Raccolta di teoremi

Emanuele Paolini

10 aprile 2025

Il seguente testo è una raccolta incrementale di teoremi in cui mi sono imbattuto per vari motivi (principalmente per la didattica) e di cui ho voluto scrivermi la dimostrazione (o a volte solo l'enunciato). Questo testo dunque non è completo, manca di molte definizioni e di molti argomenti ma forse si estenderà in futuro (presumibilmente senza mai raggiungere la completezza). L'unica utilità che penso potrebbe avere è per consultazione (per me ha, di fatto, tale utilità).

Indice

1 Analisi Reale **5**
1.1 Insiemistica . . . 5
1.2 I numeri reali . . . 5
1.3 I numeri complessi . . . 6
1.4 Successioni di numeri reali . . . 7
1.5 Teoremi sulle serie . . . 7
1.6 Funzioni di una variabile reale . . . 9
1.7 Uniforme continuità . . . 14
1.7.1 Funzioni uniformemente continue, Lipschitziane, α-Hölderiane 14
1.7.2 Successioni di Cauchy . . . 16
1.7.3 Estensioni . . . 17
1.7.4 Modulo di continuità . . . 18
1.8 Funzioni convesse . . . 19

2 Topologia **21**
2.1 Spazi topologici . . . 21
2.2 Funzioni continue . . . 24
2.3 Spazi metrici . . . 25
2.4 Limiti . . . 26
2.5 Retta reale estesa . . . 28
2.6 Successioni in spazi topologici . . . 28
2.7 Teoremi di compattezza . . . 28
2.8 Teoremi di punto fisso . . . 32

3 Spazi vettoriali topologici **33**
3.1 Il differenziale . . . 33
3.2 Convergenza uniforme . . . 34
3.3 Serie e derivata discreta . . . 35
3.4 L'esponenziale di matrici . . . 36
3.5 Algebra lineare . . . 38
3.6 Invertibilità locale . . . 39
3.7 Equazioni differenziali . . . 42
3.7.1 Sistemi di equazioni differenziali lineari del primo ordine a coefficienti costanti . . . 43
3.7.2 Equazioni differenziali lineari di ordine n . . . 43
3.7.3 Equazioni non lineari . . . 46
3.8 L'integrale di Lebesgue . . . 48
3.9 Spazi L^p . . . 52
3.10 Semi-continuità . . . 52
3.11 Serie di Fourier . . . 53
3.12 Trasformate di Fourier . . . 56

Capitolo 1

Analisi Reale

1.1 Insiemistica

Proposizione 1.1.1. *Sia $f\colon X \to Y$ una funzione. Se $A \subset X$, $B \subset Y$, $\mathcal{F}$ è una famiglia di sottoinsiemi di X e $\mathcal{G}$ è una famiglia di sottoinsiemi di Y si ha:*

1. $A \subset f^{-1}(f(A))$;

2. $B \supset f(f^{-1}(B))$;

3. $f(\bigcup \mathcal{F}) = \bigcup f(\mathcal{F})$;

4. $f(\bigcap \mathcal{F}) = \bigcap f(\mathcal{F})$;

Dimostrazione.

1. Dato $x \in A$ si ha $f(x) \in f(A)$ e quindi $x \in f^{-1}(f(A))$.

   D'altra parte possiamo trovare un esempio in cui l'inclusione è stretta, basta considerare $X = \{1, 2\}$, $Y = \{3\}$ e f la funzione con grafico $\{(1, 3), (2, 3)\} \subset X \times Y$. Scelto $A = \{1\}$ si ha $f(A) = Y$ e quindi $f^{-1}(f(A)) = X \neq A$.

2. Sia dato $b \in f(f^{-1}(B))$. Allora esiste $a \in f^{-1}(B)$ tale che $f(a) = b$. Dunque esiste $b' \in b$ tale che $f(a) = b'$. Dunque $b = b' \in B$.

   D'altra parte costuiamo un esempio in cui l'inclusione è stretta. Scegliamo $X = \{1\}$, $Y = \{2, 3\}$ e f la funzione con grafico $\{(1, 2)\} \subset X \times Y$. Scelto $B = Y$ si ha $f^{-1}(B) = X$ ma $f(f^{-1}(B)) = f(X) = \{2\} \neq B$.

3. Dato $y \in f(\bigcup \mathcal{F})$ esiste $x \in \bigcup \mathcal{F}$ tale che $f(x) = y$. Inoltre esiste $F \in \mathcal{F}$ tale che $x \in F$. Dunque $y = f(x) \in f(F) \subset \bigcup f(\mathcal{F})$.

   Dato $y \in \bigcup f(\mathcal{F})$ esiste $F \in \mathcal{F}$ tale che $y \in f(F)$. Esiste dunque $x \in F$ tale che $f(x) = y$. Chiaramente $y = f(x) \in f(F) \subset f(\bigcup \mathcal{F})$.

□

1.2 I numeri reali

Denoteremo con $\mathbb{R}$ un insieme dotato delle operazioni $S\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (somma), $P\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (prodotto) e dalla relazione $M \subset \mathbb{R} \times \mathbb{R}$ (relazione d'ordine) tale che i seguenti assiomi siano validi. Per semplicità si useranno le notazioni $x + y = S(x, y)$, $xy = x \cdot y = P(x, y)$, $x \leq y \Leftrightarrow y \geq x \Leftrightarrow (x, y) \in M$, $x < y \Leftrightarrow y > x \Leftrightarrow (x, y) \in M \wedge x \neq y$.

1. Per ogni $x, y, z \in \mathbb{R}$ vale $x + y = y + x$, $xy = yx$ (proprietà commutativa), $x + (y + z) = (x + y) + z$, $x(yz) = (xy)z$ (proprietà associativa), $x(y + z) = xy + xz$ (proprietà distributiva), $x \le y \vee y \le x$ (ordinamento totale), $x \le y \Rightarrow x + z \le y + z$ (monotonia), esistono $0, 1 \in \mathbb{R}$ tali che $0 + x = x$, $1x = x$ (esistenza degli elementi neutri).

2. Per ogni $x \in \mathbb{R}$ esiste $y \in \mathbb{R}$ (denoteremo y con $-x$) tale che $x+y = 0$, per ogni $x \in \mathbb{R} \setminus \{0\}$ esiste $y \in \mathbb{R}$ (denoteremo y con $1/x$) tale che $xy = 1$ (esistenza di opposto e inverso).

3. Ogni insieme superiormente limitato ammette estremo superiore (completezza).

Per comodità scriveremo $x - y = x + (-y)$ (dove $-y$ è l'opposto di y) e se $y \neq 0$ scriveremo $x/y = \frac{x}{y} = x \cdot (1/y)$ (dove $1/y$ è l'inverso di y).

Un sottoinsieme X di $\mathbb{R}$ si dice *induttivo* se $0 \in X$ e $x \in X \Rightarrow x + 1 \in X$. Il più piccolo sottoinsieme induttivo di $\mathbb{R}$ (ovvero l'intersezione di tutti i sottoinsiemi induttivi di $\mathbb{R}$) viene chiamato $\mathbb{N}$ (numeri naturali).

Proposizione 1.2.1 (principio di Archimede)**.** *Dato $x \in \mathbb{R}$ esiste $n \in \mathbb{N}$ tale che $x \le n$ ovvero $\mathbb{N}$ non è superiormente limitato.*

Dimostrazione. Se $\mathbb{N}$ fosse superiormente limitato ammetterebbe estremo superiore $s \in \mathbb{R}$. Allora $s - 1$ non sarebbe un maggiorante, cioè esisterebbe $n \in \mathbb{N}$ tale che $s - 1 \le n$. Dunque si avrebbe $s \le n + 1$ e quindi s non sarebbe un maggiorante. □

Definizione 1.2.2 (intervallo)**.** *Sia $I \subset \mathbb{R}$ (o di $\bar{\mathbb{R}}$). Diremo che I è un* intervallo *se dati $x, y \in I$ e dato qualunque z con $x \le z \le y$ si ha $z \in I$.*

Gli estremi di un intervallo I sono $a = \inf I$ (estremo inferiore) e $b = \sup I$ (estremo superiore). Se I è limitato allora $a, b \in \mathbb{R}$. Gli estremi dell'intervallo (quando sono finiti) possono essere o non essere inclusi nell'intervallo. Se l'estremo inferiore (rispettivamente superiore) non è incluso nell'intervallo si dice che l'intervallo è aperto a sinistra (rispettivamente a destra). Se invece gli estremi sono inclusi nell'intervallo si dice che l'intervallo è chiuso.

Un intervallo viene quindi univocamente determinato conoscendo i suoi estremi (finiti o infiniti) e sapendo se gli estremi sono contenuti oppure no nell'intervallo. Per denotare un intervallo useremo quindi scrivere gli estremi tra parentesi quadre e separati da una virgola. A seconda dell'orientamento delle parentesi quadre gli estremi vanno inclusi o esclusi. Gli intervalli di $\mathbb{R}$ sono dunque i seguenti: $]-\infty, +\infty[= \mathbb{R}$, $[a, +\infty[= \{x \in \mathbb{R} \colon a \le x\}$, $]a, +\infty[= \{x \in \mathbb{R} \colon a < x\}$, $]-\infty, b] = \{x \in \mathbb{R} \colon x \le b$, $]-\infty, b[= \{x \in \mathbb{R} \colon x < b$, $[a, b] = \{x \in \mathbb{R} \colon a \le x \le b\}$, $[a, b[= \{x \in \mathbb{R} \colon a \le x < b\}$, $]a, b] = \{x \in \mathbb{R} \colon a < x \le b\}$, $]a, b[= \{x \in \mathbb{R} \colon a < x < b\}$, $\emptyset$.

Gli intervalli di $\bar{\mathbb{R}}$ possono anche includere i punti $+\infty$ e $-\infty$.

Se I è un intervallo, denoteremo con $\overset{\circ}{I}$ l'intervallo aperto con gli stessi estremi e con $\bar{I}$ l'intervallo chiuso.

1.3 I numeri complessi

Esercizio 1.3.1. Trovare la lunghezza del lato del pentagono. Ovvero determinare $\sin(2\pi/5)$ e $\cos(2\pi/5)$.

Dimostrazione. I cinque vertici del pentagono inscritto nel cerchio di raggio 1 e orientato opportunamente sono le cinque soluzioni complesse dell'equazione $z^5 = 1$. Ovviamente se z è soluzione, allora $|z|^5 = |z^5| = 1$ cioè $z\bar{z} = |z|^2 = 1$.

Ovviamente 1 è soluzione e dividendo per $z-1$ si ottiene $z^5-1=(z-1)(z^4+z^3+z^2+z+1)$. Dobbiamo quindi trovare le quattro soluzioni di $z^4+z^3+z^2+z+1=0$. Sapendo che $z\bar z=1$ moltiplichiamo tutto per $\bar z^2$ ottenendo $z^2+z+1+\bar z+\bar z^2=(z+\bar z)^2+z+\bar z-1$. Con la sostituzione $w=z+\bar z$ si ottiene l'equazione di secondo grado $w^2+w-1=0$ che ha le due soluzioni $w=\frac{1\pm\sqrt5}{2}$. Se z è una delle due soluzioni diverse da 1 e con parte reale positiva, il lato del pentagono è dato da $l=|z-1|=\sqrt{(z-1)(\bar z-1)}=\sqrt{z\bar z-z-\bar z+1}=\sqrt{2-w}$ dove w sarà la soluzione positiva di $w^2+w-1=0$. Dunque si ottiene $l=\sqrt{\frac{3-\sqrt5}{2}}$. Per $\theta=2\pi/5$ si ha poi $\cos\theta=\Re z=\frac{z+\bar z}{2}=w/2=\frac{-1+\sqrt5}{4}$; mentre $\sin\theta=\sqrt{1-\cos^2\theta}=\sqrt{\frac{5+\sqrt5}{8}}$. □

1.4 Successioni di numeri reali

Teorema 1.4.1 (Bolzano-Weierstrass)**.** *Sia (x_k) una successione limitata di numeri reali. Allora è possibile estrarre una sottosuccessione (x_{k_j}) convergente.*

Dimostrazione. Sia $[a,b]$ un intervallo che contiene tutti i termini della successione. Per induzione troveremo una sottosuccessione x_{k_j} e una successione di intervalli $[a_j,b_j]$ con le seguenti proprietà: $b_j-a_j=2(b-a)/2^j$, $a_{j+1}\ge a_j$, $b_{j+1}\le b_j$, $x_{k_j}\in[a_j,b_j]$ e $\{k\colon x_k\in[a_j,b_j]\}$ è un insieme infinito.

Per $j=1$ scegliamo $a_1=a$, $b_1=b$ e $x_{k_1}=x_1$. Tutti i termini della successione stanno in $[a_1,b_1]$ e si ha $b_1-a_1=b-a$.

Siccome per ipotesi induttiva l'intervallo $[a_j,b_j]$ contiene infiniti punti della successione x_k, allora almeno uno dei due intervalli $[a_j,(a_j+b_j)/2]$ e $[(a_j+b_j)/2,b_j]$ contiene infiniti punti della successione. Sia $[a_{j+1},b_{j+1}]$ tale intervallo e sia $k_{j+1}>k_j$ tale che $x_{k_{j+1}}\in[a_{j+1},b_{j+1}]$. Chiaramente $a_{j+1}\ge a_j$, $b_{j+1}\le b_j$ e $b_{j+1}-a_{j+1}=(b_j-a_j)/2$.

È poi facile verificare che la sottosuccessione x_{k_j} è convergente. Infatti $a_j\le x_{k_j}\le b_j$. Le due successioni a_j e b_j sono monotòne dunque convergono. Essendo poi $|b_j-a_j|\to0$ le due successione a_j e b_j convergono allo stesso limite e quindi, per confronto, anche x_{k_j} converge a tale limite. □

1.5 Teoremi sulle serie

Teorema 1.5.1. *Siano a_k e b_k due successioni. Sia A_k una successione tale che $a_k=A_k-A_{k-1}$ (ad esempio $A_k=a_1+\ldots+a_k$ con $A_0=0$).*

Allora si ha

$$\sum_{k=p}^{q}a_kb_k=\sum_{k=p}^{q-1}A_k(b_k-b_{k+1})-A_{p-1}b_p+A_qb_q.$$

Dimostrazione. Si ha

$$\begin{aligned}\sum_{k=p}^{q}a_kb_k &= \sum_{k=p}^{q}(A_k-A_{k-1})b_k\\ &= \sum_{k=p}^{q}A_kb_k-\sum_{k=p-1}^{q-1}A_kb_{k+1}\\ &= \sum_{k=p}^{q-1}A_k(b_k-b_{k+1})+A_pb_p-A_{p-1}b_p.\end{aligned}$$

Infatti ponendo $j = k-1$ si ha

$$\sum_{k=p}^{q} A_{k-1} b_k = \sum_{j=p-1}^{q-1} A_j b_{j+1}$$

□

Teorema 1.5.2. *Siano a_k, b_k due successioni tali che $\sup_n \sum_{k=0}^n a_k < +\infty$ e b_k è decrescente e tende a zero. Allora la serie $\sum a_k b_k$ converge.*

Dimostrazione. Sia $S_n = \sum_{k=0}^n a_k b_k$ la successione delle somme parziali. Applicando la formula di somma per parti si ha

$$|S_q - S_{p-1}| = \left|\sum_{k=p}^{q} a_k b_k\right| \le \sum_{k=p}^{q-1} |A_k|(b_k - b_{k+1}) + |A_{p-1}|b_p + |A_q|b_q$$

dove $A_n = \sum_{k=0}^n a_k$. Posto $M = \sup_k A_k$ per ipotesi $M < +\infty$ e si ha quindi

$$\left|\sum_{k=p}^{q} a_k b_k\right| \le M \sum_{k=p}^{q-1} (b_k - b_{k+1}) + Mb_p + Mb_q = M(b_p - b_q) + Mb_p + Mb_q = 2Mb_p.$$

Essendo $b_p \to 0$ scelto comunque $\varepsilon > 0$ esiste N tale che per ogni $p > N$ si ha $2Mb_p < \varepsilon$ e quindi per ogni $q > p > N$ si trova $S_q - S_{p-1} < \varepsilon$. In conlusione S_n è una successione di Cauchy e quindi converge. □

Teorema 1.5.3. *Sia a_k una successione tale che $\sum_k a_k b_k$ converge per ogni $b \in c_0$ ($b_k \to 0$). Allora $a \in \ell^1$ ($\sum_k |a_k| < \infty$).*

Dimostrazione. Supponiamo per assurdo che $\sum_k |a_k| = \infty$. Sia k_n una successione crescente di naturali con $k_1 = 1$ e tale che $\sum_{k=k_n+1}^{k_{n+1}} |a_k| > 1$. Definiamo b_k in modo che $b_k = a_k/(n|a_k|)$ per $k = k_n + 1, \dots, k_{n+1}$. Chiaramente $b_k \to 0$, d'altra parte si ha

$$\sum_k a_k b_k = \sum_n \frac{1}{n} \sum_{k=k_n+1}^{k_{n+1}} |a_k| \ge \sum_n 1/n = \infty$$

che contraddice l'ipotesi. □

Teorema 1.5.4. *Supponiamo che la serie di potenze $\sum_{k=0}^\infty a_k z^k$ sia convergente per un certo $z \in \mathbb{C}$. Allora posto $f_n\colon [0,1] \to \mathbb{C}$, $f_n(t) = \sum_{k=0}^n a_k (tz)^k$ la successione (f_n) converge uniformemente. In particolare si ha*

$$\lim_{t \to 1-} \sum_{k=0}^{\infty} a_k (tz)^k == \sum_{k=0}^{\infty} a_k z^k.$$

Dimostrazione. Sia $r \in [0, +\infty]$ il raggio di convergenza della serie in questione. Siccome la serie converge in z abbiamo che $|z| \le r$. Inoltre notiamo subito che se fosse $|z| < r$ già sappiamo dai teoremi precedenti che la convergenza della serie è uniforme, dunque il caso interessante è $|z| = r$.

Dunque per ogni $t \in [0, 1[$ si ha $|tz| < r$ e quindi sappiamo che la serie

$$f(t) = \sum_{k=0}^{n} a_k (tz)^k$$

converge assolutamente per ogni $t \in [0,1[$ inoltre per ipotesi sappiamo che converge (e quindi la funzione f è ben definita) anche per ogni $t \in [0,1]$. Vogliamo mostrare che f_n converge a f uniformemente.

Dalla regola di somma per parti dato $q > p$ si ha

$$|f_q(t) - f_p(t)| = \left| \sum_{k=p+1}^{q} (a_k z^k) t^k \right| \leq \sum_{k=p+1}^{q-1} f_k(1)$$

□

Teorema 1.5.5. *Si ha $(\ell^1)^* = \ell^\infty$ e $(c_o)^* = \ell^1$.*

Dimostrazione. La prima affermazione è facile, la seconda segue dal teorema precedente. □

1.6 Funzioni di una variabile reale

Teorema 1.6.1 (degli zeri)**.** *Sia $f\colon [a,b] \to \mathbb{R}$ una funzione continua e supponiamo che $f(a) \leq 0$ e $f(b) \geq 0$. Allora esiste un punto $\xi \in [a,b]$ tale che $f(\xi) = 0$.*

Dimostrazione. Definiamo per induzione due successioni a_k e b_k in modo che valgano le seguenti proprietà:

1. $a_k, b_k \in [a,b]$;
2. $f(a_k) \leq 0$, $f(b_k) \geq 0$;
3. a_k crescente, b_k decrescente;
4. $|a_k - b_k| = |a - b|/2^{k-1}$.

Poniamo $a_1 := a$, $b_1 := b$. Supponendo di aver definito a_k e b_k definiamo ora i punti a_{k+1} e b_{k+1}. Posto $c_k = (a_k + b_k)/2$ distinguiamo due casi.

Se $f(c_k) > 0$ poniamo $a_{k+1} := a_k$ e $b_{k+1} = c_k$. Se invece $f(c_k) \leq 0$ poniamo $a_{k+1} := c_k$ e $b_{k+1} = b_k$. È facile verificare che le successioni, così definite, soddisfano le proprietà richieste.

Essendo monotone le due successioni convergono, $a_k \to \xi$, $b_k \to \eta$, ed essendo $|a_k - b_k| \to 0$ si ha $\xi = \eta$. Inoltre, essendo f continua si ha

$$\lim_{k\to\infty} f(a_k) = f(\xi) = \lim_{k\to\infty} f(b_k).$$

D'altra parte essendo $f(a_k) \leq 0$ e $f(b_k) \geq 0$ si deve avere contemporaneamente $f(\xi) \leq 0$ e $f(\xi) \geq 0$ e quindi $f(\xi) = 0$. □

Teorema 1.6.2 (dei valori intermedi)**.** *Sia $I \subset \mathbb{R}$ un intervallo e sia $f\colon I \to \mathbb{R}$ una funzione continua. Allora $f(I) \subset \mathbb{R}$ è un intervallo. Ossia: l'immagine di un intervallo tramite una funzione continua, è un intervallo.*

Dimostrazione. Siano y_1, y_2 due punti qualunque di $f(I)$ e sia $y \in [y_1, y_2]$. Dobbiamo dimostrare che $y \in f(I)$. Siano x_1 e x_2 i punti di I tali che $f(x_1) = y_1$ e $f(x_2) = y_2$. Senza perdere di generalità possiamo supporre che sia $x_1 \leq x_2$ (altrimenti scambiamo i due punti dati inizialmente) e che sia $y_1 \leq y_2$ (altrimenti consideriamo $-f$ al posto di f). Consideriamo dunque la funzione $g(x) = f(x) - y$. Essendo $y_1 \leq y \leq y_2$ si ottiene $f(x_1) \leq 0$ e $f(x_2) \geq 0$. Applicando dunque il Teorema degli zeri otteniamo che esiste un punto $x \in [x_1, x_2]$ tale che $g(x) = 0$ e quindi $f(x) = y$. Dunque $y = f(x) \in f(I)$ come volevamo dimostrare. □

Esempio 1.6.3. Vogliamo dimostrare che l'equazione $x^n = \alpha$ ha almeno una soluzione non negativa per ogni n naturale e per ogni $\alpha \geq 0$. Si consideri la funzione $f(x) = x^n$ sull'intervallo $I = [0, +\infty)$. Notiamo che $f(0) = 0$ e che $\lim_{x\to\infty} f(x) = +\infty$. Dunque $\sup f(I) = +\infty$. Siccome f è continua, sappiamo che $f(I)$ è un intervallo, e visto che $0 \in f(I)$ e che $+\infty = \sup f(I)$ deduciamo che $f(I) = [0, +\infty)$. In particolare $\alpha \in f(I)$ e quindi esiste $\bar{x} \geq 0$ tale che $f(\bar{x}) = \alpha$ ovvero $\bar{x}^2 = \alpha$.

Esempio 1.6.4. Sia $p(x)$ un qualunque polinomio di grado dispari. Supponiamo che il coefficiente del termine di grado massimo sia positivo. Allora si ha

$$\lim_{x\to-\infty} p(x) = -\infty, \qquad \lim_{x\to+\infty} p(x) = +\infty.$$

In particolare deduciamo che $\sup p(R) = +\infty$ e $\inf p(\mathbb{R}) = -\infty$. Visto che $p\colon \mathbb{R} \to \mathbb{R}$ è una funzione continua, sappiamo anche che $p(\mathbb{R})$ è un intervallo. Dunque $p(\mathbb{R}) = \mathbb{R}$ ovvero p è surgettiva. In particolare esiste $\bar{x} \in \mathbb{R}$ tale che $p(\bar{x}) = 0$.

Definizione 1.6.5 (massimo e minimo)**.** *Sia $f\colon X \to \mathbb{R}$ una funzione definita su un insieme X qualunque. Se esiste un punto $\bar{x} \in X$ tale che per ogni $x \in X$ si ha $f(\bar{x}) \geq f(x)$ allora si dice che f* ammette massimo *su X, $\bar{x}$ si dice* punto di massimo *per f su X e $f(\bar{x})$ si dice* valore massimo *di f su X e si indica con* $\max f(X)$ *o* $\max_{x\in X} f(x)$*. Analogamente si definiscono i* punti di minimo *e il* valore minimo *quando $f(\bar{x}) \leq f(x)$ per ogni $x \in X$. Il valore minimo si può indicare con* $\min f(X)$ *o* $\min_{x\in X} f(x)$*.*

Bisogna notare che non tutte le funzioni ammettono massimo o minimo. Ad esempio $f\colon \mathbb{R} \to \mathbb{R}$ con $f(x) = x$ non ammette nè massimo nè minimo. Inoltre quando il valore massimo esiste esso è unico e coincide con l'estremo superiore: $\max f(X) = \sup f(X)$. Analogamente per il minimo: $\min f(X) = \inf f(X)$.

Il punto di massimo (o minimo) invece può anche non essere unico. Ad esempio per una funzione costante ogni punto è sia di massimo che di minimo.

Teorema 1.6.6 (Weierstrass)**.** *Sia $f\colon [a, b] \to \mathbb{R}$ una funzione continua. Allora f ammette massimo e minimo.*

Dimostrazione. Sia $M = \sup f([a, b])$ e $m = \inf f([a, b])$. Per definizione di sup e inf esistono due successioni (x_k) e (y_k) tali che $f(x_k) \to M$ e $f(y_k) \to m$. Per il Teorema di compattezza possiamo estrarre due sottosuccessioni convergenti $(x_{k_j}) \to x$ e $y_{k_j} \to y$. Chiaramente $x, y \in [a, b]$ ed essendo f continua si ha $f(x) = \lim_k f(x_{k_j}) = M$ e $f(y) = \lim_k f(y_{k_j}) = m$. Dunque x e y sono rispettivamente punti di minimo e di massimo. □

Teorema 1.6.7 (Fermat)**.** *Sia $f\colon\ ]a, b[\ \to \mathbb{R}$. Se $x_0 \in\ ]a, b[$ è un minimo (o un massimo) per f e se f è derivabile in x_0, allora $f'(x_0) = 0$.*

Dimostrazione. Supponiamo che $f(x_0)$ sia un minimo (se è un massimo la dimostrazione è analoga). Allora $f(x) \geq f(x_0)$ per ogni $x \in\ ]a, b[$. Posto $m(h) = \frac{f(x_0+h)-f(x_0)}{h}$ si ha dunque che $m(h) \geq 0$ se $h > 0$ e $m(h) \leq 0$ se $h < 0$. Siccome f è derivabile in x_0 la derivata destra e sinistra in x_0 esistono e sono uguali:

$$f'(x_0) = \lim_{h\to0^+} m(h) \geq 0, \quad f'(x_0) = \lim_{h\to0^-} m(h) \leq 0.$$

Dunque si ottiene $f'(x_0) = 0$. □

Teorema 1.6.8 (Rolle)**.** *Sia $f\colon [a, b] \to \mathbb{R}$ una funzione continua su $[a, b]$ e derivabile su $]a, b[$. Se $f(a) = f(b)$ allora esiste un punto $\xi \in\ ]a, b[$ tale che $f'(\xi) = 0$.*

Dimostrazione. Per il Teorema di Weierstrass f ammette punto di minimo x_0 su $[a, b]$. Se $x_0 \in]a, b[$ allora per il Teorema di Fermat si conclude $f'(x_0) = 0$. Se invece $x_0 \in \{a, b\}$ scegliamo un punto di massimo x_1. Se $x_1 \in]a, b[$ abbiamo concluso altrimenti, essendo $f(a) = f(b)$ si ottiene $f(x_0) = f(x_1)$. In tal caso, essendo il valore massimo uguale al minimo necessariamente f è costante e quindi $f'(x) = 0$ per ogni $x \in]a, b[$. □

Teorema 1.6.9 (Lagrange)**.** *Sia $f : [a, b] \to \mathbb{R}$ una funzione continua su $[a, b]$ e derivabile su $]a, b[$. Allora esiste un punto $\xi \in]a, b[$ tale che*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Dimostrazione. Posto $m = \frac{f(b)-f(a)}{b-a}$ consideriamo la funzione $g(x) = f(x) - mx$. Si verifica facilmente che $g(b) = g(a)$ e chiaramente la funzione g è continua dove f è continua e derivabile dove f è derivabile. Dunque g verifica le ipotesi del Teorema di Rolle e quindi esiste $\xi \in]a, b[$ tale che $g'(\xi) = 0$. Cioè

$$g'(\xi) = f'(\xi) - m = 0$$

che è la tesi del teorema. □

Teorema 1.6.10 (Cauchy)**.** *Siano $f, g : [a, b] \to \mathbb{R}$ due funzioni continue su $[a, b]$ e derivabili su $]a, b[$. Allora esiste un punto $\xi \in]a, b[$ tale che*

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

Se inoltre $g'(x) \neq 0$ per ogni $x \in]a, b[$ si ha necessariamente $g(b) \neq g(a)$ (per il Teorema di Rolle) e possiamo scrivere la relazione precedente come

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Dimostrazione. Consideriamo la funzione $F(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a))$. Si verifica facilmente che $F(b) = F(a)$ e quindi applicando il Teorema di Rolle si ottiene, l'esistenza di un punto ξ tale che $F'(\xi) = 0$. Quest'ultima relazione è proprio il risultato cercato. □

Ci sarà utile, in seguito, avere i precedenti teoremi enunciati in una versione un poco più generale. Si noti infatti che le funzioni risultano definite su un generico intervallo I che può essere anche non limitato. Se I è un intervallo generico, gli estremi dell'intervallo sono $\inf I$ e $\sup I$. A seconda che l'intervallo I sia aperto, chiuso o semi-aperto, gli estremi possono essere o non essere contenuti in I.

Teorema 1.6.11 (Rolle esteso)**.** *Sia $I \subset \mathbb{R}$ un intervallo aperto e non vuoto, e sia $f : I \to \mathbb{R}$ una funzione derivabile su I. Posti $a = \inf I$ e $b = \sup I$ se i seguenti limiti esistono e sono uguali*

$$\lim_{x\to a^+} f(x) = m = \lim_{x\to b^- f(x)}$$

(con $m \in [-\infty, +\infty]$) allora esiste un punto $\xi \in I$ tale che $f'(\xi) = 0$.

Dimostrazione. Se m è finito e $f = m$ identicamente su I allora $f' = 0$ identicamente e il risultato è provato. In caso contrario esisterà un punto $x_0 \in I$ tale che $f(x_0) \neq m$. Supponiamo, per esempio, che $y_0 = f(x_0) > m$ (l'altro caso si tratterà in maniera analoga). Dalla definizione di limite (o meglio, dalla permanenza del segno) sappiamo dunque esistere due punti $a', b' \in I$ con $a < a' < x_0 < b' < b$ tali che

$f(a) < y_0$ e $f(b) < y_0$. Ora per il teorema di Weierstrass sappiamo che f ristretta all'intervallo $[a', b'] \subset I$ ammette massimo in un punto ξ. Essendo x_0 un punto di questo intervallo si avrà dunque $f(\xi) \geq f(x_0) = y_0$ e dunque ξ non può essere nè il punto a' nè il punto b'. In particolare ξ è un punto interno all'intervallo $]a', b'[$ e quindi per il Teorema di Fermat si deve avere $f'(\xi) = 0$. □

Teorema 1.6.12 (Cauchy esteso)**.** *Sia $I \subset \mathbb{R}$ un intervallo aperto e non vuoto e siano $f, g \colon I \to \mathbb{R}$ due funzioni derivabili. Posti $a = \inf I$ e $b = \sup I$, se le funzioni f e g ammettono limite finito agli estremi a e b dell'intervallo I allora esiste un punto $\xi \in I$ tale che*

$$f'(\xi)(g_b - g_a) = g'(\xi)(f_b - f_a)$$

dove f_a, f_b, g_a e g_b sono rispettivamente i limiti delle funzioni f e g nei punti a e b.

Se inoltre $g'(x) \neq 0$ per ogni $x \in\,]a, b[$ si ha necessariamente $g_b \neq g_a$ (per il Teorema di Rolle esteso) e possiamo scrivere la relazione precedente come

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f_b - f_a}{g_b - g_a}.$$

Dimostrazione. Consideriamo la funzione $F(x) = f(x)(g_b - g_a) - g(x)(f_b - f_a)$. Si verifica facilmente che la funzione F verifica le ipotesi del Teorema di Rolle esteso. Dunque sappiamo esistere un punto ξ tale che $F'(\xi) = 0$. Quest'ultima relazione è proprio il risultato cercato. □

Teorema 1.6.13 (L'Hôpital forma "0/0")**.** *Sia $x_0 \in [-\infty, +\infty]$ e I un intervallo aperto con x_0 come punto di accumulazione (ovvero $x_0 \in [\inf I, \sup I]$). Siano f e g due funzioni reali definite e derivabili su $I \setminus \{x_0\}$ con $g'(x) \neq 0$ per ogni $x \in I$. Se $\lim_{x \to x_0} f(x) = 0$, $\lim_{x \to x_0} g(x) = 0$ ed esiste il limite*

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

allora anche il seguente limite esiste e vale

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

Dimostrazione. Dato $x \in I$, dal teorema di Cauchy esteso sappiamo esistere $\xi_x \in\,]x_0, x[$ tale che

$$\frac{f(x)}{g(x)} = \frac{f(x) - f_{x_0}}{g(x) - g_{x_0}} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

Siccome per $x \to x_0$ si ha $\xi_x \to x_0$, il limite seguente esiste e vale

$$\lim_{x \to x_0} \frac{f'(\xi_x)}{g'(\xi_x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

da cui segue il risultato desiderato. □

Notiamo che il precedente teorema è valido anche per limiti destri e sinistri e per limiti all'infinito.

Facciamo alcuni esempi che mettono in evidenza l'importanza di tutte le ipotesi del teorema.

Esempio 1.6.14. Sia $f(x) = x$, $g(x) = x + 1$, $x_0 = 0$. Nonostante che

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = 1$$

si ha

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Esempio 1.6.15. Sia $f(x) = x^2 \sin \frac{1}{x}$, $g(x) = x$ e $x_0 = 0$. Nonostante che il limite

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

non esista, il limite seguente esiste e vale

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Teorema 1.6.16 (monotonia). *Sia $I \subset \mathbb{R}$ un intervallo e sia $f \colon I \to \mathbb{R}$ una funzione continua su I e derivabile su $\mathring{I}$.*

1. *Se per ogni $x \in \mathring{I}$ si ha $f'(x) \geq 0$ allora f è crescente;*
2. *se per ogni $x \in \mathring{I}$ si ha $f'(x) > 0$ allora f è strettamente crescente;*
3. *se per ogni $x \in \mathring{I}$ si ha $f'(x) \leq 0$ allora f è decrescente;*
4. *se per ogni $x \in \mathring{I}$ si ha $f'(x) < 0$ allora f è strettamente decrescente.*

Dimostrazione. Dimostriamo il primo risultato, gli altri sono analoghi. Se per assurdo f non fosse crescente esisterebbero $x, y \in I$, $x < y$ tali che $f(x) > f(y)$. Ma allora per il Teorema di Lagrange troveremmo un punto $\xi \in \,]x, y[\, \subset \mathring{I}$ in cui $f'(\xi) < 0$. □

Si confronti la seguente affermazione con il Teorema dei valori intermedi.

Proposizione 1.6.17 (continuità delle funzioni monotone). *Sia $f \colon I \to \mathbb{R}$ una funzione monotona definita su un intervallo $I \subset \mathbb{R}$. Allora i limiti*

$$\lim_{x \to x_0^+} f(x), \qquad \lim_{x \to x_1^-} f(x)$$

esistono per ogni $x_0 \in [\inf I, \sup I)$ e per ogni $x_1 \in (\inf I, \sup I]$. Inoltre per ogni $x_0 \in I$ si ha

$$\lim_{x \to x_0^-} f(x) \leq f(x_0) \leq \lim_{x \to x_0^+} f(x).$$

e supponiamo che $f(I)$ sia un intervallo. Allora f è continua.

Dimostrazione. Notiamo innanzitutto che una funzione monotona ammette sempre limite destro e limite sinistro in ogni punto. Possiamo supporre, senza perdere generalità, che la funzione f sia monotona crescente. Sia x_0 un punto dell'intervallo I e sia $I^+ = \{x \in I \colon x > x_0\}$. Supponiamo che x_0 non sia l'estremo destro dell'intervallo I cosicchè si ha $I^+ \neq \emptyset$. Allora posto $m = \inf f(I^+)$, per la definizione di estremo inferiore si ha

$$f(y) \geq m, \quad \forall y \in I^+; \qquad \forall \varepsilon > 0 \, \exists x \in I^+ \colon f(x) < m + \varepsilon.$$

Essendo però f crescente, dato $y \in (x_0, x)$ si ha $f(y) < f(x)$ e dunque ponendo $\delta = x - x_0$, si ottiene

$$\forall \varepsilon > 0 \, \exists \delta > 0 \forall y \in (x_0, x_0 + \delta) \colon f(y) \in [m, m + \varepsilon]$$

che è esattamente la proprietà che definisce

$$\lim_{x \to x_0^+} f(x) = m.$$

Analogamente si può dimostrare che esiste il limite sinistro in ogni punto $x_0 \in I$, (tranne che nel caso in cui x_0 è l'estremo sinistro dell'intervallo I).

Questo già ci dice che la funzione è continua negli estremi dell'intervallo (se l'intervallo contiene gli estremi). Preso un punto x_0 interno all'intervallo I dobbiamo invece dimostrare che il limite destro m e il limite sinistro M coincidono. Essendo f crescente si nota però che vale $f(x) \geq m$ per ogni $x > x_0$ e $f(x) \leq M$ per ogni $x < x_0$ ovvero $f(I) \subset (-\infty, M] \cup \{f(x_0)\} \cup [m, +\infty)$. Chiaramente deve essere $M \leq m$ in quanto la funzione f è crescente. D'altra parte sappiamo per ipotesi che $f(I)$ è un intervallo e quindi deve contenere interamente l'intervallo $[M, m]$ e questo può succedere solo se $M = m$. □

Proposizione 1.6.18 (continuità della funzione inversa)**.** *Sia I un intervallo di $\mathbb{R}$ e sia $f\colon I \to \mathbb{R}$ una funzione strettamente monotona e continua. Allora posto $J = f(I)$, J è un intervallo, la funzione $f\colon I \to J$ è invertibile e la funzione inversa $f^{-1}\colon J \to I$ è strettamente monotona e continua.*

Dimostrazione. Il fatto che $J = f(I)$ sia un intervallo è garantito dal Teorema dei valori intermedi. Per definizione di J si ha che $f\colon I \to J$ è surgettiva, inoltre f è iniettiva in quanto è strettamente monotona. Dunque esiste la funzione inversa $f^{-1}\colon J \to I$. Sensa perdere di generalità supponiamo ora che f sia strettamente crescente (in caso contrario possiamo applicare il teorema a $-f$). Vogliamo ora mostrare che f^{-1} è strettamente crescente. Dati $x, y \in J$ sia $a = f^{-1}(x)$ e sia $b = f^{-1}(y)$ cosicchè $f(a) = x$ e $f(b) = y$. Se $x < y$ allora necessariamente $a < b$ in quanto f è strettamente crescente. Dunque abbiamo mostrato che se $x < y$ allora $f^{-1}(x) = a < b = f^{-1}(y)$ cioè f^{-1} è pure strettamente crescente.

□

Teorema 1.6.19. *Sia $f\colon\ ]x_0 - \varepsilon, x_0 + \varepsilon[\ \to \mathbb{R}$ continua in x_0 e derivabile in ogni punto $x \neq x_0$. Se esiste $\lim_{x \to x_0} f'(x) = m$ allora f è derivabile in x_0 e vale $f'(x_0) = m$.*

Dimostrazione. Siccome f è continua in x_0 si ha $\lim_{h \to 0} f(x_0 + h) - f(x_0) = 0$ e dunque applicando l'Hôpital si ottiene

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0} \frac{f'(x_0 + h)}{1} = m.$$

□

1.7 Uniforme continuità

1.7.1 Funzioni uniformemente continue, Lipschitziane, α-Hölderiane

Definizione 1.7.1. *Una funzione $f\colon A \to \mathbb{R}$*

- *si dice* continua *se*

$$\forall x \in A\ \forall \varepsilon > 0\ \exists \delta\ \forall y \in A \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

- *si dice* uniformemente continua *se*

$$\forall \varepsilon > 0\ \exists \delta\ \forall x, y \in A \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

- *si dice* Lipschitziana *se esiste una costante $L > 0$ tale che*

$$\forall x, y \in A \quad |f(x) - f(y)| \leq L|x - y|$$

- *si dice* α-Hölderiana *(con $\alpha > 0$) se esiste una costante $C > 0$ tale che*

$$\forall x, y \in A \quad |f(x) - f(y)| \le C|x - y|^\alpha$$

Chiaramente ogni funzione uniformemente continua è anche continua. Si noti che 1-Hölderiano e Lipschitziano sono sinonimi. Notiamo anche che se f è α-Hölderiana con $\alpha > 1$ si ha

$$\frac{|f(x) - f(x_0)|}{|x - x_0|} \le C|x - x_0|^{\alpha-1} \to 0 \quad \text{per } x \to x_0$$

e quindi $f'(x_0) = 0$ per ogni $x_0 \in A$. Dunque se A è un intervallo allora f è costante. Per questo le funzioni α-Hölderiane sono interessanti solo quando $\alpha < 1$.

Teorema 1.7.2. *Se $f\colon [a, b] \to \mathbb{R}$ è α-Hölderiana per un certo $\alpha > 0$ allora f è β-Hölderiana per ogni $\beta \le \alpha$. In particolare se f è Lipschitziana allora f è α-Hölderiana per ogni $\alpha < 1$.*

Dimostrazione. Supponiamo dunque che valga

$$|f(x) - f(y)| \le C|x - y|^\alpha \quad \forall x, y \in [a, b].$$

Dato $\beta \le \alpha$ poniamo $C' = C|b - a|^{\alpha-\beta}$. Allora se $x, y \in [a, b]$ si ha

$$C|x - y|^\alpha \le C'|x - y|^\beta$$

e dunque f è β-Hölderiana. □

Teorema 1.7.3. *Per verificare che una funzione $f\colon A \to \mathbb{R}$ è uniformemente continua è sufficiente mostrare che esiste una funzione $G\colon \mathbb{R}^+ \to \mathbb{R}$ che*

- $\forall x, y \in A \quad |f(x) - f(y)| \le G(|x - y|)$*;*
- $\lim_{t \to 0^+} G(t) = 0$*.*

Dimostrazione. Essendo $G(t) \to 0$ si ha

$$\forall \varepsilon > 0\ \exists \delta \quad t < \delta \Rightarrow G(t) < \varepsilon$$

dunque

$$\forall \varepsilon > 0\ \exists \delta\ \forall x, y \in A \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| \le G(|x - y|) < \varepsilon$$

cioè f è uniformemente continua. □

Se ne deduce facilmente che ogni funzione Lipschitziana o α-Hölderiana è uniformemente continua.

Teorema 1.7.4. *Per verificare che una funzione $f\colon A \to \mathbb{R}$ non è uniformemente continua è sufficiente trovare due successioni $x_k, y_k \in A$ tali che*

$$\lim_{k\to\infty} |x_k - y_k| = 0$$

ma

$$\lim_{k\to\infty} |f(x_k) - f(y_k)| > 0.$$

Dimostrazione. Siccome $|f(x_k) - f(y_k)|$ è una successione positiva ma non infinitesima possiamo trovare un $\varepsilon > 0$ e una sottosuccessione indicizzata tramite k_j tale che $|f(x_{k_j}) - f(y_{k_j})| \geq \varepsilon$ per ogni j. D'altra parte essendo $|x_k - y_k| \to 0$ si ha

$$\forall \delta > 0\ \exists k \quad |x_k - y_k| < \delta$$

e quindi per ogni $\delta > 0$ è possibile trovare k_j tale che posto $x = x_{k_j}$ e $y = y_{k_j}$ si abbia $|f(x) - f(y)| \geq \varepsilon$ mentre $|x - y| < \delta$. Dunque f verifica la negazione della proprietà di uniforme continuità

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x, y \in A \quad |x - y| < \delta \text{ ma } |f(x) - f(y)| \geq \varepsilon.$$

□

Esercizi

1. Mostrare che $f(x) = x^2$ non è uniformemente continua.
2. Mostrare che $f(x) = 1/x$ non è uniformemente continua.
3. Mostrare che $f(x) = \sin(1/x)$ non è uniformemente continua.
4. Mostrare che $f(x) = |x|$ è Lipschitziana.
5. Mostrare che $f(x) = \sqrt{|x|}$ è 1/2-Hölderiana ma non Lipschitziana.
6. Trovare una funzione uniformemente continua ma non α-Hölderiana per alcun $\alpha > 0$.

1.7.2 Successioni di Cauchy

Definizione 1.7.5. *Una successione x_k si dice* di Cauchy *se vale la seguente proprietà*

$$\forall \varepsilon > 0\ \exists N\ \forall k, j \geq N \quad |x_k - x_j| < \varepsilon$$

Teorema 1.7.6. *Ogni successione convergente è di Cauchy.*

Dimostrazione. Sia $x_k \to x$ una successione convergente. Dunque vale

$$\forall \varepsilon > 0\ \exists N\ \forall k \geq N \quad |x_k - x| < \varepsilon$$

e quindi presi $k, j \geq N$ si ha

$$|x_k - x_j| \leq |x_k - x| + |x - x_j| < 2\varepsilon$$

cioè

$$\forall \varepsilon > 0\ \exists N\ \forall k, j \geq N \quad |x_k - x_j| < 2\varepsilon$$

che è equivalente alla definizione delle successioni di Cauchy. □

Teorema 1.7.7. *Ogni successione di Cauchy converge.*

Dimostrazione. Data una successione di Cauchy x_k definiamo

$$A = \{x \in \mathbb{R} \colon \exists N\ \forall k \geq N \quad x_k \leq x\}.$$

Chiaramente se $x \in A$ e $y \geq x$ anche $y \in A$. Dunque A è un intervallo illimitato superiormente. Se A fosse anche illimitato inferiormente (cioè $A = \mathbb{R}$) si avrebbe

$$\forall M\ \exists N\ \forall k \geq N \quad x_k \leq -M$$

che significa $x_k \to -\infty$. Questo significa in particolare che dato qualunque $\varepsilon > 0$, qualunque N e qualunque $j \geq N$ esiste $k \geq N$ tale che $x_k < x_j - \varepsilon$ cioè $|x_k - x_j| > \varepsilon$. Ma questo va contro l'ipotesi che x_k sia di Cauchy.

Dunque A è inferiormente limitato e quindi ammette estremo inferiore. Posto $a = \inf A$ si ha

$$\forall \varepsilon > 0\ \exists x \in A \quad a + \varepsilon > x$$

e quindi $a + \varepsilon \in A$ per ogni $\varepsilon > 0$. Inoltre per ogni $\varepsilon > 0$ si ha $a - \varepsilon \notin A$. Da $a + \varepsilon \in A$ ricaviamo che

$$\exists N_1\ \forall k \geq N_1 \quad x_k \leq a + \varepsilon$$

e da $a - \varepsilon \notin A$ troviamo

$$\forall N\ \exists j \geq N \quad x_j \geq a - \varepsilon$$

dalla proprietà di Cauchy si ha invece

$$\exists N_2\ \forall k, j \geq N_2 \quad x_j < x_k + \varepsilon.$$

Mettendo assieme queste proprietà e scegliendo $N = \max\{N_1, N_2\}$ si ottiene

$$\forall \varepsilon > 0\ \exists N\ \exists j \geq N\ \forall k \geq N \quad a - \varepsilon < x_j < x_k + \varepsilon \leq a + 2\varepsilon$$

da cui

$$\forall \varepsilon > 0\ \exists N\ \forall k \geq N a - 2\varepsilon < x_k \leq a + \varepsilon$$

che significa

$$\lim_{k \to \infty} x_k = a.$$

□

Teorema 1.7.8. *Se $f\colon A \to \mathbb{R}$ è uniformemente continua e x_k è una successione di Cauchy in A allora la successione $y_k = f(x_k)$ è di Cauchy.*

Dimostrazione. Sia x_k una successione di Cauchy e sia f uniformemente continua. Dall'uniforme continuità di f

$$\forall \varepsilon > 0\ \exists \delta \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

troviamo

$$\forall \varepsilon > 0\ \exists \delta_\varepsilon \quad |x_k - x_j| < \delta \Rightarrow |y_k - y_j| < \varepsilon.$$

Essendo x_k di Cauchy si ha

$$\forall \varepsilon > 0\ \exists N\ \forall k, j \geq N \quad |x_k - x_j| < \delta_\varepsilon$$

da cui

$$\forall \varepsilon > 0\ \exists N\ \forall k, j \geq N \quad |y_k - y_j| < \varepsilon$$

cioè y_k è di Cauchy. □

1.7.3 Estensioni

Teorema 1.7.9. *Sia $f\colon\ ]a, b]$ una funzione uniformemente continua. Allora esiste finito il limite*

$$\lim_{x \to a^+} f(x).$$

Dimostrazione. Sia $x_k > a$ una qualunque successione convergente ad a. Allora x_k è una successione di Cauchy e quindi $f(x_k)$ è a sua volta una successione di Cauchy che quindi converge ad un certo valore l. D'altra parte il valore l trovato non dipende dalla successione scelta. Infatti prese x_k e y_k due diverse successioni convergenti ad a, sappiamo che $f(x_k) \to l_1$ e $f(y_k) \to l_2$. Ma allora anche la successione

$$z_k = \begin{cases} x_k & \text{per } k \text{ pari} \\ y_k & \text{per } k \text{ dispari} \end{cases}$$

converge ad a e quindi anche $f(z_k)$ converge. Ma i termini pari di $f(z_k)$ convergono a l_1 e i termini dispari convergono a l_2. Dunque per l'unicità del limite $l_1 = l_2$. In conclusione esiste $l \in \mathbb{R}$ tale che per qualunque successione $x_k \to a^+$ si ha $f(x_k) \to l$ e quindi

$$\lim_{x \to a^+} f(x) = l.$$

□

Notiamo dunque che una funzione uniformemente continua $f\colon\ ]a, b] \to \mathbb{R}$ può essere con continuità ad una funzione continua $\bar{f}\colon [a, b] \to \mathbb{R}$. D'altra parte la funzione $\bar{f}$ risulta essere anche uniformemente continua per il teorema di Cantor (o per verifica diretta).

Teorema 1.7.10. *Sia $A \subset \mathbb{R}$ e $f\colon A \to \mathbb{R}$ una funzione Lipschitziana. Allora è possibile estendere f ad una funzione lipschitziana $\bar{f}\colon \mathbb{R} \to \mathbb{R}$ mantenendo la stessa costante di Lipschitz L.*

1.7.4 Modulo di continuità

Definizione 1.7.11. *Data una funzione $f\colon A \to \mathbb{R}$, chiamiamo* modulo di continuità *di f la funzione $M\colon \mathbb{R}^+ \to \mathbb{R} \cup \{+\infty\}$ definita da*

$$M(t) = \sup\{|f(x) - f(y)|\colon x, y \in A,\ |x - y| \le t\}$$

Dalla definizione si verifica facilmente che $M(t)$ è una funzione crescente. Infatti se l'insieme di cui si calcola l'estremo superiore diventa più grande (rispetto all'inclusione di insiemi) al crescere di t e quindi anche l'estremo superiore cresce al crescere di t.

Teorema 1.7.12. *Se $f\colon A \to \mathbb{R}$ è una funzione qualunque e M è il suo modulo di continuità allora:*

- *f è uniformemente continua se e solo se $M(t) \to 0$ per $t \to 0^+$;*
- *f è lipschitziana se esiste $L > 0$ tale che $M(t) \le Lt$;*
- *f è α-Hölderiana se esiste $C > 0$ tale che $M(t) \le Ct^\alpha$.*

Dimostrazione. Innanzitutto si verifica direttamente dalla definizione di M che si ha la seguente equivalenza

$$\begin{gathered} M(\delta) \le \varepsilon \\ \Updownarrow \\ \forall x, y \in A \quad |x - y| \le \delta \Rightarrow |f(x) - f(y)| \le \varepsilon \end{gathered}$$

e dalla definizione di limite (ricordando che $M(t)$ è positiva e crescente) si ha anche

$$\lim_{t \to 0^+} M(t) = 0 \quad \Leftrightarrow \quad \forall \varepsilon > 0\ \exists \delta\ M(\delta) \le \varepsilon.$$

Mettendo assieme le due precedenti equivalenze si ottiene

$$\forall \varepsilon > 0\ \exists \delta\ \forall x, y \in A \quad |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon$$

che è equivalente alla definizione di uniforme continuità per f.

Chiaramente se M è il modulo di continuità di f allora si ha

$$|f(x) - f(y)| \leq M(|x - y|) \quad \forall x, y \in A.$$

Dunque se $M(t) \leq Lt$ si trova che f è lipschitziana e se $M(t) \leq Lt^\alpha$ si trova che f è α-Hölderiana.

D'altro canto se f è Lipschitziana si ha

$$M(t) = \sup_{|x-y|\leq t} |f(x) - f(y)| \leq \sup_{|x-y|\leq t} L|x - y| \leq Lt.$$

Risultato analogo si ottiene per le funzione α-Hölderiane.

□

1.8 Funzioni convesse

Definizione 1.8.1. *Sia V uno spazio vettoriale e sia $A \subset V$. Diciamo che A è un insieme convesso se vale*

$$tx + (1 - t)y \in A \qquad \forall x, y \in A \forall t \in [0, 1].$$

Si noti che il segmento di estremi x e y si scrive nella forma

$$[x, y] := \{tx + (1 - t)y \colon t \in [0, 1]\}$$

dunque possiamo affermare che un insieme è convesso se contiene tutti i segmenti i cui estremi stanno nell'insieme stesso.

Si noti anche che nel caso $V = \mathbb{R}$, un insieme $A \subset \mathbb{R}$ è convesso se e solo se è un intervallo.

Definizione 1.8.2. *Sia V uno spazio vettoriale e sia $f \colon A \to \mathbb{R}$ una funzione definita su un sottoinsieme convesso A di V. Diciamo allora che f è convessa se*

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) \qquad \forall x, y \in A, \forall t \in [0, 1].$$

Lemma 1.8.3. *Sia $f \colon I \to \mathbb{R}$ con I intervallo di $\mathbb{R}$ una funzione convessa. Allora il rapporto incrementale*

$$R(x, y) := \frac{f(y) - f(x)}{y - x}$$

è una funzione crescente sia in x che in y.

Dimostrazione. Notiamo che $R(x, y) = R(y, x)$ e dunque è sufficiente mostrare che fissato x la funzione $y \mapsto R(x, y)$ è crescente. Mostriamo innanzitutto che se $y_2 > y_1 > x$ si ha $R(x, y_2) \geq R(x, y_1)$. Per far questo scegliamo

$$t = \frac{y_1 - x}{y_2 - x}$$

in modo che $t \in [0, 1]$ e $y_1 = tx + (1 - t)y_2$. Applicando la definizione di convessità si ottiene dunque

$$f(y_1) \leq \frac{y_1 - x}{y_2 - x} f(y_2) + \frac{y_2 - y_1}{y_2 - x} f(x)$$

sottraendo ad ambo i membri $f(x)$ si ottiene

$$f(y_1) - f(x) \leq \frac{y_1 - x}{y_2 - x} f(y_2) - \frac{y_1 - x}{y_2 - x} f(x)$$

e dividendo per $y_1 - x$ si conclude

$$\frac{f(y_1) - f(x)}{y_1 - x} \leq \frac{f(y_2) - f(x)}{y_2 - x}$$

cioè $R(x, y_1) \leq R(x, y_2)$. $\square$

In particolare per ogni $x \in I$ il rapporto incrementale

$$h \mapsto \frac{f(x+h) - f(x)}{h}$$

è una funzione crescente e dunque (se x ha un intorno destro in I) esiste la derivata destra

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h}$$

e analogamente esiste la derivata sinistra se x ha un intorno sinistro in I.

Sia ora X uno spazio vettoriale reale e $f \colon X \to \mathbb{R}$ una funzione convessa.

Allora, si vede che per ogni $v \in X$ e $x \in X$ la funzione $t \mapsto f(x + tv)$ è una funzione convessa da $\mathbb{R} \to \mathbb{R}$.

Dunque per il Lemma precedente esistono le derivate direzionali

$$\frac{\partial f}{\partial v}(x) := \lim_{h \to 0^+} \frac{f(x + hv) - f(x)}{h}$$

in ogni punto $x \in X$ e ogni direzione $v \in X$.

Notiamo inoltre che si avrà (in quanto il rapporto incrementale è crescente):

$$f(x + v) \geq f(x) + \frac{\partial f}{\partial v}(x).$$

Definiamo allora il *sottodifferenziale* di f in $x \in X$ come l'insieme

$$\partial^- f(x) := \{\xi \in X^* \colon f(x + v) \geq f(x) + \langle \xi, v \rangle\}.$$

Notiamo che questo insieme è convesso.

Capitolo 2

Topologia

2.1 Spazi topologici

Definizione 2.1.1 (spazio topologico). *Sia X un insieme e $\tau \subset 2^X$ una famiglia di sottoinsiemi di X. Diremo che (X, τ) è uno* spazio topologico *o che τ è una* topologia su X *se valgono le seguenti proprietà:*

1. *$\emptyset \in \tau, X \in \tau$;*
2. *se $A, B \in \tau$ allora $A \cap B \in \tau$;*
3. *se I è un qualunque insieme di indici e $A_i \in \tau$ per ogni $i \in I$ allora $\bigcup_{i \in I} A_i \in \tau$.*

Definizione 2.1.2 (aperto, chiuso). *Sia (X, τ) uno spazio topologico. Dato $A \subset X$ diremo che A è* aperto *in X (rispetto alla topologia τ) se $A \in \tau$ e diremo che A è* chiuso *in X (rispetto alla topologia τ) se $X \setminus A \in \tau$.*

Dalla definizione di topologica seguono le seguenti proprietà per i chiusi di uno spazio topologico (X, τ):

1. $\emptyset$ e X sono chiusi;
2. se A e B sono chiusi allora $A \cup B$ è chiuso;
3. se I è un qualunque insieme di indici e A_i è chiuso per ogni $i \in I$ allora $\bigcap_{i \in I} A_i$ è chiuso.

Ricordiamo infatti che $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ e $X \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \setminus A_i)$.

Definizione 2.1.3 (intorno). *Sia (X, τ) uno spazio topologico. Dati $U \subset X$ e $x \in X$ diremo che U è un* intorno *di x se U contiene un aperto che contiene x (cioè esiste $A \in \tau$ tale che $x \in A \subset U$).*

Notiamo che $A \subset X$ è aperto se e solo è un intorno di ogni suo punto (A è intorno di x per ogni $x \in A$).

Definizione 2.1.4 (punto interno, parte interna). *Sia (X, τ) uno spazio topologico, $A \subset X$ e $x \in X$. Si dice che x è un* punto interno *ad A se A è un intorno di x (ovvero esiste $U \in \tau$ tale che $x \in U \subset A$).*

Dato $A \subset X$ chiamiamo parte interna *di A l'insieme $\overset{\circ}{A}$ dei punti interni ad A:*

$$\overset{\circ}{A} := \{x \in X \colon \exists U \in \tau,\ x \in U \subset A\}.$$

Definizione 2.1.5 (punto aderente, chiusura). *Sia X uno spazio topologico, $A \subset X$ e $x \in X$. Si dice che x è un* punto aderente *ad A (o* punto di aderenza*) se ogni intorno di x interseca A (ovvero per ogni $U \in \tau$ si ha $x \in U \Rightarrow U \cap A \neq \emptyset$).*

Dato $A \subset X$ chiamiamo chiusura *di A l'insieme $\overline{A}$ dei punti aderenti ad A:*

$$\overline{A} := \{x \in X \colon \forall U \in \tau,\ x \in U \Rightarrow U \cap A \neq \emptyset\}.$$

Definizione 2.1.6 (bordo). *Sia X uno spazio topologico, $A \subset X$. Definiamo il* bordo *come l'insieme $\partial A = \overline{A} \setminus \overset{\circ}{A}$ dei punti aderenti ma non interni ad A.*

Definizione 2.1.7 (denso). *Sia X uno spazio topologico, $A \subset X$. Diciamo che A è* denso *in X se $\overline{A} = X$.*

Definizione 2.1.8 (punti isolati, punti di accumulazione). *Sia (X, τ) uno spazio topologico, $A \subset X$ e $x \in X$. Si dice che x è un* punto isolato *di A se $x \in A$ e se esiste un intorno di x che non contiene altri punti di A al di fuori di x (ovvero esiste $U \in \tau$ tale che $x \in U$ e $U \cap A = \{x\}$).*

Si dice che x è un punto di accumulazione *per A se x è un punto di aderenza per $A \setminus \{x\}$ (ovvero per ogni $U \in \tau$, se $x \in U$ allora $(A \setminus \{x\}) \cap U \neq \emptyset$). Equivalentemente x è un punto di accumulazione per A se x è un punto di aderenza ma non un punto isolato di A.*

Teorema 2.1.9 (parte interna, chiusura, bordo). *Sia (X, τ) uno spazio topologico e $A \subset X$. Allora*

1. *$X \setminus \overset{\circ}{A} = \overline{X \setminus A}$ e $X \setminus \overline{A} = \overset{\circ}{\overline{X \setminus A}}$;*
2. *la parte interna di A è un aperto, anzi è il più grande aperto contenuto in A;*
3. *la chiusura di A è un chiuso, anzi è il più piccolo chiuso contenente A;*
4. *il bordo di A è un chiuso.*

Dimostrazione.

1. Sia $x \in X \setminus \overset{\circ}{A}$. Allora dovrà esistere un aperto $U \in \tau$ tale che $x \in U$ ma $U \cap A = \emptyset$. Dunque $U \subset X \setminus A$ e cioè x è un punto aderente di $X \setminus A$. D'altra parte dato $x \in \overline{X \setminus A}$ se $U \in \tau$ è un qualunque aperto tale che $x \in U$ si ha $U \cap (X \setminus A) \neq \emptyset$. Dunque non può essere $U \subset A$ e quindi x non è un punto interno di A. Abbiamo quindi mostrato la prima uguaglianza.

   La seconda uguaglianza segue direttamente dalla prima sostituendo A con $X \setminus A$.

2. Siccome i punti di $\overset{\circ}{A}$ sono tutti i punti interni ad A, si ha che per ogni $x \in \overset{\circ}{A}$ esiste $U_x \in \tau$ tale che $x \in U \subset A$. Naturalmente dato $y \in U_x$ essendo $y \in U_x \subset A$ troviamo che anche y è un punto interno ad A e quindi $U_x \subset \overset{\circ}{A}$. Dunque $\overset{\circ}{A} = \bigcup_{x \in \overset{\circ}{A}} U_x$ è unione di aperti ed è quindi a sua volta aperto. Abbiamo anche visto che se U è un aperto contenuto in A allora essendo tutti i punti di U interni ad A si ha $U \subset \overset{\circ}{A}$.

3. Segue dai due punti precedenti. Infatti $\overline{A}$ è il complementare della parte interna di $X \setminus A$ dunque essendo il complementare di un aperto è un chiuso. Inoltre dato un qualunque chiuso C che contiene A il complementare di C è un aperto contenuto nel complementare di A e dunque è contenuto nella parte interna del complementare di A. Quindi $C \supset \overline{A}$.

4. Essendo $\partial A = \overline{A} \cap \overline{(X \setminus A)}$ l'intersezione di due chiusi, è esso stesso chiuso. $\square$

Definizione 2.1.10 (base della topologia)**.** *Sia (X,τ) uno spazio topologico. Si dice che una famiglia di aperti σ ($\sigma \subset \tau$) è una* base *per la topologia τ se per ogni aperto $A \in \tau$ esiste una famiglia $\{A_i\}$ di aperti di σ tale che $A = \bigcup_i A_i$.*

Definizione 2.1.11 (base di intorni)**.** *Sia (X,τ) uno spazio topologico e sia $x \in X$. Una famiglia $\mathcal{U}$ di intorni di x viene chiamata* base di intorni *se contiene intorni "arbitrariamente piccoli" cioè se dato un qualunque intorno U di x esiste un intorno $V \in \mathcal{U}$ tale che $V \subset U$.*

La conoscenza di una base di intorni per ogni punto di uno spazio topologico garantisce la possibilità di ricostruire l'intera topologia dello spazio. Infatti un insieme $A \subset X$ è aperto se e solo se per ogni punto $x \in X$ esiste un elemento U della base di intorni $\mathcal{U}_x$ di x tale che $U \subset A$.

Teorema 2.1.12 (base di intorni)**.** *Sia X un insieme e per ogni $x \in X$ sia $\mathcal{U}_x$ una famiglia di sottoinsiemi di X con le seguenti proprietà:*

1. *per ogni $x \in X$ la famiglia $\mathcal{U}_x$ non è vuota e per ogni $U \in \mathcal{U}_x$ si ha $x \in U$;*
2. *se $U, V \in \mathcal{U}_x$ allora esiste $W \in \mathcal{U}_x$ tale che $W \subset U \cap V$.*

Posto allora $\tau = \{A \subset X \colon \forall x \in A \, \exists U \in \mathcal{U}_x \; U \subset A\}$ si ha che (X,τ) è uno spazio topologico e $\mathcal{U}_x$ è una base di intorni di x per ogni $x \in X$.

Dimostrazione. Chiaramente $\emptyset \in \tau$. Anche $X \in \tau$ in quanto per ogni $x \in X$ si può scegliere un qualunque $U \in \mathcal{U}_x$ e questo è possibile in quanto $\mathcal{U}_x$ è non vuoto.

Siano ora $A, B \in \tau$. Vogliamo mostrare che $A \cap B \in \tau$. Sia quindi dato $x \in A \cap B$. Siccome $x \in A$ e $A \in \tau$ deve esistere $U \in \mathcal{U}_x$ tale che $x \in U \subset A$. Allo stesso modo deve esistere $V \in \mathcal{U}_x$ tale che $x \in V \subset B$. Dalle proprietà di $\mathcal{U}_x$ sappiamo però che esiste $W \in \mathcal{U}_x$ tale che $x \in W \subset U \cap V \subset A \cap B$. Siccome questo è vero per ogni $x \in A \cap B$ ne risulta che $A \cap B \in \tau$.

Ci resta da dimostrare che se $A_i \in \tau$ allora $\bigcup_i A_i \in \tau$. Questo è ovvio in quanto dato $x \in \bigcup_i A_i$ esiste $\bar{i}$ per cui $x \in A_{\bar{i}}$. Essendo poi $A_{\bar{i}} \in \tau$ esiste $U \in \mathcal{U}_x$ tale che $U \subset A_{\bar{i}} \subset \bigcup_i A_i$. □

Definizione 2.1.13 (separabile, base numerabile)**.** *Uno spazio topologico (X,τ) si dice* separabile *se esiste un sottoinsieme $D \subset X$ denso e numerabile.*

Uno spazio topologico (X,τ) soddisfa il primo assioma di numerabilità *se ogni punto $x \in X$ ammette una base numerabile degli intorni $\mathcal{U}(x)$.*

Uno spazio topologico (X,τ) si dice a base numerabile *o che* soddisfa il secondo assioma di numerabilità *se esiste una base numerabile $\mathcal{B}$ della topologia τ.*

Definizione 2.1.14 (prima e seconda categoria)**.** *Un sottoinsieme A di uno spazio topologico X si dice* mai denso *o anche* magro *se la parte interna della sua chiusura è vuota.*

Un sottoinsieme A di uno spazio topologico X si dice di prima categoria *se è unione numerabile di insiemi mai densi. Ovvero se è contenuto nell'unione numerabile di chiusi con parte interna vuota. Un insieme A si dice invece* di seconda categoria *se non è di prima categoria.*

Teorema 2.1.15 (Baire)**.** *Lo spazio topologico $\mathbb{R}$ è di seconda categoria (come sottoinsieme di sè stesso).*

Definizione 2.1.16 (topologia indotta sui sottoinsiemi)**.** *Sia (X,τ) uno spazio topologico e $Y \subset X$. Sia $\sigma \subset 2^Y$ definito come segue:*

$$\sigma := \{A \cap Y \colon A \in \tau\}.$$

Si verifica facilmente che (Y,σ) è uno spazio topologico. Questa topologia viene chiamata topologia indotta *da (X,τ) su Y.*

Definizione 2.1.17 (compatto)**.** *Uno spazio topologico X si dice* compatto *se da ogni ricoprimento aperto è possibile estrarre un sottoricoprimento finito*

Definizione 2.1.18 (connesso)**.** *Uno spazio topologico X si dice* connesso *se gli unici sottoinsiemi di X che sono contemporaneamente aperti e chiusi sono $\emptyset$ e X stesso.*

2.2 Funzioni continue

Definizione 2.2.1 (continuità)**.** *Siano X e Y spazi topologici e sia $f\colon X \to Y$ una funzione. Diremo che f è* continua *se per ogni aperto A in Y l'insieme controimmagine $f^{-1}(A)$ è aperto in X.*

Teorema 2.2.2 (proprietà delle funzioni continue)**.** *Siano X e Y spazi topologici e sia $f\colon X \to Y$ una funzione continua.*

1. *Se X è compatto allora $f(X)$ è compatto;*
2. *se X è sequenzialmente compatto allora $f(X)$ è sequenzialmente compatto;*
3. *se X è connesso allora $f(X)$ è connesso;*
4. *se X è connesso per archi allora $f(X)$ è connesso per archi;*

Dimostrazione.

1. Sia $\mathcal{G}$ un ricoprimento aperto di $f(X)$. Definiamo $\mathcal{F} := \{f^{-1}(G)\colon G \in \mathcal{G}\}$. Siccome f essendo $G \in \mathcal{G}$ aperto troviamo che ogni $F = f^{-1}(G)$ è aperto. Inoltre $\bigcup \mathcal{F} = f^{-1}(\bigcup \mathcal{G}) = X$ e dunque $\mathcal{F}$ è un ricoprimento aperto di X. Essendo X compatto possiamo estrarre un sottoricoprimento finito $\mathcal{F}' \subset \mathcal{F}$. La famiglia corrispondente $\mathcal{G}' := \{f(F)\colon F \in \mathcal{F}'\}$ è una sottofamiglia finita di $\mathcal{F}$ ed è un ricoprimento di $f(X)$ in quanto $\bigcup \mathcal{G}' = f(\bigcup \mathcal{F}') = f(X)$. Dunque da ogni ricoprimento aperto di $f(X)$ possiamo estrarre un sottoricoprimento finito.
2. Sia (y_k) una successione in $f(X)$. Allora esiste (x_k) tale che $f(x_k) = y_k$. Siccome X è sequenzialmente compatto possiamo estrarre una sottosuccessione convergente $x_{k_j} \to x \in X$. Ma essendo f continua abbiamo $y_{k_j} = f(x_{k_j}) \to f(x) \in f(X)$. Dunque abbiamo trovato una sottosuccessione convergente in $f(X)$.
3. Sia $B \subset f(X)$ un insieme aperto e chiuso in $f(X)$. Allora $A := f^{-1}(B)$ è aperto e chiuso in X (in quanto f è continua). Essendo X connesso ne ricaviamo che o $A = \emptyset$ o $A = X$ e quindi o $B = \emptyset$ oppure $B = f(X)$. Dunque $f(X)$ è connesso.
4. Siano a, b due punti qualunque di $f(X)$. Esistono allora $x, y \in X$ tali che $f(x) = a$ e $f(y) = b$. Essendo poi X connesso per archi deve esistere una curva continua γ che congiunge x a y in X. Essendo f continua deduciamo che $f \circ \gamma$ è una curva continua che congiunge a a b e dunque anche $f(X)$ è connesso per archi.

□

Definizione 2.2.3 (continuità in un punto)**.** *Sia $f\colon X \to Y$ una funzione tra due spazi topologici, $x_0 \in X$. Diciamo che f è* continua in x_0 *se per ogni intorno V di $f(x_0)$ esiste un intorno U di x_0 tale che $f(U) \subset V$.*

Proposizione 2.2.4 (continuità)**.** *Una funzione $f\colon X \to Y$ è continua se e solo se essa è continua in ogni punto $x \in X$.*

Dimostrazione. Mostriamo che se f è continua allora è continua in ogni punto. Sia $x_0 \in X$ e sia V un intorno di $f(x_0)$. Allora esiste un aperto V' tale che $f(x_0) \in V' \subset V$. Sappiamo allora che $U = f^{-1}(V')$ è aperto. Chiaramente $x_0 \in U$ e $f(U) = V' \subset V$.

Supponiamo ora che f sia continua in ogni punto $x \in X$. Sia V un qualunque aperto in Y e sia $A = f^{-1}(V)$. Dato $x_0 \in A$, essendo f continua in x_0 ed essendo V un intorno di $f(x_0)$ deve esistere un intorno U di x_0 tale che $f(U) \subset V$. Inoltre $U \subset A$ in quanto $A = f^{-1}(V)$ e $f(U) \subset V$. Dunque $x_0 \in U \subset V$ e quindi x_0 è un punto interno a V. Questo è vero per ogni $x_0 \in V$ e quindi V è aperto. □

Teorema 2.2.5 (continuità in un punto della funzione composta)**.** *Siano X, Y, Z spazi topologici e siano $f: X \to Y$ e $g: Y \to Z$. Se f è continua nel punto $x_0 \in X$ e g è continua nel punto $y_0 = f(x_0)$ allora la funzione composta $g \circ f: X \to Z$ è continua nel punto x_0.*

Dimostrazione. Dato un intorno V di $z_0 = g(f(x_0))$ dobbiamo mostrare l'esistenza di un intorno U di x_0 tale che $g(f(U)) \subset V$. Per la continuità di g in $y_0 = f(x_0)$ deve esistere un intorno W di y_0 tale che $G(W) \subset V$. Per la continuità di f deve esistere un intorno U di x_0 tale che $f(U) \subset W$. Dunque, come voluto, $g(f(U)) \subset g(W) \subset V$. □

Teorema 2.2.6 (continuità della funzione composta)**.** *Siano X, Y, Z spazi topologici e siano $f: X \to Y$ e $g: Y \to Z$ funzioni continue. Allora $g \circ f: X \to Z$ è una funzione continua.*

Dimostrazione. Se A è un aperto in Z allora $g^{-1}(A)$ è aperto in Y essendo g continua. Ma allora anche $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$ è aperto in quanto anche f è continua. □

2.3 Spazi metrici

Definizione 2.3.1 (spazio metrico)**.** *Dato un insieme X, una funzione $d: X \times X \to \mathbb{R}$ si dice che (X, d) è uno* spazio metrico *se per ogni $x, y, z \in X$ valgono le seguenti proprietà:*

1. *$d(x, z) \leq d(x, y) + d(y, z)$ (disuguaglianza triangolare);*
2. *$d(x, y) = d(y, x)$ (simmetria);*
3. *$d(x, y) = 0 \Leftrightarrow x = y$.*

La funzione d viene chiamata distanza.

Sia (X, d) uno spazio metrico. Allora, per ogni $x, y \in X$ si ha

$$0 = d(x, x)/2 \leq (d(x, y) + d(y, x))/2 = d(x, y)$$

da cui ricaviamo che $d(x, y) \geq 0$.

Definizione 2.3.2 (palla, disco, sfera)**.** *Se (X, d) è uno spazio metrico, $x \in X$ e $\rho > 0$, definiamo rispettivamente la* palla, il disco e la sfera centrati in x e di raggio *ρ gli insiemi*

$$\begin{aligned} \mathbb{B}_\rho(x) &= \{y \in X: d(y, x) < \rho\}, \\ \mathbb{D}_\rho(x) &= \{y \in X: d(y, x) \leq \rho\}, \\ \mathbb{S}_\rho(x) &= \{y \in X: d(y, x) = \rho\}. \end{aligned}$$

Su uno spazio metrico (X, d) si definisce in maniera canonica una topologia. Questa topologia è la più piccola topologia che rende aperte tutte le palle di X. Tutte le definizioni e le proprietà date per gli spazi topologici si estendono quindi naturalmente agli spazi metrici.

Teorema 2.3.3 (topologia indotta da una metrica)**.** *Sia (X, d) uno spazio metrico e sia τ la famiglia di tutti i sottoinsiemi A di X tali che*

$$\forall x \in A \exists \rho > 0 \colon \mathbb{B}_\rho(x) \subset A.$$

Allora (X, τ) è uno spazio topologico.

Si noti anche che dato $x \in X$ la famiglia $\mathcal{U}_x = \{\mathbb{B}_\rho(x) \colon \rho > 0\}$ è una base di intorni di x.

Definizione 2.3.4. *Sia $(x_k)_k$ una successione in uno spazio topologico X. Se $x \in X$ si dice che la successione x_k* converge *a x se per ogni intorno U di x esiste $N > 0$ tale che per ogni $k > N$ si ha $x_k \in U$ (ovvero la successione è contenuta in U definitivamente).*

2.4 Limiti

Definizione 2.4.1 (limite)**.** *Siano X e Y spazi topologici e siano $A \subset X$, $B \subset Y$. Sia data una funzione $f \colon A \to B$ sia $x_0 \in X$ un punto di aderenza di A e $y_0 \in Y$. Diremo allora che f ha limite (o converge a) y_0 in x_0 o che $f(x)$ tende ad y_0 al tendere di x a x_0 e scriveremo*

$$\lim_{x \to x_0} f(x) = y_0, \text{ o anche } f(x) \to y_0 \text{ se } x \to x_0$$

se per ogni intorno V di y_0 in Y esiste un intorno U di x_0 in X tale che $f(U \cap (A \setminus \{x_0\})) \subset V$.

Diremo inoltre che il limite

$$\lim_{x \to x_0} f(x)$$

esiste se esiste $y_0 \in Y$ tale che $\lim_{x \to x_0} f(x) = y_0$.

Notiamo che si richiede che x_0 sia un punto aderente ad A in caso contrario la definizione di limite sarebbe sempre verificata (per qualunque scelta di y_0). Vedremo invece in seguito che con poche ipotesi sulla topologia di X il limite, quando esiste, è unico. Notiamo anche che y_0 necessariamente risulta essere un punto aderente a B.

I due casi $x_0 \in A$ e $x_0 \notin A$ nella precedente definizione hanno significati leggermente diversi. Notiamo comunque che anche se $x_0 \in A$ il valore di f in x_0 non assume alcuna rilevanza nella definizione di limite.

Se $x_0 \notin A$ e $\lim_{x \to x_0} f(x) = y_0$ è abbastanza naturale estendere la funzione $f \colon A \to B$ ad una funzione $\tilde{f} \colon A \cup \{x_0\} \to B \cup \{y_0\}$ definendo $\tilde{f}(x_0) = y_0$ e $\tilde{f}(x) = f(x)$ se $x \neq x_0$. Se invece $x_0 \in A$ la funzione $\tilde{f}$ definita come sopra, risulta essere una modifica di f. In ogni caso vale il seguente risultato.

Proposizione 2.4.2 (limite e continuità)**.** *Se $\tilde{f}$ è l'estensione (o modifica) di f in un punto x_0 in cui f ammette limite y_0, allora $\tilde{f}$ risulta essere continua in x_0 (rispetto alle topologie indotte da X e da Y)*

In particolare se $f \colon A \to B$ e $x_0 \in A$ vale la seguente proprietà:

$$\lim_{x \to x_0} f(x) = f(x_0) \Leftrightarrow f \text{ è continua in } x_0.$$

Dimostrazione. Sia V un intorno di y_0 in $B \cup \{y_0\}$ (rispetto alla topologia indotta da Y su $B \cup \{y_0\}$). Allora V contiene un aperto V' contenente y_0. Siccome V' è aperto rispetto alla topologia indotta deve esistere V'' aperto in Y tale che $V' = V'' \cap (B \cup \{y_0\})$. Dunque V'' è un intorno di y_0 in Y e per la definizione di limite esiste un intorno U di x_0 in X tale che $f(U \cap (A \setminus \{x_0\})) \subset V''$. Dato che $f(A) \subset B$ e dato che $\tilde{f}(x_0) = y_0 \in V$ si trova dunque che $\tilde{f}(U \cap (A \cup \{x_0\}) \subset V$. Essendo $U \cap (A \cup \{x_0\})$ un intorno di x_0 in $A \cup \{x_0\}$ abbiamo mostrato che $\tilde{f}$ è continua in x_0.

Se poi $x_0 \in A$ e $y_0 = f(x_0)$ allora $\tilde{f} = f$ e quindi abbiamo mostrato che f è continua in x_0. Viceversa se f è continua in x_0, dato un intorno V di $f(x_0)$ in Y allora $V \cap B$ è un intorno di $f(x_0)$ in B e quindi (essendo f continua in x_0) esiste un intorno U di x_0 in A tale che $f(U) \subset V$. Ma se U è un intorno di x_0 in A allora U contiene un aperto U' in A che contiene x_0 ed essendo la topologia di A la topologia indotta da X sappiamo esistere un aperto U'' di X tale che $U' = U'' \cap A$. Dunque U'' è un intorno di x_0 in X e vale $f(U'' \cap A) = f(U') \subset f(U) \subset V$. Dunque $\lim f = f(x_0)$ per $x \to x_0$. □

Definizione 2.4.3 (frequentemente, definitivamente). *Sia X uno spazio topologico, $A \subset X$ e $x_0 \in X$ un punto di aderenza di A. Sia $P(x)$ una proprietà definita per ogni $x \in A$. Diremo che $P(x)$ vale* frequentemente *per $x \to x_0$ se per ogni intorno U di x_0 esiste $x \in A \cap U$ tale che $P(x)$ è verificata. Diremo che $P(x)$ vale* definitivamente *per $x \to x_0$ se esiste un intorno U di x_0 tale che $P(x)$ è verificata per ogni $x \in U \cap (A \setminus \{x_0\})$.*

Possiamo quindi dire che $\lim_{x\to x_0} f(x) = y_0$ se per ogni V intorno di y_0 si ha $f(x) \in V$ definitivamente per $x \to x_0$.

Proposizione 2.4.4 (limiti in spazi metrici). *Siano X, Y spazi metrici, $A \subset X$, $B \subset Y$, $f \colon A \to B$, $x_0 \in X$, $y_0 \in Y$. Allora*

$$\lim_{x\to x_0} f(x) = y_0$$

è equivalente alla seguente proprietà:

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in A \setminus \{x_0\}\ d(x, x_0) < \delta \Rightarrow d(f(x), y_0) < \varepsilon.$$

Dimostrazione. Supponiamo che $\lim f = y_0$ per $x \to x_0$. Allora dato $\varepsilon > 0$ la palla $V = \mathbb{B}_\varepsilon(y_0)$ è un intorno di y_0. Dunque esiste un intorno U di x_0 tale che $f(U \cap (A \setminus \{x_0\}) \subset V$. Essendo U un intorno di x_0 esiste un $\delta > 0$ tale che $\mathbb{B}_\delta(x_0) \subset U$. Dunque $f(\mathbb{B}_\delta(x_0) \cap (A \setminus \{x_0\})) \subset V$ il che significa che se $x \in A \setminus \{x_0\}$ e $d(x, x_0) < \delta$ allora $f(x) \in V$ cioè $d(y, y_0) < \varepsilon$.

Supponiamo viceversa che valga la seconda proprietà. Se V è un qualunque intorno di y_0 in Y allora Y contiene una palla $\mathbb{B}_\varepsilon(y_0)$ per qualche $\varepsilon > 0$. Dunque esiste $\delta > 0$ tale che $f(\mathbb{B}_\delta(x_0) \cap (A \setminus \{x_0\})) \subset \mathbb{B}_\varepsilon(y_0)$. Dunque posto $U = \mathbb{B}_\delta(x_0)$ vale $\lim f(x) = y_0$ per $x \to x_0$. □

Teorema 2.4.5 (cambio di variabile nei limiti). *Siano X, Y, Z spazi topologici, $A \subset X$, $B \subset Y$, $C \subset Z$ e siano $x_0 \in X$ un punto aderente ad A, $y_0 \in Y$ un punto aderente a B, e $z_0 \in Z$ e siano $f \colon A \to B$, $g \colon B \to C$ tali che*

$$\lim_{x\to x_0} f(x) = y_0, \quad \lim_{y\to y_0} g(y) = z_0.$$

Allora

$$\lim_{x\to x_0} f(g(x)) = z_0.$$

Dimostrazione. Si considerino le estensioni $\tilde{f}\colon A\cup\{x_0\} \to B\cup\{y_0\}$ e $\tilde{g}\colon B\cup\{y_0\} \to C\cup\{z_0\}$. Siccome i limiti di f e g esistono in x_0 e y_0 allora, $\tilde{f}$ e $\tilde{g}$ sono continue in x_0 e y_0. Per il teorema della continuità della funzione composta, ricaviamo che $\tilde{f}\circ\tilde{g}$ è continua nel punto x_0. Ma tale funzione è proprio l'estensione della funzione $f\circ g$ al punto x_0 e dunque anche il limite di $f\circ g$ esiste e vale $\tilde{f}(\tilde{g}(x_0)) = z_0$. □

2.5 Retta reale estesa

Sarà comodo, nel seguito, considerare oltre ai numeri reali $\mathbb{R}$ i due *infiniti*: $+\infty$ e $-\infty$. Denoteremo con $\bar{\mathbb{R}} = \mathbb{R}\cup\{+\infty,-\infty\}$ l'insieme dei *numeri reali estesi* (o *retta reale estesa*). Su questo insieme possiamo estendere la relazione d'ordine $\le$ definita su $\mathbb{R}$ in modo da rendere anche $\bar{\mathbb{R}}$ totalmente ordinato. Poniamo infatti $-\infty \le x$ e $x \le +\infty$ per ogni $x\in\bar{\mathbb{R}}$. Possiamo estendere anche la somma a tutte le coppie $(x,y)\in\bar{\mathbb{R}}\times\bar{\mathbb{R}}\setminus\{(+\infty,-\infty),(-\infty,+\infty)\}$.

2.6 Successioni in spazi topologici

Definizione 2.6.1 (successione)**.** *Sia X uno spazio topologico. Una* successione *di punti in X è una funzione $a\colon\mathbb{N}\to X$. Alternativamente si può dire che una successione è un elemento di $X^{\mathbb{N}}$. Spesso la successione a si indica con $(a_k)_{k\in\mathbb{N}}$ o semplicemente con a_k intendendo che k varia tra i numeri naturali.*

Consideriamo l'insieme $\bar{\mathbb{N}} = \mathbb{N}\cup\{\infty\}$; questo insieme viene naturalmente dotato di una topologia come segue. Dato un punto $n\in\mathbb{N}$ consideriamo come base di intorni di n semplicemente il singoletto $\mathcal{U}_n = \{\{n\}\}$. Una base di intorni di ∞ sia invece $\mathcal{U}_\infty = \{\{j\in\mathbb{N}\colon j>k\}\cup\{\infty\}\colon k\in\mathbb{N}\}$. Notiamo quindi che ogni $n\in\mathbb{N}$ è un punto isolato di $\bar{\mathbb{N}}$ al contrario $\infty\in\bar{\mathbb{N}}$ è un punto di accumulazione di $\mathbb{N}\subset\bar{\mathbb{N}}$. Una successione è dunque una funzione definita su $\mathbb{N}\subset\bar{\mathbb{N}}$ e ha senso quindi farne il limite per $k\to\infty$.

Definizione 2.6.2. *Sia X uno spazio topologico e sia x_k una successione di punti di X e $x\in X$. Diremo che x_k converge a x se*

$$\lim_{k\to\infty} x_k = x.$$

Esplicitando la definizione degli intorni di ∞ in $\bar{\mathbb{N}}$ si vede che la proprietà $\lim_{k\to\infty} x_k = x$ è equivalente a richiedere che per ogni V intorno di x esista $N\in\mathbb{N}$ tale che per ogni $k>N$ vale $a_k\in V$.

2.7 Teoremi di compattezza

Definizione 2.7.1 (compattezza sequenziale)**.** *Uno spazio topologico si dice* sequenzialmente compatto *se da ogni successione si può estrarre una sottosuccessione convergente.*

Definizione 2.7.2 (totale limitatezza)**.** *Uno spazio metrico (X,d) si dice* totalmente limitato *se per ogni $\varepsilon>0$ esiste un insieme finito $Y\subset X$ tale che*

$$X \subset \bigcup_{y\in Y} \mathbb{B}_\varepsilon(y).$$

L'insieme Y viene chiamato ε-reticolo *di X.*

Definizione 2.7.3 (completezza)**.** *Sia (X, d) uno spazio metrico. Una successione (x_n) in X si dice* successione di Cauchy *se vale la seguente proprietà:*

$$\forall \varepsilon > 0\ \exists N > 0\ \forall n, n' > N \quad d(x_n, x_{n'}) < \varepsilon.$$

Uno spazio metrico (X, d) si dice completo *se ogni successione di Cauchy in X è convergente.*

Teorema 2.7.4. *Se (X, d) è uno spazio metrico le seguenti proprietà sono equivalenti:*

(i) X è sequenzialmente compatto;

(ii) X è completo e totalmente limitato;

(iii) X è compatto.

Dimostrazione. Dimostriamo che (i) implica (ii). Sia dunque X sequenzialmente compatto. È facile mostrare che di conseguenza X è completo, infatti data una successione di Cauchy x_n esiste una sottosuccessione convergente $x_{n_k} \to \bar{x} \in X$. Ma allora si vede facilmente che l'intera successione x_n converge a $\bar{x}$.

Mostriamo ora che X è totalmente limitato costruendo un ε-reticolo. Dato $\varepsilon > 0$ costruiamo una successione x_n. Prendiamo $x_1 \in X$ qualsiasi. Poi scegliamo $x_2 \in X \setminus \mathbb{B}_\varepsilon(x_1)$ e, allo stesso modo $x_n \in X \setminus \bigcup_{k=1}^{n-1} \mathbb{B}_\varepsilon(x_k)$. Continuiamo così finché $X \setminus \bigcup_{k=1}^{n-1} \mathbb{B}_\varepsilon(x_k)$ non è vuoto. Se tale insieme non fosse mai vuoto, riuscirei infatti a costruire una successione x_n con la proprietà $d(x_n, x_m) > \varepsilon$ per ogni n, m. È chiaro che tale successione non ammette nessuna sottosuccessione convergente e quindi ciò non può succedere essendo X sequenzialmente compatto.

Dimostriamo che (ii) implica (i). Per ogni $k \in \mathbb{N}$ sia R_k un $1/k$-reticolo finito di X. Definiremo ora, induttivamente, una successione di punti (y_k) con la proprietà $y_k \in R_k$ e tale che $\mathbb{B}_{1/k}(y_k)$ contiene infiniti punti della successione. Definiremo inoltre una sottosuccessione x_{n_k} di x_n tale che $x_{n_k} \in \mathbb{B}_{1/k}(y_k)$.

Essendo R_1 un insieme finito, ed essendo le palle $\{\mathbb{B}_1(y)\}_{y \in R_1}$ un ricoprimento di X, esiste y_1 tale che la palla $\mathbb{B}_1(y_1)$ contiene infiniti punti della successione x_n. Sia x_{n_1} uno di questi punti.

Avendo definito y_k e x_{n_k} sappiamo che $\mathbb{B}_k(y_k)$ contiene infiniti punti della successione, tra cui x_{n_k}. Da questi infiniti punti, tolti quelli con indice minore di n_k rimangono comunque infiniti punti. Come prima sappiamo che una delle palle $\mathbb{B}_{1/k+1}(y)$ al variare di $y \in R_{k+1}$ contiene infiniti di questi punti della successione. Sia $x_{n_{k+1}}$ uno di questi.

Questo procedimento, dunque, ci ha permesso di definire una sottosuccessione x_{n_k}, verifichiamo ora che questa successione è di Cauchy. E infatti abbiamo che se $k' > k$ si ha $d(x_{n_k}, x_{n'_k}) \le d(x_{n_k}, y_k) + d(x_{n'_k}, y_k) \le 2/k$ in quanto x_{n_k} e tutti i termini successivi della successione sono conenuti in $\mathbb{B}_{1/k}(y_k)$.

Dimostriamo che se valgono sia (i) che (ii) allora vale (iii). Essendo X totalmente limitato possiamo trovare, per ogni n, un $1/n$-reticolo R_n per X. Sia $R = \bigcup_{n \in \mathbb{N}} R_n$; essendo gli R_n insiemi finiti R è numerabile, poniamo $R = \{x_n : n \in \mathbb{N}\}$. Notiamo ora che R è denso in X, infatti per ogni $x \in X$ esiste un punto $x_n \in R_n$ che dista meno di $1/n$ da x; dunque $x_n \to x$ e $x_n \in R$.

Sia $\mathcal{F} = \{A_I\}_{i \in I}$ un ricoprimento aperto di X. Per ogni $n \in \mathbb{N}$ si potrà quindi trovare un aperto $A_n \in \mathcal{F}$ che contiene x_n. L'unione $\bigcup_{n \in \mathbb{N}} A_n$ è un aperto che contiene il denso, quindi deve contenere tutto X. Dunque $\{A_n\}_{n \in \mathbb{N}}$ è un sottoricoprimento numerabile di $\mathcal{F}$.

Ora scegliamo una sottosuccessione $(A_{n_k})_{k \in \mathbb{N}}$ di $(A_n)_{n \in \mathbb{N}}$ nel seguente modo. Prendiamo $n_0 = 0$ e, una volta definiti $n_0, \dots, n_k$ definiamo n_{k+1} come il più piccolo intero maggiore di n_k tale che l'aperto $A_{n_{k+1}}$ non sia contenuto nell'unione dei precedenti: $\bigcup_{j=0}^{k} A_{n_j}$. Per concludere la dimostrazione è sufficiente mostrare che questo procedimento termina dopo un numero finito N di passi, perché in tal caso concludo che $A_{n_0} \cup \ldots \cup A_{n_N}$ contiene tutti gli aperti della successione $(A_n)_{n \in N}$ e quindi è un ricoprimento finito di X.

Se il procedimento non terminasse, otterrei una successione di aperti $(A_{n_k})_{k \in \mathbb{N}}$ tale che ogni aperto della successione non è contenuto nell'unione dei precedenti. Inoltre questa successione costituirebbe ancora un ricoprimento di X in quanto gli aperti che ho tolto dalla successione (A_n) sono tutti contenuti nell'unione di aperti di (A_{n_k}). Poniamo ora, per semplicità $A_k = A_{n_k}$.

Possiamo dunque scegliere un punto $x_k \in A_k \setminus \bigcup_{j=0}^{k-1}$. Essendo X sequenzialmente compatto la successione x_k ammette una sottosuccessione convergente ad un punto $\bar{x} \in X$. Dovrà allora esistere un $\bar{k}$ tale che $\bar{x} \in A_{\bar{k}}$ (in quanto $\{A_k\}_{k \in \mathbb{N}}$ è un ricoprimento di X). Ma questo è assurdo perché $x_k \in A_{\bar{k}}$ soltanto quando $k \leq \bar{k}$ e quindi $A_{\bar{k}}$ è un intorno di $\bar{x}$ che contiene solamente un numero finito di punti della successione x_k.

Dimostriamo che (iii) implica (i). Consideriamo una successione x_n di punti di X. Se esiste un punto $x \in X$ tale che ogni intorno di x contiene infiniti punti della successione, allora posso trovare una sottosuccessione di x_n che converge ad x. Supponiamo quindi per ogni $x \in X$ è possibile trovare un intorno aperto B_x di x tale che $\{n \colon x_n \in B_x\}$ è un insieme finito. Ovviamente gli insiemi B_x al variare di $x \in X$ formano un ricoprimento aperto di X. Essendo X compatto posso estrarre un sottoricoprimento finito $B_1, \ldots, B_m$. Ma ogni B_k contiene solo un numero finito di elementi della successione e quindi nell'unione dei B_k è contenuto solo un numero finito di elementi di x_n il che è assurdo.

□

Teorema 2.7.5 (Ascoli Arzelà, versione astratta)**.** *Siano (X, d) e (Y, d) due spazi metrici totalmente limitati. Sia $F \subset \mathcal{C}(X, Y)$ una famiglia di funzioni equicontinue cioè con la seguente proprietà:*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, y \in X \ \forall f \in F \quad d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon.$$

Allora F è totalmente limitato (rispetto alla metrica della convergenza uniforme indotta da $\mathcal{C}(X, Y)$).

Prima di fare la dimostrazione mettiamo in evidenza il seguente corollario, che è un enunciato più concreto del precedente teorema.

Corollario 2.7.6 (Ascoli Arzelà)**.** *Sia $\Omega \subset \mathbb{R}^n$ un insieme limitato e $f_k \colon \Omega \to \mathbb{R}^m$ una successione di funzioni tali che*

1. *esiste $M > 0$ tale che per ogni k e per ogni $x \in \Omega$ si ha $|f_k(x)| \leq M$ (le f_k sono equilimitate);*

2. *per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $x, y \in \Omega$ e per ogni k si ha (le f_k sono equicontinue)*

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Allora esiste una sottosuccessione che converge uniformemente.

Dimostrazione. Sia $Y = [-M, M]$ e $X = \Omega$. Siccome Y e $\overline{\Omega}$ sono compatti in uno spazio metrico separabile essi sono anche totalmente limitati. Poniamo $F = \{f_k \colon k = 1, 2, \ldots\}$. Siccome le funzioni sono equicontinue esse sono, in particolare, continue. Possiamo dunque applicare il teorema precedente per concludere che F è totalmente limitato. Dunque $\overline{F}$ è sequenzialmente compatto e quindi dalla successione data possiamo estrarre una sottosuccessione convergente. □

Vediamo ora la dimostrazione del teorema.

Dimostrazione. Dato $\varepsilon > 0$ intendiamo trovare un 4ε-reticolo per F. Sia $\delta > 0$ dato in corrispondenza di ε nella definizione di equicontinuità per F. Sia X_δ un δ-reticolo per X e Y_ε un ε-reticolo per Y. Ricordando ora che $Y_\varepsilon^{X_\delta}$ è l'insieme delle funzioni da X_δ in Y_δ, definiamo $G_\varepsilon \subset Y_\varepsilon^{X_\delta}$ come

$$G_\varepsilon = \{g \in Y_\varepsilon^{X_\delta} : \exists f \in F\ \forall_{x\in X_\delta} \quad d(f(x), g(x)) < \varepsilon\}.$$

Siccome $Y_\varepsilon^{X_\delta}$ è un insieme finito, anche G_ε lo è, poniamo $G_\varepsilon = \{g_1, \dots, g_N\}$. Sia ora $F_\varepsilon \subset F$, $F_\varepsilon = \{f_1, \dots, f_N\}$ dove $f_k \colon X \to Y$ è una funzione di F tale che vale (come garantito dalla definizione di G_ε) $d(f_k(x), g_k(x)) < \varepsilon$ per ogni $x \in X_\delta$.

Dimostriamo ora che F_ε è un 4ε-reticolo per F. Data $f \in F$ scegliamo $g \in Y_\varepsilon{}^{X_\delta}$ tale che per ogni $x \in X_\delta$ si abbia $d(f(x), g(x)) < \varepsilon$ (questo è possibile in quanto per ogni $x \in X_\delta$ esiste $y \in Y_\varepsilon$ con $d(f(x), y) < \varepsilon$). Ne risulta che $g \in G_\varepsilon$ e quindi $g = g_k$ per un certo $k \in \{1, \dots, N\}$. Notiamo poi che per ogni $x \in X_\delta$ si ha quindi $d(f(x), f_k(x)) \le d(f(x), g_k(x)) + d(g_k(x), f_k(x)) < 2\varepsilon$.

Scelto ora un $x \in X$ qualunque sappiamo esistere $x_\delta \in X_\delta$ tale che $d(x, x_\delta) < \delta$. Si ha quindi, sfruttando l'equicontinuità di F,

$$d(f(x), f_k(x)) \le d(f(x), f(x_\delta)) + d(f_k(x), f_k(x_\delta)) + d(f(x_\delta), f_k(x_\delta)) \le 4\varepsilon.$$

□

Teorema 2.7.7 (compattezza in L^p)**.** *Sia $\Omega \subset \mathbb{R}^n$ un aperto. Un insieme $X \subset L^p(\Omega)$ ($1 \le p < \infty$) è relativamente compatto se e solo se:*

(i) $\sup_{f\in X} ||f||_{L^p(\Omega)} < +\infty$ *(equilimitatezza);*

(ii) dato $\varepsilon > 0$ per ogni aperto $U \subset \Omega$ tale che $\overline{U}$ è compatto esiste $\delta > 0$, $\delta < \operatorname{dist}(U, \mathbb{R}^n \setminus \Omega)$ tale che per ogni $h \in \mathbb{R}^n$ con $|h| < \delta$ e per ogni $f \in X$ si ha

$$||\tau_h f - f||_{L_p(U)} < \varepsilon$$

dove $\tau_h f$ è la funzione in $L^p(U)$ definita da $(\tau_h f)(x) = f(x + h)$ (equicontinuità);

(iii) dato $\varepsilon > 0$ esiste un compatto $K \subset \Omega$ tale che $\sup_{f\in X} ||f||_{L^p(\Omega\setminus K)} < \varepsilon$ (equitensione).

Dimostrazione. Vedi Brezis IV.5. □

Teorema 2.7.8 (compattezza in ℓ^p)**.** *Un insieme $X \subset \ell^p$ ($1 \le p < \infty$?) è relativamente compatto se e solo se:*

(i) $\sup_{x\in X} ||x||_{\ell^p} < +\infty$ *(equilimitatezza);*

(ii) dato $\varepsilon > 0$ esiste N tale che per ogni $x \in X$ si ha (equitensione):

$$\sum_{k>N} |x_k|^p < \varepsilon.$$

Dimostrazione. Dimostriamo che X è totalmente limitato. Fissiamo $\varepsilon > 0$. Sia N l'intero dato dalla proprietà di equitensione di X per $\frac{\varepsilon}{2}$. Chiaramente l'insieme $\bar{X} := \{(x_1, \dots, x_N) \colon x \in X\}$ è equilimitato, essendo un insieme limitato di $\mathbb{R}^N$. Esiste dunque un $\frac{\varepsilon}{2}$-reticolo $\bar{R}$ per $\bar{X}$. Verifichiamo che $R := \{(x_1, \dots, x_N, 0, 0, \dots) \colon x \in \bar{R}\}$ è un ε reticolo per X. Dato $x \in X$ esiste infatti $\bar{x} \in \bar{R}$ tale che

$$\sum_{k=0}^{N} |x_k - \bar{x}_k|^p < \frac{\varepsilon}{2}$$

d'altra parte per ipotesi
$$\sum_{k>N} |x_k|^p < \frac{\varepsilon}{2}$$
da cui segue $||x - \bar{x}||_{\ell^p} < \varepsilon$. □

2.8 Teoremi di punto fisso

Teorema 2.8.1 (contrazioni)**.** *Sia (X, d) uno spazio metrico completo e sia $T \colon X \to X$ una contrazione ossia una funzione che verifica:*
$$\exists \lambda < 1 \ \forall x, y \in X \qquad d(T(x), T(y)) \leq \lambda d(x, y).$$
Allora esiste unico un punto $\bar{x} \in X$ tale che $T(\bar{x}) = \bar{x}$.

Dimostrazione. Scegliamo un punto qualunque $x_0 \in X$ e consideriamo la seguente successione definita per ricorrenza:
$$x_{k+1} = T(x_k) \quad k > 0.$$
Si ha
$$d(x_{k+1}, x_k) \leq \lambda d(x_k, x_{k-1}) \leq \ldots \leq \lambda^k d(x_1, x_0)$$
e quindi se $m > n > N$ si ha
$$d(x_m, x_n) \leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq \sum_{k=n}^{m-1} \lambda^k d(x_1, x_0) \leq \lambda^N d(x_1, x_0) \sum_{k=0}^{\infty} \lambda^k = \lambda^N d(x_1, x_0) \frac{1}{1-\lambda}.$$
Dunque per $N \to \infty$ si ha $d(x_m, x_n) \to 0$ ovvero la successione $(x_k)_k$ è di Cauchy. Essendo X completo la successione converge ad un punto $\bar{x} \in X$.

Dunque fissato a piacere $\varepsilon > 0$ è possibile trovare un n per cui $d(x_n, \bar{x}) < \varepsilon$ e $d(x_{n+1}, \bar{x}) < \varepsilon$ da cui si trova
$$d(T(\bar{x}), \bar{x}) \leq d(T(\bar{x}), x_{n+1}) + d(x_{n+1}, \bar{x}) \leq \lambda d(\bar{x}, x_n) + d(x_{n+1}, \bar{x}) \leq 2\varepsilon$$
da cui (essendo la precedente disuguaglianza vera per ogni $\varepsilon > 0$) si ha $T(\bar{x}) = \bar{x}$.

Supponiamo ora che $T(\bar{y}) = \bar{y}$. Si ha allora
$$d(\bar{x}, \bar{y}) \leq d(\bar{x}, T(\bar{x})) + d(T(\bar{x}), T(\bar{y})) + d(T(\bar{y}), y) \leq \lambda d(\bar{x}, \bar{y})$$
da cui, essendo $\lambda < 1$, si ottiene $\bar{y} = \bar{x}$. □

Teorema 2.8.2 (Brouwer o Schauder)**.** *Sia X uno spazio di Banach, $K \subset X$ un insieme compatto convesso e non vuoto, $f : K \to K$ una funzione continua. Allora f ha un punto fisso.*

Capitolo 3

Spazi vettoriali topologici

3.1 Il differenziale

Siano X e Y spazi di Banach e sia $f\colon X \to Y$. Diremo che f è differenziabile in un punto $x_0 \in X$ se esiste una applicazione lineare e continua $L \in \mathcal{L}(X,Y)$ tale che

$$\lim_{x\to x_0} \frac{\|f(x)-f(x_0)-L(x-x_0)\|_Y}{\|x-x_0\|_X} = 0.$$

Tale applicazione L viene chiamato *differenziale di f nel punto x_0* e si indica con df_{x_0} o $Df(x_0)$.

Proposizione 3.1.1 (Lagrange)**.** *Siano X, Y spazi di Banach e $f\colon X \to Y$ una funzione differenziabile nel punto $x_0 \in X$. Allora vale*

$$\limsup_{x\to x_0} \frac{\|f(x)-f(x_0)\|_Y}{\|x-x_0\|_X} \le \|Df(x_0)\|_{\mathcal{L}(X,Y)}.$$

Dimostrazione. Si ha

$$\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|} \le \frac{\|f(x)-f(x_0)-Df(x_0)(x-x_0)\|}{\|x-x_0\|} + \|Df(x_0)\|$$

e passando al $\limsup$ si ottiene la tesi. □

Teorema 3.1.2 (differenziale della funzione composta)**.** *Siano X, Y, Z spazi di Banach e sia $f\colon X \to Y$ una funzione differenziabile in un punto $x_0 \in X$ e $g\colon Y \to Z$ una funzione differenziabile nel punto $y_0 = f(x_0)$. Allora la funzione composta $g \circ f\colon X \to Z$ è differenziabile nel punto x_0 e si ha*

$$D(g\circ f)(x_0) = Dg(f(x_0))Df(x_0).$$

Dimostrazione. Sia $L = Df(x_0)$ e $M = Dg(y_0)$. Bisogna verificare che vale

$$\lim_{x\to x_0} \frac{g(f(x))-g(f(x_0))-ML(x-x_0)}{\|x-x_0\|} = 0.$$

Si ha

$$\begin{aligned} & \frac{g(f(x))-g(f(x_0))-ML(x-x_0)}{\|x-x_0\|} \\ \le\quad & \frac{g(f(x))-g(y_0)-M(f(x)-y_0)}{\|f(x)-f(x_0)\|}\frac{\|f(x)-y_0\|}{\|x-x_0\|} \\ & +\|M\|\frac{\|f(x)-f(x_0)-L(x-x_0)\|}{\|x-x_0\|}. \end{aligned}$$

Il primo fattore del primo addendo tende a zero per la differenziabilità di g nel punto y_0 facendo il cambio di variabili $y = f(x)$ e notando che essendo f differenziabile in x_0 si ha che $y \to y_0$ per $x \to x_0$. Il secondo fattore del primo addendo risulta al limite limitato da $\|L\|$ per la disuguaglianza di Lagrange. Il secondo addendo tende a zero per la differenziabilità di f nel punto x_0. □

3.2 Convergenza uniforme

Sia f_k una successione di funzioni $f_k \colon X \to Y$ dove X è uno spazio topologico e Y uno spazio metrico. Data $f \colon X \to Y$ diremo che f_k *converge uniformemente* a f se vale

$$\forall \varepsilon > 0 \exists N > 0 \colon \forall k > N \forall x \in X \quad d_Y(f_k(x), f(x)) < \varepsilon.$$

Si può definire una funzione $d \colon Y^X \times Y^X \to \mathbb{R} \cup +\infty$ definita da

$$d(f, g) = \sup_{x \in X} d(f(x), g(x)).$$

Data questa definizione, si nota che la convergenza uniforme di f_k verso f è equivalente alla seguente proprietà

$$\lim_{k \to \infty} d(f_k, f) = 0.$$

Notiamo che d verifica tutte le proprietà di una distanza a parte il fatto che può verificarsi $d(f, g) = +\infty$. Se però Y è limitato o se invece di tutte le funzioni Y^X consideriamo solo il sottospazio delle funzioni limitate, allora d risulta essere effettivamente una distanza.

Teorema 3.2.1 (continuità del limite uniforme)**.** *Sia X uno spazio topologico, Y uno spazio metrico, $f_k \colon X \to Y$ funzioni continue, e $f \colon X \to Y$ una funzione qualsiasi. Se $d(f_k, f) \to 0$ (f_k converge uniformemente a f) allora anche f è una funzione continua.*

Dimostrazione. Fissiamo un punto $x_0 \in X$ in cui vogliamo dimostrare che f è continua. Dato $\varepsilon > 0$ dobbiamo trovare un intorno U di x_0 tale che $f(U) \subset B_\varepsilon(f(x_0))$.

Siccome $f_k \to f$ uniformemente sappiamo esistere un indice N tale che $d(f_N, f) < \varepsilon/3$ cioè per ogni $x \in X$ si ha $d(f_N(x), f(x)) < \varepsilon/3$. Essendo poi f_N continua in x_0 possiamo trovare un intorno U di x_0 tale che $f_N(U) \subset B_{\varepsilon/3}(f_N(x_0))$ ovvero per ogni $x \in U$ si ha $d(f_N(x), F_N(x_0)) < \varepsilon/3$. Dunque per ogni $x \in U$ si ha

$$d(f(x), f(x_0)) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(x_0)) + d(f_N(x_0), f(x_0)) \leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

cioè, come voluto, $f(x) \in B_\varepsilon(f(x_0))$. □

Sfruttando il teorema precedente si può ottenere il seguente fondamentale risultato

Teorema 3.2.2. *Sia X uno spazio topologico e Y uno spazio metrico completo. Lo spazio $C_b(X, Y)$ delle funzioni continue e limitate $f \colon X \to Y$ risulta essere completo rispetto alla metrica d della convergenza uniforme.*

Dimostrazione. Sia f_k una successione di Cauchy per la metrica d in $C_b(X, Y)$. Siccome per ogni $x \in X$ si ha $d_Y(f_k(x), f_j(x)) \leq d(f_k, f_j)$, fissato $x \in X$ la successione $f_k(x)$ è di Cauchy in Y. Essendo Y completo $f_k(x)$ converge in Y ad un punto che chiameremo $f(x)$. Abbiamo dunque definito una funzione $f \colon X \to Y$ tale che la successione di funzioni f_k converge puntualmente a f. Sarà sufficiente dimostrare che la convergenza $f_k \to f$ è uniforme. Infatti per il teorema precedente questo ci garantisce che la funzione limite f è continua (chiaramente è anche limitata) e dunque la successione f_k converge a f in $C_b(X,Y)$.

Dimostriamo dunque che $f_k \to f$ uniformemente. Essendo f_k una successione di Cauchy dato $\varepsilon > 0$ sappiamo esistere N tale che per ogni $k, j > N$ si ha $d(f_k, f_j) < \varepsilon/2$. D'altra parte visto che f_k converge puntualmente a f sappiamo che dati $\varepsilon > 0$, $x \in X$ e k esisterà $j > k$ tale che $d(f_j(x), f(x)) < \varepsilon/2$. Dunque mettendo assieme le due proprietà otteniamo che per ogni $\varepsilon > 0$ esisterà N tale che per ogni $k > N$ e per ogni $x \in X$ esiste $j > k$ tale che contemporaneamente:

$$d(f_k, f_j) < \varepsilon/2, \quad d_Y(f_j(x), f(x)) < \varepsilon/2$$

e quindi

$$d(f_k(x), f(x)) \leq d(f_k, f_j) + d(f_j(x), f(x)) < \varepsilon.$$

Quest'ultima disuguaglianza non dipende più da j e quindi abbiamo mostrato che per ogni $\varepsilon > 0$ esiste N tale che per ogni $k > N$ e per ogni $x \in X$ si ha $d(f_k(x), f(x)) < \varepsilon$. Questo significa esattamente che $d(f_k, f) \to 0$. □

3.3 Serie e derivata discreta

Sia $\mathbb{R}^{\mathbb{N}}$ lo spazio (spazio vettoriale) delle successioni a valori in $\mathbb{R}$. Possiamo definire su questo spazio due operatori $\Delta, \Sigma\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ che avranno un ruolo analogo agli operatori di derivata e integrale per le funzioni. Data una successione a_n definiamo

$$\Delta a_n = \begin{cases} a_n - a_{n-1} & \text{se } n > 1 \\ a_1 & \text{se } n = 1, \end{cases} \qquad \Sigma a_n = \sum_{k=1}^n a_k.$$

Chiaramente Δ e Σ sono operatori lineari, e si verifica facilmente che sono uno l'inverso dell'altro.

Come per le derivate, il calcolo di Δ risulta usualmente più semplice di quello di Σ che verrà di conseguenza calcolato come inverso di Δ.

Si ha, per esempio:

$$\Delta n = n - (n-1) = 1, \quad \Delta n^2 = n^2 - (n-1)^2 = 2n - 1$$

da cui si ottiene

$$n = \frac{\Delta n^2 + \Delta n}{2} = \Delta \frac{n^2 + n}{2}.$$

In pratica siamo riusciti a trovare la nota formula

$$\sum_{j=1}^n j = \Sigma n = \frac{n^2 + n}{2} = \frac{n(n+1)}{2}.$$

E da

$$\Delta n^3 = n^3 - (n-1)^3 = 3n^2 - 3n + 1 = 3n^2 - \frac{3}{2}(\Delta n^2 + \Delta n) + \Delta n$$

si ottiene

$$n^2 = \frac{1}{3}\Delta n^3 + \frac{1}{2}\Delta n^2 + \frac{1}{6}\Delta n$$

da cui

$$\sum_{j=1}^n j^2 = \Sigma n^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n.$$

Tralasciando i passaggi intermedi si trova

$$n^3 = \frac{1}{4}\Delta n^4 + \frac{1}{2}\Delta n^3 + \frac{1}{4}\Delta n^2$$

da cui

$$\sum_{j=1}^{n} j^3 = \Sigma n^3 = \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}n^2 = \left(\frac{n^2+n}{2}\right)^2 = (\Sigma n)^2.$$

Si può proseguire induttivamente per trovare la somma di ogni potenza di n. Si noterà che Δn^k è un polinomio completo di grado $k-1$ i cui coefficienti sono i coefficienti binomiali a segni alterni. La potenza n^{k-1} di ordine maggiore avrà coefficiente k. Le potenze n^j di ordine minore $j < k-1$ potranno essere integrate tramite le formule già trovate per Σn^j.

3.4 L'esponenziale di matrici

Sia $\mathbb{M} = \mathbb{M}^{n,n}$ lo spazio delle matrici $n \times n$ a coefficienti reali o complessi. Su $\mathbb{M}$ possiamo mettere la norma delle applicazioni lineari: $\|A\| = \max_{|v|\le 1} |Av|$; con questa norma $\mathbb{M}$ risulta essere uno spazio di Banach (normato completo).

Teorema 3.4.1 (esponenziale di matrici)**.** *È ben definita la funzione* $\exp\colon \mathbb{M} \to \mathbb{M}$*:*

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Inoltre valgono le seguenti proprietà.

1. *Le matrici A e $\exp A$ commutano.*
2. *per ogni s, t scalari si ha $\exp((t+s)A) = \exp(tA)\exp(sA)$, in particolare $\exp(A)$ è invertibile e la sua inversa è $\exp(-A)$.*
3. *Fissato $A \in \mathbb{M}$, la funzione $E(t) = \exp(At)$ è l'unica soluzione del problema di Cauchy:*

$$\begin{cases} E'(t) = AE(t) \\ E(0) = \mathrm{id}. \end{cases}$$

Dimostrazione. Bisogna innanzitutto provare che la serie in questione converge. Per fare questo verifichiamo che la successione delle somme parziali è di Cauchy ovvero che la serie è assolutamente convergente:

$$\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\| \le \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = \exp\|A\|.$$

Verifichiamo ora che $\mathrm{d}/\mathrm{d}t \exp(tA) = A\exp(tA)$. Per far questo notiamo che la serie

$$\exp(tA) = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

ha come serie delle derivate

$$\sum_{k=1}^{\infty} \frac{kA^k t^{k-1}}{k!} = A\sum_{k=0}^{\infty} \frac{A^k t^k}{k!} = A\exp(tA).$$

Entrambe le serie convergono totalmente sui compatti $|t| \leq C$ e dunque la serie delle derivate converge alla derivata della serie. Notando anche che $\exp(0) = \mathrm{id}$, abbiamo verificato che $\exp(tA)$ è una soluzione del problema di Cauchy dell'enunciato.

Verifichiamo ora che $\exp(tA)\exp(-tA) = \mathrm{id}$. Innanzitutto questa uguaglianza è banalmente vera per $t = 0$ in quanto $\exp(0) = \mathrm{id}$. Dunque è sufficiente mostrare che le derivate sono uguali. E infatti troviamo $(\exp(tA)\exp(-tA))' = A\exp(tA)\exp(-tA) + \exp(tA)(-A)\exp(-tA) = 0$, dove nell'ultima uguaglianza abbiamo sfruttato il fatto che A e $\exp(sA)$ commutano, come si può facilmente notare dalla definizione di exp.

Verifichiamo ora che $\exp(tA)$ è l'unica soluzione del problema di Cauchy enunciato nel teorema. Sia $E(t)$ una soluzione qualunque del problema di Cauchy in questione. Dunque $E'(t) - AE(t) = 0$. Moltiplicando a sinistra per $\exp(-tA)$ otteniamo la derivata di un prodotto: $(\exp(-tA)E(t))' = 0$. Siccome per $t = 0$ i due fattori sono entrambi uguali a id otteniamo, per ogni t: $\exp(-tA)E(t) = \mathrm{id}$. Moltiplicando a sinistra per $\exp(tA)$ (e ricordando che $\exp(tA)$ è l'inversa di $\exp(-tA)$) otteniamo $E(t) = \exp(tA)$ confermando quindi l'unicità della soluzione.

Dimostriamo ora che $\exp((t+s)A) = \exp(tA)\exp(sA)$. Sia $E(t) = \exp(-sA)\exp((t+s)A)$. Notiamo che $E(0) = \mathrm{id}$ e che $E'(t) = \exp(-sA)A\exp((t+s)A) = AE(t)$. Dunque $E(t)$ è soluzione del solito problema di cauchy e per l'unicità della soluzione abbiamo $E(t) = \exp(tA)$ che è proprio quello che volevamo mostrare. □

Cerchiamo ora alcuni strumenti per calcolare più facilmente l'esponenziale di una matrice. Innanzitutto dalla definizione si evince facilmente che se M è invertibile allora

$$\exp(M^{-1}AM) = M^{-1}(\exp A)M.$$

Inoltre se A è una matrice a blocchi del tipo

$$A = \begin{pmatrix} E & 0 \\ 0 & F \end{pmatrix} \Rightarrow \exp A = \begin{pmatrix} \exp E & 0 \\ 0 & \exp F \end{pmatrix}$$

infatti si nota che i blocchi di A^k sono E^k e F^k e quindi il risultato segue facilmente dalla definizione. In particolare, se $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ è una matrice diagonale si ottiene che $\exp(A)$ è anch'essa una matrice diagonale, in particolare $\exp(A) = \mathrm{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n))$.

Ricordiamo però che ogni matrice complessa può essere ridotta in forma di Jordan, ovvero è equivalente ad una matrice a blocchi di Jordan. Per calcolare l'esponenziale di una matrice anche non diagonalizzabile, sarà sufficiente saper calcolare l'esponenziale dei blocchi di Jordan. Sia dunque J^n_λ il generico blocco di Jordan:

$$J^n_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}.$$

Teorema 3.4.2 (esponenziale dei blocchi di Jordan)**.** *Se $J = J^n_\lambda$ è un blocco di Jordan si ha*

$$\exp(tJ) = e^{\lambda t}M(t)$$

dove

$$M(t) = \begin{pmatrix} 1 & t & t^2/2 & \cdots & t^n/n! \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & t^2/2 \\ \vdots & \ddots & \ddots & \ddots & t \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

Dimostrazione. Per comodità poniamo $m_k(t) = t^k/k!$ per $k \geq 0$ e $m_k(t) = 0$ per $k < 0$; in questo modo abbiamo $M_{ij}(t) = m_{j-i}(t)$. Vogliamo ora dimostrare che $M'(t) + \lambda M(t) = JM(t)$. Si ha $M'_{ij}(t) = m'_{j-i}(t) = m_{j-i-1}(t)$. Dunque $M'_{ij}(t) + \lambda M_{ij}(t) = m_{j-i-1}(t) + \lambda m_{ij}(t)$. D'altra parte si ha

$$(JM(t))_{ij} = \sum_k J_{ik} M_{kj}(t) = J_{ii} M_{ij}(t) + J_{i(i+1)} M_{(i+1)j}(t) = \lambda m_{j-i}(t) + m_{j-i-1}(t)$$

e quindi l'asserto è dimostrato. Notiamo quindi che $(M(t)e^{\lambda t})' = JM(t)e^{\lambda t}$ da cui si conclude che $M(t)e^{\lambda t} = \exp(tJ)$. □

Una applicazione dell'esponenziale di matrici la troveremo in seguito nella risoluzione dei sistemi lineari di equazioni differenziali a coefficienti costanti.

3.5 Algebra lineare

Norme sulle matrici quadrate. Sia A una matrice $n \times n$. Denoteremo con A_k la k-esima riga di A e con A^j la j-esima colonna di A.

Definiamo

$$||A||_2 = \sqrt{\sum_{j=1}^n \sum_{k=1}^n (A_k^j)^2}.$$

Notiamo che

$$||A||_2 = \sqrt{\sum_{k=1}^n |A^k|^2} = \sqrt{\sum_{j=1}^n |A_j|^2}$$

E si ha anche

$$||A||_2 = \sqrt{\operatorname{tr}(AA^*)}$$

(dove A^* è la matrice trasposta di A: $(A^*)_k^j = A_j^k$) infatti

$$\operatorname{tr}(AA^*) = \sum_k (AA^*)_k^k = \sum_k \sum_j (A_k^j (A^*)_j^k) = \sum_k \sum_j (A_k^j)^2 = ||A||_2^2.$$

Un'altra norma che è possibile mettere sulle matrici è la norma uniforme delle applicazioni lineari:

$$||A|| = \sup_{v \in \mathbb{R}^n \setminus 0} \frac{|Av|}{|v|}.$$

Proposizione 3.5.1. *Sia A una matrice $n \times n$. Si ha*

$$||A|| \leq ||A||_2 \leq \sqrt{n}||A||.$$

Dimostrazione. Per la prima disuguaglianza notiamo che, preso un vettore v con $|v| = 1$, si ha (ricordiamo che $|v^* w|^2 \leq |v|^2 |w|^2$ per $v, w \in \mathbb{R}^n$)

$$|Av|^2 = \sum_k |A_k v|^2 \leq \sum_k |A_k|^2 |v|^2 = \sum_k |A_k|^2 = ||A||_2^2.$$

Per l'altra disuguaglianza, si ha (posto e_k il versore unitario con un 1 al k-esimo posto)

$$||A||_2^2 = \sum_k |A_k|^2 = \sum_k |Ae_k|^2 \leq \sum_k ||A||^2 \leq n||A||^2.$$

Notiamo inoltre che è possibile fare semplici esempi in cui valgono le uguaglianze.

□

Lemma 3.5.2 (Lagrange). *Sia $f\colon \mathbb{R}^n \to \mathbb{R}^n$ una funzione derivabile. Allora*

$$|f(x) - f(y)| \leq \sqrt{n} \sup_{z\in[x,y]} ||Df(z)||_2 |x-y|.$$

Dimostrazione. Applichiamo il teorema di Lagrange alle funzioni

$$t \mapsto f^k(tx + (1-t)y)$$

ottenendo che esistono $\xi_k \in [0,1]$ tali che

$$f^k(x) - f^k(y) = \frac{d}{dt}_{|t=\xi_k} f^k(tx + (1-t)y) = Df^k(\xi_k x + (1-\xi_k)y)(x-y),$$

da cui

$$|f^k(x) - f^k(y)|^2 \leq \sup_{z\in[x,y]} |Df^k(z)|^2 |x-y|^2 \leq \sup_{z\in[x,y]} ||Df(z)||_2^2 |x-y|^2$$

e, sommando su k,

$$|f(x) - f(y)|^2 \leq n \sup_{z\in[x,y]} ||Df(z)||_2 |x-y|^2.$$

□

3.6 Invertibilità locale

Teorema 3.6.1 (criterio di Lipschitz). *Sia $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ una funzione di classe $\mathcal{C}^1$ definita su un insieme convesso Ω e supponiamo che si abbia $L = \sup_{x\in\Omega} \|Df(x)\| < +\infty$. Allora la funzione f è L-Lipschitziana, ossia*

$$|f(x') - f(x)| \leq L|x'-x| \qquad \forall x, x' \in \Omega.$$

Dimostrazione. Siano $x, x' \in \Omega$ fissati e consideriamo la funzione $g(t) = f(tx' + (1-t)x)$. Chiaramente $g\colon [0,1] \to \mathbb{R}^m$ è di classe $\mathcal{C}^1$ e si ha

$$g'(t) = Df(tx' + (1-t)x)\,(x'-x).$$

Dalla formula fondamentale del calcolo integrale si ottiene

$$f(x') - f(x) = g(1) - g(0) = \int_0^1 g'(t)\,dt = \int_0^1 Df(tx' + (1-t)x)\,dt\,(x'-x)$$

da cui segue

$$|f(x') - f(x)| \leq \int_0^1 \|Df(tx' + (1-t)x)\|\,dt\,|x'-x| \leq L|x'-x|$$

come volevasi dimostrare. □

Teorema 3.6.2 (inversione locale)**.** *Sia Ω un aperto di $\mathbb{R}^n$, $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$, sia $x_0 \in \Omega$ e supponiamo che $Df(x_0)$ sia una matrice invertibile. Allora esiste un intorno U di x_0 e un intorno V di $f(x_0)$ per cui $f(U) = V$ e*

$$f_{|U} \colon U \to V$$

è bigettiva, la sua inversa $f^{-1} \colon V \to U$ è differenziabile e se $f(x) = y$ (con $x \in U$) si ha

$$D(f^{-1}(y)) = (Df(x))^{-1}.$$

Dimostrazione. Sia $A = (Df(x_0))^{-1}$. Siccome Ω è aperto e $Df \colon \Omega \to \mathbb{R}^n$ è continuo, possiamo scegliere $\rho > 0$ in modo che

$$\overline{\mathbb{B}_\rho(x_0)} \subset \Omega \qquad \text{e} \qquad \|Df(x) - Df(x_0)\| \leq \frac{1}{2\|A\|} \quad \forall x \in \overline{\mathbb{B}_\rho(x_0)}.$$

Poniamo $r = \frac{\rho}{2\|A\|}$ e $y_0 = f(x_0)$. Fissato $y \in \mathbb{B}_r(y_0)$ consideriamo l'applicazione

$$\begin{aligned} T \colon \overline{\mathbb{B}_\rho(x_0)} &\to \mathbb{R}^n \\ T(x) &= x + A(y - f(x)). \end{aligned}$$

Si ha, per $x \in \overline{\mathbb{B}_\rho(x_0)}$,

$$\|DT(x)\| = \|\mathrm{id} - ADf(x)\| \leq \|A\| \cdot \|Df(x_0) - Df(x)\| \leq \frac{1}{2}$$

e quindi, per il criterio di Lipschitz, la funzione T è $\frac{1}{2}$-Lipschitziana (ossia una contrazione).

Verifichiamo anche che $T(\overline{\mathbb{B}_\rho(x_0)}) \subset \overline{\mathbb{B}_\rho(x_0)}$. Infatti, per $x \in \overline{\mathbb{B}_\rho(x_0)}$, si ha

$$\begin{aligned} |T(x) - x_0| &\leq |T(x) - T(x_0)| + |T(x_0) - x_0| \\ &\leq \frac{1}{2}|x - x_0| + |A(y - y_0)| \leq \frac{\rho}{2} + \|A\| r \leq \rho. \end{aligned}$$

Dunque $T \colon \overline{\mathbb{B}_\rho(x_0)} \to \overline{\mathbb{B}_\rho(x_0)}$ è una contrazione, $\overline{\mathbb{B}_\rho(x_0)}$ è uno spazio metrico completo e quindi, per il Teorema delle contrazioni, esiste un unica soluzione dell'equazione

$$T(x) = x,$$

cioè esiste un unico punto x in $\overline{\mathbb{B}_\rho(x_0)}$ per cui $f(x) = y$.

Dato $y \in \mathbb{B}_r(y_0)$ possiamo dunque trovare un $x \in \overline{\mathbb{B}_\rho(x_0)}$ che risolve $f(x) = y$, chiamiamo g l'applicazione $y \mapsto x$

$$g \colon \mathbb{B}_r(y_0) \to \overline{\mathbb{B}_\rho(x_0)}, \qquad f(g(y)) = y.$$

Poniamo dunque $V = \mathbb{B}_r(y_0)$ e $U = g(V)$. Chiaramente $f_{|U} \colon U \to V$ è bigettiva e ha g come inversa.

Vogliamo ora dimostrare che g è differenziabile in ogni punto $y \in \mathbb{B}_r(y_0)$ e che il suo differenziale $Dg(y)$ è uguale a $Df(x)^{-1}$ con $x = g(y)$; ossia

$$\lim_{y' \to y} \frac{g(y') - g(y) - Df(x)^{-1}(y' - y)}{|y' - y|} = 0.$$

Dato $y' \in \mathbb{B}_r(y_0)$, posto $x' = g(y')$ stiamo considerando la seguente quantità

$$\frac{g(y') - g(y) - Df(x)^{-1}(y' - y)}{|y' - y|} = \frac{x' - x - Df(x)^{-1}(f(x') - f(x))}{|y' - y|}$$

che per la differenziabilità di f in x possiamo scrivere come

$$\begin{aligned} &= \frac{x'-x-Df(x)^{-1}(Df(x)(x'-x)+o(|x'-x|))}{|y'-y|} \\ &= \frac{Df(x)^{-1}o(|x'-x|)}{|y'-y|} = \frac{Df(x)^{-1}o(|x'-x|)}{|x'-x|}\cdot\frac{|x'-x|}{|y'-y|}. \end{aligned} \tag{3.1}$$

Vogliamo ora dimostrare che per $y' \to y$ si ha $x' \to x$ cosicché il primo termine di quest'ultimo prodotto tende a zero per la definizione di o "¡piccolo"¿. Contemporaneamente dimostreremo che il secondo termine è limitato.

Per fare questo consideriamo la mappa T definita in precedenza e per la quale si ha $T(x) = x$. Ricordando che T è 1/2-lipschitziana si ha

$$\begin{aligned} \frac{1}{2}|x'-x| &\geq |T(x')-T(x)| = |x'+Df(x)(y-f(x'))-x| \\ &\geq |x'-x|-|Df(x)(y-f(x'))| = |x'-x|-\|Df(x)\||y-y'| \end{aligned}$$

da cui

$$|x'-x| \leq \frac{|y'-y|}{2\|Df(x)\|}.$$

Questo da un lato ci dice che se $y' \to y$ allora anche $x' \to x$ e dall'altro ci dice che il rapporto $|x'-x|/|y'-y|$ è limitato. Dunque la quantità in (3.1) tende a zero, la funzione g è differenziabile in y e il suo differenziale è $Dg(y) = Df(x)^{-1}$.

Per concludere che $g \in \mathcal{C}^1$ rimane solo da provare che il differenziale Dg è una funzione continua. Ma dato che $Dg(y) = Df(g(y))^{-1}$, si osserva che essendo g continua (in quanto differenziabile), essendo Df continua (per ipotesi) ed essendo continua anche l'operazione di inversione di una matrice, la funzione Dg deve essere continua.

L'ultima osservazione è che U risulta essere un intorno di x_0. Infatti $U = g^{-1}(V)$, ed essendo V aperto e g continua, concludiamo che anche U è aperto. □

Teorema 3.6.3 (funzione implicita, Dini)**.** *Sia Ω un aperto di $\mathbb{R}^n \times \mathbb{R}^m$ e sia $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$. Sia $(x_0, y_0) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^m$. Se la matrice $D_y f(x_0, y_0)$ delle derivate di $f(x, y)$ rispetto alle variabili y*

$$D_y f(x_0,y_0) = \left(\frac{\partial f^j}{\partial y^k}(x_0,y_0)\right)_{j,k} \qquad j=1,\dots,m \quad k=1,\dots,m$$

è invertibile, allora esiste un intorno $U \subset \mathbb{R}^n$ di x_0 e una funzione $g \in \mathcal{C}^1(U, \mathbb{R}^m)$ tale che

$$f(x,g(x)) = f(x_0,y_0) \qquad \forall x \in U.$$

Inoltre, per $y = g(x)$ si ha

$$Dg(x) = -D_y f(x,y)^{-1} D_y f(x,y).$$

Dimostrazione. Consideriamo la funzione $\tilde f \in \mathcal{C}^1(\Omega, \mathbb{R}^n \times \mathbb{R}^m)$

$$\tilde f(x,y) = (x, f(x,y)).$$

si trova facilmente che

$$D\tilde f(x,y) = \left(\begin{array}{c|c} \mathrm{id} & 0 \\ \hline D_x f & D_y f \end{array}\right)$$

dove id è la matrice identità $n \times n$. Dunque essendo $D_y f(x_0, y_0)$ invertibile, anche $D\tilde f(x_0, y_0)$ lo è e posso applicare il teorema di invertibilità locale a $\tilde f$. Dunque esisterà una funzione g di classe $\mathcal{C}^1$ in un intorno di $(x_0, f(x_0, y_0)$ tale che

$\tilde{f}(\tilde{g}(x,y)) = (x,y)$. Posto $\tilde{g}(x,y) = (g_1(x,y), g_2(x,y))$ con g_1 a valori in $\mathbb{R}^n$, e g_2 a valori in $\mathbb{R}^m$ si ha dunque

$$(x,y) = \tilde{f}(g_1(x,y), g_2(x,y)) = (g_1(x,y), f(g_1(x,y), g_2(x,y)))$$

da cui si ottiene $g_1(x,y) = x$ e $f(x, g_2(x,y)) = y$. Basterà dunque porre $g(x) = g_2(x, f(x_0, y_0))$ per ottenere $f(x, g(x)) = f(x_0, y_0)$.

Per dimostrare la formula di calcolo di Dg è sufficiente differenziare l'equazione

$$f(x, g(x)) = f(x_0, y_0)$$

per ottenere

$$D_x f(x, g(x)) + D_y f(x, g(x)) Dg(x) = 0$$

da cui la tesi. □

3.7 Equazioni differenziali

Teorema 3.7.1 (soluzione dell'equazione lineare del prim'ordine in forma normale)**.** *Sia I un intervallo di $\mathbb{R}$ e siano f, g funzioni continue su I. Tutte le soluzioni dell'equazione differenziale*

$$u'(x) = f(x)u(x) + g(x)$$

sono definite su I e sono della forma

$$u(x) = e^{F(x)}(G(x) + c)$$

dove F e G sono univocamente determinate dalle condizioni $F'(x) = f(x)$, $G'(x) = g(x)e^{-F(x)}$ e c è una costante arbitraria.

Dimostrazione. Moltiplicando l'equazione differenziale per $e^{-F(x)}$ otteniamo

$$u'(x)e^{-F(x)} - f(x)u(x)e^{-F(x)} = g(x)e^{-F(x)}$$

cioè

$$(u(x)e^{-F(x)})' = G'(x)$$

essendo questa relazione vera su un insieme connesso I si ricava, per qualche costante c

$$u(x)e^{-F(x)} = G(x) + c$$

che è la tesi del teorema □

Teorema 3.7.2 (indipendenza delle soluzioni)**.** *Siano $u_1, \dots, u_n$ funzioni derivabili n-volte su un intervallo I, soluzioni dell'equazione differenziale*

$$u^{(n)}(x) = a_{n-1}(x)u^{(n-1)}(x) + \ldots + a_1(x)u'(x) + a_0(x),$$

e si definisca quindi la matrice wronskiana

$$W(x) = \begin{pmatrix} y_1(x) & \cdots & y_n(x) \\ y_1'(x) & \cdots & y_n'(x) \\ \vdots & & \vdots \\ y_1^{(n-1)} & \cdots & y_n^{(n-1)} \end{pmatrix}.$$

Allora sono equivalenti

1. *le funzioni* $u_1, \dots, u_n$ *sono vettori indipendenti dello spazio* $\mathcal{C}^n(I)$;
2. *esiste un punto* $x_0 \in I$ *tale che* $\det W(x_0) \neq 0$;
3. *per ogni* $x \in I$ *si ha* $\det W(x) \neq 0$.

Dimostrazione. Mostriamo innanzitutto che se $\det W(x_0) \neq 0$ allora $u_1, \dots, u_n$ sono indipendenti (si noti che a questo scopo non è necessario supporre che le $u_1, \dots u_n$ siano soluzioni dell'equazione differenziale). Supponiamo quindi di avere una combinazione lineare

$$c_1 u_1(x) + \ldots + c_n u_n(x) = 0.$$

Derivando $n-1$ volte e valutando le n equazioni in x_0 si ottiene il sistema

$$W(x_0) \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0.$$

Essendo $\det W(x_0) \neq 0$ allora il sistema ha come unica soluzione $c_1 = \ldots = c_n = 0$.

Ci resta ora solo da dimostrare che se $u_1, \dots u_n$ sono indipendenti allora per ogni $x \in I$ si ha $\det W(x) \neq 0$. Per far questo mi sa che è proprio necessario usare il teorema di Cauchy sull'esistenza e unicità... □

Lemma 3.7.3 (rapprsentazione reale delle soluzioni). *Dati* $\alpha, \beta \in \mathbb{R}$

$$\operatorname{span}_{\mathbb{C}}(e^{(\alpha+i\beta)x}, e^{(\alpha+i\beta)x}) = \operatorname{span}_{\mathbb{C}}(e^{\alpha x}\cos(\beta x), e^{\alpha x}\sin(\beta x)).$$

3.7.1 Sistemi di equazioni differenziali lineari del primo ordine a coefficienti costanti

Teorema 3.7.4 (soluzione generale sistema lineare a coefficienti costanti). *Il sistema*

$$u'(x) = Au(x)$$

ha come soluzioni

$$u(x) = \exp(xA)u_0$$

dove u_0 *è un qualunque vettore costante.*

Dimostrazione. Notiamo innanzitutto che le funzioni $u(x) = \exp(xA)u_0$ sono tutte soluzioni, infatti $u'(x) = A\exp(xA)u_0 = Au(x)$. Sia ora $u(x)$ una generica soluzione di $u'(x) = Au(x)$. Notiamo che $(\exp(-xA)u(x))' = -A\exp(-xA)u(x) + \exp(xA)Au(x) = 0$, ed essendo $\exp(0)u(0) = u(0)$ si ha $u(x) = \exp(xA)u(0)$. Dunque $u(x)$ è una delle soluzioni già trovate. □

3.7.2 Equazioni differenziali lineari di ordine n

In questa sezione porremo $X = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, uno spazio vettoriale di dimensione infinita sul campo $\mathbb{C}$. Dato $\lambda \in \mathbb{C}$ e un intero non negativo m definiamo i seguenti sottospazi vettoriali di X:

$$V_\lambda^m = \{p(x)e^{\lambda x} \colon p(x) \in \mathbb{C}[x],\ \deg p < m\}.$$

Notiamo che V_λ^m ha dimensione m; notiamo inoltre che V_0^m è lo spazio dei polinomi complessi di grado strettamente minore di m.

Nel seguito consideremo l'operatore $(D-\lambda)\colon X \to X$ definito da $(D-\lambda)u(x) = u'(x) - \lambda u(x)$ per ogni $u \in X$; Notiamo fin da ora che vale la seguente proprietà:

$$(D-\lambda)p(x)e^{\lambda x} = e^{\lambda x}Dp(x).$$

Lemma 3.7.5. *Presi $\mu, \lambda \in C$, $\mu \neq \lambda$ e n, m interi non negativi, si ha*

$$(D-\lambda)^m(V_\mu^n) = V_\mu^n;$$

ovvero $(D-\lambda)^m$ ristretto a V_μ^n è un automorfismo di V_μ^n.

Dimostrazione. Notiamo innanzitutto che preso $u \in V_\mu^n$ si ha $(D-\lambda)u \in V_\mu^n$. Posto $u(x) = p(x)e^{\mu x}$ si ha $(D-\lambda)p(x)e^{\mu x} = e^{\mu x}(D-\lambda+\mu)p(x) = e^{\mu x}q(x) \in V_\mu^n$ essendo il grado di q non maggiore di quello di p. D'altra parte preso $u \in V_\mu^n$, $u(x) = p(x)e^{\mu x}$ l'equazione $(D-\lambda)u = 0$ è equivalente a $(D-\lambda+\mu)p(x) = 0$ cioè $p'(x) = (\lambda-\mu)p(x)$. Se p fosse diverso da zero il polinomio a sinistra dell'uguaglianza avrebbe grado strettamente minore di quello a destra e dunque necessariamente $p = 0$ e quindi $u = 0$. Dunque $(D-\lambda)$ è iniettivo su V_μ^n e di conseguenza è anche surgettivo. Iterando m volte si ottiene che anche $(D-\lambda)^m$ ristretto a V_μ^n è iniettivo e surgettivo. □

Lemma 3.7.6.

$$\ker(D-\lambda)^m = V_\lambda^m.$$

Dimostrazione. Facciamo la dimostrazione per induzione su m. Per $m = 0$ notiamo che $V_\lambda^0 = \{0\}$ e infatti $(D-\lambda)^0$ è per convenzione l'identità.

Supponiamo ora di sapere, per induzione, che $\ker(D-\lambda)^{m-1} = V_\lambda^{m-1}$.

Mostriamo che $\ker(D-\lambda)^m \subset V_\lambda^m$. Preso $u \in \ker(D-\lambda)^m$ si ha $(D-\lambda)(D-\lambda)^{m-1}u = 0$ da cui $(D-\lambda)u \in \ker(D-\lambda)^{m-1} = V_\lambda^{m-1}$ e quindi $(D-\lambda)u(x) = p(x)e^{\lambda x}$ con p polinomio di grado minore di $m-1$. Moltiplicando ambo i membri per $e^{-\lambda x}$ si ottiene $(u(x)e^{-\lambda x})' = p(x)$ da cui $u(x)e^{-\lambda x} = q(x)$ dove q è un polinomio di grado minore di m tale che $q' = p$. In conclusione abbiamo trovato che $u(x) = q(x)e^{\lambda x}$ e quindi $u \in V_\lambda^m$.

Preso invece $u \in V_\lambda^m$ potremo scrivere $u(x) = q(x)e^{\lambda x}$. Troviamo quindi $(D-\lambda)u(x) = q'(x)e^{\lambda x} \in V_\lambda^{m-1} = \ker(D-\lambda)^{m-1}$ da cui $u \in \ker(D-\lambda)^m$. □

Lemma 3.7.7. *Sia X uno spazio vettoriale, V un sottospazio vettoriale di X di dimensione finita, $T\colon X \to X$ un operatore lineare tale che $T(V) = V$ (T ristretto a V è un automorfismo di V). Allora $T^{-1}(V) = \ker T \oplus V$.*

Dimostrazione. Sia $u \in T^{-1}(V)$. Poniamo $v = T(u)$, $v \in V$. Essendo V di dimensione finita esiste un unico $w \in V$ tale che $T(w) = v$. Dunque $T(u-w) = T(u)-T(w) = 0$ da cui $w-u \in \ker T$. Abbiamo quindi verificato che $u = (u-w)+w$ sta in $\ker T + V$. Prendendo ora $v \in \ker T \cap V$; si ha ovviamente $T(v) = 0$, d'altra parte T ristretto a V è un automorfismo di V e quindi $v = 0$. Dunque $\ker T \cap V = \{0\}$ da cui $\ker T + V = \ker T \oplus V$. □

Lemma 3.7.8. *Siano $\lambda_1, \ldots, \lambda_k$ numeri complessi distinti e $m_1, \ldots, m_k$ interi non negativi. Allora*

$$\ker \prod_{j=1}^{k}(D-\lambda_j)^{m_j} = \bigoplus_{j=1}^{k} V_{\lambda_j}^{m_j}.$$

Dimostrazione. Facciamo la dimostrazione per induzione su k. Da un lemma precedente sappiamo che $\ker(D-\lambda_k)^{m_k} = V_{\lambda_k}^{m_k}$. Posto $T = \prod_{j=1}^{k-1}(D-\lambda_j)^{m_j}$ e $V = \bigoplus_{j=1}^{k-1} V_{\lambda_j}^{m_j}$ si ha, applicando il lemma precedente,

$$\begin{aligned}\ker((D-\lambda_k)^{m_k}T) &= \{u \in X\colon (D-\lambda_k)^{m_k}u \in \ker T\} \\ &= (D-\lambda_k)^{-m_k}(V) = \ker(D-\lambda_k)^{m_k} \oplus V.\end{aligned}$$

□

Teorema 3.7.9 (soluzione generale dell'omogenea)**.** *Tutte e sole le soluzioni complesse dell'equazione lineare omogenea di grado n, a coefficienti complessi costanti*

$$u^{(n)}(x) + a_{n-1}u^{(n-1)}(x) + \ldots + a_1 u'(x) + a_0 u(x) = 0$$

sono della forma

$$u(x) = p_1(x)e^{\lambda_1 x} + \ldots + p_k(x)e^{\lambda_k x}$$

dove $\lambda_1, \ldots, \lambda_k$ sono le soluzioni complesse distinte del polinomio associato

$$p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0,$$

$m_1, \ldots, m_k$ sono le rispettive molteplicità ($m_1 + \ldots + m_k = n$) e $p_1, \ldots, p_k$ sono polinomi qualunque a coefficienti complessi di grado strettamente minore, rispettivamente, di $m_1, \ldots, m_k$.

Dimostrazione. Essendo $p(z) = \prod_{j=1}^k (z-\lambda_j)^{m_j}$ l'equazione differenziale diventa

$$\prod_{j=1}^k (D-\lambda_k)^{m_j} u = 0$$

e quindi il lemma precedente dà esattamente il risultato desiderato. □

Dato $\lambda \in \mathbb{C}$ e m intero non negativo definiamo l'operatore $T_\lambda^m \colon X \to X$ come

$$(T_\lambda^m u)(x) = (D-\lambda)^m (x^m u(x)).$$

Lemma 3.7.10.

$$T_\lambda^m(V_\lambda^n) = V_\lambda^n.$$

Dimostrazione. Prendiamo $u \in V_\lambda^n$, $u(x) = p(x)e^{\lambda x}$. Si ha $T_\lambda^m u(x) = (D-\lambda)^m(x^m p(x)e^{\lambda x} = e^{\lambda x} D^m x^m p(x) = q(x)e^{\lambda x}$ da cui si vede che $T_\lambda^m(u) \in V_\lambda^n$ in quanto il grado di $q(x) = D^m x^m p(x)$ è uguale al grado di p. D'altra parte se $T_\lambda^m(u) = 0$ dovrà essere $q = 0$ e quindi $D^m x^m p(x) = 0$. Se p fosse diverso da zero si avrebbe un assurdo in quanto l'operatore $p(x) \mapsto D^m x^m p(x)$ non fa diminuire il grado. Concludiamo che T_λ^m ristretto a V_λ^n è iniettivo e quindi è anche surgettivo. □

Teorema 3.7.11 (soluzione particolare della non omogenea)**.** *L'equazione differenziale*

$$u^{(n)}(x) + a_{n-1}u^{(n-1)}(x) + \ldots + a_1 u'(x) + a_0 u(x) = \bar{p}(x)e^{\lambda x}$$

dove $\bar{p}(x)$ è un polinomio, ha una soluzione particolare del tipo

$$u(x) = x^m \bar{q}(x)e^{\lambda x}$$

dove $\bar{q}$ è un polinomio dello stesso grado di $\bar{p}$ e m è la molteplicità di λ come radice del polinomio $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$ ($m = 0$ se $p(\lambda) \neq 0$).

Dimostrazione. Sia $p(z) = (z-\lambda)^m \prod_{j=1}^k (z-\lambda_j)^{m_j}$. Dobbiamo dimostrare che esiste $v \in V_\lambda^n$ tale che $u(x) = x^m v(x)$ è soluzione di $p(D)u(x) = \bar{p}(x)e^{\lambda x}$ dove $n = \deg \bar{p}$. Ovvero dobbiamo dimostrare che

$$\prod_{j=1}^k (D-\lambda_j)^{m_j} T_\lambda^m(V_\lambda^n) = V_\lambda^n.$$

Il lemma precedente ci dice che $T_\lambda^m(V_\lambda^n) = V_\lambda^n$ mentre da un lemma ancora precedente già sapevamo che $(D-\lambda_j)^{m_j}(V_\lambda^n) = V_\lambda^n$ essendo $\lambda_j \neq \lambda$. □

3.7.3 Equazioni non lineari

Lemma 3.7.12. *Siano dati $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$, $\delta > 0$, $R \in]0, +\infty]$, $L > 0$. Posto $I = [x_0, x_0 + \delta]$ (oppure $I = [x_0, x_0 - \delta]$) sia data $f \colon I \times \mathbb{B}_R(y_0) \to \mathbb{R}$ (se $R = +\infty$ intendiamo $\mathbb{B}_R(y_0) = \mathbb{R}^n$). Posto $M = \max\{|f(x, y_0)| \colon x \in I\}$ supponiamo inoltre che valgano le seguenti proprietà:*

- *f è continua;*
- *per ogni $x \in I$, $y_1, y_2 \in \mathbb{B}_R(y_0)$ vale $|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2|$ (lipschitzianità rispetto a y uniforme rispetto a x);*
- *$\delta < R/M$.*

Allora esiste una unica funzione derivabile $y \colon I \to \mathbb{B}_R(y_0)$ che risolve il problema di Cauchy

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0. \end{cases}$$

Dimostrazione. Innanzitutto verifichiamo che y è soluzione del problema di Cauchy se e solo se y verifica la seguente equazione integrale

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t)) \mathrm{d}t.$$

Infatti se y è una soluzione del problema di Cauchy, integrando l'equazione $y'(t) = f(t, y(t))$ tra x_0 e x e ricordando che $y(x_0) = y_0$ si ottiene l'equazione integrale.

Se invece y è soluzione dell'equazione integrale, notiamo innanzitutto che y è continua. Infatti il termine di destra è continuo per la continuità dell'integrale rispetto agli estremi di integrazione. Essendo y continua ne segue che anche $f(t, y(t))$ è una funzione continua in t. Dunque possiamo applicare il teorema fondamentale del calcolo integrale derivando membro a membro per ottenere l'equazione differenziale in questione. Si verifica inoltre banalmente che vale $y(x_0) = y_0$.

Definiamo ora per induzione una successione di funzioni $y_k \colon I \to \mathbb{B}_R(y_0)$ nel seguente modo:

$$y_0(x) = y_0, \quad y_{k+1}(x) = y_0 + \int_{x_0}^{x} f(t, y_k(t)) \mathrm{d}t.$$

Dobbiamo dimostrare che questa è una buona definizione, nel senso che per ogni $x \in I$ dovrà essere $y_k(x) \in \mathbb{B}_R(y_0)$. Dimostriamo questo per induzione. Per $k = 0$ si ha $y_0(x) = y_0 \in \mathbb{B}_R(y_0)$. Supponendo ora che $y_k(x) \in \mathbb{B}_R(y_0)$ per ogni $x \in I$ notiamo che

$$|y_{k+1}(x) - y_0| \le \left| \int_{x_0}^{x} f(t, y_k(t)) \mathrm{d}t \right| \le \int_{x_0}^{x} |f(t, y_k(t))| \mathrm{d}t \le M|x - x_0| \le M\delta < R$$

da cui segue $y_{k+1}(x) \in \mathbb{B}_R(y_0)$.

Dimostriamo ora per induzione la seguente proprietà:

$$|y_{k+1}(x) - y_k(x)| \le \frac{ML^k}{(k+1)!} |x - x_0|^{k+1}.$$

Per $k = 0$ si ha

$$|y_1(x) - y_0(x)| = \left| \int_{x_0}^{x} f(t, y_0) \mathrm{d}t \right| \le \int_{x_0}^{x} |f(t, y_0)| \mathrm{d}t \le M|x - x_0|.$$

Supposta vera la disuguaglianza per k otteniamo per $k+1$:

$$\begin{aligned}
|y_{k+1}(x) - y_k(x)| &\leq \int_{x_0}^x |f(t, y_k(t)) - f(t, y_{k-1}(t))| \mathrm{d}t \leq \int_{x_0}^x L|y_k(t) - y_{k-1}(t)| \mathrm{d}t \\
&\leq L \int_{x_0}^x \frac{ML^{k-1}}{k!} |t - x_0|^k \mathrm{d}t = \frac{ML^k}{k!} \frac{1}{1+k} |x - x_0|^{k+1}
\end{aligned}$$

che è la disuguaglianza cercata. Questo ci dice che la successione y_k è di Cauchy rispetto alla convergenza uniforme su I. Infatti notiamo che la serie

$$\sum_{k=0}^{\infty} \|y_{k+1} - y_k\| \leq \frac{M}{L} \sum_{k=0}^{\infty} \frac{(L\delta)^{i+1}}{(i+1)!} = \frac{M}{L}(\mathrm{e}^{L\delta} - 1) < +\infty$$

è convergente. Dunque i resti k-esimi della serie sono infinitesimi il che significa che $\|y_{k+j} - y_k\|$ è infinitesimo in k.

Dunque abbiamo provato che la successione y_k converge ad una funzione continua $\bar{y} \colon I \to \mathbb{B}_R(y_0)$. D'altra parte l'operatore $y \mapsto T(y)$, che alla funzione y associa la funzione $T(y)$ definita da $T(y)(x) = \int_{x_0}^x f(t, y(t)) dt$ è continuo rispetto alla convergenza uniforme (infatti $\|T(y_1) - T(y_2)\| \leq L\delta \|y_1 - y_2\|$) ed essendo per ogni k $y_{k+1} = y_0 + T(y_k)$ si ottiene, passando al limite, $\bar{y} = y_0 + T(\bar{y})$. Dunque $\bar{y}$ è una soluzione del problema integrale.

Verifichiamo ora che tale soluzione è unica. Supponiamo che esista un'altra soluzione y del problema integrale diversa da $\bar{y}$. Sia $\bar{x} = \max\{x \in I \colon y(x) = \bar{y}(x)\}$. Sicuramente questo massimo esiste, essendo y e $\bar{y}$ funzioni continue, ed è $\bar{x} > x_0$. Scegliamo poi $x \in [\bar{x}, \bar{x} + 1/(2L)]$ in modo che sia $|y(x) - \bar{y}(x)| = \max\{|y(t) - \bar{y}(t)| \colon t \in [\bar{x}, \bar{x} + 1/(2L)]\}$.

Allora

$$\begin{aligned}
|y(x) - \bar{y}(x)| &\leq \int_{x_0}^x |f(t, y(t)) - f(t, \bar{y}(t))| \mathrm{d}t = \int_{\bar{x}}^x |f(t, y(t)) - f(t, \bar{y}(t))| \mathrm{d}t \\
&\leq |x - \bar{x}| L \max_{t \in [\bar{x}, x]} |y(t) - \bar{y}(t)| = L|x - \bar{x}| \cdot |y(x) - \bar{y}(x)| \\
&\leq |y(x) - \bar{y}(x)|/2
\end{aligned}$$

il che è assurdo. □

Teorema 3.7.13 (Cauchy, esistenza e unicità locale)**.** *Sia Ω un aperto di $\mathbb{R} \times \mathbb{R}^n$ e $f \colon \Omega \to \mathbb{R}$ una funzione continua. Supponiamo che f sia lipschitziana rispetto alla seconda variabile uniformemente rispetto alla prima, cioè che esista $L > 0$ tale che*

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

per ogni $(x, y_1), (x, y_2) \in \Omega$. Sia poi $(x_0, y_0) \in \Omega$.

Allora esistono $\delta > 0$, $R > 0$ tali che $[x_0 - \delta, x_0 + \delta] \times \mathbb{B}_R(y_0) \subset \Omega$ e il problema di Cauchy

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0. \end{cases}$$

ha una unica soluzione $y \colon [x_0 - \delta, x_0 + \delta] \to \mathbb{B}_R(y_0)$ nell'intervallo $[x_0 - \delta, x_0 + \delta]$.

Dimostrazione. Sia $K \subset \Omega$ un intorno compatto del punto (x_0, y_0) e sia $M = \max_K |f|$. Scegliamo ora $R > 0$ e $\delta > 0$ abbastanza piccoli in modo che si abbia $[x_0 - \delta, x_0 + \delta] \times \mathbb{B}_R(y_0) \subset \Omega$ e $\delta < R/M$. Applicando il lemma precedente otteniamo il risultato desiderato. □

Teorema 3.7.14 (Cauchy, esistenza e unicità globale). *Sia $[a,b]$ un intervallo di $\mathbb{R}$, $x_0 \in [a,b]$ e sia $f\colon [a,b] \times \mathbb{R}^n \to \mathbb{R}$ una funzione continua e L-lipschitziana rispetto alla seconda variabile uniformemente rispetto alla prima. Allora il solito problema di Cauchy ha una unica soluzione definita su tutto $[a,b]$.*

Dimostrazione. Basta applicare il lemma con $R = +\infty$. □

La seguente è una dimostrazione alternativa del teorema di Cauchy.

Teorema 3.7.15 (Cauchy, esistenza e unicità). *Siano $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}$ e sia $f\colon \mathbb{B}_r(x_0) \times \mathbb{B}_R(y_0) \to \mathbb{R}$ con le seguenti proprietà:*

(i) f è continua;

(ii) $f(x,y)$ è L-lipschitziana in y uniformemente rispetto a x ovvero:

$$|f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2| \qquad \forall x \in \mathbb{B}_r(x_0)\ \forall y_1, y_2 \in \mathbb{B}_R(y_0).$$

Allora, posto $M = \|f\|_\infty$, e $\rho = \min\{R, 1/L, R/M\}$ esiste una unica funzione $y \in \mathcal{C}^1(\mathbb{B}_\rho(x_0), \mathbb{B}_R(y_0))$ che verifica

$$\begin{cases} y'(x) &= f(x,y(x)) \qquad \forall x \in \mathbb{B}_\rho(x_0) \\ y(x_0) &= y_0. \end{cases} \tag{3.2}$$

Si noti che se al posto di (i) ponessimo solamente $f(x,y)$ continua rispetto a x, per la proprietà (ii) si avrebbe comunque che f è continua.

Dimostrazione. Sia $X = \mathcal{C}(\mathbb{B}_\rho(x_0), \overline{\mathbb{B}_R(y_0)})$, e si consideri la mappa $T\colon X \to \mathcal{C}(\mathbb{B}_\rho(x_0), \mathbb{R})$ definita da

$$T(y)(x) = x_0 + \int_{x_0}^{x} f(t, y(t))\mathrm{d}t.$$

Osserviamo che $T(y) = y$ se e solo se y verifica il sistema (3.2). Infatti se $T(y) = y$, chiaramente si ha $y(x_0) = y_0$ e derivando si ottiene $y'(x) = f(x,y(x))$. D'altra parte se y verifica (3.2) allora integrando si ottiene proprio $y = Ty$.

Verifichiamo dunque che T è una contrazione:

$$|(T(y_1) - T(y_2))(x)| \le \int_{x_0}^{x} |f(t,y_1(t)) - f(t,y_2(t))|\mathrm{d}t \le \rho L \|y_1 - y_2\|_\infty$$

ed essendo $\rho L < 1$ ne consegue che T è effettivamente una contrazione.

Verifichiamo ora che $T(X) \subset X$. Sia dunque $y \in X$, cioè $|y(x) - y_0| \le R$ per ogni $x \in \mathbb{B}_\rho(x_0)$. Si avrà allora

$$|T(y)(x) - y_0| \le \int_{x_0}^{x} |f(t,y(t))|\mathrm{d}t \le \rho M \le R.$$

Dunque T ammette un unico punto fisso y che è la soluzione del nostro problema. □

3.8 L'integrale di Lebesgue

Per quanto riguarda l'integrale di Riemann si ha il seguente risultato

Teorema 3.8.1. *Sia $\{f_n\}$ una successione di funzioni limitate, Riemann-integrabili, definite su un intervallo limitato I a valori in $\mathbb{R}$. Se f_n converge uniformemente ad una certa funzione f, allora f è Riemann-integrabile e vale*

$$\lim_{n\to\infty} \int_I f_n(x)\mathrm{d}x = \int_I f(x)\mathrm{d}x.$$

I seguenti esempi mostrano che, nel teorema precedente, la limitatezza dell'intervallo e delle funzioni sono necessarie:

$$\begin{aligned} f_n\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \frac{1}{n}\chi_{[0,n]}(x) \end{aligned}$$

$$\begin{aligned} f_n\colon [0,1] &\to \mathbb{R} \\ x &\mapsto n\chi_{]0,\frac{1}{n}]}(x). \end{aligned}$$

La necessità di avere migliori proprietà di passaggio al limite per gli integrali è stata la principale motivazione della nascita dell'integrale di Lebesgue. Vediamo brevemente la definizione di integrale di Lebesgue.

Sia (X,μ) uno spazio misurabile. Diremo che una funzione $s\colon X \to [-\infty,+\infty]$ è una funzione semplice se la sua immagine è un insieme finito, cioè se esistono $E_1,\dots,E_n \subset X$ e $s_1,\dots,s_n \in [-\infty,+\infty]$ tali che

$$s(x) = \sum_{k=1}^n s_k \chi_{E_k}(x).$$

Se s è una funzione semplice misurabile allora gli insiemi E_k sono misurabili e possiamo quindi definire

$$I(s) = \sum_{k=1}^n s_k \mu(E_k).$$

Data una funzione misurabile $f\colon X \to [0,+\infty]$ definiamo l'integrale di Lebesgue di f su X come

$$\int_X f(x)\mathrm{d}\mu(x) = \sup_{0\le s\le f} I(s)$$

dove s varia tra tutte le funzioni semplici misurabili $0 \le s \le f$. Si noti che $\int_X f(x)\mathrm{d}\mu(x) \in [0,+\infty]$. La definizione non poteva essere data con l'inf degli integrali delle funzioni semplici maggiori di f, perché funzioni come $f\colon]0,1[\to \mathbb{R}$, $f(x) = x^{-\frac{1}{2}}$ non sono maggiorate da alcuna funzione semplice! Però se $\mu(x) < \infty$ e $f(x) \le M$, sapendo che $\int f = \int M - \int (M-f)$ si trova facilmente che $\int_X f(x)\mathrm{d}\mu(x) = \inf_{s\ge f} I(s)$. Dunque non è necessario definire gli integrali superiore e inferiore come si faceva con l'integrale di Riemann.

Se $f\colon X \to [-\infty,+\infty]$ diciamo che $f \in L^1(X)$ se $\int_X |f(x)|\mathrm{d}\mu(x) < +\infty$ e posto $f = f^+ - f^-$ con $f^+, f^- \ge 0$ definiamo

$$\int_X f(x)\mathrm{d}\mu(x) = \int_X f^+(x)\mathrm{d}\mu(x) - \int_X f^-(x)\mathrm{d}\mu(x).$$

Si noti dunque che se $f \in L^1(X)$ allora $\int_X f\mathrm{d}\mu \in]-\infty,+\infty[$.

Teorema 3.8.2 (Beppo Levi)**.** *Sia (X,μ) uno spazio misurabile, e sia $\{f_n\}$ una successione crescente di funzioni misurabili su X a valori in $[0,+\infty]$. Allora si ha*

$$\lim_{n\to\infty}\int_X f_n(x)\mathrm{d}\mu(x) = \int_X \lim_{n\to\infty} f_n(x)\mathrm{d}\mu(x).$$

Dimostrazione. Innanzi tutto, essendo la successione $n \mapsto f_n$ crescente e l'integrale monotono, entrambi i limiti esistono e coincidono con gli estremi superiori. Posto quindi

$$f(x) = \lim_{n\to\infty} f_n(x) = \sup_n f_n(x)$$

si nota immediatamente che f è una funzione misurabile, in quanto estremo superiore di funzioni misurabili. Inoltre essendo
$f_n \leq f$ per ogni n, per la monotonia dell'integrale si ha,

$$\sup_n \int_X f_n \mathrm{d}\mu(x) \leq \int_X f(x)\mathrm{d}\mu(x). \tag{3.3}$$

Prendiamo una qualsiasi funzione semplice misurabile s, con $0 \leq s \leq f$. Scelto $\alpha < 1$ poniamo poi

$$E_n = \{x \in X \colon f_n(x) \geq \alpha s(x)\}.$$

Gli insiemi E_n sono misurabili e sono una successione crescente che (essendo $\sup_n f_n(x) = f(x) \geq s(x) > \alpha s(x)$) invade tutto lo spazio X, inoltre si ha

$$\int_X f_n(x)\mathrm{d}\mu(x) \geq \int_{E_n} f_n(x)\mathrm{d}\mu(x) \geq \alpha \int_{E_n} s(x)\mathrm{d}\mu(x). \tag{3.4}$$

Essendo s una funzione semplice è facile verificare che la funzione $E \mapsto \int_E s(x)\mathrm{d}\mu(x)$ è una misura quindi si ha

$$\lim_{n\to\infty} \int_{E_n} s(x)\mathrm{d}\mu(x) = \int_X s(x)\mathrm{d}\mu(x)$$

e passando al limite nella disuguaglianza (3.4) si ottiene

$$\sup_n \int_X f_n(x)\mathrm{d}\mu(x) \geq \alpha \int_X s(x)\mathrm{d}\mu(x).$$

Siccome quest'ultima disuguaglianza è vera per ogni scelta di $\alpha < 1$ e per ogni funzione semplice misurabile $s \leq f$ si ha

$$\sup_n \int_X f_n(x)\mathrm{d}\mu(x) \geq \int_X f(x)\mathrm{d}\mu(x)$$

che insieme a (3.3) ci dà la tesi. □

Teorema 3.8.3 (Lemma di Fatou)**.** *Sia* (X, μ) *uno spazio misurabile e sia* $\{f_n\}$ *una successione di funzioni misurabili su* X *a valori in* $[0, +\infty]$*. Allora*

$$\int_X \liminf_{n\to\infty} f_n(x)\mathrm{d}\mu(x) \leq \liminf_{n\to\infty} \int_X f_n(x)\mathrm{d}\mu(x).$$

Dimostrazione. Posto $f(x) = \liminf_{n\to\infty} f_n(x)$ e $g_n(x) = \inf_{k\geq n} f_k(x)$, dalla definizione di limite superiore si ha

$$f(x) = \sup_n g_n(x).$$

Siccome g_n è una successione crescente di funzioni misurabili possiamo applicare il Teorema di Beppo Levi, ottenendo

$$\int_X f(x)\mathrm{d}\mu(x) = \lim_{n\to\infty} \int_X g_n(x)\mathrm{d}\mu(x)$$

ma, d'altra parte, essendo $g_n \leq f_n$, si conclude notando che

$$\lim_{n\to\infty} \int_X g_n(x)\mathrm{d}\mu(x) = \liminf_{n\to\infty} \int_X g_n(x)\mathrm{d}\mu(x) \leq \liminf_{n\to\infty} \int_X f_n(x)\mathrm{d}\mu(x).$$

□

Un esempio di come nella tesi del teorema precedente possa non esserci l'uguaglianza è dato dalla successione

$$f_n(x) = \chi_{[n,n+1]}(x).$$

Teorema 3.8.4 (Lebesgue, convergenza dominata)**.** *Sia (X,μ) uno spazio misurabile e sia $\{f_n\}$ una successione di funzioni misurabili $f_n\colon X \to [-\infty,+\infty]$ e supponiamo che per ogni x esista il limite*

$$f(x) = \lim_{n\to\infty} f_n(x).$$

Supponiamo inoltre che esista una funzione $g \in L^1(X)$ tale che $|f_n| \le g$ per ogni n. Allora $f \in L^1(X)$ e

$$\lim_{n\to\infty} \int_X |f_n(x) - f(x)| \mathrm{d}\mu(x) = 0.$$

Dimostrazione. Dalle stabilità delle funzioni misurabili per passaggi al limite sappiamo che f è una funzione misurabile. Si ha, sfruttando il Lemma di Fatou,

$$\begin{aligned}
\lim_{n\to\infty} \int_X |f(x) - f_n(x)| \mathrm{d}\mu(x) &\le \limsup_{n\to\infty} \int_X |f(x) - f_n(x)| \mathrm{d}\mu(x) \\
&= -\liminf_{n\to\infty} \int_X -|f(x) - f_n(x)| \mathrm{d}\mu(x) \\
&= \int_X 2g(x) \mathrm{d}\mu(x) - \liminf_{n\to\infty} \int_X 2g(x) - |f(x) - f_n(x)| \mathrm{d}\mu(x) \\
&\le \int_X 2g(x) \mathrm{d}\mu(x) - \int_X 2g(x) - \limsup_{n\to\infty} |f(x) - f_n(x)| \mathrm{d}\mu(x) \\
&= \int_X 2g(x) \mathrm{d}\mu(x) - \int_X 2g(x) \mathrm{d}\mu(x) = 0.
\end{aligned}$$

□

Teorema 3.8.5 (scambio derivata e integrale)**.** *Sia (Y,μ) uno spazio misurabile, $\Omega \subset \mathbb{R}^n$ un aperto e sia data $f\colon X \times Y \to \mathbb{R}$ tale che per ogni $x \in \Omega$ la funzione $y \mapsto f(x,y)$ sia misurabile.*

Sia

$$F(x) = \int_E f(x,y) \mathrm{d}y.$$

Se esiste $g \in L^1(Y)$ tale che per ogni $x \in \Omega$ $|f(x,y)| \le g(y)$ e se $x \mapsto f(x,y)$ è continua per ogni $y \in Y$, allora $F\colon \Omega \to \mathbb{R}$ è continua.

Se inoltre $x \mapsto f(x,y)$ è differenziabile in Ω per ogni $y \in Y$ e esistono $g_1,\dots,g_n \in L^1(Y)$ tali che $\left|\frac{\partial f}{\partial x_k}(x,y)\right| \le g_k(y)$ allora $F \in \mathcal{C}^1(\Omega)$ e

$$\frac{\partial F}{\partial x_k}(x) = \int_E \frac{\partial f}{\partial x_k}(x,y) \mathrm{d}y \qquad \forall x \in \Omega.$$

Dimostrazione. Sia $(x_h)_h$ una successione in Ω tale che $x_h \to x \in \Omega$. Applicando il teorema di Lebesgue alle funzioni $f_h(y) = f(x_h,y)$ si ottiene $F(x_h) \to F(x)$. Dunque F è continua.

Dimostriamo ora la seconda parte. Fissiamo un punto $x \in \Omega$, una successione $\varepsilon_h \to 0$ e sia e_k il k-esimo vettore della base canonica di $\mathbb{R}^n$. Sia poi $\xi_h(y) \in [0,\varepsilon_h]$ tale che (per il teorema di Lagrange)

$$\frac{f(x+\varepsilon_h e_k, y) - f(x,y)}{\varepsilon_h} = \frac{\partial f}{\partial x_k}(x + \xi_h(y) e_k, y).$$

Posto $f_h(y) = \frac{\partial f}{\partial x_k}(x + \xi_h(y), y)$ si ha che $|f_h(y)| \leq g_k(y)$ per ogni h e $f_h(y) \to \frac{\partial f}{\partial x_k}(x, y)$. Applicando dunque il teorema di Lebesgue alle f_h si ottiene dunque

$$\frac{F(x + \varepsilon_h e_k, y) - F(x, y)}{\varepsilon_h} = \int_Y f_h(y) \mathrm{d}y \to \int_Y \frac{\partial f}{\partial x_k}(x, y) \mathrm{d}y.$$

Dunque F è derivabile nella direzione k, e inoltre $\partial F/\partial x_k$ è continua. Ne consegue che F è differenziabile su tutto Ω come volevamo dimostrare. □

3.9 Spazi L^p

Proposizione 3.9.1 (disuguaglianza di Young)**.** *Siano $a, b > 0$ e $p, q \in]0, \infty[$ con $\frac{1}{p} + \frac{1}{q} = 1$. Allora*

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

Dimostrazione. Per la concavità del logaritmo abbiamo:

$$\frac{1}{p} \log a^p + \frac{1}{q} \log b^q \leq \log\left(\frac{1}{p} a^p + \frac{1}{q} b^q\right).$$

Facendo l'esponenziale di ambo i membri otteniamo la tesi. □

Sia (X, μ) uno spazio misurabile. Per $p \in [1, \infty[$ definiamo lo spazio $L^p(X, \mu)$ come l'insieme delle funzioni misurabili $f : X \to \mathbb{R}$ tali che

$$\|f\|_p := \left(\int_X |f|^p \right)^{\frac{1}{p}} < +\infty.$$

Teorema 3.9.2 (disuguaglianza di Hölder)**.** *Sia (X, μ) uno spazio misurabile e siano $f \in L^p(X, \mu)$ e $g \in L^q(X, \mu)$ con $p, q \in]1, \infty[$, $\frac{1}{p} + \frac{1}{q} = 1$. Allora*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Dimostrazione. Si ha, applicando la disuguaglianza di Young,

$$\frac{\int_X |fg| \mathrm{d}\mu}{\|f\|_p \|g\|_q} = \int_X \frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q} \mathrm{d}\mu \leq \frac{1}{p} \int_X \frac{|f|^p}{\|f\|_p^p} \mathrm{d}\mu + \frac{1}{q} \int_X \frac{|g|^q}{\|g\|_q^q} \mathrm{d}\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

□

3.10 Semi-continuità

Teorema 3.10.1. *Sia X uno spazio di Banach e $F : X \to \mathbb{R}$ una funzione convessa e continua. Allora F è debolmente s.c.i.*

Dimostrazione. Sia $A_a := F^{-1}(]a, +\infty[)$. Essendo F continua e convessa $X \setminus A_a$ è chiuso e convesso. Dunque $X \setminus A_a$ è anche debolmente chiuso (la topologia debole coincide con la forte sui convessi) e quindi A_a è debolmente aperto. Siccome gli insiemi $]a, +\infty[$ sono una base della topologia della semicontinuità inferiore su $\mathbb{R}$, abbiamo dimostrato che F è debolmente semicontinua inferiormente. □

Teorema 3.10.2. *Sia Ω un aperto di $\mathbb{R}^n$, e $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, $(x, u, p) \mapsto f(x, u, p)$ tale che*

(i) $x \mapsto f(x, u, p)$ è misurabile per ogni $u \in \mathbb{R}$ e ogni $p \in \mathbb{R}^n$;

(ii) $u \mapsto f(x,u,p)$ *è continua per quasi ogni* $x \in \Omega$ *e per ogni* $p \in \mathbb{R}^n$*;*

(iii) $p \mapsto f(x,u,p)$ *è convessa per quasi ogni* $x \in \Omega$ *e per ogni* $u \in \mathbb{R}$.

Allora il funzionale $F\colon W^{1,p}(\Omega) \to \mathbb{R}$ *(*$p > 1$*)*

$$F(u) = \int_\Omega f(x, u(x), Du(x)) \mathrm{d}x$$

è debolmente semicontinuo inferiormente.

Sarà vero?!?!?

3.11 Serie di Fourier

Sia H uno spazio di Hilbert e sia $(\cdot,\cdot)$ il prodotto scalare su H.

Teorema 3.11.1 (proiezione ortogonale)**.** *Sia* $E \subset H$ *un insieme convesso chiuso e non vuoto. Allora esiste unico un elemento di* E *con minima norma. Cioè esiste unico* $x \in E$ *tale che* $|x| \le |y|$ *per ogni* $y \in E$.

In particolare, se V *è un sottospazio vettoriale chiuso di* H*, possiamo definire l'applicazione lineare*

$$\Pi_V \colon H \to V$$

che manda un punto w *di* H *nel punto di* V *con minima distanza da* w*. Inoltre* $\Pi(w)$ *è caratterizzato dalla proprietà:*

$$(w - \Pi(w), v) = 0 \quad \forall v \in V.$$

cioè Π_v *è una* proiezione ortogonale*.*

Dimostrazione. Si verifica facilmente che per $x, y \in H$ vale la regola del parallelogramma:

$$|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2.$$

Sia ora $d = \inf_{x \in E} |x|$, e siano $x, y \in E$. Essendo E convesso, anche $\frac{x+y}{2} \in E$ e si ha dunque:

$$|x-y|^2 = 2|x|^2 + 2|y|^2 - 4\left|\frac{x+y}{2}\right|^2 \le 2|x|^2 + 2|y|^2 - 4d^2.$$

Dunque se avessimo due punti x, y di minima distanza, cioè tali che $|x| = |y| = d$ si avrebbe allora $|x - y| = 0$ e quindi $x = y$ da cui l'unicità del punto di minima distanza. Vediamo ora l'esistenza. Sia $x_k \in E$ tale che $|x_k| \to d$. Dall'ultima disuguaglianza si ha

$$|x_k - x_j|^2 \le 2|x_k|^2 + 2|x_j|^2 - 4d$$

e dunque $|x_k - x_j| \to 0$ per $k, j \to \infty$. Questo significa che x_k è una successione di Cauchy, ed essendo H completo $x_k \to x$.

Siccome E è chiuso si avrà $x \in E$, ed essendo $|\cdot|\colon H \to \mathbb{R}$ una funzione continua si ottiene $|x| = d$, ovvero x è un punto di norma minima.

Ora, se V è un sottospazio chiuso di H, e $w \in H$ se prendiamo $E = V - w$ e chiamiamo $v - w$ l'elemento di minima norma di E, otteniamo che v è l'unico punto di minima distanza di V da w. Possiamo dunque definire la mappa $\Pi_V\colon H \to V$ che associa al punto $w \in H$ il punto v di V di minima distanza da w.

Notiamo ora che se $v \in V$ si ha

$$|w - (\Pi_V(w) + v)|^2 = |w - \Pi_V(w)|^2 + |v|^2 + 2(w - \Pi(w), v)$$

e dovendo essere $|w-(\Pi_V(w)+v)| \geq |w-\Pi_V(w)|$ si ottiene $(w-\Pi_V(w),v) \leq 0$. Analogamente si ottiene anche $w-\Pi_V(w), -v) \leq 0$ da cui $(w-\Pi_V(w),v)=0$. D'altra parte se per un certo $v' \in V$ si ha $(w-\overline{w}, v) \leq 0$ per ogni $v \in V$, si ha

$$|w-v'| \leq |w-v| \quad \forall v \in V$$

da cui $v' = \Pi_V(w)$.

Ci resta ora solo da verificare che Π_V è lineare. Il fatto che $\Pi_V(\lambda w) = \lambda \Pi_V(w)$ deriva direttamente dall'osservazione che $d(\lambda w, V) = |\lambda| d(w,V)$ e che $\Pi_V(w) = \Pi_V(-w)$.

Verifichiamo ora che $\Pi_V(w+w') = \Pi_V(w) + \Pi_V(w')$. Basta notare che, preso $v \in V$,

$$(w+w'-(\Pi_V(w)+\Pi_V(w')), v) = (w-\Pi_V(w), v) + (w'-\Pi_V(w'), v) = 0$$

e quindi per la caratterizzazione delle proiezioni si ottiene $\Pi_V(w)+\Pi_V(w') = \Pi_V(w+w')$. □

Un insieme $\{e_k\}$ si dice un *sistema ortonormale completo* di H se:

1. $(e_j, e_k) = 0$ per $j \neq k$ (ortogonalità);
2. $(e_k, e_k) = 1$ (normalità);
3. $\forall k\ (v, e_k) = 0 \Rightarrow v = 0$ per $v \in H$ (completezza).

Teorema 3.11.2 (Serie di Fourier)**.** *Sia H uno spazio di Hilbert, e sia $\{e_k\}$, $k \in \mathbb{N}$ un sistema ortonormale completo. Allora, per ogni $v \in H$ si ha*

$$v = \sum_k (v, e_k) e_k.$$

Inoltre si ha

$$\|v\|_H^2 = \sum_k |(v, e_k)|^2.$$

Dimostrazione. Consideriamo lo spazio vettoriale $V \subset H$ generato dai vettori $e_1, \dots, e_n, \dots$, ovvero l'insieme delle combinazioni lineari finite di tali vettori:

$$V = \{\sum_{k=1}^n \lambda^k e_k \colon n \in \mathbb{N},\ \lambda^k \text{scalare}\}.$$

L'insieme $\overline{V}$ é un sottospazio vettoriale chiuso di H, verifichiamo ora che $\overline{V} = H$. Consideriamo la proiezione $\Pi = \Pi_{\overline{V}} \colon H \to \overline{V}$ come data dal teorema di proiezione ortogonale. Preso ora $w \in H$, si avrà per la caratterizzazione della proiezione ortogonale, $(w - \Pi(w), e_k) = 0$ per ogni k e dunque, per la complettezza del sistema (e_k), si ottiene $w - \Pi(w) = 0$ ovvero $w \in \overline{V}$, come volevamo dimostrare.

Dunque ogni elemento $v \in H$ può essere approssimato da elementi $v_n \in V_n$ dove

$$V_n = \{\sum_{k=1}^n \lambda^k e_k \colon \lambda^k \text{scalare}\} = \langle e_1, \dots, e_n \rangle.$$

Dunque per ogni $\varepsilon > 0$ esiste $n \in \mathbb{N}$ per cui $d(v, V_n) < \varepsilon$ e cioè, posto $\Pi_n = \Pi_{V_n}$, si ha $|v - \Pi_n(v)| < \varepsilon$. Essendo $\Pi_n(v) \in V_n$ si avrà $\Pi_n(v) = \sum_{k=1}^n \lambda^k e_k$, d'altra parte essendo $(v - \Pi_n(v), e_k) = 0$ per $k = 1, \dots, n$ si ottiene $\lambda^k = (\Pi_n(v), e_k) = (v, e_k)$. Dunque la successione

$$v_n = \sum_{k=1}^n (v, e_k) e_k$$

converge a v, cioè

$$v = \sum_{k=1}^{\infty} (v, e_k) e_k.$$

D'altra parte sapendo che anche $|v_n|$ converge a $|v|$ e che

$$|v_n|^2 = \sum_{k=1}^{n} |(v, e_k)|^2$$

si ottiene

$$|v|^2 = \sum_{k=1}^{\infty} |(v, e_k)|^2.$$

$\square$

A questo punto possiamo costruire un esempio concreto di sistema ortonormale completo, che ci permetterà di sviluppare in serie di Fourier le funzioni periodiche.

Consideriamo dunque lo spazio di Hilbert $H = L^2([0, 2\pi], \mathbb{C})$. Vogliamo dimostrare che i vettori

$$e_k(x) = \frac{\mathrm{e}^{ikx}}{\sqrt{2\pi}}$$

per $k \in \mathbf{Z}$ formano un sistema ortonormale completo.

Innanzitutto notiamo che

$$(e_k, e_k) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{ikx} \mathrm{e}^{-ikx} \mathrm{d}x = \frac{1}{2\pi} \int_0^{2\pi} 1 \mathrm{d}x = 1$$

e

$$(e_j, e_k) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{ijx} \mathrm{e}^{-ikx} \mathrm{d}x = \frac{1}{2\pi} \left[\frac{\mathrm{e}^{i(j-k)x}}{i(j-k)} \right]_0^{2\pi} = 0.$$

Poi si sceglie

$$Q_k(t) = c_k \left(\frac{1 + \cos t}{2} \right)^k$$

....

Dunque se $f \in L^2([0, 2\pi])$ si ha

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbf{Z}} c_k e^{ikx} \quad \text{in } L^2$$

con

$$c_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-ikx} \mathrm{d}x$$

e

$$\|f\|_2 = \sum_{k \in \mathbf{Z}} |a_k|^2.$$

Se f è reale (ma tutto funziona anche se f è complessa!) può essere più comodo sviluppare la serie con seni e coseni, notando che il sistema $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}} \cos(kx)$, $\frac{1}{\sqrt{\pi}} \sin(kx)$ è ortonormale e completo ($k = 1, 2, \ldots$).

Dunque posto

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(x) \mathrm{d}x, \\ a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) \mathrm{d}x, \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) \mathrm{d}x, \end{aligned}$$

si ottiene

$$f(x) = \frac{a_0}{\sqrt{2\pi}} + \sum_{k=1}^{\infty} \frac{a_k}{\sqrt{\pi}} \cos(kx) + \sum_{k=1}^{\infty} \frac{b_k}{\sqrt{\pi}} \sin(kx)$$

con

$$\|f\|_2 = |a_0|^2 + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2).$$

Teorema 3.11.3 (convergenza uniforme)**.** *Se $f \in W^{1,2}([0, 2\pi])$ allora la serie di Fourier di f converge ad f uniformemente.*

Teorema 3.11.4 (Dirichlet-Jordan, convergenza puntuale)**.** *Se $f \in BV([0, 2\pi])$, allora*

1. *la serie di Fourier di f converge a f in ogni punto di continuità e a $\frac{f(x^+)+f(x^-)}{2}$ negli altri;*
2. *la convergenza è uniforme sui compatti su cui f è $\mathcal{C}^0$;*
3. *le somme parziali sono uniformemente limitate.*

3.12 Trasformate di Fourier

Per $f \in L^1(\mathbb{R}; \mathbb{C})$ possiamo definire la trasformata di Fourier come:

$$\hat{f}(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-itx} \mathrm{d}x$$

definiamo poi il prodotto di convoluzione come

$$(f * g)(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y) g(y) \mathrm{d}y.$$

Per $\lambda > 0$ definiamo

$$H_\lambda(t) := e^{-|\lambda t|}.$$

Lemma 3.12.1. *Si ha*

$$\hat{H}_\lambda(x) = \sqrt{\frac{2}{\pi}} \frac{\lambda}{\lambda^2 + x^2}$$

e inoltre

$$\frac{1}{\sqrt{2\pi}} \int \hat{H}_\lambda(x) = 1.$$

Dimostrazione. Infatti

$$\begin{aligned} \sqrt{2\pi} H_\lambda(x) &= \int e^{-|\lambda t|} e^{-itx} \mathrm{d}t = \int_0^\infty e^{-t(\lambda+ix)} + e^{-t(\lambda-ix)} \mathrm{d}t \\ &= \left[\frac{-1}{\lambda+ix} + \frac{-1}{\lambda-ix} \right]_0^\infty = \frac{1}{(\lambda+ix)(\lambda-ix)} = \frac{2\lambda}{\lambda^2+x^2}. \end{aligned}$$

Per la seconda affermazione si ha

$$\frac{\pi}{\sqrt{2\pi}} \int \hat{H}_\lambda(x) \mathrm{d}x = \int \frac{\lambda}{\lambda^2 + x^2} \mathrm{d}x = \int \frac{\frac{1}{\lambda}}{1 + \left(\frac{x}{\lambda}\right)^2} \mathrm{d}x = [\arctan(x/\lambda)]_{-\infty}^{+\infty} = \pi.$$

$\square$

Cerchiamo ora di dimostrare che $\bar{\hat{\hat{f}}} = f$ quando $f, \hat{f} \in L^1$.

Visto che direttamente non ci si riesce tentiamo di usare H_λ e $\hat{H}_\lambda$ per approssimare f.

Il primo passo è dato dal seguente

Lemma 3.12.2. *Per $f \in L^1$ si ha*

$$\overline{\widehat{\hat{f} H_\lambda}} = \hat{f} * \hat{H}_\lambda.$$

Dimostrazione. Si ha infatti

$$\begin{aligned}
\overline{\widehat{\hat{f} H_\lambda}}(x) &= \frac{1}{2\pi} \int \overline{\int f(y) e^{-ity} \mathrm{d}y H_\lambda(t) e^{-itx}} \mathrm{d}t = \frac{1}{2\pi} \int \int \overline{f(y)} e^{-it(x-y)} \mathrm{d}y H_\lambda(t) \mathrm{d}t \\
&= \frac{1}{2\pi} \int \int \overline{f(x-y)} e^{-itx} \mathrm{d}y H_\lambda(t) \mathrm{d}t = \frac{1}{2\pi} \int \int H_\lambda(t) e^{-ity} \mathrm{d}t \overline{f(x-y)} \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}} \int \hat{H}_\lambda(y) \overline{f(x-y)} \mathrm{d}y = (\bar{f} * \hat{H}_\lambda)(x).
\end{aligned}$$

□

A questo punto si nota che $H_\lambda(t) \to 1$ per $\lambda \to 0$. Vediamo in che senso $\hat{H}_\lambda(x) \to \delta_0$.

Lemma 3.12.3. *Se $f \in L^p(\mathbb{R}, \mathbb{C})$, $p \in [0, \infty[$, si ha*

$$f * \hat{H}_\lambda \to f$$

in $L^p(\mathbb{R}, \mathbb{C})$, per $\lambda \to 0$.

Dimostrazione. Dal Lemma 3.12.2 sappiamo che $\frac{1}{\sqrt{2\pi}} H_\lambda(x) \mathrm{d}x$ è una misura di probabilità a cui possiamo applicare la disuguaglianza di Jensen. Si ha dunque

$$\begin{aligned}
(2\pi)^{\frac{p}{2}} \int |f * \hat{H}_\lambda - f|^p &= \int \left| \int f(x-y) \hat{H}_\lambda(y) \mathrm{d}y - f(x) \right|^p \mathrm{d}x \\
&= \int \left| \int [f(x-y) - f(x)] \hat{H}_\lambda(y) \mathrm{d}y \right|^p \mathrm{d}x \\
&\leq \int \int |f(x-y) - f(x)|^p \hat{H}_\lambda(y) \mathrm{d}x \mathrm{d}y.
\end{aligned}$$

Chiamando ora $(\tau_y f)(x) := f(x-y)$, sappiamo che $\tau_y \colon L^p \to L^p$ è un operatore continuo. Dunque fissato $\varepsilon > 0$ esiste $\delta > 0$ per cui se $|y| < \delta$ si ha $\|\tau_y f - f\|_p^p < \varepsilon$. D'altra parte, per λ sufficientemente piccolo si ha anche che $\int_{\mathbb{R} \setminus \mathbb{B}_\delta} \hat{H}_\lambda(x) \mathrm{d}x < \varepsilon$.

Dunque, proseguendo le disuguaglianze, si ha

$$\int_{\mathbb{R}} \|\tau_y f - f\|_p^p \hat{H}_\lambda(y) \mathrm{d}y \leq \int_{\mathbb{B}_\delta} \varepsilon \hat{H}_\lambda(y) \mathrm{d}y + \int_{\mathbb{R} \setminus \mathbb{B}_\delta} \|2f\|_p^p H_\lambda(y) \mathrm{d}y \leq \varepsilon + \|2f\|_p^p \varepsilon$$

che può essere reso arbitrariamente piccolo. □

Teorema 3.12.4. *Sia $f \in L^1(\mathbb{R}, \mathbb{C})$, $\cap f \in L^1(\mathbb{R}, \mathbb{C})$. Allora*

$$\bar{\hat{\hat{f}}} = f \qquad (\text{in } L^1).$$

Dimostrazione. Vogliamo applicare il Lemma 3.12.2. Da un lato, siccome $H_\lambda(t) \to 1$ (per $\lambda \to 0$) e $|H_\lambda(t)| \leq e^{-|t|}$ per $\lambda \leq 1$, per convergenza dominata si ha, per ogni x,

$$\lim_{\lambda \to 0} \hat{\overline{\hat{f} H_\lambda}}(x) = \lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} \int \overline{\hat{f}(t) H_\lambda(t)} e^{-itx} \mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int \overline{\hat{f}(t)} e^{-itx} \mathrm{d}t = \hat{\overline{\hat{f}}}(x).$$

D'altro canto, per il Lemma 3.12.3, sappiamo che $\bar{f} * \hat{H}_\lambda \to \bar{f}$ in L^1 e quindi, esiste una successione $\lambda_k \to 0$ su cui si ha convergenza quasi ovunque, cioè per quasi ogni x si ha

$$\lim_{k \to \infty} (\bar{f} * \hat{H}_{\lambda_k})(x) = \bar{f}(x).$$

Mettendo insieme le due cose si ottiene la tesi. $\square$