  $n$ ,  $f_n \geq \lambda \chi_{\{f_n \geq \lambda\}}$ , e dunque

$$\int_X f_n \geq \lambda \mu(\{f_n \geq \lambda\})$$

Ora,  $\bigcup_n \{f_n \geq \lambda\} = \{f \geq \lambda\} = X$ , dato che  $f \geq \phi = 1$ . Dunque per ogni  $\lambda < 1$

$$\lim_{n \rightarrow +\infty} \int_X f_n \geq \lambda \mu(X)$$

Passando al sup su  $\lambda < 1$  si ottiene la tesi:

$$\lim_{n \rightarrow +\infty} \int_X f_n \geq \mu(X) = \int_X \phi$$

□

Applicando Beppo Levi si vede che l'integrale di  $f$  si può definire alternativamente come

$$\int_X f = \lim_{n \in \mathbb{N}} \int_X \phi_n$$

dove le  $\phi_n$  sono una successione di funzioni semplici che convergono crescendo a  $f$ . Questo fatto torna utile nel seguente lemma:

**Lemma 9.2.4.** *L'integrale di funzioni positive è additivo nell'integrando. In simboli, se  $f, g \geq 0$ ,*

$$\int_X f + \int_X g = \int_X f + g$$

*Dimostrazione.* Mostriamo dapprima che l'integrale è additivo sulle funzioni semplici. Date  $\phi = \sum a_i \chi_{A_i}$  e  $\psi = \sum b_j \chi_{B_j}$ , WLOG in forma canonica, la loro somma è una funzione semplice sulla partizione  $(A_i \cap B_j)_{i,j}$ , e dunque

$$\int_X \phi + \psi = \sum_{i,j} (a_i + b_j) \mu(A_i \cap B_j) = \sum_i a_i \sum_j \mu(A_i \cap B_j) + \sum_j b_j \sum_i \mu(A_i \cap B_j) = \int_X \phi + \int_X \psi$$

Ora, se  $\phi_n \uparrow f$  e  $\psi_n \uparrow g$  allora  $(\phi_n + \psi_n) \uparrow (f + g)$ , perciò per quanto visto prima

$$\int_X f + \int_X g = \lim_n \int_X \phi_n + \lim_n \int_X \psi_n = \lim_n \int_X \phi_n + \psi_n = \int_X f + g$$

□

**Definizione 58** (Funzione integrabile). *Una funzione misurabile  $f : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathbb{B})$  si dice integrabile se  $\int_X |f| d\mu < +\infty$ . In tal caso, data una qualsiasi coppia di funzioni positive  $g, h$  integrabili tali che  $f = g - h$ , definiamo l'integrale di  $f$  come*

$$\int_X f = \int_X g - \int_X h$$

La definizione non dipende dalla scelta della decomposizione, perché se  $g - h = g' - h'$  per additività dell'integrale delle funzioni positive

$$\int_X g d\mu + \int_X h' d\mu = \int_X (g + h') d\mu = \int_X (g' + h) d\mu = \int_X g' d\mu + \int_X h d\mu$$

E dato che tutte le funzioni hanno integrale finito è possibile fare la differenza, da cui la tesi. In generale è comodo scegliere la scomposizione  $f = f_+ - f_-$ , che soddisfa la richiesta perché  $|f_{\pm}| \leq |f|$  e dunque le due funzioni sono integrabili. Con tale scelta si mostra facilmente il seguente:

**Lemma 9.2.5.** *Una funzione misurabile è limite puntuale di una successione di funzioni semplici.*

*Dimostrazione.* Se  $\phi_n \uparrow f_+$  e  $\psi_n \uparrow f_-$ , allora  $\phi_n - \psi_n \rightarrow f$ , e la somma è ancora una funzione semplice.  $\square$

Poniamo  $\mathcal{L}^1(X, \mu) = \{f : X \rightarrow \mathbb{R} \text{ integrabili}\}$ . Riciclando i risultati dimostrati per le funzioni positive, con pochissimo sforzo ricaviamo che  $\mathcal{L}(X, \mu)$  è uno spazio vettoriale reale e l'applicazione  $\int_X$  è un suo funzionale lineare monotono rispetto all'ordine indotto da  $\mathbb{R}$ . Su tale spazio possiamo definire la seminorma  $L^1$

$$\|f\|_1 = \int_X |f| dx$$

Non è ancora una norma, dato che esistono funzioni non nulle il cui integrale è nullo, (ad esempio la funzione  $\chi_{x=0}$ ). Allora possiamo definire la relazione di equivalenza su  $\mathcal{L}^1$

$$f \sim_{L^1} g \iff \|f - g\|_1 = 0$$

Il quoziente si denota con  $L^1(X, \mu)$ . Si osservi che la seminorma  $L^1$  passa al quoziente<sup>1</sup> e definisce effettivamente una norma, detta sempre norma  $L^1$ .

Spesso è comodo sapere che

**Lemma 9.2.6.** *Sia  $f \in \mathcal{L}^1$ .  $\|f\|_1 = 0$  se e solo se  $f(x) = 0$  quasi ovunque. Detto in altri termini,  $f \sim_{L^1} g \iff f = g$  quasi ovunque.*

Si osservi che  $f$  potrebbe valere  $+\infty$  su un insieme trascurabile, irrilevante ai fini dell'integrazione.

*Dimostrazione.* Se  $f$  è nulla quasi ovunque, possiamo decomporre  $f = g - h$ , con  $g$  ed  $h$  nulle quasi ovunque. Una funzione semplice minore di  $g$  deve essere nulla su un insieme di misura piena, dunque ovviamente l'integrale di  $g$  è nullo. Ripetendo il ragionamento per  $h$  otteniamo che  $\|f\|_1 = 0$ .

Viceversa per monotonia

$$0 = \int_X |f| d\mu \geq \frac{1}{n} \int_X \chi_{\{|f| \geq \frac{1}{n}\}} = \frac{1}{n} \mu \left( \left\{ |f| \geq \frac{1}{n} \right\} \right)$$

Dunque per ogni  $n$  vale che  $\mu \left( \left\{ |f| \geq \frac{1}{n} \right\} \right) = 0$ . Allora per subadditività della misura

$$\mu(\{|f| \neq 0\}) \leq \sum_{n \in \mathbb{N}} \mu \left( \left\{ |f| \geq \frac{1}{n} \right\} \right) = 0$$

$\square$

Nella dimostrazione abbiamo usato il seguente fatto, di per sé degno di nota per quanto autoevidente.

**Lemma 9.2.7** (Disuguaglianza di Markov). *Sia  $f$  una funzione misurabile positiva. Allora per ogni  $\lambda > 0$  vale*

$$\int f d\mu \geq \lambda \mu(\{f \geq \lambda\})$$

<sup>1</sup>In realtà questo è un caso particolare della costruzione della norma quoziente rispetto al sottospazio vettoriale degli elementi su cui la seminorma è nulla. Vedasi 3.3.2

### 9.2.1 Lebesgue contro Riemann

L'integrale di Lebesgue risulta essere un'estensione dell'integrale secondo Riemann, come mostra il seguente teorema.

**Teorema 9.2.8.** *Se  $f : [a, b] \rightarrow \mathbb{R}$  è Riemann-integrabile allora  $f \in \mathcal{L}^1$  e i due integrali coincidono.*

*Dimostrazione.* Indicheremo con  $\int_*$  e  $\int^*$  gli integrali inferiori e superiori secondo Riemann, e con  $\int_X$  l'integrale di Lebesgue sull'insieme  $X$ .

Consideriamo una funzione Riemann integrabile su  $[a, b]$ , che dunque deve essere limitata per definizione. A meno di spezzarla nella somma di parte positiva e negativa possiamo supporre che  $f \geq 0$ . Data una partizione  $p$  di  $[a, b]$ , la somma inferiore  $s(f, p) = \sum_{k=1}^n |I_k| \inf_{I_k} f$  è l'integrale di Lebesgue della funzione semplice  $\phi = \sum_{k=1}^n \chi_{I_k} \inf_{I_k} f$ . Scegliamo dunque una successione di partizioni  $p_n \subseteq p_{n+1}$  tali che

$$\int_* f dx = \sup_p s(f, p) = \sup_{n \in \mathbb{N}} s(f, p_n)$$

e associamo a ogni somma inferiore la corrispondente funzione semplice  $\phi_n$ . Per costruzione tali funzioni sono minori di  $f$ , e convergono crescendo (dato che le partizioni si fanno progressivamente più fini) a una funzione  $f_* \leq f$ . Dunque per Beppo Levi (o per definizione dell'integrale)

$$\int_* f dx = \sup_{n \in \mathbb{N}} \int_{[a,b]} \phi_n = \int_{[a,b]} f_*$$

Il ragionamento si può ripetere considerando le somme superiori. In tal caso le funzioni semplici convergono decrescendo a una funzione  $f^* \geq f$ , dominate dalla prima, dunque per il ConDom

$$\int^* f dx = \int_{[a,b]} f^*$$

Se  $f$  è Riemann-integrabile, i due integrali coincidono, e dunque  $\int_{[a,b]} f^* - f_* = 0$ . Dato che l'integrando è positivo dobbiamo avere che  $f^* = f_*$  quasi ovunque, e poiché  $f_* \leq f \leq f^*$  anche  $f = f^*$  quasi ovunque. Dunque  $f$  risulta integrabile e il suo integrale di Lebesgue coincide con l'integrale di Riemann.  $\square$

La classe delle funzioni Lebesgue-misurabili è molto più ampia, dato che comprende anche funzioni molto discontinue. Si pensi alla funzione di Dirichlet, cioè  $\chi_{\mathbb{Q} \cap [0,1]}$ , che ha integrale zero secondo Lebesgue (in quanto è nulla eccetto su un insieme trascurabile<sup>2</sup>) ma che non è integrabile secondo Riemann dato che è discontinua su tutto  $[0, 1]$ .

## 9.3 Convergenza di integrali

Osserviamo innanzitutto che

**Lemma 9.3.1.** *Se  $f_n \rightarrow f$  in norma  $L^1$ , la successione degli integrali converge a  $\int_X f$ .*

*Dimostrazione.* Basta notare che per monotonia dell'integrale

$$\left| \int_X f_n - \int_X f \right| \leq \int_X |f_n - f| = \|f_n - f\|_1 \rightarrow 0$$

$\square$

**Lemma 9.3.2 (Fatou).** *Sia  $(f_n)_{n \in \mathbb{N}} : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathbb{B})$  una successione di funzioni misurabili non negative. Allora*

$$\liminf_{n \rightarrow +\infty} \int_X f_n d\mu \geq \int_X \liminf_{n \rightarrow +\infty} f_n d\mu$$

<sup>2</sup>Vedasi 8.6.3.1.

In pratica, questa proprietà è la semicontinuità inferiore dell'integrale di funzioni positive rispetto alla convergenza puntuale.<sup>3</sup> Infatti se  $f_n \rightarrow f$ , allora  $\liminf_{n \rightarrow +\infty} f_n = f$ , e dunque

$$\liminf_{n \rightarrow +\infty} \int_X f_n d\mu \geq \int_X f d\mu$$

*Dimostrazione.* Sia  $g_n = \inf_{k \geq n} f_k$ . Allora  $g_n$  è una successione di funzioni misurabili non negative che converge crescendo a  $\sup_{n \in \mathbb{N}} \inf_{k \geq n} f_k = \liminf_{n \rightarrow +\infty} f_n$ . Inoltre, se  $k \geq n$  per costruzione  $f_k \geq g_n$ , e dunque  $\int_X f_k \geq \int_X g_n$ . Perciò

$$\inf_{k \geq n} \int_X f_k \geq \int_X g_n$$

Passando al sup su  $n$  e applicando il teorema di Beppo Levi alle  $g_n$  otteniamo

$$\liminf_{n \rightarrow +\infty} \int_X f_n = \sup_{n \in \mathbb{N}} \inf_{k \geq n} \int_X f_k \geq \sup_{n \in \mathbb{N}} \int_X g_n = \int_X \sup_{n \in \mathbb{N}} g_n = \int_X \liminf_{n \rightarrow +\infty} f_n$$

□

In generale non vale l'uguaglianza. Ad esempio, scegliendo  $f_n = \chi_{[n, n+1]}$ , l'integrale degli elementi della successione è costantemente 1 ma il limite puntuale delle  $f_n$  è identicamente nullo. Questo ci dice in particolare che la convergenza puntuale non implica la convergenza degli integrali, e dunque nemmeno la convergenza in norma  $L^1$  per quanto ci dice il lemma 9.3.1.

D'altra parte, non si può nemmeno dire che la convergenza  $L^1$  implichi la convergenza puntuale quasi ovunque. Consideriamo infatti la successione di intervalli  $I_0 = [0, 1]$ ,  $I_1 = [0, 1/2]$ ,  $I_2 = [1/2, 1]$ ,  $I_3 = [0, 1/3]$  ... e definiamo  $f_n = \chi_{I_n}$ . Poiché l'ampiezza degli intervalli tende a zero le funzioni convergono a zero in norma  $L^1$ . Tuttavia non si ha convergenza in nessun punto, perché ogni  $x \in [0, 1]$  appartiene ad almeno uno degli intervalli della suddivisione in  $n$  parti (dunque  $f_n(x) = 1$  frequentemente), e non può appartenere a più di due di essi (perciò frequentemente  $f_n(x) = 0$ ).

Ora tiriamo fuori un "grosso" teorema che abbiamo già visto in passato.

**Teorema 9.3.3** (ConDom di Lebesgue). *Sia  $(f_n)_{n \in \mathbb{N}}$  una successione di funzioni integrabili convergenti puntualmente a  $f$ . Se esiste una funzione  $g$  integrabile e positiva che domina la successione, ossia che  $|f_n| \leq g$  per ogni  $n$ , allora  $f_n \rightarrow f$  in norma  $L^1$ , e dunque in particolare*

$$\lim_{n \rightarrow +\infty} \int_X f_n = \int_X f$$

*Dimostrazione.* Applichiamo il lemma di Fatou alle due successioni  $g - f_n$  e  $g + f_n$ , che sono non negative e misurabili per ogni  $n$ . Dunque

$$\begin{aligned} \int_X g + \liminf_{n \rightarrow +\infty} \int_X f_n &= \liminf_{n \rightarrow +\infty} \int_X g + f_n \geq \int_X g + f = \int_X g + \int_X f \\ \int_X g - \limsup_{n \rightarrow +\infty} \int_X f_n &= \liminf_{n \rightarrow +\infty} \int_X g - f_n \geq \int_X g - f = \int_X g - \int_X f \end{aligned}$$

Allora, sottraendo l'integrale finito di  $g$ , otteniamo

$$\limsup_{n \rightarrow +\infty} \int_X f_n \leq \int_X f \leq \liminf_{n \rightarrow +\infty} \int_X f_n$$

Da cui la convergenza degli integrali. Se applichiamo quanto appena ottenuto alla successione  $|f - f_n| \leq 2g$  che converge puntualmente a zero otteniamo

$$\|f - f_n\|_1 = \int_X |f - f_n| \rightarrow 0$$

Dunque la convergenza è in norma  $L^1$ .

□

<sup>3</sup>Vedasi la definizione di semicontinuità data in 1.2.

Possiamo leggermente sgonfiare le ipotesi del risultato precedente:

**Corollario 9.3.3.1** (ConDom bucherellato). *Sia  $f_n$  una successione di funzioni integrabili convergenti puntualmente quasi ovunque a una funzione  $f$  e dominate quasi ovunque da una funzione integrabile positiva  $g$ . Allora  $f_n \rightarrow f$  in norma  $L^1$ .*

*Dimostrazione.* Sia  $A \subseteq X$  l'insieme di misura piena su cui valgono sia la convergenza sia la dominazione. Per il lemma 9.2.6 possiamo sostituire ogni funzione  $h$  che compare nell'enunciato con  $h\chi_A$ , che coincide con  $h$  quasi ovunque. In tal caso valgono le ipotesi del teorema di Lebesgue, dunque si ha la tesi.  $\square$

### 9.3.1 Integrazione per serie

**Teorema 9.3.4.** *Sia  $(f_n)_{n \in \mathbb{N}}$  una successione di funzioni integrabili, tali che  $\sum_{n \in \mathbb{N}} \|f_n\|_1 < +\infty$ . Allora la serie  $\sum_{n \in \mathbb{N}} f_n$  converge, sia in norma 1 sia quasi ovunque, ed è dominata quasi ovunque da una funzione integrabile. In particolare*

$$\int_X \sum_{n \in \mathbb{N}} f_n = \sum_{n \in \mathbb{N}} \int_X f_n$$

**Corollario 9.3.4.1.**  $L^1(X, \mu)$  è uno spazio di Banach, per criterio di convergenza totale.

*Dimostrazione.* Sia  $g(x) = \sum_{n \in \mathbb{N}} |f_n(x)|$ .  $g$  è una funzione (eventualmente a valori nella retta estesa) misurabile e non negativa. Inoltre per Beppo Levi

$$\int_X g(x) = \sum_{n \in \mathbb{N}} \int_X |f_n(x)| = \sum_{n \in \mathbb{N}} \|f_n\|_1 < +\infty$$

Allora  $g(x)$  è finita quasi ovunque, o il suo integrale divergerebbe. In altri termini la serie  $\sum_{n \in \mathbb{N}} f_n(x)$  converge assolutamente quasi ovunque a una certa  $F$ . Ora applichiamo il teorema del ConDom alla successione delle somme parziali, dominate quasi ovunque da  $g$ , ottenendo la convergenza in norma  $L^1$ .  $\square$

Possiamo mostrare una sorta di inverso del teorema del ConDom:

**Teorema 9.3.5.** *Sia  $(f_n)_{n \in \mathbb{N}}$  una successione di funzioni integrabili convergente in  $L^1$  a  $f$ . Allora esiste una sottosuccessione dominata che converge q.o. a  $f$ .*

*Dimostrazione.* Per ipotesi  $\|f_n - f\| \rightarrow 0$ , dunque possiamo estrarre una sottosuccessione  $f_{n(k)}$  per cui  $\|f_{n(k)} - f\| \leq 2^{-k}$ . Allora

$$\|f_{n(k)} - f_{n(k+1)}\| \leq \|f_{n(k)} - f\| + \|f_{n(k+1)} - f\| \leq 2^{-(k-1)}$$

Dunque la serie  $\sum_{k=0}^{\infty} (f_{n(k+1)} - f_{n(k)})$  converge normalmente, e per quanto visto nel teorema precedente converge quasi ovunque a  $f$  ed è dominata quasi ovunque dalla serie  $|f_{n(0)}| + \sum_{k=0}^{\infty} |f_{n(k+1)} - f_{n(k)}|$ .  $\square$

In generale non possiamo richiedere nulla di più, in quanto esiste una successione di funzioni convergenti a zero in  $L^1$  e puntualmente ma che non può essere dominata. Infatti, poniamo  $x_0 = 1$  e definiamo induttivamente  $x_{n+1}$  in modo che  $\int_{x_{n+1}}^{x_n} \frac{dx}{x} = \frac{1}{n}$ , cosa che è sempre possibile dato che l'integrale di  $\frac{1}{x}$  diverge. Poniamo dunque  $f_n = \frac{1}{x} \chi_{[x_{n-1}, x_n]}$ . Per costruzione  $\int_{[0,1]} f dx = \frac{1}{n}$ , e ovviamente la successione converge puntualmente a zero. Tuttavia, ogni funzione  $g$  positiva che maggiora la successione deve maggiorare anche

$$\sup f_n = \begin{cases} \frac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

Dunque l'integrale di  $g$  deve divergere.

Un'altra conseguenza del teorema 9.3.5 è che la convergenza quasi ovunque non può essere indotta da alcuna topologia. Infatti, consideriamo una successione di funzioni convergenti a zero in  $L^1$  ma non quasi ovunque (ad esempio, le indicatrici degli intervalli  $I_0 = [0, 1]$ ,  $I_1 = [0, 1/2]$ ,  $I_2 = [1/2, 1]$ ,  $I_3 = [0, 1/3]$ ... di cui abbiamo parlato nella sezione precedente). Ogni sottosuccessione converge a zero in  $L^1$ , dunque per il teorema è possibile estrarre una sotto-sottosuccessione convergente a zero quasi ovunque. Questo significa che esiste una successione con la proprietà di Urysohn<sup>4</sup> ma non convergente, il che esclude la possibilità che la convergenza sia data da una topologia.

<sup>4</sup>Vedasi 1.3.1.

## 9.4 Integrali doppi

Chiameremo “integrale sequenziale” una scrittura del tipo

$$\int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x)$$

I prossimi teoremi definiranno il legame tra gli integrali sequenziali e l'integrale sullo spazio prodotto.

**Teorema 9.4.1** (Tonelli). *Siano  $(X, \mathcal{A}, \mu)$  e  $(Y, \mathcal{B}, \nu)$  spazi di misura  $\sigma$ -finiti. Sia  $f : X \times Y \rightarrow [0, +\infty]$  misurabile rispetto all'algebra prodotto. Allora*

1. La mappa  $x \rightarrow \int_Y f(x, y) d\nu(y)$  è misurabile e positiva, e dunque l'integrale sequenziale è ben definito.
2. Vale che

$$\int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_{X \times Y} f(x, y) d\mu \otimes \nu(x, y)$$

**Corollario 9.4.1.1.** *Nelle ipotesi del teorema 9.4.1 l'ordine di integrazione è indifferente, ossia*

$$\int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y)$$

*Dimostrazione.* Innanzitutto osserviamo che l'integrando del punto 1. è misurabile come funzione su  $Y$ , in quanto è ottenuta componendo le funzioni misurabili  $f \circ j_x$ , e dunque l'integrale ha senso.

Assumiamo ora che  $\mu$  e  $\nu$  siano finite. Consideriamo le funzioni

$$\mathcal{F} = \{f : X \times Y \rightarrow [0, +\infty] \text{ per cui valgono 1. e 2.}\}$$

Ora, l'insieme  $\{E \in \mathcal{A} \otimes \mathcal{B} \text{ t.c. } \chi_E \in \mathcal{F}\}$  è un sistema di Dynkin. Infatti:

- Ovviamente  $\chi_\emptyset \in \mathcal{F}$ .
- Se  $\chi_E \in \mathcal{F}$ , allora  $\chi_{E^c} = 1 - \chi_E$ , dunque sicuramente la funzione che manda  $x$  nell'integrale  $\int_Y \chi_{E^c}(x, y) d\nu(y) = \nu(Y) - \int_Y \chi_E(x, y) d\nu(y)$  è misurabile e positiva e il suo integrale coincide con l'integrale sul prodotto. Si osservi che qui abbiamo usato l'ipotesi di finitezza di  $\nu$ .
- Se  $\chi_{E_n} \in \mathcal{F}$  per una famiglia numerabile disgiunta  $(E_n)_{n \in \mathbb{N}} \in \mathcal{A} \otimes \mathcal{B}$ , dato che  $\chi_{\sqcup_n E_n} = \sum_n \chi_{E_n}$  applicando Beppo Levi possiamo scambiare serie e integrale più o meno sempre:

$$\begin{aligned} \int_{X \times Y} \sum_n \chi_{E_n}(x, y) d\mu \otimes \nu(x, y) &= \sum_n \int_{X \times Y} \chi_{E_n}(x, y) d\mu \otimes \nu(x, y) = \sum_n \int_X \left( \int_Y \chi_{E_n}(x, y) d\nu(y) \right) d\mu(x) \\ &= \int_X \sum_n \left( \int_Y \chi_{E_n}(x, y) d\nu(y) \right) d\mu(x) \end{aligned}$$

Dunque  $\chi_{\sqcup_n E_n} \in \mathcal{F}$ .

Ora, tale  $\lambda$ -sistema contiene il  $\pi$ -sistema dei prodotti. Infatti, per ogni  $A \in \mathcal{A}$ , per ogni  $B \in \mathcal{B}$  abbiamo  $\chi_{A \times B} \in \mathcal{F}$ , dato che  $\chi_{A \times B}(x, y) = \chi_A(x) \chi_B(y)$ , dunque

$$\int_X \left( \int_Y \chi_A(x) \chi_B(y) d\nu(y) \right) d\mu(x) = \int_X \nu(B) \chi_A(x) d\mu(x) = \mu(A) \nu(B) = \int_{X \times Y} \chi_{A \times B}(x, y) d\mu \otimes \nu(x, y)$$

Dunque per il lemma di Dynkin  $\chi_E \in \mathcal{F}$  per ogni  $E \in \mathcal{A} \otimes \mathcal{B}$ . Ora, per linearità la tesi è vera per le funzioni semplici. Inoltre applicando ripetutamente la proprietà di Beppo Levi si vede che, se  $\phi_n \uparrow f$ , allora  $f \in \mathcal{F}$ . Dunque ogni funzione misurabile e positiva appartiene a  $\mathcal{F}$ , da cui la tesi.

Consideriamo ora il caso generale in cui  $X$  e  $Y$ , e dunque anche il loro prodotto, sono  $\sigma$ -finiti. Siano  $X_n \uparrow X$  e  $Y_n \uparrow Y$  finiti. Dunque per ogni  $m$  e  $n$  la tesi è vera per  $f|_{X_m \times Y_n}$  con la misura ristretta. Allora, applicando compulsivamente Beppo Levi otteniamo

$$\int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_X \lim_n \left( \int_{Y_n} f(x, y) d\nu(y) \right) d\mu(x) = \lim_n \int_X \left( \int_{Y_n} f(x, y) d\nu(y) \right) d\mu(x) =$$

$$= \lim_n \lim_m \int_{X_m} \left( \int_{Y_n} f(x, y) d\nu(y) \right) d\mu(x) = \lim_n \lim_m \int_{X_m \times Y_n} f(x, y) d\mu \otimes \nu(x, y) = \int_{X \times Y} f(x, y) d\mu \otimes \nu(x, y)$$

□

L'ipotesi di  $\sigma$ -finitezza è necessaria, come mostra il seguente controesempio. Siano  $X = Y = [0, 1]$  con i boreliani e siano  $\mu = \lambda$  e  $\nu$  la misura contapunti (che associa a un insieme la sua cardinalità, eventualmente infinita), ovviamente non  $\sigma$ -finita. Sia  $\Delta = \{(x, x), x \in [0, 1]\}$  la diagonale. Allora

$$\int_X \left( \int_Y \chi_\Delta(x, y) d\nu(y) \right) dx = \int_X dx = 1$$

Mentre

$$\int_Y \left( \int_X \chi_\Delta(x, y) dx \right) d\nu(y) = 0$$

**Teorema 9.4.2** (Fubini). *Siano  $(X, \mathcal{A}, \mu)$  e  $(Y, \mathcal{B}, \nu)$  spazi di misura  $\sigma$ -finiti. Sia  $f \in L^1(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ . Allora*

1. *La mappa  $x \rightarrow \int_Y f(x, y) d\nu(y)$  è definita  $\mu$ -quasi ovunque;*

2. *Vale che*

$$\int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_{X \times Y} f(x, y) d\mu \otimes \nu(x, y)$$

*Dimostrazione.* Sia  $f = f_+ - f_-$ . Per il teorema di Tonelli la tesi è vera per  $f_+$ , ed essendo  $f$  integrabile abbiamo che

$$\int_X \left( \int_Y f_+(x, y) d\nu(y) \right) d\mu(x) = \int_{X \times Y} f_+(x, y) d\mu \otimes \nu(x, y) < +\infty$$

Ma allora  $\int_Y f_+(x, y) d\nu(y) < +\infty$  per  $\mu$ -quasi ogni  $x \in X$ . Perciò poniamo

$$g_+(x) = \chi_{(\int_Y f_+(x, y) d\nu(y) < +\infty)}(x) \int_Y f_+(x, y) d\nu(y)$$

Tale mappa è misurabile perché è il prodotto della funzione integrale (misurabile per il teorema di Tonelli) per la caratteristica di un insieme misurabile. Possiamo fare lo stesso discorso per  $f_-$  e concludere l'esistenza di una certa  $g_-$  misurabile e positiva che soddisfa l'uguaglianza

$$\int_Y f(x, y) d\nu(y) = g_+(x) - g_-(x) \quad \mu\text{-q.o.}$$

Abbiamo dunque ottenuto il primo punto della tesi, e il secondo segue applicando il teorema di Tonelli a  $g_+$  e  $g_-$  e sommando i risultati. □

Risulta necessario richiedere che  $f$  sia integrabile, in quanto si possono costruire funzioni il cui integrale sequenziale dipende dall'ordine di integrazione. Consideriamo ad esempio  $X = Y = [0, 1]$  con i boreliani e siano  $\mu = \nu = \lambda$ . Sia

$$f(x, y) = \operatorname{sgn}(x - y)(\max(x, y))^{-2}$$

Essendo la  $f$  antisimmetrica negli argomenti,

$$\int_{[0,1]} \left( \int_{[0,1]} f dx \right) dy = - \int_{[0,1]} \left( \int_{[0,1]} f dy \right) dx$$

Perciò gli integrali differiscono se uno dei due è non nullo (e dunque entrambi lo sono). Se  $x \neq 0$  abbiamo che

$$\int_0^1 \operatorname{sgn}(x - y)(\max(x, y))^{-2} dy = \int_0^x x^{-2} dy - \int_x^1 y^{-2} dy = x^{-1} + 1 - x^{-1} = 1$$

Da cui, a meno di escludere l'insieme  $\{0\}$  di misura nulla

$$\int_{(0,1]} \left( \int_{[0,1]} f dy \right) dx = \int_{(0,1]} dx = 1 \neq 0$$

### 9.4.1 Insiemi trascurabili nel prodotto

**Lemma 9.4.3** (Principio di Cavalieri). *Siano  $(X, \mathcal{A}, \mu)$  e  $(Y, \mathcal{B}, \nu)$  spazi di misura  $\sigma$ -finiti e sia  $N \in \mathcal{A} \otimes \mathcal{B}$ . Allora*

$$\mu \otimes \nu(N) = \int_{X \times Y} \chi_N(x, y) \mu \otimes \nu(x, y) = \int_X \left( \int_Y \chi_N d\nu(y) \right) d\mu(x) = \int_X \nu(j_x^{-1}(N)) d\mu(x)$$

La dimostrazione è una banale applicazione del teorema di Tonelli.

**Corollario 9.4.3.1.** *Nelle ipotesi del lemma 9.4.3  $N$  è trascurabile se e solo se l'insieme*

$$\{x \in X \text{ t.c. } \nu(j_x^{-1}(N)) > 0\}$$

*è  $\mu$ -trascurabile.*

In altri termini, le sezioni di un insieme di misura nulla nel prodotto sono quasi tutte trascurabili nei fattori.

### 9.4.2 Affettare gli integrali

Ho fatto una decina di fette, che faccio lascio?

---

*Un salumiere semplice*

Sia  $f$  misurabile e positiva sullo spazio di misura  $X$   $\sigma$ -finito. Dato che ovviamente per ogni  $c \in \mathbb{R}^+$  si ha  $c = \int_0^{+\infty} \chi_{[0,c]} dx$ , possiamo scrivere l'integrale di  $f$  come

$$\int_X f(x) d\mu(x) = \int_X \left( \int_0^{+\infty} \chi_{[0,f(x)]}(t) dt \right) d\mu$$

Applicando il teorema di Tonelli possiamo scambiare l'ordine di integrazione, ottenendo

$$\int_X f(x) d\mu(x) = \int_0^{+\infty} \left( \int_X \chi_{[0,f(x)]}(t) d\mu(x) \right) dt = \int_0^{+\infty} \left( \int_X \chi_{[t,+\infty]}(f(x)) d\mu(x) \right) dt = \int_0^{+\infty} \mu(\{f(x) \geq t\}) dt$$

Questo giustifica l'immagine popolare secondo la quale l'integrale di Lebesgue viene fatto integrando "a strisce orizzontali".

## 9.5 Densità in $L^1$

Cosa gli ha cazzo preso?

---

*Pmajer si scaglia contro la Jamboard, che conduce l'incontro.*

In questa sottosezione parleremo di un paio di classi di funzioni "maneggevoli" che sono dense in  $L^1$ . Si noti che lavoreremo nello spazio quoziente, quindi parleremo sempre di classi di equivalenza e non di vere e proprie funzioni.

**Teorema 9.5.1.** *Sia  $(X, \mathcal{A}, \mu)$  uno spazio di misura. Le classi di equivalenza delle funzioni semplici sono dense in  $L^1(X)$ .*

Per dimostrare il teorema è sufficiente provare che

**Lemma 9.5.2.** *Sia  $\phi_n$  una successione crescente di funzioni semplici positive, convergente a  $f$ . Se  $f$  è integrabile la convergenza è in norma  $L^1$ .*

Se il lemma è vero, basta scrivere  $f = f_+ - f_-$  e sommare le due successioni di funzioni semplici.

*Dimostrazione.* Per teorema di Beppo Levi,  $\int_X \phi_n d\mu \uparrow \int_X f d\mu < +\infty$ , dunque

$$\|f - \phi_n\|_1 = \int_X (f - \phi_n) d\mu = \int_X f d\mu - \int_X \phi_n d\mu \rightarrow 0$$

□

Nel caso reale è possibile dire di più:

**Teorema 9.5.3.** *Le funzioni semplici ottenute dalle funzioni caratteristiche di pluri-intervalli sono dense in  $L^1(\mathbb{R}^n)$ . In simboli  $\bar{V} = L^1(\mathbb{R}^n)$ , dove*

$$V = \text{Span} \left\{ [\chi_I] \text{ t.c. } I = \prod_{i=1}^n [a_i, b_i] \right\}$$

*Dimostrazione.* Mostriamo innanzitutto che  $\bar{V}$  contiene le funzioni caratteristiche degli aperti di misura finita. Infatti ognuno di questi si può scrivere come unione numerabile di pluri-intervalli aperti  $I_k$  perché la topologia euclidea è secondo-numerabile. Ora, sostituiamo progressivamente  $I_k$  con  $J_k = I_k \setminus \bigcap_{i=1}^{k-1} I_i$ , in modo che l'unione risulti disgiunta, e ogni volta che sottraiamo un elemento scriviamo la differenza come somma finita di pluri-intervalli (non per forza tutti aperti). Al termine del processo si sarà ottenuta una famiglia numerabile disgiunta di pluri-intervalli la cui unione è  $A$ . Dunque la somma delle funzioni caratteristiche di tali pluri-intervalli converge in norma  $L^1$  alla caratteristica di  $A$ , se quest'ultimo ha misura finita.

Ora, se  $A$  è un insieme Lebesgue-misurabile di misura finita, esiste una successione di aperti di misura finita (WLOG innestati) che ricopre  $A$  e la cui intersezione coincide con  $A$  a meno di un insieme di misura nulla.<sup>5</sup> Dunque le classi delle funzioni caratteristiche di tali aperti convergono in norma  $L^1$  alla classe della funzione caratteristica di  $A$ , e perciò  $\bar{V}$  contiene le classi di tutte le funzioni caratteristiche.

Per linearità le classi delle funzioni semplici sono contenuti in  $\bar{V}$ , dunque si conclude per il lemma 9.5.2. □

Affinando il teorema si mostra che

**Corollario 9.5.3.1.**  *$L^1(\mathbb{R}^n)$  è separabile. Più precisamente, l'insieme*

$$V = \text{Span}_{\mathbb{Q}} \left\{ [\chi_I] \text{ t.c. } I = \prod_{i=1}^n [a_i, b_i], a_i, b_i \in \mathbb{Q} \right\}$$

*è numerabile e denso in  $L^1(\mathbb{R}^n)$ .*

*Dimostrazione.* Basta osservare che ogni pluri-intervallo si approssima a piacere con pluri-intervalli di estremi razionali, e che ogni funzione semplice è approssimata da combinazioni lineari a coefficienti razionali. □

Si può verificare con metodi simili che gli spazi  $L^p$  sono misurabili per  $p \neq +\infty$ . In quest'ultimo caso l'asserto falso, dato che è possibile determinare un sottoinsieme non numerabile di funzioni la cui distanza reciproca è sempre 1 (ad esempio, le funzioni caratteristiche dei pluri-intervalli).

## 9.6 Continuità uniforme in $L^1$

**Lemma 9.6.1.** *Sia  $f \in \mathcal{L}^1(\mathbb{R}^n)$ . Allora*

$$\lim_{h \rightarrow 0} \|f(x) - f(x-h)\|_1 = 0$$

*Dimostrazione.* Sia  $\mathcal{F}$  il sottospazio vettoriale delle funzioni  $f \in \mathcal{L}^1(\mathbb{R}^n)$  per cui vale la tesi. Se mostriamo che contiene le funzioni caratteristiche di pluri-intervalli e che è chiuso possiamo concludere che coincide con  $\mathcal{L}^1(\mathbb{R}^n)$  per il teorema 9.5.3. La prima affermazione è vera perché, se  $I = \prod_{i=1}^n [a_i, b_i]$ , consideriamo l'ispessimento della frontiera  $J(h) = \prod_{i=1}^n [a_i - \|h\|, b_i + \|h\|] \setminus \prod_{i=1}^n [a_i + \|h\|, b_i - \|h\|]$ . Allora

$$|\chi_I(x) - \chi_I(x-h)| = |\chi_I(x) - \chi_{I+h}(x)| = \chi_{I \Delta (I+h)}(x) \leq \chi_{J(h)}(x)$$

<sup>5</sup>Vedasi 8.6.4.

Da cui

$$\|\chi_I(x) - \chi_I(x-h)\|_1 \leq \prod_{i=1}^n (b_i - a_i + 2\|h\|) - \prod_{i=1}^n (b_i - a_i - 2\|h\|)$$

che tende a zero per  $h \rightarrow 0$ .

Per mostrare la chiusura sia  $f \in \overline{\mathcal{F}}$ . Per ogni  $g \in \mathcal{F}$  si ha

$$\|f(x) - f(x-h)\|_1 \leq \|f(x) - g(x)\|_1 + \|g(x) - g(x-h)\|_1 + \|g(x-h) - f(x-h)\|_1 \leq 2\|f(x) - g(x)\|_1 + \|g(x) - g(x-h)\|_1$$

Dunque

$$\limsup_{h \rightarrow 0} \|f(x) - f(x-h)\|_1 \leq 2\|f(x) - g(x)\|_1$$

e prendendo l'estremo inferiore su  $g$

$$\limsup_{h \rightarrow 0} \|f(x) - f(x-h)\|_1 \leq \inf_{g \in \mathcal{F}} 2\|f(x) - g(x)\|_1 = 0$$

Allora  $f \in \mathcal{F}$ , cioè  $\mathcal{F}$  coincide con la propria chiusura.  $\square$

Il teorema precedente dà senso alla seguente definizione:

**Definizione 59.** Sia  $\omega : [0, +\infty] \rightarrow [0, +\infty]$  una funzione crescente, continua nell'origine e ivi nulla.  $\omega$  è un modulo di continuità  $L^1$  per  $f \in \mathcal{L}^1(\mathbb{R}^n)$  se per ogni  $h \in \mathbb{R}^n$

$$\|f(x) - f(x-h)\|_1 \leq \omega(\|h\|)$$

Come modulo si può scegliere sempre

$$\omega(\delta) = \sup_{\|h\| \leq \delta} \|f(x) - f(x-h)\|_1$$

Con ragionamenti simili è possibile definire un modulo di continuità  $L^p$  per ogni  $p \in [1, +\infty]$ . Si osservi che il modulo  $L^\infty$  coincide con il modulo di continuità classico.

Un'applicazione del teorema precedente fornisce una caratterizzazione degli insiemi di misura non nulla.

**Teorema 9.6.2** (Steinhaus). Sia  $E \subseteq \mathbb{R}^n$  misurabile secondo Lebesgue e di misura positiva. Allora

$$E - E = \{x - y \text{ t.c. } x, y \in E\}$$

è un intorno dell'origine.

*Dimostrazione.* Per il lemma 9.6.1  $\lambda(E \Delta (E+h)) = \|\chi_E - \chi_{E+h}\|_1 \rightarrow 0$  per  $h \rightarrow 0$ . Dunque esiste un  $\delta > 0$  tale che per ogni  $h \in B(0, \delta)$

$$\lambda(E) > \lambda(E \Delta (E+h)) \geq \lambda(E \setminus (E+h))$$

Questo in particolare significa che  $\lambda(E \cap (E+h)) = \lambda(E) - \lambda(E \setminus (E+h)) > 0$ , e perciò l'intersezione è non vuota. Allora esiste  $a \in E \cap (E+h)$ , ossia  $a \in E$  e  $a-h \in E$ . Dunque per ogni  $h \in B(0, \delta)$   $h \in E - E$ , da cui la tesi.  $\square$

### 9.6.1 Approssimazioni continue

Sia  $f \in \mathcal{L}^1(\mathbb{R}^n)$  e fissiamo un  $\varepsilon > 0$ . Consideriamo la funzione che associa a ogni punto  $x$  la media integrale di  $f$  sulla palla di centro  $x$  e raggio  $\varepsilon$

$$\bar{f}_\varepsilon = \frac{1}{\lambda(B(x, \varepsilon))} \int_{B(x, \varepsilon)} f(y) dy$$

**Lemma 9.6.3.**  $\bar{f}_\varepsilon$  è uniformemente continua, di modulo  $\frac{1}{\lambda(B(x, \varepsilon))} \omega$  dove  $\omega$  è il modulo di continuità  $L^1$  di  $f$ .

*Dimostrazione.* Dati  $x, x' \in \mathbb{R}^n$ , si ha che

$$|\bar{f}_\varepsilon(x) - \bar{f}_\varepsilon(x')| = \frac{1}{\lambda(B(0, \varepsilon))} \left| \int_{B(0, \varepsilon)} f(x+h) - f(x'+h) dh \right|$$

Facciamo il cambio di variabile<sup>6</sup>  $x+h = u$  e allarghiamo il dominio di integrazione a tutto  $\mathbb{R}^n$ , ottenendo che

$$|\bar{f}_\varepsilon(x) - \bar{f}_\varepsilon(x')| \leq \frac{1}{\lambda(B(0, \varepsilon))} \int_{\mathbb{R}^n} |f(u) - f(u + (x' - x))| du \leq \frac{1}{\lambda(B(0, \varepsilon))} \omega(\|x - x'\|)$$

Da cui la tesi.  $\square$

Si osservi che il modulo di continuità diverge per  $\varepsilon \rightarrow 0$ . Questo perché le funzioni  $\bar{f}_\varepsilon$  non possono essere equicontinue al variare di  $\varepsilon$ , o per teorema di Ascoli-Arzelà una loro sottosuccessione convergerebbe uniformemente e dunque  $f$  dovrebbe essere continua.

Un'altra proprietà è la seguente:

**Lemma 9.6.4.**  $\bar{f}_\varepsilon$  tende a zero all'infinito.

*Dimostrazione.* Per ogni  $x \in \mathbb{R}^n$  sia  $R = \|x\| - 2\varepsilon$ . Per tale scelta

$$|\bar{f}_\varepsilon(x)| = \frac{1}{\lambda(B(0, \varepsilon))} \left| \int_{B(x, \varepsilon)} f dx \right| \leq \frac{1}{\lambda(B(0, \varepsilon))} \int_{\mathbb{R}^n \setminus B(0, R)} |f|$$

Per il teorema di convergenza dominata

$$\int_{\mathbb{R}^n \setminus B(0, R)} |f| = \int_{\mathbb{R}^n} |f| \chi_{\mathbb{R}^n \setminus B(0, R)} \rightarrow 0$$

per  $R \rightarrow +\infty$ , in quanto l'integrando è dominato dalla funzione integrabile  $|f|$ . Allora se  $\|x\| \rightarrow +\infty$   $\bar{f}_\varepsilon(x)$  tende a zero.  $\square$

**Lemma 9.6.5.**  $\|\bar{f}_\varepsilon\|_1 \leq \|f\|_1$ , e se  $f$  è non negativa si ha l'uguaglianza.

*Dimostrazione.* Scriviamo

$$\int_{\mathbb{R}^n} |\bar{f}_\varepsilon| dx = \int_{\mathbb{R}^n} \frac{1}{\lambda(B(0, \varepsilon))} \left| \int_{\mathbb{R}^n} f(y) \chi_{B(0, \varepsilon)}(x-y) dy \right| dx \leq \frac{1}{\lambda(B(0, \varepsilon))} \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(y)| \chi_{B(0, \varepsilon)}(x-y) dy \right) dx$$

Per il teorema di Tonelli possiamo scambiare l'ordine degli integrali, ottenendo

$$\int_{\mathbb{R}^n} |\bar{f}_\varepsilon| dx \leq \frac{1}{\lambda(B(0, \varepsilon))} \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \chi_{B(0, \varepsilon)}(x-y) dx \right) |f(y)| dy = \int_{\mathbb{R}^n} |f(y)| dy = \|f\|_1$$

$\square$

**Teorema 9.6.6** (Densità delle funzioni continue). *Le (classi di equivalenza delle) funzioni uniformemente continue sono dense in  $L^1(\mathbb{R}^n)$ . Più precisamente, per ogni  $f$  integrabile la successione  $\bar{f}_\varepsilon$  converge a  $f$  in norma  $L^1$  per  $\varepsilon \rightarrow 0$ .*

*Dimostrazione.*  $\mathcal{F} = \{f \in \mathcal{L}^1(\mathbb{R}^n) \text{ t.c. } \bar{f}_\varepsilon \rightarrow f\}$  è un sottospazio vettoriale di  $\mathcal{L}^1$  per ovvi motivi. Inoltre contiene le funzioni caratteristiche dei pluri-intervalli, perché se  $f = \chi_I$  dove  $I = \prod_{i=1}^n [a_i, b_i]$  si ha che  $|\chi_I(x) - \bar{\chi}_{I_\varepsilon}(x)| = 0$  se  $d(x, \partial I) > \varepsilon$ . Infatti, se un tale  $x$  si trova all'esterno la palla di centro  $x$  e raggio  $\varepsilon$  non interseca il plurirettangolo, e quindi l'integrale mediato è nullo, mentre se  $x$  è all'interno la palla è totalmente contenuta nel pluriintervallo, dunque la media fa 1 come il valore della caratteristica. Inoltre,  $0 \leq \bar{\chi}_{I_\varepsilon}(x) \leq 1$ , perché è la misura di un sottoinsieme della palla  $B(x, \varepsilon)$  fratto la misura di quest'ultima. Allora

$$\|\chi_I - \bar{\chi}_{I_\varepsilon}(x)\|_1 \leq \int_{\mathbb{R}^n} |\chi_I(x)| + |\bar{\chi}_{I_\varepsilon}(x)| \leq 2 \lambda(\{x \text{ t.c. } d(x, \partial I) \leq \varepsilon\}) = o(1)$$

<sup>6</sup>Questo cambio di variabile è completamente innocuo, come vedremo in 9.8.4.

Mostriamo che  $\mathcal{F}$  è chiuso. Sia  $f \in \overline{\mathcal{F}}$  e  $g \in \mathcal{F}$ . Allora

$$\|f - \overline{f}_\varepsilon\|_1 \leq \|f - g\|_1 + \|g - \overline{g}_\varepsilon\|_1 + \|\overline{g}_\varepsilon - \overline{f}_\varepsilon\|_1$$

Che per linearità è uguale a

$$\|f - \overline{f}_\varepsilon\|_1 \leq \|f - g\|_1 + \|g - \overline{g}_\varepsilon\|_1 + \|(\overline{g - f})_\varepsilon\|_1$$

Per quanto visto nel lemma 9.6.5

$$\|f - \overline{f}_\varepsilon\|_1 \leq 2\|f - g\|_1 + \|g - \overline{g}_\varepsilon\|_1$$

Si conclude passando all'estremo inferiore per  $g \in \mathcal{F}$  e poi prendendo il limite superiore per  $\varepsilon \rightarrow 0$ .  $\square$

Il teorema precedente, combinato con il teorema 9.3.5, ci dice che esiste una sottosuccessione delle funzioni  $\overline{f}_\varepsilon$  che converge a  $f$  quasi ovunque. In altri termini

**Corollario 9.6.6.1.** *A meno di sottosuccessioni, per quasi ogni  $x \in \mathbb{R}^n$  la successione delle medie integrali sulle palle di raggio  $\varepsilon$  e centro  $x$  converge a  $f(x)$*

In realtà con molto più dispendio energetico si può mostrare che

**Teorema 9.6.7** (Differenziazione di Lebesgue).  $\overline{f}_\varepsilon \rightarrow f$  quasi ovunque. In altri termini per quasi ogni  $x \in \mathbb{R}^n$

$$\frac{1}{\lambda(B(x, \varepsilon))} \int_{B(x, \varepsilon)} f(y) dy \rightarrow f(x)$$

per  $\varepsilon \rightarrow 0$ .

In dimensione 1 questo fatto è l'analogo del teorema fondamentale del calcolo integrale, perché si sta dicendo che per quasi ogni  $x$

$$\lim_{\varepsilon \rightarrow 0} \frac{\int_{x-\varepsilon}^{x+\varepsilon} f(x) dx}{2\varepsilon} = f(x)$$

Si noti che “la derivata dell'integrale” è definita quasi ovunque (e non potevamo aspettarci di meglio, dopotutto stiamo lavorando con classi di equivalenza di funzioni a meno di insiemi trascurabili).

## 9.7 Compattezza in $L^1$ (incompleto)

Sia  $f \in \mathcal{L}^1(\mathbb{R}^n)$ . Per quanto risulta dalla dimostrazione del lemma 9.6.4 si ha che, per  $R \rightarrow +\infty$ ,

$$\int_{\mathbb{R}^n \setminus B(0, R)} |f| dx \rightarrow 0$$

Allora possiamo trovare una funzione  $\varepsilon : [0, +\infty) \rightarrow [0, +\infty)$  infinitesima all'infinito tale che  $\int_{B(0, R)^c} |f| dx \leq \varepsilon(R)$ , che con molta fantasia chiameremo modulo di finitezza.

Diamo una sfilza di definizioni tutte uguali, che useremo nel teorema di Fréchet-Kolmogorov:

**Definizione 60** (Famiglia equicontinua in  $L^1$ ). *Un insieme di funzioni  $L^1$  si dice equicontinuo se esiste un modulo di continuità  $L^1$  comune a tutta la famiglia.*

**Definizione 61** (Famiglia equiconcentrata in  $L^1$ ). *Un insieme di funzioni  $L^1$  si dice equiconcentrato (in inglese “equitight”) se esiste un modulo di finitezza comune a tutta la famiglia.*

**Teorema 9.7.1** (Fréchet-Kolmogorov). *Una famiglia di funzioni  $L^1(\mathbb{R}^n)$  è relativamente compatta se e solo se è equicontinua ed equiconcentrata.*

La dimostrazione è stata lasciata come esercizio, ma ne riparlamo quando la mia voglia avrà misura non nulla. Per gli interessati rimando a [Wikipedia](#). A volte si aggiunge anche l'ipotesi che la famiglia sia equilimitata (ossia che esista una costante  $C$  che maggia tutte le norme  $L^1$  delle funzioni), ma è stato mostrato che tale ipotesi discende dalle altre.

## 9.8 Cambio di variabili

### 9.8.1 Immagine continua di insiemi misurabili

**Lemma 9.8.1** (Ivan Bioli). *Sia  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione continua.  $f$  manda Lebesgue-misurabili in Lebesgue-misurabili se e solo se manda insiemi trascurabili in insiemi trascurabili.*

*Dimostrazione.* ( $\Rightarrow$ ) Se  $f$  mandasse un insieme trascurabile  $N$  in un insieme di misura positiva, per il teorema 8.7.5 esisterebbe un sottoinsieme  $W \subseteq f(N)$  non misurabile, che dovrebbe essere l'immagine di un sottoinsieme di  $N$ .

( $\Leftarrow$ ) Un Lebesgue è un'unione numerabile di compatti cui si aggiunge un sottoinsieme di un Borel-nullo, per quanto osservato in 8.6.4. Dunque anche l'immagine di un misurabile sarà fatta in tal modo, perché immagine continua di un compatto è compatta e per ipotesi  $f$  manda insiemi trascurabili in insiemi trascurabili.  $\square$

In generale, non tutte le funzioni continue preservano l'algebra di Lebesgue in avanti. Ad esempio, esistono omeomorfismi che mappano insiemi di misura nulla in insiemi non misurabili. Ad esempio, sia  $C : [0, 1] \rightarrow [0, 1]$  la funzione di Cantor<sup>7</sup>, che sappiamo essere continua e debolmente crescente, e consideriamo l'omeomorfismo  $g(x) = x + C(x)$  che manda  $[0, 1]$  in  $[0, 2]$ . Ora, il complementare dell'insieme di Cantor  $C$  è un'unione numerabile disgiunta di intervalli  $J_k$ , su ciascuno dei quali la funzione di Cantor è costante. Allora  $g(J_k) = J_k + c$  per una certa costante, e dunque

$$\lambda g([0, 1] \setminus C) = \sum_k \lambda(g(J_k)) = \sum_k \lambda(J_k) = 1$$

Dunque  $g(C)$  ha misura 1 per differenza, e procedendo come nel teorema si trova un sottoinsieme dell'insieme di Cantor la cui immagine non è misurabile.

Per le funzioni lipschitziane è possibile dare una stima della misura esterna degli insiemi di arrivo.

**Lemma 9.8.2.** *Sia  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$  lipschitziana. Allora esiste una costante  $C$  tale che, per ogni  $E \in \mathbb{R}^n$ ,  $\lambda^*(g(E)) \leq C\lambda^*(E)$ .*

*Dimostrazione.* Fissato  $\varepsilon > 0$ , dato che la misura di Lebesgue è una restrizione della misura esterna generata dal volume dei pluri-intervalli, per la definizione 43 possiamo determinare una famiglia numerabile di rettangoli  $E \subseteq \bigcup J_k$  tali che  $\sum \lambda(J_k) \leq \lambda^*(E) + \varepsilon$ . Possiamo supporre che i  $J_k$  abbiano tutti i lati della stessa lunghezza. Infatti, a meno di ingrandirli in modo controllato possiamo supporre che i  $J_k$  abbiano estremi razionali, e che dunque ognuno di essi si possa suddividere in un numero finito di cubetti tutti uguali. I cubi sono le palle nella norma uniforme, e dato che tutte le norme sono equivalenti in  $\mathbb{R}^n$   $g$  è lipschitziana nella norma uniforme, supponiamo di costante  $L$ . Dunque, se  $J_k = B_\infty(a_k, r_k)$  si ha che  $g(J_k) \subseteq B_\infty(g(a_k), Lr_k)$ , che ha misura  $L^n \lambda(J_k)$ . Ora basta osservare che

$$\lambda(g(E)) \leq \sum \lambda(B_\infty(g(a_k), Lr_k)) \leq L^n \sum \lambda(J_k) \leq L^n(\lambda^*(E) + \varepsilon)$$

e passare all'estremo inferiore su  $\varepsilon$ .  $\square$

Questo ci mostra in particolare che

**Corollario 9.8.2.1.** *Una funzione localmente lipschitziana manda insiemi di misura nulla in insiemi di misura nulla, e dunque preserva l'algebra di Lebesgue in avanti.*

*Dimostrazione.* Dato un insieme di misura nulla  $E$ , la sua immagine si può vedere come

$$g(E) = \bigcup_{k \in \mathbb{N}} g(E \cap \overline{B}(0, k))$$

Ora, dato che una funzione localmente lipschitziana è lipschitziana sui compatti connessi, per ogni  $k$  esiste una costante  $C_k$  per la quale si ha che

$$\lambda(g(E \cap \overline{B}(0, k))) \leq C_k \lambda(E \cap \overline{B}(0, k)) = 0$$

Dunque  $E$  è trascurabile in quanto unione numerabile di pezzi di misura nulla.  $\square$

<sup>7</sup>[Mangioni, 2019].

### 9.8.2 Comportamento della misura di Lebesgue sotto applicazioni lineari

**Teorema 9.8.3.** *Sia  $T \in GL(n, \mathbb{R})$ . Allora per ogni  $E \in \mathcal{L}(\mathbb{R}^n)$  si ha che*

$$\lambda(T(E)) = |\det T| \lambda(E)$$

*Dimostrazione.* Intanto osserviamo che un'applicazione lineare su  $\mathbb{R}^n$  è lipschitziana, dunque manda misurabili in misurabili per il corollario 9.8.2.1 e ha senso prendere la misura dell'immagine. Osserviamo poi che  $\mu(E) = \frac{\lambda(T(E))}{\lambda(T(I^n))}$ , dove  $I^n$  è il cubo unitario, è una misura invariante per traslazioni, dato che

$$\frac{\lambda(T(E+h))}{\lambda(T(I^n))} = \frac{\lambda(T(E) + T(h))}{\lambda(T(I^n))} = \frac{\lambda(T(E))}{\lambda(T(I^n))}$$

per invarianza della misura di Lebesgue. Inoltre  $\mu$  è normalizzata a 1 sul cubo unitario, e quindi per il teorema 8.6.6 la nuova misura coincide con la misura di Lebesgue sui boreliani, e dunque sui Lebesgue misurabili in virtù del teorema 8.3.7. Si ha dunque che

$$\lambda(T(E)) = \lambda(T(I^n))\lambda(E)$$

Mostriamo ora che  $\lambda(T(I^n)) = |\det T|$ . Innanzitutto osserviamo che  $\lambda(T(I^n))$  è moltiplicativo, in quanto date due applicazioni lineari  $L$  e  $T$  si ha

$$\lambda(TL(I^n)) = \lambda(T(L(I^n))) = \lambda(T(I^n))\lambda(L(I^n))$$

Ora, se  $T \in \mathcal{O}(n, \mathbb{R})$   $T$  preserva la palla euclidea  $B$ , dunque

$$\lambda(T(B)) = \lambda(B) = \lambda(T(I^n))\lambda(B)$$

Perciò la costante per le applicazioni ortogonali è  $1 = |\det T|$ . Ricordiamo poi che ogni matrice invertibile ammette una decomposizione polare  $T = OS$ , dove  $O \in \mathcal{O}(n, \mathbb{R})$  e  $S$  è simmetrica e definita positiva; a sua volta quest'ultima si diagonalizza ortogonalmente per il teorema spettrale, dunque possiamo scrivere  $S = PDP^T$ , con  $D$  diagonale e  $P \in \mathcal{O}(n, \mathbb{R})$ . Allora  $T = OPDP^T$ , dunque per moltiplicatività di  $\lambda(T(I^n))$  abbiamo che

$$\lambda(T(E)) = \lambda(D(I^n))\lambda(E)$$

Si verifica facilmente che la costante associata a una matrice diagonale a entrate positive è il suo determinante, che è il modulo del determinante di  $T$ . Segue dunque la tesi.  $\square$

### 9.8.3 Mettiamo assieme i pezzi

Simpatica questa dimostrazione,  
non l'ho capita ma mi piace.

*Pmajer dopo la sua dimostrazione*

**Teorema 9.8.4** (Cambio di variabile). *Sia  $g : U \rightarrow g(U)$  un diffeomorfismo  $C^1$  tra aperti di  $\mathbb{R}^n$ . Allora per ogni  $u \in \mathcal{L}^1(g(U))$  si ha che*

$$\int_{g(U)} u(y) dy = \int_U u(g(x)) |\det d_x g| dx$$

*Dimostrazione.* Ci serve innanzitutto mostrare un lemma generale:

**Lemma 9.8.5.** *Sia  $g : U \rightarrow U'$  una mappa  $C^1$  tra aperti di  $\mathbb{R}^n$  (non per forza un diffeomorfismo). Allora per ogni insieme misurabile  $E \subseteq U$  vale che*

$$\lambda(g(E)) \leq \int_E |\det d_x g| dx$$

*Dimostrazione.* Possiamo supporre che  $E$  sia compatto. Infatti per quanto detto in 8.6.4.1 ogni insieme misurabile secondo Lebesgue è un'unione numerabile di compatti  $E_n$  a meno di un insieme nullo  $N$ , dunque

$$g(E) = g\left(\bigcup_{n \in \mathbb{N}} E_n \sqcup N\right) = g\left(\bigcup_{n \in \mathbb{N}} E_n\right) \cup g(N)$$

Dato che  $N$  ha misura nulla e  $g$  è localmente lipschitziana la sua immagine deve avere misura nulla, perciò possiamo trascurarlo. Dunque se mostriamo la tesi per i compatti per le proprietà della misura avremo che

$$\lambda(g(E)) = \sup_{N \in \mathbb{N}} \lambda\left(g\left(\bigcup_{n=0}^N E_n\right)\right) \leq \sup_{N \in \mathbb{N}} \int_{\bigcup_{n=0}^N E_n} |\det d_x g| dx = \int_E |\det d_x g| dx$$

Siano ora

$$\delta < \delta_0 < \min\{\|x - y\|_\infty, x \in E, y \in U^C\}$$

Il minimo esiste per compattezza di  $E$  e dunque è non nullo. Poniamo inoltre  $I_\delta = [0, \delta)$  e consideriamo la decomposizione di  $\mathbb{R}^n$  in una griglia di cubi semiaperti ottenuti traslando  $I_\delta^n$ . Sia  $E^\delta$  la chiusura dell'unione finita dei cubi di tale sorta che ricoprono  $E$ . Dato che  $\delta < \delta_0$   $E^\delta$  è ancora contenuta in  $U$ . Ora,  $E \subseteq E^\delta \subseteq U$  per ogni  $\delta$ ; d'altro canto  $\bigcap_{\delta > 0} E^\delta = E$ , e dato che  $E^\delta$  ha misura finita essendo compatto possiamo dire che  $\lambda(E^\delta) \downarrow \lambda(E)$ . Inoltre, detti  $Q_1, \dots, Q_k$  i cubi la cui unione dà  $E^\delta$ ,

$$\lambda(g(E)) \leq \lambda(g(E^\delta)) \leq \sum_{i=1}^k \lambda(g(Q_i)) = \int_{\mathbb{R}^n} \sum_{i=1}^k \frac{\lambda(g(Q_i))}{\lambda(Q_i)} \chi_{Q_i}(x) dx$$

Sia  $\omega$  un modulo di continuità per il differenziale di  $g$  sul compatto  $E^{\delta_0}$ , rispetto alla norma infinito. Dunque per ogni coppia di punti  $x, y$  in uno dei cubi  $Q_i$  di lato  $\delta < \delta_0$  abbiamo che

$$\|g(y) - g(x) - d_x g(y - x)\|_\infty = (\|g(x + t(y - x)) - t d_x g(y - x)\|_\infty)_{t=0}^{t=1}$$

e per il teorema del valor medio

$$\|g(y) - g(x) - d_x g(y - x)\|_\infty \leq \|y - x\|_\infty \sup_{t \in [0,1]} \|d_{x+t(y-x)} g - d_x g\|_\infty \leq \delta \omega(\delta)$$

Allora per ogni  $x$  in  $Q_i$

$$g(Q_i) \subseteq g(x) + d_x g(Q_i - x) + \overline{B}_\infty(0, \delta \omega(\delta))$$

A sua volta,  $\overline{B}_\infty(0, \delta \omega(\delta))$  è un cubo chiuso di lato  $2\delta \omega(\delta)$ . Dunque, detto  $I^n$  il cubo chiuso unitario

$$\lambda(g(Q_i)) \leq \lambda(d_x g(I^n) + 2\delta \omega(\delta) I^n)$$

Operando un'omotetia di costante  $\delta$  otteniamo, per quanto visto in 9.8.3,

$$\lambda(g(Q_i)) \leq \delta^n \lambda(d_x g(I^n) + 2\omega(\delta) I^n) = \lambda(Q_i) \lambda(d_x g(I^n) + 2\omega(\delta) I^n)$$

Dunque sostituendo nell'integrale<sup>8</sup>

$$\lambda(g(E)) \leq \int_{\mathbb{R}^n} \sum_{i=1}^k \chi_{Q_i}(x) \lambda(d_x g(I^n) + 2\omega(\delta) I^n) dx = \int_{E^\delta} \lambda(d_x g(I^n) + 2\omega(\delta) I^n) dx$$

E passando al limite

$$\lambda(g(E)) \leq \inf_{\delta > 0} \int_{E^\delta} \lambda(d_x g(I^n) + 2\omega(\delta) I^n) dx$$

Ora, per ogni  $x$  si ha

$$\bigcap_{k \in \mathbb{N}^+} d_x g(I^n) + 2\omega(1/k) I^n = d_x g(I^n)$$

<sup>8</sup>Ha senso integrare perché la funzione che manda  $x$  in  $\lambda(d_x g(I^n) + 2\omega(\delta) I^n)$  è misurabile in quanto continua.

Dunque, dato che ciascuno di questi insiemi ha misura finita,

$$\lambda(d_x g(I^n) + 2\omega(1/k)I^n) \downarrow \lambda(d_x g(I^n)) = |\det d_x g| \lambda(I^n) = |\det d_x g|$$

Per mostrare che gli integrali convergono osserviamo che

$$d_x g(I^n) + 2\omega(1/k)I^n \subseteq \overline{B}_\infty(0, \|dg\|_{\infty, E^{\delta_0}} + 2\omega(\delta_0))$$

Dunque la sua misura è dominata da una certa costante  $M$  uniformemente in  $x$  e  $\delta$ . Allora essendo la misura di  $E^{\delta_0}$  finita la costante  $M$  è integrabile, perciò possiamo applicare il teorema del ConDom e ottenere

$$\lambda(g(E)) \leq \inf_{\delta > 0} \int_{E^\delta} \lambda(d_x g(I^n) + 2\omega(\delta)I^n) dx = \int_E |\det d_x g| dx$$

□

Ora, la tesi del lemma precedente si riscrive in questo modo:

$$\int_{g(U)} \chi_{g(E)}(y) \leq \int_E \chi_{g(E)}(g(x)) |\det d_x g| dx$$

Dunque abbiamo mostrato una proprietà per le funzioni semplici di insiemi della forma  $g(E)$  (cioè tutti perché  $g$  è invertibile). Per linearità possiamo estendere l'enunciato alle funzioni semplici, e per Beppo Levi otteniamo che

**Lemma 9.8.6.** *Nelle ipotesi del teorema, per ogni  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  positiva si ha*

$$\int_{g(U)} u(y) \leq \int_E u(g(x)) |\det d_x g| dx$$

Applicando la disuguaglianza sostituendo  $g$  con  $g^{-1}$  otteniamo

$$\int_{g(U)} u(y) dy \leq \int_U g(x) |\det d_x g| dx \leq \int_{g(U)} u(y) |\det d_{g^{-1}(y)} g| |\det d_y g^{-1}| dy = \int_{g(U)} u(y) dy$$

Perché

$$|\det d_{g^{-1}(y)} g| |\det d_y g^{-1}| = |\det(d_{g^{-1}(y)} g \circ d_y g^{-1})| = |\det(d_y(g \circ g^{-1}))| = 1$$

Dunque abbiamo provato la tesi per le funzioni positive. Spezzando una funzione in  $L^1$  in parte positiva e negativa si ottiene la tesi nel caso generale. □

**Lemma 9.8.7** (Sard). *Sia  $g : U \rightarrow \mathbb{R}^n$  una funzione  $C^1$ . Sia*

$$C_g = \{x \in U \text{ t.c. } \det d_x g = 0\}$$

*l'insieme dei punti critici di  $g$ . Allora l'insieme dei valori critici  $g(C_g)$  ha misura nulla.*

Basta infatti applicare il lemma 9.8.5 all'insieme  $C_g$ .

Con questo risultato possiamo estendere il teorema di cambio di variabile:

**Teorema 9.8.8** (Cambio di variabile, forma perfetta). *Sia  $g : U \rightarrow \mathbb{R}^n$  una funzione  $C^1$  iniettiva. Allora per ogni funzione  $u \in \mathcal{L}^1(\mathbb{R}^n)$  vale che*

$$\int_{g(U)} u(y) dy = \int_U u(g(x)) |\det d_x g| dx$$

*Dimostrazione.* Osserviamo che la restrizione di  $g$  a  $V = U \setminus C_f$  è un diffeomorfismo  $C^1$ , per teorema di invertibilità locale. Dunque, dato che  $g(U)$  e  $g(V)$  differiscono per un sottoinsieme del trascurabile  $g(C_g)$  possiamo scrivere

$$\int_{g(U)} u(y) dy = \int_{g(V)} u(y) dy = \int_V u(g(x)) |\det d_x g| dx = \int_U u(g(x)) |\det d_x g| dx$$

L'ultima uguaglianza è vera perché su  $C_g$  la funzione  $u(g(x)) |\det d_x g|$  è identicamente nulla. □

### 9.8.4 Radon-Nikodym

In questa sottosezione vogliamo citare alcuni strumenti più avanzati che permettono una dimostrazione alternativa della formula di cambio variabile.

**Definizione 62** (AC). *Siano  $\mu, \nu$  due misure su  $(X, \mathcal{A})$ . Si dice che  $\nu$  è assolutamente continua rispetto a  $\mu$ , e si scrive  $\nu \ll \mu$  se*

- ogni insieme  $\mu$ -trascurabile è  $\nu$ -trascurabile;
- ogni insieme  $\mu$ -finito è  $\nu$ -finito.

**Lemma 9.8.9.** *Data una funzione  $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$  positiva<sup>9</sup>, la funzione di insieme*

$$A \in \mathcal{A} \rightarrow \int_X \chi_A(x) f(x) d\mu(x)$$

è una misura finita e assolutamente continua rispetto a  $\mu$ . Tale misura si indica con  $f\mu$  e si dice definita per densità da  $f$ .

*Dimostrazione.* Mostriamo che la funzione definita sopra è  $\sigma$ -additiva. Data un'unione numerabile disgiunta  $\sqcup_{n \in \mathbb{N}} E_n$  possiamo scrivere

$$f\mu(\sqcup_{n \in \mathbb{N}} E_n) = \int_X \sum_{n \in \mathbb{N}} f(x) \chi_{E_n}(x) d\mu(x) = \sum_{n \in \mathbb{N}} \int_X f(x) \chi_{E_n}(x) d\mu(x) = \sum_{n \in \mathbb{N}} f\mu(E_n)$$

dove l'ultimo passaggio è il teorema di Beppo Levi.

Si vede facilmente che  $f\mu$  è finita, perché lo è l'integrale di  $f$ . Per l'assoluta continuità basta osservare che se  $A$  è  $\mu$ -trascurabile la funzione  $\chi_A(x)f(x)$  è nulla quasi ovunque, e perciò ha integrale nullo.  $\square$

Faremo uso del seguente teorema, che non dimostreremo non essendo nelle nostre corde:

**Teorema 9.8.10** (Radon-Nikodym). *Sia  $(X, \mathcal{A}, \mu)$  uno spazio di misura  $\sigma$ -finito e sia  $\nu \ll \mu$  su  $\mathcal{A}$  finita. Allora esiste una funzione  $\mu$ -integrabile e positiva  $f$  per cui  $f\mu = \nu$ , ed è unica a meno di uguaglianze quasi ovunque. Tale funzione si dice derivata di Radon-Nikodym di  $\nu$  rispetto a  $\mu$  e si indica con  $\frac{d\nu}{d\mu}$ .*

Il teorema ci dice che, per ogni insieme misurabile  $A$ ,

$$\int_X \chi_A d\nu = \int_X \chi_A \frac{d\nu}{d\mu} d\mu$$

Generalizzando con gli usuali ragionamenti di densità possiamo dire che:

**Corollario 9.8.10.1.** *Nelle ipotesi del teorema di Radon-Nikodym, per ogni funzione  $u \in \mathcal{L}^1(\nu)$  vale l'uguaglianza*

$$\int_X u d\nu = \int_X u \frac{d\nu}{d\mu} d\mu$$

In particolare  $u \frac{d\nu}{d\mu} \in \mathcal{L}^1(\mu)$ .

Ora, sia  $\nu$  una misura finita sui boreliani di  $\mathbb{R}^n$  e assolutamente continua rispetto a  $\lambda$ , e sia  $\nu = f\lambda$ . Per determinare  $f$  osserviamo che

$$\frac{\nu(B(x, r))}{\lambda(B(x, r))} = \frac{1}{\lambda(B(x, r))} \int_{B(x, r)} f(y) dy = \bar{f}_r(x)$$

Inoltre il teorema di differenziazione<sup>10</sup> ci dice che  $\bar{f}_r \rightarrow f$  in  $L^1$  e quasi ovunque per  $r \rightarrow 0$ . Dunque

$$\frac{\nu(B(x, r))}{\lambda(B(x, r))} \rightarrow f(x) \text{ quasi ovunque}$$

Questo giustifica intuitivamente la notazione  $f = \frac{d\nu}{d\lambda}$ , perché è come se stessimo considerando un "rapporto incrementale".

Dimostriamo ora la formula di cambio variabile con gli strumenti visti finora.

<sup>9</sup>La costruzione funziona anche con una  $f$  di segno qualsiasi.

<sup>10</sup>Vedasi 9.6.7

*Dimostrazione.* Sia  $U \subseteq \mathbb{R}^n$  aperto e sia  $g : U \rightarrow \mathbb{R}^n$   $C^1$  e iniettiva. Consideriamo

$$\nu(A) = \lambda(g(A))$$

Dato che  $g$  è localmente lipschitziana l'immagine di un misurabile è misurabile, perciò la funzione è ben definita. Inoltre è una misura, perché data un'unione disgiunta  $\sqcup_{n \in \mathbb{N}} A_n$  si ha che

$$\nu(\sqcup_{n \in \mathbb{N}} A_n) = \lambda(\sqcup_{n \in \mathbb{N}} g(A)) = \sum_{n \in \mathbb{N}} \lambda(g(A))$$

Per costruzione  $\nu$  è finita sui compatti, dato che lo è la misura di Lebesgue, ed è assolutamente continua rispetto a  $\lambda$ , perché una funzione localmente lipschitziana manda insiemi trascurabili in insiemi trascurabili. Allora per Radon-Nikodym esiste una funzione  $f$  positiva e integrabile sui compatti<sup>11</sup> tale che

$$\nu(A) = \lambda(g(A)) = \int_A f(x) dx$$

Ora, per quasi ogni  $x \in U$   $f(x)$  è il limite della successione  $\frac{\lambda(g(B(x,r)))}{\lambda(B(x,r))}$ , e in modo analogo a quanto visto nell'altra dimostrazione<sup>12</sup> si può far vedere che tale limite è  $|\det d_x g|$ .  $\square$

### 9.8.5 Integrazione in coordinate polari

La funzione

$$(-\pi, \pi) \times \mathbb{R}^+ \xrightarrow{g} \mathbb{R}^2 \setminus (\mathbb{R}_0^- \times \{0\})$$

$$(\theta, \rho) \longrightarrow \rho(\cos \theta, \sin \theta)$$

è un diffeomorfismo  $C^1$  con differenziale

$$d_{(\theta, \rho)} g = \begin{bmatrix} -\rho \sin \theta & \cos \theta \\ \rho \cos \theta & \sin \theta \end{bmatrix}$$

Dunque per ogni funzione  $u \in \mathcal{L}^1(\mathbb{R}^2)$  abbiamo (a meno di escludere la semiretta reale negativa, che ha misura nulla)

$$\iint_{\mathbb{R}^2} u(x, y) dx dy = \int_{-\pi}^{\pi} \int_0^{+\infty} u(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta$$

Ad esempio possiamo calcolare l'integrale della gaussiana:

$$\begin{aligned} \left( \int_{\mathbb{R}} e^{-x^2} dx \right)^2 &= \int_{\mathbb{R}} e^{-y^2} \left( \int_{\mathbb{R}} e^{-x^2} dx \right) dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy \\ &= \int_{-\pi}^{\pi} \int_0^{+\infty} e^{-\rho^2} \rho d\rho d\theta = -\pi \int_0^{+\infty} e^{-\rho^2} (-2\rho) d\rho = -\pi (e^{-\rho^2})_{\rho=0}^{\rho=+\infty} = \pi \end{aligned}$$

Dunque

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

## 9.9 Continuità di operatori di sostituzione

**Definizione 63** (Operatore di sostituzione). *Sia  $(X, \mathcal{A}, \mu)$  uno spazio di misura e sia  $f : X \times \mathbb{R} \rightarrow \mathbb{R}$ . L'operatore di sostituzione associato a  $f$  è la mappa definita come*

$$\Phi : \mathbb{R}^X \longrightarrow \mathbb{R}^X$$

$$u \longrightarrow f(\cdot, u(\cdot))$$

<sup>11</sup>Più precisamente, stiamo applicando il teorema di Radon-Nikodym alla restrizione di  $\nu$  ai compatti, e poi unendo le funzioni così ottenute a dare una  $f$  definita su  $\mathbb{R}^n$ . Questo si può fare perché le funzioni parziali sono determinate a meno di uguaglianze quasi ovunque, perciò ogni estensione deve coincidere quasi ovunque con le funzioni sugli insiemi più piccoli.

<sup>12</sup>...

Ci chiediamo quando questa operazione definisca una mappa continua da  $L^1$  in sé.

**Definizione 64** (Ipotesi di Carathéodory). *Una funzione  $f : X \times \mathbb{R} \rightarrow \mathbb{R}$  soddisfa le ipotesi di Carathéodory se  $\forall x \in X$  e  $\forall t \in \mathbb{R}$  vale che*

1.  $f(x, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$  è continua;
2.  $f(\cdot, t) : X \rightarrow \mathbb{R}$  è misurabile;
3.  $|f(x, t)| \leq h(x) + c|t|$  per una qualche  $h$  integrabile positiva e una certa costante  $c$ .

Si osservi che una funzione con tali proprietà è misurabile nella coppia  $(x, t)$  rispetto alla  $\sigma$ -algebra prodotto  $\mathcal{A} \otimes \mathcal{L}(\mathbb{R})$ . Infatti, consideriamo la successione di funzioni

$$f_m(x, t) = f\left(x, \frac{\lfloor mt \rfloor}{m}\right) = \sum_{k \in \mathbb{Z}} f\left(x, \frac{k}{m}\right) \chi_{[0,1)}(mt - k) = \sum_{k \in \mathbb{Z}} f\left(x, \frac{k}{m}\right) \chi_{\left[\frac{k}{m}, \frac{k}{m} + 1\right)}(t)$$

Ciascuna  $f_m$  è misurabile rispetto all'algebra prodotto, in quanto si ottiene moltiplicando una funzione misurabile in  $x$ , indipendente da  $t$ , e una funzione caratteristica di  $t$  nella quale non compare  $x$ . Inoltre, poiché  $f(t, x)$  è continua in  $t$  per quasi ogni  $x$ ,  $f$  risulta il limite (quasi ovunque) delle  $f_n$  e quindi è misurabile.

**Teorema 9.9.1.** *Se  $f$  soddisfa le ipotesi di Carathéodory, l'operatore di sostituzione associato induce una mappa continua da  $L^1$  in sé.*

*Dimostrazione.* Mostriamo dapprima che se  $u : X \rightarrow \mathbb{R}$  è misurabile lo è anche  $\Phi(u)$ . Proviamolo innanzitutto per una funzione semplice  $\phi = \sum_{k=1}^n c_k \chi_{A_k}$ . A meno di porla in forma canonica e di aggiungere un insieme su cui  $\phi = 0$  possiamo supporre che gli  $A_k$  definiscano una partizione di  $X$ . Allora vale che

$$\Phi(\phi)(x) = f(x, \phi(x)) = \sum_{k=1}^n f(x, c_k) \chi_{A_k}(x)$$

Dunque  $\Phi(\phi)$  è misurabile perché per ipotesi lo è ogni  $f(\cdot, c_k)$ . Inoltre, se  $u$  è una funzione misurabile e  $(\phi_n)_{n \in \mathbb{N}}$  è una successione di funzioni semplici che vi converge puntualmente, per quasi ogni  $x$  fissato  $f(x, \phi_n(x)) \rightarrow f(x, u(x)) = \Phi(u)(x)$ , dunque quest'ultima è misurabile rispetto all'algebra di Lebesgue perché si ottiene come limite quasi ovunque di funzioni misurabili.

Ora supponiamo che  $u \in \mathcal{L}^1$ . Per ipotesi  $|f(x, u(x))| \leq h(x) + c|u(x)|$ , e poiché la RHS è integrabile lo è anche  $f(x, u(x))$ . Dunque l'operatore manda  $\mathcal{L}^1$  in sé.

Mostriamo che  $\Phi$  è sequenzialmente continua. Fissata una funzione  $u \in \mathcal{L}^1$ , ci basta mostrare<sup>13</sup> che da ogni successione  $u_n \rightarrow u$  in norma  $L^1$  si estrae una sottosuccessione  $u_{n(k)}$  tale che  $\Phi(u_{n(k)}) \rightarrow \Phi(u)$  in norma  $L^1$ . Questo è vero perché, per l'inverso del teorema del ConDom, esiste una sottosuccessione  $u_{n(k)} \rightarrow u$  convergente quasi ovunque e dominata da una qualche  $g \in \mathcal{L}^1$  positiva. Allora  $f(x, u_{n(k)}(x)) \rightarrow f(x, u(x))$  per quasi ogni  $x$ , dato che  $f(x, \cdot)$  è continua, e inoltre  $|f(x, u_{n(k)}(x))| \leq h(x) + cg(x)$ . Dunque per teorema del ConDom  $\Phi(u_{n(k)}) \rightarrow \Phi(u)$ .

Ora,  $\Phi$  passa al quoziente  $L^1$ , perché, se  $u(x) = v(x)$  per  $\mu$ -q.o.  $x \in X$  allora  $f(x, u(x)) = f(x, v(x))$  per gli stessi  $x$ , e cioè  $\Phi(u) = \Phi(v)$  quasi ovunque.  $\square$

### 9.9.1 Call of DUTIS

**Teorema 9.9.2** (Dutis per l'integrale di Lebesgue). *Sia  $(X, \mathcal{A}, \mu)$  uno spazio di misura  $\sigma$ -finito. Sia  $f : X \times \mathbb{R} \rightarrow \mathbb{R}$  con le seguenti proprietà, che estendono le ipotesi di Carathéodory:*

1.  $f(x, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$  è  $C^1$  per quasi ogni  $x \in X$ ;
2.  $f(\cdot, t) : X \rightarrow \mathbb{R}$  è misurabile per ogni  $t \in \mathbb{R}$ ;
3.  $|f(x, t)| + |\partial_t f(x, t)| \leq h(x)$  per una funzione integrabile  $h \in \mathcal{L}^1(X)$ .

<sup>13</sup>Questa è una conseguenza della proprietà di Urysohn, di cui si parla nella sezione 1.3.1.

Allora la funzione  $F(t) = \int_X f(x, t) d\mu(x)$  è di classe  $C^1$  e la sua derivata è

$$\frac{d}{dt} \left( \int_X f(x, t) d\mu(x) \right) = \int_X \partial_t f(x, t) d\mu(x)$$

*Dimostrazione.* Mostriamo dapprima che la funzione è continua, cosa che si può far vedere per successioni. Se  $t_n \rightarrow t$  allora per continuità la successione  $f(x, t_n)$  converge puntualmente quasi ovunque a  $f(x, t)$ . Dunque, dato che  $|f(x, t)| \leq h(x)$ , per convergenza dominata abbiamo che

$$\lim_{n \rightarrow +\infty} F(t_n) = \lim_{n \rightarrow +\infty} \int_X f(x, t_n) d\mu(x) = \int_X \lim_{n \rightarrow +\infty} f(x, t_n) d\mu(x) = \int_X f(x, t) d\mu(x) = F(t)$$

Ora, la funzione  $\partial_t f(x, t)$  è misurabile a  $t$  fissato, in quanto possiamo scriverla come il seguente limite puntuale (quasi ovunque) di funzioni misurabili:

$$\partial_t f(x, t) = \lim_{h \rightarrow 0} \frac{f(x, t+h) - f(x, t)}{h}$$

Dunque gli stessi ragionamenti fatti per  $F$  ci dicono che la funzione  $G(t) = \int_X \partial_t f(x, t) d\mu(x)$  è ben definita e continua. Inoltre  $\partial_t f(x, t)$  è misurabile nella coppia  $(x, t)$  poiché rispetta le ipotesi di Carathéodory<sup>14</sup>, e perciò  $G$  è integrabile in  $t$  per quanto segue dal teorema di Fubini.

Ora, vogliamo mostrare che

$$\int_{t_0}^t G(s) ds = F(t) - F(t_0)$$

Questo ci permetterà di concludere che  $G$  è la derivata di  $F$ , perché l'integrale della funzione continua  $G$  coincide con quello di Riemann, per il quale vale il teorema fondamentale del calcolo integrale. Osserviamo dunque che

$$\int_{t_0}^t \left( \int_X \partial_t f(x, t) d\mu(x) \right) ds = \int_X \left( \int_{t_0}^t \partial_t f(x, t) ds \right) d\mu(x) = \int_X (f(x, t) - f(x, t_0)) d\mu(x) = F(t) - F(t_0)$$

□

Si potrebbe generalizzare il teorema a funzioni tra spazi di Banach, ma servirebbe la teoria di Bochner sull'integrazione di funzioni a valori vettoriali e penso che non sia il caso indagare oltre.

## 9.10 L'integrale come funzionale continuo (e viceversa)

Consideriamo una misura  $\mu$  sui boreliani dell'intervallo  $I = [0, 1]$ . Essa determina un funzionale lineare sulle funzioni continue  $C^0(I, \mathbb{R})$ , il quale associa a  $f$  il suo integrale rispetto a  $\mu$ . In simboli, esiste un funzionale  $F \in (C^0(I, \mathbb{R}))^*$  tale che

$$F(f) = \int_0^1 f(x) d\mu(x)$$

Si osservi che  $F$  è positivo, perché l'integrale di una funzione  $f \geq 0$  è non negativo. Generalizzando, una misura boreliana di segno qualsiasi  $\mu$ , che possiamo scrivere come differenza di due misure positive  $\mu_+ - \mu_-$ , determina un funzionale  $F$  che corrisponde all'integrazione in  $d\mu$ .

Si può mostrare il viceversa:

**Teorema 9.10.1** (Riesz-Markov-Kakutani). *Ogni funzionale lineare positivo  $F \in (C^0(I))^*$  è l'integrale rispetto a un'unica misura finita sui boreliani.*

*Dimostrazione.* La seguente dimostrazione è INCOMPLETE! Consiglio di saltarla, tanto era stata lasciata come esercizio.

Consideriamo, al variare delle funzioni continue  $u, v \in C^0(I)$ , i sottoinsiemi di  $I \times \mathbb{R}$  della forma

$$[u, v) = \{x, t \in I \times \mathbb{R} \text{ t.c. } u(x) \leq t < v(x)\}$$

<sup>14</sup>Vedasi il commento dopo la definizione 64.

Di fatto  $[u, v)$  è la regione di piano compresa tra il grafico di  $u$  (incluso) e quello di  $v$  (escluso). Questa famiglia di parti, che chiameremo  $\mathcal{F}$ , è un quasi-anello, perché

$$[u, v) \cap [u', v') = \{x, t \in I \times \mathbb{R} \text{ t.c. } \max(u(x), u'(x)) \leq t < \min(v(x), v'(x))\}$$

$$[u, v) \cap [u', v') = [\max(u(x), u'(x)), \min(v(x), v'(x)))$$

e per quanto riguarda la differenza

$$[u, v) \setminus [u', v') = \{x, t \in I \times \mathbb{R} \text{ t.c. } u(x) \leq t < v(x), t < u'(x) \vee t \geq v'(x)\} = [u, \min(v, u')) \cup [\max(u, v'), v)$$

Inoltre  $\mathcal{F}$  genera i boreliani di  $I \times \mathbb{R}$ . Per mostrarlo basta osservare che un rettangolo della forma  $(a, b) \times (c, d)$  si ottiene come unione di una famiglia numerabile di insiemi di  $\mathcal{F}$ , ad esempio considerando una successione  $[u_n, v_n)$  dove  $u_n \downarrow c\chi(a, b)$  puntualmente e  $v_n \uparrow d\chi(a, b)$ .<sup>15</sup> Ora, la funzione che manda  $[u, v) \in \mathcal{F}$  in  $\langle F, v - u \rangle$  è  $\sigma$ -additiva. Infatti, se  $\bigcup [u_k, v_k) = [U, V)$  Dunque per teorema di Carathéodory possiamo estendere tale funzione a una misura  $m$  su  $I \times \mathbb{R}$ . Questa risulta invariante per traslazioni verticali, ossia per ogni boreliano  $A \subseteq I \times \mathbb{R}$  e ogni numero reale  $t$  si ha  $m(A) = m(A + t)$ . Infatti questa proprietà vale per gli insiemi del  $\pi$ -sistema  $\mathcal{F}$ , e dunque basta mostrare che gli insiemi per cui è vera formano un sistema di Dynkin.

- $m(\emptyset) = m(\emptyset + t) = 0$ .

- 

- 

Per quanto visto nel teorema 8.7.1 ogni misura boreliana su  $[0, 1]$  è determinata dalla sua funzione di ripartizione  $\psi$  e dal valore  $c$  che assume sul singoletto  $\{1\}$ . Poniamo dunque

$$\psi(x) = \sup\{F(u) \text{ t.c. } 0 \leq u \leq \chi_{[0, x)}, u \in C^0(I)\}$$

$$c = F(1) - \psi(1)$$

Vorremmo poter scrivere  $\phi(x) = F(\chi_{[0, x)})$ , nello spirito di identificare  $\phi(x)$  con  $\mu([0, x))$ , ma non possiamo applicare il funzionale a delle funzioni discontinue.

Ora, ovviamente  $\psi$  è crescente e si annulla in  $x = 0$ . Per la continuità a sinistra osserviamo che

$$\begin{aligned} \sup_{y \leq x} \psi(y) &= \sup_{y \leq x} \sup\{F(u) \text{ t.c. } 0 \leq u \leq \chi_{[0, y)}, u \in C^0(I)\} = \\ &= \sup\{F(u) \text{ t.c. } \exists y < x \text{ t.c. } 0 \leq u \leq \chi_{[0, y)}, u \in C^0(I)\} = \\ &= \sup\{F(u) \text{ t.c. } 0 \leq u \leq \chi_{[0, x)}, u \in C^0(I)\} = \psi(x) \end{aligned}$$

Sia dunque  $\mu$  la misura la cui funzione di ripartizione è  $\psi$  e sia  $f \in C^0(I)$ , WLOG non negativa. Mostriamo che  $\int_0^1 f(x) d\mu(x) = F(f)$ . Essendo  $f$  continua,  $f$  è limite puntuale di una successione di funzioni semplici  $\phi$  ottenute sommando funzioni caratteristiche di intervalli (ad esempio, si possono scegliere le funzioni i cui integrali sono le somme integrali inferiori di  $f$ ). Dunque per come è costruito l'integrale

$$\int_0^1 f(x) d\mu(x) = \sup_{\phi \leq f} \int_0^1 \phi(x) d\mu(x)$$

Se  $\phi = \sum_{k=0}^K a_k \chi_{[x_k, x_{k+1})} + \phi(1)\chi_{\{1\}}$  abbiamo

Dunque  $F$  si estende a un funzionale lineare positivo sulle funzioni boreliani limitate, perché  $\int_I f d\mu$  è ben definito anche per le funzioni boreliane. Dovrebbe valere anche il viceversa: se si estende allora è l'integrale per qualche misura.  $\square$

Generalizzando a misure con segno otteniamo il seguente teorema:

**Corollario 9.10.1.1** (Riesz-Markov-Kakutani). *Ogni funzionale lineare  $F \in (C^0(I))^*$  è l'integrale rispetto a un'unica misura finita (con segno) sui boreliani.*

<sup>15</sup>Si può fare, sfruttando il raccordo  $C^\infty$  tra funzioni arbitrarie. A riguardo vedasi il capitolo 6 di [SNS, 2015].

### 9.10.1 Problema dei momenti di Hausdorff

**Definizione 65.** Sia  $F \in (C^0(I))^*$ , su un intervallo  $I \subseteq \mathbb{R}$  qualsiasi. Il  $k$ -esimo momento di  $F$ , con  $k \in \mathbb{N}$ , è definito come

$$M_F(k) = F(x^k) = \langle F, x^k \rangle$$

dove abbiamo adottato la notazione  $F(f) = \langle F, f \rangle$ . In particolare, se  $F$  è l'integrale rispetto alla misura  $\mu$  poniamo

$$M_\mu(k) = \int_I x^k d\mu(x)$$

In questa sezione ci porremo il problema di determinare se una successione di numeri reali non negativi sia o meno la successione dei momenti di un funzionale positivo (e dunque di una misura, in virtù del teorema di Riesz-Markov), e in tal caso come si possa determinare quest'ultimo. Lavoreremo su  $I = [0, 1]$ , perché il problema si fa molto meno banale se l'intervallo considerato è infinito. Si osservi che sotto questa ipotesi un funzionale  $F$  è determinato univocamente dalla successione dei suoi momenti, in quanto conoscere il valore di  $F$  sui monomi ci permette di ricavare quanto fa sui polinomi, e dunque su tutte le funzioni in virtù del teorema di approssimazione di Weierstrass<sup>16</sup> e della continuità di  $F$  rispetto alla convergenza uniforme.

Stabiliamo innanzitutto una condizione necessaria affinché una successione possa rappresentare i momenti di un funzionale positivo  $F$ . Per ogni  $k, n \in \mathbb{N}$  si ha  $x^k(1-x)^n \geq 0$ , dunque

$$0 \geq \langle F, x^k(1-x)^n \rangle = \left\langle F, x^k \sum_{j=0}^n \binom{n}{j} (-1)^j x^j \right\rangle = \sum_{j=0}^n \binom{n}{j} (-1)^j M_F(k+j)$$

Per dare un nome a questa proprietà ci serviranno un paio di strumenti:

**Definizione 66** (Shift). L'operatore lineare  $S : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$  tale che  $S(x)(k) = x(k+1)$  si dice *shift*.

In altri termini, lo shift manda la successione  $(x_0, x_1, \dots)$  in  $(x_1, x_2, \dots)$ . Si osservi che su  $l_\infty$  lo shift è lipschitziano di norma 1, perché il sup dei moduli della successione scalata a sinistra è minore o uguale del sup dei moduli della successione originale.

**Definizione 67** (Completa monotonia). Una successione reale  $y_n$  si dice *completamente monotona* se per ogni coppia di naturali  $k, n \in \mathbb{N}$  si ha che

$$\sum_{j=0}^n \binom{n}{j} (-1)^j y_{k+j} \geq 0$$

In altri termini una successione è *completamente monotona* se per ogni  $n \in \mathbb{N}$  la successione  $(I-S)^n(y)$  è non negativa.

Dunque la successione  $M_F$  è completamente monotona. Mostriamo che tale condizione è anche sufficiente.

**Teorema 9.10.2** (Hausdorff). Sia  $y$  una successione completamente monotona. Definiamo i coefficienti

$$c_{k,n} = \binom{n}{k} S^k(I-S)^{n-k}(y)(0) = \binom{n}{k} (I-S)^{n-k}(y)(k)$$

Sia inoltre  $\delta_a \in (C^0(I))^*$  il funzionale definito da  $\langle \delta_a, f \rangle = f(a)$ . La successione di funzionali lineari positivi  $F_n = \sum_{k=0}^n c_{k,n} \delta_{\frac{k}{n}}$  converge puntualmente a un funzionale  $\bar{F} \in (C^0(I))^*$  la cui successione dei momenti è  $y$ .

*Dimostrazione.* Sia  $B_n : C^0(I) \rightarrow \mathbb{R}_n[x]$  l'operatore che associa a ogni funzione continua il suo polinomio di Bernstein  $n$ -esimo.<sup>17</sup> Osserviamo che, data una funzione continua  $f$ , si ha

$$\langle F_n, f \rangle = B_n(f)[S](y)(0)$$

<sup>16</sup>Vedasi 5.2.1.

<sup>17</sup>Vedasi 5.2.2.

dove si intende che il polinomio  $B_n(f)$  è calcolato nell'operatore  $S$  e poi applicato a  $u$ . Infatti risulta

$$\langle F_n, f \rangle = \sum_{k=0}^n c_{k,n} f\left(\frac{k}{n}\right) = \left( \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} S^k (I - S)^{n-k} \right) (y)(0) = B_n(f)[S](y)(0)$$

Ora, dato che  $F_n$  è positivo, la sua norma operatoriale si ottiene calcolandolo nella funzione costante 1, come abbiamo visto nella sezione 5.2.3. Dunque

$$\|F_n\| = \langle F_n, 1 \rangle = B_n(1)[S](y)(0) = y(0)$$

Perciò la nostra successione di operatori è equilipschitziana. Inoltre, se  $f(x) = x^k$ ,  $B_n(x^k)$  è un polinomio di grado  $k$  che converge uniformemente a  $x^k$ . Scriviamo allora

$$B_n(x^k) = x^k + \varepsilon_{k,n}(x)$$

con  $\varepsilon_{k,n}(x) \in \mathbb{R}_k[x]$  e  $\varepsilon_{k,n} \rightarrow 0$  uniformemente. Dato che tutte le norme in  $\mathbb{R}^{k+1} \cong \mathbb{R}_k[x]$  sono equivalenti sappiamo anche che la norma uno del vettore dei coefficienti è infinitesima, e cioè che la somma dei moduli dei coefficienti di  $\varepsilon_{k,n}(x)$  tende a zero. Dunque

$$B_n(x^k)[S] = S^k + \varepsilon_{k,n}[S]$$

con  $\varepsilon_{k,n}[S] \rightarrow 0$  in norma operatoriale (perché quest'ultima si stima dall'alto con la somma dei moduli dei coefficienti, in quanto per ogni  $j$  si ha  $\|S^j\| = 1$ ). Allora

$$\langle F_n, x^k \rangle = B_n(x^k)[S](y)(0) \rightarrow S^k(y)(0) = y(k)$$

Perciò  $F_n$  converge puntualmente sui monomi, e per linearità converge sui polinomi. Ma una successione di funzioni equicontinue tra spazi di Banach converge in un chiuso, perciò per densità dei polinomi in  $C^0(I)$  le  $F_n$  convergono puntualmente su tutto lo spazio a un certo funzionale  $F$ . Quest'ultimo risulta ancora lineare e positivo, essendo queste caratteristiche conservate dal limite puntuale. Mostriamo che  $F$  ha come momento  $k$ -esimo proprio  $y(k)$ .

$$\langle F, x^k \rangle = \lim_{n \rightarrow +\infty} \langle F_n, x^k \rangle = \lim_{n \rightarrow +\infty} B_n(x^k)[S](y)(0) = y(k)$$

□

# Capitolo 10

## (Gran) Varietà

### 10.1 Varietà in $\mathbb{R}^n$

**Definizione 68.** Un sottoinsieme  $M \subseteq \mathbb{R}^n$  è una varietà differenziabile di classe  $C^l$  e di dimensione  $k$  se per ogni  $x \in M$  esistono due intorni aperti  $U \in \mathbb{R}^k$  e  $V \in \mathbb{R}^n$ , con  $x \in V$ , e un diffeomorfismo di classe  $C^l$   $\phi : U \rightarrow V \cap M$ , detto parametrizzazione locale.

**Definizione 69.** Lo spazio tangente al punto  $x \in M$  è il sottospazio lineare  $T_x(M) = d\phi_{\phi^{-1}(x)}(\mathbb{R}^k)$ .

La definizione precedente non è chiarissima. Si può facilmente dimostrare che è equivalente a quest'altra (che si può dimostrare essere molto meglio perché l'ho aggiunta io):

**Definizione 70.** Dato  $x \in M$ , consideriamo l'insieme delle curve  $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$   $C^1$  tali che  $\gamma(0) = x$ . Lo spazio tangente è lo spazio vettoriale generato dalle velocità in zero di tali curve.

In soldoni, la seconda definizione ci dice che il tangente è l'insieme delle possibili direzioni spaziali nelle quali ci si può muovere rimanendo in  $M$ .

### 10.2 Generalizzazione ai Banach

**Definizione 71.** Un sottoinsieme di uno spazio di Banach  $M \subseteq E$  si dice sottovarietà differenziabile di classe  $C^l$  se per ogni  $x \in M$  esistono:

- un aperto  $V \subseteq E$  contenente  $x$ ;
- una coppia di spazi di Banach  $E_1 \times E_2$ ;
- un aperto  $U \subseteq E_1 \times E_2$  (per comodità contenente l'origine);
- un diffeomorfismo (di classe  $C^l$ )  $\varphi : U \rightarrow V$ , detto parametrizzazione locale, tale che  $\varphi(0) = x$  e che

$$\varphi((E_1 \times \{0\}) \cap U) = M \cap V$$

Inoltre, se  $\dim(E_1) = k$  come spazio vettoriale, la sottovarietà si dice di dimensione  $k$ .

In parole povere, la sottovarietà è diffeomorfa a  $E_1$  e ha un intorno che eredita la struttura di prodotto di  $E_1 \times E_2$ . La definizione è un po' più complessa rispetto al caso di  $\mathbb{R}^n$ , dove possiamo pensare una varietà  $M$  come un insieme che è "localmente  $\mathbb{R}^k$ " e quindi in modo implicito decompone localmente  $\mathbb{R}^n$  nel prodotto di  $M$  e "dell'ortogonale".

Osserviamo inoltre che  $E_1 \times E_2$  è isomorfo a  $E$  tramite il differenziale di  $\phi$  nell'origine, che è un isomorfismo lineare. Dunque di fatto una sottovarietà è strettamente legata a una decomposizione dello spazio in cui vive.

**Definizione 72.** Con la notazione della definizione precedente, si definisce lo spazio tangente a una sottovarietà  $M$  in un suo punto  $x$  come

$$T_x(M) = d\varphi_0(E_1 \times \{0\})$$

Il tangente è in realtà indipendente da più o meno tutto quello che abbiamo usato per definirlo. In altre parole, il tangente dipende solo dalla “forma” della varietà, non dalla parametrizzazione. Consideriamo infatti due parametrizzazioni  $\varphi : U \subseteq E_1 \times E_2 \rightarrow V$  e  $\psi : W \subseteq F_1 \times F_2 \rightarrow V$ . Supponiamo WLOG che i due aperti di arrivo siano gli stessi, a meno di restringere le applicazioni. Ora,  $\varphi = \psi(\psi^{-1}\varphi)$ , e la quantità tra parentesi è un diffeomorfismo tra  $E_1 \times E_2$  e  $F_1 \times F_2$  che manda  $E_1 \times \{0\} \cap U$  in  $F_1 \times \{0\} \cap W$ . Perciò, considerando i differenziali,

$$T_x(M, \varphi) = d\varphi_0(E_1 \times \{0\}) = d\psi_0(d(\psi^{-1}\varphi)_0(E_1 \times \{0\})) = d\psi_0(F_1 \times \{0\}) = T_x(M, \psi)$$

Possiamo dire di più:

**Lemma 10.2.1.** *Se  $x, y \in M$  appartengono alla stessa componente connessa, i due tangenti sono isomorfi. In altre parole, la dimensione di una varietà connessa è ben definita.*

*Dimostrazione.* Dato  $x \in M$ , sia  $C$  l'insieme degli elementi della componente connessa di  $x$  che hanno tangente isomorfo a  $T_x(M)$ . Ora,  $C$  è aperto. Infatti, consideriamo un punto  $y$  il cui tangente è isomorfo a  $T_x(M)$ . Il suo tangente è  $d\varphi_0(E_1 \times \{0\})$  per una certa parametrizzazione  $\varphi$ . Quest'ultima è un diffeomorfismo, e in particolare in un intorno dell'origine il suo differenziale è ancora invertibile. Allora esiste un intorno di  $y$  i cui punti hanno tangente  $d\varphi_z(E_1 \times \{0\})$  per certi  $z \in E_1 \times E_2$ , e questi tangenti sono tutti isomorfi a  $E_1 \times \{0\}$  e quindi al tangente in  $y$ .

Con ragionamenti analoghi si mostra che anche il complementare di  $C$  in  $M$  è aperto, dunque per connessione  $M = C$  essendo  $C$  non vuoto.  $\square$

## 10.3 Come spammare sottovarietà

Quest'uomo ha guadagnato 300 sottovarietà in 16 minuti!

*Pubblicità sui siti di streaming*

### 10.3.1 Grafici di funzioni

Consideriamo una funzione  $f : \Omega \subseteq E \rightarrow F$  di classe  $C^1$ . Il suo grafico è una sottovarietà. Infatti la funzione  $\varphi : \Omega \times E \rightarrow \Omega \times F$  definita come

$$\varphi(x, y) = (x, f(x) - y)$$

è un diffeomorfismo e manda  $\Omega \times \{0\}$  nel grafico di  $f$ . Inoltre, il tangente al grafico è il grafico del differenziale:

$$T_{(x, f(x))}(\text{graph}(f)) = d\varphi_{(x, f(x))}(E_1 \times \{0\}) = \begin{bmatrix} 1 & 0 \\ df_x & -1 \end{bmatrix} \begin{bmatrix} E_1 \\ 0 \end{bmatrix} = \begin{bmatrix} E_1 \\ df_x(E_1) \end{bmatrix} = \text{graph}(df_x)$$

Ora, consideriamo il grafico di una funzione  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , che è dunque una  $n$ -varietà. Il suo tangente nel punto  $(x, f(x))$  è allora un iperpiano, definito come l'insieme dei vettori ortogonali a un certo  $w$ . Possiamo scegliere come  $w$  il versore  $\frac{(-\vec{\nabla} f(x), 1)}{\sqrt{1 + \|\vec{\nabla} f(x)\|^2}}$ , in quanto

$$\left\langle \left( v, \vec{\nabla} f(x) \cdot v \right), \left( -\vec{\nabla} f(x), 1 \right) \right\rangle = 0$$

### 10.3.2 Controimmagini di valori regolari

**Teorema 10.3.1.** *Sia  $f : \Omega \subseteq E_1 \times E_2 \rightarrow F$  una funzione di classe  $C^1$  e sia  $p \in F$ . Se in ogni punto  $x \in f^{-1}(p)$  il differenziale nella seconda variabile è invertibile, allora  $f^{-1}(p)$  è una varietà localmente diffeomorfa a un aperto di  $E_1$ . Inoltre*

$$T_x(f^{-1}(p)) = \ker df_x$$

*Dimostrazione.* Per il teorema della funzione implicita, esiste una funzione  $g$  tale che la controimmagine di  $p$  è localmente il grafico di  $g$ , e dunque in particolare è una varietà per il teorema precedente. Ora,

$$T_x(f^{-1}(p)) = T_x(\text{graph}(g)) = \text{graph}(dg_x) = \{(v, dg_x(v)), v \in E_1\}$$

Sviluppando  $dg_x$  per quanto sappiamo dal teorema della funzione implicita:

$$T_x(f^{-1}(p)) = \{(v, -d_2 f_x^{-1} d_1 f_x(v)), v \in E_1\}$$

Ora, poiché  $df_x(v, w) = d_1 f_x(v) + d_2 f_x(w)$ , dove  $(v, w) \in E_1 \times E_2$ , è facile osservare che

$$T_x(f^{-1}(p)) = \ker df_x$$

Da cui la tesi. □

Consideriamo ora il caso particolare in cui  $f : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$ . Traduciamo le ipotesi del teorema precedente nella seguente definizione.

**Definizione 73.** *Un punto  $x \in E$  si dice punto regolare per  $f : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$  se  $df_x$  è surgettivo.<sup>1</sup> Un punto  $p \in \mathbb{R}^{n-k}$  si dice valore regolare se tutte le sue controimmagini sono punti regolari.*

Ora, se  $p$  è un valore regolare  $\mathbb{R}^n$  si decompone in due sottospazi in somma diretta:

$$\mathbb{R}^n = \ker df_x \times \mathbb{R}^{n-k} = E_1 \times E_2$$

dove il secondo spazio è diffeomorfo all'immagine tramite  $f$ . Allora, a meno di cambiare base opportunamente<sup>2</sup> possiamo supporre che  $f$  sia la proiezione sulle ultime  $n - k$  coordinate, e che quindi il suo differenziale nella seconda variabile sia un isomorfismo. Giungiamo dunque al seguente:

**Corollario 10.3.1.1.** *Sia  $f : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$ . La controimmagine di un valore regolare è una  $k$ -varietà il cui tangente è  $T_x(f^{-1}(p)) = \ker df_x$ .*

A meno di modificare  $f = (f_1, \dots, f_{n-k})$ , se supponiamo che il valore regolare sia il vettore nullo abbiamo identificato la varietà come l'insieme di livello di  $n - k$  vincoli  $f_1, \dots, f_{n-k}$ . Si osservi che la condizione di regolarità impone che i gradienti delle  $f_i$  siano tutti linearmente indipendenti (la dimostrazione è lasciata al lettore). In quest'ottica possiamo enunciare un analogo del teorema dei moltiplicatori di Lagrange.

### 10.3.3 Differenziabilità

**Definizione 74** (Funzioni di classe  $C^k$ ). *Siano  $\Sigma, \Gamma \subseteq \mathbb{R}^N$  due varietà di classe  $C^k$ . Una funzione  $g : \Sigma \rightarrow \Gamma$  si dice di classe  $C^k$  se per ogni punto  $p \in \Sigma$  esiste una parametrizzazione  $\phi$  di un intorno di  $p$  tale che  $g \circ \phi$  è di classe  $C^k$  (come funzione tra aperti di  $\mathbb{R}^N$ ).*

Si osservi che la definizione non dipende dalla parametrizzazione, perché se  $g \circ \phi$  è di classe  $C^k$  e  $\psi$  è un'altra parametrizzazione di un intorno di  $p$  allora  $g \circ \psi = g \circ \phi \circ \phi^{-1} \circ \psi$  è di classe  $C^k$  nell'intersezione degli intorni, dato che lo sono  $\phi$  e  $\psi$ .

**Definizione 75** (Differenziale). *Sia  $g : \Sigma \rightarrow \Gamma$  una funzione  $C^1$  tra varietà e sia  $p \in \Sigma$ . Il differenziale di una funzione tra varietà è l'applicazione tra i tangenti delle due varietà che manda la derivata di una curva nella derivata della curva immagine. In simboli:*

$$d_p g : T_p(\Sigma) \longrightarrow T_{g(p)}(\Gamma)$$

$$\gamma'(0) \longrightarrow (g \circ \gamma)'(0)$$

<sup>1</sup>Volendo si potrebbe dare una definizione analoga per spazi di Banach a dimensione qualsiasi, ma si dovrebbe richiedere che  $df_x$  sia l'inversa sinistra di un'applicazione lineare continua, che è leggermente più restrittivo. Non ci addenteremo in questi campi.

<sup>2</sup>In realtà quanto stiamo facendo ha come giustificazione un teorema non banale di geometria differenziale. Io quel teorema l'ho visto e fa un po' schifo, quindi glissiamo che è meglio. Tranquilli, funziona tutto.

La definizione è ben posta. Infatti, se un vettore  $v \in T_p \Sigma$  è la velocità in zero di due curve  $\gamma$  e  $\delta$ , detta  $\phi$  una parametrizzazione locale abbiamo

$$(g \circ \gamma)'(0) = (g \circ \phi \circ \phi^{-1} \circ \gamma)'(0) = d(g \circ \phi) \circ d\phi^{-1}(\gamma'(0)) = d(g \circ \phi) \circ d\phi^{-1}(\delta'(0)) = (g \circ \delta)'(0)$$

Il ragionamento mostra altresì che il differenziale è un'applicazione lineare, perché è la composizione di due differenziali classici. Inoltre per costruzione il differenziale di una composizione è la composizione dei differenziali, perché

$$d(f \circ g)(\gamma'(0)) = (f \circ g \circ \gamma)'(0) = (f \circ (g \circ \gamma))'(0) = df(dg(\gamma'(0)))$$

Si osservi che il differenziale classico  $D$  di un'applicazione tra aperti di  $\mathbb{R}^n$  si può vedere come differenziale  $d$  di una funzione tra varietà. Infatti, sia  $f : U \rightarrow V$  una mappa  $C^1$  nel senso usuale, e dato un punto  $x \in U$  e un vettore  $v \in \mathbb{R}^n$  consideriamo la curva  $\gamma(t) = x + tv$ , la cui derivata in zero è  $v$ . Allora

$$d_x f(v) = (f \circ \gamma)'(0) = f(x + tv)'(0) = Df_x[v]$$

In questo caso il tangente all'aperto  $U$  nel punto  $x$  è tutto  $\mathbb{R}^n$ , stavolta pensato come spazio vettoriale e non come spazio ambiente in cui vive  $x$ .

## 10.4 Punti estremanti su sottovarietà

**Teorema 10.4.1.** *Sia  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione di classe  $C^1$  e sia  $\Sigma \subseteq \mathbb{R}^n$  una sottovarietà di classe  $C^1$ . Se  $x_0 \in \Sigma \cap \Omega$  è un punto estremante locale per  $f|_{\Sigma}$ , allora  $\vec{\nabla} f_{x_0} \perp T_{x_0}(\Sigma)$ , ossia  $\vec{\nabla} f_{x_0}$  appartiene allo spazio normale  $N_{x_0}(\Sigma) = T_{x_0}(\Sigma)^\perp$ .*

*Dimostrazione.* Per definizione di sottovarietà, esistono due aperti  $U$  e  $V$  e un diffeomorfismo  $\varphi$  tale che  $\varphi^{-1}(V \cap \Sigma) = \mathbb{R}^k \times \{0\} \cap U$ . Ora, i punti di  $\mathbb{R}^k \times \{0\}$  soddisfano i vincoli  $g^i(x) = x_i = 0$  per ogni  $i = k+1, \dots, n$ . Perciò  $V \cap \Sigma$  è l'intersezione dei luoghi di zeri delle funzioni  $g^i \circ \varphi^{-1}$ , di classe almeno  $C^1$ . Di ognuna di queste funzioni possiamo considerare il gradiente, il quale dev'essere ortogonale al tangente in  $x_0$ . Infatti applicando a  $T_{x_0}(\Sigma)$  il differenziale di  $g^i \circ \varphi^{-1}$  otteniamo

$$d(g^i \circ \varphi^{-1})_{x_0}(T_{x_0}(\Sigma)) = dg^i_{\varphi^{-1}(x_0)} \circ d\varphi_{x_0}(T_{x_0}(\Sigma)) = dg^i_{\varphi^{-1}(x_0)}(\mathbb{R}^k \times \{0\}) = 0$$

Ma il differenziale di  $g^i \circ \varphi^{-1}$  è il prodotto scalare con il suo gradiente, perciò ogni vettore del tangente è ortogonale al gradiente di ogni  $g^i \circ \varphi^{-1}$ .

Per concludere osserviamo che, per principio dei moltiplicatori di Lagrange,  $\vec{\nabla} f_{x_0}$  deve essere contenuto in  $\text{Span}\{\vec{\nabla}(g^i \circ \varphi^{-1})_{x_0}\}_{i=k+1, \dots, n}$ , e dunque deve essere ortogonale al tangente in  $x_0$ .  $\square$

L'analogia con il teorema dei moltiplicatori sta nel fatto che il gradiente di  $f$  appartiene allo spazio normale, che nel caso in cui la varietà è la controimmagine di un valore regolare coincide con lo spazio generato dai gradienti dei vincoli.

## 10.5 Misura su sottovarietà

Sia  $\Sigma \subseteq \mathbb{R}^n$  una  $k$ -varietà  $C^1$  e sia  $\phi : U \subseteq \mathbb{R}^k \rightarrow \Sigma$  una parametrizzazione locale. Per ogni  $A \subseteq U$  misurabile secondo Lebesgue definiamo la funzione di insieme

$$\mu(\phi(A)) = \int_A \sqrt{\det((d_x \phi)^T d_x \phi)} dx$$

La definizione è ben posta, perché per ogni  $x$  la matrice  $(d_x \phi)^T d_x \phi$  è simmetrica e semidefinita positiva, e dunque il suo determinante è non negativo. Inoltre  $\mu$  è una misura, e più precisamente

$$\mu = \phi_{\#} \left( \sqrt{\det((d_x \phi)^T d_x \phi)} dx \right)$$

dove  $\phi_{\#}$  indica la misura push-forward tramite  $\phi$  della misura definita dalla densità  $\sqrt{\det((d_x\phi)^T d_x\phi)}$  a partire dalla misura di Lebesgue.<sup>3</sup> Si osservi come nel caso monodimensionale la misura che abbiamo definito generalizza la nozione di lunghezza di una curva differenziabile, che sappiamo essere data da

$$L(\gamma, a, b) = \int_{[a,b]} \|\gamma'\| dx = \mu(\gamma([a, b]))$$

In analogia col caso delle curve possiamo mostrare che la misura è indipendente dalla parametrizzazione. Dato un cambio di carta  $\psi : U' \subseteq \mathbb{R}^k \rightarrow U$  e un insieme misurabile  $A \in U'$  si ha

$$\mu(\phi \circ \psi(A)) = \int_A \sqrt{\det((d_x\psi)^T (d_{\psi(x)}\phi)^T d_{\psi(x)}\phi d_x\psi)} dx = \int_A \sqrt{\det((d_{\psi(x)}\phi)^T d_{\psi(x)}\phi)} |\det d_x\psi| dx$$

e per la formula di cambio di variabile

$$\mu(\phi \circ \psi(A)) = \int_{\psi(A)} \sqrt{\det((d_x\phi)^T d_x\phi)} dx = \mu(\phi(\psi(A)))$$

Unendo gli intorni parametrizzati otteniamo un ricoprimento aperto<sup>4</sup>  $(U_i)_{i \in \mathbb{N}}$  a cui è associata una classe di misure locali  $(\sigma_i)_{i \in \mathbb{N}}$ . Si osservi che ciascuna coppia di misure coincide sull'intersezione degli aperti di definizione, in virtù dell'invarianza sotto cambi di carta. Da tale ricoprimento estraiamo la partizione  $V_i = U_i \setminus \bigcap_{j < i} U_j$ . Ora, dato un boreliano  $A \subseteq \Sigma$  definiamo la sua misura "globale"

$$\sigma(A) = \sum_{i \in \mathbb{N}} \sigma_i(A \cap V_i)$$

Si osservi che le restrizioni di  $\sigma$  agli aperti  $U_i$  coincidono con le misure locali. Questo ci dice in particolare che  $\sigma$  è unica (e dunque non dipende nemmeno dalla partizione scelta per costruirla) in quanto se un'altra misura  $\mu$  coincide con  $\sigma$  su tutti gli intorni locali deve coincidere con  $\sigma$  ovunque.<sup>5</sup>

### 10.5.1 Lemma di Sard per varietà

**Definizione 76** (Punto critico). *Sia  $f : \Sigma \rightarrow \Gamma$  una funzione  $C^1$  tra varietà della stessa dimensione. Un punto  $p \in \Sigma$  si dice critico se  $d_p f$  non è surgettivo. L'immagine di un punto critico si dice valore critico.*

**Lemma 10.5.1** (Sard per varietà). *Sia  $f : \Sigma \rightarrow \Gamma$  una funzione  $C^1$  tra varietà. I valori critici hanno misura nulla.*

*Dimostrazione.*<sup>6</sup> Dimostriamo il caso in cui le due varietà hanno la stessa dimensione. Consideriamo una famiglia numerabile di aperti  $V'$  di  $\Gamma$  ben parametrizzati e scegliamo un ricoprimento numerabile di  $\Sigma$  con aperti ben parametrizzati  $V$ , in modo che l'immagine di ciascuno di essi giaccia in uno dei  $V'$ . Ci basta mostrare che in ciascuno dei  $V$  i punti critici hanno misura nulla e poi concludere per  $\sigma$ -additività. Siano allora rispettivamente  $\phi : U \rightarrow V$  e  $\psi : U' \rightarrow V'$  le parametrizzazioni di due aperti tali che  $f(V) \subseteq V'$ . Per come è definita la misura un sottoinsieme di  $V$  è trascurabile se lo è la sua preimmagine tramite  $\psi$ . Consideriamo dunque la funzione  $\psi^{-1} \circ f \circ \phi$ . I suoi valori critici sono i punti per cui  $\det d(\psi^{-1} \circ f \circ \phi) = 0$ , che in particolare, per moltiplicatività del differenziale, contengono i punti in cui  $df$  non è invertibile. La tesi segue dunque applicando il lemma di Sard per funzioni tra aperti e ricordando che il differenziale classico si può intendere come il differenziale di una funzione tra varietà.  $\square$

<sup>3</sup>Vedasi 53. Si osservi che la parametrizzazione  $\phi$  è misurabile secondo Lebesgue. Infatti essendo un diffeomorfismo la sua inversa è localmente lipschitziana, e dunque manda misurabili in misurabili.

<sup>4</sup>Possiamo prendere un ricoprimento numerabile dato che la topologia della varietà è quella indotta dall'euclidea, che ha la proprietà di Lindelöf.

<sup>5</sup>Questo fatto è vero in generale: se due misure sui boreliani di uno spazio di Lindelöf  $X$  coincidono localmente (ossia per ogni punto esiste un intorno sui boreliani del quale coincidono) allora esse coincidono sui boreliani di  $X$ .

<sup>6</sup>Questo è un mio tentativo, è una delle miriadi di cose che sono state lasciate come esercizio.

### 10.5.2 Cambio di variabile

Si può dimostrare il seguente teorema, che discende dalla formula di cambio di variabile per funzioni tra aperti di  $\mathbb{R}^n$ :

**Teorema 10.5.2** (Cambio di variabile tra varietà). *Siano  $\Sigma$  e  $\Gamma$  due varietà della stessa dimensione, siano  $\sigma$  e  $\gamma$  le rispettive misure superficiali e sia  $g : \Sigma \rightarrow \Gamma$  un diffeomorfismo  $C^1$  tra varietà (nel senso della definizione 75). Allora per ogni  $u \in L^1(\gamma)$  vale l'uguaglianza*

$$\int_{\Gamma} u(y) d\gamma(y) = \int_{\Sigma} u(g(x)) \sqrt{\det((dg_x)^T dg_x)} d\sigma(x)$$

### 10.5.3 Misura su grafici

Se  $\Sigma = \text{graph}(u)$  con  $u : U \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$ , scegliendo come parametrizzazione la funzione  $g(x) = (x, u(x)) : U \rightarrow \mathbb{R}^{k+1}$  abbiamo che

$$d_x g = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \partial_1 u(x) & \partial_2 u(x) & \dots & \partial_k u(x) & 1 \end{bmatrix}$$

Dunque  $(d_x g)^T d_x g = Id + \vec{\nabla} u \otimes \vec{\nabla} u$ , il cui determinante è  $1 + \|\vec{\nabla} u\|^2$  per quanto sappiamo dall'algebra lineare numerica<sup>7</sup>. Perciò la misura indotta sulla varietà risulta

$$\mu(g(A)) = \int_A \sqrt{\det(d_x g)^T d_x g} dx = \int_A \sqrt{1 + \|\vec{\nabla} u\|^2} dx$$

Consideriamo ora la sfera  $n$ -dimensionale  $\partial B(0, r) \subseteq \mathbb{R}^{n+1}$ . L'emisfero superiore aperto  $E^+$  è un grafico, dunque possiamo definire su di esso la misura  $\sigma_+$ . Facendo lo stesso con l'emisfero inferiore aperto  $E_-$  possiamo definire una misura sui Lebesgue-misurabili della sfera data da

$$\sigma(A) = \sigma_+(A \cap E^+) + \sigma_-(A \cap E^-)$$

Al fine di esplicitare tale misura, detta  $u(x) = \sqrt{1 - \|x\|^2}$  la parametrizzazione di  $E^+$ , supponiamo per semplicità che  $A = g(U) \subseteq E^+$ . La sua misura è data da

$$\sigma(u(U)) = \int_U \sqrt{1 + \frac{\|x\|^2}{1 - \|x\|^2}} dx = \int_U \frac{1}{\sqrt{1 - \|x\|^2}} dx = \int_U \frac{1}{u(x)} dx$$

Si verifica<sup>8</sup> inoltre che la misura sopra definita coincide con

$$\sigma(A) = \frac{n+1}{R} \lambda_{n+1}(C(A))$$

dove  $C(A)$  è il cono centrato nell'origine passante per  $A$ , ossia la combinazione convessa dell'insieme  $A$  con l'origine.

Segue allora banalmente che  $\sigma$  è invariante per rotazioni. Infatti, se  $T$  è una rotazione, dato che il cono di  $T(A)$  è ottenuto ruotando il cono di  $A$  otteniamo

$$\begin{aligned} \sigma(T(A)) &= \frac{n+1}{R} \lambda_{n+1}(C(T(A))) = \frac{n+1}{R} \lambda_{n+1}(T(C(A))) = \frac{n+1}{R} \lambda_{n+1}(T(C(A))) = \\ &= \frac{n+1}{R} |\det T| \lambda_{n+1}(C(A)) = \frac{n+1}{R} \lambda_{n+1}(C(A)) = \sigma(A) \end{aligned}$$

Inoltre, osservando che il cono generato dal bordo di una sfera è la sfera stessa si ottiene la seguente relazione tra la misura del bordo e la derivata della misura:

$$\sigma(\partial_t B(0, t)) = \frac{n+1}{t} \lambda(B(0, t)) = \partial_t \lambda(B(0, t))$$

<sup>7</sup>Vedasi [Bini et al., 1992].

<sup>8</sup>Era un esercizio che non ho voluto o saputo approfondire.

Perché la misura della palla unitaria è della forma  $\alpha t^n$  per omogeneità. La misura appena costruita ci permette estendere l'integrazione in coordinate polari al caso  $n$ -dimensionale. Sia data infatti una funzione integrabile  $f \in \mathcal{L}^1(\mathbb{R}^{n+1})$ . A meno di spezzarla in due parti, possiamo supporre che la funzione sia nulla per gli  $x$  la cui ultima coordinata è negativa. Allora con il cambio di variabile  $\mathbb{R}^+ \times B_n(0, 1) \rightarrow \mathbb{R}^{n+1} \setminus \{0\}$  che manda  $(r, x)$  in  $r(x, \sqrt{1 - \|x\|}) = rg(x)$  otteniamo

$$\int_{\mathbb{R}^{n+1}} f(y)dy = \int_{\mathbb{R}^+ \times \mathbb{S}^{n-1}} f(rx, r\sqrt{1 - \|x\|}) |\det d_{(r,x)}(rg(x))| dr dx$$

Ora,

$$d_{(r,x)}(rg(x)) = \begin{bmatrix} x_1 & r & & & \\ x_2 & & r & & \\ \vdots & & & \ddots & \\ x_n & & & & r \\ \sqrt{1 - \|x\|^2} & -\frac{rx_1}{\sqrt{1 - \|x\|^2}} & -\frac{rx_2}{\sqrt{1 - \|x\|^2}} & \cdots & -\frac{rx_n}{\sqrt{1 - \|x\|^2}} \end{bmatrix}$$

Il cui determinante è  $\frac{r^{n-1}}{\sqrt{1 - \|x\|^2}}$ , cosa che vi invito a verificare<sup>9</sup>. Dunque per come è definita la misura su  $E^+$  si ha

$$\int_{\mathbb{R}^{n+1}} f(y)dy = \int_{\mathbb{R}^+ \times \mathbb{S}^{n-1}} f(rx, r\sqrt{1 - \|x\|}) \frac{r^{n-1}}{\sqrt{1 - \|x\|^2}} dr dx = \int_{\mathbb{R}^+ \times E_n^+} f(rz) r^{n-1} dr d\sigma(z)$$

E per teorema di Fubini

$$\int_{\mathbb{R}^{n+1}} f(y)dy = \int_{\mathbb{R}^+} \left( \int_{E_n^+} f(rz) d\sigma(z) \right) r^{n-1} dr$$

Dunque in generale

$$\int_{\mathbb{R}^{n+1}} f(y)dy = \int_{\mathbb{R}^+} \left( \int_{\mathbb{S}^{n-1}} f(rz) d\sigma(z) \right) r^{n-1} dr$$

### 10.5.4 Insight geometrico, senza le dovute prove

La costruzione vista nel paragrafo precedente è un caso particolare delle più generali misure di Hausdorff, definite sugli spazi metrici a partire dalla funzione diametro. Più precisamente, dato uno spazio metrico  $(X, d)$  e un numero reale positivo  $\alpha$  consideriamo la funzione  $\alpha_\delta(E) = \text{diam}(E)^\alpha$ , definita sull'anello dei sottoinsiemi  $E \subseteq X$  di diametro minore di  $\delta$ . Chiamiamo  $H_\delta^\alpha$  la misura esterna indotta da  $\alpha_\delta$ . Osserviamo che al restringersi di  $\delta$  tale misura cresce, perché la famiglia degli insiemi di diametro al più  $\delta$  si restringe e la misura esterna è ottenuta come l'estremo inferiore sui ricoprimenti con insiemi della famiglia considerata. Allora poniamo  $H^\alpha = \sup_{\delta > 0} H_\delta^\alpha$  che è ancora una misura esterna. Non è banale mostrare che tra gli insiemi misurabili secondo Carathéodory vi sono i boreliani, e che dunque è ben definita una misura  $\alpha$ -dimensionale su  $(X, \mathbb{B}(X))$ .

Nel caso in cui  $X = \mathbb{R}^n$  dalla definizione si ricava che  $H^\alpha$  è invariante per traslazioni ed  $\alpha$ -omogenea, ossia tale che

$$H^\alpha(\rho E) = \rho^\alpha H^\alpha(E)$$

Inoltre se  $\alpha = k \in \mathbb{N}$  la restrizione ai sottospazi lineari di dimensione  $\alpha$  della misura  $\alpha$ -dimensionale coincide con la misura di Lebesgue, perché si può mostrare che vale 1 sul cubo unitario.

## 10.6 Teorema della divergenza

**Definizione 77** (Funzione differenziabile su un chiuso). *Sia  $B \subseteq \mathbb{R}^n$  un chiuso. Diciamo che una funzione  $u : B \rightarrow \mathbb{R}^m$  è di classe  $C^k$  su  $B$  se è la restrizione di una funzione di classe  $C^k$  definita su un aperto  $A \supseteq B$ .*<sup>10</sup>

<sup>9</sup>Il conto è facile, ma scriverlo in L<sup>A</sup>T<sub>E</sub>X è una gran pena. Spero che il lettore perdonerà la mia accidia.

<sup>10</sup>Si potrebbero dare definizioni meno stringenti, ma vogliamo evitare di preoccuparci troppo del bordo di  $B$ .

**Definizione 78** (Bordo  $C^k$ ). Un aperto  $\Omega \subseteq \mathbb{R}^n$  ha bordo  $C^k$  se per ogni  $p \in \partial\Omega$  esiste un suo intorno  $U \subseteq \mathbb{R}^n$  tale che  $U \cap \partial\Omega$  è il grafico di una funzione  $C^k$  di cui  $U \cap \Omega$  è il sottografico.

**Teorema 10.6.1.** Sia  $\Omega \subseteq \mathbb{R}^n$  un aperto limitato a bordo  $C^1$  (a meno di un insieme trascurabile rispetto alla misura  $(n-1)$ -dimensionale su  $\mathbb{R}^n$ ). Sia  $u \in C^1(\overline{\Omega})$ . Detta  $\{e_i\}_{i \leq n}$  la base canonica e detto  $w_{ext}(p)$  il versore normale al tangente del bordo nel punto  $p \in \partial\Omega$ , in direzione uscente rispetto all'interno della varietà, si ha

$$\int_{\Omega} \partial_i u(y) dy = \int_{\partial\Omega} u(p)(w_{ext}(p))_i d\sigma(p)$$

Si osservi come nel caso 1-dimensionale il teorema coincida con il teorema fondamentale del calcolo integrale, dato che la misura di Lebesgue in dimensione zero è la contapunti e il versore normale vale 1 per l'estremo di destra e  $-1$  per l'estremo di sinistra.

*Dimostrazione.* Consideriamo dapprima il caso in cui deriviamo rispetto all'ultima variabile e  $\Omega$  è della forma

$$\Omega = \{(x, t) \text{ t.c. } x \in U, f_-(x) < t < f_+(x)\}$$

dove  $U \subseteq \mathbb{R}^{n-1}$  è un aperto limitato con bordo  $C^1$  e  $f_+, f_- : U \rightarrow \mathbb{R}$  sono funzioni di classe  $C^1$ , con  $f_- \leq 0 \leq f_+$ . Il bordo di  $\Omega$  risulta allora una sorta di "cilindro" le cui basi sono i grafici delle due funzioni. Sia  $u \in C^1(\overline{\Omega}, \mathbb{R})$ . Posto  $y = (x, t)$  si ha

$$\int_{\Omega} \partial_n u(y) dy = \int_{U \times \mathbb{R}} \partial_t u(x, t) \chi_{\Omega}(x, t) d(x, t)$$

L'integrando è di fatto integrabile, in quanto la funzione continua  $\partial_t u$  è limitata in modulo sul compatto  $\overline{\Omega}$ . Allora possiamo applicare il teorema di Fubini e ottenere

$$\begin{aligned} \int_{\Omega} \partial_n u(y) dy &= \int_U \left( \int_{\mathbb{R}} \partial_t u(x, t) \chi_{(f_-(x), f_+(x))}(t) dt \right) dx = \\ &= \int_U \left( \int_{f_-(x)}^{f_+(x)} \partial_t u(x, t) dt \right) dx = \int_U u(x, f_+(x)) dx - \int_U u(x, f_-(x)) dx \end{aligned}$$

Sia ora  $\Sigma_+ = \text{graph}(f_+)$  e sia  $w(x, f_+(x)) = \frac{(-\vec{\nabla} f_+(x), 1)}{\sqrt{1 + \|\vec{\nabla} f_+(x)\|^2}}$  il versore ortogonale al tangente. Si osservi che esso punta verso l'alto, dato che la sua ultima coordinata è positiva, e perciò è diretto verso l'esterno di  $\Omega^{11}$ . Visto che  $(w(x, f_+(x)))_n = \frac{1}{\sqrt{1 + \|\vec{\nabla} f_+(x)\|^2}}$ , possiamo scrivere

$$\int_U u(x, f_+(x)) dx = \int_U u(x, f_+(x))(w(x, f_+(x)))_n \sqrt{1 + \|\vec{\nabla} f_+\|^2} dx$$

Si osservi che  $\sqrt{1 + \|\vec{\nabla} f_+\|^2}$  è la densità che definisce la misura  $d\sigma$  su  $\Sigma_+$ , come si è visto in 10.5.3. Possiamo fare lo stesso ragionamento per  $f_-$ , osservando però che il versore ortogonale al grafico di  $f_-$  che punta verso l'esterno è l'opposto di  $w(x, f_-(x)) = \frac{(-\vec{\nabla} f_-(x), 1)}{\sqrt{1 + \|\vec{\nabla} f_-(x)\|^2}}$ . Dunque otteniamo

$$\int_{\Omega} \partial_n u(y) dy = \int_U u(x, f_+(x)) dx - \int_U u(x, f_-(x)) dx = \int_{\partial\Omega} u(y)(w_{ext}(y))_n d\sigma(y)$$

Dove con  $w_{ext}$  si intende il versore diretto verso l'esterno di  $\Omega$  (che per  $f_+$  coincide con  $w$ , mentre per  $f_-$  è  $-w$ ). Si osservi che abbiamo esteso l'integrale a tutto il bordo di  $\Omega$ . Infatti il contributo della faccia laterale del cilindro<sup>12</sup> è nulla, in quanto l'ultima coordinata del versore normale in questo caso è sempre zero.

<sup>11</sup>Vedasi 10.3.1.

<sup>12</sup>Ha senso parlarne perché anche il bordo laterale è una  $(n-1)$ -varietà, ottenuta "estrudendo" il bordo di  $U$ .

Ora, estendiamo il teorema al caso in cui  $\Omega$  è un qualsiasi aperto limitato a bordo  $C^1$ . WLOG sia  $i = n$ , e considerando  $\Omega$  come un sottoinsieme di  $\mathbb{R}^{n-1} \times \mathbb{R}$  sia  $P$  il proiettore su  $\mathbb{R}^{n-1}$ . Allora, detto

$$C = \{x \in \partial\Omega \text{ t.c. } w_{ext} \perp e_n\}$$

per il lemma di Sard<sup>13</sup> la proiezione  $P(C)$  è trascurabile rispetto alla misura  $(n-1)$ -dimensionale, in quanto è composta dalle immagini dei punti per i quali la proiezione non è una bigezione tra il tangente della varietà e il tangente di  $\mathbb{R}^{n-1}$ . Dunque possiamo procedere come se  $\partial\Omega$  non contenesse punti dove  $w_{ext} \perp e_n$ . Infatti, da un lato l'integrale su  $\partial\Omega$  è identicamente nullo sui punti critici  $C$ , dall'altro i punti di  $\Omega$  la cui proiezione è un valore critico sono contenuti nell'insieme  $P(C) \times \mathbb{R}$ , trascurabile in  $\mathbb{R}^n$ .

Ora, sia  $V$  una componente connessa di  $P(\Omega)$ . Per ogni  $x \in V$  l'intersezione  $P^{-1}(x) \cap \partial\Omega$  è un sottoinsieme compatto e discreto di  $\mathbb{R}$ . Infatti per ogni punto  $p$  dell'intersezione esiste un intorno  $U$  dove  $\partial\Omega \cap U$  è il grafico di una funzione di cui  $\Omega \cap U$  è il sottografico. In tale intorno il grafico si approssima localmente con il suo tangente, che interseca la retta  $P^{-1}(x)$  in un punto perché  $(w_{ext}(p))_n \neq 0$ . Dunque l'intersezione è discreta, ed essendo compatta ha cardinalità finita. Si osservi poi che esiste un intorno aperto  $x \in A \subseteq V$  tale che  $P^{-1}(A) \cap \Omega$  è un'unione di aperti della forma  $\{f_- < y < f_+\}$ . Dunque il compatto  $V$  si ricopre con un numero finito di questi aperti, che possiamo considerare disgiunti a meno di eliminarne il bordo (trascurabile rispetto alla misura  $(n-1)$ -dimensionale). In altri termini  $P^{-1}(V)$  è un'unione finita e disgiunta di aperti della forma del caso base, per cui vale la tesi. Si conclude osservando che le componenti connesse di  $P(\Omega)$  hanno parte interna non vuota, e dunque sono numerabili in quanto si può associare a ognuna di esse un diverso vettore a coordinate razionali. Perciò l'integrale su tutto  $\Omega$  è la serie degli integrali sulle componenti connesse (per teorema di convergenza dominata, dato che sul compatto  $\bar{\Omega}$  tutte le nostre funzioni sono derivabili in quando continue), da cui la tesi.  $\square$

Per arrivare alla formulazione classica del teorema della divergenza ci servirà un'altra definizione.

**Definizione 79** (Divergenza). *Sia  $v \in C^1(\bar{\Omega}, \mathbb{R}^n)$  un campo vettoriale. La sua divergenza è lo scalare*

$$\operatorname{div}(v)(x) = \sum_{i=1}^n \partial_i v_i(x) = \operatorname{tr} d_x v$$

Sommando le relazioni ottenute nel teorema precedente si ricava il seguente:

**Corollario 10.6.1.1** (Teorema della divergenza). *Sia  $v \in C^1(\bar{\Omega}, \mathbb{R}^n)$  un campo vettoriale. Allora*

$$\int_{\Omega} \operatorname{div}(v)(x) dx = \int_{\partial\Omega} \langle v, w_{ext} \rangle(p) d\sigma(p)$$

La RHS si interpreta fisicamente come il flusso del campo attraverso la superficie  $\partial\Omega$ . In tal senso la divergenza di un punto è una misura di "quante" linee di campo vengano generate o assorbite da esso. Quella del corollario non è l'unica formula che si può ottenere smanettando con il teorema di questa sezione. Ad esempio, se  $f, g$  sono funzioni scalari e  $v$  un campo vettoriale, valgono le seguenti uguaglianze:

- $\int_{\Omega} \vec{\nabla} u dx = \int_{\partial\Omega} f(p) w_{ext}(p) d\sigma(p)$ ;
- $\int_{\Omega} f \operatorname{div}(v)(x) \langle v, \vec{\nabla} f \rangle(x) dx = \int_{\partial\Omega} f(p) \langle v, w_{ext} \rangle(p) d\sigma(p)$ ;
- $\int_{\Omega} \nabla^2 f(x) dx = \int_{\partial\Omega} \langle \vec{\nabla} f, w_{ext} \rangle(p) d\sigma(p)$ , dove il laplaciano  $\nabla^2$  è la divergenza del gradiente;
- $\int_{\Omega} f(x) \nabla^2 g(x) dx + \int_{\Omega} \langle \vec{\nabla} f, \vec{\nabla} g \rangle(x) dx = \int_{\partial\Omega} f(p) \langle \vec{\nabla} g, w_{ext} \rangle(p) d\sigma(p)$  (prima identità di Green);
- $\int_{\Omega} f(x) \nabla^2 g(x) dx - \int_{\Omega} g(x) \nabla^2 f(x) dx = \int_{\partial\Omega} f(p) \langle \vec{\nabla} g, w_{ext} \rangle(p) d\sigma(p) - \int_{\partial\Omega} g(p) \langle \vec{\nabla} f, w_{ext} \rangle(p) d\sigma(p)$  (seconda identità di Green, null'altro che la differenza tra la prima e la simmetrica);

<sup>13</sup>Vedasi 10.5.1.



# Capitolo 11

## Equazioni differenziali ordinarie lineari (EDOL)

Ci occuperemo ora di sviscerare vita morte e miracoli di equazioni del tipo

$$u'(t) = A(t)u(t) + b(t)$$

con  $t$  che varia in un intervallo  $I \subseteq \mathbb{R}$ ,  $u, b$  funzioni continue da  $I$  a  $\mathbb{C}^n$  e  $A : I \rightarrow M(n, \mathbb{C})$  una funzione continua a valore nelle matrici complesse, munite della topologia euclidea.<sup>1</sup> Sarebbe possibile, ma leggermente più tecnico, lavorare in spazi di Banach, perciò evitiamo per quanto possibile.

### 11.1 Esponenziale di matrici

Data  $A \in M(n, \mathbb{C})$ , definiamo la sua matrice esponenziale come

$$\exp(A) = \sum_{k=0}^{+\infty} \frac{A^k}{k!}$$

Osserviamo innanzitutto che la definizione è ben posta, ossia che la serie delle somme parziali ammette limite. Infatti, passando a una qualsiasi norma operatoriale, osserviamo che

$$\sum_{k=0}^{+\infty} \frac{\|A^k\|}{k!} \leq \sum_{k=0}^{+\infty} \frac{\|A\|^k}{k!} = e^{\|A\|}$$

Dunque la successione delle somme parziali converge totalmente, perciò converge anche nella distanza indotta dalla norma in quanto lo spazio delle funzioni lineari continue è di Banach.

Ora,  $\exp : M(n, \mathbb{C}) \rightarrow M(n, \mathbb{C})$  è una funzione continua. Infatti le somme parziali  $S_N$  sono ovviamente continue, in quanto sono “polinomi” di matrici, e convergono uniformemente sui limitati a  $\exp$ . Questo perché, considerando la norma uniforme della differenza su una palla di matrici  $B(0, r) \subseteq (M(n, \mathbb{C}), \|\cdot\|)$ , abbiamo

$$\|\exp - S_N\|_{\infty, B(0, r)} \leq \sup_{\|A\| \leq r} \sum_{k=N+1}^{+\infty} \frac{\|A\|^k}{k!} \leq \sum_{k=N+1}^{+\infty} \frac{r^k}{k!} \rightarrow 0$$

Dunque  $\exp$  è (localmente) limite uniforme di funzioni continue, e dunque è continua.

Osserviamo infine che  $\exp$  è anche differenziabile in ogni punto. Infatti è limite localmente uniforme delle  $S_n$ , che sono differenziabili e convergono in  $\exp(0) = I$ . Allora

$$d\exp_A(H) = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{j=0}^{k-1} A^j H A^{k-1-j}$$

---

<sup>1</sup>Osserviamo che tale topologia è quella indotta da una qualsiasi norma operatoriale, perché in uno spazio vettoriale a dimensione finita tutte le norme sono equivalenti (vedasi 2.4.1).

In particolare, se  $A$  e  $H$  commutano, il differenziale assume una forma abbastanza carina:

$$d \exp_A(H) = H \sum_{k=1}^{+\infty} \frac{A^{k-1}}{(k-1)!} = H \exp(A)$$

Da quanto detto finora segue anche che la funzione  $f : I \rightarrow M(n, \mathbb{R})$  definita come  $f(t) = \exp(tA)$  è derivabile con derivata  $f'(t) = A \exp(tA)$ .

Infine, praticamente con la stessa dimostrazione vista per il caso unidimensionale si può mostrare che

$$\exp(A) = \lim_{n \rightarrow +\infty} \left( I + \frac{A}{n} \right)^n$$

### 11.1.1 Proprietà dell'esponenziale

1. Se  $AB = BA$ , allora  $\exp(A)B = B \exp(A)$ , in quanto la moltiplicazione per  $B$  è un'applicazione lineare continua e dunque si può portare sotto il segno di sommatoria. Mostriamo ora che, sotto le stesse ipotesi,  $\exp(A) \exp(B) = \exp(A+B)$ .

Consideriamo innanzitutto la funzione  $f(t) = e^{tA} e^{tB} e^{-t(A+B)}$ . La sua derivata è

$$f'(t) = A e^{tA} e^{tB} e^{-t(A+B)} + e^{tA} B e^{tB} e^{-t(A+B)} - e^{tA} e^{tB} (A+B) e^{-t(A+B)}$$

Poiché tutto commuta, la derivata diventa

$$f'(t) = (A+B-A-B) e^{tA} e^{tB} e^{-t(A+B)} = 0$$

Allora la funzione considerata è costante in  $t$  e deve quindi valere che

$$f(1) = e^A e^B e^{-(A+B)} = f(0) = I$$

Osserviamo che, se nel caso precedente poniamo  $B = 0$ , otteniamo  $e^A e^{-A} = \text{cost} = I$ , cioè  $e^A$  è invertibile per ogni matrice  $A$  e ha come inversa  $e^{-A}$ . Allora possiamo invertire la relazione e ottenere  $e^A e^B = e^{A+B}$ , come voluto.

2. Se  $P \in GL$ , allora  $e^{PAP^{-1}} = P e^A P^{-1}$ , perché l'operatore di coniugio è lineare e continuo.
3. Se  $A$  è a blocchi, della forma  $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$ , allora  $e^A = \begin{bmatrix} e^B & 0 \\ 0 & e^C \end{bmatrix}$ , ovviamente.
4. Mettendo insieme i pezzi, possiamo scrivere una forma esplicita per l'esponenziale di una matrice complessa. Infatti, sia  $J = PAP^{-1}$  la forma di Jordan di  $A$ . Possiamo limitarci a studiare  $J$ , per la proprietà 2, e considerare cosa succede a un singolo blocco di Jordan, per la proprietà 3. Ora, posto  $J = \lambda I + N$ , poiché l'identità commuta con tutto abbiamo che

$$\exp(J) = \exp(\lambda I + N) = \exp(\lambda I) \exp(N)$$

Il primo pezzo è semplicemente  $e^\lambda I$ . Il secondo è una somma finita perché  $N$  è nilpotente, ed è dato da una matrice triangolare superiore la cui diagonale  $k$ -esima ha tutti gli elementi pari a  $1/k!$ . Si ottiene allora

$$\exp(J) = e^\lambda \begin{bmatrix} 1 & 1 & \frac{1}{2} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & 1 & \dots & \frac{1}{(n-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

5. Ricordando che la traccia è la somma degli autovalori, dalla proprietà 4 si ottiene una formula per il determinante dell'esponenziale:

$$\det(\exp(A)) = \det(\exp(J)) = e^{\text{tr}(A)}$$

### 11.1.2 Immagine della mappa esponenziale

**Teorema 11.1.1.** *La mappa  $\exp : M(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C})$  è surgettiva, ossia ogni matrice invertibile è l'esponenziale di qualcosa.*

*Dimostrazione.* Innanzitutto, poiché  $P \exp(A) P^{-1} = \exp(PAP^{-1})$  l'immagine dell'esponenziale è chiusa per coniugio, dunque basta mostrare che un blocco di Jordan  $J = \lambda I + N$ , con  $\lambda \neq 0$ , è l'esponenziale di qualcosa. Allora scriviamo

$$\lambda I + N = \lambda \left( I + \frac{1}{\lambda} N \right) = e^\mu \exp(L)$$

Per un certo numero complesso  $\mu$ , che sappiamo esistere perché l'esponenziale complessa è surgettiva, e per una certa matrice  $L$ , che possiamo scegliere essere la serie logaritmica di  $\frac{1}{\lambda} N$ :

$$L = \log \left( I + \frac{1}{\lambda} N \right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k} \left( \frac{1}{\lambda} N \right)^k$$

Questa serie converge perché la somma è finita (dato che  $N$  è nilpotente), ed è un'inversa dell'esponenziale (sebbene non sia l'unica). Infatti, l'uguaglianza  $\exp(\log(1+t)) = 1+t$  è vera come identità tra serie di potenze, a prescindere dal valore che assuma  $t$ .  $\square$

Per quanto riguarda le matrici reali, dovrebbe essere vero il seguente teorema<sup>2</sup>, che Majer si ostina a procrastinare nonostante le mie incalzanti domande:

**Teorema 11.1.2.** *Data  $A \in GL(n, \mathbb{R})$ , esiste  $B$  tale che  $A = \exp(B)$  se e solo se esiste  $C$  tale che  $A = C^2$ .*

## 11.2 Sistemi di EDO a coefficienti costanti

Innanzitutto, osserviamo che possiamo restringerci ai sistemi di equazioni lineari, perché un'equazione di ordine  $n$ , della forma

$$p(D)u = u^{(n)} + \sum_{k=0}^{n-1} a_k u^{(k)} = f$$

Si può ridurre allo studio del sistema lineare di  $n+1$  incognite

$$\begin{cases} x'_0 = x_1 \\ x'_1 = x_2 \\ \dots \\ x'_{n-1} = -\sum_{k=0}^{n-1} a_k x_k + f \end{cases}$$

Osserviamo che la matrice che rappresenta questo sistema è la matrice compagna del polinomio  $p(x)$ , ossia

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}$$

In generale, supponiamo di voler risolvere il problema di Cauchy

$$\begin{cases} u'(t) = Au(t) + b(t) \\ u(t_0) = u_0 \end{cases}$$

Dove  $A \in M(\mathbb{C}, n)$ ,  $b$  è una funzione continua da un intervallo aperto  $I \subseteq \mathbb{R}$  in  $\mathbb{C}^n$ ,  $u$  è una funzione da un sottointervallo di  $I$  in  $\mathbb{C}^n$  e  $u_0 \in \mathbb{C}^n$ .

<sup>2</sup>Per gli interessati rimando alle referenze della pagina [https://en.wikipedia.org/wiki/Logarithm\\_of\\_a\\_matrix](https://en.wikipedia.org/wiki/Logarithm_of_a_matrix).

**Teorema 11.2.1.** *Nelle ipotesi dette sopra, la soluzione  $u(t)$  esiste ed è unica ed è definita globalmente sul dominio di  $b(t)$ .*

*Dimostrazione.* Moltiplichiamo a sinistra entrambi i membri per  $e^{-tA}$ . Otteniamo

$$\begin{aligned} e^{-tA}u'(t) &= e^{-tA}Au(t) + e^{-tA}b(t) \\ (e^{-tA}u)' &= e^{-tA}b(t) \\ e^{-tA}u(t) - e^{-t_0A}u(t_0) &= \int_{t_0}^t e^{-xA}b(x)dx \\ u(t) &= e^{(t-t_0)A}u(t_0) + e^{tA} \int_{t_0}^t e^{-xA}b(x)dx \end{aligned}$$

Questo prova in un colpo solo l'esistenza globale di una soluzione (perché effettivamente questa funzione è ben definita e soddisfa l'equazione) e la sua unicità (se esiste, coincide con quella trovata).  $\square$

**Definizione 80.** *Il flusso di un'equazione differenziale (definita globalmente per tempi  $t \in I$  e condizioni iniziali  $u_0 \in X$ ) è la funzione*

$$\begin{aligned} \Phi : I^2 \times X &\longrightarrow X \\ (t, t_0, u_0) &\longrightarrow u(t) \end{aligned}$$

Dove  $u$  è la soluzione del problema di Cauchy associato di condizione iniziale  $u(t_0) = u_0$ .

Nel caso appena trattato il flusso è

$$u(t) = e^{(t-t_0)A}u(t_0) + e^{tA} \int_{t_0}^t e^{-xA}b(x)dx$$

Osserviamo che, a  $t$  fissato, questa funzione è continua in tutte le variabili e nei dati del problema  $A, b$ . In altri termini, perturbando le condizioni iniziali e le equazioni del sistema si ottengono soluzioni "vicine puntualmente" alla soluzione originale. Ciò però non significa che le soluzioni sono vicine uniformemente in  $t$ . Ad esempio, mentre le soluzioni di  $u' = \varepsilon u$  con  $\varepsilon = 0$  sono costanti, basta modificare  $\varepsilon$  per ottenere soluzioni infinite o infinitesime per  $t \rightarrow +\infty$ .

### 11.3 Stabilità delle soluzioni omogenee

**Teorema 11.3.1** (Soluzioni infinitesime). *Sia data l'equazione  $u'(t) = Au(t)$ . Le seguenti affermazioni sono equivalenti:*

1. *Tutte le soluzioni, al variare dei parametri iniziali, sono infinitesime per  $t \rightarrow +\infty$ ;*
2. *Lo spettro di  $A$  ha solo autovalori di parte reale negativa;*
3.  *$\|e^{tA}\| \rightarrow 0$  per  $t \rightarrow +\infty$  per una qualsiasi norma matriciale.*

*Dimostrazione.* Mostriamo nell'ordine le seguenti implicazioni.

- (1  $\implies$  2) Se esiste un autovalore di parte reale non negativa, sia esso  $\lambda$  e  $v$  un suo autovettore non nullo. Allora la funzione  $u(t) = e^{\lambda t}v$  risolve il problema di Cauchy

$$\begin{cases} u'(t) = Au(t) \\ u(0) = v \end{cases}$$

ed è dunque una soluzione complessa del problema, la quale tuttavia diverge per  $t \rightarrow +\infty$ . Questo conclude se  $A$  è complessa. Nel caso reale dobbiamo esibire una soluzione reale con lo stesso comportamento. Osserviamo che la coniugata  $\overline{u(t)}$  è soluzione dell'equazione, dato che, posto  $u = a + ib$  con  $a, b$  reali,

$$\overline{u}' = (a - ib)' = a' - ib' = \overline{u}' = \overline{Au} = A\overline{u}$$

Dunque per linearità  $\Re(u(t)) = \frac{1}{2}(u(t) + \overline{u(t)})$  è una soluzione reale, che ha lo stesso comportamento di  $u$  all'infinito perché

$$\Re(e^{t\lambda}) = e^{t\Re(\lambda)}$$

- (2  $\implies$  3) Se  $J = PAP^{-1}$  è la forma di Jordan di  $A$ , vale che

$$\|e^{tA}\| \leq \|P\| \|P^{-1}\| \|e^{tJ}\|$$

Stimiamo la norma di  $e^{tJ}$  con la somma delle norme dei singoli blocchi, che sono della forma

$$e^{tJ} = e^{t\lambda I} e^{tN} = e^{\lambda t} \begin{bmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Con un po' di disuguaglianza triangolari si mostra che la norma di questa matrice è limitata, e dunque che la norma del singolo blocco va come  $e^{\lambda t}$ . Quest'ultimo è infinitesimo per  $t \rightarrow +\infty$  perché l'esponente ha parte reale negativa.

- (3  $\implies$  1) Basta osservare che  $\|u(t)\| \leq \|e^{tA}\| \|u_0\|$ .

□

In maniera del tutto analoga si mostrano i seguenti teoremi, che caratterizzano le altre soluzioni stabili:

**Teorema 11.3.2** (Soluzioni stabili in avanti). *Le seguenti affermazioni sono equivalenti:*

1. Tutte le soluzioni, al variare dei parametri iniziali, sono limitate per  $t \in [0, +\infty[$ ;
2. Lo spettro di  $A$  ha solo autovalori di parte reale negativa o nulla. Inoltre, gli autovalori immaginari puri sono semisemplici, ossia hanno molteplicità algebrica uguale a quella geometrica<sup>3</sup>;
3.  $\|e^{tA}\|$  è limitata per  $t \in [0, +\infty[$  per una qualsiasi norma matriciale.

Caratterizzando analogamente le soluzioni stabili all'indietro, otteniamo il seguente:

**Teorema 11.3.3** (Soluzioni globalmente limitate). *Le seguenti affermazioni sono equivalenti:*

1. Tutte le soluzioni, al variare dei parametri iniziali, sono limitate per  $t \in ]-\infty, +\infty[$ ;
2. Lo spettro di  $A$  ha solo autovalori immaginari puri semisemplici, e dunque in particolare  $A$  è diagonalizzabile in  $\mathbb{C}$ ;
3.  $\|e^{tA}\|$  è limitata per  $t \in ]-\infty, +\infty[$  per una qualsiasi norma matriciale.

## 11.4 Sistema lineare generico

**Teorema 11.4.1.** *Il problema di Cauchy per il sistema lineare*

$$\begin{cases} u'(t) = A(t)u(t) + b(t) \\ u(t_0) = u_0 \end{cases}$$

*ammette una e una sola soluzione definita globalmente sul dominio  $I$  di  $t$ .*

<sup>3</sup>Questa aggiunta è necessaria perché altrimenti la norma del blocco di Jordan sarebbe dominata dai  $t$  che compaiono sulla sopradiagonale nell'espansione di  $e^{tN}$ .

### 11.4.1 Unicità

Supponiamo di avere due soluzioni  $v$  e  $w$ . Allora la loro differenza  $u = v - w$  soddisfa il problema di Cauchy

$$\begin{cases} u'(t) = A(t)u(t) \\ u(t_0) = 0 \end{cases}$$

È allora sufficiente mostrare che tale sistema ha una sola soluzione.

Se  $u$  risolve, il suo luogo di zeri è sia chiuso (cosa vera per ogni luogo di zeri) che aperto. Infatti, se  $u(y) = 0$ , scegliamo un intorno  $J \subseteq I$  di  $y$  tale che  $\|A\|_{\infty, J}|J| < 1$ . Un tale intorno esiste perché  $A$  è continua, e dunque localmente limitata. Adesso, possiamo scrivere la soluzione nella forma integrale equivalente<sup>4</sup>

$$u(t) = u(y) + \int_y^t A(x)u(x) dx = \int_y^t A(x)u(x) dx$$

Passando alla norma uniforme su  $J$  e stimando in modo bovino l'integrale otteniamo

$$\|u(t)\|_{\infty, J} \leq \|A(x)\|_{\infty, J}|J|\|u(x)\|_{\infty, J}$$

Poiché  $\|A\|_{\infty, J}|J| < 1$  dobbiamo concludere che  $\|u(t)\|_{\infty, J} = 0$ , cioè che  $u$  è identicamente nulla su  $J$ . In altri termini, abbiamo appena esibito un intorno  $J$  di  $y$  totalmente contenuto nel luogo di zeri di  $u$ , e dunque tale insieme è aperto. Poiché  $I$  è connesso, il suo unico sottoinsieme aperto e chiuso non vuoto è lui stesso, perciò  $u \equiv 0$  su tutto  $I$ .

### 11.4.2 Matrice di Wronski

É uno strumentopolo misterioso  
che ci servirà più tardi.

---

*Topolino*

Per mostrare esplicitamente una soluzione del problema di Cauchy, costruiremo una matrice che fa il gioco della funzione flusso, definita in 11.2.

**Teorema 11.4.2.** *Esiste una e una sola funzione  $W \in C^0(I \times I, M(n, \mathbb{C}))$ , detta matrice di Wronski, che soddisfa il sistema*

$$\begin{cases} \partial_t W(t, s) = A(t)W(t, s) \\ W(s, s) = Id \end{cases}$$

*Dimostrazione.* Mostriamo che, comunque si scelga  $J \subseteq I$  compatto, esiste una e una sola soluzione definita su  $J$ . Questo risultato ci permetterà poi di raccordare le soluzioni per definire una funzione globale su  $I$ . Ora, il problema è equivalente al seguente problema integrale:

$$W(t, s) = Id + \int_s^t A(x)W(x, s) dx$$

In altri termini, stiamo cercando un punto fisso dell'operatore  $\Gamma$  da  $C^0(J \times J, M(n, \mathbb{C}))$  in sé definito come

$$\Gamma(B)(t, s) = Id + \int_s^t A(x)B(x, s) dx$$

Ora,  $C^0(J \times J, M(n, \mathbb{C})) = BC(J \times J, M(n, \mathbb{C}))$ , perché il dominio è compatto e quindi ogni funzione ha immagine compatta in  $M(n, \mathbb{C})$ , cioè in particolare di norma limitata. Questo significa che  $\Gamma$  è definita su uno spazio metrico completo con la distanza uniforme, e dunque per mostrare che ammette un punto fisso

---

<sup>4</sup>Si osservi che le soluzioni del problema integrale sono le stesse che risolvono il problema di Cauchy, in quanto in entrambi i casi si richiede che le soluzioni abbiano regolarità  $C^1$ .

ci basta far vedere che  $\Gamma^k$  è una contrazione per qualche  $k \in \mathbb{N}$ .<sup>5</sup> Ora, date  $X, Y \in BC(J \times J, M(n, \mathbb{C}))$ , vale che

$$\Gamma^k(X)(t, s) - \Gamma^k(Y)(t, s) = \int_s^t \dots \int_s^{x_2} \int_s^{x_1} A(x_{k-1}) \dots A(x_0)(X(x_0, s) - Y(x_0, s)) dx_0 dx_1 \dots dx_{k-1}$$

Infatti, la formula è evidentemente vera per  $k = 1$ , e se è vera per  $k = n$  vale che

$$\begin{aligned} \Gamma^{n+1}(X)(t, s) - \Gamma^{n+1}(Y)(t, s) &= \Gamma^n(\Gamma(X))(t, s) - \Gamma^n(\Gamma(Y))(t, s) = \\ &= \int_s^t \dots \int_s^{x_1} A(x_{n-1}) \dots A(x_0)(\Gamma(X)(x_0, s) - \Gamma(Y)(x_0, s)) dx_0 dx_1 \dots dx_{n-1} = \\ &= \int_s^t \dots \int_s^{x_1} \int_s^{x_0} A(x_{n-1}) \dots A(x_0) A(x) X(x, s) - \Gamma(Y)(x, s) dx dx_0 dx_1 \dots dx_{n-1} \end{aligned}$$

Che è la formula cercata, a meno di cambiare nome alle variabili mute. Ora, la norma uniforme della differenza tra le immagini si stima come

$$\begin{aligned} \|\Gamma^k(X) - \Gamma^k(Y)\| &\leq \|X - Y\| \|A\|^k \sup_{t, s \in J} \int_s^t \dots \int_s^{x_2} \int_s^{x_1} dx_0 dx_1 \dots dx_{k-1} = \\ &= \|X - Y\| \|A\|^k \sup_{t, s \in J} \frac{(t-s)^k}{k!} = \|X - Y\| \|A\|^k \frac{|J|^k}{k!} \end{aligned}$$

Definitivamente in  $J$ , la quantità  $\|A\|^k \frac{|J|^k}{k!}$  è strettamente minore di 1, da cui segue che  $\Gamma^k$  è una contrazione.  $\square$

Con facili verifiche si ottiene la soluzione generale dell'equazione omogenea:

**Corollario 11.4.2.1** (Soluzione dell'omogenea). *Il sistema*

$$\begin{cases} u'(t) = A(t)u(t) \\ u(t_0) = u_0 \end{cases}$$

Ha come unica soluzione

$$u(t) = W(t, t_0)u_0$$

In altri termini, il flusso  $\Phi(t, t_0, u_0) = W(t, t_0)u_0$  è lineare nel valore iniziale  $u_0$ .

Teniamo a mente la costruzione della matrice Wronskiana come punto fisso, perché ripeteremo esattamente gli stessi passaggi nella dimostrazione del teorema di Cauchy-Lipschitz.

### 11.4.3 Proprietà della matrice Wronskiana

Prima che il lettore salti a piedi uniti questa parte, ci tengo ad osservare la cosa seguente:

**Lemma 11.4.3.** *Se  $A(t)$  e  $A(s)$  commutano per ogni  $t, s \in I$ , la matrice Wronskiana assume la forma*

$$W(t, s) = \exp\left(\int_s^t A(x) dx\right)$$

In particolare, se  $A$  è costante, ritroviamo il flusso dell'equazione a coefficienti costanti  $u(t) = \exp(tA)u_0$ .

*Dimostrazione.* La funzione delle ipotesi risolve il problema che definisce la matrice Wronskiana. Infatti tale soluzione è l'identità per  $t = s$  e la sua derivata parziale in  $t$  vale

$$\partial_t \exp\left(\int_s^t A(x) dx\right) = \sum_{k=0}^{+\infty} \frac{\partial_t \left(\int_s^t A(x) dx\right)^k}{k!} = \sum_{k=0}^{+\infty} \frac{1}{k!} \sum_{j=0}^k \left(\int_s^t A(x) dx\right)^j A(t) \left(\int_s^t A(x) dx\right)^{k-j-1}$$

<sup>5</sup>Vedasi 3.5.1.

Poiché le  $A(x)$  commutano per ogni tempo  $x \in I$ , possiamo raccogliere  $A(t)$  e ottenere

$$\partial_t \exp \left( \int_s^t A(x) dx \right) = A(t) \exp \left( \int_s^t A(x) dx \right)$$

Per unicità della soluzione del problema di punto fisso segue la tesi.  $\square$

Ora che ho l'attenzione del lettore, ecco una lista di proprietà più o meno utili di 'sto coso, che potremo tradurre in proprietà dell'esponenziale.

1.  $W(t, r)W(r, s) = W(t, s)$

Infatti, comunque si scelga un vettore  $u_0$ , le funzioni  $u(t) = W(t, s)u_0$  e  $v(t) = W(t, r)W(r, s)u_0$  sono soluzioni del problema di Cauchy omogeneo, come si verifica facendo le derivate parziali in  $t$ , e sono uguali per  $t = r$ . Dunque per unicità devono coincidere per ogni  $t$ , e l'uguaglianza passa alle matrici poiché  $W(t, r)W(r, s)u_0 = W(t, s)u_0$  per ogni  $u_0$ .

Quanto abbiamo mostrato ci dice che le soluzioni del problema omogeneo si "prolungano": se si parte da  $u_0$  al tempo  $s$ , fermarsi al tempo  $t$  è esattamente come fermarsi al tempo  $r$  e poi ripartire dal punto  $W(r, s)u_0$  per fermarsi definitivamente al tempo  $t$ .

2.  $W(t, s)^{-1} = W(s, t)$

Segue dalla proprietà 1.

3.  $\partial_t W(t, s) = A(t)W(t, s)$

4.  $\partial_s W(t, s) = -W(s, t)^{-1}A(s)$

Infatti osserviamo che

$$\partial_s W(t, s) = \partial_s (W(s, t)^{-1})$$

Detta  $inv : GL(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C})$  l'operatore che associa a una matrice la sua inversa, il cui differenziale è  $d inv_A(H) = A^{-1}HA^{-1}$ ,<sup>6</sup> possiamo scrivere, per composizione dei differenziali:

$$\partial_s W(t, s) = \partial_s (W(s, t)^{-1}) = -W(s, t)^{-1}(A(s)W(s, t))W(s, t)^{-1} = -W(s, t)^{-1}A(s)$$

5.  $W_{-A^T} = (W_A^{-1})^T$

Dove  $W_A$  la matrice Wronskiana ottenuta da  $A$ . Ciò è vero perché entrambe risolvono il problema

$$\begin{cases} \partial_t W(t, s) = -A^T(t)W(t, s) \\ W(s, s) = Id \end{cases}$$

Questo perché, derivando rispetto a  $T$  la seconda, si ottiene

$$\partial_t (W_A^{-1}(t, s))^T = \partial_t (W_A(s, t))^T = (-W_A(s, t)A(t))^T = -A^T(t)(W_A^{-1}(t, s))^T$$

Abbiamo usato che la trasposizione è un'operatore lineare, e in quanto tale ha come differenziale se stessa.

6. Dalla proprietà precedente segue che, se  $A(t)$  è antisimmetrica per ogni  $t$ ,  $W_A$  è ortogonale e dunque lo è anche  $e^A$ .

Un'ulteriore proprietà è rappresentata dal seguente:

**Teorema 11.4.4** (Formula di Liouville). *Vale che*

$$\det(W(t, s)) = \exp \left( \int_s^t \text{tr}(A(x)) dx \right)$$

Dove  $\text{tr}(A)$  è la traccia. In particolare,  $W(t, s)$  è invertibile per ogni scelta di  $t$  e  $s$ .

<sup>6</sup>Vedasi 4.6.

*Dimostrazione.* Mostriamo che, fissato  $s$ , il determinante come funzione di  $t$  risolve il problema di Cauchy

$$\begin{cases} \lambda'(t) = \operatorname{tr}(A(t))\lambda(t) \\ \lambda(s) = \det Id = 1 \end{cases}$$

Ciò è equivalente alla tesi perché la soluzione a tale problema esiste ed è unica ed è  $\exp\left(\int_s^t \operatorname{tr}(A(x)) dx\right)$ , per quanto sappiamo già dalla teoria delle equazioni differenziali a valori in  $\mathbb{C}$ . Ora, per calcolare la derivata del determinante osserviamo che

$$\det W(t+h, s) = \det(W(t+h, t)W(t, s)) = \det W(t, s) \det(W(t, t) + h\partial_s W(s, t)|_{s=t} + o(h))$$

$$\det W(t+h, s) = \det W(t, s) \det(W(t, t) + hA(t)W(t, t) + o(h)) = \det W(t, s) \det(Id + hA(t) + o(h))$$

Ora, detta  $B(t) = A(t) + o(h)/h = A(t) + o(1)$ , osserviamo che, detto  $p_B(t)$  il polinomio caratteristico di  $B$ ,

$$\det(Id + hB) = h^n \det\left(B - \frac{(-1)}{h}I\right) = h^n p_B\left(\frac{-1}{h}\right) = 1 + \operatorname{tr}(B)h + o(h)$$

Unendo i pezzi abbiamo che

$$\det W(t+h, s) = \det W(t, s)(1 + \operatorname{tr}(A(t) + o(1))h + o(h)) = \det W(t, s)(1 + \operatorname{tr}(A(t))h + o(h))$$

Da cui  $\det(W(t, s))' = \det W(t, s)\operatorname{tr}(A(t))$ , come voluto.  $\square$

#### 11.4.4 Soluzione dell'equazione generale

**Teorema 11.4.5.** *Il problema di Cauchy*

$$\begin{cases} u'(t) = A(t)u(t) + b(t) \\ u(t_0) = u_0 \end{cases}$$

Ha come unica soluzione

$$u(t) = W(t, t_0)u_0 + \int_{t_0}^t W(t, s)b(s) ds$$

Tale soluzione è definita globalmente sul dominio di  $b$ .

*Dimostrazione.* Scriviamo  $u'(s) = A(s)u(s) + b(s)$  e moltiplichiamo a sinistra per  $W(t, s)$ . Otteniamo

$$W(t, s)u'(s) - W(t, s)A(s)u(s) = \partial_s(W(t, s)u(s)) = W(t, s)b(s)$$

Integrando otteniamo

$$W(t, t)u(t) - W(t, t_0)u_0 = \int_{t_0}^t W(t, s)b(s) ds$$

Da cui la tesi.  $\square$

### 11.5 Equazioni lineari periodiche (teoria di Floquet)

Consideriamo ora il caso specifico in cui la matrice  $A(t)$  che definisce il sistema è periodica di periodo principale  $T$ . Potremo dare delle condizioni di stabilità per le soluzioni del sistema  $u'(t) = A(t)u(t)$ , cosa che in generale non è possibile.

Innanzitutto, consideriamo la matrice wronskiana associata  $W_A(t, t_0)$ . Per semplicità di notazione fissiamo  $t_0 = 0$ , e d'ora in poi indicherò  $W_A(t, 0)$  con  $W(t)$ . In generale non è detto che  $W(t)$ , e con essa la soluzione  $u(t)$ , sia periodica: basti pensare che l'equazione  $y' = y$  è a coefficienti costanti (e quindi periodici) ma la soluzione è un esponenziale. Tuttavia, possiamo osservare che  $W(t+T) = W(t)M_T$  per una certa matrice costante  $M_T$  detta matrice di monodromia. Infatti, entrambe le scritture, se derivate, risolvono l'equazione  $X'(t) = A(t)X(t)$ , perché  $W(t+T)' = A(t+T)W(t+T) = A(t)W(t+T)$ . Dunque

per unicità delle soluzioni le due scritture coincidono se sono uguali in  $t = 0$ . Questo ci porta a dire che la matrice di monodromia è  $M = W(T)$ , e dunque

$$W(t + T) = W(t)M_T = W(t)W(T)$$

Abbiamo usato la notazione  $M_T$  perché la matrice di monodromia dipende dal periodo considerato. Ad esempio,  $A(t)$  è periodica anche di  $kT$ , con  $k \in \mathbb{N}$ , e la matrice di monodromia  $M_{kT}$  si ottiene considerando che

$$W(t + kT) = W(t)M_{kT} = W(t)(W(T))^k = W(t)M_T^k$$

Osserviamo inoltre che la matrice di monodromia è invertibile, perché  $M_T = W(t + T)W(t)^{-1}$  e il wronskiano è sempre invertibile.

**Lemma 11.5.1.** *Il sistema ammette una soluzione periodica di periodo almeno  $T$  se e solo se  $M_T$  ha un autovettore di autovalore 1.*

*Dimostrazione.* La soluzione generale è  $u(t) = W(t)u_0$ , che è periodica se e solo se

$$u(t + T) = W(t + T)u_0 = W(t)M_T u_0 = W(t)u_0 = u(t)$$

□

Analogamente si mostra che

**Lemma 11.5.2.** *Il sistema ammette una soluzione periodica di periodo almeno  $kT$  se e solo se  $M_T^k$  ha un autovettore di autovalore 1, cioè se  $M_T$  ha come autovalore una radice  $k$ -esima dell'unità.*

## 11.6 Stabilità dei sistemi periodici

Innanzitutto, dati due sistemi  $u'(t) = A(t)u(t)$  e  $v'(t) = B(t)v(t)$ , se  $W_B(t) = P(t)W_A(t)$  con un cambio di variabile  $P(t) \in GL(n)$  di norma limitata, allora le soluzioni dei due sistemi hanno lo stesso comportamento all'infinito, cioè lo studio della stabilità delle soluzioni di un sistema è equivalente allo studio della stabilità delle soluzioni dell'altro. Questa considerazione ci fa sperare di ricondurre le equazioni periodiche alle equazioni a coefficienti costanti, di cui sappiamo tutto.

Restringiamoci al caso in cui  $A(t)$  e  $B(t)$  sono  $T$ -periodiche, e cerchiamo un cambio di variabile anch'esso  $T$ -periodico, che esiste se e solo se

$$P(t) = W_B(t)W_A(t)^{-1} = W_B(t + T)W_A(t + T)^{-1} = W_B(t)M_T(B)M_T(A)^{-1}W_A(t)^{-1}$$

Quindi condizione necessaria e sufficiente per la coniugazione dei due sistemi è che abbiano la stessa matrice di monodromia.

Adesso, se  $B$  è costante, la sua matrice di monodromia è  $M_T(B) = \exp(TB)$ . Quindi il problema si riduce a cercare una  $B$  costante tale che  $M_T(A) = \exp(TB)$ , che esiste poiché la matrice di monodromia è invertibile.<sup>7</sup> Riassumendo abbiamo ottenuto il seguente:

**Lemma 11.6.1.** *Il sistema  $u'(t) = A(t)u(t)$  si comporta ai limiti come  $v'(t) = Bv(t)$ , con  $\exp(TB) = M_T$ .*

**Definizione 81.** *Data una  $A(t)$   $T$ -periodica, i suoi moltiplicatori di Floquet sono gli autovalori di  $M_T$*

Osserviamo che, per ogni  $\lambda$  autovalore per  $M_T$ , esiste un autovalore<sup>8</sup>  $\mu$  per  $B$  tale che  $\lambda = e^{T\mu}$ . Possiamo dunque tradurre le condizioni sugli autovalori di  $B$  in condizioni sugli autovalori di  $M_T$ , ricordando che la norma di  $e^{a+bi}$  dipende solo dalla parte reale dell'esponente.

<sup>7</sup>Questa dimostrazione determina una  $B$  a coefficienti complessi. Tuttavia, non è necessario aggiungere postille per il caso reale, tanto le condizioni di stabilità sono le stesse per matrici reali e complesse.

<sup>8</sup>In generale, data una funzione  $f : M(n, \mathbb{C}) \rightarrow M(n, \mathbb{C})$ , per ogni  $\lambda$  autovalore per  $f(B)$ , esiste un  $\mu$  autovalore per  $B$  tale che  $\lambda = f(\mu)$ . Infatti, dato un autovalore  $\mu$  per  $B$ ,  $f(\mu)$  è un autovalore per  $f(B)$ , e in dimensione finita la molteplicità algebrica di  $\mu$  è minore o uguale alla molteplicità dell'immagine. Poiché la somma delle molteplicità algebriche degli autovalori è la dimensione  $n$  dello spazio, non ci possono essere altri autovalori di  $f(B)$ , perché le molteplicità algebriche degli autovalori di  $B$  sommano già a  $n$ . Si osservi che questa dimostrazione è falsa in dimensione infinita, nella quale non ha senso parlare di molteplicità.

**Teorema 11.6.2** (Stabilità dei sistemi periodici). *Sia  $A(t)$   $T$ -periodica. Le soluzioni di  $u'(t) = A(t)u(t)$  sono*

- *infinitesime per  $t \rightarrow +\infty$  se e solo se tutti i moltiplicatori di Floquet hanno modulo minore di 1;*
- *limitate per  $t \rightarrow +\infty$  se e solo se tutti i moltiplicatori di Floquet hanno modulo minore o uguale a 1, e gli autovalori di norma 1 sono semisemplici;*
- *limitate globalmente se e solo se  $M_T$  è diagonalizzabile con tutti gli autovalori di norma 1.*



## Capitolo 12

# Equazioni differenziali non lineari

Non è chiarissimo come definire un'equazione differenziale, dato che vorremmo chiamare con questo nome davvero un sacco di cose. Generalmente è meglio definire che cosa sia una soluzione di un'equazione differenziale, lasciando all'immaginazione del lettore gli ulteriori dettagli.

**Definizione 82.** Dato un aperto  $\Omega \subseteq \mathbb{R}^{n+1}$  e una funzione  $f : \Omega \rightarrow \mathbb{R}^n$ , una curva  $u : \mathbb{R} \rightarrow \mathbb{R}^n$  risolve l'equazione differenziale in forma esplicita  $u'(t) = f(u(t), t)$  se soddisfa tale relazione e il suo grafico è contenuto in  $\Omega$ .

Come nel caso lineare, è rilevante lo studio dei problemi di Cauchy,<sup>1</sup> della forma

$$\begin{cases} u'(t) = f(u(t), t) \\ u(t_0) = u_0 \end{cases}$$

In generale, non è detto che la soluzione esista, sia unica e sia definita globalmente, e si danno facili esempi di tutti e tre i casi.

### 12.1 Teorema di Cauchy-Lipschitz(-Picard-Lindelöf-compagnia bella)

**Definizione 83** (Ipotesi di Lipschitz). Una funzione  $f : \Omega \subseteq \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  si dice localmente lipschitziana nella variabile  $x \in \mathbb{R}^n$  uniformemente nella variabile  $t \in \mathbb{R}$  se per ogni  $(t_0, x_0) \in \Omega$  esiste un  $L(x_0) > 0$  e un intorno rettangolare  $[t_0 - \delta, t_0 + \delta] \times D(x_0, \varepsilon)$  su cui  $f$  è  $L(x_0)$ -lipschitziana nella variabile  $x$ , ossia per ogni  $t \in [t_0 - \delta, t_0 + \delta]$  e per ogni  $x, y \in D(x_0, \varepsilon)$  si ha che

$$\|f(t, x) - f(t, y)\| \leq L(x_0)\|x - y\|$$

Si osservi che questa proprietà è, in un certo senso, uniforme in  $t$ , nel senso che la costante  $L(x_0)$  dipende soltanto da  $x_0$  e non da  $t_0$ .

Osserviamo che una funzione continua e differenziabile nella prima variabile rispetta le ipotesi di Lipschitz, che quindi non sono così assurde né difficili da ottenere. Infatti, una funzione differenziabile su un convesso è lipschitziana, e dunque basta restringere  $f$  a una palla convessa contenuta in  $\Omega$  e alla quale appartiene il punto considerato.

Il seguente lemma ci permette di ampliare l'intorno in cui vale la lipschitzianità a un qualsiasi rettangolo compatto che contiene  $(t_0, x_0)$

**Lemma 12.1.1.** Sia  $f$  localmente lipschitziana in  $x$  su un aperto  $\Omega$ . Allora  $f$  è Lipschitziana nella variabile  $x$  su ogni rettangolo compatto  $R = [t_0 - \delta, t_0 + \delta] \times D(x_0, \varepsilon)$ .

*Dimostrazione.* Ricopriamo  $R$  con le parti interne dei rettangoli  $R(t, x)$  dati dalla locale lipschitzianità. Per compattezza, possiamo estrarre un sottoricoprimento finito. Sia  $L$  la massima costante di Lipschitz associata a un rettangolo del sottoricoprimento.

<sup>1</sup>Mi scuso in anticipo se, ogni due per tre, scambierò l'ordine di  $u$  e  $t$  negli argomenti della  $f$ .

Ora, fissiamo  $t, x, y$  tali che  $(t, x), (t, y) \in R$  e consideriamo la retta che collega i due punti, che per convessità è totalmente contenuta in  $R$ . Tale retta intercetta un numero finito di rettangoli del sottoricoprimento, e ciascuna intersezione della retta con un rettangolo è un segmento per convessità di rettangoli. Dunque è possibile trovare un cammino finito  $x = x_0, x_1, \dots, x_n = y$  di modo che il segmento  $\{(t, sx_k + (1-s)x_{k+1}), s \in [0, 1]\}$  sia totalmente contenuto in un rettangolo. Allora, detta  $L_k$  la costante di Lipschitz relativa al tratto  $k$ -esimo, risulta che

$$\|f(t, x) - f(t, y)\| \leq \sum_{k=0}^{n-1} \|f(t, x_k) - f(t, x_{k+1})\| \leq \sum_{k=0}^{n-1} L_k \|x_k - x_{k+1}\| \leq L \sum_{k=0}^{n-1} \|x_k - x_{k+1}\| = L \|x - y\|$$

□

**Teorema 12.1.2** (Cauchy-Lipschitz). *Sia  $f$  continua e localmente lipschitziana nella prima variabile (uniformemente nella seconda). Allora per ogni  $(x_0, t_0) \in \Omega$  esiste un  $\delta$  tale che il problema di Cauchy*

$$\begin{cases} u'(t) = f(u(t), t) \\ u(t_0) = x_0 \end{cases}$$

ammette una e una sola soluzione locale definita per  $t \in (t_0 - \delta, t_0 + \delta)$ .<sup>2</sup>

*Dimostrazione.* Innanzitutto riscriviamo il problema nella forma integrale equivalente

$$u(t) = x_0 + \int_{t_0}^t f(u(s), s) ds$$

Sia  $R = D(x_0, \varepsilon) \times [t_0 - \delta, t_0 + \delta]$  un qualsiasi intorno rettangolare compatto contenuto in  $\Omega$  e sia  $L$  la costante di Lipschitz ottenuta come nel lemma precedente. A meno di restringere  $\delta$  si può supporre che  $\delta \|f\|_{\infty, R} < \varepsilon$ , dove  $\|\cdot\|_{\infty, R}$  è la norma uniforme su  $R$ .

Ora, consideriamo l'operatore  $\Gamma : C^0([t_0 - \delta, t_0 + \delta], D(x_0, \varepsilon)) \rightarrow C^0([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$  definito come

$$\Gamma(u)(t) = x_0 + \int_{t_0}^t f(u(s), s) ds$$

Il problema si è ridotto a trovare un punto fisso di  $\Gamma$ . Osserviamo che  $\Gamma$  manda  $C^0([t_0 - \delta, t_0 + \delta], D(x_0, \varepsilon))$  in sé, cioè che  $\Gamma(u)$  ha sempre immagine contenuta in  $D(x_0, \varepsilon)$ . Infatti

$$\|\Gamma(u)(t) - x_0\| = \left\| \int_{t_0}^t f(u(s), s) ds \right\| \leq \delta \|f\|_{\infty, R} < \varepsilon$$

Ora,  $C^0([t_0 - \delta, t_0 + \delta], D(x_0, \varepsilon))$  è completo perché i suoi elementi sono le funzioni continue e limitate tra i due spazi considerati. Allora se mostriamo che  $\Gamma^k$  è una contrazione per un certo  $k \in \mathbb{N}$  abbiamo concluso. Questo è vero perché per ogni  $t$  vale la stima

$$\|\Gamma^k(u)(t) - \Gamma^k(v)(t)\| \leq \|u - v\|_{\infty, [t_0 - \delta, t_0 + \delta]} \frac{L^k |t - t_0|^k}{k!}$$

Mostriamolo per induzione su  $k$ . Se  $k = 0$  la tesi è banalmente vera. Ora, se è vera per  $k$  abbiamo che

$$\|\Gamma^{k+1}(u)(t) - \Gamma^{k+1}(v)(t)\| = \left\| \int_{t_0}^t (f(\Gamma^k(u)(s), s) - f(\Gamma^k(v)(s), s)) ds \right\|$$

Per lipschitzianità locale vale che

$$\left\| \int_{t_0}^t (f(\Gamma^k(u)(s), s) - f(\Gamma^k(v)(s), s)) ds \right\| \leq \left\| \int_{t_0}^t L \|\Gamma^k(u)(s) - \Gamma^k(v)(s)\| ds \right\|$$

<sup>2</sup>In teoria, le ipotesi vere di questo teorema sono molto più precise, perché è possibile determinare un  $\delta$  che vada bene per ogni condizione iniziale  $(x, t)$  all'interno di un rettangolo contenuto in  $\Omega$ . Per gli interessati rimando al capitolo 8 di [SNS, 2015].

Per ipotesi induttiva l'integrando si maggiora con

$$\left\| \int_{t_0}^t L^{k+1} \|u - v\|_{\infty, [t_0 - \delta, t_0 + \delta]} \frac{|s - t_0|^k}{k!} ds \right\| = \|u - v\|_{\infty, [t_0 - \delta, t_0 + \delta]} \frac{L^{k+1} |t - t_0|^{k+1}}{(k+1)!}$$

Ora, passando al sup su  $t$  otteniamo che

$$\|\Gamma^k(u) - \Gamma^k(v)\|_{\infty, [t_0 - \delta, t_0 + \delta]} \leq \|u - v\|_{\infty, [t_0 - \delta, t_0 + \delta]} \frac{L^k \delta^k}{k!}$$

Allora esiste un  $k$  sufficientemente grande da rendere  $\Gamma^k$  una contrazione. Segue la tesi.  $\square$

Si osservi che abbiamo dimostrato anche un criterio per valutare l'ampiezza dell'intervallo di tempo su cui è definita la soluzione locale:

**Corollario 12.1.2.1.** *Sia  $f : \Omega \subseteq \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione che soddisfa le ipotesi di Lipschitz. Data una condizione iniziale  $(t_0, x_0)$ , se esistono  $\varepsilon$  e  $\delta$  per cui  $R = D(x_0, \varepsilon) \times [t_0 - \delta, t_0 + \delta] \subseteq \Omega$  e  $\delta \|f\|_{\infty, R} < \varepsilon$ , allora la soluzione  $u(t)$  è definita almeno nell'intervallo  $[t_0 - \delta, t_0 + \delta]$ .*

## 12.2 Equazioni autonome

**Definizione 84.** *Un'equazione differenziale si dice autonoma se  $f$  non dipende da  $t$  ma solo da  $u(t)$ .*

La proprietà rilevante delle equazioni autonome è che sono invarianti per traslazioni temporali. Questa tautologia sta a significare che, se  $u(t)$  è la soluzione che in  $t_0$  vale  $x_0$ , allora per ogni  $s \in \mathbb{R}$   $u(t - s)$  è la soluzione che in  $s$  vale  $x_0$ , ossia la proprietà di composizione del flusso si esprime come

$$\Phi(t, s, x_0) = \Phi(t - s, 0, x_0)$$

In altri termini, un'equazione autonoma permette di riscalarare sempre il tempo in modo da partire da  $t_0 = 0$ .

**Lemma 12.2.1.** *Un'equazione differenziale in forma esplicita si riconduce a un'equazione autonoma.*

*Dimostrazione.* Consideriamo il problema di Cauchy

$$\begin{cases} u'(t) = f(u(t), t) \\ u(t_0) = u_0 \end{cases}$$

Poniamo  $y(t) = (u(t), t)$  e osserviamo che possiamo riscrivere il problema nella seguente forma autonoma:

$$\begin{cases} y'(t) = (f(y(t)), 1) = F(y(t)) \\ y(t_0) = (u_0, t_0) \end{cases}$$

$\square$

Bello questo lemma, eh? Bene, è inutile. Infatti, le ipotesi del teorema di Cauchy-Lipschitz (cioè quelle dell'unico teorema rilevante di questo capitolo) si esprimono bene separando le variabili, mentre un enunciato per le equazioni autonome richiederebbe un loro potenziamento (in pratica, la lipschitzianità in  $y(t) = (u(t), t)$ , cioè in entrambe le variabili). Tuttavia, potrebbe tornarci utile nella dimostrazione del teorema di esistenza di Peano, che richiede solo la continuità di  $f$ .

## 12.3 Esistenza locale per funzioni continue

**Teorema 12.3.1** (Peano). *Sia  $f \in C^0(B, \mathbb{R})$  definita sulla palla chiusa  $B = D(0, r) \subseteq \mathbb{R}^n$ . Il problema di Cauchy*

$$\begin{cases} u' = f(u) \\ u(0) = 0 \end{cases}$$

*ammette soluzione  $u \in C^1(I, \mathbb{R}^n)$ , dove  $I = [-a, a]$  è tale che  $ar = \|f\|_{\infty, D}$ .*

Osserviamo che, per quanto già visto, ci siamo ridotti a studiare un'equazione autonoma.

*Dimostrazione.* Sia  $f_k \in C^0(D, \mathbb{R})$ , con  $k \in \mathbb{N}$  una successione di funzioni lipschitziane convergenti uniformemente a  $f$ . Ad esempio, poiché  $D$  è compatto possiamo applicare le ipotesi del teorema di Stone e approssimare  $f$  con funzioni polinomiali.<sup>3</sup> Inoltre posso supporre che  $\|f_k\|_{\infty, B} \leq \|f\|_{\infty, B}$ . Difatti se ciò non è vero posso sostituire  $f_k$  con la funzione  $\frac{\min\{\|f_k\|_{\infty, B}, \|f\|_{\infty, B}\}}{\|f_k\|_{\infty, B}} f_k$ , che continua a convergere a  $f$  perché il coefficiente tende a 1.

Ora, consideriamo la soluzione del problema di Cauchy approssimato

$$\begin{cases} u'_k = f_k(u_k) \\ u(0) = 0 \end{cases}$$

Tale funzione esiste perché valgono le ipotesi di Lipschitz, ed è definita su  $I = [-a, a]$  in quanto  $a\|f_k\|_{\infty, B} \leq a\|f\|_{\infty, B} \leq r$ .<sup>4</sup>

Ora, le  $u_k$  sono equilipschitziane (in quanto la loro derivata è limitata in modulo da  $\|f\|_{\infty, D}$ ) e a valori in un compatto, dunque il teorema di Ascoli-Arzelà<sup>5</sup> ci garantisce la loro convergenza a una certa  $u \in C^0(I, \mathbb{R})$ , a meno di estrarre una sottosuccessione. Mostriamo che  $f_k(u_k) = u'_k$  converge uniformemente a  $f(u)$ . Ricordando che  $f$  è uniformemente continua, essendo definita su un compatto, abbiamo

$$\|f_k(u_k) - f(u)\|_{\infty, I} \leq \|f_k(u_k) - f(u_k)\|_{\infty, I} + \|f(u_k) - f(u)\|_{\infty, I} \leq \|f_k - f\|_{\infty, B} + \omega_f(\|u_k - u\|_{\infty, I}) = o(1)$$

Ora, poiché  $u'_k \rightarrow f(u)$  e  $u_k(0) \rightarrow u(0)$ , il teorema di limite sotto il segno della derivata<sup>6</sup> ci garantisce che il limite  $u$  è derivabile con derivata continua  $f(u)$ , da cui la tesi. □

## 12.4 Estensione di soluzioni

Per i teoremi seguenti richiederemo solo l'esistenza della soluzione garantita dal teorema di Peano, senza aggiungere le ipotesi di unicità.

**Teorema 12.4.1.** *Sia  $f : \Omega \rightarrow \mathbb{R}^n$  continua. Sia  $u : I = ]a, b[ \rightarrow \mathbb{R}^n$  una soluzione dell'equazione  $u'(t) = f(t, u(t))$ . Se esiste il limite  $\lim_{t \rightarrow b} u(t) = \bar{x}$  e  $(b, \bar{x}) \in \Omega$ , allora  $u$  si estende a una soluzione definita su un intervallo della forma  $]a, b + \delta[$ .*

*Dimostrazione.* Sia  $v$  una soluzione locale del problema

$$\begin{cases} v' = f(t, v(t)) \\ v(b) = \bar{x} \end{cases}$$

Sia  $]b - \varepsilon, b + \delta[$  il dominio di definizione di  $v$ . Allora la funzione

$$\bar{u}(t) = \begin{cases} u(t) & \text{se } a < t < b \\ v(t) & \text{se } b \leq t < b + \delta \end{cases}$$

è soluzione dell'equazione differenziale. Infatti, se  $t \neq b$ , la funzione coincide con una soluzione. In  $t = b$   $\bar{u}$  è continua, poiché esiste  $\lim_{t \rightarrow b} u(t) = \bar{x} = v(b)$ , ed è derivabile in quanto esistono i limiti da destra e da sinistra della derivata. Infatti per continuità di  $f$  abbiamo che

$$\bar{u}'(b^-) = \lim_{t \rightarrow b^-} u'(t) = \lim_{t \rightarrow b^-} f(t, u(t)) = f(b, \bar{x})$$

E d'altra parte, allo stesso modo,

$$\bar{u}'(b^+) = \lim_{t \rightarrow b^+} v'(t) = \lim_{t \rightarrow b^+} f(t, v(t)) = f(b, \bar{x})$$

□

<sup>3</sup>Vedasi 5.3.3.

<sup>4</sup>Infatti la condizione sull'ampiezza dell'intorno di esistenza locale è esattamente questa, come abbiamo visto in 12.1.2.

<sup>5</sup>Vedasi 5.4.1.

<sup>6</sup>Vedasi 6.12.1.

Il teorema precedente si può indebolire, richiedendo solo che esista una successione  $t_n \rightarrow b$  le cui immagini mediante  $u$  convergono a un certo valore.

**Teorema 12.4.2.** *Sia  $f : \Omega \rightarrow \mathbb{R}^n$  continua. Sia  $u : ]a, b[ \rightarrow \mathbb{R}^n$  una soluzione dell'equazione  $u'(t) = f(t, u(t))$ . Se esiste una successione  $t_n \rightarrow b$  le cui immagini ammettono limite  $\lim_{n \rightarrow +\infty} u(t_n) = \bar{x}$  e  $(b, \bar{x}) \in \Omega$ , allora  $u$  si estende a una soluzione definita su un intervallo della forma  $]a, b + \delta[$ .*

*Dimostrazione.* Se mostriamo che  $u$  ammette limite per  $t \rightarrow b$  possiamo applicare il risultato precedente. Supponiamo allora per assurdo che  $\lim_{t \rightarrow b} u(t) \neq \bar{x}$ . Allora esistono un  $\varepsilon$  e una successione  $s_k \rightarrow b$  tale che  $\|u(s_k) - \bar{x}\| \geq \varepsilon$  per ogni  $k$  naturale. A meno di restringere  $\varepsilon$ , supponiamo che  $[b - \varepsilon, b] \times D(\bar{x}, \varepsilon) \subseteq \Omega$  e consideriamo la sottosuccessione dei  $t_n > b - \varepsilon$ .

Ora, per ogni  $n$  naturale sia  $t_n^* = \min\{t \in [t_n, b[ \text{ t.c. } \|u(t) - \bar{x}\| \geq \varepsilon\}$ . Tale  $t_n^*$  è ben definito perché l'insieme di cui è minimo è non vuoto, dovendo contenere definitivamente la successione  $s_k$ . Inoltre  $t_n \leq t_n^* < b$  e  $t_n \rightarrow b$ , dunque anche  $t_n^* \rightarrow b$ . Abbiamo allora due successioni convergenti a  $b$  tali che

$$\begin{cases} \|u(t_n) - \bar{x}\| \rightarrow 0 \\ \|u(t_n^*) - \bar{x}\| > \varepsilon \end{cases}$$

Dunque esiste un  $N$  tale che, per ogni  $n > N$   $t_n \neq t_n^*$ , poiché prima o poi la successione  $u(t_n)$  sta a meno di  $\varepsilon$  dal suo limite. Inoltre, per come sono definiti i  $t_n^*$ , se  $t_n \leq s \leq t_n^*$  allora  $\|u(s) - \bar{x}\| \leq \varepsilon$ .

Ora, restringendoci alla sottosuccessione con  $n > N$ , per ipotesi

$$\left\| \frac{u(t_n^*) - u(t_n)}{t_n^* - t_n} \right\| \rightarrow +\infty$$

Tuttavia, per teorema del valor medio per curve, otteniamo

$$\left\| \frac{u(t_n^*) - u(t_n)}{t_n^* - t_n} \right\| \leq \sup_{t_n^* < s < t_n} \|u'(s)\| = \sup_{t_n^* < s < t_n} \|f(s, u(s))\| \leq \max_{(s,x) \in [b-\varepsilon, b] \times D(\bar{x}, \varepsilon)} \|f(s, x)\| < +\infty$$

Dove l'ultima maggiorazione deriva dal fatto che  $\|f\|$  è continua su un compatto. Ciò porta a un evidente assurdo.  $\square$

Ora, notiamo che le ipotesi del teorema precedente sono banalmente verificate se  $u$  è limitata, o meglio se esiste un compatto che contiene il grafico di  $u$ , perché è sempre possibile estrarre una successione  $t, u(t)$  convergente a qualcosa. Dunque se una soluzione ha grafico in un compatto la si può estendere. In altri termini, vale il seguente teorema:

**Teorema 12.4.3** (Fuga dai compatti). *Sia  $f : \Omega \rightarrow \mathbb{R}^n$  continua e sia  $u : ]a, b[ \rightarrow \mathbb{R}^n$  una soluzione massimale dell'equazione  $u(t) = f(t, u(t))$ . Per ogni  $K \subseteq \Omega$  compatto esistono  $a' < b' \in ]a, b[$  tali che  $(t, u(t)) \notin K$  per ogni  $t \in ]a, a'] \cup [b', b[$ .*

*Dimostrazione.* Per assurdo, supponiamo che esista un compatto  $K$  cui frequentemente appartiene il grafico di  $u$ . In altri termini, supponiamo che  $a$  e  $b$  siano finiti, e senza perdere generalità che esista  $t_n \rightarrow b$  tale che  $(t_n, u(t_n)) \in K$ . Da quest'ultima si può estrarre una sottosuccessione convergente a  $(b, \bar{x})$  per compattezza di  $K$ , e dunque si può estendere la soluzione a un intervallo  $]a, b + \delta[$  strettamente più grande di  $]a, b[$ . Ciò viola l'ipotesi di massimalità della soluzione.  $\square$

Se le ipotesi su  $f$  garantiscono l'esistenza e unicità locali (ad esempio se  $f$  è localmente lipschitziana nella seconda variabile) per ogni  $(t_0, u_0) \in \Omega$  è possibile definire l'intervallo "massimo" di esistenza della soluzione al problema di Cauchy con tale dato iniziale. Infatti possiamo considerare l'insieme  $\Gamma$  delle coppie  $(u, I)$ , dove  $u$  è una soluzione del problema di Cauchy definita sull'intervallo aperto  $I$  contenente  $t_0$ . Allora definiamo una soluzione sull'unione dei domini ponendo  $U(t) = u_\lambda(t)$  se  $t \in \lambda$ .<sup>7</sup> Tale soluzione è ben definita, perché in ogni punto vale l'unicità locale, e definisce l'unica soluzione massimale del problema.

<sup>7</sup>Per chi se lo chiedesse, non serve l'assioma della scelta per definire tale funzione. Infatti, basta dire che, per ogni  $t \in \bigcup I_\lambda$ , esiste un  $I_\lambda$  che lo contiene, per definizione di unione, e definire  $U(t) = u(t_\lambda)$ .

In questi termini, sia  $] \tau_*(t_0, x_0), \tau^*(t_0, x_0)[$  l'intervallo massimale della soluzione al problema con condizione iniziale  $u(t_0) = x_0$ . Sia poi

$$\Sigma = \{(t, t_0, x_0) \in \mathbb{R} \times \Omega \text{ t.c. } t \in ] \tau_*(t_0, x_0), \tau^*(t_0, x_0)[\}$$

Allora si può definire il flusso di un'equazione differenziale generica come la funzione

$$\begin{aligned} \Phi_f : \Sigma &\longrightarrow \mathbb{R}^n \\ (t, t_0, u_0) &\longrightarrow u(t) \end{aligned}$$

Ove  $u(t)$  è la soluzione massima del problema di Cauchy con condizioni iniziali  $u(t_0) = u_0$ . D'ora in poi, indicheremo con  $\Phi_f(t_0, u_0)$  la soluzione del problema con condizione iniziale  $u(t_0) = u_0$ , definita sull'intervallo massimale. Si osservi che, per unicità della soluzione, il flusso ha la seguente proprietà di "prolungamento":

$$\Phi(t, r, x) = \Phi(t, s, \Phi(s, r, x))$$

In altri termini, il risultato che si ottiene partendo da  $x$  al tempo  $r$  equivale a quello ottenuto "facendo tappa" al tempo  $s$  in  $\Phi(s, r, x)$ .

## 12.5 Criterio di esistenza globale

**Teorema 12.5.1.** *Sia  $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  che soddisfa le ipotesi di Lipschitz. Se  $f$  è sublineare in  $x$ , cioè esistono una norma<sup>8</sup>  $\|\cdot\|$  e due funzioni reali  $a, b \in C^0(I, \mathbb{R})$  tali che*

$$\|f(t, x)\| \leq a(t)\|x\| + b(t)$$

Allora tutte le soluzioni di  $u' = f(t, u)$  sono definite globalmente su  $I$ .

*Dimostrazione.* Fissato un  $\varepsilon > 0$  e un intervallo compatto  $J = [\alpha, \beta] \subseteq I = (\alpha_0, \beta_0)$ , per quanto osservato nel corollario al teorema di Cauchy-Lipschitz<sup>9</sup> ci basta mostrare l'esistenza di un  $\delta$  uniforme tale che per ogni condizione iniziale  $(t_0, x_0) \in J \times \mathbb{R}^n$  esista un  $\varepsilon > 0$  per cui  $R = [t_0 - \delta, t_0 + \delta] \times D(x_0, \varepsilon) \subseteq I \times \mathbb{R}^n$  e inoltre  $\delta \|f\|_{\infty, R} < \varepsilon$ . Infatti, se ogni soluzione locale esiste per un tempo  $\delta$  ed è ivi limitata è sempre possibile costruire una soluzione il cui intervallo massimale contenga  $J$ , unendo le soluzioni parziali (cosa che si può fare perché queste ultime non possono divergere essendo limitate).

La prima condizione è banalmente verificata per ogni  $\delta$  per cui  $[t_0 - \delta, t_0 + \delta] \subseteq I$ , dunque basta imporre che  $\delta \leq \frac{1}{2} \min\{\alpha - \alpha_0, \beta_0 - \beta\}$ . Per la seconda, per teorema di Weierstrass applicato all'intervallo  $[\alpha - \delta, \beta + \delta]$  esistono due costanti  $A, B$  positive e valide per ogni  $(t_0, x_0) \in J \times \mathbb{R}^n$  per cui

$$\|f(t, x)\| \leq A\|x\| + B$$

Basta allora imporre che, fissato  $x_0 \in \mathbb{R}^n$ , esista un  $\varepsilon > 0$  per cui  $\delta \|f\|_{\infty, R} \leq \delta(A(\|x_0\| + \varepsilon) + B) < \varepsilon$ . Ora, la funzione  $\frac{\varepsilon}{A(\|x_0\| + \varepsilon) + B}$  tende a  $1/A$  per  $\varepsilon \rightarrow +\infty$ , dunque definitivamente è maggiore di  $1/2A$ . Perciò si può scegliere

$$\delta \leq \frac{1}{2} \min \left\{ \alpha - \alpha_0, \beta_0 - \beta, \frac{1}{A} \right\}$$

□

## 12.6 Dipendenza continua dai dati iniziali

L'obiettivo che ci prefiggiamo ora è mostrare che, sotto le ipotesi di Lipschitz, l'insieme  $\Sigma$  è aperto e il flusso è continuo nei dati iniziali. Iniziamo con un lemma che sembra totalmente a caso.

<sup>8</sup>In realtà si può scegliere una qualsiasi norma, per l'equivalenza delle norme sugli spazi vettoriali di dimensione finita. Infatti basta mostrare il risultato per una norma specifica e poi, scelta una seconda norma, inglobare nella definizione di  $A$  la costante di bilipschitzianità che la lega alla prima.

<sup>9</sup>vedasi 12.1.2.1.

**Lemma 12.6.1.** *Sia  $u = \Phi_f(t_0, x_0)$  con  $f$  che soddisfa le ipotesi di Lipschitz. Dato un intervallo  $I = [t_0, t_1] \subseteq \text{dom}(u)$ , esistono due costanti  $\rho > 0$  e  $C > 0$  tali che, per ogni  $y_0 \in B(x_0, \rho)$  la soluzione  $v = \Phi_f(t_0, y_0)$  è definita almeno fino a  $t_1$  e vale che  $\|u - v\|_{\infty, I} \leq C\|x_0 - y_0\|$ . Simili risultati si danno considerando l'intervallo  $J = [t_2, t_0]$ .*

In soldoni, il lemma ci sta dicendo che per ogni coppia di tempi  $(t_0, t_1)$  l'insieme

$$\Sigma_{t_1, t_0} = \{x \in \mathbb{R}^n \text{ t.c. } (t_1, t_0, x) \in \Sigma\}$$

è aperto, perché contiene la palla  $B(x, \rho)$ . Inoltre, la funzione

$$\begin{array}{ccc} \Sigma_{t_1, t_0} & \longrightarrow & \mathbb{R}^n \\ x & \longrightarrow & \Phi(t_1, t_0, x) \end{array}$$

è localmente lipschitziana di costante  $C$ , e dunque il flusso è continuo se si fissano  $t_1$  e  $t_0$ . Ciò significa che, se si fissa il tempo iniziale e il tempo di osservazione, piccole variazioni della condizione iniziale generano piccole variazioni del risultato.

*Dimostrazione.* Fissato  $\varepsilon > 0$ , consideriamo l'insieme

$$T_\varepsilon = \{(t, x) \in I \times \mathbb{R}^n \text{ t.c. } \|u(t) - x\| \leq \varepsilon\}$$

Questo insieme è un intorno tubolare compatto del grafico di  $u$  tra i punti  $t_0$  e  $t_1$ . Possiamo scegliere  $\varepsilon$  in modo che  $T_\varepsilon \subseteq \Omega$ . Infatti, per ogni  $t \in I$  esiste una palla  $B((t, u(t)), \varepsilon(t)) \subseteq \Omega$ . Queste palle ricoprono  $u(I)$  e dunque per compattezza possiamo estrarne un numero finito e prendere il minimo dei raggi come  $\varepsilon$ .

Ora,  $f$  è lipschitziana nella prima variabile sul compatto  $T_\varepsilon$ , come osservato in 1.4.5. Detta  $L$  la costante di Lipschitz, mostriamo che le costanti cercate possono essere  $\rho = \frac{1}{2}\varepsilon e^{-L(t_1-t_0)}$  e  $C = e^{L(t_1-t_0)}$ . Sia infatti  $y_0$  tale che  $(t_0, y_0) \in \Omega$  e sia  $v = \Phi_f(t_0, y_0)$ . Per la proprietà di fuga dai compatti esiste un tempo  $t > t_0$  tale che  $(t, v(t)) \notin T_\varepsilon$ . Consideriamo adesso  $t_* = \inf\{t > t_0 \text{ t.c. } t \in \text{dom}(v), (t, v(t)) \notin T_\varepsilon\}$ . Per definizione l'intervallo  $[t_0, t_*] \subseteq \text{dom}(v)$  ha immagine contenuta in  $T_\varepsilon$ , e per continuità e poiché  $T_\varepsilon$  è chiuso anche  $t_*$  ha immagine contenuta in  $T_\varepsilon$ . Dunque, per ogni  $t \in [t_0, t_*]$  vale che

$$u(t) - v(t) = u(t_0) - v(t_0) + \int_{t_0}^t (f(s, u(s)) - f(s, v(s))) ds$$

Passando alle norme, abbiamo che la differenza  $\|u(t) - v(t)\|$ , che intendiamo stimare uniformemente con  $C\|x_0 - y_0\|$ , è maggiorata nel modo seguente:

$$\|u(t) - v(t)\| \leq \|x_0 - y_0\| + \int_{t_0}^t \|f(s, u(s)) - f(s, v(s))\| ds \leq \|x_0 - y_0\| + L \int_{t_0}^t \|u(s) - v(s)\| ds$$

Abbiamo usato che  $f$  è lipschitziana su  $T_\varepsilon$ . Indichiamo ora con  $\phi(t) = \int_{t_0}^t \|u(s) - v(s)\| ds$ . La disuguaglianza di sopra si può scrivere come una "disequazione differenziale"

$$\phi'(t) \leq \|x_0 - y_0\| + L\phi(t)$$

Da cui, moltiplicando per  $e^{-Lt} > 0$ , otteniamo

$$(\phi'(t) - \|x_0 - y_0\| - L\phi(t))e^{-Lt} = \left( \left( \frac{\|x_0 - y_0\|}{L} + \phi(t) \right) e^{-Lt} \right)' \leq 0$$

Dunque la funzione tra parentesi è debolmente decrescente in  $t$ . Dato che  $\phi(t_0) = \int_{t_0}^{t_0} \|u(s) - v(s)\| ds = 0$  abbiamo che

$$\left( \frac{\|x_0 - y_0\|}{L} + \phi(t) \right) e^{-Lt} \leq \frac{\|x_0 - y_0\|}{L} e^{-Lt_0}$$

Riprendendo la disequazione iniziale abbiamo che

$$\|u(t) - v(t)\| = \phi'(t) \leq \|x_0 - y_0\| + L\phi(t) \leq e^{L(t-t_0)} \|x_0 - y_0\|$$

Da cui, passando al sup su  $t \in [t_0, (\min(t_*, t_1))]$ ,  $C = e^{L(\min(t_*, t_1) - t_0)}$ .

Ci serve ora mostrare che, sotto certe condizioni su  $\|x_0 - y_0\|$ , rappresentate dal bound superiore tramite  $\rho$ , si ha  $t^* \geq t_1$ , in modo che la soluzione  $v$  sia definita almeno fino a  $t_1$ . Supponiamo allora che  $\|x_0 - y_0\| \leq \rho = \frac{1}{2}\varepsilon e^{-L(\min(t_*, t_1) - t_0)}$ , da cui per ogni  $t \in [t_0, t_*]$

$$\|u(t) - v(t)\| \leq e^{L(\min(t_*, t_1) - t_0)} \|x_0 - y_0\| \leq \frac{1}{2}\varepsilon$$

In particolare, se per assurdo  $t_* < t_1$ , avrei che  $\|u(t_*) - v(t_*)\| \leq \frac{1}{2}\varepsilon$ . Questo significa che  $(t_*, v(t_*))$  sta nella parte interna di  $T_\varepsilon$ , perché  $t_0 < t_* < t_1$  e  $v(t_*) \in B(u(t_*), \varepsilon)$ . Questo viola l'ipotesi che  $t_*$  sia l'estremo inferiore dei valori per cui  $(t, v(t)) \notin T_\varepsilon$ .

Dunque riassumendo possiamo scegliere  $C = e^{L(t_1 - t_0)}$  e  $\rho = \frac{1}{2}\varepsilon e^{-L(t_1 - t_0)}$ .  $\square$

Da sola questa dimostrazione non conclude molto, tuttavia credo che Majer si sia dimenticato di svolgere la seconda parte. Ammucchiamo (anche) questa polvere sotto al tappeto (tanto, ormai, fa qualche differenza?) e procediamo.

**Teorema 12.6.2** (Differenziabilità del flusso). *Sia  $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$  aperto e sia  $f : \Omega \rightarrow \mathbb{R}^n$  continua e con differenziale parziale in  $x$  continuo. Allora il flusso  $\Phi : \Sigma \rightarrow \mathbb{R}^n$  è di classe  $C^1$*

*Dimostrazione.* Se mostriamo che i tre differenziali parziali esistono continui in ogni punto  $(t, s, x) \in \Sigma$  possiamo concludere per il teorema del differenziale totale. Trattiamo i tre casi separatamente. Nel seguito, per snellire la notazione, ometterò i pedici che indicano i punti di applicazione dei differenziali, limitandomi a indicare la variabile rispetto alla quale si differenzia.

1. Per come è definita l'equazione differenziale si ha che  $\partial_t \Phi(t, s, x) = f(t, \Phi(t, s, x))$ , che è continua per ipotesi.
2. Occupiamoci della differenziabilità rispetto a  $x$ . Per capire chi possa essere il differenziale parziale  $d^x$  osserviamo che, se esiste ed è continuo, deve soddisfare la seguente proprietà:

$$d^x f(t, \Phi(t, s, x)) = d^2 f(t, \Phi(t, s, x)) \circ d^x \Phi(t, s, x)$$

Inoltre, il primo termine si riscrive come

$$d^x f(t, \Phi(t, s, x)) = d^x \partial_t \Phi(t, s, x) = \partial_t d^x \Phi(t, s, x)$$

L'ultima uguaglianza è un'applicazione del teorema di Schwarz, in quanto il differenziale misto  $d^x \partial_t$  esiste ed è continuo (in quanto è dato da  $d^2 f(t, t, \Phi(t, s, x)) \circ d^x \Phi(t, s, x)$ ). Mettendo insieme le due uguaglianze possiamo concludere che, se esiste, il differenziale in  $x$  deve risolvere l'equazione

$$u'(t) = \partial_t d^x \Phi(t, s, x) = d^2 f(t, \Phi(t, s, x)) \circ d^x \Phi(t, s, x) = A(t)u(t)$$

con condizione iniziale  $u(s) = Id$  (in quanto  $\Phi(s, s, x) = x$ ). Allora il differenziale dev'essere  $W_A(t, s)$ , con  $A(t)$  come sopra. Si osservi che la forma trovata garantisce anche la continuità in tutte le variabili, per dipendenza continua di  $W$  da  $A(t) = d^2 f(t, \Phi(t, s, x))$ ,  $t$  ed  $s$ . Mostriamo che effettivamente il differenziale è quello cercato, cioè che, fissati  $s$  e  $x$  e posto

$$z(t) = \Phi(t, s, y) - \Phi(t, s, x) - W_A(t, s)[y - x]$$

la funzione  $z(t)$  è  $o(\|x - y\|)$  se si tiene fisso  $t$  e si manda  $y \rightarrow x$ . Osserviamo preliminarmente che  $z(s) = 0$  per ovvi motivi. Derivando rispetto a  $t$  otteniamo

$$z'(t) = f(t, \Phi(t, s, y)) - f(t, \Phi(t, s, x)) - A(t)W_A(t, s)[y - x]$$

mentre componendo a sinistra per  $A(t) = d^2 f(t, \Phi(t, s, x))$  si ha

$$A(t)z(t) = d^2 f(t, \Phi(t, s, x))[\Phi(t, s, y) - \Phi(t, s, x)] - A(t)W_A(t, s)[y - x]$$

Sottraendo le due equazioni membro a membro otteniamo la seguente equazione differenziale per  $z$ :

$$\begin{cases} z'(t) - A(t)z(t) = f(t, \Phi(t, s, y)) - f(t, \Phi(t, s, x)) - d^2 f(t, \Phi(t, s, x))[\Phi(t, s, y) - \Phi(t, s, x)] = h(t) \\ z(s) = 0 \end{cases}$$

Ora, poniamo  $\Phi(t, s, y) = v_1$  e  $\Phi(t, s, x) = v_0$ . Per teorema del valor medio applicato alla funzione  $f(t, \lambda v_1 + (1 - \lambda)v_0) - \lambda d^2 f(t, v_0)$ ,  $h$  si stima in norma come

$$\|h(t)\| \leq \|f(t, v_1) - f(t, v_0) - d^2 f(t, v_0)[v_1 - v_0]\| \leq \sup_{0 < \lambda < 1} \|d^2 f(t, \lambda v_1 + (1 - \lambda)v_0) - d^2 f(t, v_0)\| \|v_1 - v_0\|$$

Ora, la norma infinito di  $h$  su  $[t, s]$  si stima come

$$\sup_{\tau \in [t, s]} \|\tau\| \leq \sup_{0 < \lambda < 1, \tau \in [t, s]} \|d^2 f(\tau, \lambda \Phi(\tau, s, y) + (1 - \lambda)\Phi(\tau, s, x)) - d^2 f(\tau, \Phi(\tau, s, x))\| \|\Phi(\tau, s, y) - \Phi(\tau, s, x)\|$$

Quando  $y \rightarrow x$ ,  $d^2 f(\tau, \lambda \Phi(\tau, s, y) + (1 - \lambda)\Phi(\tau, s, x))$  converge a  $d^2 f(\tau, \Phi(\tau, s, x))$  uniformemente in  $\tau$ . Questo perché la funzione che a  $\tau$  associa la distanza  $\lambda \Phi(\tau, s, y) + (1 - \lambda)\Phi(\tau, s, x)$  è continua e dunque sul compatto  $[t, s]$  ammette un minimo. Tale minimo è strettamente decrescente al tendere di  $y$  a  $x$ , perciò la convergenza è uniforme.

Con tutto 'sto casino abbiamo mostrato (spero)<sup>10</sup> che  $\|h\|_{\infty, [t, s]} = o(\|y - x\|)$ . Allora, dato che  $z(t) = \int_s^t W_a(t, \tau)h(\tau)d\tau$ , la norma di  $z$  si stima come

$$\|z\|_{\infty, [t, s]} \leq |t - s| \|W_A\|_{\infty, [t, s]^2} \|h\|_{\infty, [t, s]} = o(\|y - x\|)$$

Che è esattamente quello che volevamo.

3. Consideriamo infine il differenziale nel tempo iniziale  $s$ . Osserviamo che  $\Phi(s, t, \Phi(t, s, x)) = x$ , in quanto andare avanti e tornare indietro nel tempo in modo reversibile è concesso nelle EDO<sup>11</sup>. Poniamo  $u(s) = \Phi(t, s, x)$  e, a  $t$  fissato, sia  $G(s, u) = \Phi(s, t, u) - x$ , che si annulla per  $u = u(s)$ . Ora, il luogo di zeri di  $G$  soddisfa localmente le ipotesi del teorema della funzione implicita, che ci permettono di esplicitare  $u$  in funzione di  $s$ . Infatti, il differenziale in  $u$  è sempre invertibile, perché

$$d_u G(s, u) = d_u \Phi(s, t, u) = W_{A(t)}(t, s) \in GL$$

dove  $A$  è la matrice definita nel punto precedente. Inoltre, la funzione  $\Phi(s, t, u)$  è di classe  $C^1$ , perché per quanto mostrato finora  $\Phi$  ammette differenziali parziali continui nella prima e nella terza entrata (che sono le uniche variabili nel nostro caso). Dunque il teorema di invertibilità locale ci garantisce che  $u(s) = \Phi(t, s, x)$  è di classe  $C^1$ .

□

<sup>10</sup>Ho la vaga impressione che questa cosa non sia stata davvero mostrata, perciò vi invito a scrivermi se sapete come si faccia.

<sup>11</sup>E Martin McFly muto.



## Capitolo 13

# Metodi analitici in topologia

Questo capitolo raccoglie un paio di risultati offtopic che mi sembrava uno spreco gettare nel dimenticatoio del capitolo precedente.

### 13.1 Teorema di non retrazione e teorema di Brouwer

**Teorema 13.1.1** (Non retrazione). *Non esiste una retrazione  $r : B_n(0, 1) \rightarrow \mathbb{S}^{n-1} = \partial B_n(0, 1)$ , ossia una funzione continua che su  $\mathbb{S}^{n-1}$  sia l'identità.*

Occupiamoci dapprima del caso in cui  $f$  è di classe  $C^\infty$ , in quanto la trattazione generale si potrà vedere come una conseguenza del teorema di Brouwer.

*Dimostrazione.* Per assurdo esista una retrazione  $r \in C^\infty$ . La retrazione è per deformazione, perché possiamo costruire un'omotopia  $g$  tra la retrazione e  $id_{B_n}$  nel seguente modo:

$$g(x, t) = tr(x) + (1 - t)x$$

Consideriamo inoltre la funzione reale

$$\phi(t) = \int_{B_n} \det d_x(g(\cdot, t)) dx$$

Osserviamo che  $\phi$  assume due valori distinti agli estremi. Infatti  $g(\cdot, 0) = id_{B_n}$  ha differenziale costantemente pari a 1, e dunque  $\phi(0) = \int_{B_n} dx = \lambda(B_n)$ . D'altra parte  $g(\cdot, 1) = r$  ha valori in  $\mathbb{S}^{n-1}$  che è una  $(n-1)$ -varietà, e dunque il suo differenziale deve avere determinante nullo perché altrimenti per teorema di invertibilità locale avremmo un diffeomorfismo tra un aperto di  $\mathbb{R}^{n-1}$  e un aperto di  $\mathbb{R}^n$ . Ciò tuttavia è impossibile, perché l'immagine di tale diffeomorfismo dovrebbe essere composta solo da valori critici, e dunque per Sard dovrebbe avere misura nulla. Da questo detour ricaviamo che  $\phi(1) = 0$ .

Vogliamo dunque ottenere un assurdo mostrando che  $\phi$  dovrebbe essere costante. Per DUTIS la sua derivata vale

$$\phi'(t) = \int_{B_n} \partial_t \det d_x(g(t, \cdot)) dx$$

Ora, per la formula di Leibnitz <sup>1</sup>

$$\det d_x g(\cdot, t) = \det[\partial_1 g(x, t), \dots, \partial_n g(x, t)] = \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) \prod_{i=1}^n \partial_i g(x, t)_{\sigma(i)}$$

---

<sup>1</sup>Si osservi che, data una qualsiasi forma multilineare  $\psi$ , lo stesso ragionamento permette di dire che

$$\psi'(a_1, \dots, a_n) = \sum_{j=1}^n \psi(a_1, \dots, a'_j, \dots, a_n)$$

Allora derivando e ricomponendo i determinanti otteniamo

$$\partial_t \det d_x(g(t, \cdot)) = \sum_{j=1}^n \det[\partial_1 g(x, t), \dots, \partial_t \partial_j g(x, t), \dots, \partial_n g(x, t)] = \sum_{j=1}^n \det[\partial_1 g(x, t), \dots, \partial_j \partial_t g(x, t), \dots, \partial_n g(x, t)]$$

Ora,

$$\begin{aligned} & \det[\partial_1 g(x, t), \dots, \partial_j \partial_t g(x, t), \dots, \partial_n g(x, t)] = \\ & = \partial_j \det[\partial_1 g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)] - \sum_{i \neq j} \det[\partial_1 g(x, t), \dots, \partial_j \partial_i g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)] \end{aligned}$$

Dunque  $\partial_t \det d_x(g(t, \cdot))$  è dato da

$$\sum_{j=1}^n \partial_j \det[\partial_1 g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)] - \sum_{j=1}^n \sum_{i \neq j} \det[\partial_1 g(x, t), \dots, \partial_j \partial_i g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)]$$

Il secondo pezzo è nullo, perché gli addendi  $i, j$  e  $j, i$  sono opposti dato che il determinante cambia di segno se si scambiano due colonne. Allora, sostituendo nell'integrale abbiamo

$$\phi'(t) = \sum_{j=1}^n \int_{B_n} \partial_j \det[\partial_1 g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)] dx$$

Poniamo  $x = x_j e_j + y \in \mathbb{R} \times \mathbb{R}^{n-1}$ . Allora possiamo applicare Fubini-Tonelli, dato che le funzioni sono tutte continue e dunque integrabili sul compatto  $B_n$ . Otteniamo dunque

$$\begin{aligned} \phi'(t) &= \sum_{j=1}^n \int_{B_{n-1}} \left( \int_{-\sqrt{1-\|y\|^2}}^{\sqrt{1-\|y\|^2}} \partial_j \det[\partial_1 g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)] dx \right) dy = \\ &= \sum_{j=1}^n \int_{B_{n-1}} (\det[\partial_1 g(x, t), \dots, \partial_t g(x, t), \dots, \partial_n g(x, t)])_{x_j=\sqrt{1-\|y\|^2}}^{x_j=-\sqrt{1-\|y\|^2}} dy \end{aligned}$$

Ora, il membro tra parentesi è nullo. Infatti  $x = (y_1, \dots, y_{j-1}, \pm\sqrt{1-\|y\|^2}, y_{j+1}, \dots, y_n) \in \mathbb{S}^{n-1}$ , e dunque  $g(x, t) = \text{cost} = x$  in quanto l'omotopia è costante sul bordo. Ma allora la colonna  $j$ -esima del differenziale è nulla, e perciò lo è il determinante. Dunque  $\phi$  è costante, da cui l'assurdo.  $\square$

**Teorema 13.1.2** (Brouwer). *Ogni funzione continua  $f : B_n(0, 1) \rightarrow B_n(0, 1)$  ha un punto fisso.*

*Dimostrazione.* Possiamo supporre che  $f$  sia di classe  $C^\infty$ . Infatti, se esistesse una funzione  $f \in C^0$  priva di punti fissi, potremmo considerare una mappa liscia<sup>2</sup>  $g \in C^\infty$  tale che  $\|f - g\|_\infty < \frac{1}{2} \min_{x \in B_n} \|x - f(x)\|$ . Allora anche  $g$  non avrebbe punti fissi, dato che per ogni  $x \in B_n$

$$\|x - g(x)\| \geq \|x - f(x)\| - \|g(x) - f(x)\| \geq \frac{1}{2} \min_{x \in B_n} \|x - f(x)\| > 0$$

Dunque, consideriamo la semiretta di origine  $f(x)$  e passante per  $x$ , che è ben definita perché i due punti non coincidono, e definiamo  $r(x)$  come l'intersezione di tale semiretta con il bordo di  $B_n$ . Tale funzione è ovviamente  $C^\infty$  ed è una retrazione di  $B_n$  sul bordo, cosa che viola il teorema di non retrazione.  $\square$

Dal teorema di Brouwer segue banalmente la generalizzazione del teorema di non retrazione alle funzioni continue. Se infatti esistesse una retrazione continua  $r : B_n \rightarrow \mathbb{S}^{n-1}$  componendo con la mappa antipodale su  $\mathbb{S}^{n-1}$  otterremmo una funzione continua da  $B_n$  in  $\mathbb{S}^{n-1} \subseteq B_n$  priva di punti fissi (dato che i punti del bordo andrebbero nel loro antipodale, mentre i punti interni andrebbero nel bordo).

<sup>2</sup>Bisognerebbe precisare che la funzione approssimante, ottenuta ad esempio con i polinomi di Bernstein, potrebbe non essere definita a valori in  $B_n$ , bensì eventualmente in una palla leggermente più grossa. In tal caso possiamo comporre  $g$  con un'omotopia per farla entrare in  $B_n$ , ed è possibile farlo in modo che la funzione riscalata sia ancora arbitrariamente vicina a  $f$  in norma uniforme.

### 13.1.1 Teorema di Perron-Frobenius

**Teorema 13.1.3** (Perron-Frobenius, enunciato parziale). *Sia  $A \in M(n, \mathbb{R})$  una matrice a entrate non negative. Allora  $A$  ammette un autovalore non negativo  $\lambda$  e un autovettore per  $\lambda$  di coordinate non negative.*

*Dimostrazione.* Sia  $\Delta = \{x \in \mathbb{R}^n \text{ t.c. } 0 \leq x_i \leq 1 \forall i \leq n, \sum_{i=1}^n x_i = 1\}$  il simplesso standard. Se  $\ker A \cap \Delta \neq \emptyset$  abbiamo finito, perché esiste un autovettore a entrate non negative per l'autovalore zero. Altrimenti è ben definita la mappa da  $\Delta$  in sé che manda  $x$  in  $\frac{Ax}{\|Ax\|_1}$ . L'immagine sta ancora in  $\Delta$  perché una matrice a entrate non negative manda vettori a entrate non negative in vettori a entrate non negative, e dunque  $\frac{Ax}{\|Ax\|_1}$  ha entrate tutte positive e ha norma 1 unitaria. Poiché  $\Delta \cong B_{n-1}(0, 1)$  il teorema di Brouwer ci assicura l'esistenza di un punto fisso, ossia di un vettore tale che  $Ax = \|Ax\|_1 x$ . Segue dunque la tesi.  $\square$

Il teorema vero in realtà è molto più corposo, ma non vi auguro di scoprirlo.



# Capitolo 14

## Delirio e voli pindarici

Ah, ma c'è anche questo fatterello interessante. . .

---

*Pietro Majer, tipo sempre.*

Qui ho ammassato, con pochissimo tatto, tutti gli argomenti cui non sono riuscito a trovare una collocazione all'interno del libro. Non aspettatevi alcuna coerenza (peraltro, assente anche in tutti i capitoli precedenti).

### 14.1 Proiettili vaganti

Non mi avrete mai, str\*nzi!

---

La polizia, sulla Fiat Punto d'ordinanza, insegue un ladro di auto che non ha trovato niente di meglio di una Panda. Entrambe le macchine si muovono alla stessa (misera) velocità, e la Madama spara al malvivente per forargli le gomme. Supponendo che il criminale abbia appena eseguito una svolta a novanta gradi e che l'area attorno all'incrocio sia priva di palazzi, dove si deve trovare un onesto cittadino per evitare di essere gambizzato per sbaglio?

La situazione può essere tradotta come segue. Supponendo che la distanza tra le due macchine, costante in quanto si muovono alla stessa velocità, sia 1, l'area critica è l'unione dei segmenti della forma

$$\left\{ (x, y) \in (\mathbb{R}^+)^2 \mid \frac{x}{t} + \frac{y}{1-t} = 1 \right\}_{t \in ]0,1[}$$

Infatti essi rappresentano le rette che incontrano l'asse delle ascisse in  $t$  e l'asse delle ordinate in  $1-t$ , che sono rispettivamente la distanza dall'incrocio del malvivente e della pubblica sicurezza. Allora la zona da evitare è il sottografico dell'involuppo delle curve, che nel nostro caso è dato da

$$f(x) = \sup_{t \in ]0,1[} \frac{x(t-1)}{t} + 1 - t = \sup_{t \in ]0,1[} x - \frac{x}{t} + 1 - t$$

Imponendo la derivata uguale a zero si ha

$$\frac{x}{t^2} = 1 \quad \longrightarrow \quad t = \sqrt{x}$$

Si ricava così che l'involuppo è dato da

$$f(x) = \frac{x(\sqrt{x}-1)}{\sqrt{x}} + 1 - \sqrt{x} = (\sqrt{x}-1)^2$$

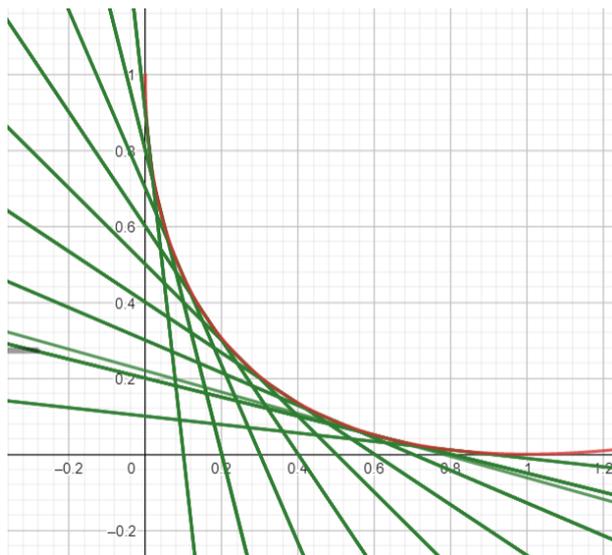


Figura 14.1: Rappresentazione dell'involuppo delle rette a somma delle intercette fissata

### 14.1.1 Generalizzazione del caso precedente

Dato un numero reale  $a$ , dimostriamo che l'involuppo delle rette

$$\left\{ (x, y) \in (\mathbb{R}^+)^2 \mid \frac{x}{t} + \frac{y}{s} = 1 \right\}_{t, s \in ]0, 1[} \quad \text{con } t^a + s^a = 1$$

è della forma  $x^b + y^b = 1$  per qualche  $b \in \mathbb{R}$ .

Potremmo procedere brutalmente come nella sezione precedente, ma generalmente otterremmo funzioni le cui derivate sono orrende e dagli zeri incalcolabili. Seguiamo allora un'altra strada. Comunque si scelgano  $\alpha$  e  $\beta$  reali si ha che

$$\frac{x}{t} + \frac{y}{s} = 1 \quad \implies \quad \left( \frac{x}{t} + \frac{y}{s} \right)^\alpha (t^a + s^a)^\beta = 1$$

Ora, l'obiettivo è ricavare  $x^b + y^b$  da qualche parte. Per farlo, possiamo provare ad applicare la disuguaglianza di Hölder<sup>1</sup>, scegliendo un  $\alpha$  che risolve  $\alpha + \alpha/a = 1$ . Infatti con un po' di conti che vi risparmio si ha che

$$1 = \left( \frac{x}{t} + \frac{y}{s} \right)^\alpha (t^a + s^a)^{\alpha/a} \geq \left( \frac{x}{t} \right)^\alpha t^\alpha + \left( \frac{y}{s} \right)^\alpha s^\alpha = x^\alpha + y^\alpha$$

In altri termini, tutti i punti delle rette stanno sotto il grafico di  $x^\alpha + y^\alpha$ . Viceversa, si prova facilmente che la tangente al grafico in ogni punto è una retta della famiglia considerata, e dunque il grafico è l'involuppo delle rette date.

## 14.2 Coniugio di contrazioni

**Teorema 14.2.1.** *Siano  $f, g : E \rightarrow E$  omeomorfismi e contrazioni, dove  $E$  è uno spazio di Banach. Se  $d_\infty(f, g) \leq +\infty$ , esiste un'unica funzione limitata  $u \in B(E)$  tale che  $(I + u)$  coniuga  $f$  e  $g$ , cioè*

$$(I + u) \circ f = g \circ (I + u)$$

*Inoltre  $(I + u)$  è un omeomorfismo.*

*Dimostrazione.* Consideriamo la funzione  $T$  su  $B(E)$  definita come

$$T(u) = g \circ (I + u) \circ f^{-1} - I$$

<sup>1</sup>vedi 2.3.2.

Osserviamo che  $T$  manda  $B(E)$  in sé. Infatti

$$\|T(u)\|_\infty = \sup_{x \in E} \|(g \circ (I + u) - f) \circ f^{-1}(x)\|$$

Ora, poiché  $f^{-1}$  è bigettiva, è equivalente fare il *sup* su  $y \in f(E) = E$ . Allora possiamo eliminare  $f^{-1}$  e ottenere

$$\|T(u)\|_\infty = \|g \circ (I + u) - f\|_\infty = \|g \circ (I + u) - g + g - f\|_\infty \leq \|g \circ (I + u) - g\| + \|g - f\|_\infty \leq K_g \|u\|_\infty + d_\infty(f, g)$$

Da cui segue che anche  $T(u)$  è limitata.

Ora, mostriamo che  $T$  è una contrazione, cosa che ci permette di concludere l'esistenza e unicità del punto fisso (che coniuga  $f$  e  $g$ ) perché  $B(E)$  è completo. Infatti, ragionando in modo identico a prima, si giunge a

$$\|T(u) - T(v)\|_\infty \leq K_g \|u - v\|$$

Ora, mostriamo che  $(I + u)$  è un omeomorfismo. Possiamo innanzitutto osservare che la soluzione  $u$  deve essere continua, dato che  $BC(E)$  è un chiuso non vuoto  $T$ -invariante, e perciò lo è anche  $(I + u)$ .

Per quanto riguarda l'invertibilità, scambiando  $f$  e  $g$  nell'ipotesi del teorema otteniamo un'unica funzione  $(I + v)$  tale che  $(I + v) \circ g = f \circ (I + v)$ . Allora, se applichiamo  $(I + u) \circ (I + v)$  a  $g$  otteniamo

$$(I + u) \circ (I + v) \circ g = (I + u) \circ f \circ (I + v) = g \circ (I + u) \circ (I + v)$$

Allora  $(I + u) \circ (I + v)$ , che è della forma  $(I + w)$  con  $w \in B(E)$ , coniuga  $g$  con se stessa. Per quanto visto prima, esiste un'unica funzione di questo tipo, e dunque  $(I + u) \circ (I + v) = I$  poiché ovviamente  $I \circ g = g \circ I$ . Reiterando il ragionamento con  $f$  otteniamo anche  $(I + v) \circ (I + u) = I$ , e da queste condizioni segue che  $(I + u) = (I + v)^{-1}$ . Peraltro, anche  $(I + v)$  è continua per simmetria, dunque  $(I + u)$  è un omeomorfismo.  $\square$

## 14.3 Derivate del Dini

**Definizione 85.** Data una funzione reale e continua  $f$ , si definisce la sua derivata destra superiore del Dini come

$$D^+ f(x) = \limsup_{y \rightarrow x^+} \frac{f(y) - f(x)}{h} = \inf_{x \in U} \sup_{\text{intorno } y \in U \setminus \{x\}} \frac{f(y) - f(x)}{h}$$

Analogamente si definisce la derivata destra inferiore come

$$D_+ f(x) = \liminf_{y \rightarrow x^+} \frac{f(y) - f(x)}{h} = \sup_{x \in U} \inf_{\text{intorno } y \in U \setminus \{x\}} \frac{f(y) - f(x)}{h}$$

Il lettore è invitato (ma anche no) a immaginare le definizioni delle altre derivate del Dini.

## 14.4 Luogo di zeri di una derivata

**Teorema 14.4.1.** Data una funzione  $f : I \rightarrow \mathbb{R}$  derivabile ovunque, il luogo di zeri della sua derivata è un  $G_\delta$ .

*Dimostrazione.* Sia  $A_n$  l'unione degli intervalli aperti della forma  $]a, b[ \subseteq I$ , di ampiezza al più  $1/n$  e il cui rapporto incrementale agli estremi è minore di  $1/n$  in modulo. Mostriamo che

$$\{x \in I \text{ t.c. } f'(x) = 0\} = \bigcap_{n \in \mathbb{N}} A_n$$

Mostriamo la doppia inclusione. Se  $x$  ha derivata nulla, per definizione  $\lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t} = 0$ . Allora fissato  $\varepsilon = 1/2n$  esiste un  $\delta$  per cui, se  $|t| < \delta$ , il rapporto incrementale è minore di  $1/2n$ . Allora scegliamo  $a_n < x < b_n$  in modo che  $b_n - a_n < \min 1/n, \delta$ . Si ha dunque

$$\left| \frac{f(b_n) - f(a_n)}{b_n - a_n} \right| \leq \left| \frac{f(b_n) - f(x)}{b_n - a_n} \right| + \left| \frac{f(x) - f(a_n)}{b_n - a_n} \right| \leq \left| \frac{f(b_n) - f(x)}{b_n - x} \right| + \left| \frac{f(x) - f(a_n)}{x - a_n} \right| \leq \frac{1}{n}$$

Perciò, per ogni  $n \in \mathbb{N}$ ,  $x \in A_n$ .

Viceversa, se  $x$  ha derivata non nulla  $d$ , vale che

$$f(x+t) - f(x) = dt + o(t)$$

Perciò, comunque si scelgano  $a < x < b$  a distanza minore di  $1/n$ , vale

$$f(b) - f(x) = d(b-x) + o(b-x) = d(b-x) + o(1/n)$$

$$f(x) - f(a) = d(x-a) + o(x-a) = d(x-a) + o(1/n)$$

Sommando le due scritte e dividendo per  $(b-a)$  otteniamo

$$\frac{f(b) - f(a)}{b-a} = d + o(1/n)$$

Allora, scegliendo  $n$  tale che  $|d| > 1/n$ , otteniamo un intorno di  $x$  in cui nessun rapporto incrementale può essere piccolo a piacere.  $\square$

## 14.5 Derivate di Pompeiu

Sappiamo che, se una funzione è integrabile secondo Riemann, allora per il teorema fondamentale del calcolo è la derivata della sua funzione integrale. Tuttavia, il seguente esempio del 1900 mostra una funzione derivabile ma con derivata non Riemann-integrabile.

Sia  $\{a_k\}$  una successione di numeri reali positivi la cui serie converge e sia  $q_k$  una numerazione dei razionali in  $[0, 1]$ . Consideriamo la funzione

$$f(x) = \sum_{k=0}^{\infty} a_k (x - q_k)^{1/3} - C$$

Dove  $C$  è una costante che serve a porre  $f(0) = 0$ . Inoltre, a meno di riscalare gli  $a_k$ , supponiamo  $f(1) = 1$ .

La funzione limite è continua perché limite uniforme di funzioni continue, ed è strettamente crescente perché lo sono tutti i termini della somma. Allora è un omeomorfismo di  $[0, 1]$  in sé per un teorema visto ad Analisi 1.

Osserviamo ora che, in senso lato, la derivata della funzione esiste ed è uguale alla serie delle derivate  $g(x)$ . Per mostrarlo distinguiamo due casi.

- Se  $g(x) = +\infty$ , allora essendo la serie a termini positivi

$$\forall M \exists N \text{ t.c. } \sum_{k=0}^N \frac{a_k}{3} |x - q_k|^{-2/3} > M$$

Ora, consideriamo la derivata inferiore del Dini di  $f$  e tronchiamo la serie al termine  $N$ -esimo, così da poter portare il  $\liminf$  all'interno.

$$\liminf_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} \geq \sum_{k=0}^N \frac{a_k}{3} |x - q_k|^{-2/3} > M$$

Cioè possiamo definire  $f'(x) = +\infty$ .

- Se  $g(x)$  è finito, mostriamo che  $f(y) - f(x) - g(x)(y-x) = o(y-x)$ . Dato che le quantità a sinistra dell'uguale sono delle serie assolutamente convergenti, possiamo portarle tutte sotto lo stesso segno di sommatoria. Otteniamo allora, con un po' di manipolazioni algebriche, che

$$f(y) - f(x) - g(x)(y-x) = \sum_{k=0}^{\infty} (y-x)a_k(x-q_k)^{-2/3} \left( (x-q_k)^{2/3} \frac{(y-q_k)^{1/3} - (x-q_k)^{1/3}}{(y-q_k) - (x-q_k)} - \frac{1}{3} \right) =$$

$$= (y-x) \sum_{k=0}^{\infty} a_k (x-q_k)^{-2/3} \left( \frac{1}{\left(\frac{y-q_k}{x-q_k}\right)^{2/3} - \left(\frac{y-q_k}{x-q_k}\right)^{1/3} + 1} - \frac{1}{3} \right)$$

Ora, il termine tra parentesi è infinitesimo per  $y \rightarrow x$ , e dunque è limitato da una costante. Allora la serie è dominata termine a termine dalla serie che definisce  $g(x)$ , assolutamente convergente per ipotesi. Perciò per convergenza dominata il limite della serie è la serie dei limiti, cioè è infinitesima. Allora l'espressione diventa  $(y-x)o(1) = o(y-x)$ , come voluto.

Ora, la derivata di  $f$  è uniformemente limitata dal basso dalla serie degli  $a_k/3$ , come si verifica banalmente. Meno ovvio è, invece, che l'insieme dei punti di continuità di  $f'$  coincide con l'insieme su cui essa è infinita. Infatti, se  $f'(x) = +\infty$ , per ogni  $M$  esiste un  $N$  tale che la somma parziale  $N$ -esima è maggiore di  $N$ . Ora, quest'ultima è una somma finita di funzioni continue, perciò esiste un intorno in cui la somma parziale è maggiore di  $M$ . Dunque anche la derivata, che è maggiore della somma parziale, è maggiore di  $M$  in quell'intorno. Invece, se  $f'(x) < +\infty$ ,  $f'$  non è ivi continua, perché per densità dei razionali in ogni intorno di  $x$  esiste un valore in cui la derivata diverge.

Adesso, consideriamo la funzione inversa  $F = f^{-1}$ . Per quanto osservato finora,  $F$  è un omeomorfismo strettamente crescente la cui derivata è  $F'(x) = 1/g(F(x))$ . Dunque  $F$  è derivabile ovunque nel dominio con derivata limitata, e perciò è lipschitziana.

Tuttavia,  $F'$  non è Riemann-integrabile. Se lo fosse, dovrebbe essere continua a meno di un insieme trascurabile. Ora, l'insieme di continuità di  $F'$  sono i punti dove si annulla, e l'integrale di una funzione nulla quasi ovunque è nullo. Ma allora per teorema fondamentale del calcolo integrale,  $F \equiv 0$ , perché dovrebbe coincidere con l'integrale della sua derivata.

### 14.5.1 Ma non ne bastava una sola?

**Definizione 86.** Una funzione reale di variabile reale è una derivata di Pompeiu se è la derivata di una funzione derivabile ovunque e se il suo luogo di zeri è denso.

Questo obbrobrio di funzioni è un sottospazio vettoriale delle funzioni da  $I$  in  $\mathbb{R}$ . Infatti:

- La funzione identicamente nulla è una derivata di Pompeiu.
- Moltiplicando una funzione per uno scalare non si varia il suo luogo di zeri;
- La somma di una derivata di Pompeiu è una derivata di Pompeiu, in quanto il luogo di zeri della somma è almeno l'intersezione del luogo di zeri delle singole funzioni, che sono  $G_\delta^2$  densi e dunque hanno intersezione densa.

Si può dire di più. Se  $I$  è un intervallo compatto, le derivate di Pompeiu sono un chiuso di  $B(I, \mathbb{R})$  (e dunque sono uno spazio di Banach con la distanza uniforme). Infatti, se una successione di derivate converge, possiamo imporre che le antiderivate abbiano tutte lo stesso valore in un  $x_0 \in I$  e concludere che la funzione limite è anch'essa derivata di qualcosa, per teorema di passaggio al limite sotto il segno di differenziale.<sup>3</sup> Inoltre, il limite ha un insieme di zeri denso, perché l'intersezione numerabile di  $G_\delta$  densi è anch'essa un  $G_\delta$  ed è densa per completezza di  $I$ .

## 14.6 Formula di Stirling, rivisitata e corretta

Con i cannoni di questo e altri corsi possiamo (ri)dimostrare la formula di approssimazione del fattoriale:

**Teorema 14.6.1** (Formula di Stirling). Per  $n \rightarrow +\infty$

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

<sup>2</sup>Vedasi 14.4.

<sup>3</sup>Vedasi 6.12.1. Il teorema si può applicare perché l'intervallo su cui sono definite le funzioni è limitato.

*Dimostrazione.* Ricordiamo che, per quanto visto a Geometria 2, se  $f : B(0, R) \rightarrow \mathbb{C}$  è olomorfa,

$$\frac{f^n(0)}{n!} = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z^{n+1}} dz$$

dove  $\gamma(t) = re^{it}$ , con  $r < R$  e  $t \in [-\pi, \pi]$ . In particolare, applicando il ragionamento alla funzione  $f(z) = e^z$ , otteniamo

$$\frac{1}{n!} = \frac{1}{2\pi i r^{n+1}} \int_{-\pi}^{\pi} e^{re^{it}} e^{-(n+1)it} r i e^{it} dt = \frac{1}{2\pi r^n} \int_{-\pi}^{\pi} e^{re^{it}} e^{-nit} dt$$

Dato che  $e^z$  è intera possiamo scegliere  $r = n$ . Dunque, con una innocua manipolazione formale,

$$\frac{1}{n!} = \frac{1}{2\pi n^n} \int_{-\pi}^{\pi} e^{n - \frac{nt^2}{2} + n(e^{it} - 1 - it + \frac{t^2}{2})} dt = \frac{1}{2\pi} \left(\frac{e}{n}\right)^n \int_{-\pi}^{\pi} e^{-\frac{nt^2}{2} + n(e^{it} - 1 - it + \frac{t^2}{2})} dt$$

e con il cambio di variabile  $s = \sqrt{nt}$  (che per comodità applichiamo solo a parte dell'integrando, capiremo dopo perché)

$$\frac{1}{n!} = \frac{1}{2\pi\sqrt{n}} \left(\frac{e}{n}\right)^n \int_{\mathbb{R}} e^{-\frac{s^2}{2} + n(e^{it} - 1 - it + \frac{t^2}{2})} \chi_{[-\pi\sqrt{n}, \pi\sqrt{n}]} ds$$

Per dimostrare il teorema di Stirling basta ora far vedere che la successione degli integrali converge a  $\sqrt{2\pi} = \int_{\mathbb{R}} e^{-\frac{s^2}{2}} ds$  per  $n \rightarrow +\infty$ . Il pezzo tra parentesi nell'esponente è il resto dello sviluppo al secondo ordine della funzione  $e^{it}$ , che sappiamo essere un  $o(t^2) = o\left(\frac{s^2}{n}\right)$ . Dunque l'integrando converge puntualmente a  $e^{-\frac{s^2}{2}}$  per  $n \rightarrow +\infty$ . Per ottenere la convergenza degli integrali useremo il ConDom. Infatti, per quanto sappiamo sul modulo dell'esponenziale possiamo dire che

$$\left| e^{n(e^{it} - 1 - it + \frac{t^2}{2})} \right| = e^{n\Re(e^{it} - 1 - it + \frac{t^2}{2})} = e^{n(\cos t - 1 + \frac{t^2}{2})}$$

e applicando la formula del resto di Lagrange al quarto ordine alla funzione  $\cos t$

$$\cos t - 1 + \frac{t^2}{2} = \cos \theta \frac{t^4}{24} \leq \frac{t^4}{24} \leq \frac{\pi^2 t^2}{24} < \frac{t^2}{2}$$

Quindi l'integrando si domina con la funzione  $e^{s^2(\frac{\pi^2}{24} - \frac{1}{2})}$ , integrabile dato che l'esponente è negativo. Allora

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} e^{-\frac{s^2}{2} + n(e^{it} - 1 - it + \frac{t^2}{2})} \chi_{[-\pi\sqrt{n}, \pi\sqrt{n}]} ds = \int_{\mathbb{R}} e^{-\frac{s^2}{2}} ds = \sqrt{2\pi}$$

Da cui

$$\frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{n!} \rightarrow 1$$

□

## 14.7 Quoziente di Rayleigh

**Definizione 87.** Data una funzione  $f : E \rightarrow F$ , si dice che  $x \in E$  è un punto critico se  $df_x$  non è iniettivo. In tal caso  $f(x)$  si dice valore critico.

Sia  $A$  una matrice reale simmetrica. Definiamo il suo quoziente di Rayleigh<sup>4</sup> come la funzione

$$q(x) = \frac{x^T A x}{x^T x}$$

**Teorema 14.7.1.** Un numero complesso  $\lambda$  è autovalore per  $A$  di autovettore  $x$  se e solo se  $x$  è un punto critico di  $q$ , di valore critico  $\lambda$ .

<sup>4</sup>No, non è quello di One Piece.

*Dimostrazione.* Osserviamo che, nel caso in cui l'immagine è  $\mathbb{R}$ , essere un punto critico significa avere gradiente nullo. Scriviamo allora<sup>5</sup>

$$\vec{\nabla} \frac{x^T Ax}{x^T x} = \frac{x^T x \vec{\nabla} x^T Ax - x^T Ax \vec{\nabla} x^T x}{(x^T x)^2} = 0$$

Ora, osserviamo che

$$(x+u)^T A(x+u) = x^T Ax + 2Ax \cdot u + o(\|u\|)$$

Per cui il gradiente di  $x^T Ax$  è  $2Ax$ . Allora, facendo lo stesso ragionamento con  $A = Id$ , otteniamo che  $\vec{\nabla} x^T x = 2x$ . Otteniamo allora

$$(x^T x)Ax - (x^T Ax)x = 0$$

Cioè

$$Ax - q(x)x = 0$$

Allora, posto  $q(x) = \lambda$ , segue la tesi in entrambi i versi (perché tutte le implicazioni sono doppie).  $\square$

## 14.8 Equazioni ai commutatori

**Definizione 88.** Dati due operatori lineari  $A, B$ , si definisce il loro commutatore come

$$[A, B] = AB - BA$$

Un'equazione differenziale matriciale del tipo

$$\begin{cases} X'(t) = [A(t), X(t)] \\ X(t_0) = X_0 \end{cases}$$

si dice equazione ai commutatori. Di questa sappiamo tutto, perché si verifica che la soluzione generale è  $X(t) = W_A(t, t_0)X_0W_A(t_0, t)$ . Osserviamo in particolare che, dato che  $W_A(t, t_0) = W_A(t_0, t)^{-1}$ , la soluzione è un cammino di coniugati di  $X_0$ , e dunque ne condivide il polinomio minimo e il polinomio caratteristico. Questo ci dice che, se la matrice iniziale ha qualche proprietà spettrale particolare, questa passa a tutta la soluzione.

### 14.8.1 Proiettori e commutatori

I have a P, I have a P', OH! PP'!

---

**Definizione 89.** Un operatore lineare idempotente  $P = P^2$  si dice proiettore.

Poiché l'essere idempotenti è una proprietà che dipende solo dal polinomio caratteristico, se  $X(t)$  risolve un'equazione ai commutatori con condizione iniziale idempotente allora  $X(t)$  è un cammino di proiettori. Il teorema seguente mostra che vale anche il viceversa.

**Teorema 14.8.1.** Data una curva  $P: \mathbb{R} \rightarrow M(n, \mathbb{R})$  di classe  $C^1$ , se  $P(t)$  è un proiettore a ogni tempo  $t$  esiste una matrice  $Q$  tale che  $P'(t) = [Q, P] = QP - PQ$ . In altri termini, un cammino di proiettori risolve una particolare equazione ai commutatori.

*Dimostrazione.* Poiché  $P = P^2$ , derivando otteniamo  $P' = PP' + P'P$ . Moltiplicando per  $P$  abbiamo

$$PP' = PPP' + PP'P = PP' + PP'P$$

Da cui  $PP'P = 0$ . Ora, poniamo  $Q = [P', P] = P'P - PP'$  e verifichiamo che funziona.

$$[Q, P] = P'PP - PP'P - (PP'P - PPP') = P'P + PP' = P'$$

$\square$

---

<sup>5</sup>Faremo uso della seguente regola di calcolo, molto simile a quella per le derivate:  $\vec{\nabla}(u/v) = (v\vec{\nabla}u - u\vec{\nabla}v)/(v^2)$ .



# Bibliografia

- [Vergamini-Marchetti-Traino, 2020] M. Vergamini, A. Marchetti, L. Traino, *Geometria 2* (2020)  
[https://uz.sns.it/~marco\\_vergamini/appunti.html](https://uz.sns.it/~marco_vergamini/appunti.html)
- [Mangioni, 2019] G. Mangioni, *Gran varietà matematico* (2019)  
<https://poisson.phc.dm.unipi.it/~mangioni/>
- [SNS, 2015] AA.VV., *Dispense del corso di Complementi di Matematica* (2015)  
[https://serse.sns.it/sns/\\_DOCs/CRS\\_MATERIALE\\_DIDATTICO/972df8fe86d43dc55378736e125fd8de.pdf](https://serse.sns.it/sns/_DOCs/CRS_MATERIALE_DIDATTICO/972df8fe86d43dc55378736e125fd8de.pdf)
- [Ambrosio-Da Prato-Mennucci, 2011] L. Ambrosio, G. Da Prato, A. Mennucci, *An introduction to measure theory and probability* (2011)
- [Pontryagin, 1962] L.S. Pontryagin, *Ordinary differential equations* (1962)
- [Cartan, 1981] H. Cartan, *Differential calculus on normed space* (1981)
- [De Marco, 1999] G. De Marco, *Analisi 2, teoria ed esercizi* (1999)
- [Wheeden-Zigmund, 1977] R. Wheeden, A. Zigmund, *Measure and integral*, (1977)
- [Bini et al., 1992] R. Bevilacqua, D. Bini, M. Capovani, O. Menchi, *Metodi Numerici*, (1992)