Diario del Corso di Analisi Funzionale - Mod. I

Corso di Laurea: Matematica - Magistrale
Docente: Sisto Baldo

ATTENZIONE: Il presente Diario del Corso vuole essere un riassunto abbastanza dettagliato di quello che è stato detto in aula, e come tale può essere un utile sussidio per chi voglia sistemare i propri appunti, o per chi sia stato assente e voglia ricostruire i contenuti di una lezione. D'altra parte, queste brevi paginette NON possono sostituire completamente un libro di testo, la lezione in aula o un'interazione diretta con il docente o l'esercitatore: siete quindi invitati a servirvi ANCHE di queste altre opportunità per approfondire le vostre conoscenze!

Indice

1 Lezione del 2/10/2009 (2 ore)

Breve presentazione del programma del corso: contenuti, modalità d'esame, materiale di studio...

La prima parte di questo corso è dedicata all'introduzione della teoria della misura e dell'integrazione secondo Lebesgue. Contestualmente, e con poco o nessuno sforzo aggiuntivo, avremo modo di familiarizzarci anche con la teoria della misura (e dell'integrazione) astratte.

Probabilmente, nei precedenti corsi di Analisi abbiamo tutti incontrato la *misura di Peano-Jordan*, che è probabilmente uno dei modi più semplici di definire in modo rigoroso l'area di un sottinsieme del piano (il volume di un sottinsieme dello spazio...

Ricordiamo alcune definizioni rilevanti:

DEFINIZIONE: Un *intervallo* o *rettangolo* in $\mathbf{R}^n$ è un sottinsieme $I \subset \mathbf{R}^n$ che sia prodotto cartesiano di intervalli unidimensionali: $I = (a_1, b_1) \times (a_2, b_2) \times \ldots \times (a_n, b_n)$. Gli intervalli unidimensionali di cui si fa il prodotto possono essere anche chiusi, oppure chiusi in una sola delle due estremità. La *misura* di un intervallo I è per definizione il numero

$$|I| = \prod_{i=1}^{n} (b_i - a_i).$$

Si vede subito che per $n = 2$ il nostro intervallo è un rettangolo con lati paralleli agli assi, e la sua misura coincide con l'area. Invece, per $n = 3$, I sarà un parallelepipedo e la sua misura coincide con il volume.

Gli insiemi *misurabili secondo Peano-Jordan* sono insiemi la cui area si approssima bene, sia da fuori che da dentro, con unioni finite di intervalli.

DEFINIZIONE (Insieme misurabile secondo Peano-Jordan): Un sottinsieme $A \subset \mathbf{R}^n$ si dice misurabile secondo Peano-Jordan se è limitato e per ogni $\varepsilon > 0$ esistono un numero finito di intervalli $I_1, \ldots, I_N, J_1, \ldots, J_K \subset \mathbf{R}^n$ tali che gli I_i hanno due a due in comune solo punti della frontiera, i J_i hanno due a due in comune solo punti della frontiera,

$$\bigcup_{i=1}^{N} I_i \subset A \subset \bigcup_{i=1}^{K} J_i$$

e infine

$$\sum_{i=1}^{K} |J_i| - \sum_{i=1}^{N} |I_i| \leq \varepsilon.$$

In tal caso, la *misura di Peano-Jordan* di A si definisce come

$$\begin{aligned}|A| = \quad & \sup\{\sum_{i=1}^{N}|I_i| : I_i \textit{ due a due con interni disgiunti}, \bigcup_{i=1}^{N} I_i \subset A\} \\ = \quad & \inf\{\sum_{i=1}^{K}|J_i| : J_i \textit{ due a due con interni disgiunti}, \bigcup_{i=1}^{K} J_i \supset A\}.\end{aligned}$$

È facile vedere che un rettangolo è misurabile secondo Peano-Jordan, mentre l'insieme dei punti a coordinate razionali di un rettangolo non lo è. Nel piano, il *trapezoide* sotteso ad una funzione di una variabile integrabile secondo Riemann è misurabile secondo Peano-Jordan, e la sua misura è data proprio dall'integrale. Sono anche misurabili secondo Peano-Jordan gli insiemi dati dalla parte di piano compresa tra i grafici di due funzioni di una variabile integrabili secondo Riemann:

ESERCIZIO: Siano $g, h : [a, b] \to \mathbf{R}$ due funzioni di una variabile, integrabili secondo Riemann e con $g(x) \le h(x)$ per ogni $x \in [a, b]$. Consideriamo l'insieme $A = \{(x, y) \in \mathbf{R}^2 : \ x \in [a, b], \ g(x) \le y \le h(x)$. Mostrare che A è misurabile secondo Peano-Jordan e si ha

$$|A| = \int_a^b (h(x) - g(x))\, dx.$$

Un insieme A di questo tipo si chiama *semplice rispetto all'asse delle* x... Gli insiemi semplici rispetto all'asse delle y si definiscono in modo analogo, e ci sono anche naturali generalizzazioni in dimensione più alta.

La misura di Peano-Jordan è un ottimo oggetto, che però si comporta male rispetto ad operazioni *numerabili*: se è vero che l'unione di un numero finito di insiemi misurabili secondo P.-J. rimane misurabile, questo non è vero per unioni numerabili (un unione numerabile di *punti* può dare un insieme non misurabile: un esempio è l'insieme dei punti con coordinate razionali in un rettangolo). Per questa ed altre ragioni, risulta utile definire una nozione più generale di misura, che sarà appunto la misura di Lebesgue.

Prima di dare la definizione, ricordiamo un utile fatterello topologico.

La seguente proposizione mostra come qualunque aperto di $\mathbf{R}^n$, comunque complicato, possa essere ottenuto facendo un'unione numerabile di intervalli.

PROPOSIZIONE: Ogni aperto $A \subset \mathbf{R}^n$ è unione numerabile di intervalli aperti.

DIM.: Consideriamo la famiglia $\mathcal{F}$ costituita da tutti i cubi di $\mathbf{R}^n$ del tipo $(q_1 - r, q_1 + r) \times (q_2 - r, q_2 + r) \times \ldots \times (q_n - r, q_n + r)$, dove tutti i q_i ed r sono razionali. Questa è una famiglia numerabile di intervalli.

Mostriamo che A è unione degli elementi della famiglia numerabile di intervalli

$$\mathcal{F}' = \{I \in \mathcal{F} : I \subset A\}.$$

Infatti, poiché A è aperto, per ogni $x \in A$ esiste una palla aperta $B_{r(x)}(x) \subset A$. Dentro questa palla possiamo trovare un cubo centrato in x dentro il quale, grazie alla densità dei razionali, c'è un elemento $I_x \in \mathcal{F}$ che contiene x. Per costruzione, $I_x \in \mathcal{F}'$: abbiamo mostrato che per ogni $x \in A$ c'è un elemento della famiglia numerabile $\mathcal{F}'$ che lo contiene. Dunque $A = \bigcup_{I \in \mathcal{F}'} I$.
Q.E.D.

Siamo ora in grado di definire la *misura esterna di Lebesgue* di un sottinsieme di $\mathbf{R}^n$: l'idea è molto simile a quella della definizione della misura di Peano-Jordan, solo che useremo unioni numerabili anziché unioni finite di intervalli.

DEFINIZIONE (Misura esterna di Lebesgue): Se $A \subset \mathbf{R}^n$, la sua *misura esterna di Lebesgue* si definisce come

$$m(A) = \inf\{\sum_{i=1}^{\infty} |I_i| : \ I_i \textit{ intervalli}, \ \bigcup_{i=1}^{\infty} I_i \supset A\}.$$

Si noti che non richiediamo che gli intervalli abbiano parti interne disgiunte. Inoltre, consideriamo anche l'insieme vuoto come intervallo degenere, in modo da poter considerare anche ricoprimenti finiti.

La misura esterna di Lebesgue gode delle seguenti proprietà elementari:

TEOREMA (Proprietà elementari della misura esterna di Lebesgue): Sia $m : \mathcal{P}(\mathbf{R}^n) \to [0, +\infty]$ *la misura esterna di Lebesgue*[1]*. Valgono i fatti seguenti:*

(i) $m(\emptyset) = 0$, $m(\{x\}) = 0$ *per ogni* $x \in \mathbf{R}^n$.

(ii) Se $A \subset \bigcup_{i=1}^{\infty} A_i$, *con* $A, A_1, A_2, \ldots \subset \mathbf{R}^n$, *allora*

$$m(A) \leq \sum_{i=1}^{\infty} m(A_i)$$

(numerabile subadditività della misura di Lebesgue). In particolare, se $A \subset B$ *vale* $m(A) \leq m(B)$ *(monotonia della misura di Lebesgue).*

[1] $\mathcal{P}(\mathbf{R}^n)$ denota l'insieme delle parti di $\mathbf{R}^n$, ossia l'insieme di tutti i sottinsiemi di $\mathbf{R}^n$.

(iii) Nella definizione della misura esterna di Lebesgue, non è restrittivo chiedere che gli intervalli I_i siano tutti aperti.

(iv) $m(I) = |I|$ per ogni intervallo $I \subset \mathbf{R}^n$. Inoltre, $m(\mathbf{R}^n) = +\infty$.

DIM.: La (i) è lasciata come facile esercizio. Per quanto riguarda la (ii), è importante fare un'osservazione preliminare che ricorre in tutta la teoria della misura: la somma di una serie di numeri non negativi (che ovviamente può essere $+\infty$) non cambia se si permuta l'ordine degli addendi della serie (per esercizio si provi a dimostrare questo fatto, che è falso per le serie a termini di segno qualunque che non siano assolutamente convergenti).

Fissiamo $\varepsilon > 0$ e un indice i: per definizione di inf possiamo trovare una successione di intervalli $\{I^i_j\}_j$ tali che $\bigcup_{j=1}^{\infty} I^i_j \supset A_i$ e

$$\sum_{j=1}^{\infty} |I^i_j| < m(A_i) + \frac{\varepsilon}{2^i}.$$

Allora $\{I^i_j\}_{i,j}$ è un ricoprimento numerabile di A fatto di intervalli, e per definizione di misura di Lebesgue abbiamo

$$m(A) \leq \sum_{i,j=1}^{\infty} |I^i_j| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |I^i_j| \leq \sum_{i=1}^{\infty} (m(A_i) + \frac{\varepsilon}{2^i}) = \sum_{i=1}^{\infty} m(A_i) + \varepsilon,$$

da cui segue (ii) perché ε può essere preso arbitrariamente piccolo.

La monotonia è conseguenza immediata della subadditività numerabile.

Dimostriamo (iii): se $A \subset \mathbf{R}^n$, per ogni $\varepsilon > 0$ possiamo trovare degli intervalli I_j tali che $\bigcup_{j=1}^{\infty} I_j \supset A$ e

$$\sum_{j=1}^{\infty} |I_j| < m(A) + \frac{\varepsilon}{2}.$$

Per ogni $j = 1, 2, \ldots$ sia $I'_j \supset I_j$ un intervallo *aperto* di poco più grande, scelto in modo che $|I'_j| < |I_j| + \frac{\varepsilon}{2^{j+1}}$. Allora

$$\sum_{j=1}^{\infty} |I'_j| < \sum_{j=1}^{\infty} (|I_j| + \frac{\varepsilon}{2^{j+1}}) < m(A) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

e (iii) è dimostrata.

Sorpendentemente, la (iv) è la proprietà più difficile da dimostrare. Grazie alla (iii), essa segue immediatamente dalla seguente

AFFERMAZIONE: Se I è un intervallo, allora per ogni successione di intervalli I_j aperti con $\bigcup_{j=1}^{\infty} I_j \supset I$ si ha

$$(*) \quad |I| \leq \sum_{j=1}^{\infty} |I_j|.$$

La (*) è dimostrabile abbastanza facilmente se gli I_j sono in numero finito, lo è meno nel caso generale di un ricoprimento numerabile. Se però $J \subset I$ è un intervallo *chiuso e limitato*, possiamo usare la compattezza per dire che esiste un numero finito di intervalli $I_1, I_2, \dots, I_N$ del nostro ricoprimento di I tali che $J \subset \bigcup\limits_{j=1}^{N} I_j$. Poiché la (*) è vera per i ricoprimenti finiti, se ne deduce che

$$|J| \leq \sum_{j=1}^{N} |I_j| \leq \sum_{j=1}^{\infty} |I_j|.$$

Poiché la misura di J può essere presa vicina quanto si vuole alla misura di I, (*) risulta dimostrata. Q.E.D.

Come immediata conseguenza del nostro teorema, vediamo che un sottinsieme *numerabile* di $\mathbf{R}^n$ ha misura zero: infatti, un punto di $\mathbf{R}^n$ ha evidentemente misura di Lebesgue zero e la nostra affermazione segue dalla numerabile subadditività.

La misura esterna di Lebesgue è un importante caso particolare di un oggetto più generale, chiamato misura esterna:

DEFINIZIONE (Misura esterna): Una *misura esterna* su un insieme X è una funzione $\mu : \mathcal{P}(\mathbf{R}^n) \to [0, +\infty]$ tale che $\mu(\emptyset) = 0$ e che sia numerabilmente subadditiva: se $A, A_1, A_2, A_3, \ldots \subset X$ e $A \subset \bigcup\limits_{j=1}^{\infty} A_j$, allora

$$\mu(A) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

Dalla numerabile subadditività segue che μ è monotona: se $A \subset B$ allora $\mu(A) \leq \mu(B)$.

Un esempio di misura esterna diversa dalla misura di Lebesgue è la *restrizione* della misura di Lebesgue a un sottinsieme $A_0 \subset \mathbf{R}^n$: questa è la misura $\tilde{m}$ definita da

$$\tilde{m}(A) := m(A \cap A_0).$$

Un altro esempio è la misura δ_0 (*delta di Dirac centrata in* 0), misura su $\mathbf{R}^n$ definita da

$$\delta_0(A) = \begin{cases} 1 & se\ 0 \in A, \\ 0 & altrimenti. \end{cases}$$

Ancora, è una misura esterna la *"misura che conta"* definita da

$$\#(A) = \begin{cases} numero\ degli\ elementi\ di\ A & se\ A\ \grave{e}\ finito, \\ +\infty & altrimenti. \end{cases}$$

In generale, si può dire che la misura di Lebesgue non ha buone proprietà su *tutti* i sottinsiemi di $\mathbf{R}^n$: essa mostra un comportamento assai più simpatico e desiderabile su una particolare classe di insiemi, detti *misurabili*:

DEFINIZIONE (Insiemi misurabili secondo Lebesgue, definizione di Caratheodory): Un sottinsieme $A \subset \mathbf{R}^n$ si dice *misurabile secondo Lebesgue* o *m-misurabile* se vale l'ugualgianza

$$m(T) = m(T \cap A) + m(T \setminus A)$$

per ogni sottinsieme $T \subset \mathbf{R}^n$. In sostanza, chiediamo che A "spezzi bene" la misura di ogni insieme di $\mathbf{R}^n$.

Si noti che grazie alla numerabile subadditività della misura esterna abbiamo sempre $m(T) \leq m(T \cap A) + m(T \setminus A)$: è quindi sufficiente verificare che valga la disuguaglianza opposta

$$m(T) \geq m(T \cap A) + m(T \setminus A) \qquad \forall T \subset \mathbf{R}^n.$$

Analogamente, data una misura esterna μ, A si dice μ-misurabile se $\mu(T) = \mu(T \cap A) + \mu(T \setminus A)$ per ogni $T \subset \mathbf{R}^n$.

OSSERVAZIONE: In seguito ci sarà utile il seguente fatto: se $A \subset \mathbf{R}^n$ è misurabile secondo Lebesgue e $\tilde{m}$ denota la restrizione della misura di Lebesgue ad un *qualunque* insieme $A_0 \subset \mathbf{R}^n$, allora A è anche $\tilde{m}$-misurabile. Se infatti $T \subset \mathbf{R}^n$ abbiamo

$$\tilde{m}(T) = m(T \cap A_0) = m((T \cap A_0) \cap A) + m((T \cap A_0) \setminus A) =$$
$$m((T \cap A) \cap A_0) + m((T \setminus A) \cap A_0) = \tilde{m}(T \cap A) + \tilde{m}(T \setminus A).$$

Questo stesso fatto rimane vero, con identica dimostrazione, anche se m e $\tilde{m}$ vengono sostituite da una generica misura esterna μ e dalla sua restrizione $\tilde{\mu}$ all'insieme A_0.

Il seguente teorema mostra due cose: innanzitutto, se partiamo da insiemi misurabili e facciamo operazioni di unione numerabile, complementazione e intersezione numerabile, rimaniamo sempre nell'ambito degli insiemi misurabili. Inoltre, la misura di Lebesgue, o una qualunque misura esterna μ, se ristrette agli insiemi misurabili hanno buone proprietà, la principale delle quali è la *numerabile additività*: la misura dell'unione di una famiglia numerabile di insiemi due a due disgiunti è uguale alla somma delle loro misure.

TEOREMA (Proprietà degli insiemi misurabili e della misura sugli insiemi misurabili): Sia μ una misura esterna su un insieme X. Valgono i seguenti fatti

(i) Se A è μ-misurabile, allora $A^C = X \setminus A$ è μ-misurabile. Inoltre, se $\mu(A) = 0$ allora A è μ-misurabile.

(ii) Unione o intersezione numerabile di insiemi μ-misurabili è μ-misurabile.

(iii) Se $\{A_i\}_i$ è una famiglia di insiemi μ-misurabili due a due disgiunti e $A = \bigcup_{i=1}^{\infty} A_i$, allora

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$$

(numerabile additività della misura di Lebesgue sui misurabili).

(iv) Se $\{A_i\}$ è una successione crescente di insiemi μ-misurabili, cioè se $A_1 \subset A_2 \subset A_3 \subset \dots$, e $A = \bigcup_{i=1}^{\infty} A_i$ allora

$$\mu(A) = \lim_{i \to +\infty} \mu(A_i).$$

(v) Se $\{A_i\}$ è una successione decrescente di insiemi μ-misurabili, cioè se $A_1 \supset A_2 \supset A_3 \supset \dots$, se $\mu(A_1) < +\infty$ e se infine $A = \bigcap_{i=1}^{\infty} A_i$, allora

$$\mu(A) = \lim_{i \to +\infty} \mu(A_i).$$

DIM.: La (i) è ovvia se si osserva che la condizione di misurabilità può essere riscritta:

$$\mu(T) \geq \mu(T \cap A) + \mu(T \cap A^C) \qquad \forall T \subset X.$$

Dimostreremo il resto la prossima volta!

2 Lezione del 6/10/2009 (2 ore)

Concludiamo la dimostrazione del teorema sulle proprietà degli insiemi misurabili e della misura sui misurabili.

Che un insieme di misura nulla sia misurabile è immediato. Da questo segue in particolare che $\emptyset$ e X sono misurabili.

Mostriamo una versione indebolita di (ii): se A e B sono misurabili, allora $A \cup B$ e $A \cap B$ sono misurabili. Infatti, se $T \subset X$ si ha

$$\begin{aligned}&\mu(T) = \mu(T \cap A) + \mu(T \setminus A) = \\ &\mu((T \cap A) \cap B) + \mu((T \cap A) \setminus B) + \mu((T \setminus A) \cap B) + \mu((T \setminus A) \setminus B).\end{aligned}$$

Si osservi l'ultima riga: l'unione degli insiemi nei primi tre addendi è esattamente $T \cap (A \cup B)$ per cui, per la subadditività della misura, la somma dei primi tre addendi è $\geq \mu(T \cap (A \cup B))$. Invece, l'insieme nell'ultimo addendo non è altro che $T \setminus (A \cup B)$: si ha allora

$$\mu(T) \geq \mu(T \cap (A \cup B)) + \mu(T \setminus (A \cup B)),$$

e $A \cup B$ è misurabile. Da questo e da (i) segue la misurabilità di $A \cap B$ perché $A \cap B = (A^C \cup B^C)^C$. Per induzione, segue anche che unione e intersezione *finita* di insiemi misurabili è misurabile (alle unioni e intersezioni numerabili arriveremo solo alla fine, dopo aver dimostrato tutto il resto!).

Cominciamo a dimostrare (iii): essa è vera per l'unione di *due* insiemi misurabili e disgiunti in quanto $\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B)$. Per induzione, ne deriva che (iii) è vera per l'unione di una famiglia finita di insiemi misurabili due a due disgiunti.

Nel caso generale di una famiglia *numerabile* di insiemi misurabili due a due disgiunti, la numerabile subadditività della misura fornisce $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$, mentre la monotonia assicura che per ogni $N \in \mathbf{N}$

$$\mu(A) \geq \mu\left(\bigcup_{i=1}^{N}(A_i)\right) = \sum_{i=1}^{N} \mu(A_i),$$

dove l'ultima uguaglianza vale per quanto osservato sulle unioni finite di insiemi misurabili disgiunti. Passando al sup su N si ricava

$$\mu(A) \geq \sum_{i=1}^{\infty} \mu(A_i),$$

e (iii) è dimostrata.

Dimostriamo (iv): basta applicare (iii) alla successione di insiemi due a due disgiunti data da $B_1 = A_1$, $B_i = A_i \setminus A_{i-1}$ $(i \geq 2)$. Si ha

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{N \to +\infty} \sum_{i=1}^{N} \mu(B_i) = \lim_{N \to +\infty} \mu(A_N).$$

Dimostriamo (v): Definiamo la successione crescente di insiemi $B_i = A_1 \setminus A_i$, $i = 2, 3, \ldots$. Allora

$$A_1 = A \cup \bigcup_{i=2}^{\infty} B_i$$

e per (iv) si ha

$$\mu(A_1) \leq \mu(A) + \lim_{i \to +\infty} [\mu(A_1) - \mu(A_i)],$$

da cui $\lim\limits_{i \to +\infty} \mu(A_i) \leq \mu(A)$. La disuguaglianza opposta vale per monotonia, per cui la (v) è dimostrata.

A questo punto il teorema è quasi dimostrato: manca solo la (ii).

Sia $A = \bigcup\limits_{i=1}^{\infty} A_i$, con gli A_i tutti misurabili Dobbiamo mostrare che A è misurabile.

Sia $T \subset \mathbf{R}^n$. Consideriamo la successione crescente di insiemi misurabili $B_N := \bigcup\limits_{i=1}^{N} A_i$: essi sono misurabili anche per la misura esterna $\tilde{\mu}$ data dalla restrizione di μ all'insieme T (cioè la misura definita da $\tilde{\mu}(A) := \mu(T \cap A)$ per ogni $A \subset \mathbf{R}^n$). Per la monotonia della misura abbiamo:

$$(***) \quad \mu(T) = \mu(T \cap B_N) + \mu(T \setminus B_N) \geq \mu(T \cap B_N) + \mu(T \setminus A)$$

D'altra parte, per (iv) applicata alla misura esterna $\tilde{\mu}$ abbiamo

$$\lim_{N \to +\infty} \mu(T \cap B_N) = \lim_{N \to +\infty} \tilde{\mu}(B_N) = \tilde{\mu}(A) = \mu(T \cap A)$$

e la misurabilità di A segue passando al limite per $N \to +\infty$ in (***). La misurabilità di $\bigcap\limits_{i=1}^{\infty} A_i$ segue al solito scrivendo

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^C \right)^C .$$

Q.E.D.

La famiglia dei misurabili per una misura esterna forma una *σ-algebra.* Inoltre, una *misura esterna* ristretta ai soli insiemi misurabili è quel che si chiama una *misura*:

DEFINIZIONE (σ-algebra, misura): Dato un insieme X, una famiglia $\mathcal{A} \subset \mathcal{P}(X)$ di sottinsiemi di X si dice σ-algebra se $X \in \mathcal{A}$, $A^C \in \mathcal{A}$ ogni volta che $A \in \mathcal{A}$, e se $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ ogniqualvolta $A_i \in \mathcal{A}$ per ogni $i = 1, 2, \ldots$.

Dato X ed una σ-algebra $\mathcal{A}$ su X, una misura è una funzione $\mu : \mathcal{A} \to [0, +\infty]$ tale che $\mu(\emptyset) = 0$ e che sia numerabilmente additiva: se $A_1, A_2, A_3, \ldots$ sono elementi di $\mathcal{A}$ due a due disgiunti, allora

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Dopo questo risultato generale valido per tutte le misure esterne, torniamo allo specifico della misura di Lebesgue: il seguente teorema mostra che gli insiemi misurabili secondo Lebesgue abbondano.

TEOREMA (Regolarità della misura di Lebesgue): I sottinsiemi aperti e i sottinsiemi chiusi di $\mathbf{R}^n$ sono misurabili secondo Lebesgue. Inoltre, se A è un insieme misurabile secondo Lebesgue, allora per ogni $\varepsilon > 0$ esistono B aperto e C chiuso, con $C \subset A \subset B$ e $m(B \setminus C) < \varepsilon$.

DIM.: È un esercizio relativamente semplice verificare che gli intervalli sono insiemi misurabili secondo Lebesgue: un intervallo si ottiene come intersezione finita di *semispazi.* A sua volta, un semispazio S è misurabile secondo Lebesgue: se T è un insieme test, fissiamo $\varepsilon > 0$ e sia $\{I_i\}$ una famiglia numerabile di intervalli che ricopre T tale che $\sum_{i=1}^{\infty} |I_i| < m(T) + \varepsilon$. Definiamo poi $I'_i = I_i \cap S$, $I''_i = I_i \cap (\mathbf{R}^n \setminus S)$: questi sono ancora intervalli (eventualmente vuoti), la somma delle cui misure è esattamente $|I_i|$. Inoltre, la famiglia $\{I'_i\}$ ricopre $T \cap S$, mentre $\{I''_i\}$ ricopre $T \cap S^C$: dunque

$$m(T) + \varepsilon > \sum_{i=1}^{\infty} |I'_i| + \sum_{i=1}^{\infty} |I''_i| \geq m(T \cap S) + m(T \cap S^C)$$

e la misurabilità di S segue perché ε è arbitrario.

Di conseguenza gli intervalli sono misurabili, e lo sono anche gli aperti perché possono essere ottenuti come unione numerabile di intervalli.

I chiusi sono misurabili perché i loro complementari sono aperti e quindi misurabili.

Sia ora A misurabile, $\varepsilon > 0$: mostriamo che esiste un aperto $B \supset A$ con $m(B \setminus A) < \varepsilon/2$. Supponiamo dapprima che A abbia misura finita. Per definizione di misura di Lebesgue, possiamo trovare una famiglia numerabile di intervalli $I_1, I_2, \dots$ con $\bigcup\limits_{i=1}^{\infty} I_i \supset A$ e $\sum\limits_{i=1}^{\infty} |I_i| \leq m(A) + \varepsilon/2$. Abbiamo già visto che non è restrittivo supporre che gli I_i siano tutti aperti. Se $B = \bigcup\limits_{i=1}^{\infty} I_i$, allora B è aperto e per subadditività

$$m(B) \leq \sum_{i=1}^{\infty} m(I_i) \leq m(A) + \varepsilon/2,$$

da cui $m(B \setminus A) = m(B) - m(A) \leq \varepsilon/2$.

Mostriamo che anche un insieme misurabile A con $m(A) = +\infty$ si approssima "da fuori" con insiemi aperti: prendiamo $\varepsilon > 0$ e mostriamo che esiste $B \supset A$, B aperto, tale che $m(B \setminus A) < \varepsilon$.

A tal fine consideriamo gli insiemi misurabili $A_N = A \cap B_N(0)$, $N = 1, 2, \dots$: essi hanno tutti misura finita e la loro unione è A. Per ciascuno di questi possiamo trovare $B_N \supset A_N$, B_N aperto tale che $m(B_N \setminus A_N) < \frac{\varepsilon}{2^{N+1}}$: definiamo $B = \bigcup\limits_{N=1}^{\infty} B_N$.

Ora, B è un aperto che contiene A, e inoltre $B \setminus A \subset \bigcup\limits_{N=1}^{\infty} (B_N \setminus A_N)$: per subadditività numerabile ricaviamo $m(B \setminus A) \leq \sum\limits_{N=1}^{\infty} m(B_N \setminus A_N) < \frac{\varepsilon}{2}$.

Mostriamo infine che dato A misurabile e $\varepsilon > 0$, esiste un chiuso $C \subset A$ con $m(A \setminus C) < \varepsilon/2$: questo concluderà la nostra dimostrazione. A questo fine, scegliamo un aperto $F \supset A^C$ tale che $m(F \setminus A^C) < \varepsilon/2$. Allora $C = F^C$ è un chiuso, $C \subset A$, e $m(A \setminus C) = m(F \setminus A^C) < \varepsilon/2$. Q.E.D.

3 Lezione del 7/10/2009 (2 ore)

Proponiamo subito un esercizio collegato al teorema sulla regolarità della misura di Lebesgue:

ESERCIZIO (Involucro boreliano): Ricordo che la *sigma-algebra di Borel* in $\mathbf{R}^n$ (o in uno spazio topologico) è, per definizione, la più piccola σ-algebra che contiene gli aperti. Gli elementi della σ-algebra di Borel si chiamano *boreliani*. Mostrare che dato un insieme misurabile secondo Lebesge, esistono dei boreliani B, C tali che $C \subset A \subset B$ e $m(B \setminus C) = 0$. Di conseguenza, un insieme misurabile secondo Lebesge si può sempre scrivere come unione di un boreliano e di un insieme di misura nulla.

Nonostante vi siano moltissimi insiemi misurabili secondo Lebesgue, non tutti i sottinsiemi di $\mathbf{R}^n$ lo sono:

ESEMPIO (Insieme non misurabile di Vitali): Mettiamoci nel caso $n = 1$, e consideriamo l'intervallo $(0,1) \subset \mathbf{R}$. Definiamo la seguente relazione di equivalenza su $(0,1)$: diciamo che $x \sim y$ se e solo se $x - y \in \mathbf{Q}$. La nostra relazione di equivalenza partiziona l'intervallo $(0,1)$ in infinite classi di equivalenza: definiamo un insieme A che contenga esattamente un elemento per ogni classe di equivalenza[2]. Mostreremo che l'insieme A non è misurabile secondo Lebesgue.

Per ogni $q \in \mathbf{Q} \cap [0,1)$ definiamo gli insiemi $A_q = \{x + q : \ x \in A\}$. Siccome la misura di Lebesgue è invariante per traslazione (questo è ovvio per come è definita: la misura di un intervallo è invariante per traslazione!) abbiamo che $m(A_q) = m(A)$. Poiché gli intervalli sono misurabili secondo Lebesgue abbiamo anche $m(A) = m(A_q) = m(A_q \cap (0,1)) + m(A_q \setminus (0,1))$. Se $B_q = A_1 \setminus (0,1)$, definiamo $\tilde{B}_q = \{x : \ x + 1 \in B_q\}$: evidentemente $m(\tilde{B}_q) = m(B_q)$ per l'invarianza per traslazioni della misura di Lebesgue.

Definiamo infine $\tilde{A}_q = (A_q \cap (0,1)) \cup \tilde{B}_q$. Per quanto visto sopra, abbiamo $m(\tilde{A}_q) = m(A)$. Ora, è facile vedere che gli insiemi A_q sono due a due disgiunti al variare di $q \in \mathbf{Q} \cap [0,1)$ e che $\bigcup_q \tilde{A}_q = (0,1)$. Se A fosse misurabile, lo sarebbero anche gli insiemi $\tilde{A}_q$ e per additività numerabile avremmo

$$1 = m([0,1)) = \sum_{n=1}^{\infty} m(\tilde{A}_q) = \sum_{n=1}^{\infty} m(A).$$

Questo è assurdo: infatti la misura di A è nulla oppure positiva. Se fosse $m(A) = 0$, l'espressione di destra varrebbe 0, mentre se fosse $m(A) > 0$ essa varrebbe $+\infty$: in tutti e due i casi essa non può essere uguale a 1. Dunque A non è misurabile secondo Lebesgue.

In vista della definizione dell'integrale di Lebesgue, occorre definire un'importante classe di funzioni: le funzioni *misurabili.*

DEFINIZIONE (funzione misurabile): Sia $A \subset X$, $f : A \to \overline{\mathbf{R}}$. Il simbolo $\overline{\mathbf{R}}$ denota l'insieme $\mathbf{R} \cup \{+\infty\} \cup \{-\infty\}$: in questo contesto, in futuro useremo la "strana" convenzione che $0 \cdot \pm\infty = 0$, mentre la somma $+\infty - \infty$ rimarrà non definita, come è giusto che sia!

La funzione f si dice *misurabile* (rispetto ad una fissata misura esterna, per esempio la misura di Lebesgue su $\mathbf{R}^n$) se per ogni $a \in \mathbf{R}$ gli insiemi $f^{-1}((a, +\infty]) = \{x \in A : f(x) > a\}$ sono misurabili.

Una caratterizzazione equivalente della misurabilità, di sapore un po' più topologico, è data dalla seguente

[2]Per poter definire questo insieme, dobbiamo assumere la validità dell'assioma della scelta!

PROPOSIZIONE (Caratterizzazione delle funzioni misurabili): Una funzione $f : A_0 \to \mathbf{R} \cup \{+\infty\}$ (con $A_0 \subset X$) è misurabile se e solo se $f^{-1}(\{+\infty\})$, $f^{-1}(\{-\infty\})$ sono misurabili e $f^{-1}(U)$ è misurabile per ogni aperto $U \subset \mathbf{R}$.

DIM.: Se sappiamo che $f^{-1}(\{+\infty\})$, $f^{-1}(\{-\infty\})$ sono misurabili e $f^{-1}(U)$ è misurabile per ogni aperto $U \subset \mathbf{R}$, allora f è misurabile perché $f^{-1}((a,+\infty]) = f^{-1}((a,+\infty)) \cup f^{-1}(\{+\infty\})$.

Viceversa, supponiamo che f sia misurabile. Possiamo scrivere

$$f^{-1}(\{+\infty\}) = \bigcap_{N=1}^{\infty} f^{-1}((N,+\infty]),$$

per cui $f^{-1}(\{+\infty\})$ è misurabile in quanto intersezione numerabile di misurabili.

Dall'ipotesi di misurabilità di f segue allora che $f^{-1}((a,+\infty))$ è misurabile per ogni $a \in \mathbf{R}$. Dimostriamo che anche gli insiemi $f^{-1}([a,+\infty))$, $f^{-1}((-\infty,a))$ e $f^{-1}((-\infty,a])$ sono tutti misurabili per ogni $a \in \mathbf{R}$. Infatti, $f^{-1}([a,+\infty)) = \bigcap_{N=1}^{\infty} f^{-1}((a-\frac{1}{N},+\infty))$ è misurabile in quanto intersezione numerabile di misurabili. Le controimmagini di semirette "sinistre" del tipo $f^{-1}([-\infty,a))$ e $f^{-1}([-\infty,a])$ sono misurabili in quanto sono complementari di controimmagini di semirette "destre". Ne segue che $f^{-1}(\{-\infty\})$ è misurabile: $f^{-1}(\{-\infty\}) = \bigcap\limits_{N=1}^{\infty} f^{-1}([-\infty,-N])$...e sono misurabili anche le controimmagini di semirette "sinistre" senza $-\infty$.

Allora, anche le controimmagini di intervalli aperti sono misurabili, infatti $f^{-1}((a,b)) = f^{-1}((-\infty,b)) \cap f^{-1}(a,+\infty))$. Se poi $U \subset \mathbf{R}$ è aperto, scriviamo $U = \bigcup\limits_{i=1}^{\infty} I_i$, con $I_i \subset \mathbf{R}$ intervalli aperti. Allora $f^{-1}(U) = \bigcup\limits_{i=1}^{\infty} f^{-1}(I_i)$ è misurabile. Q.E.D.

Come conseguenza di questo teorema, il dominio A di f deve necessariamente ssere misurabile (perché $A = f^{-1}(\mathbf{R}) \cup f^{-1}(\{+\infty\}) \cup f^{-1}(\{-\infty\})$. Inoltre, una funzione continua a valori reali, definita su un aperto di $\mathbf{R}^n$, è certamente misurabile secondo Lebesgue. Perché?

Le funzioni misurabili sono "stabili" per tutta una serie di operazioni algebriche e di limite:

PROPOSIZIONE (Stabilità delle funzioni misurabili): Supponiamo che f, g siano misurabili, $\lambda \in \mathbf{R}$ e che $\{f_n\}$ sia una successione di funzioni misurabili. Allora

(i) l'insieme $\{x : \ f(x) > g(x)\}$ è misurabile;

(ii) se $\phi : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ è continua, allora $\phi \circ f$ è misurabile (sul suo dominio);

(iii) le funzioni $f+g$, λf, $|f|$, $\max\{f,g\}$, $\min\{f,g\}$ *e* fg *sono tutte misurabili nel loro dominio;*

(iv) le funzioni $\sup f_n$, $\inf f_n$, $\limsup f_n$, $\liminf f_n$ *e* $\lim f_n$ *sono tutte misurabili nel loro dominio.*

DIM.: Per verificare la (i), osserviamo che se $f(x) > g(x)$, allora esiste un razionale q compreso tra $g(x)$ e $f(x)$. Allora il nostro asserto è vero in quanto possiamo scrivere

$$\{x: \ f(x) > g(x)\} = \bigcup_{q\in\mathbf{Q}} \left(f^{-1}((q,+\infty]) \cap g^{-1}([-\infty,q))\right),$$

per cui abbiamo espresso il nostro insieme come unione numerabile di insiemi misurabili.

La (ii), nel caso di funzioni a valori reali, è ovvia grazie alla nostra caratterizzazione delle funzioni misurabili: sappiamo infatti che la controimmagine di un aperto secondo ϕ è un aperto. Nel caso di funzioni a valori reali estesi, occorre precisare cosa vuol dire che Φ è continua: significa che la controimmagine di ogni aperto di $\overline{\mathbf{R}}$ è un aperto in $\overline{\mathbf{R}}$. A loro volta, gli aperti di $\overline{\mathbf{R}}$ sono gli insiemi che si possono ottenere prendendo unioni (è sufficiente prenderle numerabili) di intervalli aperti di $\mathbf{R}$ e di semirette "intorno di $\pm\infty$", cioè del tipo $(a,+\infty]$ oppure $[-\infty,a)$. È allora un semplice esercizio verificare che la composizione è ancora misurabile.

Vediamo la (iii): siano f, g misurabili e consideriamo la funzione somma $f+g$ (essa è definita sull'intersezione dei domini, privata dei punti in cui la somma si presenta nella forma $+\infty-\infty$ o $-\infty+\infty$). Essa è misurabile in quanto

$$(f+g)^{-1}((a,+\infty]) = \{x : f(x) > a - g(x)\}$$

è misurabile grazie a (i): la funzione $a-g(x)$ è infatti banalmente misurabile. Da (ii) segue poi la misurabilità di λf, di $|f|$ e di f^2 (che si ottengono da f componendo con una funzione continua). Se f, g sono a valori reali possiamo poi scrivere $\max\{f(x),g(x)\} = \frac{1}{2}(f(x)+g(x)+|f(x)-g(x)|)$, $\min\{f(x),g(x)\} = \frac{1}{2}(f(x)+g(x)-|f(x)-g(x)|)$, $f(x)g(x) = \frac{1}{2}((f(x)+g(x))^2 - f^2(x) - g^2(x))$, il che ci fornisce la misurabilità di $\max\{f,g\}$, $\min\{f,g\}$ e fg. Nel caso generale di funzioni a valori reali estesi, il ragionamento appena fatto ci fornisce la misurabilità della restrizione delle funzioni che ci interessano all'insieme, evidentemente misurabile, dove sia f che g sono finite.

Tutto il resto è facilmente decomponibile in pochi pezzi misurabili, su ciascuno dei quali le funzioni in esame sono costanti: per esempio, $f(x)g(x)$ vale identicamente $+\infty$ sull'insieme (misurabile) $\{x \in X : \ f(x) = +\infty, \ g(x) > 0\}$, vale 0 sull'insieme $\{x \in X : \ f(x) = +\infty, \ g(x) = 0\}$, etc. In conclusione, se ne deduce facilmente che la funzione prodotto è misurabile.

Per quanto riguarda la (iv), sia $f(x) = \sup\{f_n(x) : n = 1, 2, \ldots\}$. Si ha $f^{-1}((a, +\infty)) = \bigcup_n f_n^{-1}((a, +\infty))$, per cui f è misurabile essendolo le f_n. Analogamente, $\inf_n f_n(x)$ è misurabile.

La funzione $\liminf_{n\to+\infty} f_n(x)$ è misurabile in quanto $\liminf_{n\to+\infty} f_n(x) = \sup_n \inf\{f_m(x) : m \geq n\}$. Analogamente, $\limsup_{n\to+\infty} f_n(x)$ è misurabile. L'insieme dove le due funzioni misurabili $\liminf_n f_n$ e $\limsup_n f_n$ coincidono è misurabile: tale insieme è proprio quello in cui esiste $\lim_n f_n$, che quindi è misurabile. Q.E.D.

Un'importante sottoclasse delle funzioni misurabili è quella delle *funzioni semplici*: nella definizione di integrale di Lebesgue esse giocheranno lo stesso ruolo che le funzioni a scala avevano in quella dell'integrale di Riemann.

Ricordiamo che, dato $A \subset X$, la sua *funzione caratteristica* è la funzione

$$\mathbf{1}_A(x) = \begin{cases} 1 & \textit{se } x \in A, \\ 0 & \textit{se } x \notin A. \end{cases}$$

DEFINIZIONE: Una *funzione semplice* $\phi : \mathbf{R}^n \to \mathbf{R}$ è una combinazione lineare finita di funzioni caratteristiche di insiemi misurabili. In altre parole, ϕ è semplice se esistono un numero finito di insiemi misurabili $A_1, A_2, \ldots, A_N$ e dei numeri reali $c_1, c_2, \ldots, c_N$ tali che $\phi(x) = \sum_{i=1}^{N} c_i \mathbf{1}_{A_i}(x)$. Evidentemente, non è restrittivo supporre che gli A_i siano due a due disgiunti. In modo equivalente, possiamo dire che una funzione semplice è una funzione *misurabile* la cui immagine è un *insieme finito.*

Se $\phi(x) \geq 0$ per ogni x, definiamo in modo naturale l'integrale (di Lebesgue) di ϕ rispetto alla misura μ come

$$\int_{\mathbf{R}^n} \phi(x)\, dx = \sum_{i=1}^{N} c_i \ \mu(A_i).$$

Osserviamo che una funzione a scala è una funzione semplice in cui gli insiemi A_i sono intervalli. Per questo tipo di funzioni, la nuova definizione di integrale coincide con la precedente. Inoltre, non è difficile vedere che l'integrale sulle funzioni semplici gode delle usuali proprietà di linearità e monotonia.

Come vedremo, l'integrale di Lebesgue di una funzione misurabile non negativa f si definisce in maniera del tutto analoga all'integrale (inferiore) di Riemann, sostituendo le funzioni a scala con le funzioni semplici: $\int f(x)\, d\mu(x) = \sup\{\int \phi(x)\, d\mu(x) : \ \phi \textit{ semplice}, \ \phi \leq f\}$.

Tuttavia, per provare che quest'oggetto gode di tutte le proprietà che ci aspettiamo, sarà necessario provare un risultato di approssimazione: il prossimo, fondamentale teorema dice che ogni funzione misurabile *non negativa* può essere approssimata da sotto con una successione di funzioni semplici:

TEOREMA (Approssimazione di funzioni misurabili con funzioni semplici): Sia $f : X \to [0, +\infty]$ *una funzione misurabile. Allora esiste una successione* $\phi_k : X \to [0, +\infty)$ *di funzioni semplici tali che* $f \geq \phi_{k+1} \geq \phi_k$ *(*$k = 1, 2, 3, \ldots$*) e tali che*

$$\lim_{k \to +\infty} \phi_k(x) = f(x) \quad \forall x \in X.$$

DIM.: Per ogni fissato $k = 1, 2, \ldots$ e $j = 0, 2, \ldots, k2^k - 1$ definiamo gli insiemi misurabili $E_{k,j} = f^{-1}([\frac{j}{2^k}, \frac{j+1}{2^k}))$, mentre poniamo $E_{k,k2^k} = f^{-1}([k2^k, +\infty])$.

Consideriamo poi le funzioni semplici

$$\phi_k(x) = \sum_{j=0}^{k2^k} \frac{j}{2^k} \mathbf{1}_{E_{k,j}}(x).$$

Per costruzione, queste funzioni sono misurabili e sono tutte minori o uguali a f. Inoltre, esse formano una successione crescente: basta osservare che per ogni k e per ogni $j = 1, \ldots, k2^k - 1$ si ha $E_{k,j} = E_{k+1,2j} \cup E_{k+1,2j+1}$. Quanto poi a $E_{k,k2^k}$, questo verrà suddiviso al passo successivo in $2^k + 1$ insiemi...

È poi facile vedere che $\phi_k(x) \to f(x)$ per ogni x: se $f(x) < +\infty$, per k abbastanza grande si ha $f(x) - \phi_k(x) \leq 2^{-k}$, mentre se invece $f(x) = +\infty$ si ha $x \in E_{k,2^k}$ per ogni k e quindi $\phi_k(x) = k \to +\infty$. Q.E.D.

4 Lezione del 8/10/2009 (2 ore)

È finalmente giunto il momento di introdurre l'integrale di Lebesgue di una funzione misurabile non negativa: la definizione è quella anticipata ieri.

DEFINIZIONE: L'*integrale di Lebsegue* di una funzione misurabile $f : X \to [0, +\infty]$ si definisce come

$$\int_X f(x)\, d\mu(x) = \sup\{\int_X \phi(x)\, d\mu(x) : \ \phi \textit{ semplice}, \ \phi \leq f\}.$$

Se poi $f : A \to [0, +\infty]$ è misurabile, definiamo $\int_A f(x)\, d\mu(x)$ come $\int_X \tilde{f}(x)\, d\mu(x)$, dove $\tilde{f} : X \to [0, +\infty]$ si ottiene estendendo f ponendola uguale a 0 fuori da A.

Il prossimo risultato di convergenza integrale si rivelerà importantissimo per la teoria dell'integrale di Lebesgue, grazie anche al risultato di approssimazione con funzioni semplici che abbiamo dimostrato prima.

TEOREMA (di Beppo Levi o della convergenza monotona): Sia $\{f_k\}$ *una successione di funzioni misurabili non negative,* $f_k : X \to [0, +\infty]$*, e supponiamo che la successione sia anche crescente:* $f_{k+1}(x) \geq f_k(x)$ *per ogni* $x \in X$ *e per ogni* $k = 1, 2, 3, \ldots$*. Allora, se* $f(x) = \lim_{k\to+\infty} f_k(x)$*, si ha*

$$\int_{\mathbf{R}^n} f(x)\, d\mu(x) = \lim_{k\to+\infty} \int_{\mathbf{R}^n} f_k(x)\, d\mu(x).$$

DIM.: Notiamo innanzitutto che la funzione f è misurabile in quanto limite (sup) di funzioni misurabili. Inoltre, la successione $k \mapsto \int_X f_k(x)\, d\mu(x)$ è crescente: indichiamo con α il suo limite. Evidentemente, essendo $f \geq f_k$ per ogni k, si ha $\int_X f(x)\, d\mu(x) \geq \alpha$: in particolare, se $\alpha = +\infty$, il teorema è dimostrato. Se invece $\alpha \in \mathbf{R}$, ci rimane da dimostrare la disuguaglianza opposta

$$\int_X f(x)\, d\mu(x) \leq \alpha.$$

A tal fine, fissiamo $c \in (0, 1)$ e una funzione semplice $s : X \to [0, +\infty)$ con $s \leq f$. La funzione semplice s può essere scritta $s(x) = \sum_{j=1}^N s_j \mathbf{1}_{A_j}(x)$, con A_j insiemi misurabili due a due disgiunti. Definiamo $E_k = \{x \in X : f_k(x) \geq cs(x)\}$. Grazie al fatto che le f_k tendono a f e che $c < 1$, abbiamo che $\bigcup_{k=1}^{\infty} E_k = X$, e inoltre la successione di insiemi misurabili E_k è crescente perché lo sono le f_k. Definiamo poi $A_{j,k} = A_j \cap E_k$: grazie alla continuità della misura sulle successioni crescenti abbiamo $\mu(A_{j,k}) \to \mu(A_j)$ per $k \to +\infty$. Allora:

$$\alpha = \lim_{k\to+\infty} \int_X f_k(x)\, d\mu(x) \geq$$

$$\lim_{k\to+\infty} \int_{E_k} f_k(x)\, d\mu(x) \geq \lim_{k\to+\infty} \int_{E_k} c\, s(x)\, d\mu(x) =$$

$$\lim_{k\to+\infty} c \sum_{j=1}^N s_j\, \mu(A_{j,k}) = c \sum_{j=1}^{\infty} s_j\, \mu(A_j) = c \int_X s(x)\, d\mu(x).$$

Passando al sup su tutte le funzioni semplici $s \leq f$ e su tutti i $c < 1$, si ottiene la disuguaglianza che ci mancava. Q.E.D.

Vediamo subito alcune importanti conseguenze del teorema di Beppo Levi:

(i) *Additività dell'integrale rispetto alla funzione integranda:* Siano $f, g : X \to [0, +\infty]$ funzioni misurabili. Allora

$$\int_X (f(x) + g(x))\, d\mu(x) = \int_X f(x)\, d\mu(x) + \int_X g(x)\, d\mu(x).$$

Infatti possiamo trovare due successioni crescenti di funzioni semplici, $\{s_k\}$, $\{u_k\}$ con $s_k \to f$, $u_k \to g$. L'integrale delle funzioni semplici è evidentemente additivo: il teorema di Beppo-Levi ci consente di passare al limite e ottenere l'identità voluta.

(i) *Numerabile additività dell'integrale rispetto all'insieme di integrazione:* Se $\{A_i\}$ è una successione di insiemi misurabili due a due disgiunti e f è una funzione misurabile non negativa definita su $A = \bigcup\limits_{i=1}^{\infty} A_i$, allora

$$\int\limits_A f(x)\, d\mu(x) = \sum_{i=1}^{\infty} \int\limits_{A_i} f(x)\, d\mu(x).$$

Basta infatti considerare la successione crescente di funzioni misurabili $g_k(x) = \sum\limits_{i=1}^{k} f(x)\, \mathbf{1}_{A_i}(x)$, che converge alla funzione $g(x) = f(x)\, \mathbf{1}_A(x)$. Un'applicazione del teorema di Beppo Levi dimostra subito la tesi.

(iii) *Integrazione per serie:* Se $\{f_k\}$ è una successione di funzioni misurabili non negative definite su A, allora

$$\int\limits_A \sum_{i=1}^{\infty} f_k(x)\, d\mu(x) = \sum_{i=1}^{\infty} \int\limits_A f_k(x)\, d\mu(x).$$

Basta applicare il teorema di Beppo Levi e l'additività dell'integrale rispetto alla funzione integranda alle somme parziali della serie.

DEFINIZIONE (Integrale di funzioni di segno qualunque): Che fare se abbiamo una funzione misurabile di *segno qualunque* $f : A \to \overline{\mathbf{R}}$? Definiamo la parte positiva e la parte negativa di f nel modo seguente:

$$f^+(x) := \max\{0, f(x)\}, \qquad f^-(x) := -\min\{0, f(x)\}.$$

Evidentemente, si ha $f(x) = f^+(x) - f^-(x)$ e $|f(x)| = f^+(x) + f^-(x)$. Se gli integrali di f^+ e f^- non sono *entrambi* $+\infty$, f si dice *integrabile secondo Lebesgue* e definiamo

$$\int_A f(x)\, d\mu(x) := \int_A f^+(x)\, d\mu(x) - \int_A f^-(x)\, d\mu(x).$$

Se poi i due integrali della parte positiva e della parte positiva sono *entrambi finiti*, allora $f(x)$ si dice *sommabile*, ed ha integrale finito. Evidentemente, una funzione misurabile f è sommabile se e solo se il suo modulo ha integrale finito.

Si noti che le proprietà di additività rispetto alla funzione integranda, e di numerabile additività rispetto all'insieme di integrazione, continuano a valere per funzioni sommabili. In particolare, poiché è evidente dalla definizione che le costanti si possono "portare fuori dall'integrale", l'integrale di Lebesgue è lineare sullo spazio vettoriale delle funzioni sommabili.

Vediamo gli enunciati di altri due celebri teoremi di convergenza sotto il segno di integrale: il Lemma di Fatou e il Teorema della convergenza dominata. La dimostrazione arriverà la prossima volta!

TEOREMA (Lemma di Fatou): Sia $f_k : X \to [0, +\infty]$ una successione di funzioni misurabili non negative, $f(x) = \liminf_{k\to+\infty} f_k(x)$. Allora

$$\int_X f(x)\, d\mu(x) \leq \liminf_{k\to+\infty} \int_X f_k(x)\, d\mu(x).$$

Un paio di osservazioni: il lemma di Fatou è in generale falso per funzioni di segno qualunque. Si prenda per esempio $A = \mathbf{R}$, $\mu = m$ (misura di Lebesgue), $f_k(x) = -1/k$ (funzioni costanti). Allora $f_k(x) \to 0$, ma

$$\int_{\mathbf{R}} f_k(x)\, dx = -\infty, \qquad \int_{\mathbf{R}} 0\, dx = 0^3.$$

Siccome la nostra successione di costanti cresce, lo stesso esempio mostra che il teorema di Beppo Levi non vale per funzioni di segno qualunque. Infine, le stesse funzioni *cambiate di segno* mostrano che nella tesi del Lemma di Fatou può valere la disuguaglianza stretta.

Probabilmente il più celebre risultato di convergenza integrale nel quadro della teoria di Lebesgue è il seguente:

[3] Quando la misura è quella di Lebesgue, si usa scrivere dx anziché $dm(x)$

TEOREMA (Della convergenza dominata di Lebesgue): Sia $f_k : X \to \overline{\mathbf{R}}$ una successione di funzioni misurabili, e supponiamo che esista una funzione sommabile $\phi : X \to [0, +\infty]$ tale che $|f_k(x)| \leq \phi(x)$ per ogni k e per ogni x. Se esiste il limite $f(x) = \lim_{k\to+\infty} f_k(x)$, allora

$$\lim_{k\to+\infty} \int_X |f_k(x) - f(x)| \, d\mu(x) = 0,$$
$$\int_X f(x) \, d\mu(x) = \lim_{k\to+\infty} \int_X f_k(x) \, d\mu(x).$$

5 Lezione del 14/10/2009 (2 ore)

DIM. del Lemma di Fatou: Sappiamo già che f è misurabile non negativa. Si ha $f(x) = \lim_{k\to+\infty} g_k(x)$, dove $g_k(x) = \inf\{f_h(x) : \ h \geq k\}$. Poiché le g_k sono una successione crescente di funzioni misurabili non negative abbiamo per Beppo Levi

$$\int_X f(x) \, d\mu(x) = \lim_{k\to+\infty} \int_X g_k(x) \, d\mu(x).$$

La tesi segue allora grazie alla monotonia dell'integrale di Lebesgue, poiché si ha evidentemente $g_k(x) \leq f_k(x)$. Q.E.D.

DIM. del Teorema della convergenza dominata: La funzione limite f è misurabile, ed è anche sommabile perché il suo modulo è dominato da ϕ. Inoltre, $|f_k(x) - f(x)| \leq |f_k(x)| + |f(x)| \leq 2\phi(x)$. Ne segue che la successione di funzioni $2\phi(x) - |f_k(x) - f(x)|$ è non negativa e tende puntualmente alla funzione sommabile $2\phi(x)$. Dal Lemma di Fatou segue allora che

$$\liminf_{k\to+\infty} \int_X (2\phi(x) - |f_k(x) - f(x)|) \, d\mu(x) \geq \int_X 2\phi(x) \, d\mu(x),$$

da cui semplificando l'integrale di $2\phi(x)$:

$$\limsup_{k\to+\infty} \int_X |f_k(x) - f(x)| \, d\mu(x) \leq 0,$$

che è la prima parte della tesi. La seconda parte segue perchè

$$\left| \int_X f_k(x) \, d\mu(x) - \int_X f(x) \, d\mu(x) \right| \leq \int_X |f_k(x) - f(x)| \, d\mu(x).$$

Q.E.D.

Per convincerci fino in fondo della ragionevolezza della teoria di Lebesgue, è estremamente importante fare il confronto tra l'integrale di Riemann e quello di Lebesgue.

Prima di farlo, è utile introdurre una comoda terminologia: si dice che una certa proprietà è vera *per quasi ogni* $x \in \mathbf{R}^n$ *(o q.o.* $x \in A$*, con* A *misurabile)* se l'insieme degli x per cui la proprietà è falsa ha misura nulla. Per esempio, date due funzioni $f, g : \mathbf{R}^n \to \overline{\mathbf{R}}$, diremo che esse sono *quasi ovunque uguali* se $\mu(\{x : \ f(x) \neq g(x)\}) = 0$.

È un semplice esercizio verificare che una funzione quasi ovunque uguale ad una funzione misurabile è essa stessa misurabile (infatti tutti gli insiemi di misura nulla sono misurabili). Inoltre, due funzioni quasi ovunque uguali hanno lo stesso integrale . Vale anche il seguente risultato:

PROPOSIZIONE: Se $f : A \to [0, +\infty]$ *è misurabile e* $\int\limits_A f(x)\ d\mu(x) = 0$*, allora* $f = 0$ *quasi ovunque in* A*.*

DIM.: Possiamo scrivere

$$\{x \in A : \ f(x) > 0\} = \bigcup_{n=1}^{\infty} \{x \in A : \ f(x) > \frac{1}{n}\}.$$

Tutti gli insiemi a destra hanno misura nulla: se fosse infatti $\mu(E_{\overline{n}}) > 0$, con $E_{\overline{n}} = \{x \in A : \ f(x) > \frac{1}{\overline{n}}\}$ avremmo

$$\int\limits_A f(x)\ d\mu(x) \geq \int\limits_{E_{\overline{n}}} f(x)\ d\mu(x) \geq \mu(E_{\overline{n}})/\overline{n} > 0,$$

contro l'ipotesi. Q.E.D.

Il seguente teorema mostra che l'integrale di Riemann coincide con l'integrale di Lebesgue, fatto ovviamente rispetto alla misura di Lebesgue, sull'insieme delle funzioni integrabili secondo Riemann (limitate su un insieme limitato: per gli integrali impropri la faccenda è leggermente più complicata[4]).

[4]Si dimostri per esercizio che l'integrale improprio secondo Riemann di una funzione *non negativa*, se esiste, coincide col suo integrale di Lebesgue: basta usare opportunamente il Teorema di Bebbo Levi e il teorema di confronto tra integrale di Riemann e di Lebesgue enunciato in questa pagina. Stessa cosa se la funzione è assolutamente integrabile nel senso di Riemann (si usi il risultato per le funzioni non negative ed il teorema della convergenza dominata). Invece, se f è integrabile in senso improprio con integrale finito, ma l'integrale del suo valore assoluto diverge a $+\infty$, si vede facilmente che gli integrali della parte positiva e della parte negativa sono $+\infty$: per questo motivo, la funzione *non è integrabile* nel senso di Lebesgue.

Enunciamo e dimostriamo il teorema in dimensione 1: la generalizzazione a dimensione superiore si dimostra allo stesso modo.

TEOREMA: Sia $f : [a, b] \to \mathbf{R}$ una funzione limitata ed integrabile secondo Riemann. Allora f è misurabile secondo Lebesgue, e il suo integrale di Lebesgue coincide con l'integrale di Riemann.

DIM.: Ai fini della dimostrazione, dobbiamo provvisoriamente distinguere l'integrale di Riemann da quello di Lebesgue: data $f : [a, b] \to \mathbf{R}$, conveniamo che $\int_a^b f(x)\, dx$ rappresenti il suo integrale di Lebesgue, mentre indicheremo con $\mathcal{R}\int_a^b f(x)\, dx$ il suo integrale di Riemann (purché esistano)... Ricordiamo anche che l'integrale di Lebesgue delle funzioni a scala coincide per definizione con il loro integrale di Riemann.

Per definizione di integrale (superiore ed inferiore) secondo Riemann, è possibile trovare due successioni di funzioni a scala $\{\psi_n\}$ e $\{\phi_n\}$, con $\psi_n \geq f \geq \phi_n$ e

$$\lim_{n\to+\infty} \int_a^b \psi_n\, dx = \lim_{n\to+\infty} \int_a^b \phi_n\, dx = \mathcal{R}\int_a^b f\, dx.$$

Siano ora $\overline{\psi}(x) = \inf\{\psi_n(x) : \ n = 1, 2, \ldots\}$, $\underline{\phi}(x) = \sup\{\phi_n(x) : \ n = 1, 2, \ldots\}$. Queste due funzioni sono misurabili, e $\underline{\phi} \leq f \leq \overline{\psi}$. Per la monotonia dell'integrale sarà

$$\int_a^b \psi_n(x)\, dx \geq \int_a^b \overline{\psi}(x)\, dx,$$

da cui passando al limite

$$\mathcal{R}\int_a^b f(x)\, dx \geq \int_a^b \overline{\psi}(x)\, dx,$$

e analogamente

$$\mathcal{R}\int_a^b f(x)\, dx \leq \int_a^b \underline{\phi}(x)\, dx.$$

Siccome $\overline{\psi} \geq \underline{\phi}$, se ne deduce che $\int_a^b(\overline{\psi} - \underline{\phi})\, dx = 0$, da cui $\overline{\psi} - \underline{\phi} = 0$ quasi ovunque, ossia $\overline{\psi} = \underline{\phi} = f$ quasi ovunque in $[a, b]$. Ne segue immediatamente che f è misurabile e che il suo integrale di Lebesgue coincide con quello di Riemann. Q.E.D.

In realtà, si può dimostrare che una funzione limitata è integrabile secondo Riemann se e soltanto se essa è *quasi ovunque continua* (Teorema di Vitali). Per motivi di tempo, non dimostreremo questo teorema.

Passiamo ad un importantissimo teorema, che migliora di gran lunga il teorema di riduzione degli integrali doppi che abbiamo visto per l'integrale di Riemann:

TEOREMA (di Fubini e Tonelli): Sia $f : \mathbf{R}^2 \to \overline{\mathbf{R}}$ *una funzione misurabile. Allora*

(i) Se $f \geq 0$*, allora per quasi ogni* $y \in \mathbf{R}$ *la funzione* $x \mapsto f(x,y)$ *è misurabile sulla retta reale. Inoltre, la funzione* $y \mapsto \int_{\mathbf{R}} f(x,y)\,dx$ *è misurabile e si ha*

$$(*) \quad \int_{\mathbf{R}^2} f(x,y)\,dx\,dy = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x,y)\,dx \right)\,dy.$$

Ovviamente, le stesse cose valgono anche scambiando il ruolo di x *e* y*.*

(ii) Se f *è di segno qualunque e* $\int_{\mathbf{R}} \left(\int_{\mathbf{R}} |f(x,y)|\,dx \right)\,dy < +\infty$*, allora* f *è sommabile. La stessa cosa vale anche scambiando l'ordine di integrazione.*

(iii) Se f *è di segno qualunque e sommabile, continua a valere l'enunciato del punto (i).*

Questo teorema vale anche, con le ovvie modifiche, per integrali su $\mathbf{R}^n$. Lo dimostreremo tra un po', per misure più generali della misura di Lebesgue: questo richiederà tuttavia la definizione di misura prodotto.

Per il momento osserviamo che se la funzione f è di segno qualunque e non è sommabile, l'enunciato non è più vero e i due integrali iterati possono essere diversi, come mostrano esempi anche semplici[5]. Invece, il teorema si generalizza a dimensione superiore: lo spazio ambiente può essere $\mathbf{R}^n \times \mathbf{R}^k$, e possiamo supporre $x \in \mathbf{R}^n$, $y \in \mathbf{R}^k$...

Visto che siamo in vena di dare enunciati, tanto vale dare anche quello del teorema di cambio di variabile per gli integrali multipli: una versione non ottimale è la seguente.

TEOREMA (Del cambiamento di variabili negli integrali multipli): Sia $\Phi : A \to B$ *un diffeomorfismo, con* A *e* B *aperti di* $\mathbf{R}^n$*,* $f : B \to \overline{\mathbf{R}}$ *una funzione integrabile. Allora si ha*

$$\int_B f(y)\,dy = \int_A f(\Phi(x))|\det(\nabla\Phi(x))|\,dx.$$

[5]Si consideri per esempio la funzione $f(x,y) = (x-y)/(x+y)^3$: i suoi integrali iterati sono finiti e diversi sul quadrato $[0,1] \times [0,1]$.

Torniamo al teorema di Fubini: cominciamo col dare la definizione di *misura prodotto* di due misure. Nel seguito, presento in realtà qualche dettaglio in più di quelli forniti a lezione!

DEFINIZIONE: Sia μ una misura su $\mathbf{R}^n$, ν una misura su $\mathbf{R}^m$[6]. La *misura prodotto* $\mu \times \nu$ è un'applicazione $\mu \times \nu : \mathbf{R}^{n+m} \to [0, +\infty]$ definita per ogni $S \subset \mathbf{R}^{n+m}$ da

$$\begin{aligned}(\mu \times \nu)(S) \qquad &= \inf\{\sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) : \\ &S \subset \bigcup_{i=1}^{\infty}(A_i \times B_i),\ A_i\ \mu - misurabile,\ B_i\ \nu - misurabile\}.\end{aligned}$$

È un semplice esercizio verificare che questa è effettivamente una misura esterna (e quindi il nome è giustificato): si procede allo stesso modo che per la misura di Lebesgue...

Inoltre, facendo il prodotto delle misure di Lebesgue su due spazi euclidei, si ottiene proprio la misura di Lebesgue sullo spazio prodotto:

OSSERVAZIONE: Denotiamo con m_n, m_m, m_{n+m} la misura di Lebesgue su $\mathbf{R}^n$, $\mathbf{R}^m$, $\mathbf{R}^{n+m}$ rispettivamente. Allora $m_{n+m} = m_n \times m_m$. Infatti, se nei ricoprimenti che definiscono la misura prodotto mettessimo l'ulteriore condizione che A_i e B_i siano intervalli, otterremmo esattamente la misura di Lebesgue sullo spazio prodotto (perchè ogni intervallo di $\mathbf{R}^{n+m}$ è prodotto di un intervallo di $\mathbf{R}^n$ e di un intervallo di $\mathbf{R}^m$, e la misura è il prodotto delle misure): se ne deduce che $m_n \times m_m \leq m_{n+m}$.

Per ottenere la disuguaglianza opposta, prendiamo un ricoprimento di $S \subset \mathbf{R}^{n+m}$ come nella definizione della misura prodotto. Si noti che non è restrittivo supporre che $m(A_i) \leq 1$, $m(B_i) \leq 1$ per ogni i (se così non fosse, basterebbe suddividere gli insiemi che formano il ricoprimento in insiemi misurabili più piccoli due a due disgiunti: ne basta una quantità numerabile, e la somma che compare nella definizione di misura prodotto non cambia grazie all'additività della misura di Lebesgue).

Fissiamo $\varepsilon > 0$: per definizione di misura di Lebesgue, per ogni fissato i troviamo un ricoprimento di A_i con intervalli $I_{i,j}$ tali che $\sum_j m_n(I_{i,j}) < m_n(A_i) + \varepsilon/2^i$, e analogamente un ricoprimento di B_i con intervalli $J_{i,k}$ tali

[6] Più in generale, μ può essere una misura su uno spazio X, ν una misura su uno spazio Y. La misura prodotto è allora definita su $X \times Y$. Tutti i risultati che seguono valgono in questo caso più generale (tranne, ovviamente, quanto verrà detto sulla misura di Lebesgue).

che $\sum_k m_m(J_{i,k}) < m_m(B_i) + \varepsilon/2^i$. Gli intervalli $I_{i,j} \times I_{i,k}$ formano ancora un ricoprimento numerabile di S, ed è facile vedere che le somme delle loro misure non superano $\sum_i m_n(A_i)m_m(B_i) + 3\varepsilon$: dunque fare l'inf su ricoprimenti fatti di prodotti di insiemi misurabili o di intervalli è la stessa cosa, per cui $m_{n+m} = m_n \times m_m$.

Il prossimo teorema è la versione generale del teorema di Fubini. Per comprendere l'enunciato è necessaria la definizione di σ-finitezza:

DEFINIZIONE: Una misura μ su $\mathbf{R}^n$ si dice *σ-finita* se $\mathbf{R}^n$ si può scrivere come unione numerabile di insiemi misurabili di misura finita. Più in generale, un insieme si dice σ-finito rispetto ad una misura μ, se è misurabile e può essere decomposto in una quantità numerabile di insiemi misurabili di misura finita: è abbastanza immediato vedere che questi insiemi si possono scegliere due a due disgiunti.

Una funzione $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ si dice σ-finita se $\mathbf{R}^n$ si può spezzare in un'unione numerabile di insiemi misurabili, su ciascuno dei quali f è sommabile.

ESERCIZIO: Si dimostri per esercizio che la misura di Lebesgue è σ-finita. Si mostri poi che una funzione misurabile secondo Lebesgue $f : \mathbf{R}^n \to [0, +\infty]$ è σ-finita a patto che $m(\{x \in \mathbf{R}^n : f(x) = +\infty\}) = 0$[7].

TEOREMA (di Fubini, versione generale): Sia μ una misura su $\mathbf{R}^n$, ν una misura su $\mathbf{R}^m$. Allora la misura prodotto $\mu \times \nu$ ha la seguente proprietà: se $S \subset \mathbf{R}^{n+m}$ esiste un insieme $(\mu \times \nu)$-misurabile S' tale che $S \subset S'$ e $(\mu \times \nu)(S) = (\mu \times \nu)(S')$ (S' si chiama involucro misurabile di S). Se A è μ-misurabile e B è ν-misurabile allora $A \times B$ è $(\mu \times \nu)$-misurabile e si ha $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$.

Se poi $S \subset \mathbf{R}^{n+m}$ è σ-finito rispetto a $\mu \times \nu$, allora le sezioni

$$S_y = \{x \in \mathbf{R}^n : \ (x, y) \in S\}, \qquad S_x = \{y \in \mathbf{R}^m : \ (x, y) \in S\}$$

sono rispettivamente μ-misurabili per ν-quasi ogni fissato y, e ν-misurabili per μ-quasi ogni fissato x. Inoltre

$$(\mu \times \nu)(S) = \int_{\mathbf{R}^m} \mu(S_y)\, d\nu(y) = \int_{\mathbf{R}^n} \nu(S_x)\, d\mu(x).$$

Se poi $f : \mathbf{R}^{n+m} \to \overline{\mathbf{R}}$ è integrabile e σ-finita rispetto alla misura $\mu \times \nu$ (in particolare, se f è sommabile), allora le applicazioni

$$\phi : x \mapsto \int_{\mathbf{R}^m} f(x, y)\, d\nu(y), \quad \psi : y \mapsto \int_{\mathbf{R}^n} f(x, y)\, d\mu(x)$$

[7] Si considerino gli insiemi $A_i = \{x \in \mathbf{R}^n : i \leq f(x) < i+1\}$: tolto l'insieme dove f vale $+\infty$, formano una partizione di $\mathbf{R}^n$. Se $m(A_i) < +\infty$, è facile vedere che $\int_{A_i} f(x)\, dx < +\infty$. Altrimenti, grazie alla σ-finitezza della misura, A_i può essere decomposto in una successione di insiemi di misura finita...

sono ben definite per μ-quasi ogni x, ν-quasi ogni y, e sono rispettivamente μ-integrabile e ν-integrabile. Infine vale il teorema di riduzione degli integrali doppi:

$$\int_{\mathbf{R}^{n+m}} f(x,y)\ d(\mu\times\nu)(x,y) = \int_{\mathbf{R}^n} \phi(x)\ d\mu(x) = \int_{R^m} \psi(y)\ d\nu(y).$$

OSSERVAZIONE: Da questo enunciato si recupera subito quello che abbiamo dato per la misura di Lebesgue: l'unica cosa che non è immediatamente ovvia è cosa succede se facciamo l'integrale iterato di una funzione $f : \mathbf{R}^2 \to [0, +\infty]$ che sia misurabile ma *non σ-finita* rispetto alla misura di Lebesgue nel piano. Dobbiamo far vedere che l'integrale iterato di questa funzione vale $+\infty$...e quindi "predice" correttamente che essa non è sommabile.

Vediamo come svolgere questo esercizio: la nostra ipotesi implica che l'insieme $S = \{(x,y) :\ f(x,y) = +\infty\}$ ha misura di Lebesgue positiva (altrimenti f sarebbe σ-finita. Ora, la misura di S si ottiene integrando la misura delle sezioni S_y. Dunque l'insieme $\{y :\ m(S_y) > 0\}$ deve avere misura positiva. Per ogni y in tale insieme si ha evidentemente $\int_{\mathbf{R}} f(x,y)\ dy = +\infty$ (perché $f \equiv +\infty$ su S).

In conclusione, quando calcoliamo l'integrale iterato di f su S dobbiamo integrare la costante $+\infty$ su un insieme di misura positiva: il risultato è $+\infty$.

In classe non abbiamo visto la dimostrazione del teorema di Fubini...ma se qualcuno ci tiene, eccola!

DIM. del Teorema di Fubini: Si consideri la famiglia di insiemi

$$\mathcal{F} = \{S \subset \mathbf{R}^{n+m} :\ S_y \text{ è } \mu\text{−}misurabile\ per\ \nu\text{−}q.o.\ y \in \mathbf{R}^m,\ y \mapsto \mu(S_y) \text{ è } \nu\text{−}misurabile\}.$$

Dato $S \in \mathcal{F}$, definiamo

$$\rho(S) = \int_{\mathbf{R}^m} \mu(S_y)\ d\nu(y) = \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} \mathbf{1}_S(x,y)\ d\mu(x)\right)\ d\nu(y).$$

Grazie alle proprietà della misura e degli integrali è facile verificare che ρ è numerabilmente subadditiva sugli elementi di $\mathcal{F}$, ed è additiva sulle famiglie numerabili disgiunte di elementi di $\mathcal{F}$.

Definiamo poi tre altre famiglie di sottinsiemi di $\mathbf{R}^{n+m}$:

$$\begin{aligned}
\mathcal{P}_0 &= \{A \times B :\ A\ \mu - misurabile,\ B\ \nu - misurabile\},\\
\mathcal{P}_1 &= \{\bigcup_{i=1}^{\infty} S_i :\ S_i \in \mathcal{P}_0\}, \quad \mathcal{P}_2 = \{\bigcap_{i=1}^{\infty} S_i :\ S_i \in \mathcal{P}_1\}.
\end{aligned}$$

È immediato verificare che $\mathcal{P}_0 \subset \mathcal{F}$ e se $A \times B \in \mathcal{P}_0$, allora $\rho(A \times B) = \mu(A)\nu(B)$.

L'*intersezione* di due elementi di $\mathcal{P}_0$ è ancora un elemento di $\mathcal{P}_0$: $(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$. Se ne deduce facilmente che la *differenza* di due elementi di $\mathcal{P}_0$ si può scrivere come unione di due elementi *disgiunti* di $\mathcal{P}_0$: precisamente, si ha

$$(A_1 \times B_1) \setminus (A_2 \times B_2) = ((A_1 \cap A_2) \times (B_1 \setminus B_2)) \cup ((A_1 \setminus A_2) \times B_1).$$

Di conseguenza, gli elementi di $\mathcal{P}_1$ sono unioni di successioni *disgiunte* di elementi di $\mathcal{P}_{\prime}$: in particolare, $\mathcal{P}_1 \subset \mathcal{F}$. Inoltre, gli elementi di $\mathcal{P}_1$ sono precisamente quelli che compaiono nella definizione della misura prodotto: ne segue che non è restrittivo supporre che gli $A_i \times B_i$ che compaiono in quella definizione siano disgiunti.

Mostriamo ora che per ogni $S \subset \mathbf{R}^{n+m}$ si ha

$$(*)\ (\mu \times \nu)(S) = \inf\{\rho(R) :\ S \subset R \in \mathcal{P}_1\}.$$

Sia infatti R come in $(*)$, e scriviamo $R = \bigcup_{i=1}^{\infty} A_i \times B_i$ (dove tutti i prodotti sono in $\mathcal{P}_0$ e due a due disgiunti). Allora

$$\rho(R) = \sum_i \rho(A_i \times B_i) = \sum_i \mu(A_i)\nu(B_i).$$

Passando all'inf si ha allora

$$(\mu \times \nu)(S) = \inf\{\rho(R) :\ S \subset R \in \mathcal{P}_1\}.$$

Se poi $A \times B \in \mathcal{P}_0$ si ha, per definizione di misura prodotto: $(\mu \times \nu)(A \times B) \leq \mu(A)\nu(B) = \rho(A \times B) \leq \rho(R)$ per ogni $R \in \mathcal{P}_1$, $A \times B \subset R$. Passando all'inf su R e usando $(*)$ si ha allora

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B).$$

Mostriamo che $A \times B$ è anche $(\mu \times \nu)$-misurabile: sia $T \subset \mathbf{R}^{n+m}$ e prendiamo $R \in \mathcal{P}_1$ con $T \subset R$. Gli insiemi $R \setminus (A \times B)$, $R \cap (A \times B)$ appartengono ancora a $\mathcal{P}_1$ e formano una partizione di R: usando ancora $(*)$ si ha

$$\begin{aligned}&(\mu \times \nu)(T \cap (A \times B)) + (\mu \times \nu)(T \setminus (A \times B)) \leq \\ &\rho(R \cap (A \times B)) + \rho(R \setminus (A \times B)) = \rho(R)\end{aligned}$$

e passando all'inf su R

$$(\mu \times \nu)(T \cap (A \times B)) + (\mu \times \nu)(T \setminus (A \times B)) \leq (\mu \times \nu)(T),$$

per cui $A \times B$ è misurabile. Ne segue subito che anche gli elementi di $\mathcal{P}_1$, $\mathcal{P}_2$ sono $(\mu \times \nu)$-misurabili. Si noti anche che $(\mu \times \nu)(S) = \rho(S)$ per ogni $S \in \mathcal{P}_1$: basta scrivere S come unione di elementi due a due disgiunti di $\mathcal{P}_0$, e usare la numerabile additività di ρ e della misura prodotto.

Vediamo come si trova l'involucro misurabile di un insieme. Sempre grazie a $(*)$, dato $S \subset \mathbf{R}^{n+m}$ possiamo trovare una successione $\{R_i\} \subset \mathcal{P}_1$ tale che $S \subset R_i$ per ogni i e $\rho(R_i) < (\mu \times \nu)(S) + 1/i$. Questa successione possiamo poi sceglierla decrescente (l'intersezione di due elementi di $\mathcal{P}_1$ è ancora un elemento di $\mathcal{P}_1$...). Se $(\mu \times \nu)(S) < +\infty$, il teorema della convergenza dominata ci dice allora che

$$\rho(\bigcap_i R_i) = \lim_{i \to +\infty} \rho(R_i) = (\mu \times \nu)(S).$$

Poiché poi $\rho(R_i) = (\mu \times \nu)(R_i)$ e $(\mu \times \nu)(R_i) \to (\mu \times \nu)(\bigcap_{i=1}^{\infty} R_i)$ per una ben nota proprietà della misura, abbiamo trovato un elemento di $\mathcal{P}_2$ che ha la stessa misura di S. Se per disgrazia fosse $(\mu \times \nu)(S) = +\infty$, basta invece prendere $R = \mathbf{R}^{n+m} \in \mathcal{P}_2$.

Sia poi S tale che $(\mu \times \nu)(S) = 0$. Allora esiste $R \supset S$, $R \in \mathcal{P}_2$ tale che $\rho(R) = 0$. Ne segue subito che $S \in \mathcal{F}$ e $\rho(S) = 0$.

Se invece $S \subset \mathbf{R}^{n+k}$ è un insieme misurabile con $(\mu \times \nu)(S) < +\infty$, prendiamo un suo involucro misurabile $R \in \mathcal{P}_2$. Evidentemente, $(\mu \times \nu)(R \setminus S) = 0$, da cui come abbiamo visto segue $\rho(R \setminus S) = 0$. Ma allora $\mu(S_y) = \mu(R_y)$ per ν-quasi ogni $y \in \mathbf{R}^m$ per cui

$$(**)\ (\mu \times \nu)(S) = \rho(R) = \int_{\mathbf{R}^m} \mu(S_y)\, d\nu(y).$$

La $(**)$ vale poi anche per un insieme S σ-finito: basta decomporlo in una successione disgiunta di insiemi di misura finita, per ciascuno dei quali la formula è vera, e applicare poi la numerabile additività.

Per concludere la dimostrazione, osserviamo che la $(**)$ non è che l'ultimo enunciato della tesi nel caso in cui $f(x, y) = \mathbf{1}_S(x, y)$. Per linearità, la tesi è vera anche per funzioni semplici non negative (purché gli "scalini" siano tutti σ-finiti). Da questo si ottiene l'enunciato per $f : \mathbf{R}^{n+m} \to [0, +\infty]$ purché f sia misurabile e σ-finita: basta usare il teorema di approssimazione con una successione crescente di funzioni semplici e il teorema di Beppo Levi (le funzioni approssimanti soddisfano automaticamente l'ipotesi di σ-finitezza). Infine, per una funzione di segno qualunque basta separare la parte positiva e la parte negativa. Q.E.D.

6 Lezione del 15/10/2009 (2 ore)

Cominciamo finalmente a dedicarci allo studio dei rudimenti dell'Analisi Funzionale lineare: concentreremo in particolare il nostro interesse sugli spazi di

Banach e sugli spazi di Hilbert, vedendo numerosi esempi e - temo - qualche complemento di teoria della misura.

In realtà, soprattutto all'inizio quel che faremo non è altro che uno studio accurato dell'*algebra lineare* in spazi di dimensione infinita! Probabilmente saprete già che, dal punto di vista puramente algebrico, quando si passa dalla dimensione finita alla dimensione infinita le novità non sono molte: se ammettiamo l'assioma della scelta, uno spazio vettoriale su **R** o su **C** ammette sempre una base, ossia un insieme massimale di vettori linearmente indipendenti. Inoltre, ogni elemento dello spazio si scrive come combinazione lineare *finita* di elementi della base...

Dal punto di vista dell'analista, tuttavia, questo non è sufficiente: per esempio, come minimo abbiamo bisogno di una nozione di continuità (e quindi di una topologia...o ancora meglio, se possibile, di una metrica), che sia *compatibile* con le operazioni di spazio vettoriale. Vogliamo cioè, come richiesta minima assolutamente irrinunciabile, una topologia tale che le *operazioni di spazio vettoriale* (somma e prodotto per scalare) siano continue.

DEFINIZIONE: Sia X uno spazio vettoriale su **R** o su **C**, dotato di una topologia τ. Lo spazio (X, τ) si dice *spazio vettoriale topologico* se e soltanto se le operazioni di spazio vettoriale (somma e prodotto per scalare) sono continue rispetto alle ovvie topologie prodotto indotte da τ.

Gli esempi più semplici di spazi vettoriali topologici, nonché (quasi) i soli di cui ci occuperemo in questo corso, sono gli *spazi normati*:

DEFINIZIONE (Norma, spazio normato): Sia X uno spazio vettoriale su **R** (le modifiche per spazi vettoriali su **C** sono ovvie...). Una *norma* su X è un'applicazione $\|\cdot\| : X \mapsto \mathbf{R}$ tale che

(i) $\|x\| \geq 0 \ \forall x \in X$, $\|x\| = 0$ sse $x = 0$;

(ii) $\|\lambda x\| = |\lambda|\|x\|$ per ogni $x \in X$, $\lambda \in \mathbf{R}$ (omogeneità);

(iii) $\|x + y\| \leq \|x\| + \|y\|$ per ogni $x, y \in X$ (disuguaglianza triangolare).

Uno spazio vettoriale su cui è definita una norma si chiama *spazio normato*: esso è uno spazio metrico con la distanza definita da

$$d(x, y) := \|x - y\|, \quad x, y \in X.$$

È un semplice esercizio verificare che uno spazio normato (con la distanza naturale appena introdotta) è uno spazio vettoriale topologico.

Detto a margine, visto che uno spazio normato è uno spazio metrico, vale la pena di osservare che è perfettamente lecito testare la continuità con le successioni: per esempio una funzione $f : X \to \mathbf{R}$ è continua se e solo se per ogni $\overline{x}$ e per ogni successione $x_k \to \overline{x}$ si ha $f(x_k) \to f(\overline{x})$.

Vediamo subito alcuni esempi di spazi normati:

1. Lo spazio $\mathbf{R}^n$ è uno spazio normato con l'usuale norma euclidea: se $x = (x_1, \ldots, x_n)$, essa è definita da $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$. Altre norme su $\mathbf{R}^n$ sono anche $|x|_1 = \sum_{i=1}^n |x_i|$ e $|x|_\infty = \max\{|x_i| : \ i = 1, \ldots, n\}$: è un semplice ed utile esercizio verificarlo, e può anche essere istruttivo disegnare le palle della metrica indotta da queste norme.

   Più in generale, sono norme su $\mathbf{R}^n$ quelle definite nel modo seguente:
   $$|x|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p < +\infty.$$
   In questo caso, la verifica della disuguaglianza triangolare è meno ovvia: avremo modo di vederla nel seguito del corso.

2. Un altro spazio normato che dovrebbe essere ben noto, ed è forse l'esempio più semplice di spazio normato di dimensione infinita, è lo spazio delle funzioni continue $C^0([a, b])$ (o più in generale $C^0(K)$, con K spazio metrico compatto) dotato della norma
   $$\|u\|_\infty = \sup\{|u(x)| : \ x \in [a, b]\}.$$

3. Altro esempio di spazio normato in dimensione infinita: $C^0([a, b])$ con la norma
   $$\|u\|_1 = \int_a^b |u(x)| \, dx$$
   (lo si verifichi!).

4. Utili esempi di spazi normati di dimensione infinita sono gli *spazi di successioni* ℓ^p, $1 \leq p \leq +\infty$, che si definiscono nel modo seguente. Sia $1 \leq p < +\infty$, e si consideri una successione di numeri reali $\{x_n\}_{n \in N}$: la sua *norma* ℓ^p si definisce come
   $$\|\{x_n\}\|_p := \left(\sum_{n=1}^\infty |x_n|^p \right)^{1/p}.$$
   Se poi $p = +\infty$, si definisce
   $$\|\{x_n\}\|_\infty = \sup\{|x_n| : \ n = 1, 2, 3, \ldots\}.$$

Infine, per $1 \le p \le +\infty$ definiamo

$$\ell^p = \{\{x_n\}_{n\in\mathbf{N}} \in \mathbf{R}^{\mathbf{N}} : \ \|\{x_n\}\| < +\infty\}.$$

Questo è uno spazio normato con la norma $\|\cdot\|_p$... (Se diamo per buono che le p-norme che abbiamo definito negli spazi euclidei (al punto 1.) siano, appunto, delle norme, la verifica della disuguaglianza triangolare si ottiene semplicemente passando al limite...).

5. Altri esempi estremamente importanti di spazi normati sono gli *spazi di Lebesgue* L^p. Se Ω è un sottinsieme misurabile di $\mathbf{R}^n$, $1 \le p < +\infty$ e $u : \Omega \to \overline{\mathbf{R}}$ è una funzione misurabile secondo Lebesgue, definiamo

$$\|u\|_{L^p(\Omega)} := \left(\int_\Omega |u(x)|^p \, dx \right)^{1/p}.$$

Lo spazio $L^p(\Omega)$ si definisce poi come l'insieme delle funzioni misurabili tali che $\|u\|_{L^p} < +\infty$, quozientato rispetto alla relazione di equivalenza

$$u \sim v \Leftrightarrow u(x) = v(x) \textit{ per q.o. } x \in \Omega.$$

Quella scritta sopra è allora una norma su L^p, come vedremo più avanti nel corso.

Più in generale, se Ω è un insieme e μ è una misura esterna su Ω, data una funzione μ-misurabile $u : \Omega \to \overline{\mathbf{R}}$ possiamo definire

$$\|u\|_{L^p(\mu)} := \left(\int_\Omega |u(x)|^p \, d\mu(x) \right)^{1/p}.$$

Lo spazio $L^p(\mu)$ si definisce poi come prima.

Ci sono poi anche gli spazi $L^\infty(\Omega)$, $L^\infty(\mu)$, che si definiscono in modo leggermente più complicato: vedremo più avanti!

E' un esercizio relativamente semplice verificare che gli spazi di successioni ℓ^p sono spazi di quest'ultimo tipo, ottenuti prendendo $\Omega = \mathbf{N}$ e come misura μ la counting measure (cioè la misura tale che $\mu(A)$ è il numero di elementi di A se A è finito, $+\infty$ altrimenti). A questo scopo, basta far vedere che data una successione $\{a_n\}_{n\in\mathbf{N}}$ di numeri reali non negativi si ha

$$\int_{\mathbf{N}} a_n \, d\mu(n) = \sum_{n=1}^{\infty} a_n.$$

Indichiamo infatti con $\{a_n^N\}_n$ la successione "troncata" ai primi N termini, cioè $a_n^N = a_n$ se $1 \leq n \leq N$, mentre $a_n^N = 0$ altrimenti. Per $N \to +\infty$, queste successioni convergono puntualmente e in modo crescente alla successione originale. Il teorema di Beppo Levi permette allora di dire che l'integrale di $\{a_n\}$ è il limite degli integrali delle successioni troncate: ma gli integrali delle successioni troncate sono semplicemente le somme parziali della serie (tali successioni sono "funzioni semplici" per la nostra misura!).

Applicando poi il teorema della convergenza dominata alle stesse successioni troncate, si vede che l'enunciato continua a valere se $\{a_n\}$ è una successione a termini di segno qualunque la cui serie è *assolutamente convergente*: per quanto appena visto questo implica $\{a_n\} \in L^1(\mu)$. Le nostre successioni troncate sono allora dominate da $\{|a_n|\}$ e convergono puntualmente ad a_n per $N \to +\infty$: ancora una volta, il loro integrale non è altro che la somma parziale n-esima della serie. Il teorema della convergenza dominata permette allora di concludere che l'integrale del limite è il limite degli integrali, cioè la serie.

Esattamente come succedeva negli spazi euclidei, una proprietà altamente desiderabile per uno spazio normato è la *completezza*:

DEFINIZIONE (Spazio di Banach): Uno spazio vettoriale normato X si dice *di Banach* se è completo, ossia se ogni successione di Cauchy a valori in X converge.

Riprendiamo per un momento i nostri esempi di spazi normati e vediamo se sono completi:

1. È ben noto che lo spazio $\mathbf{R}^n$ con la metrica euclidea è completo. Lo è anche con tutte le altre metriche indicate: vedremo tra poco che su $\mathbf{R}^n$ *tutte le metriche provenienti da una norma sono equivalenti*, ossia inducono la stessa topologia (hanno gli stessi insiemi aperti, per cui la convergenza rispetto ad una delle norme è equivalente a quella rispetto ad un'altra) e hanno le stesse successioni di Cauchy.

2. Anche $(C^0([a,b]), \|\cdot\|_\infty)$ è un esempio di spazio di Banach. Sia infatti $\{f_n\}$ una successione di Cauchy nel nostro spazio: per ogni $\varepsilon > 0$ esiste $\nu \in \mathbf{N}$ tale che $\|f_n - f_m\|_\infty < \varepsilon$ per ogni $m, n \geq \nu$. Allora, per ogni $x \in [a,b]$ si ha
$$(**) |f_m(x) - f_n(x)| < \varepsilon$$

per $m, n \geq \nu$, e la successione reale $\{f_n(x)\}$ è di Cauchy. Per la completezza di $\mathbf{R}$, esiste un numero reale $f(x)$ tale che $f_n(x) \to f(x)$. Passando al limite per $m \to +\infty$ nella (**) otteniamo

$$|f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq \nu,$$

e passando al sup su $x \in [a, b]$ $\|f_n - f\|_\infty \leq \varepsilon$ per $n \geq \nu$. Dunque $f_n \to f$ uniformemente. La funzione f è poi continua perché sappiamo già che limite uniforme di funzioni continue è continuo.

3. Anche tutti gli spazi ℓ^p, $L^p(\Omega)$, $L^p(\mu)$ sono onesti spazi di Banach: li studieremo più in dettaglio durante il corso!

Nel nostro discorsetto sulla completezza di $\mathbf{R}^n$ con le varie norme, è saltato fuori il concetto di *norme equivalenti*: due norme su uno stesso spazio vettoriale si dicono equivalenti se inducono la stessa topologia, ossia se gli insiemi aperti indotti dalle due norme sono gli stessi.

Supponiamo dunque di avere uno spazio vettoriale X e due norme $\|\cdot\|_1$, $\|\cdot\|_2$ su di esso.Verifichiamo che le due norme sono equivalenti *se e solo se* esistono due costanti $c, C > 0$ tali che

$$(*) \; c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \forall x \in X.$$

Infatti, come ben sappiamo in ogni spazio metrico gli aperti si definiscono a partire dalle palle aperte: dunque, è immediato verificare che la topologia indotta è la stessa se e solo se, data comunque una palla "di uno dei due tipi" (cioè una palla rispetto ad una delle due metriche), è possibile scegliere una palla "dell'altro tipo", di raggio convenientemente piccolo, avente lo stesso centro ed in essa contenuta. Grazie all'omogeneità della norma ed alla definizione della metrica, le palle in uno spazio normato si ottengono tutte dalla palla unitaria centrata nell'origine tramite un'omotetia ed una traslazione....et voila: le due disuguaglianze (*) implicano che $B^{(1)}_{1/C}(0) \subset B^{(2)}_1(0)$ e che $B^{(2)}_c(0) \subset B^{(1)}_1(0)$.

Vediamo che vale anche il viceversa: se le due norme inducono la stessa topologia, allora deve valere (*). Infatti, la palla $B^{(1)}_1(0)$ è un aperto sia per la prima che per la seconda norma: esiste dunque $r > 0$ tale che $B^{(2)}_r(0) \subset B^{(1)}_1(0)$. Dato $x \in X$, $x \neq 0$ si ha evidentemente che $y = r/2\frac{x}{\|x\|_2} \in B^{(2)}_r(0)$, da cui $\|y\|_1 < 1$. Ma allora $r/2\|x\|_1 \leq \|x\|_2$. L'altra disuguaglianza si ottiene in modo analogo.

Grazie a (*), vediamo anche che due norme equivalenti hanno le stesse successioni di Cauchy: la completezza o meno dello spazio rimane invariata se si passa ad una norma equivalente. Viceversa, le due norme $\|\cdot\|_\infty$ e $\|\cdot\|_1$

sullo spazio delle funzioni continue $C^0([a,b])$ non sono equivalenti (perché la prima è una norma completa, la seconda non lo è). In effetti, la topologia indotta dalla norma $\|\cdot\|_\infty$ è strettamente *più forte* (ha più aperti) di quella indotta dalla norma $\|\cdot\|_1$... Le palle aperte rispetto alla norma $\|\cdot\|_\infty$ non sono aperte rispetto alla norma $\|\cdot\|_1$, mentre il viceversa è vero.

7 Lezione del 16/10/2009 (2 ore)

Vediamo qualche altro esempio di spazi in carne ed ossa:

- Lo spazio $(C^0([a,b]), \|\cdot\|_1)$ non è uno spazio di Banach perché non è completo. Prendiamo per esempio $[a,b] = [-1,1]$ e consideriamo la successione

  $$u_n(x) = \begin{cases} -1 & sex \leq -1/n, \\ 1 & se\ x \geq 1/n, \\ nx & se\ -1/n < x < 1/n. \end{cases}$$

  Questa successione è di Cauchy rispetto alla norma $\|\cdot\|_1$, ma non converge ad alcuna funzione continua: in effetti, è immediato verificare che converge in norma alla funzione discontinua $\mathrm{sgn}(x)$. Durante il corso, studieremo diffusamente il completamento di questo spazio: si tratta dello spazio di Banach $L^1([a,b])$ delle funzioni sommabili secondo Lebesgue dotato della norma $\|\cdot\|_1$.

- Lo spazio $\mathcal{C}^1([a,b])$ dotato della norma $\|f\|_{\mathcal{C}^1} := \|f\|_\infty + \|f'\|_\infty$ è uno spazio di Banach.

  Evidentemente basta verificare che è completo: che quella sia una norma è immediato.

  Ora, sia $\{f_n\}$ una successione di Cauchy in $\mathcal{C}^1$. Allora, per come è definita la norma, le successioni $\{f_n\}$ e $\{f'_n\}$ sono di Cauchy per la norma $\|\cdot\|_\infty$, che è una norma completa sullo spazio $\mathcal{C}^0$: ne segue che esistono due funzioni continue f, g tali che $f_n \to f$ e $f'_n \to g$ in norma $\|\cdot\|_\infty$. Ora, per ogni $x \in [a,b]$ si ha

  $$f_n(x) = f_n(a) + \int_a^x f'_n(t)\,dt.$$

  Passando al limite per $n \to +\infty$ otteniamo

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

da cui $f \in C^1$ con $f' = g$: da questo segue evidentemente che $f_n \to f$ in norma $\| \cdot \|_{\mathcal{C}^1}$.

- Al contrario, lo spazio $\mathcal{C}^1([a,b])$ con la norma $\| \cdot \|_\infty$ non è completo: esistono successioni di funzioni di classe $\mathcal{C}^1$ che convergono uniformemente a funzioni che non sono derivabili. Un controesempio si ottiene prendendo $[a,b] = [-1,1]$ e $f_n(x) = \sqrt{x^2 + 1/n}$: è facile verificare che questa successione di funzioni converge uniformemente a $|x|$.

Abbiamo già accennato al fatto che tutte le norme su $\mathbf{R}^n$ sono equivalenti: dimostriamolo!

TEOREMA: Tutte le norme sullo spazio $\mathbf{R}^n$ sono equivalenti.

DIM.: Poiché l'equivalenza tra norme gode della proprietà transitiva, sarà sufficiente mostrare che la norma euclidea $|\cdot|$ è equivalente ad una qualunque altra fissata norma $\|\cdot\|$. Dobbiamo far vedere che vale la coppia di disuguaglianze (*): esistono due costanti $c, C > 0$ tali che $c|x| \leq \|x\| \leq C|x|$ per ogni $x \in \mathbf{R}^n$, $x \neq 0$. Dividendo il tutto per $|x|$ e sfruttando l'omogeneità della norma, vediamo che la nostra coppia di disuguaglianze è vera se e solo se vale

$$(**)\ c \leq \|x\| \leq C \qquad \forall x \in \mathbf{R}^n, |x| = 1.$$

Ora, osserviamo che la funzione $x \mapsto \|x\|$ è continua rispetto alla topologia euclidea: essa è addirittura lipschitziana, come mostra la seguente catena di disuguaglianze:

$$\begin{aligned} |\|x\| - \|y\|| \quad &\leq \quad \|x - y\| = \|\sum_{i=1}^{n}(x_i - y_i)e_i\| \leq \sum_{i=1}^{n} |x_i - y_i|\|e_i\| \leq \\ &|x - y|(\sum_{i=1}^{n} \|e_i\|), \end{aligned}$$

in cui gli e_i sono i vettori della base canonica di $\mathbf{R}^n$ e si è usata l'omogeneità della norma e la disuguaglianza triangolare.

D'altra parte, la sfera unitaria $S = \{x \in \mathbf{R}^n :\ |x| = 1\}$ è un compatto euclideo: allora la (**) è vera a patto di prendere $c = \min\{\|x\| :\ x \in S\}$, $C = \max\{\|x\| :\ x \in S\}$, e questi oggetti esistono finiti grazie al teorema di Weierstrass (e $c \neq 0$ grazie al fatto che la norma $\|\cdot\|$ è non degenere). Q.E.D.

Da questo risultato segue che tutte le norme su uno spazio vettoriale reale di dimensione finita sono equivalenti, e che ogni isomorfismo lineare tra un tale spazio normato e $\mathbf{R}^n$ con la norma euclidea è anche un omeomorfismo (esercizio!). In sostanza, tutti gli spazi vettoriali (reali) normati di dimensione finita sono equivalenti allo spazio euclideo.

Invece, in dimensione infinita si possono mettere norme *non equivalenti* sullo stesso spazio: abbiamo visto che $(C^0([a,b]), \|\cdot\|_\infty)$ e $(C^0([a,b]), \|\cdot\|_1)$ non hanno la stessa topologia. Precisamente, la volta scorsa abbiamo costruito una successione di funzioni continue che converge a 0 nella norma $\|\cdot\|_1$ ma non nella norma $\|\cdot\|_\infty$. Questo ha una conseguenza a prima vista sorprendente: la *mappa identica* $Id : (C^0([a,b]), \|\cdot\|_1) \to (C^0([a,b]), \|\cdot\|_\infty)$ non è continua nell'origine. Abbiamo dunque un esempio di mappa *lineare e invertibile* tra due spazi normati di dimensione infinita che non è nemmeno continua!

Questo fatto non è legato alla scelta di norme esotiche e non complete (come la norma $\|\cdot\|_1$), ma è intrinseco della dimensione infinita:

PROPOSIZIONE: Sia $(X, \|\cdot\|)$ *uno spazio vettoriale normato (su* $\mathbf{R}$*) di dimensione infinita. Allora esistono sempre funzionali lineari* $T : X \to \mathbf{R}$ *che sono discontinui.*

DIM.: Esibiremo esplicitamente un funzionale lineare su X (un elemento del *duale algebrico* di X) che è discontinuo.

Prendiamo una base $\mathcal{B} = \{x_\alpha\}_{\alpha \in I}$ del nostro spazio: essa è evidentemente infinita. Non è restrittivo (normalizzando opportunamente gli elementi della base) supporre che sia $\|x_\alpha\| = 1$ per ogni $\alpha \in I$. Sappiamo bene dal corso di algebra lineare che per specificare un'applicazione lineare $T : X \to \mathbf{R}$ basta definirla sugli elementi della base: per linearità, essa sarà allora ben definita su tutto lo spazio. A questo fine, scegliamo un sottinsieme numerabile $\mathcal{B}' = \{\tilde{x}_n\} \subset \mathcal{B}$ e poniamo $T(\tilde{x}_n) = n$, $n = 1, 2, \ldots$, mentre poniamo $T(x_\alpha) = 0$ se $x_\alpha \in \mathcal{B} \setminus \mathcal{B}'$. Il funzionale lineare così definito risulta discontinuo: la successione $y_n = x_n/\sqrt{n}$ tende a 0 in norma, ma $T(y_n) = \sqrt{n} \to +\infty$. Q.E.D.

La devastante conseguenza della proposizione precedente è la seguente: quando un analista parla di *duale* di uno spazio vettoriale normato X (o più in generale di uno s.v. topologico), non si riferisce *mai* al duale "dei geometri", cioè alla totalità dei funzionali lineari su X (*duale algebrico*), ma alla totalità dei funzionali lineari continui:

DEFINIZIONE (Duale topologico): Dato uno spazio vettoriale normato (spazio vettoriale topologico) X, il suo *duale topologico* (o più semplicemente il suo duale) è lo spazio vettoriale dei funzionali lineari *continui* $T : X \to \mathbf{R}$. Esso si denota con X'.

La continuità associata alla linearità permette la seguente caratterizzazione degli elementi del duale di uno spazio normato:

TEOREMA (Caratterizzazione dei funzionali lineari continui su uno spazio normato): Sia $(X, \|\cdot\|)$ uno spazio normato, $T : X \to \mathbf{R}$ un funzionale lineare. I seguenti fatti sono equivalenti:

(i) T è continuo;

(ii) T è continuo nell'origine;

(iii) T è limitato: esiste una costante $C > 0$ tale che

$$|T(x)| \leq C\|x\| \quad \forall x \in X;$$

(iv) Il nucleo di T, $\ker(T)$ è un sottospazio chiuso di X.

DIM.: È ovvio che $(i) \Rightarrow (ii)$, mentre che $(ii) \Rightarrow (i)$ è conseguenza del fatto che $|T(x) - T(\overline{x})| = |T(x - \overline{x})|$ grazie alla linearità. È anche ovvio che $(iii) \Rightarrow (ii)$.

Mostriamo viceversa che $(ii) \Rightarrow (iii)$: grazie alla definizione di continuità nell'origine (con $\varepsilon = 1$), esiste $\delta > 0$ tale che $|T(x)| \leq 1$ ogniqualvolta $\|x\| \leq \delta$. Allora, per ogni $y \in X$, $y \neq 0$, si ha

$$|T(y)| = |T(\frac{\|y\|}{\delta} \cdot y\frac{\delta}{\|y\|})| = \frac{\|y\|}{\delta}|T(y\frac{\delta}{\|y\|})| \leq \frac{1}{\delta}\|y\|,$$

che è proprio la limitatezza di T.

È poi ovvio che $(i) \Rightarrow (iv)$: la controimmagine del chiuso $\{0\}$ secondo la funzione continua T è un chiuso.

Per concludere la dimostrazione, la prossima volta faremo vedere che $(iv) \Rightarrow (ii)$.

8 Lezione del 21/10/2009 (2 ore)

Concludiamo la dimostrazione del teorema che caratterizza la continuità di un funzionale lineare, mostrando che $(iv) \Rightarrow (ii)$. Se $T \equiv 0$ non c'è nulla da dimostrare. Altrimenti, sia $\ker(T)$ chiuso e supponiamo per assurdo che T non sia continuo in 0. Allora è possibile trovare una successione $\{x_n\} \subset X$ con $\|x_n\| \to 0$ e tale che $T(x_n) \not\to 0$. Estraendo eventualmente una sottosuccessione, questo implica che esiste una costante $c > 0$ tale che $|T(x_n)| \geq c$ per ogni n.

Sia poi $y \in X$ qualunque, e consideriamo la successione $y_n = y - \frac{x_n}{T(x_n)}T(y)$: è immediato verificare che $y_n \in \ker(T)$ e che $y_n \to y$ in norma. Ne deriva che $y \in \overline{\ker(T)} = \ker(T)$, da cui $\ker(T) = X$ (in quanto y è arbitrario) e $T \equiv 0$, contro la nostra ipotesi che il funzionale fosse non nullo. Q.E.D.

Abbiamo visto dunque che un funzionale lineare $T : X \to \mathbf{R}$ (con X spazio vettoriale normato) è continuo se e solo se T è limitato, cioè se esiste una costante $C > 0$ tale che

$$|T(x)| \leq C\|x\| \quad \forall x \in X.$$

La *norma* di $T \in X'$ si definisce come la *più piccola* delle costanti C per cui la disuguaglianza è vera: precisamente, definiamo

$$(*)\ \|T\|_{X'} := \sup\{\frac{|T(x)|}{\|x\|} :\ x \in X,\ x \neq 0\} = \sup\{|T(x)| :\ x \in X,\ \|x\| \leq 1\}.$$

TEOREMA (duale di uno spazio normato): Sia X uno spazio vettoriale normato su $\mathbf{R}$ (non necessariamente completo). Consideriamo il duale topologico X' di X, dotato della norma duale definita in (*). Allora $(X', \|\cdot\|_{x'})$ è uno spazio di Banach.

DIM.: La verifica che $\|\cdot\|_{X'}$ è una norma su X è molto semplice ed è lasciata per esercizio.

Rimane da verificare la completezza: sia dunque $\{T_n\} \subset X'$ una successione di Cauchy rispetto alla norma $\|\cdot\|_{X'}$. Fissato $\varepsilon > 0$, esiste $\nu \in \mathbf{N}$ tale che $\|T_n - T_m\|_{X'} \leq \varepsilon$ per $m, n \geq \nu$. Allora, per ogni $x \in X$ e per $m, n \geq \varepsilon$ si ha

$$(**)\ |T_n(x) - T_m(x)| \leq \varepsilon\|x\|,$$

per cui la successione di numeri reali $\{T_n(x)\}$ è di Cauchy e converge ad un numero reale che chiameremo $T(x)$. È immediato verificare la linearità di questo limite puntuale $T : X \to \mathbf{R}$. Ci rimane da verificare che T è continuo, e che $T_n \to T$ nella norma di X'.

Passando al limite per $m \to +\infty$ in $(**)$ si ottiene

$$|T_n(x) - T(x)| \leq \varepsilon\|x\| \quad \forall x \in X,\ \forall n \geq \nu,$$

cioè $\|T_n - T\|_{X'} \leq \varepsilon$ per $n \geq \nu$. È anche ovvio che T è un funzionale lineare limitato in quanto, grazie alla disuguaglianza precedente, abbiamo

$$\|T\|_{X'} \leq \|T_\nu\|_{X'} + \|T_\nu - T\|_{X'} < +\infty.$$

Q.E.D.

OSSERVAZIONE/ESERCIZIO: Procedendo alla stessa maniera, potete dimostrare il seguente importante fatto. Siano $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ due spazi

normati, il secondo dei quali è di Banach. Allora lo spazio vettoriale $\mathcal{L}(X;Y)$ delle applicazioni lineari e continue di X in Y è uno spazio di Banach con la norma
$$\|T\|_{\mathcal{L}(X;Y)} = \sup\{\|T(x)\|_Y : \ x \in X, \ \|x\|_X \leq 1\}.$$

ESEMPIO: Mostriamo con un esempio che il sup nella definizione di norma duale non è necessariamente un massimo. Consideriamo lo spazio vettoriale $\ell^1 = \{\{x_n\}_{n\in\mathbf{N}} : \ \sum_{n=1}^{\infty} |x_n| < +\infty\}$, dotato della norma
$$\|\{x_n\}\|_{\ell^1} = \sum_{n=1}^{\infty} |x_n|.$$
Vedremo che questo è uno spazio di Banach.

Consideriamo poi il funzionale lineare $T : \ell^1 \to \mathbf{R}$ definito da
$$T(\{x_n\}) = \sum_{n=1}^{\infty}(1 - 1/n)x_n.$$
Si verifica facilmente che questo funzionale è ben definito su ℓ^1 (perché la serie converge assolutamente) ed è lineare. È poi limitato:
$$(***)|T(\{x_n\})| \leq \sum_{n=1}^{\infty}(1 - 1/n)|x_n| \leq \|\{x_n\}\|_{\ell^1},$$
e si noti che *l'ultima disuguaglianza è stretta se* $\{x_n\}$ *è una successione non identicamente nulla.*

La (***) dice che $\|T\|_{(\ell^1)'} \leq 1$. D'altra parte, la norma duale vale proprio 1: facciamo vedere che esiste una successione e_k di elementi di ℓ^1 (una successione di successioni...) tale che $\|e_k\|_{\ell^1} = 1$ e $T(e_k) \to 1$. Basta prendere come e_k il "k-esimo elemento della base canonica", cioè la successione il cui k-esimo elemento vale 1, mentre tutti gli altri valgono 0: si ha dunque $T(e_k) = 1 - 1/k$ e ne deduciamo che
$$\|T\|_{(\ell^1)'} = 1.$$
Siccome però la disuguaglianza (***) è stretta per le successioni non nulle, deduciamo che il sup nella definizione di norma duale non è un massimo.

Uno dei risultati più utili (e famosi) per lo studio del duale, è il teorema di Hahn-Banach:

TEOREMA (Hahn-Banach): Sia X uno spazio vettoriale reale, $p : X \to [0, +\infty)$ un'applicazione tale che *(i)* $p(\lambda x) = \lambda p(x)$ *per ogni* $x \in X$*, per ogni* $\lambda > 0$ *(positiva omogeneità);*

(ii) $p(x + y) \leq p(x) + p(y)$ *per ogni* $x, y \in X$ *(subadittività).*

Sia poi Y *un sottospazio vettoriale proprio di* X*,* $T : Y \to \mathbf{R}$ *un funzionale lineare su* Y *tale che* $T(x) \leq p(x)$ *per ogni* $x \in Y$*. Allora esiste un funzionale lineare* $\tilde{T} : X \to \mathbf{R}$ *che estende* T *(cioè* $\tilde{T}(x) = T(x)$ *per ogni* $x \in Y$*) e tale che* $\tilde{T}(x) \leq p(x)$ *per ogni* $x \in X$*.*

Come caso particolare, se X *è normato ogni funzionale lineare continuo su* Y *si estende ad un funzionale lineare continuo definito su tutto* X *avente la stessa norma.*

DIM.: La parte più corposa della dimostrazione consiste nel provare il seguente asserto: *se* Z *è un sottospazio proprio di* X *e* $T : Z \to \mathbf{R}$ *è lineare con* $T(x) \leq p(x)$ *per ogni* $x \in Z$*, il funzionale si può estendere ad un sottospazio strettamente più grosso in modo che la disuguaglianza continui a valere.*

Sia infatti $x_0 \in X \setminus Z$: estendiamo T ad un funzionale $\tilde{T}$ definito sullo spazio $Z \oplus \mathbf{R}\{x_0\}$. Per linearità avremo, per ogni $x \in Z$ e per ogni $t \in \mathbf{R}$:

$$\tilde{T}(x + tx_0) = T(x) + t\alpha,$$

dove $\alpha = T(x_0)$ è un numero reale da scegliere in modo che valga

$$T(x) + t\alpha = p(x + tx_0) \quad \forall x \in Z,\ \forall t \in \mathbf{R}.$$

Usando la linearità di T e la positiva omogeneità di p, e distinguendo i casi $t > 0$ e $t < 0$, si vede facilmente che la nostra disuguaglianza è equivalente alla coppia di disuguaglianze

$$\begin{aligned} T(x) + \alpha \leq p(x + x_0)\ \forall x \in Z, \\ T(y) - \alpha \leq p(y - x_0)\ \forall y \in Z. \end{aligned}$$

cioè α deve soddisfare

$$T(y) - p(y - x_0) \leq \alpha \leq p(x + x_0) - T(x) \quad \forall x, y \in Z.$$

Questo è possibile a patto che il membro di sinistra sia sempre minore o uguale del membro di destra, per ogni scelta di $x, y \in Z$. Ma questo è vero perché $T(x) + T(y) = T(x + y) \leq p(x + y) \leq p(x + x_0) + p(y - x_0)$ per ogni $x, y \in Z$. Il nostro asserto risulta dunque provato.

La dimostrazione del teorema segue abbastanza facilmente dall'asserto usando una delle versioni equivalenti del Lemma di Zorn. Consideriamo infatti la famiglia $\mathcal{F}$ dei funzionali lineari definiti su sottospazi Z di X contenenti

Y, che siano estensioni del nostro funzionale originale T e siano dominati dalla funzione p. Mettiamo una relazione d'ordine su $\mathcal{F}$: se $R : Z_1 \to \mathbf{R}$ e $S : Z_2 \to \mathbf{R}$ sono due elementi di $\mathcal{F}$, diciamo che $R \preceq S$ se $Z_1 \subset Z_2$ e S estende R. Il principio di massimalità di Hausdorff ci assicura che possiamo trovare un sottinsieme totalmente ordinato massimale $\mathcal{G} \subset \mathcal{F}$,

$$\mathcal{G} = \{S_\sigma : Z_\sigma \to \mathbf{R}, \sigma \in I\}.$$

Questo insieme ammette un maggiorante: si tratta del funzionale S definito sul sottospazio

$$Z = \bigcup_{\sigma \in I} Z_\sigma$$

(verificare che si tratta di un sottospazio!) ponendo $S(x) = S_\sigma(x)$ se $x \in S_\sigma$ (questa è una buona definizone perché $\mathcal{G}$ è totalmente ordinato).

Possiamo ora dire che $Z = X$ e S è l'estensione cercata: se così non fosse, l'asserzione dimostrata prima ci permetterebbe di estendere S ad un sottospazio strettamente più grande, contro la massimalità dell'insieme totalmente ordinato $\mathcal{G}$: questo conclude la dimostrazione dell'enunciato principale del teorema.

Il "caso particolare" in uno spazio normato segue subito ponendo $p(x) = \|T\|_{Y'} \, \|x\|$. Q.E.D.

La prossima volta vedremo alcune conseguenze del Teorema di Hahn-Banach: spero di convincervi che valgono ampiamente lo sforzo usato per la dimostrazione!

9 Lezione del 28/10/2009 (2 ore)

Una delle molte conseguenze del Teorema di Hahn-Banach è la seguente: *se $x \in X$ è possibile trovare $T \in X'$ tale che $\|T\| = 1$ e $T(x) = \|x\|$.*

Infatti, il funzionale può essere definito sulla retta $\mathbf{R}\{x\}$ ponendo $T(tx) = t\|x\|$ (ed ha evidentemente norma 1), e può poi essere esteso a tutto X grazie al Teorema di Hahn-Banach.

Questo implica anche che per ogni $x \in X$ vale

$$\|x\| = \max\{T(x) : \ T \in X', \ \|T\| \leq 1\} :$$

la disuguaglianza è ovvia, l'uguale si ottiene usando il funzionale appena costruito!

Altra conseguenza dell'asserto appena dimostrato è il fatto che *il duale di uno spazio normato X separa i punti*: dati due punti $x, y \in X$ qualunque, è sempre possibile trovare un elemento del duale che assume *valori diversi* su di essi. Per convincersene, basta prendere un funzionale che vale $\|x-y\|$ sul vettore $x-y$: allora $T(x)-T(y) = T(x-y) \neq 0$. Val la pena di dire che questa proprietà non è vera per uno spazio vettoriale topologico generale: esistono onestissimi s.v.t. (metrizzabili e persino completi) il cui duale contiene solo il funzionale identicamente nullo!

ESERCIZIO: In generale, dato un funzionale lineare T limitato su un sottospazio Y di $(X, \|\cdot\|)$, la sua estensione ad un elemento $\tilde{T} \in X'$ con la stessa norma, data dal Teorema di Hahn-Banach, non è evidentemente unica: è facilissimo fare degli esempi anche in dimensione finita.

Si mostri però che se Y è un sottospazio *denso* di X, ossia se $\overline{Y} = X$, allora l'estensione è unica. Vi viene in mente un modo per dimostrare l'esistenza di questa estensione senza ricorrere al teorema di Hahn-Banach?

La nostra analisi sulle conseguenze del Teorema di Hahn-Banach è ben lungi dall'essere conclusa. Prima di procedere, però, è utile fare una digressione un po' più concreta, in modo da "fabbricare" un numero sufficiente di esempi di spazi di Banach in dimensione infinita. In particolare, nella prima lezione abbiamo introdotto gli spazi ℓ^p senza dimostrare nulla: vogliamo dunque verificare che la norma $\|\cdot\|_{\ell^p}$ è una norma e che si tratta di spazi completi. Faremo poi lo stesso anche per gli spazi $L^p(\Omega)$ (con Ω aperto di $\mathbf{R}^n$) e $L^p(\mu)$, con μ misura esterna qualunque.

A questo fine, cominciamo a dimostrare una semplice disuguaglianza in $\mathbf{R}$.

Dato un numero reale $p \in (1, +\infty)$, il suo esponente coniugato q si definisce come l'unico numero reale tale che

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Come naturale estensione, l'esponente coniugato di $p = 1$ è per definizione $q = +\infty$, e viceversa. Vale la seguente *disuguaglianza di Young*:

LEMMA: Se p, q sono esponenti coniugati, $1 < p < +\infty$, allora

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q \qquad \forall a > 0, b > 0.$$

DIM.: Usando la concavità della funzione logaritmo otteniamo:

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

da cui segue immediatamente la tesi.

Prima di dimostrare il prossimo teorema, è utile ricordare la definizione della norma $L^\infty(\mu)$, con μ una generica misura esterna.

DEFINIZIONE (Spazi $L^\infty(\Omega)$): Se μ è una misura esterna su X e $u : X \to \overline{\mathbf{R}}$ è misurabile, definiamo

$$\|u\|_{L^\infty(\mu)} = \text{esssup}\{|u(x)| : \ x \in X\} := \inf\{t \in \mathbf{R} : \ \mu(\{x \in X : \ |u(x)| > t\}) = 0\}.$$

Si tratta sostanzialmente del sup di $|u(x)|$ *a meno di insiemi di misura nulla.* Naturalmente, $L^\infty(\mu)$ sarà lo spazio delle funzioni con norma finita, quozientato rispetto alla solita relazione di equivalenza.

Uno strumento utilissimo per lo studio degli spazi ℓ^p, L^p è la disuguaglianza di Hölder: essa ci servirà in particolare per dimostrare che si tratta di spazi normati.

PROPOSIZIONE (Disuguaglianza di Hölder): Dato un numero reale $p \in (1, +\infty)$, il suo esponente coniugato q è l'unico numero reale tale che

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Come naturale estensione, l'esponente coniugato di $p = 1$ è $q = +\infty$, e viceversa.

Se $1 \leq p \leq +\infty$ e q è il suo esponente coniugato si ha, per ogni coppia di funzioni μ-misurabili $u, v : X \to \overline{\mathbf{R}}$ (rispetto a μ, misura esterna su X):

$$\int_X |u(x)v(x)| \, d\mu(x) \leq \|u\|_{L^p(\mu)} \, \|v\|_{L^q(\mu)}.$$

DIM.: Affrontiamo dapprima il caso $1 < p < +\infty$: il caso $p = 1, \ q = \infty$ sarà studiato in seguito.

La disuguaglianza di Hölder è una semplice conseguenza della disuguaglianza di Young. Innanzitutto, notiamo che tutte le nostre p-norme e q-norme sono omogenee: moltiplicando l'argomento per $\lambda \in \mathbf{R}$, la norma risulta moltiplicata per $|\lambda|$. Per questo motivo, possiamo supporre senza perdita di generalità he sia $\|u\|_{L^p(\mu)} = \|v\|_{L^q(\mu)} = 1$. In tal caso, grazie a Young:

$$|u(x)v(x)| \leq \frac{1}{p}|u(x)|^p + \frac{1}{q}|v(x)|^q,$$

da cui integrando su X rispetto alla misura μ

$$\int_X \|uv|\ d\mu \leq \frac{1}{p} + \frac{1}{q} = 1,$$

che è la disuguaglianza voluta[8].

Rimane da vedere il caso limite $p = 1$, $q = +\infty$: possiamo supporre $\|v\|_{L^\infty} < +\infty$ (altrimenti la disuguaglianza è ovvia). Se $t > \|v\|_\infty$, allora $A = \{x \in X :\ |v(x)| > t\}$ ha misura 0: l'integrale di una qualunque funzione misurabile non negativa su X o su $X \setminus A$ hanno lo stesso valore. Allora:

$$\int\limits_X |u(x)v(x)|\ d\mu(x) \leq t \int\limits_X |u(x)|\ d\mu(x) = t\|u\|_{L^1},$$

da cui segue la disuguaglianza voluta facendo tendere t a $\|v\|_\infty$. Q.E.D.

Dalla disuguaglianza di Hölder segue facilmente che le p-norme sono tali:

TEOREMA: Se $1 \leq p \leq +\infty$, *gli spazi* $(L^p(\mu), \|\cdot\|_{L^p(\mu)})$ *sono spazi vettoriali normati.*

DIM.: Le nostre "candidate norme" sono evidentemente non negative ed omogenee. Inoltre, sono finite nei nostri spazi e sono non degeneri: è un semplice esercizio - che tra l'altro abbiamo già svolto - verificare che

$$\int\limits_X |u(x)|^p\ d\mu(x) = 0 \Rightarrow u(x) = 0 \ \textit{per}\ \mu - \textit{q.o.}\ x \in X,$$

per cui la norma L^p è non degenere grazie alla relazione di equivalenza che abbiamo messo sulle funzioni misurabili. Rimane da dimostrare la disuguaglianza triangolare (che in questo caso particolare prende il nome di *disuguaglianza di Minkowski*). Ancora una volta, questa è quasi ovvia nei casi $p = 1$ e $p = +\infty$.

Proviamo la disuguaglianza di Minkowski per $1 < p < +\infty$. Prendiamo due funzioni misurabili u, v e usiamo la disuguaglianza di Hölder:

$$\int\limits_X |u+v|^p \leq \int\limits_X |u||u+v|^{p-1} + \int\limits_X |v||u+v|^{p-1} \leq (\|u\|_p + \|v\|_p) \left(\int\limits_X |u+v|^p \right)^{(p-1)/p}.$$

La disuguaglianza voluta si ottiene allora dividendo ambo i membri per $\left(\int\limits_\Omega |u+v|^p \right)^{(p-1)/p}$. Occorre però sapere che questa quantità è finita non

[8] I casi in cui una delle due norme vale 0, oppure se una o entrambe valgono $+\infty$, sono ovvi!

appena lo sono $\|u\|_p$ e $\|v\|_q$: questo segue immediatamente dalla convessità della funzione $s \mapsto s^p$ in quanto

$$\left|\frac{|u(x)|+|v(x)|}{2}\right|^p \le \frac{1}{2}|u(x)|^p + \frac{1}{2}|v(x)|^p.$$

Q.E.D.

OSSERVAZIONE: Quando detto per gli spazi $L^P(\mu)$ si applica ovviamente agli spazi ℓ^p con le loro norme: abbiamo visto infatti che essi non sono altro che un caso particolare dei primi, con $X = \mathbf{N}$ e μ la counting measure.

A questo punto, occorre dimostrare che gli spazi $L^p(\mu)$ con le rispettive norme sono spazi di Banach:

TEOREMA (di Riesz-Fischer): Sia μ una misura esterna su un insieme X, $1 \le p \le +\infty$. Allora gli spazi $L^p(\mu)$ sono completi. Inoltre, data una successione $\{f_n\}$ convergente ad una funzione f in norma L^p, è possibile estrarre una sottosuccessione $\{f_{n_k}\}$ tale che $f_{n_k}(x) \to f(x)$ per quasi ogni $x \in X$.

DIM.: Dimostriamo dapprima il caso $1 \le p < +\infty$: il caso $p = +\infty$ è diverso (e fortunatamente più facile!) e verrà trattato a parte.

Sia dunque $\{f_n\}$ una successione di Cauchy in $L^p(\mu)$: è allora facile costruire una successione crescente di interi n_k tale che

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p} \le \frac{1}{2^k}, \quad k = 1, 2, 3, \dots$$

Consideriamo poi le funzioni

$$g_K(x) = \sum_{k=1}^{K} |f_{n_{k+1}}(x) - f_{n_k}(x)|, \quad g(x) = \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|$$

(dove l'ultima serie ha perfettamente senso perché i suoi termini sono non negativi). Usando la disuguaglianza triangolare e la nostra scelta di n_k vediamo subito che $\|g_K\|_{L^p} \le 1$ per ogni K. Inoltre, per il teorema di Beppo Levi si ha $\|g_K\|_{L^p} \to \|g\|_{L^p}$, da cui si deduce che $g \in L^p(\mu)$ e di conseguenza $|g(x)| < +\infty$ per quasi ogni $x \in X$.

Consideriamo ora le somme telescopiche

$$\sum_{k=1}^{K} (f_{n_{k+1}}(x) - f_{n_k}(x)) = f_{n_K}(x) - f_{n_1}(x)$$

e la corrispondente serie $\sum_{k=1}^{\infty}(f_{n_{k+1}}(x)-f_{n_k}(x))$: quest'ultima è assolutamente convergente per quasi ogni x (la serie dei moduli converge infatti a $g(x)$), per cui per quasi ogni $x \in X$ esiste il limite

$$\lim_{k\to+\infty} f_{n_k}(x) =: f(x).$$

Nell'insieme di misura nulla sul quale non c'è convergenza possiamo definire f a piacimento, per esempio $f(x) = 0$.

Inoltre, dalle disuguaglianze precedenti segue subito che

$$|f_{n_k}(x)| \leq (f_{n_1}(x) + g(x)) \in L^p(\mu),$$

e il teorema della convergenza dominata di Lebesgue ci permette di concludere che $f_{n_k} \to f$ in $L^p(\Omega)$: lo applichiamo una prima volta a $|f_{n_k}|^p$ per dedurre che $f \in L^p$, poi a $|f_{n_k} - f|^p$ per dedurre la convergenza in norma.

È poi un semplicissimo esercizio verificare che se una successione di Cauchy ha una sottosuccessione convergente, in realà *tutta* la successione converge (questo fatto vale in uno spazio metrico qualunque).

Rimane da vedere il caso $p = +\infty$: lo vedremo domani.

10 Lezione del 29/10/2009 (2 ore)

Vediamo il caso $p = +\infty$ nel Teorema di Riesz-Fischer. Sia dunque $\{f_n\}$ di Cauchy in L^∞. Per ogni $k \in \mathbf{N}$ esiste un indice n_k tale che $\|f_m - f_n\|_\infty < \frac{1}{k}$ per ogni $m, n \geq n_k$. Allora gli insiemi

$$A_k = \{x \in X : |f_n(x) - f_m(x)| > 1/k \textit{ per qualche } m, n \geq n_k\}, \quad A = \bigcup_{k=1}^{\infty} A_k$$

hanno tutti misura nulla (per definizione di norma L^∞). Se ne deduce che per ogni $x \in X \setminus A$ la successione $f_n(x)$ è di Cauchy in $\mathbf{R}$, e converge ad un limite $f(x)$ che al solito estendiamo ponendo $f(x) = 0$ per ogni $x \in A$. Passando al limite per $m \to +\infty$ nella disuguaglianza

$$|f_n(x) - f_m(x)| \leq 1/k \quad \forall x \in X \setminus A,\ \forall m, n \geq n_k$$

si ottiene

$$|f_n(x) - f_(x)| \leq 1/k \quad \forall x \in X \setminus A,\ \forall n \geq n_k,$$

da cui $\|f_n - f\|_\infty \leq 1/k$ per ogni $n > n_k$ e $f_n \to f$ in L^∞.

Per concludere, ci resta l'ultimo asserto sul fatto che la convergenza $f_n \to f$ in L^p implica la convergenza puntuale quasi ovunque di una sottosuccessione. Evidentemente, se f_n converge in L^p, essa è una successione di Cauchy. Il teorema appena dimostrato fornisce una sottosuccessione che converge puntualmente quasi ovunque ed in L^p ad una certa funzione: quest'ultima deve coincidere con f quasi ovunque per l'unicità del limite in L^p. Q.E.D.

OSSERVAZIONE: Abbiamo visto dunque che, in generale, se $1 \le p < +\infty$, la convergenza in $L^p(\mu)$ non implica la convergenza puntuale quasi ovunque di *tutta* la successione. Questa implicazione è invece vera nel caso di L^∞ (e segue dalla nostra dimostrazione del teorema di completezza!).

Un esempio di successione $\{u_i\}$ che converge a 0 il $L^p([0,1])$ (per ogni $1 \le p < +\infty$) ma non converge puntualmente in nessun punto dell'intervallo $[0,1]$ si può costruire come segue: se $i \in [2^k, 2^{k+1}-1] \cap \mathbf{N}$, poniamo

$$u_i(x) = \mathbf{1}_{[0,2^{-k}]}\left(x - \frac{i-2^k}{2^k}\right).$$

ESERCIZIO: Si faccia vedere che se Ω ha misura di Lebesgue finita, allora una funzione in $L^p(\Omega)$ appartiene a $L^r(\Omega)$ per ogni $r \in [1,p]$ (cioè gli spazi L^p diventano sempre più piccoli all'aumentare dell'esponente). Si mostri con un esempio che questo non è vero se la misura di Ω è infinita. *[SUGG.: Applicare la disuguaglianza di Hölder al prodotto $|u(x)|^r \cdot 1$, prendendo come esponenti p/r ed il coniugato...]*

Ad esercitazioni avete studiato esplicitamente i duali degli spazi di successioni ℓ^p:

PROPOSIZIONE: Sia $1 \le p \le +\infty$, q coniugato a p, $\{y_k\} \in \ell^q$. Definiamo un'applicazione lineare $T_{\{y_k\}} : \ell^p \to \mathbf{R}$ nel modo seguente:

$$(*)\ T_{\{y_k\}}(\{x_k\}) = \sum_{k=1}^{\infty} y_k x_k.$$

Allora l'applicazione $\Phi : \{y_k\} \mapsto T_{\{y_k\}}$ è una ben definita isometria lineare di ℓ^q in $(\ell^p)'$. Se poi $1 \le p \le +\infty$, allora la mappa Φ è suriettiva, e costituisce dunque un'isomorfismo isometrico tra ℓ^q ed il duale di ℓ^p.

OSSERVAZIONE/ESERCIZIO: Possiamo costruire in modo analogo anche un'iniezione isometrica di $L^q(\Omega)$ nel duale di $L^p(\Omega)$ (con questa notazione indichiamo gli spazi rispetto alla misura di Lebesgue...): se $1 \le p \le +\infty$

e q è l'esponente coniugato, ad una funzione $v \in L^q(\Omega)$ possiamo associare l'elemento $T_v \in (L^p)'$ definito da

$$T_v(u) = \int\limits_\Omega u(x)v(x)\,dx \quad \forall u \in L^p(\Omega).$$

Mostriamo che $\Phi : v \mapsto T_v$ è un'isometria lineare tra $L^q(\Omega)$ e $(L^p(\Omega))'$.

Infatti, dalla disuguaglianza di Hölder segue subito che $\|T_v\|_{(L^p)'} \leq \|v\|_{L^q}$. Se poi $q < +\infty$, poniamo $u(x) = sgn(v(x))|v(x)|^{q-1}/\|v\|_{L^q}^{q-1}$: questa funzione ha norma L^p unitaria e $T_v(u) = \|v\|_{L^q}$, per cui vi è uguaglianza delle norme.

Se infine $q = +\infty$, osserviamo che per ogni $\varepsilon > 0$ esiste un sottinsieme $A \subset \Omega$ tale che $0 < m(A) < +\infty$ e $|v(x)| \geq \|v\|_{L^\infty} - \varepsilon$ per ogni $x \in A$. Si consideri la funzione

$$u(x) = \begin{cases} \mathrm{sgn}(v(x))/m(A) & \textit{se } x \in A, \\ 0 & \textit{altrimenti} \end{cases}$$

Allora $\|u\|_{L^1} = 1$ e $\int\limits_\Omega u(x)v(x)\;dx \geq \|v\|_{L^\infty} - \varepsilon$. Dall'arbitrarietà di ε deduciamo l'uguaglianza delle norme.

Lo stesso risultato vale per $L^p(\mu)$ con μ misura esterna qualunque...con un minimo di ipotesi in più nel caso $p = 1$, $q = +\infty$ (per esempio, tutto funziona se μ è σ-finita).

OSSERVAZIONE/ESERCIZIO: Anche l'isomorfismo di cui al teorema precedente vale anche per gli spazi $L^p(\Omega)$: il duale di $L^p(\Omega)$ è $L^q(\Omega)$ per $1 \leq p < +\infty$. Stessa cosa per $L^p(\mu)$, con qualche attenzione sulla misura nel caso $p = 1$.

La dimostrazione richiede risultati un po' più raffinati di teoria della misura, e la vedremo in seguito.

Invece, anche in questo caso il duale di $L^\infty(\Omega)$ è strettamente più grande di $L^1(\Omega)$. Mostriamo subito un esempio di un funzionale lineare $T \in (L^\infty([-1,1]))'$ per il quale *non esiste* alcuna funzione $v(x) \in L^1([-1,1])$ tale che

$$T(u) = \int_{-1}^{1} uv\,dx \quad \forall u \in L^\infty.$$

Definiamo dapprima T sul sottospazio $C^0([-1,1])$ ponendo $T(u) = u(0)$ per ogni funzione continua u. Questo è un funzionale lineare di norma 1 (verificarlo!), che può essere esteso ad un funzionale di norma 1 definito su tutto lo spazio L^∞ con il teorema di Hahn-Banach. Si verifichi che se per assurdo esistesse v come sopra, essa dovrebbe essere nulla per quasi ogni $x \in [-1,1]$ (Si usi il seguente fatto: data la funzione $\mathrm{sgn}(v(x))$, è possibile costruire una successione di funzioni continue u_h con $u_h(x) \to \operatorname{sgn}(v(x))$ per quasi ogni $x \in [-1,1]$, $u_h(0) = 0$ e $-1 \le u_h(x) \le 1$ per ogni $x \in [-1,1]$. L'esistenza di una tale successione viene da teoremi di approssimazione di funzioni L^p con funzioni regolari, che dimostreremo verso la fine del corso. Si usi poi il teorema della convergenza dominata.). Questo è contraddittorio perché $T \not\equiv 0$.

Nel nostro studio dei duali degli spazi ℓ^p (e $L^p(\Omega)$) abbiamo a volte fatto uso del teorema di Hahn-Banach. È giunto ora il momento di vederne un paio di conseguenze "geometriche" relative alla possibilità di separare insiemi convessi con un iperpiano chiuso.

A questo scopo, si rivela fondamentale la definizione di *funzionale di Minkowski* associato a un convesso:

PROPOSIZIONE (Funzionale di Minkowski): Sia $(X, \|\cdot\|)$ uno spazio normato, C un sottinsieme aperto e convesso di X contenente l'origine. Definiamo il funzionale di Minkowski associato a C nel modo seguente:

$$p(x) = \inf\{t > 0 : \ \frac{x}{t} \in C\}.$$

La funzione $p(x)$ è una ben definita funzione reale positivamente omogenea e subadditiva[9]*. Vale inoltre, per un'opportuna costante $K > 0$,*

$$(*) \ p(x) \le K\|x\| \quad \forall x \in X,$$

e infine

$$(**) \ C = \{x \in X : \ p(x) < 1\}.$$

DIM.: Sia $r > 0$ tale che $B_r(0) \subset C$ (esiste perché C è aperto): per ogni $x \in X$ si ha allora che $r\frac{x}{2\|x\|} \in C$, da cui $p(x) \le \frac{2}{r}\|x\|$ e la (*) è dimostrata. In particolare, $p(x)$ è sempre finito.

È poi immediato verificare che $p(x)$ è positivamente omogenea. Dimostriamo (**): se $x \in C$, grazie al fatto che C è aperto esiste $r > 0$ tale che $(1+r)x \in C$. Ne segue che $p(x) \le \frac{1}{1+r} < 1$. Viceversa, se $p(x) < 1$ posso trovare $0 \le t < 1$ tale che $\frac{x}{t} \in C$. Ma allora $x = t(\frac{x}{t}) + (1-t)0 \in C$ grazie alla convessità di C.

Rimane da provare la subadditività (la disuguaglianza triangolare): Siano $x, y \in X$. Per definizione del funzionale di Minkowski, abbiamo che per ogni $\varepsilon > 0$ vale

$$\frac{x}{p(x)+\varepsilon} \in C, \quad \frac{y}{p(y)+\varepsilon} \in C.$$

[9] Si veda l'enunciato del Teorema di Hahn-Banach

Prendiamo una combinazione convessa di questi due punti di C con

$$t = \frac{p(x) + \varepsilon}{p(x) + p(y) + 2\varepsilon}, \quad (1 - t) = \frac{p(y) + \varepsilon}{p(x) + p(y) + 2\varepsilon}.$$

Ne deduciamo immediatamente (grazie alla convessità di C) che

$$\frac{x + y}{p(x) + p(y) + 2\varepsilon} \in C,$$

da cui

$$p(x + y) \leq p(x) + p(y) + 2\varepsilon.$$

Grazie all'arbitrarietà di ε otteniamo la subadditività. Q.E.D.

ESERCIZIO: Si mostri che se $C = B_1(0)$ è la palla aperta unitaria del nostro spazio normato, allora $p(x) = \|x\|$.

Le conseguenze geometriche del teorema di Hahn-Banach riguardano la possibilità di separare due convessi disgiunti con un iperpiano. Ci servirà un lemma che ha un certo interesse anche indipendentemente:

LEMMA: Sia C un convesso aperto e non vuoto di X, $x_0 \in X \setminus C$. Allora esiste $T \in X'$ tale che

$$T(x) < T(x_0) \quad \forall x \in C.$$

DIM.: Scegliamo $\overline{y} \in C$ e poniamo $\tilde{C} = C - \{\overline{y}\}$, $\overline{x} = x_0 - \overline{y}$. A questo punto, $\tilde{C}$ è un aperto convesso contenente l'origine e $\overline{x} \notin C$. Se p è il funzionale di Minkowski associato a $\tilde{C}$, si ha $p(\overline{x}) \geq 1$ mentre $\tilde{C} = \{x : \ p(x) < 1\}$.

Definiamo sul sottospazio unidimensionale $Y = \mathbf{R}\{\overline{x}\}$ il funzionale lineare $T \in Y'$ tale che $T(t\overline{x}) = tp(\overline{x})$. Usando la positiva omogeneità del funzionale di Minkowski si vede subito che $T(t\overline{x}) \leq p(t\overline{x})$ per ogni $t \in \mathbf{R}$.

Per il teorema di Hahn-Banach, possiamo estendere T ad un funzionale lineare definito su tutto X in modo che sia $T(x) \leq p(x)$ per ogni $x \in X$. Si ha poi $T \in X'$ grazie alla (*) della proposizione precedente.

Abbiamo allora $T(x) < 1$ per ogni $x \in \tilde{C}$, mentre $T(\overline{x}) = p(\overline{x}) \geq 1$. Aggiungendo $\overline{y}$ e usando la linearità di T si deduce che $T(x) < 1 + T(\overline{y})$ per ogni $x \in C$, mentre $T(x_0) \geq 1 + T(\overline{y})$. Q.E.D.

11 Lezione del 30/10/2009 (2 ore)

Siamo ora in grado di enunciare e dimostrare la "versione geometrica" del teorema di Hahn-Banach:

TEOREMA (Hahn-Banach, versione geometrica): Siano $(X, \|\cdot\|)$ *uno spazio normato,* A*,* B *sottinsiemi convessi disgiunti e non vuoti di* X*. Allora*

(i) Se A *è aperto, esistono* $T \in X'$*,* $T \neq 0$*,* $\alpha \in \mathbf{R}$ *tali che*

$$T(x) \leq \alpha \leq T(y) \quad \forall x \in X, \ \forall y \in Y.$$

Geometricamente, possiamo dire che l'iperpiano (affine) chiuso di equazione $T(x) = \alpha$ *separa i convessi* A *e* B *in senso largo.*

(ii) Se A *e* B *sono chiusi e* A *è anche compatto, allora esistono* $T \in X'$*,* $T \neq 0$*,* $\alpha \in \mathbf{R}$ *e* $\varepsilon > 0$ *tali che*

$$T(x) \leq \alpha - \varepsilon \ \forall x \in A, \quad \alpha + \varepsilon \leq T(y) \ \forall y \in B.$$

Geometricamente, possiamo dire che l'iperpiano (affine) chiuso di equazione $T(x) = \alpha$ *separa i convessi* A *e* B *in senso stretto.*

.

DIM. DEL TEOREMA: Cominciamo a dimostrare (i): poniamo $C = A - B = \{x - y : \ x \in A, \ y \in B\}$ (attenzione: stiamo facendo una differenza "algebrica" e non insiemistica!).

È immediato verificare che C è convesso. È poi aperto in quanto esprimibile come unione di aperti: $A - B = \bigcup_{y \in B} (A - \{y\})$.

Applichiamo il lemma con $x_0 = 0$: si noti infatti che $0 \notin C$ perché A e B sono disgiunti. Troviamo $T \in X'$ tale che

$$T(x) < 0 = T(0) \quad \forall x \in C,$$

da cui (linearità di T)

$$T(x) < T(y) \quad \forall x \in A, \ \forall y \in B.$$

La prima parte del Teorema è allora provata se poniamo $\alpha = \sup\{T(x) : \ x \in A\}$.

Per provare la (ii), poniamo $A_\varepsilon = A + B_\varepsilon(0)$, $B_\varepsilon = B + B_\varepsilon(0)$. È immediato verificare che si tratta di insiemi aperti convessi.

Dico che $A_\varepsilon \cap B_\varepsilon = \emptyset$ per ε abbastanza piccolo: se così non fosse, potremmo trovare una successione $z_n \in A_{\frac{1}{n}} \cap B_{\frac{1}{n}}$. Ora, è possibile scrivere $z_n = a_n + w_n = b_n + w'_n$, con $a_n \in A$, $b_n \in B$ e $\|w_n\| < 1/n$, $\|w'_n\| < 1/n$. Grazie alla compattezza di A e a meno di sottosuccessioni, abbiamo $a_n \to \overline{a} \in A$. Ma allora (visto che $w_n \to 0$ e $w'_n \to 0$) si ha anche $b_n \to \overline{a}$, da cui $\overline{a} \in B$ (per la chiusura di B). Questo è assurdo perché per ipotesi $A \cap B = \emptyset$.

Prendiamo allora ε abbastanza piccolo, e applichiamo la parte (i) del teorema ai convessi A_ε, B_ε: troviamo $T \in X'$ e $\alpha \in \mathbf{R}$ tali che

$$\begin{aligned} T(a+w) &\leq \alpha \quad \forall a \in A,\ w \in B_\varepsilon(0) \\ T(b+w') &\geq \alpha \quad \forall b \in B,\ w' \in B_\varepsilon(0). \end{aligned}$$

Passando al sup su w e all'inf su w' in queste due disuguaglianze si ottiene

$$\begin{aligned} T(a) + \varepsilon\|T\| &\leq \alpha \quad \forall a \in A \\ T(b) - \varepsilon\|T\| &\geq \alpha \quad \forall b \in B. \end{aligned}$$

Q.E.D.

OSSERVAZIONE: Senza qualche ipotesi sui convessi disgiunti A,B, in generale non è possibile trovare un iperpiano che li separi: si pensi al caso in cui A è un sottospazio vettoriale denso di uno spazio di dimensione infinita X, mentre $B = \{x_0\}$ contiene un unico punto non in A. Se esistesse un funzionale T come in (i), esso sarebbe superiormente limitato (da $T(x_0)$...). Poiché A è un sottospazio vettoriale, T deve allora essere identicamente nullo su A. Infine, poiché A è denso, T deve essere il funzionale nullo.

Ecco un'utile conseguenza di questo teorema "geometrico":

COROLLARIO: Sia $(X, \|\cdot\|)$ uno spazio normato, Y un suo sottospazio vettoriale. Allora Y è denso se e soltanto se ogni funzionale lineare $T \in X'$ che si annulla identicamente su Y è identicamente nullo.

DIM.: È ovvio che se Y è denso, un funzionale lineare continuo che si annulli su Y deve essere identicamente nullo.

Viceversa, supponiamo che Y non sia denso, ossia che $\overline{Y}$ sia un sottospazio chiuso proprio di X. Mostriamo che c'è un funzionale lineare non identicamente nullo che si annulla su Y. A questo scopo, prendiamo $x_0 \in X \setminus \overline{Y}$ e applichiamo l'enunciato (ii) del teorema precedente agli insiemi convessi $\overline{Y}$ e $\{x_0\}$ (il secondo dei quali è compatto). Esiste allora $T \in X'$ tale che $T(x) < T(x_0)$ per ogni $x \in \overline{Y}$. Dalla linearità di T e dal fatto che Y è un sottospazio, segue subito che $T \equiv 0$ in Y (un funzionale lineare non nullo non ha MAI immagine superiormente limitata!). Dunque, $T(x) = 0$ per ogni $x \in Y$, mentre $T(x_0) \neq 0$. Q.E.D.

Un altro risultato importante per lo studio del duale di uno spazio normato è il seguente

TEOREMA (di Banach-Steinhaus): Sia $(X, \| \cdot \|)$ *uno spazio di Banach,* $\{T_k\} \subset X'$ *una successione di funzionali lineari continui su* X *che sia puntualmente limitata, cioè tale che*

$$\sup\{|T_k(x)| : k \in \mathbf{N}\} < +\infty \quad \forall x \in X.$$

Allora la successione è uniformemente limitata, ossia esiste $C > 0$ *tale che* $\|T_k\|_{X'} \leq C$ *per ogni* $k \in \mathbf{N}$.

Per dimostrare il Teorema di Banach-Steinhaus, useremo il seguente importante risultato:

TEOREMA (di Baire): Sia (X, d) *uno spazio metrico completo. Se* $\{F_k\}$ *è una successione di insiemi chiusi in* X *privi di punti interni, allora* $\bigcup_{k \in \mathbf{N}} F_k$ *è privo di punti interni.*

Passando ai complementari, possiamo dire che in uno spazio metrico completo intersezione numerabile di aperti densi è un insieme denso.

DIM.: Sia $\Omega \subset X$ un qualunque sottinsieme aperto non vuoto: dobbiamo far vedere che $\Omega \setminus \bigcup_{k \in \mathbf{N}} F_k$ è non vuoto.

Scegliamo $x_1 \in X$, $1 > r_1 > 0$ tali che $B_{r_1}(x_1) \subset \Omega$ e $B_{r_1}(x_1) \cap F_1 = \emptyset$: questo è possibile perché il complementare di F_1 è un aperto denso. A questo punto, notiamo che esistono punti del complementare di F_2 dentro $B_{r_1}(x_1)$ (perché F_2 non ha punti interni): possiamo dunque scegliere x_2, r_2 tali che $B_{r_2}(x_2) \cap F_2 = \emptyset$, $\overline{B_{r_2}(x_2)} \subset B_{r_1}(x_1)$ e $r_2 < 1/2$.

Procedendo in questo modo, possiamo costruire una successione $\{x_k\} \subset X$ e una successione di numeri reali positivi $\{r_k\}$ tali che

$$B_{r_k}(x_k) \cap F_k = \emptyset,\ \overline{B_{r_k}(x_k)} \subset B_{r_{k-1}}(x_{k-1}),\ r_k < 1/k.$$

Grazie al fatto che il raggio di queste palle tende a zero, e per come esse sono "inscatolate", si vede subito che $\{x_k\}$ è una successione di Cauchy e che il suo limite $\overline{x}$ è tale che

$$\overline{x} \in \bigcap_{k \in \mathbf{N}} B_{r_k}(x_k).$$

D'altra parte,

$$\bigcap_{k \in \mathbf{N}} B_{r_k}(x_k) \subset \Omega \setminus \bigcup_{k \in \mathbf{N}} F_k,$$

per cui quest'ultimo insieme è non vuoto. Q.E.D.

La prossima volta vedremo la dimostrazione del teorema di Banach-Steinhaus.

12 Lezione del 4/11/2009 (2 ore)

Passiamo alla dimostrazione del teorema di Banach-Steinhaus:

DIM. del Teorema di Banach-Steinhaus): Poniamo

$$F_k = \{x \in X : \ |T_h(x)| \leq k \ \forall h \in \mathbf{N}\}.$$

Questi insiemi sono chiusi, e la loro unione è tutto X grazie all'ipotesi di limitatezza puntuale. Per il Teorema di Baire, deve esistere un indice $\overline{k} \in \mathbf{N}$ tale che $F_{\overline{k}}$ ha parte interna non vuota. Scegliamo dunque $\overline{x} \in X$, $r > 0$ tali che $B_r(\overline{x}) \subset F_{\overline{k}}$: si ha allora

$$|T_h(\overline{x} + y)| \leq \overline{k} \quad \forall x \in X, \|x\| < r, \ \forall h \in \mathbf{N}.$$

Se poi $x \in X$, $\|x\| \leq 1$ si ha per ogni h:

$$|T_h(x)| = \frac{2}{r}|T_h\left(\frac{r}{2}x\right)| = \frac{2}{r}(|T_h\left(\overline{x} + \frac{r}{2}x - \overline{x}\right)| \leq$$
$$\frac{2}{r}\left(|T_h\left(\overline{x} + \frac{r}{2}x\right)| + |T_h(\overline{x})|\right) \leq \frac{4}{r}\overline{k}.$$

Passando al sup su x si ha la tesi. Q.E.D.

OSSERVAZIONE/ESERCIZIO: Vediamo subito una facile ma importante conseguenza del Teorema di Banach-Steinhaus: se $\{T_k\} \subset X'$ è una successione di funzionali lineari tale che $T_k(x) \to T(x) \in \mathbf{R}$ per ogni $x \in X$ (cioè T_n tende puntualmente ad una certa funzione reale T), allora $T \in X'$ e

$$\|T\| \leq \liminf_{k \to +\infty} \|T_k\|.$$

Si dimostri questo risultato, e si mostri con un esempio che non è detto che $T_k \to T$ in X'. *(L'enunciato è una conseguenza abbastanza diretta del Teorema di Banach-Steinhaus: la linearità di T è ovvia, mentre il teorema consente di dire che i T_k sono equilimitati in norma... dunque il limite puntuale T è limitato. La disuguaglianza sulle norme è una facile conseguenza. Infine, per il controesempio richiesto si consideri lo spazio ℓ^2 e la successione (di successioni) e^k costituita dai vettori della base canonica, vista come elemento del duale di ℓ^2...)*

Introduciamo un paio di concetti fondamentali: quelli di spazio biduale e di spazio riflessivo.

DEFINIZIONE (biduale, spazio di Banach riflessivo): Sia $(X, \|\cdot\|)$ uno spazio normato. Indichiamo con X'' il suo *biduale*, ossia lo spazio vettoriale dei funzionali lineari e continui sul duale X'.

Come forse ci è già noto dal corso di algebra lineare, esiste un modo canonico di vedere un elemento $x \in X$ come elemento del biduale. Precisamente, associamo ad x il funzionale $S_x : X' \to \mathbf{R}$ definito da

$$S_x(T) = T(x) \quad \forall T \in X'.$$

La mappa $J : X \to X''$ che manda x in S_x è un'*isometria lineare*: infatti, come conseguenza del Teorema di Hahn-Banach, abbiamo visto che si può scrivere

$$\|x\| = \max\{T(x) : \ T \in X', \ \|T\|_{X'} \leq 1\} \quad \forall x \in X,$$

da cui $\|x\| = \|S_x\|_{X''}$.

In dimensione finita, la mappa J è un isomorfismo (canonico) tra X e X''. In dimensione infinita, invece, può essere che $J(X)$ sia un sottinsieme proprio di X''. Uno spazio normato si dice *riflessivo* se $J(X) = X''$ (ossia se lo spazio "coincide col suo biduale").

Grazie al nostro studio dettagliato degli spazi ℓ^p, non è difficile vedere che ℓ^1 e ℓ^∞ non sono riflessivi, mentre gli spazi ℓ^p lo sono per $1 < p < +\infty$.

La riflessività, come la *separabilità*[10] giocano un ruolo essenziale nella teoria delle topologie deboli.

L'immersione di X nel suo biduale ci consente di ottenere un'altra conseguenza molto interessante del Teorema di Banach-Steinhaus:

OSSERVAZIONE/ESERCIZIO: Sia $(X, \|\cdot\|)$ uno spazio di Banach, A un suo sottinsieme. Allora A è limitato se e soltanto se, per ogni $T \in X'$, l'immagine $T(A)$ è un sottinsieme limitato di $\mathbf{R}$. (*SUGG.: È ovvio che se A è limitato allora $T(A)$ è limitata per ogni $T \in X'$. Viceversa, basta far vedere che ogni successione $\{x_n\} \subset A$ tale che $\{T(x_n)\}$ è limitata per ogni $T \in X'$, è limitata in norma. Per avere questo risultato basta applicare il Teorema di Banach-Steinhaus alla successione $S_{x_n} \in X''$: essa è puntualmente limitata per ipotesi, per cui deve essere limitata in norma. Possiamo allora concludere perchè $\|x_n\| = \|S_{x_n}\|_{X''}$.*)

OSSERVAZIONE: Vediamo ora come un risultato "astratto" come il Teorema di Banach-Steinhaus possa avere delle applicazioni molto concrete: motriamo che

[10] *Uno spazio metrico si dice separabile se ha un sottinsieme denso numerabile.*

esistono delle funzioni continue e 2π-periodiche la cui serie di Fourier non converge in tutti i punti.

Se $f : \mathbf{R} \to \mathbf{R}$ è una funzione continua e 2π-periodica, ricordiamo che la sua serie di Fourier è data da

$$a_0/2 + \sum_{n=1}^{\infty}[a_n \cos nx + b_n \sin nx],$$

con

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt.$$

Un noto teorema di convergenza afferma che se f è di classe C^1, allora la serie di Fourier converge in ogni punto a $f(x)$ (e converge anche uniformememente). Le ipotesi del teorema sembrano però 'eccessive": per calcolare i coefficienti di Fourier non c'è certo bisogno della derivabilità! Mostriamo però che se f è soltanto C^0, non è detto che ci sia convergenza puntuale in tutti i punti.

Per farlo, occorre sapere che per la ridotta parziale N-esima della serie di Fourier di f vale la formula seguente:

$$f_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin((N+\frac{1}{2})(y-x))}{2\sin((y-x)/2)} f(y)\, dy$$

(si vedano per esempio i miei appunti del corso di Analisi 3UD, a.a. 2004-2005).

Facciamo vedere che esistono funzioni continue e periodiche per cui non è vero che $f_N(0) \to f(0)$, cioè per cui la serie di Fourier non converge nell'origine. A questo scopo, consideriamo lo spazio di Banach $X = C^0(2\pi)$ delle funzioni continue e 2π-periodiche, dotato della solita norma $\|\cdot\|_\infty$. Se definiamo

$$T_N : f \mapsto \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin((N+\frac{1}{2})y)}{2\sin(y/2)} f(y)\, dy,$$

i funzionali T_N sono elementi ben definiti di $(L^\infty(2\pi))'$, e quindi anche di X' (perchè X è un sottospazio chiuso di L^∞). Inoltre, se poniamo

$$g_N(y) = \frac{\sin((N+\frac{1}{2})y)}{2\sin(y/2)},$$

si può verificare che $\|T_N\|_{X'} = \|g_N\|_{L^1([-\pi,\pi])}$.[11]

[11]E' ovvio che T_N ha la norma indicata come elemento del duale di L^∞, per cui $\|T_n\|_{X'} \leq \|g_N\|_{L^1}$. D'altra parte, è possibile trovare una successione σ_k di funzioni continue e periodiche tali che $\lim\limits_{k\to+\infty} \sigma_k(x) \to \mathrm{sgn}(g_N(x))$ per q.o. $x \in [-\pi, \pi]$ e tali che $\|\sigma_k\|_\infty \leq 1$. Il teorema della convergenza dominata permette di affermare che $T_N(\sigma_k) \to \|g_N\|_{L^1}$, da cui la tesi.

Ora, per come abbiamo costruito i funzionali si ha $f_N(0) = T_N(f)$. Se fosse vero che $f_N(0) \to f(0)$ per ogni $f \in X$, in particolare dovremmo avere

$$\sup_N |T_N(f)| < +\infty \quad \forall f \in X.$$

Il teorema di Banach-Steinhaus ci permetterebbe allora di concludere che le norme dei funzionali T_N sono equilimitate: peccato che questo sia falso, perché $\|g_N\|_{L^1} \to +\infty$ per $N \to +\infty$.[12]

Un altro risultato molto importante è il seguente:

TEOREMA (dell'applicazione aperta): Siano $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ *spazi di Banach,* $T : X \to Y$ *un'applicazione lineare continua e suriettiva. Allora* T *è aperta: esiste* $r > 0$ *tale che* $T(B_1(0)) \supset B_r(0)$.

A domani la dimostrazione!

13 Lezione del 5/11/2009 (2 ore)

DIMOSTRIAMO il teorema dell'applicazione aperta.

Per cominciare, facciamo vedere che esiste $r > 0$ tale che

$$(*) \qquad \overline{T(B_1(0))} \supset B_{2r}(0).$$

Infatti, grazie alla suriettività di T abbiamo $Y = \bigcup_{n=1}^{\infty} \overline{T(B_n(0))}$. Il lemma di Baire garantisce che *almeno uno* di questi chiusi ha parte interna non vuota. Siccome per l'omogeneità della norma e la linearità di T questi insiemi sono omotetici, $\overline{T(B_1(0))}$ ha parte interna non vuota.

Tale parte interna è un simmetrico convesso aperto (chiusura di un convesso simmetrico è convessa e simmetrica, stessa cosa vale per la parte interna...): se ne deduce che 0 appartiene alla parte interna di $\overline{T(B_1(0))}$ e $(*)$ è dimostrato.

Facciamo poi vedere che

$$T(B_1(0)) \supset B_r(0),$$

che è la tesi.

Sia infatti $\|y\| < r$: cerchiamo $x \in X$ tale che $\|x\| < 1$ e $T(x) = y$.

[12] Con un cambio di variabile e ricordando che $|\sin t| \leq |t|$, ci si riconduce alla non sommabilità della funzione $\frac{\sin t}{t}$ sulla semiretta $[1, +\infty)$.

Siccome per $(*)$ abbiamo $B_r(0) \subset \overline{T(B_{1/2}(0))}$, per ogni $\varepsilon > 0$ possiamo trovare $z \in X$ con $\|z\| < 1/2$ e $\|y - T(z)\| < \varepsilon$. Scegliendo $\epsilon = r/2$ troviamo dunque $z_1 \in X$ con $\|z_1\| < 1/2$ e $\|y - T(z_1)\| < r/2$.

Poichè $B_{r/2} \subset \overline{T(B_{1/4}(0))}$, ripetendo lo stesso ragionamento con $y - T(z_1)$ al posto di y ed $\varepsilon = r/4$, trovo $z_2 \in X$ con $\|z_2\| < 1/4$ e $\|y - T(z_1) - T(z_2)\| < r/4$.

Procedendo ricorsivamente allo stesso modo, costruisco una successione $\{z_n\} \subset X$ tale che $\|z_n\| < 1/2^n$ e $\|y - T(z_1 + z_2 + \ldots + z_n)\| < r/2^n$. La successione $x_n = z_1 + z_2 + \ldots + z_n$ è di Cauchy per cui $x_n \to \overline{x}$ in X. Si ha evidentemente $\|\overline{x}\| < 1$ e $y = T(\overline{x})$ per la continuità di T. Q.E.D.

OSSERVAZIONI/COROLLARI: Un'importante conseguenza del teorema è la seguente: se $T : X \to Y$ è un *isomorfismo algebrico* tra spazi di Banach e *T è continuo*, allora l'inversa $T^{-1} : Y \to X$ è continua e T è un isomorfismo di spazi di Banach. Infatti, l'apertura di T fornisce proprio la limitatezza di T^{-1}.

Altra conseguenza interessante: se $\|\cdot\|$ e $\|\cdot\|'$ sono due norme su X, entrambe di Banach, ed esiste $C > 0$ tale che $\|x\| \leq C\|x\|'$ per ogni $x \in X$, allora le due norme sono equivalenti. Basta infatti applicare l'osservazione precedente alla funzione identica tra i due spazi di Banach.

Un'altra importante conseguenza del teorema dell'applicazione aperta è il

TEOREMA (Del grafico chiuso): Siano X, Y spazi di Banach $T : X \to Y$ un'applicazione lineare. Allora T è continua se e solo se il grafico $G_T = \{(x, T(x)) : x \in X\}$ è chiuso in $X \times Y$.

DIM.: Supponiamo che T sia continua e che si abbia $(x_n, T(x_n)) \to (\overline{x}, \overline{y})$. Grazie alla continuità di T si ha $\overline{y} = T(\overline{x})$ per cui $(\overline{x}, \overline{y}) \in G_T$: il grafico è chiuso.

Viceversa, supponiamo che G_T sia chiuso. Allora G_T è un sottospazio vettoriale chiuso dello spazio di Banach $X \times Y$, ed è così esso stesso uno spazio di Banach.

La mappa $\Phi : x \mapsto (x, T(x))$ è chiaramente un isomorfismo algebrico tra X e G_T, con inversa p_1 (la proiezione sul primo fattore di $X \times Y$). Siccome tale proiezione è continua, grazie al teorema dell'applicazione aperta anche Φ lo è. Ma allora lo è anche $T = p_2 \circ \Phi$ (dove p_2 è la proiezione sul secondo fattore). Q.E.D.

Per concludere la nostra analisi generale degli spazi di Banach, vogliamo occuparci del problema della compattezza. In dimensione finita, la relativa abbondanza di insiemi compatti (tutti i chiusi limitati di $\mathbf{R}^n$ lo sono!) ci ha permesso di dimostrare parecchi teoremi. In dimensione infinita, invece, non tutti i chiusi limitati sono compatti:

TEOREMA: Sia $(X, \|\cdot\|)$ uno spazio normato. Allora la dimensione di X è finita se e solo se la palla chiusa unitaria

$$\overline{B} = \{x \in X : \ \|x\| \leq 1\}$$

è compatta.

Per dimostrare il Teorema, ci serve il seguente risultato:

LEMMA (Riesz): Sia $(X, \|\cdot\|)$ uno spazio normato, Y un suo sottospazio chiuso proprio. Allora esiste $\overline{x} \in X$ con $\|x\| = 1$ e $\mathrm{dist}(\overline{x}, Y) > 1/2$.

DIM.: Prendiamo un punto $x_0 \in X \setminus Y$. Siccome Y è chiuso, la distanza δ tra x_0 e Y è positiva. Inoltre, per definizione di distanza, esiste certamente $y_0 \in Y$ tale che $\|x_0 - y_0\| < 2\delta$. Poniamo

$$\overline{x} = \frac{x_0 - y_0}{\|x_0 - y_0\|}.$$

È chiaramente $\|\overline{x}\| = 1$, e se inoltre $y \in Y$ si ha:

$$\|\overline{x} - y\| = \frac{1}{\|x_0 - y_0\|}\|x_0 - (y_0 + y\|x_0 - y_0\|)\| < 1/2,$$

dove si è usato il fatto che $y_0 + y\|x_0 - y_0\| \in Y$. Q.E.D.

Dimostriamo il teorema sulla (non) compattezza della palla unitaria chiusa.

È chiaro che se X ha dimensione finita, allora la palla unitaria chiusa è compatta: possiamo ricondurci sempre a $X = \mathbf{R}^n$ con una norma, che sappiamo sarà equivalente a quella euclidea...

Ne segue che la palla chiusa della norma è anche un chiuso limitato euclideo, e dunque è un compatto (sia euclideo che nella norma di partenza!).

Viceversa, supponiamo che X abbia dimensione infinita. Possiamo allora costruire una successione crescente di sottospazi $Y_1 \subset Y_2 \subset Y_3 \ldots$ in modo tale che $\dim(Y_k) = k$, $k = 1, 2, \ldots$

Fissiamo $x_1 \in Y_1$, $\|x_1\| = 1$. Per il Lemma di Riesz con $X = Y_2$ e $Y = Y_1$, troviamo $x_2 \in Y_2$ con $\|x_2\| = 1$ e $\mathrm{dist}(x_2, Y_1) > 1/2$.

Continuando in questo modo troviamo una successione $\{x_k\}$ tale che $\|x_k\| = 1$, $x_k \in Y_k$ e $\mathrm{dist}(x_k, Y_{k-1}) > 1/2$. Abbiamo così costruito una successione a valori in X, tutta costituita da elementi di norma 1, i cui elementi distano l'uno dall'altro più di $1/2$. Una tale successione non ha evidentemente alcuna sottosuccessione di Cauchy: la palla unitaria chiusa non è quindi compatta. Q.E.D.

Una caratterizzazione efficace dei compatti in uno spazio metrico (e quindi in uno spazio di Banach) è data dal teorema che segue. Premettiamo una definizione:

DEFINIZIONE (Totale limitatezza): Sia (X, d) uno spazio metrico. Un sottinsieme $K \subset X$ si dice *totalmente limitato* se, per ogni $\varepsilon > 0$, è possibile ricoprire K con un *numero finito* di palle di raggio ε.

TEOREMA: Sia (X, d) uno spazio metrico, $K \subset X$. Allora K è compatto se e soltanto se è completo e totalmente limitato.

Inoltre, un sottinsieme K totalmente limitato di uno spazio metrico completo X è relativamente compatto: da ogni successione a valori in K possiamo estrarre una sottosuccessione convergente in X.

DIM.: Sappiamo gi'a che, in uno spazio metrico, compattezza per ricoprimenti e compattezza per successioni sono la stessa cosa.

Supponiamo dunque che K sia compatto: mostriamo che è completo e totalmente limitato. Per quanto riguarda la completezza, sia $\{x_k\} \subset K$ una successione di Cauchy. Per compattezza, essa ha una sottosuccessione convergente ad un punto $\overline{x} \in K$. Ma abbiamo già osservato che quando una successione di Cauchy ha una sottosuccessione convergente, in realtà *tutta* la successione converge!

Mostriamo la totale limitatezza: scelto $\varepsilon > 0$, consideriamo la famiglia di palle aperte $\{B_\varepsilon(x)\}_{x \in K}$. Questo è un ricoprimento aperto di K: per compattezza, possiamo estrarre un sottoricoprimento finito...che ci darà un numero finito di palle di raggio ε che ricopre K.

Dimostreremo domani il viceversa.

14 Lezione del 6/11/2009 (2 ore)

Viceversa, supponiamo che K sia completo e totalmente limitato e sia $\{x_k\} \subset K$ una successione a valori in K: mostriamo che è possibile estrarre una sottosuccessione che converge ad un punto di K.

Per la totale limitatezza è possibile ricoprire K con un numero finito di palle aperte di raggio 1. Necessariamente, dentro una di queste, che chiameremo B_1, cadranno infiniti elementi della successione. Sia $\{x_k^{(1)}\}$ la sottosuccessione di $\{x_k\}$ costituita dagli elementi che cadono dentro B_1.

Ricopriamo poi K con un numero finito di palle di raggio $1/2$: dentro una di queste, detta B_2, cadranno infiniti elementi di $\{x_k^{(1)}\}$. Sia $\{x_k^{(2)}\}$ la sottosuccessione di $\{x_k^{(1)}\}$ costituita dagli elementi che cadono entro B_2. Procediamo poi allo stesso modo, ricoprendo con palle di raggio $1/3$, $1/4$...

Costruiamo così per induzione una successione di sottosuccessioni tali che $\{x_k^{(n)}\}$ è caratterizzata dal fatto di essere una sottosuccessione di $\{x_k^{(n-1)}\}$ e dal fatto che i suoi elementi sono contenuti in una palla di raggio $1/n$.

Prendiamo infine la *sottosuccessione diagonale* definita da $\tilde{x}_k = x_k^{(k)}$ (cioè il k-esimo elemento della successione diagonale è il k-esimo elemento della k-esima sottosuccessione.

La successione $\{\tilde{x}_k\}$ ha la proprietà di essere una sottosuccessione di $\{x_k^{(n)}\}$ per $k \geq n$: in particolare, essa è una sottosuccessione di $\{x_k\}$, ed è evidentemente di Cauchy (perchè dall'n-esimo termine in poi è tutta contenuta in una palla di raggio $1/n$, e questo vale per ogni n). Dunque $\tilde{x}_k \to \overline{x} \in K$ grazie alla completezza, come volevasi.

La dimostrazione appena vista funziona anche se è X ad essere completo, mentre K è solo totalmente limitato: in tal caso, possiamo solo dire che $\overline{x} \in X$. Q.E.D.

OSSERVAZIONE: Una conseguenza di questa caratterizzazione dei compatti è la seguente: un compatto in uno spazio normato ha la proprietà che, per ogni $\varepsilon > 0$, esiste un sottospazio di dimensione finita Y_ε da cui tutti gli elementi di K distano meno di ε. Infatti, per la totale limitatezza, possiamo trovare un numero finito di palle $B_\varepsilon(x_1)$,...,$B_\varepsilon(x_N)$ che ricoprono K. Basta allora definire $Y_\varepsilon = \text{span}\{x_1, \ldots, x_N\}$.

In conclusione, in dimensione infinita i compatti sono abbastanza "smilzi", e le palle chiuse si guardano bene dall'essere compatte. In particolare, non è vero che da una successione limitata in norma si può estrarre una sottosuccessione convergente!

Per rimediare alla non compattezza della palla chiusa in dimensione infinita, ci si è inventati il concetto di *convergenza debole*: se siamo in uno spazio di Banach "decente", da ogni successione limitata è possibile estrarre una sottosuccessione debolmente convergente. Vediamo la definizione:

DEFINIZIONE: Sia $(X, \|\cdot\|)$ uno spazio normato. Si dice che una successione $\{x_k\} \subset X$ *converge debolmente* a $\overline{x} \in X$, e si scrive $x_k \rightharpoonup \overline{x}$, se e soltanto se

$$T(x_k) \to T(\overline{x}) \quad \forall T \in X'.$$

ESEMPI/OSSERVAZIONI/ESERCIZI: Osserviamo per prima cosa che una successione convergente in norma converge anche debolmente (grazie alla continuità di ogni $T \in X'$): la *convergenza forte* implica la *convergenza debole*.

In dimensione finita, convergenza forte (cioè convergenza in norma) e convergenza debole coincidono: infatti una successione in $\mathbf{R}^n$ converge se e solo se convergono le sue componenti.

In generale, invece, in dimensione infinita la convergenza forte e la convergenza debole non sono la stessa cosa: vedremo fra un attimo che in uno spazio riflessivo esistono sempre successioni debolmente convergenti che *non* convergono in norma.

Per fare un po' di esempi espliciti, possiamo usare gli spazi ℓ^p e $L^p(\Omega)$, per i quali conosciamo alla perfezione il duale!

Se $1 \le p < +\infty$, consideriamo per ogni $n \in \mathbf{N}$ un elemento $x^n = \{x^n_k\}_k \in \ell^p$. Per definizione, abbiamo che $x^n \rightharpoonup \overline{x} = \{\overline{x}_k\}_k$ in ℓ^p se e solo se per ogni $T \in (\ell^p)'$ si ha $T(x^n) \to T(\overline{x})$, ossia (grazie alla caratterizzazione del duale) se e soltanto se

$$\lim_{n\to+\infty} \sum_{k=1}^{\infty} x^n_k y_k = \sum_{k=1}^{\infty} \overline{x}_k y_k \quad \forall \{y_k\} \in \ell^q.$$

Per esempio, la successione e^n dei "vettori della base canonica" in ℓ^p (cioè e^n è l'elemento di ℓ^p il cui n-esimo termine vale 1, mentre tutti gli altri valgono 0), allora $e^n \rightharpoonup 0$ in ℓ^p per $1 < p < +\infty$, mentre e^n non converge debolmente a nulla in ℓ^1.

La convergenza debole in $L^p(\Omega)$ si caratterizza in maniera analoga a quella negli spazi di successioni: data $\{u_k\} \subset L^p(\Omega)$ (sempre con $1 \le p < +\infty$), avremo che $u_k \rightharpoonup u$ in L^p se e solo se

$$\lim_{n\to+\infty} \int_\Omega u_k(x)v(x)\,dx = \int_\Omega u(x)v(x)\,dx \quad \forall v \in L^q(\Omega).$$

L'utilità della convergenza debole viene in gran parte dal seguente

TEOREMA (Banach-Alaoglu): Se X è uno spazio di Banach riflessivo, la palla unitaria chiusa è debolmente (sequenzialmente) compatta. In altre parole, da ogni successione limitata in norma è possibile estrarre una sottosuccessione debolmente convergente ad un elemento di X.

Non dimostreremo questo teorema, se non (a suo tempo) nel caso particolare in cui X è uno spazio di Hilbert separabile.

OSSERVAZIONE: Se $\{x_n\} \subset X$ (con X di Banach) è una successione tale che $x_n \rightharpoonup x$, allora $\sup\{\|x_n\| : n \in \mathbf{N}\} < +\infty$: ogni successione debolmente convergente è limitata in norma. Infatti, per la definizione stessa di convergenza debole, $\{T(x_n)\}$ è un insieme limitato in $\mathbf{R}$ per ogni $T \in X'$: abbiamo visto che questo implica che $\{x_n\}$ è limitata in norma (come conseguenza del Teorema di Banach-Steinhaus).

OSSERVAZIONE: In uno spazio di Banach riflessivo di dimensione infinita esistono sempre successioni debolmente convergenti che non convergono in norma. Infatti, abbiamo visto che la palla unitaria chiusa non è fortemente compatta: esiste una successione di elementi di norma 1 che non ha sottosuccessioni convergenti in norma. D'altra parte, il teorema di Banach-Alaoglu assicura che esiste una sottosuccessione debolmente convergente!

ESERCIZIO: Sia X uno spazio di Banach, $C \subset X$ un insieme convesso e fortemente chiuso (cioè chiuso nella topologia indotta dalla norma). Allora C è debolmente sequenzialmente chiuso: se $\overline{x} \in X$ è limite debole di una successione a valori in C, allora $\overline{x} \in C$.

Sia infatti $x_n \rightharpoonup \overline{x}$ con $\{x_n\} \subset C$. Supponiamo per assurdo che $\overline{x} \notin C$. Allora possiamo applicare la seconda forma geometrica del teorema di Hahn-Banach ai convessi $\{\overline{x}\}$ e C, il primo dei quali è compatto ed il secondo chiuso. Troviamo allora $T \in X'$ e $\varepsilon > 0$ tali che $T(x) + \varepsilon < T(\overline{x})$ per ogni $x \in C$, ed in particolare $T(x_n) + \varepsilon < T(\overline{x})$ per ogni n. Questo è assurdo perché $T(x_n) \to T(\overline{x})$ per definizione di convergenza debole!

ESERCIZIO: Se X è uno spazio di Banach e $F : X \to \mathbf{R}$ è una funzione convessa e continua, allora F è (sequenzialmente) debolmente semicontinua inferiormente: per ogni successione $x_n \rightharpoonup \overline{x}$ si ha $F(\overline{x}) \leq \liminf\limits_{n\to+\infty} F(x_n)$.

In particolare, dato che la norma é una funzione convessa, si ha

$$\|\overline{x}\| \leq \liminf_{n\to+\infty} \|x_n\| \quad \forall x_n \rightharpoonup \overline{x}.$$

Sia $\ell = \liminf_{n\to+\infty} F(x_n)$. Se $\ell = +\infty$, non c'é nulla da dimostrare. Sia dunque $\ell < +\infty$: scegliamo $s \in \mathbf{R}$, $s > \ell$ e consideriamo il sottolivello

$$C_s = \{x \in X : \ F(x) \leq s\}.$$

Questo insieme è un convesso chiuso, grazie rispettivamente alla convessità ed alla continuità forte di F: per l'esercizio precedente C_s è dunque debolmente sequenzialmente chiuso. Inoltre, estraendo eventualmente una sottosuccessione possiamo supporre $F(x_n) \to \ell$: tolti allora un numero finito di termini della sottosuccessione avremo $\{x_n\} \subset C_s$. Ma allora, grazie alla chiusura debole, $\overline{x} \in C_s$ da cui $F(\overline{x}) \leq s$. Dall'arbitrarietà di $s > \ell$ segue subito la tesi.

15 Lezione del 11/11/2009 (2 ore)

Come aperitivo, facciamo un ultimo esercizio sulle convergenze deboli.

ESERCIZIO: Se X è uno spazio di Banach riflessivo, C un convesso chiuso non vuoto e $x_0 \in X$, mostrare che esiste un elemento di C di distanza minima da x_0.

A meno di traslazioni, ci riduciamo al caso $x_0 = 0$: vogliamo dimostrare che esiste un punto di C di norma minima.

Sia infatti $\{y_n\} \subset C$ una successione tale che $\|y_n\| \to \inf\{\|y\| : \ y \in C\} = \delta$. Evidentemente la nostra successione è limitata in norma: il teorema di Banach-Alaoglu garantisce allora che esiste una sottosuccessione debolmente convergente ad un punto $\overline{y}$. La debole chiusura di C ci assicura che $\overline{y} \in C$, la debole semicontinuità inferiore della norma permette di concludere che $\overline{y}$ è il punto di distanza minima cercato: $\|\overline{y}\| \leq \liminf \|y_n\| = \delta$.

ESERCIZIO (non svolto in classe): Se lo spazio di Banach X non è riflessivo, il risultato dell'esercizio precedente può essere falso. Si consideri infatti lo spazio $C^0([0,1])$ con la norma uniforme, e l'insieme

$$C = \{u \in C^0([0,1]) : \int_0^{1/2} u(x)\, dx - \int_{1/2}^1 u(x)\, dx = 1\}.$$

Questo è un chiuso convesso, e l'estremo inferiore delle norme dei suoi elementi è 1. D'altra parte, non esiste alcun elemento di C che ha norma 1: non c'è un punto di C di distanza minima dall'origine!

La convessità di C è evidente, come pure la sua chiusura: basta osservare che $\Phi : u \mapsto \int_0^{1/2} u(x)\, dx - \int_{1/2}^1 u(x)\, dx$ è un funzionale lineare continuo sul nostro spazio, per cui C è un iperpiano (affine) chiuso.

Verifichiamo innanzitutto che $\mathrm{dist}(0, C) \geq 1$: maggiorando gli integrali, si vede subito che nessuna funzione di norma minore di 1 può appartenere a C (la nostra differenza di integrali si maggiora appunto con

$$1/2\,(\mathrm{essup}\{u(x) : x \in [0,1/2]\} - \mathrm{essinf}\{u(x) : x \in [1/2,1]\})\,,$$

che è minore o uguale a $\|u\|_\infty$).

Se poi $\|u\|_\infty = 1$, questa stessa maggiorazione ci dice che per appartenere a C la funzione u dovrebbe valere quasi ovunque 1 su $[0,1/2]$, -1 su $[1/2,1]$: evidentemente, non esiste alcuna funzione continua con questa caretteristica.

Infine, che la distanza sia proprio 1 si vede prendendo la successione di funzioni continue $u_n(x)$ che vale $1+1/n$ su $[0, 1/2 - 1/(n+1)]$, vale $-1-1/n$

su $[1/2+1/(n+1),1]$ ed è lineare su $[1/2-1/(n+1),1/2+1/(n+1)]$: queste funzioni appartengono a C, con $\|u_n\|_\infty = 1+1/n \to 1$.

OSSERVAZIONE (non vista in classe): Nello spazio ℓ^1 (che non è riflessivo!) ogni successione debolmente convergente è anche fortemente convergente. Si tratta di un esempio davvero patologico, e la dimostrazione non è affatto semplice (ed infatti *non l'abbiamo vista in aula*)!

Evidentemente, ci si può ridurre al caso di una successione $\{x^n\} \subset \ell^1$ con $x^n \rightharpoonup 0$. Per quanto visto prima, $\|x^n\|_{\ell^1} \leq C$ per ogni n (la successione è limitata in norma). Dobbiamo mostrare che in realtà $\|x^n\|_{\ell^1} \to 0$. Supponiamo per assurdo che questo non sia vero: a meno di estrarre una sottosuccessione, questo implica che $\|x^n\|_{\ell^1} \geq c > 0$ per ogni n. Mostriamo che questo ci permette di estrarre un'ulteriore sottosuccessione che *non* converge debolmente a 0, il che è assurdo.

Innanzitutto, se è $x^n = \{x^n_k\}_k$, allora deve essere

$$\lim_{n\to+\infty} x^n_{\overline{k}} = 0 \quad \forall \overline{k} \in \mathbf{N} :$$

la convergenza debole a 0 implica la convergenza a 0 di tutte le componenti (basta applicare la definizione di convergenza debole con $\{y_k\} = e^{\overline{k}}$). Dunque, per ogni fissato $N \in \mathbf{N}$ si ha anche

$$\lim_{n\to+\infty} \sum_{k=1}^{N} |y^n_k| = 0.$$

Usando questo fatto e la definizione di serie, vediamo che è possibile scegliere una successione strettamente crescente di numeri naturali $k_1 < k_2 < k_3 < \dots$ ed una sottosuccessine di x^n (che denoteremo ancora con x^n) in modo che

$$\sum_{k=k_n+1}^{k_{n+1}} |x^n_k| \geq \frac{3}{4}\|x^n\|_{\ell^1},\ n = 1,2,3,\dots$$

Definiamo ora una successione $\{y_k\} \in \ell^\infty$ nel modo seguente: se $k_n+1 \leq k \leq k_{n+1}$, allora $y_k = \mathrm{sgn}(x^n_k)$. Allora, per ogni fissato n, si ha

$$T_{\{y_k\}}(x^n) = \sum_{k=k_n+1}^{k_{n+1}} |x^n_k| + \sum_{altri\ \ k} x^n_k y_k \geq \frac{3}{4}\|x^n\|_{\ell^1} - \frac{1}{4}\|x^n\|_{\ell^1} \geq \frac{1}{2}c.$$

Questa successione evidentemente non tende a zero per $n \to +\infty$, il che contraddice la convergenza debole di x^n a 0.

Oltre al teorema di Banach-Alaouglu che garantisce la compattezza debole, esistono in commercio anche dei teoremi che caratterizzano i compatti in norma.

Molto interessante è per esempio la questione della compattezza in C^0: si tratta del più classico e semplice degli spazi funzionali...e purtroppo non è riflessivo! Abbiamo però il seguente risultato di compattezza per funzioni continue:

TEOREMA (Ascoli-Arzelà): Sia $u_n : A \to B$ una successione di funzioni continue, con A e B spazi metrici compatti. Se la successione u_n è equicontinua, cioè se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $x, y \in A$ e $d_A(x, y) < \delta$ implica $d_B(u_n(x), u_n(y)) < \varepsilon$ per ogni n, allora u_n ha una sottosuccessione che converge uniformemente ad una funzione continua $u : A \to B$.

DIM.: Innanzitutto, osserviamo che l'insieme $C^0(A, B)$ delle funzioni continue da A in B diventa uno spazio metrico completo con la distanza uniforme $d(u, v) = \sup\{d_b(u(x), v(x)) : \ x \in A\}$. Per dimostrare il teorema basta far vedere che $\mathcal{F} = \{u_n : \ n \in \mathbf{N}\}$ è un sottinsieme totalmente limitato di $C^0(A, B)$. Preso $\varepsilon > 0$, si applichi la totale limitatezza di B per scrivere $B = B_1 \cup B_2 \cup \ldots \cup B_N$, con B_j palle di raggio ε. Si usi poi l'equicontinuità per trovare δ tale che $d_A(x, y) < \delta$ implichi $d_B(u_n(x), u_n(y)) < \varepsilon$, e poi la totale limitatezza di A per scrivere $A = A_1 \cup \ldots \cup A_M$, con A_i palle di raggio δ e centro a_i.

Per ogni multiindice $(j_1, j_2, \ldots j_M) \in \{1, 2, \ldots, N\}^M$ (ce ne sono un numero finito) si consideri l'insieme di funzioni

$$\mathcal{W}_{(j_1, j_2, \ldots j_M)} = \{u \in \mathcal{F} : \ f(a_i) \in B_{j_i}, \ i = 1, \ldots, M\}.$$

Ciascun elemento della successione appartiene ovviamente ad uno di questi insiemi. Inoltre, se non è vuoto, ciascuno di questi insiemi ha diametro minore di 4ε, ed è quindi contenuto in una palla di raggio 4ε: se infatti $u, v \in \mathcal{W}_{(j_1, j_2, \ldots j_M)}$ e $x \in A$, scegliamo i tale che $x \in A_i$. Usando la disuguaglianza di equicontinuità si ha allora $d_B(u(x), v(x)) \leq d_B(u(x), u(a_i)) + d_B(u(a_i), v(a_i)) + d_B(v(a_i), v(x)) < 4\varepsilon$. Siamo dunque riusciti a coprire $\mathcal{F}$ con un numero finito di palle di raggio 4ε. Q.E.D.

OSSERVAZIONE: Per funzioni a valori reali il teorema di Ascoli-Arzelà si enuncia di solito dicendo ogni successione di funzioni in $C^0(A; \mathbf{R})$ (A spazio metrico compatto) che siano *equicontinue* ed *equilimitate*, ammette una sottosuccessione che converge uniformemente ad una funzione continua.

L'equilimitatezza garantisce infatti che tutte le funzioni della successione hanno valori nel comaptto $[-M, M]$, con M numero reale opportunamente grande.

Il teorema di Ascoli-Arzelà si può applicare ad un gran numero di problemi dell'analisi: per esempio, per dimostrare il teorema di Peano che predica l'esistenza di soluzioni per il problema di Cauchy nell'ambito delle Equazioni Differenziali Ordinarie.

Archiviato il teorema di Ascoli-Arzelà, vogliamo cominciare ad occuparci della teoria degli spazi di Hilbert. Prima di darne la definizione, abbiamo bisogno di introdurre il concetto di prodotto scalare e di norma indotta da un prodotto scalare.

DEFINIZIONE: Sia X uno spazio vettoriale reale. Un *prodotto scalare* su X è un'applicazione

$$\begin{array}{rcl} <\cdot,\cdot>: X\times X & \to & \mathbf{R} \\ (x,y) & \mapsto & <x,y> \end{array}$$

che sia bilineare (cioè lineare in ciascuno dei due argomenti x e y), simmetrica (cioè $<x,y>=<y,x>$ per ogni x,y) e definita positiva (cioè $<x,x>\geq 0$, con l'uguaglianza se e solo se $x=0$).

Un prodotto scalare su X può essere usato per definire una norma nel modo seguente:

$$\|x\| :=<x,x>^{1/2} .$$

Naturalmente, bisogna verificare che la norma indotta da un prodotto scalare è proprio una norma. Questa è una delle conseguenze della seguente proposizione, che raccoglie alcuni fatti elementari ma fondamentali sui prodotti scalari:

PROPOSIZIONE: Sia $<\cdot,\cdot>$ un prodotto scalare su X, $\|\cdot\|$ la norma da esso indotta. Allora valgono i fatti seguenti

(i) Per ogni $x,y\in X$ vale la disuguaglianza di Cauchy-Schwarz

$$<x,y>\leq \|x\|\ \|y\|;$$

(ii) La norma indotta dal prodotto scalare è una norma;

(iii) Vale l'identità del parallelogramma:

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2) \quad \forall x,y\in X;$$

(iv) Vale l'identità di polarizzazione:

$$< x, y >= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \quad \forall x, y \in X.$$

DIM.: Se $x, y \in X$ e $t \in \mathbf{R}$ si ha:

$$0 \leq \|ty + x\|^2 =< ty + x, ty + x >= t^2\|y\|^2 + 2t < x, y > +\|x\|^2.$$

Ora, se questo polinomio di secondo grado in t è sempre positivo, il suo discriminante deve essere minore o uguale a zero: questa è esattamente la disuguaglianza di Cauchy-Schwarz, e (i) è dimostrato.

Proviamo (ii): innanzitutto, si vede subito che la norma è omogenea e non degenere. Rimane allora da verificare la disuguaglianza triangolare: per ogni x, y si ha, usando la disuguaglianza di Cauchy-Schwarz:

$$\begin{aligned} &\|x + y\|^2 =< x + y, x + y >= \|x\|^2 + 2 < x, y > +\|y\|^2 \leq \\ &\|x\|^2 + 2\|x\| \, \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Le identità (iii) e (iv) si dimostrano banalmente espandendo i prodotti scalari che definiscono le norme coinvolte. La (iii) si chiama *identità del parallelogramma* perché se si interpretano i vettori x e y come lati di un parallelogramma, allora $x+y$ e $x-y$ ne rappresentano le diagonali: l'identità è dunque una generalizzazione di un noto risultato di geometria euclidea. Q.E.D.

L'identità del parallelogramma caratterizza le norme che "provengono" da un prodotto scalare:

PROPOSIZIONE: Sia $(X, \| \cdot \|)$ *uno spazio normato. Allora l'applicazione*

$$a(x, y) :== \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2), \quad x, y \in X$$

è un prodotto scalare che induce la norma data se e solo se la norma verifica l'identità del parallelogramma.

DIM.: Abbiamo già visto che se la norma proviene da un prodotto scalare, allora vale l'identità del parallelogramma e il prodotto scalare si scrive come sopra grazie all'identità di polarizzazione. Viceversa, supponiamo che valga l'identità del parallelogramma e definiamo $a(x, y)$ come nell'enunciato. È allora evidente che questa funzione è simmetrica, che $a(x, x) = \|x\|^2 \geq 0$ con l'uguaglianza se e solo se $x = 0$. Inoltre, $a(x, 0) = a(0, y) = 0$ e $a(-x, y) = -a(x, y)$. È anche immediato verificare che la funzione è continua.

Prendiamo poi $x_1, x_2, y \in X$: allora, usando l'identità del parallelogramma

$$(*) \qquad a(x_1, y) + a(x_2, y) = \frac{1}{4}(\|x_1 + y\|^2 - \|x_1 - y\|^2 + \|x_2 + y\|^2 - \|x_2 - y\|^2) = \frac{1}{8}(\|x_1 + x_2 + 2y\|^2 + \|x_1 - x_2\|^2 - \|x_1 + x_2 - 2y\|^2 - \|x_1 - x_2\|^2) = \frac{1}{2}a(x_1 + x_2, 2y).$$

In particolare, se prendiamo $x_1 = x$, $x_2 = 0$, l'ultima identità diventa

$$(**)\ a(x, y) = \frac{1}{2}a(x, 2y) \quad \forall x, y.$$

Applicando (**) entro (*) otteniamo allora

$$a(x_1, y) + a(x_2, y) = a(x_1 + x_2, y) \quad \forall x_1, x_2, y.$$

Usando ancora l'additività si ottiene facilmente che $a(mx, y) = m\, a(x, y)$ per ogni $m \in \mathbf{Z}$. Allora, grazie a (**) e alla simmetria:

$$a(\frac{m}{2^n}x, y) = \frac{m}{2^n}a(x, y) \quad \forall x, y \in X,\ \forall m \in \mathbf{Z}, n \in \mathbf{N}.$$

L'insieme dei razionali binari, ossia dei numeri della forma $m/2^n$, è denso in $\mathbf{R}$: grazie alla continuità di a possiamo allora concludere che $a(tx, y) = ta(x, y)$ per ogni $x, y \in X$ e per ogni $t \in \mathbf{R}$: questo conclude la dimostrazione che $a(x, y)$ è un prodotto scalare. Q.E.D.

DEFINIZIONE (Spazio di Hilbert): Uno spazio vettoriale reale X, dotato di un prodotto scalare $< \cdot, \cdot >$ si dice *di Hilbert* se è di Banach rispetto alla norma indotta dal prodotto scalare.

ESEMPI: Tipici prototipi di spazi di Hilbert sono ℓ^2 col prodotto scalare $< \{x_k\}, \{y_k\} >:= \sum_{k=1}^{\infty} x_k y_k$ e $L^2(\Omega)$ col prodotto scalare $< u, v >:= \int\limits_{\Omega} u(x)v(x)\, dx$.

16 Lezione del 12/11/2009 (2 ore)

Il risultato che segue è una versione rafforzata di un risultato che abbiamo visto la volta scorsa in uno spazio di Banach riflessivo (in uno degli esercizi). La dimostrazione che ne daremo questa volta è però indipendente dal Teorema di Banach-Alaouglu.

TEOREMA (di proiezione su un convesso chiuso): Sia X uno spazio di Hilbert, C un sottinsieme chiuso, convesso e non vuoto di X, $x_0 \in X$. Allora esiste un unico punto $\overline{y} \in C$ tale che $\|x_0 - \overline{y}\| = \mathrm{dist}(x_0, C)$.

DIM.: Tramite una traslazione, possiamo sempre ridurci al caso in cui $x_0 = 0$: si tratta dunque di dimostrare che in C esiste un unico elemento di norma minima. Sia dunque $\delta = \inf\{\|y\| : \ y \in C\}$, e sia $\{y_n\} \subset C$ una successione tale che $\|y_n\| \to \delta$ (una tale successione esiste per definizione di estremo inferiore!).

Dimostriamo che $\{y_n\}$ è di Cauchy in X: a questo scopo, si prenda l'identità del parallelogramma con $x/2$, $y/2$ al posto di x, y... Con facili conti si arriva all'identità

$$\|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) - 4\left\|\frac{x+y}{2}\right\|^2,$$

valida per ogni $x, y \in X$. Si noti inoltre che, se x e y appartengono a C, allora $\frac{x+y}{2} \in C$: dall'identità appena provata e dalla definizione di δ abbiamo allora

$$(***) \qquad \|y_n - y_m\|^2 = 2(\|y_n\|^2 + \|y_m\|^2) - 4\left\|\frac{y_n + y_m}{2}\right\|^2 \leq$$
$$2(\|y_n\|^2 + \|y_m\|^2) - 4\delta^2.$$

L'ultima quantità tende a zero per $m, n \to +\infty$, per cui la successione $\{y_n\}$ è di Cauchy e tende a un qualche punto $\overline{y} \in X$. Essendo C chiuso, $\overline{y} \in C$. Inoltre, $\|\overline{y}\| = \delta$ per la continuità della norma: $\overline{y}$ è il punto di norma minima cercato.

Rimane da verificare che questo punto di norma minima è unica: se fosse anche $\|\tilde{y}\| = \delta$ per qualche $\tilde{y} \in C$, applichiamo la disuguaguglianza (***) con $y_n = \overline{y}$, $y_m = \tilde{y}$ e troviamo

$$\|\overline{y} - \tilde{y}\| \leq 0,$$

da cui $\overline{y} = \tilde{y}$. Q.E.D.

COROLLARIO: Sia Y un sottospazio chiuso dello spazio di Hilbert X, $x_0 \in X$. Allora esiste un unico punto $\overline{y} \in Y$ di minima distanza da x_0. Esso è caratterizzato dalla relazione di ortogonalità

$$< x_0 - \overline{y}, y >= 0 \quad \forall y \in Y.$$

DIM.: Il Teorema ci garantisce che esiste un unico punto $\overline{y}$ di minima distanza. Per definizione, $\overline{y}$ realizza la distanza minima se e soltanto se, per ogni $y \in Y$ con $\|y\| = 1$ e per ogni $t \in \mathbf{R}$, $t \neq 0$ si ha

$$0 < \|x_0 - (\overline{y} + ty)\|^2 - \|x_0 - \overline{y}\|^2 = \\ t^2\|y\|^2 - 2t < x_0 - \overline{y}, y > .$$

Si vede subito che l'ultima espressione è positiva per ogni $t \neq 0$ se e solo se $< x_0 - \overline{y}, y >= 0$. Q.E.D.

L'ultimo corollario è particolarmente importante: da esso è facile dedurre che uno spazio di Hilbert si decompone come somma diretta di un suo sottospazio chiuso e del suo spazio ortogonale, con proiezioni continue.

PROPOSIZIONE: Sia $Y \subset X$ un sottospazio chiuso non vuoto dello spazio di Hilbert X, $p : X \to X$ l'applicazione che a ogni $x \in X$ associa il punto del sottospazio Y più vicino a x. Allora p è lineare e continua, la sua restrizione a Y è l'identità. Inoltre $x - p(x)$ è ortogonale a Y, per cui possiamo scrivere $X = Y \oplus Y^\perp$, con proiezioni continue. Infine, $\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2$ per ogni $x \in X$.

DIM.: Il Corollario appena visto asserisce che $p(x)$ si caratterizza l'unico punto di Y tale che $< x - p(x), y >= 0$ per ogni $y \in Y$, cioè come l'unico punto di Y tale che $x - p(x) \in M^\perp$: per questa ragione, esso viene chiamato *proiezione ortogonale* di x su Y.

È evidente che p coincide con l'identità su Y. Mostriamo che è lineare: siano $x_1, x_2 \in X$, $t \in \mathbf{R}$. Allora si ha $0 =< x_1 - p(x_1), y >=< x_2 - p(x_2), y >$ per ogni $y \in Y$, e grazie alla bilinearità del prodotto scalare

$$< x_1 - tx_2 - (p(x_1) + tp(x_2), y >= 0 \quad \forall y \in Y,$$

da cui $p(x_1 + tx_2) = p(x_1) + tp(x_2)$.

Se poi $x \in X$, essendo $p(x) \in Y$ si ha $< x - p(x), p(x) >= 0$, da cui

$$\|x\|^2 =< p(x) + (x - p(x)), p(x) + (x - p(x) >= \|p(x)\|^2 + \|x - p(x)\|^2.$$

Questo mostra anche che p è continua, in quanto l'identità implica

$$\|p(x)\| \leq \|x\|,$$

cioè la norma di $p : X \to X$ come applicazione lineare è ≤ 1 (in realtà è uguale a 1 perché coincide con l'identità su Y). Q.E.D.

Osserviamo per prima cosa che la proiezione su un sottospazio di dimensione finita di uno spazio di Hilbert può essere scritta esplicitamente:

OSSERVAZIONE: Se Y è un sottospazio di dimensione finita di X, e $\{e_1, \ldots, e_n\}$ è una sua base ortonormale, allora possiamo scrivere esplicitamente

$$p(x) = \sum_{i=1}^{n} < x, e_i > e_i.$$

Inoltre, $\|p(x)\|^2 = \sum_{i=1}^{n} (< x, e_i >)^2$. Basta infatti verificare che $x - p(x)$ è ortogonale a tutti gli elementi di Y: questo è certamente vero se questa proprietà vale per gli elementi della base. Ora,

$$< x - p(x), e_j > = < x, e_j > - \sum_{i=1}^{n} < x, e_i >< e_i, e_j > = 0,$$

come volevasi dimostrare. L'espressione della norma di $p(x)$ è un'immediata conseguenza dell'ortonormalità degli e_i.

Si noti che questo risultato vale indipendentemente dalla completezza di X: nel teorema di proiezione su un convesso, essa serviva solo per dimostrare l'esistenza del punto di distanza minima...mentre in questo caso lo esibiamo esplicitamente!

Il seguente teorema caratterizza il duale di uno spazio di Hilbert: per ogni funzionale lineare continuo $T \in X'$ esiste un unico $y \in X$ tale che $T(x) = < y, x >$ per ogni $x \in X$. In particolare, il duale di X è isomorfo e isometrico a X:

TEOREMA (di rappresentazione di Riesz): Sia X uno spazio di Hilbert. Definiamo l'applicazione

$$\begin{aligned} \Phi : X \quad &\to \quad X' \\ y \quad &\mapsto T_y \end{aligned}$$

dove per definizione $T_y(x) := < y, x >$ per ogni $x \in X$. Allora Φ è un isomorfismo isometrico tra X e X'.

DIM.: Dalla disuguaglianza di Cauchy-Schwarz abbiamo

$$T_y(x) = < y, x > \leq \|y\| \, \|x\|,$$

per cui il funzionale lineare T_y è continuo di norma $\leq \|y\|$. D'altra parte, $T_y(\frac{y}{\|y\|}) = \|y\|$, da cui $\|T_y\|_{X'} = \|y\|$. Ne segue che l'applicazione $\Phi : X \to X'$ (che è evidentemente lineare) è ben definita ed è un'isometria.

Rimane da verificare che Φ è suriettiva, ossia che per ogni $T \in X'$ esiste $y \in X$ tale che $T = T_y$.

Sia $Y = \ker(T)$. Se $Y = X$, basta prendere $y = 0$: possiamo dunque supporre che Y sia un sottospazio chiuso proprio di X. Sia allora $x_0 \in X \setminus Y$, $\overline{y}$ la proiezione ortogonale di x_0 su Y. Fissato $x \in X$, si vede subito che

$$x - \frac{T(x)}{T(x_0 - \overline{y})}(x_0 - \overline{y}) \in Y.$$

Questo vettore deve dunque essere ortogonale a $x_0 - \overline{y}$:

$$< x_0 - \overline{y}, x - \frac{T(x)}{T(x_0 - \overline{y})}(x_0 - \overline{y}) >= 0,$$

da cui con facili semplificazioni

$$T(x) =< x, T(x_0 - \overline{y})\frac{x_0 - \overline{y}}{\|x_0 - \overline{y}\|^2} >,$$

e il nostro asserto è provato con $y = T(x_0 - \overline{y})\frac{x_0 - \overline{y}}{\|x_0 - \overline{y}\|^2}$. Q.E.D.

Per proseguire, ci serve definire la somma di una famiglia qualunque (non necessariamente numerabile) di numeri non negativi:

DEFINIZIONE: Sia $\{t_\alpha\}_{\alpha \in I}$ una famiglia (non necessariamente numerabile) di numeri reali *non negativi*. Allora, per definizione,

$$\sum_{\alpha \in I} t_\alpha = \sup\{\sum_{\alpha \in J} t_\alpha : \ J \subset I, \ J \ insieme \ finito\}.$$

La famiglia $\{t_\alpha\}$ si dice *sommabile* se la sua somma è finita.

Osserviamo che se l'insieme degli indici I è numerabile e $\{\alpha_n\}_{n \in \mathbf{N}}$ è una sua enumerazione, allora

$$\sum_{\alpha \in I} t_\alpha = \sum_{n=1}^{\infty} t_{\alpha_n}$$

(e in particolare la somma della serie non dipende dalla particolare scelta dell'enumerazione). È infatti ovvio che la somma a sinistra è maggiore o uguale della serie a destra (la prima è definita come estremo superiore di un insieme più ampio di somme di un numero finito di termini!). Viceversa, data una qualunque somma *di un numero finito* di elementi della famiglia $\{t_\alpha\}$, possiamo sempre scegliere $N \in \mathbf{N}$ abbastanza grande, in modo che tutti i termini della somma compaiano nella somma parziale $\sum_{n=1}^{N} t_{\alpha_n}$...per cui la serie è maggiore o uguale della somma a sinistra.

OSSERVAZIONE: Se $\{t_\alpha\}_{\alpha\in I}$ è sommabile, allora l'insieme $I' = \{\alpha \in I : t_\alpha > 0\}$ è al più numerabile.

Infatti, per ogni fissato $n = 1, 2, 3, \ldots$, l'insieme $I_n = \{\alpha \in I : t_\alpha > 1/n\}$ è necessariamente finito.

OSSERVAZIONE: Se $\{c_\alpha\}_{\alpha\in I}$ è una famiglia di numeri reali tale che $\sum_{\alpha\in I} |c_\alpha| < +\infty$, è ben definita (e finita) la somma $\sum_{\alpha\in I} c_\alpha$.

Un modo per definirla è prendere l'*integrale* della famiglia $\{c_\alpha\}$ rispetto alla counting measure sull'insieme degli indici I. Infatti, è facile vedere (e l'abbiamo visto a esercitazione nel caso numerabile...) che l'integrale di una famiglia *non negativa* coincide con la definizione di somma data sopra: questo rende "ragionevole" definire la somma tramite l'integrale anche nel caso generale!

DEFINIZIONE: Sia I un insieme non necessariamente numerabile di indici. Denotiamo con $\ell^2(I)$ l'insieme delle famiglie di numeri reali $\{c_\alpha\}_{\alpha\in I}$ tali che sia finita la somma

$$\sum_{\alpha\in I} c_\alpha^2.$$

Questo insieme diventa uno spazio vettoriale normato con la norma

$$\|\{c_\alpha\}_{\alpha\in I}\|_{\ell^2(I)} = (\sum_{\alpha\in I} c_\alpha^2)^{1/2},$$

e si può verificare che si tratta di uno spazio di Banach (infatti, non è altro che lo spazio $L^(\mu)$, con μ la counting measure su I). Anzi, si tratta di uno spazio di Hilbert con il prodotto scalare

$$< \{a_\alpha\}, \{b_\alpha\} >:= \sum_\alpha a_\alpha b_\alpha,$$

in cui la somma a destra è assolutamente convergente grazie alla disuguaglianza di Hölder in ℓ^2, ed è ben definita in quanto non dipende dalla particolare enumerazione dell'insieme su cui i coefficienti sono diversi da zero.

17 Lezione del 13/11/2009 (2 ore)

Torniamo agli spazi di Hilbert: diamo la fondamentale definizione di serie di Fourier astratta di un elemento $x \in X$ rispetto ad una fissata famiglia ortonormale di elementi di X.

PROPOSIZIONE (Disuguaglianza di Bessel): Sia X uno spazio di Hilbert, $\{e_\alpha\}_{\alpha\in I}$ una famiglia ortonormale di elementi di X (cioè $\|e_\alpha\| = 1$ per ogni

$\alpha \in I$ e $< e_\alpha, e_\beta >= 0$ per ogni $\alpha, \beta \in I$, $\alpha \neq \beta$). Se $x \in X$, definiamo i suoi coefficienti di Fourier rispetto a $\{e_\alpha\}$ come i numeri reali

$$c_\alpha =< x, e_\alpha >, \quad \alpha \in I.$$

Allora vale la disuguaglianza di Bessel

$$\sum_{\alpha \in I} |c_\alpha|^2 \leq \|x\|.$$

In particolare, solo una quantità numerabile dei coefficienti di Fourier di x può essere diversa da zero.

DIM.: Sia $J \subset I$ un qualunque insieme finito di indici. Allora il vettore $\sum_{\alpha \in J} < x, e_\alpha > e_\alpha$ è la proiezione ortogonale di x sullo spazio generato da $\{e_\alpha\}_{\alpha \in J}$ e si ha

$$\|x\|^2 = \| \sum_{\alpha \in J} < x, e_\alpha > e_\alpha\|^2 + \|x - \sum_{\alpha \in J} < x, e_\alpha > e_\alpha\|^2$$

da cui (usando l'ortonormalità degli e_α):

$$\|x\|^2 \geq \| \sum_{\alpha \in J} < x, e_\alpha > e_\alpha\|^2 = \sum_{\alpha \in J} \| < x, e_\alpha > e_\alpha\|^2 = \sum_{\alpha \in J} c_\alpha^2.$$

Passando al sup su tutti i sottinsiemi finiti $J \subset I$ si ha la disuguaglianza di Bessel. Q.E.D.

La disuguaglianza di Bessel ci assicura che i coefficienti di Fourier $\{c_\alpha\}$ di $x \in X$, rispetto ad una famiglia ortonormale $\{e_\alpha\}_{\alpha \in I}$, appartengono a $\ell^2(I)$. Viceversa, ogni elemento di $\ell^2(I)$ coincide con i coefficienti di Fourier di un qualche elemento di X:

TEOREMA: Sia X uno spazio di Hilbert, $\{e_\alpha\}_{\alpha \in I}$ una famiglia ortonormale. Per ogni famiglia di numeri reali $\{t_\alpha\}_{\alpha \in I} \in \ell^2(I)$ esiste un elemento $x \in X$ tale che

$$< x, e_\alpha >= c_\alpha \quad \forall \alpha \in I.$$

In altre parole, l'applicazione lineare

$$\begin{array}{rcl} \Psi : X & \to & \ell^2(I) \\ x & \mapsto & \{< x, e_\alpha >\}_{\alpha \in I} \end{array}$$

è suriettiva.

DIM.: Sappiamo che i c_α sono non nulli al più per una famiglia numerabile $I' \subset I$ di indici. Scegliamo un'enumerazione di I':

$$I' = \{\alpha_k : \ k = 1, 2, 3 \dots\}.$$

Poniamo poi

$$x_n = \sum_{k=1}^{n} c_{\alpha_k} e_{\alpha_k}.$$

Grazie alla ortonormalità degli e_α si ha

$$\|x_n - x_{n+h}\|^2 = \sum_{k=n}^{n+h} c_{\alpha_k}^2,$$

da cui si deduce che la successione $\{x_n\}$ è di Cauchy (perché la serie $\sum\limits_{k=1}^{\infty} c_{\alpha_k}^2$ è convergente). Dunque $x_n \to x \in X$. Si ha poi, grazie alla continuità del prodotto scalare,

$$< x, e_\alpha >= \lim_{n \to +\infty} < x_n, e_\alpha >= c_\alpha$$

(per verificare l'ultima uguaglianza, distinguere il caso $\alpha \in I'$ e $\alpha \in I \setminus I'$. Q.E.D.

OSSERVAZIONE: Nella dimostrazione del teorema precedente, il punto x è stato trovato come somma della serie $\sum_{k=1}^{\infty} c_{\alpha_k} e_{\alpha_k}$. Verrebbe voglia di scrivere

$$x = \sum_{\alpha \in I} c_\alpha e_\alpha,$$

Questo è vero se l'insieme ortonormale è *massimale*, e sarà una conseguenza del prossimo teorema.

Più in generale, dato un sistema ortonormale $\{e_\alpha\}_{\alpha \in I}$ in X e $\{c_\alpha\} \in \ell^2(I)$, la *serie di Fourier*

$$\sum_{\alpha \in I} c_\alpha e_\alpha$$

è ben definita e converge ad un elemento di X: precisamente, converge alla proiezione $p(x)$ di x sulla *chiusura* dello spazio generato dai vettori $\{e_\alpha\}$.

Il prossimo teorema assicura che l'applicazione Ψ definita nel teorema precedente è un isomorfismo isometrico a patto che il sistema ortonormale $\{e_\alpha\}$ sia *massimale.* In tal caso, dato $x \in X$ potremo sempre scrivere

$$x = \sum_{\alpha \in I} c_\alpha e_\alpha,$$

ove $c_\alpha =< x, e_\alpha >$ sono i coefficienti di Fourier di x.

TEOREMA (Serie di Fourier astratte): Sia X uno spazio di Hilbert, $\{e_\alpha\}_{\alpha\in I}$ una famiglia ortonormale in X. Allora i fatti seguenti sono equivalenti:

(i) La famiglia $\{e_\alpha\}_{\alpha\in I}$ è massimale, nel senso che se vi aggiungiamo un qualunque altro elemento di X, essa non è più ortonormale;

(ii) Lo spazio generato dai vettori $\{e_\alpha\}_{\alpha\in I}$ è denso in X;

(iii) Vale l'identità di Parseval: per ogni vettore $x \in X$, se denotiamo con $c_\alpha =< x, e_\alpha >$ i suoi coefficienti di Fourier, allora

$$\|x\|^2 = \sum_{\alpha\in I} |c_\alpha|^2.$$

Dunque la mappa Ψ definita nel teorema precedente è un isomorfismo isometrico tra X e $\ell^2(I)$. In particolare essa è iniettiva, per cui la serie di Fourier nella dimostrazione del teorema converge a x e non dipende dalla particolare enumerazione scelta per l'insieme dei coefficienti di Fourier non nulli.

DIM.: Mostriamo che $(i) \Rightarrow (ii)$: supponiamo per assurdo che $Y = \text{span}\{e_\alpha\}$ non sia denso, e sia $x_0 \in X \setminus \overline{Y}$. Allora, se $p(x_0)$ è la proiezione ortogonale di x_0 su $\overline{Y}$, $x_0 - p(x_0)$ è un vettore non nullo ortogonaloe a tutti gli e_α, contro l'ipotesi di massimalità.

Mostriamo poi che $(ii) \Rightarrow (iii)$: scelto $\varepsilon > 0$, per ipotesi per ogni $x \in X$ possiamo trovare una combinazione lineare finita $\lambda_1 e_{\alpha_1} + \lambda_2 e_{\alpha_2} + \ldots + \lambda_N e_{\alpha_N}$ tale che

$$\|x - \lambda_1 e_{\alpha_1} + \lambda_2 e_{\alpha_2} + \ldots + \lambda_N e_{\alpha_N}\|^2 < \varepsilon.$$

Questo implica

$$\|x - c_{\alpha_1} e_{\alpha_1} - c_{\alpha_2} e_{\alpha_2} - \ldots - c_{\alpha_N} e_{\alpha_N}\|^2 < \varepsilon$$

(per la proprietà di minimalità della proiezione ortogonale $p(x)$ su $Y = \text{span}\{e_{\alpha_1}, \ldots, e_{\alpha_N}\}$), da cui $\varepsilon > \|x - p(x)\|^2 = \|x\|^2 - \|p(x)\|^2 = \|x\|^2 - \sum_{i=1}^N c_{\alpha_i}^2$ e dunque

$$\|x\|^2 \leq \sum_{\alpha\in I} c_\alpha^2 + \varepsilon.$$

Poiché $\sum_{\alpha\in I} c_\alpha^2 \leq \|x\|^2$ (disuguaglianza di Bessel), l'identità di Parseval risulta dimostrata grazie all'arbitrarietà di ε.

Mostriamo infine che $(iii) \Rightarrow (i)$: sia $x_0 \in X$ ortogonale a tutti gli e_α. L'identità di Parseval permette allora di dire che $\|x\| = 0$, quindi $x = 0$ e la famiglia ortonormale $\{e_\alpha\}$ non risulta ulteriormente ampliabile. Q.E.D.

OSSERVAZIONE: Un insieme ortonoromale massimale in uno spazio di Hilbert si chiama *base di Hilbert.* È facile vedere che una base di Hilbert esiste sempre (lemma di Zorn): dunque, un qualunque spazio di Hilbert X è isomorfo e isometrico ad un $\ell^2(I)$ per un'opportuna scelta dell'insieme di indici I.

ESERCIZIO: Sia X uno spazio di Hilbert, $\{e_\alpha\}_{\alpha \in I}$ una famiglia ortonormale (non necessariamente massimale). Provare che per ogni $x \in X$ è ben definita la somma della serie

$$\sum_{\alpha \in I} < x, e_\alpha > e_\alpha.$$

(SUGG.: Si consideri il sottospazio $Y = \overline{\text{span}\{e_\alpha\}_{\alpha \in I}}$. Si mostri che la serie converge precisamente alla proiezione ortogonale di x su Y...)

Ricordiamo che uno spazio topologico si dice *separabile* se contiene un sottinsieme denso numerabile.

PROPOSIZIONE: Uno spazio di Hilbert X è separabile se e soltanto se possiede una base di Hilbert numerabile.

DIM.: Se X possiede una base di Hilbert numerabile, per il teorema precedente lo spazio da essa generato è denso in X. Consideriamo allora l'insieme delle *combinazioni lineari finite a coefficienti razionali* degli elementi della base di Hilbert. Questo è un insieme numerabile denso in X.

Viceversa, sia $\{x_n\} \subset X$ un insieme numerabile denso. Applichiamo il procedimento di ortonormalizzazione di Gram-Schmidt a questo insieme (avendo cura, ad ogni passo, di eliminare il vettore in esame se è linearmente dipendente dai precedenti). Otteniamo così una successione di vettori ortonormali $\{e_k\}$ che genera un sottospazio di X che contiene tutti gli x_n, ossia un sottospazio denso. Essa è dunque una base di Hilbert grazie al teorema precedente. Q.E.D.

Ci si chiederà che cosa centra la teoria astratta vista fino ad ora con le *serie di Fourier* nel tradizionale senso "trigonometrico"! Ecco la risposta:

OSSERVAZIONE: Consideriamo lo spazio

$$L^2(2\pi) = \{u : \mathbf{R} \to \mathbf{R} : \ u \ misurabile \ 2\pi - periodica, \ \int_{-\pi}^{\pi} u^2(x)\, dx < +\infty\},$$

con la consueta convenzione di considerare equivalenti funzioni che differiscono su un insieme di misura nulla. Questo è uno spazio di Hilbert col prodotto scalare

$$< u, v >= \int_{-\pi}^{\pi} u(x)v(x)\, dx.$$

Consideriamo la seguente famiglia di elementi di $L^2(2\pi)$:

$$\mathcal{F} = \{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \sin nx, \frac{1}{\sqrt{\pi}} \cos nx : \ n = 1, 2, \ldots\}.$$

Si vede facilmente che questa famiglia è ortonormale (basta usare le note relazioni di ortogonalità (!) per le funzioni trigonometriche), e che la serie di Fourier di una funzione $u \in L^2(2\pi)$ rispetto a questo sistema ortonormale è proprio la consueta serie di Fourier.

Inoltre, la famiglia $\mathcal{F}$ è massimale: lo vedremo la prossima volta. Come conseguenza, avremo ottenuto che la classica serie di Fourier di una funzione L^2 converge in L^2 alla funzione stessa!

18 Lezione del 17/11/2009 (2 ore)

Il fatto che il sistema ortonormale in $L^2(2\pi)$ dato da

$$\mathcal{F} = \left\{ \frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}} \cos nx,\ \frac{1}{\sqrt{\pi}} \sin nx,\ n = 1, 2, \ldots \right\}$$

sia *massimale*, e quindi una base di Hilbert, è una conseguenza del *Teorema di Stone-Weierstrass*, che dice che ogni funzione continua e 2π-periodica si può approssimare uniformemente con polinomi trigonometrici (che, ricordiamo, sono per definizione le combinazioni lineari finite di elementi di $\mathcal{F}$).

Poiché,come vedremo presto, ogni elemento di $L^2(2\pi)$ si approssima in norma L^2 con funzioni continue, ne segue che si approssima anche con polinomi trigonometrici, cioè con combinazioni lineari degli elementi di $\mathcal{F}$: lo spazio generato dal nostro sistema ortonormale è dunque denso in $L^2(2\pi)$, e $\mathcal{F}$ è massimale.

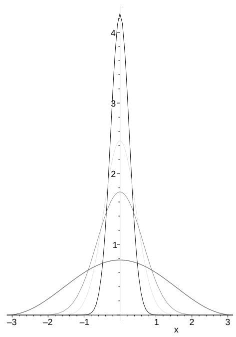
Il teorema sulle serie di Fourier astratte, applicato a questo caso concreto, ci assicura che la serie di Fourier di una funzione L^2 converge in norma L^2 alla funzione stessa!

Vediamo la dimostrazione del teorema di Stone-Weierstrass per i polinomi trigonometrici. Poiché questo fatto ci servirà, ricordiamo che l'insieme dei polinomi trigonometrici è un algebra: il prodotto di due polinomi trigonometrici è ancora tale. Questo si può vedere facilmente, per esempio, esprimendo seni e coseni tramite esponenziali complessi... Una conseguenza di questo è che un polinomio in $\sin x$, $\cos x$ è un polinomio trigonometrico!

TEOREMA (Stone-Weierstrass): Sia $u : \mathbf{R} \to \mathbf{R}$ *una funzione continua e* 2π*-periodica. Allora per ogni* $\varepsilon > 0$ *esiste un polinomio trigonometrico* $v(x)$ *tale che* $\|u - v\|_\infty < \varepsilon$.

DIM.: Per ogni numero naturale n, consideriamo le seguenti funzioni (che, per quanto detto sopra, sono polinomi trigonometrici): $\phi_n(t) = c_n \left(\frac{1+\cos t}{2}\right)^n$, ove le costanti c_n sono scelte in modo che $\int_{-\pi}^{\pi} \phi_n(t)\, dt = 1$.

Disegnando il grafico di queste funzioni, si vede che si tratta di funzioni periodiche non negative che si "concentrano" sempre più attorno ai punti del tipo $2k\pi$:



Definiamo dei polinomi trigonometrici che approssimano u, sfruttando le ϕ_n nel modo seguente:

$$u_n(x) = \int_{-\pi}^{\pi} u(x+t)\phi_n(t)\, dt.$$

Queste funzioni sono sostanzialmente ottenute facendo delle "medie pesate" di u: mostreremo che $u_n \to u$ uniformemente.

Prima di farlo, mostriamo però che le u_n sono dei polinomi trigonometrici, perché dalla definizione questo fatto non è proprio visibilissimo... Per convincersene, è sufficiente cambiare variabile nell'integrale, ponendo $s = x + t$. Sfruttando la periodicità delle funzioni coinvolte vediamo subito che le u_n possono essere scritte nel seguente modo equivalente:

$$u_n(x) = \int_{-\pi}^{\pi} u(s)\phi_n(s-x)\, ds.$$

Poichè le ϕ_n sono polinomi trigonometrici, usando le formule di addizione di seno e coseno e la linearità dell'integrale vediamo subito che le u_n sono effettivamente polinomi trigonometrici!

Per mostrare invece che $u_n \to u$ uniformemente, useremo la prima scrittura delle u_n. Grazie al fatto che le ϕ_n sono non negative ed hanno integrale 1, otteniamo:

$$(*)|u_n(x){-}u(x)| = |\int_{-\pi}^{\pi}(u(x{+}t){-}u(x))\phi_n(t)\ dt| \leq \int_{-\pi}^{\pi}|u_n(x{+}t){-}u_n(x)|\phi_n(t)\ dt.$$

Per mostrare che l'ultimo integrale tende unifomemente a 0 per $n \to +\infty$, ci serve una semplice stima delle costanti c_n che compaiono nella definizione di ϕ_n. Si ha

$$\frac{1}{c_n} = \int_{-\pi}^{\pi}\left(\frac{1+\cos t}{2}\right)^n dt \geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}}\left(\frac{1+\cos t}{2}\right)^n \geq \frac{2}{\sqrt{n}}\left(\frac{1+\cos(\frac{1}{\sqrt{n}})}{2}\right)^n,$$

dove l'ultima disuguaglianza è stata ottenuta minorando la funzione integranda col suo valore agli estremi. Poichè la quantità tra parentesi tonde nell'ultima espressione tende a $e^{-1/4}$, possiamo scrivere $c_n \leq k\sqrt{n}$ per n grande abbastanza, dove k è un'opportuna costante positiva.

A questo punto, ricordiamoci che u è continua e periodica: per questo motivo, essa è maggiorata in modulo da una costante $M > 0$, ed è anche uniformemente continua: per ogni $\varepsilon > 0$, possiamo trovare $\delta > 0$ tale che $|x-y| < \delta$ implica $|u(x) - u(y)| < \varepsilon$.

A questo punto, spezziamo l'integrale di destra in (*) sugli insiemi $[-\delta, \delta]$ e $[-\pi, -\delta] \cup [\delta, \pi]$. Per la nostra scelta di δ si ha

$$\int_{-\delta}^{\delta}|u(x+t) - u(x)|\phi_n(t)\ dt \leq \varepsilon\int_{-\delta}^{\delta}\phi_n(t)\ dt < \varepsilon.$$

D'altra parte

$$\int_{\delta}^{\pi}|u(x+t) - u(x)|\phi_n(t)\ dt \leq 2Mc_n\int_{\delta}^{\pi}\left(\frac{1+\cos t}{2}\right)^n dt \leq$$
$$4Mk\pi\sqrt{n}\left(\frac{1+\cos(1/\sqrt{\delta})}{2}\right)^n,$$

e l'ultima quantità tende a 0 per $n \to +\infty$ (in modo uniforme in x, visto che essa è indipendente da x). L'integrale su $[-\pi, -\delta]$ si stima esattamente allo stesso modo: prendendo n abbastanza grande, otteniamo dunque che $|u_n(x) - u(x)| < 2\varepsilon$ per ogni x. Q.E.D.

Per concludere la nostra discussione degli spazi di Hilbert, studiamo la convergenza debole in questo contesto. Grazie al teorema di rappresentazione di Riesz, la definizione di convergenza debole nel caso di uno spazio di Hilbert X si legge come segue: se $\{x_n\} \subset X$, allora

$$(x_n \rightharpoonup x) \Longleftrightarrow_{Def} (< y, x_n > \to < y, x > \quad \forall y \in X).$$

La seguente proposizione raccoglie un paio di fatti interessanti sulla convergenza debole in uno spazio di Hilbert:

PROPOSIZIONE: Sia X uno spazio di Hilbert. Allora

(i) Se $x_n \rightharpoonup x$, allora $\{x_n\}$ è limitata e $\|x\| \leq \liminf_{n\to+\infty} \|x_n\|$.

(ii) Se $\{x_n\}$ è una successione tale che per ogni $y \in X$ esiste finito il limite $T(y) := \lim_{n\to+\infty} < x_n, y >$, allora esiste un unico $x \in X$ tale che $x_n \rightharpoonup x$.

(iii) Se $x_n \rightharpoonup x$ e $y_n \to y$ (convergenza in norma), allora

$$< x_n, y_n > \to < x, y > .$$

(iv) Si ha $x_n \to x$ se e soltanto se $x_n \rightharpoonup x$ e $\|x_n\| \to \|x\|$.

DIM.: La (i) la abbiamo già vista in un qualunque spazio di Banach. Per la (ii), si applichi il Teorema di Banach-Steinhaus ai funzionali $T_n(y) :=< x_n, y >$, e si applichi poi il Teorema di rappresentazione di Riesz al funzionale limite $T(y)$.

Per la (iii), scriviamo

$$< x_n, y_n > - < x, y >= (< x_n, y_n > - < x_n, y >) + (< x_n, y > - < x, y >).$$

La seconda parentesi tonda tende a zero per definizione di convergenza debole. Per stimare la prima parentesi tonda, si osservi che $\{x_n\}$ è limitata in norma (grazie a (i)) e si applichi la disuguaglianza di Cauchy-Schwarza a $< x_n, y_n - y >$. Dunque, anche la prima parentesi tonda tende a 0 e (iii) risulta provata.

Per la (iv), una freccia è ovvia. Per l'altra basta scrivere

$$\|x_n - x\|^2 = \|x_n\|^2 - 2 < x_n, x > +\|x\|^2$$

e applicare la convergenza debole e la convergenza delle norme. Q.E.D.

19 Lezione del 18/11/2009 (2 ore)

Vediamo ora la dimostrazione del Teorema di Banach-Alaoglu in uno spazio di Hilbert separabile:

TEOREMA: Sia X uno spazio di Hilbert separabile, $\{x_n\}_{n\in\mathbf{N}}$ una successione limitata in X. Allora esiste $x \in X$ e una sottosuccessione x_{n_k} di x_n tali che $x_{n_k} \rightharpoonup x$.

DIM.: Sia $\{e_j\}_{j\in\mathbf{N}}$ un insieme ortonormale massimale in X (base di Hilbert). Allora per ogni n possiamo scrivere

$$x_n = \sum_{j=1}^{\infty} c_{nj} e_j,$$

dove $c_{nj} =< x_n, e_j >$ sono i coefficienti di Fourier di x_n. Inoltre, l'ipotesi di limitatezza della successione si può esprimere dicendo che esiste $C > 0$ tale che

$$\sum_{j=1}^{\infty} c_{nj}^2 \leq C \quad \forall n \in \mathbf{N}.$$

In particolare, l'ultima disuguaglianza dice che per ogni fissato j si ha $|c_{nj}|\sqrt{C}$ per ogni n.

Considero in particolare i coefficienti c_{n1}: essi formano una successione limitata in $\mathbf{R}$, in particolare posso estrarre una sottosuccessione $x_n^{(1)}$ di x_n (i cui coefficienti di Fourier saranno denotati con $c_{nj}^{(1)}$ tale che $c_{n1}^{(1)} \to \overline{c}_1 \in \mathbf{R}$.

Anche la successione $c_{n2}^{(1)}$ è limitata: posso estrarre una sottosuccessione $x_n^{(2)}$ di $x_n^{(1)}$ tale che $c_{n2}^{(2)} \to \overline{c}_2 \in \mathbf{R}$.

Procedendo in questo modo, mi costruisco per ricorrenza una successione di sottosuccessioni $x_n^{(k)}$ tali che $x_n^{(k)}$ è una sottosuccessione di $x_n^{(k-1)}$, e tali che i coefficienti di Fourier $c_{nj}^{(k)}$ soddisfino

$$\lim_{n\to+\infty} c_{nj}^{(k)} = \overline{c}_j, \; j = 1, 2, \ldots, k.$$

Prendiamo la sottosuccessione diagonale definita da $\tilde{x}_n = x_n^{(n)}$, i cui coefficienti di Fourier saranno indicati con $\tilde{c}_{nj}$. Essa è una sottosuccessione di $\{x_n\}$ con la proprietà che

$$\lim_{n\to+\infty} \tilde{c}_{nj} = \lim_{n\to+\infty} < \tilde{x}_n, e_j >= \overline{c}_j \quad \forall j \in \mathbf{N}.$$

Grazie alla linearità del limite e del prodotto scalare segue che, posto $Y = \mathrm{span}\{e_j : \; j \in \mathbf{N}\}$ (sottospazio denso), per ogni $y \in Y$ esiste finito il limite

$$\lim_{n\to+\infty} < \tilde{x}_n, y >= T(y),$$

e questo limite puntuale $T : Y \to \mathbf{R}$ è evidentemente lineare. Inoltre, T è limitato (è limite puntuale di funzionali equilimitati in norma), per cui può essere esteso con Hahn-Banach ad un funzionale limitato T definito su tutto X. Sia $\overline{x} \in X$ tale che $T(y) =< \overline{x}, y >$ per ogni $y \in X$ (esiste per il teorema di rappresentazione di Riesz...e i suoi coefficienti di Fourier sono evidentemente $\overline{c}_j$). La nostra costruzione ci dice che $< \tilde{x}_n, y > \to < \overline{x}, y >$ per ogni $y \in Y$: dimostriamo che la stessa cosa vale per ogni $y \in X$.

Sia dunque $y \in X$, $\varepsilon > 0$. Poiché Y è denso in X, esiste $\tilde{y} \in Y$ tale che $\|\tilde{y} - y\| < \varepsilon$. Allora:

$$
\begin{aligned}
&< \tilde{x}_n, y > - < \overline{x}, y >= \\
&< \tilde{x}_n, y - \tilde{y} > + (< \tilde{x}_n, \tilde{y} > - < \overline{x}, \tilde{y} >) - < \overline{x}, y - \tilde{y} > .
\end{aligned}
$$

Il modulo della quantità tra parentesi tonda è certamente minore di ε a patto di prendere n abbastanza grande. Gli altri due termini si stimano in modulo con $C\varepsilon$: prendiamo per esempio il primo e otteniamo grazie a Cauchy-Schwarz

$$
| < \tilde{x}_n, y - \tilde{y} > | \leq \|\tilde{x}_n\| \, \|y - \tilde{y}\| < C\varepsilon.
$$

Concludiamo dunque che si ha

$$
| < \tilde{x}_n, y > - < \overline{x}, y > | < (2C + 1)\varepsilon
$$

a patto di prendere n abbastanza grande. Q.E.D.

OSSERVAZIONE: Il teorema precedente si può facilmente estendere ad uno spazio di Hilbert non separabile X: se infatti $\{x_n\}$ è una successione limitata in X, poniamo $Z = \overline{\text{span}\{x_n : n \in \mathbf{N}\}}$. Questo è evidentemente uno spazio di Hilbert separabile (le combinazioni lineari a coefficienti razionali degli x_n sono un denso): per il risultato precedente troviamo $\overline{x} \in Z$ e una sottosuccessione x_{n_k} tali che $< x_{n_k}, y > \to < \overline{x}, y >$ per ogni $y \in Z$, per $k \to +\infty$. La stessa relazione di limite vale evidentemente per ogni $y \in Z^\perp$ (perché tutti i prodotti scalari si annullano!). Ma allora essa è valida per ogni $y \in X$ perché $X = Z \oplus Z^\perp$: $x_{n_k} \rightharpoonup \overline{x}$ in X.

OSSERVAZIONE: La dimostrazione fatta per gli spazi di Hilbert separabili si ripete, con poche modifiche, nel caso di uno *spazio di Banach riflessivo* il cui duale è *separabile* (ipotesi che è poi equivalente a chiedere che lo spazio sia *riflessivo e separabile*). In questo caso, la base di Hilbert $\{e_j\}$ viene sostituita da una famiglia numerabile di elementi del duale che ne genera un sottospazio denso: si noti infatti che nella dimostrazione non gioca alcun ruolo essenziale l'ortonormalità degli $\{e_j\}$.

Vediamo in dettaglio come si procede per gli spazi L^p (*NON visto in classe!*):

TEOREMA (Compattezza debole in L^p): Sia $1 < p < +\infty$ (si noti che abbiamo esplicitamente escluso $p = 1$ e $p = \infty$). Se $\{u_n\} \subset L^p([a,b])$ è una successione limitata, cioè tale che esiste $C > 0$ per cui $\|u_n\|_{L^p} \leq C$ per ogni $n \in \mathbf{N}$, allora esiste una sottosuccessione u_{n_k} e una funzione $u \in L^p$ tali che $u_{n_k} \rightharpoonup u$ debolmente in L^p per $k \to +\infty$.

DIM.: Sia q l'esponente coniugato di p. Siccome $q \neq +\infty$, lo spazio $L^q([a,b])$ è separabile: possiamo dunque trovare una successione $\{v_j\}_j \subset L^q([a,b])$ di funzioni L^q, che formino una base di uno sottospazio Y denso in L^q (basta partire da un sottinsieme denso numerabile, e sceglierví un sottinsieme massimale di funzioni linearmente indipendenti...). Dividendo queste funzioni per la loro norma, non è restrittivo supporre $\|v_j\|_{L^q} = 1$ per ogni j.

Grazie alla disuguaglianza di Hölder, si vede subito che per ogni fissato j la successione reale

$$n \mapsto \int_a^b u_n(x)v_j(x)\,dx$$

è equilimitata dalla costante C, e possiede dunque una sottosuccessione convergente.

Usando un procedimento diagonale come nella dimostrazione del caso Hilbertiano, vediamo allora che è possibile estrarre una sottosuccessione $\{u_{n_k}\}$ di $\{u_n\}$ tale che, per ogni fissato j, la successione reale $k \mapsto \int_a^b u_{n_k}(x)v_j(x)\,dx$ tende ad un certo numero reale c_j.

Se poi $v \in Y$, possiamo scrivere in modo unico $v(x) = \sum_{j=1}^J \lambda_j v_j(x)$. Definiamo allora un funzionale lineare $T : Y \to \mathbf{R}$ tramite

$$T(v) = \sum_{j=1}^J \lambda_j c_j.$$

Si ha per costruzione

$$T(v) = \lim_{k\to+\infty} \int_a^b u_{n_k}(x)v(x)\,dx$$

e da Hölder e dall'equilimitatezza delle u_n segue subito che T è un funzionale lineare continuo su Y.

Per densità, questo si estende in modo unico ad un funzionale lineare continuo $T : L^q \to \mathbf{R}$...e grazie al teorema di Riesz (caratterizzazione del duale di L^q) abbiamo che esiste $u \in L^p$ tale che $T(v) = \int_a^b u(x)v(x)\,dx$ per ogni $v \in L^q$.

Abbiamo così trovato $u \in L^p$ tale che

$$\int_a^b u_{n_k} v\, dx \to \int_a^b uv\, dx \quad \forall v \in Y.$$

Grazie alla densità di Y, questa relazione di limite si estende a tutte le $v \in L^q$. Q.E.D.

Tra gli spazi funzionali più importanti che abbiamo incontrato in questo corso ci sono gli spazi $L^p(\Omega)$.Vogliamo ora studiarne meglio alcune proprietà particolarmente importanti per le applicazioni: in particolare, il fatto che ogni funzione in $L^p(\Omega)$ (con p finito) si possa approssimare bene con funzioni regolari.

Cominciamo col seguente importante teorema, che mostra che le funzioni misurabili hanno molto a che fare con le funzioni continue:

TEOREMA (Lusin): Sia $u : \Omega \to \mathbf{R}$ una funzione misurabile, con Ω insieme limitato misurabile secondo Lebesgue. Allora, per ogni $\varepsilon > 0$, possiamo trovare un compatto $K \subset \Omega$ con $|\Omega \setminus K| < \varepsilon$ e tale che la restrizione di u a K è continua.

DIM.: Per ogni fissato $j = 1, 2, 3, \ldots$, scriviamo $\mathbf{R} = \bigcup_{i=1}^{n} I_{ij}$, con I_{ij} intervalli disgiunti di lunghezza minore di $1/j$. Fissiamo anche dei punti $y_{ij} \in I_{ij}$.

Poniamo poi $A_{ij} = u^{-1}(I_{ij})$: questi sono insiemi misurabili due a due disgiunti la cui unione è tutto Ω. La regolarità della misura di Lebesgue ci assicura poi che possiamo trovare dei compatti $K_{ij} \subset A_{ij}$ tali che $|A_{ij} \setminus K_{ij}| < \frac{\varepsilon}{2^{i+j}}$.

Evidentemente si ha $|\Omega \setminus \bigcup_{i=1}^{\infty} K_{ij}| < \frac{\varepsilon}{2^j}$, e grazie alla continuità della misura sulle successioni decrescenti possiamo scegliere $N_j \in \mathbf{N}$ tale che

$$|\Omega \setminus \bigcup_{i=1}^{N_j} K_{ij}| < \frac{\varepsilon}{2^j}.$$

Definiamo $K_j = \bigcup_{i=1}^{N_j} K_{ij}$: questo è evidentemente un compatto. Se poi definiamo $u_j : K_j \to \mathbf{R}$ ponendo $u_j(x) = y_{ij}$ per $x \in K_{ij}$, otteniamo una funzione continua (perché è costante su ogni K_{ij}, e c'è un numero finito di questi insiemi che sono a distanza positiva l'uno dall'altro) tale che evidentemente $|u_j(x) - u(x)| < 1/j$ per ogni $x \in K_j$.

Se infine definiamo $K = \bigcap_{j=1}^{\infty} K_j$, otteniamo un compatto che verifica $|\Omega \setminus K| < \varepsilon$ sul quale $u_j \to u$ uniformemente. Ne segue che la restrizione di u a K è continua in quanto limite uniforme di funzioni continue. Q.E.D.

OSSERVAZIONE: Si noti che l'enunciato del teorema di Lusin non è in contraddizione col fatto che esistano funzioni misurabili ovunque discontinue: non stiamo affermando che i punti di K siano di continuità per u, ma per la sua restrizione a K!

Il prossimo teorema è un ben noto risultato di estensione:

TEOREMA (Tietze): Sia $K \subset \mathbf{R}^n$ un compatto. Se $u : K \to \mathbf{R}$ è continua, esiste una funzione continua $\tilde{u} : \mathbf{R}^n \to \mathbf{R}$ che estende u (cioè tale che $u(x) = \tilde{u}(x)$ per ogni $x \in K$) e tale che $\|\tilde{u}\|_\infty = \|u\|_\infty$. Se poi $K \subset \Omega$, con Ω aperto di $\mathbf{R}^n$, ossiamo anche chiedere che $u \in C^0_C(\Omega)$.

DIM.: Posto $M = \|u\|_\infty$, definiamo due compatti $K_1 = u^{(-1)}([-M, -M/3])$, $K_2 = u^{(-1)}([M/3, M])$: supponiamo per un momento che siano entrambi non vuoti, e sia $\delta > 0$ la loro distanza. Allora la funzione

$$\tilde{u}_1(x) = \min\{M/3, -M/3 + \frac{2M}{3\delta}\mathrm{dist}(x, K_1)\}$$

è continua, definita dappertutto e compresa tra $-M/3$ e $M/3$. Inoltre, essa vale $-M/3$ su K_1, $M/3$ su K_2. Ne segue che $|\tilde{u}_1(x) - u(x)| \le \frac{2}{3}M$ per ogni $x \in K$. Se poi K_1 fosse vuoto, otterremmo lo stesso risultato prendendo $\tilde{u}_1(x) = M/3$ (funzione costante). Analoga cosa si può fare se K_2 è vuoto. Vedremo domani come concludere la dimostrazione...

20 Lezione del 19/11/2009 (2 ore)

Concludiamo la dimostrazione del teorema di Tietze.

Ripetiamo la stessa costruzione per la funzione $u_2 = u - \tilde{u}_1$: troviamo una funzione continua $\tilde{u}_2$ ovunque definita, con $\|\tilde{u}_2\|_\infty \le \frac{2}{9}M$ e tale che $\|u - \tilde{u}_1 - \tilde{u}_2\|_\infty < \frac{4}{9}M$. Procedendo allo stesso modo, troviamo una successione $\tilde{u}_k$ di funzioni continue tali che $\|\tilde{u}_k\|_\infty \le \frac{2^{k-1}}{3^k}M$ e tali che

$$(*)\ \|u - \tilde{u}_1 - \tilde{u}_2 - \ldots - \tilde{u}_k\|_\infty < \frac{2^k}{3^k}\ \text{in } K.$$

La serie di funzioni continue

$$\sum_{k=1}^\infty \tilde{u}_k(x)$$

converge uniformemente in $\mathbf{R}^n$ ad una funzione $\tilde{u}$ (perché converge la serie delle norme[13]). Usando la maggiorazione sulle norme degli addendi, si vede

[13] Questo fatto vi è probabilmente già noto per lo spazio di Banach delle funzioni continue e limitate su $\mathbf{R}^n$...ma vale in generale: far vedere per esercizio che se X è di Banach e se $\{x_n\} \subset X$ è tale che $\sum_{n=1}^\infty \|x_n\| < +\infty$, allora la serie $\sum_{n=1}^\infty x_n$ è convergente in X... Al solito, basta far vedere che la successione delle somme parziali è di Cauchy!

subito che la norma di $\tilde{u}$ è minore o uguale a M. Inoltre, per la (*), $\tilde{u}$ coincide con u su tutti i punti di K.

Sia poi $\Omega \supset K$ un aperto. Prendiamo due aperti limitati Ω', Ω'' tali che $K \subset \Omega' \subset\subset \Omega'' \subset\subset \Omega$, e aggiungiamo il compatto $\overline{\Omega}'' \setminus \Omega'$ a K, prescrivendo che $u = 0$ su quell'insieme. Applichiamo poi il teorema come dimostrato sopra, ed infine modifichiamo l'estensione $\tilde{u}$ ponendola uguale a 0 su $\Omega \setminus \Omega''$. Q.E.D.

Ecco finalmente il risultato di densità delle funzioni continue:

TEOREMA (Densità delle funzioni continue in L^p): Sia $1 \leq p < +\infty$, Ω un aperto di $\mathbf{R}^n$. Allora le funzioni continue a supporto compatto sono dense in $L^p(\Omega)$.

DIM. Devo far vedere che data $u \in L^p(\Omega)$ e $\varepsilon > 0$ è possibile trovare $v \in C^0_C(\Omega)$ tale che $\|u - v\|_{L^p} < \varepsilon$.

Supponiamo dapprima che Ω sia limitato, e che $\|u\|_\infty = M < +\infty$: queste restrizioni saranno rimosse in seguito.

Per il teorema di Lusin trovo un compatto $K \subset \Omega$ tale che $|\Omega \backslash K| < (\frac{\varepsilon}{2M})^p$ e la restrizione di u a K è continua. Applico il teorema di estensione di Tietze e trovo $v \in C^0_C(\Omega)$ tale che $\|v\|_\infty = M$ e $v \equiv u$ su K. Allora

$$\|u - v\|_{L^p(\Omega)} = \|u - v\|_{L^p(\Omega\backslash K)} \leq \|u\|_{L^p(\Omega\backslash K)} + \|v\|_{L^p(\Omega\backslash K)} < \varepsilon,$$

come volevasi.

Se Ω ha misura finita ma u non è limitata, osserviamo che la funzione troncata $u_M(x) = \max\{-M, \min\{M, u(x)\}\}$ diventa arbitrariamente vicina a u in norma L^p, a patto di prendere M grande abbastanza: basta osservare che la misura degli insiemi $E_M = \{x \in \Omega : |u(x)| > M$ tende a 0 quando $M \to +\infty$, e usare l'assoluta continuità dell'integrale.

Infine, se Ω non è limitato si considerino le funzioni

$$u_R(x) = \begin{cases} u(x) & se\ |x| < R, \\ 0 & se\ |x| \geq R. \end{cases}$$

Usando il teorema della convergenza dominata si vede che $u_R \to u$ in L^p per $R \to +\infty$. Le funzioni u_R hanno supporto in $\Omega \cap B_R(0)$, che è un aperto limitato, per cui i discorsi precedenti sono applicabili. Q.E.D.

OSSERVAZIONE: Il teorema di densità è falso per $p = +\infty$. Infatti, le funzioni continue formano un sottospazio chiuso di $L^\infty(\Omega)$.

Vediamo alcune conseguenze della densità delle funzioni continue in L^p:

OSSERVAZIONE: Come importante conseguenza del teorema precedente, gli spazi $L^p(\Omega)$ sono separabili per $1 \leq p < +\infty$: è infatti abbastanza facile costruire un insieme numerabile di funzioni che sia denso nelle funzioni continue. Basta prendere ad esempio le *funzioni a scala* con "scalini" delimitati da intervalli di estremi razionali e con "altezza" razionale: è chiaro (uniforme continutà) che ogni funzione continua a supporto compatto si approssima uniformemente con funzioni a scala di questo tipo!

Si mostri invece (esercizio!) che $L^\infty(\Omega)$ non è separabile.

ESERCIZIO: Un'altra importante conseguenza della densità delle funzioni continue è la *continuità delle traslazioni* in $L^p(\mathbf{R}^n)$: sia $u \in L^p(\mathbf{R}^n)$, $y \in \mathbf{R}^n$. Poniamo $u_y(x) := u(x-y)$ (funzione traslata del vettore y). Mostrare che allora

$$\lim_{y\to 0} \|u - u_y\|_{L^p} = 0.$$

(SUGG:Se $u \in C^0_C(\mathbf{R}^n)$, l'asserto è una semplice conseguenza dell'uniforme continuità. Per una funzione generica, si approssimi la funzione data con una funzione continua.)

Il teorema di densità può essere considerevolmente "migliorato" fino a dimostrare che le funzioni C^∞_C sono un denso in $L^p(\Omega)$. Il teorema si può dimostrare applicando un procedimento noto come regolarizzazione per convoluzione: è precisamente il metodo che abbiamo usato per dimostrare la completezza del sistema trigonometrico in $L^2(2\pi)$.

Cominciamo con un risultato di regolarità:

LEMMA (Regolarità del prodotto di convoluzione): Sia $u \in L^1_{loc}(\mathbf{R}^n)$, $\phi \in C^1_C(\mathbf{R}^n)$. Allora la funzione

$$v(x) = \int_{\mathbf{R}^n} u(z)\phi(x-z)\, dz$$

è di classe C^1 e si ha

$$\frac{\partial v}{\partial x_i}(x) = \int_{\mathbf{R}^n} u(z)\frac{\partial \phi}{\partial x_i}(x-z)\, dz :$$

Iterando il risultato otteniamo che se $\phi \in C^\infty_C$, allora v è C^∞.

Vedremo domani la dimostrazione.

21 Lezione del 20/11/2009 (2 ore)

Dimostriamo il lemma sulla regolarità del prodotto di convoluzione.

Siccome la funzione integranda dipende in modo C^1 da x, questo risultato è una versione del classico teorema di derivazione sotto il segno di integrale. Lo si dimostra abbastanza facilmente con il teorema della convergenza dominata.

Cominciamo col dimostrare che v è continua: sia $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$ con $|y| \leq 1$. Sia poi K un compatto che contiene il supporto delle funzioni $\phi(x + y - \cdot)$ per ogni y come sopra, L la costante di Lipschitz di ϕ.

Allora:

$$|v(x+y) - v(x)| \leq \int\limits_K |u(z)||\phi(x+y-z) - \phi(x-z)|\, dz.$$

Grazie alla continuità della ϕ la funzione integranda tende puntalmente a 0 per $v \to 0$. È poi dominata da $L|u|\mathbf{1}_K$: grazie al teorema della convergenza dominata abbiamo dunque la continuità di v

Mostriamo poi che v è derivabile parzialmente e che le derivate sono come nella tesi: la continuità delle derivate segue poi ripetendo il ragionamento appena fatto con $\frac{\partial \phi}{\partial x_i}$ al posto di ϕ. Per ogni fissato $i = 1, \dots n$ e per ogni $|h| < 1$ abbiamo:

$$\frac{v(x+he_i) - v(x)}{h} = \int\limits_K u(z) \frac{\phi(x+he_i-z) - \phi(x-z)}{h}\, dz.$$

Per $h \to 0$ l'integranda tende puntualmente a $u(z)\frac{\partial \phi}{\partial x_i}(x-z)$ e la convergenza è dominata da $Lu(z)\mathbf{1}_K$, da cui la tesi. Q.E.D.

Una funzione L^p si può approssimare con funzioni regolari ottenute facendone il prodotto di convoluzione con un'opportuna successione di funzioni C^∞_C dette *mollifiers*.

TEOREMA (Regolarizzazione per convoluzione): Sia $u \in L^p(\mathbf{R}^n)$ *con* $1 \leq p < +\infty$*. Allora esiste una successione* $\{u_k\} \subset C^\infty(\mathbf{R}^n)$ *tale che* $u_k \to u$ *in* $L^p(\mathbf{R}^n)$*. La successione può anche essere costruita in modo che* $\{u_k\} \subset C^\infty_C(\mathbf{R}^n)$.

DIM.: Sia $\phi : \mathbf{R}^n \to \mathbf{R}$ una funzione C^∞ tale che $\phi(x) \geq 0$, $\phi(x) = \phi(-x)$ per ogni x e tale che valga

$$\operatorname{spt} \phi \subset B_1(0), \qquad \int\limits_{\mathbf{R}^n} \phi(x)\, dx = 1.$$

Una tale funzione è chiamata *mollifier* o *bump function*: per esempio, possiamo prendere $\phi(x) = c \exp(-\frac{1}{1-|x|^2})$ per $|x| \leq 1$, $\phi(x) = 0$ per $|x| > 1$, dove

c è una costante positiva da scegliere in modo che l'integrale della funzione sia 1.

Definiamo poi $\phi_k(x) = k^n\phi(kx)$: queste funzioni godono di tutte le proprietà di ϕ, ma sono sempre più concentrate attorno all'origine, perché si ha $\operatorname{spt} \phi_k \subset B_{1/k}(0)$.

Poniamo
$$u_k(x) = \int_{\mathbf{R}^n} u(x-y)\phi_k(y)\, dy.$$
Con un cambiamento di variabili si vede immediatamente che possiamo scrivere equivalentemente $u_k(x) = \int_{\mathbf{R}^n} u(z)\phi_k(x-z)\, dz$. Applicando il Lemma, si vede subito che $u_k \in C^\infty(\mathbf{R}^n)$. Queste funzioni u_k si chiamano *regolarizzate per convoluzione* di u.

Mostriamo che le funzioni u_k convergono a u in L^p. Si ha
$$\|u_k - u\|^p_{L^p(\mathbf{R}^n)} = \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} (u(x-y) - u(x))\phi_k(y)\, dy \right|^p\, dx.$$
Nell'integrale interno, scriviamo $\phi_k(y) = \phi_k(y)^{1/p}\phi_k(y)^{1-1/p}$ ed usiamo la disugualianza di Hölder: ricordando che ϕ_k ha integrale 1 otteniamo che esso si maggiora con
$$\left(\int_{\mathbf{R}^n} |u(x-y) - u(x)|^p \phi_k(y)\, dy \right)^{1/p}.$$
Sostituendo nell'espressione sopra otteniamo dunque:
$$\|u_k - u\|^p_{L^p(\mathbf{R}^n)} \le \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} |u(x-y) - u(x)|^p \phi_k(y)\, dy\, dx =$$
$$\int_{B_{1/k}(0)} \phi_k(y) \int_{\mathbf{R}^n} |u(x-y) - u(x)|^p\, dx\, dy,$$
dove si è usato il teorema di Fubini e la proprietà del supporto di ϕ_k.

Sia $\varepsilon > 0$ fissato. Siccome nel nostro integrale doppio $|y| \le 1/k$, la continuità delle traslazioni in L^p mi garantisce che per k abbastanza grande l'integranda interna è minore di ε: il tutto si stima dunque con ε, e abbiamo provato che $u_k \to u$,come volevasi dimostrare.

Se poi vogliamo che le funzioni approssimanti siano a supporto compatto, osserviamo dapprima che per ogni $\varepsilon > 0$ esiste $N \in \mathbf{N}$ tale che $\|u_N - u\|_{L^p} < \varepsilon$, ove $u_N = u \cdot \mathbf{1}_{B_N(0)}$ (si veda la dimostrazione del teorema di densità delle funzioni continue). La funzione u_N è identicamente nulla fuori dalla palla di raggio N, da cui si deduce subito che le sue regolarizzate per convoluzione $(u_N)_k$ sono a supporto compatto (precismanete, il loro supporto è contenuto nella palla di centro 0 e raggio $N + 1/k$). Per k abbastanza grande, avremo $\|(u_N)_k - u_N\|_{L^p} < \varepsilon$, da cui $\|u - (u_N)_k\| < 2\varepsilon$: l'approssimazione si può ottenere con funzioni C^∞_C. Q.E.D.

Come annunciato, l'ultima parte del corso sarà dedicata a complementi di teoria della misura.

Vediamo alcune condizioni che ci consentono di dire che una data misura gode di proprietà di regolarità analoghe a quella di Lebesgue: in particolare, siamo interessati alla possibilità di approssimare la misura con aperti e/o con compatti.

Nel seguito, supporremo che X sia uno spazio metrico *localmente compatto e separabile.*

DEFINIZIONE: Se X è come sopra, la *σ-algebra di Borel* $\mathcal{B}$ è la più piccola σ-algebra che contiene tutti gli aperti di X. Una misura (misura esterna) μ si dice *di Borel* se i boreliani (cioè gli elementi di $\mathcal{B}$) sono μ-misurabili.

Una misura esterna (risp. misura) μ si dice poi *Borel regolare* se ogni insieme (risp. insieme misurabile) A è contenuto in un boreliano B tale che $\mu(A) = \mu(B)$.

Infine, μ si dice *di Radon* se è Borel-regolare e $\mu(K) < +\infty$ per ogni compatto K.

Una misura di Radon è regolare nello stesso senso della misura di Lebesgue: la misura di ogni insieme A è l'estremo superiore delle misure dei compatti contenuti in A, e l'estremo inferiore delle misure degli aperti contenenti A:

TEOREMA (Approssimazione della misura con aperti, chiusi, compatti): Sia μ una misura esterna di Borel regolare su X. Supponiamo inoltre che esista una successione di aperti $\{V_j\}$ tale che $X = \bigcup V_j$ e $\mu(V_j) < +\infty$. Allora, per ogni $A \subset X$ vale

$$\begin{aligned}(*)\ \mu(A) &= \inf\{\mu(U) :\ U \textit{ aperto},\ U \supset A\},\\ (**)\ \mu(A) &= \sup\{\mu(C) :\ C \textit{ chiuso},\ C \subset A\}.\end{aligned}$$

Se la misura μ è soltanto di Borel, le stesse relazioni sono valide per A boreliano.

Infine, se μ è di Radon, allora l'esistenza della successione di aperti V_j come sopra è automatica. Inoltre,

$$(***)\ \mu(A) = \sup\{\mu(K) :\ K \textit{ compatto},\ K \subset A\}.$$

DIM. - NON VISTA IN CLASSE: Cominciamo col dimostrare (*) per A boreliano.

Supponiamo per cominciare che sia $\mu(X) < +\infty$: quest'ipotesi supplementare sarà rimossa in seguito.

Poniamo $\mathcal{A} = \{A \textit{ boreliani} : (*) \textit{ è vera}\}$. Mostriamo che $\mathcal{A}$ è chiuso per unione ed intersezione numerabile. Infatti, se $A = \bigcup\limits_{n=1}^{\infty} A_n$ con $A_n \in \mathcal{A}$, allora per ogni $\varepsilon > 0$ possiamo trovare degli aperti U_n tali che $A_n \subset U_n$ e $\mu(U_n \setminus A_n) < \varepsilon/2^n$. Allora $U = \bigcup_{n=1} U_n$ è un aperto che contiene A e

$$\mu(U \setminus A) \leq \mu\left(\bigcup_{n=1}^{\infty}(U_n \setminus A_n)\right) < \varepsilon,$$

da cui $A \in \mathcal{A}$. Se poi $B = \bigcap\limits_{n=1}^{\infty} A_n$, allora $B \subset \bigcap_{n=1}^{\infty} U_n$ e si vede subito che $\mu(\bigcap_{n=1}^{\infty} U_n \setminus B) < \varepsilon$. Se poniamo $V_N = \bigcap_{n=1}^{N} U_n$, otteniamo una successione decrescente di aperti di misura finita che contengono B, per cui $\mu(V_N \setminus B) \to \mu(\bigcap_{n=1}^{\infty} U_n \setminus B)$. Dunque per N abbastanza grande avremo $\mu(V_N \setminus B) < \varepsilon$ e $B \in \mathcal{A}$.

Ovviamente la famiglia $\mathcal{A}$ contiene gli aperti di X. Essendo chiusa per intersezione numerabile, contiene anche i chiusi: ogni chiuso C in uno spazio metrico si scrive come intersezione numerabile di aperti,

$$C = \bigcap_{n=1}^{\infty}\{x \in X : \operatorname{dist}(x, C) < 1/n\}.$$

A questo punto basta definire $\mathcal{A}' = \{A \in \mathcal{A} : A^C \in \mathcal{A}\}$: questa è una σ-algebra di insiemi (è per definizione chiusa per complementazione, inoltre un'unione numerabile di elementi di $\mathcal{A}'$ è un'elemento di $\mathcal{A}$ il cui complementare appartiene ad $\mathcal{A}$, in quanto intersezione numerabile di elementi di $\mathcal{A}$), e $\mathcal{A}'$ contiene gli aperti. Dunque $\mathcal{A}'$ coincide con la σ-algebra di Borel, come volevasi dimostrare.

Vediamo come comportarci per misure infinite. Se $\mu(X) = +\infty$, usiamo gli aperti V_j nell'ipotesi: dato un boreliano A, applichiamo quanto appena dimostrato alle misure finite $\mu|_{V_j}$ (definite da $\mu|_{V_j}(A) = \mu(A \cap V_j)$ e per ogni $\varepsilon > 0$ troviamo aperti U_j tali che $U_j \cap V_j \supset A \cap V_j$ e $\mu(U_j \cap V_j) < \mu(A \cap V_j) + \varepsilon/2^j$. Si vede subito che l'aperto $U = \bigcup_{j=1}^{\infty}(U_j \cap V_j)$ contiene A e ne approssima la misura a meno di ε.

Abbiamo dimostrato dunque che (*) vale per i Boreliani, nella sola ipotesi che la misura sia di Borel. Se poi la misura è di Borel regolare, (*) vale evidentemente per ogni insieme.

La (**) segue immediatamente passando ai complementari.

Infine, se $\mu(K) < +\infty$ per ogni K compatto e ricordiamo che per ipotesi X è uno spazio metrico separabile localmente compatto, è facile mostrare che esiste una successione di aperti V_j come in ipotesi: se $\{x_j\}$ è un denso numerabile, troviamo raggi r_j tali che le palle chiuse $B_{r_j}(x_j)$ sono compatti che ricoprono X. Allora le palle aperte con gli stessi raggi hanno la proprietà richiesta.

La (***) segue facilmente perché ogni chiuso può essere scritto come unione di una successione crescente di compatti (si usino opportunamente le palle chiuse di cui sopra...). Q.E.D.

Quando si costruiscono misure (esterne) su uno spazio metrico, è utile avere un criterio che garantisca la misurabilità dei boreliani:

TEOREMA (Criterio di Caratheodory): Sia μ *una misura esterna su* X *tale che* $\mu(A \cup B) = \mu(A) + \mu(B)$ *ogni volta che* $\operatorname{dist}(A, B) > 0$. *Allora* μ *è una misura di Borel.*

DIM. - NON VISTA IN CLASSE: Ci basta dimostrare che ogni chiuso C è misurabile, cioè che per ogni insieme $T \subset X$ si ha

$$\mu(T) \geq \mu(T \setminus C) + \mu(T \cap C).$$

Consideriamo i chiusi $C_j = \{x \in X : \operatorname{dist}(x, C) \leq 1/j\}$: poiché $T \setminus C_j$ si trova a distanza positiva da C si ha $\mu(T) \geq \mu((T \setminus C_j) \cup (T \cap C)) = \mu(T \setminus C_j) + \mu(T \cap C)$.

Per concludere basta allora mostrare che $\mu(T \setminus C_j) \to \mu(T \setminus C)$ per $j \to +\infty$. D'altra parte, se

$$R_k = \{x \in T : \frac{1}{k+1} < \operatorname{dist}(x, C) \leq \frac{1}{k}\}$$

abbiamo $T \setminus C = (T \setminus C_j) \cup (\bigcup_{k=j}^{\infty} R_k)$ e possiamo concludere grazie alla subadditività se facciamo vedere che

$$\lim_{j \to +\infty} \sum_{k=j}^{\infty} \mu(R_k) = 0.$$

Questo è vero perché la serie $\sum_{k=1}^{\infty} \mu(R_k)$ converge. Consideriamo infatti una somma di un numero finito di termini *di indice pari*: usando l'additività della misura sugli insiemi a distanza positiva e poi la monotonia si ha $\sum_{k=1}^{N} \mu(R_{2k}) = \mu(\bigcup_{k=1}^{N} R_{2k}) \leq \mu(T)$. Analoga maggiorazione vale per la somma di un numero finito di termini dispari: quindi, le somme parziali della serie sono maggiorate da $2\mu(T)$ e la serie converge (si noti infatti che se avevamo $\mu(T) = +\infty$ non c'era nulla da dimostrare!). Q.E.D.

Osserviamo che per misure generali il teorema di Lusin, e il conseguente teorema di densità delle funzioni continue, sono un pochino delicati: essi dipendono essenzialmente dalla possibilità di approssimare la misura di un insieme con aperti che lo contengono e con compatti in esso contenuti. Grazie al risultato che abbiamo visto, tutto funziona se abbiamo una misura di Radon (al solito, su uno spazio metrico localmente compatto e separabile)...La verifica dei particolari è lunghetta ma non difficile!

Sia ora μ una misura su un insieme X, con σ-algebra dei misurabili $\mathcal{S}$. Possiamo fabbricare un gran numero di nuove misure col metodo seguente: data una funzione misurabile $u : X \to [0, +\infty]$ definitamo una misura ν sulla σ-algebra $\mathcal{S}$ ponendo

$$(****)\ \nu(A) = \int_A u(x)\, d\mu(x) \qquad \forall A \in \mathcal{S}.$$

È immediato verificare che ν è una misura positiva. Inoltre, essa ha la proprietà che $\nu(A) = 0$ ogni volta che $\mu(A) = 0$: questo fatto si esprime dicendo che ν è *assolutamente continua* rispetto a μ.

DEFINIZIONE: Date due misure μ, ν sulla stessa σ-algebra $\mathcal{S}$, diciamo che ν è assolutamente continua rispetto a μ (e scriviamo $\nu << \mu$) se $A \in \mathcal{S}$, $\mu(A) = 0$ implica $\nu(A) = 0$.

Se μ è una misura finita, tutte le misure assolutamente continue rispetto a μ possono essere scritte come in (****):

TEOREMA (di Radon-Nikodym): Sia μ una misura finita su X (cioè $\mu(X) < +\infty$), ν un'altra misura finita definita sulla stessa σ-algebra $\mathcal{S}$ tale che $\nu << \mu$. Allora esiste una funzione $w \in L^1(\mu)$, $w \geq 0$, tale che

$$\nu(A) = \int_A w(x)\, d\mu(x) \quad \forall A \in \mathcal{S}.$$

DIM.: Consideriamo la misura $\rho = \mu + \nu$. Osserviamo che, grazie alla assoluta continuità di ν rispetto a μ, due funzioni uguali quasi ovunque rispetto a μ o a ρ sono uguali quasi ovunque anche rispetto a ν.

Se $u \in L^1(\rho)$ definiamo $T(u) := \int_X u\, d\nu$. Questo funzionale è lineare: inoltre, usando la disuguaglianza di Cauchy-Schwarz abbiamo

$$T(u) \leq \int_X |u|\, d\nu \leq \int_X |u|\, d\rho \leq \|u\|_{L^2(\rho)} \rho(X)^{1/2}.$$

Ne segue che $T \in (L^2(\rho))'$: per il teorema di Riesz sul duale di uno spazio di Hilbert, esiste un'unica funzione $v \in L^2(\rho)$ tale che

$$(I) \int_X u \, d\nu = \int_X vu \, d\rho \quad \forall u \in L^2(\rho).$$

Vien voglia allora di scrivere

$$\int_X (1-v)u \, d\nu = \int_X uv \, d\mu \quad \forall u \in L^2(\rho)$$

e di prendere $u = \mathbf{1}_E \frac{1}{1-v}$: se questa funzione appartenesse a $L^2(\rho)$, avremmo la tesi con $w = v/(1-v)$. Siccome questo in generale non è vero...e rischiamo anche di aver diviso per 0, dobbiamo inventarci un'altra strada: vedremo la prossima volta come concludere in modo rigoroso la dimostrazione!

22 Lezione del 24/11/2009 (2 ore)

Concludiamo la dimostrazione del teorema di Radon-Nikodym.

Applicando (I) alla funzione $u = \mathbf{1}_E$, con E misurabile, si ottiene $\nu(E) = \int_E v \, d\rho$. Siccome poi $0 \leq \nu(E) \leq \rho(E)$, abbiamo anche

$$(II)\ 0 \leq \frac{1}{\rho(E)} \int_E v \, d\rho \leq 1 \quad \forall E \in \mathcal{S},\ \rho(E) > 0.$$

Da questo segue che $0 \leq v(x) \leq 1$ per ρ-quasi ogni $x \in X$ (se $E_n = \{x \in X : v(x) \geq 1 + 1/n\}$ avesse misura positiva, il termine centrale in (II) sarebbe strettamente maggiore di 1... Analogamente, v non può essere strettamente negativa in un insieme di misura positiva).

La (I) diventa allora

$$(III) \int_X (1-v)u \, d\nu = \int_X uv \, d\mu \quad \forall u \in L^2(\rho).$$

A questo punto, dato $E \in \mathcal{S}$, verrebbe voglia di prendere $u = \mathbf{1}_E \frac{1}{1-v}$: se questa funzione appartenesse a $L^2(\rho)$, la (III) ci darebbe la tesi con $w = v/(1-v)$. Siccome però *non* abbiamo questa informazione, occorre procedere in modo diverso.

Se $A = \{x \in X : v(x) = 1\}$, la (III) con $u = \mathbf{1}_A$ mostra che $\mu(A) = 0$, da cui $\nu(A) = 0$. Al di fuori da questo insieme di misura nulla, è ben definita la funzione $1/(1-v)$...che però non è detto sia in $L^2(\rho)$.

Se però $E \in \mathcal{S}$, allora per ogni $n \in \mathbf{N}$ le funzioni $v_n(x) = (1 + v(x) + v^2(x) + \ldots + v^n(x))\mathbf{1}_E(x)$ appartengono a $L^2(\rho)$ (sono limitate). Sostituendo in (III) si ha

$$\int_E (1 - v^{n+1}(x))\, d\nu(x) = \int_E (v(x) + v^2(x) + \ldots + v^n(x))\, d\mu(x).$$

Il membro di sinistra tende a $\nu(E)$ grazie al teorema di Beppo Levi (l'integranda tende crescendo a 1 per quasi ogni $x \in X$)... L'integranda nel membro di destra tende crescendo a $w(x) = \frac{v(x)}{1-v(x)}$, per cui (sempre grazie al teorema di Beppo Levi) l'integrale tende a $\int_E w(x)\, d\mu(x)$. Si ha insomma $\nu(E) = \int_E w(x)\, d\mu(x)$. La sommabilità di w viene dal fatto che ν è una misura finita. Q.E.D.

ESERCIZIO: Il teorema rimane vero se μ e ν sono *misure σ-finite.*[14] In questo caso, la funzione w nella tesi non è necessariamente sommabile. *(SUGG.: Non è restrittivo supporre che gli X_i nella definizione di σ-finitezza siano due a due disgiunti, e vadano bene contemporaneamente per μ e per ν. Si applichi poi il teorema alla restrizione delle misure ad X_i...)*

Il teorema è invece falso se le misure non sono σ-finite: si prenda ad esempio come ν la misura di Lebesgue su $\mathbf{R}$, come μ la counting measure sempre su $\mathbf{R}$ (ristretta alla σ-algebra dei misurabili secondo Lebesgue): allora $\nu << \mu$, ma non può esistere una funzione w come nell'enunciato (dimostrarlo per esercizio!)!

È giunto il momento di introdurre le *misure con segno*!

DEFINIZIONE (Misura con segno): Sia $\mathcal{S}$ una σ-algebra di sottinsiemi di un dato insieme X. Una *misura con segno* (finita) è una funzione $\mu : \mathcal{S} \to \mathbf{R}$ tale che $\mu(\emptyset) = 0$ e $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ se $A_n \in \mathcal{S}$ per ogni n e $A_m \cap A_n = \emptyset$ per $m \neq n$ (numerabile additività).

Si noti che la richiesta di numerabile additività è meno innocua di quanto possa sembrare: se si permutano gli insiemi A_n, la loro unione non cambia e quindi la somma della serie deve sempre essere $\mu(A)$. Questo implica che la serie deve *sempre convergere assolutamente*, per ogni decomposizione numerabile di un insieme misurabile in insiemi misurabili due a due disgiunti.

Evidentemente, una misura con segno non gode di proprietà di monotonia: $A \subset B$ non implica $\mu(A) \leq \mu(B)$. In compenso, è facile verificare che valgono

[14] Per definizione, μ è σ-finita se esiste una successione di insiemi misurabili X_i tali che $X = \bigcup_{i=1}^{\infty} X_i$ e $\mu(X_i) < +\infty$. Per esempio, la misura di Lebesgue in $\mathbf{R}^n$ è σ-finita. La counting measure su un insieme più che numerabile non lo è.

le usuali proprietà di continuità della misura sulle successioni crescenti e decrescenti di insiemi: verificatelo per esercizio!

*DEFINIZIONE (Insiemi positivi e negativi per una misura con segno):*Data una misura con segno μ, un insieme $P \in \mathcal{S}$ si dice *positivo* se $\mu(E) \geq 0$ per ogni $E \in \mathcal{S}$, $E \subset P$. In modo analogo si definisce un insieme *negativo*. Un insieme nullo sarà poi un insieme i cui sottinsiemi hanno tutti misura nulla: esso è contemporaneamente un insieme positivo e negativo!

TEOREMA (Decomposizione di Hahn di una misura con segno): Sia μ una misura con segno su X (con σ-algebra dei misurabili $\mathcal{S}$). Allora esistono un insieme positivo $P \in \mathcal{S}$ e un insieme negativo $N \in \mathcal{S}$ disgiunti e tali che $P \cup N = X$. Una tale decomposizione si chiama decomposizione di Hahn *per la misura μ. Essa è unica a meno di insiemi nulli.*

DIM.: L'unicità della decomposizione di Hahn a meno di insiemi nulli è ovvia...l'esistenza un po' meno! Procediamo per passi, facendo vedere per prima cosa che

CLAIM I: per ogni fissato insieme misurabile M si ha $\sup\{\mu(E) : E \in \mathcal{S}, E \subset M\} < +\infty$

Dimostriamo l'affermazione per assurdo. Se fosse $\sup\{\mu(E) : E \in \mathcal{S}, E \subset M\} = +\infty$, mostriamo che possiamo trovare due insiemi misurabili disgiunti A e B tali che $A \cup B = M$, $|\mu(A)| \geq 1$ e $\sup\{\mu(E) : E \in \mathcal{S}, E \subset B\} = +\infty$.

Se ci riusciamo, possiamo poi applicare lo stesso procedimento a B, decomponendolo in due insiemi con proprietà analoghe: iterando questo passo, riusciamo a costruire due successioni di insiemi misurabili A_n, B_n tali che $A_n \cap B_n = \emptyset$, $A_n \cup B_n = B_{n-1}$,$|\mu(A_n)| \geq 1$ e $\sup\{\mu(E) : E \in \mathcal{S},\ E \subset B_n\} = +\infty$. In particolare, gli A_n sono disgiunti ed hanno tutti misura di modulo maggiore o uguale a 1: questo è assurdo, perchè si dovrebbe avere

$$\sum_{n=1}^{\infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \in \mathbf{R},$$

ma una serie il cui termine generale non tende a 0 non può convergere!

Mostriamo dunque che esistono A e B con le proprietà dette. Grazie all'ipotesi che $\sup\{\mu(E) : E \in \mathcal{S}, E \subset M\} = +\infty$ possiamo scegliere un insieme misurabile B tale che $\mu(B) > 1 + |\mu(M)|$, e poniamo $A = M \setminus B$. Allora $\mu(M) = \mu(A) + \mu(B) > \mu(A) + 1 + |\mu(M)|$, da cui $\mu(A) < -1$: sia A che B hanno misura maggiore o uguale a 1 in modulo. Ora, i due estremi superiori

$$\sup\{\mu(E) : E \in \mathcal{S}, E \subset B\},$$
$$\sup\{\mu(E) : E \in \mathcal{S}, E \subset A\}$$

non possono essere entrambi finiti, altrimenti lo sarebbe anche lo stesso sup fatto su tutti i sottinsiemi di M: il nostro asserto è dunque provato a patto di scambiare A e B se necessario.

CLAIM II: per ogni $A \in \mathcal{S}$ e per ogni $\varepsilon > 0$ possiamo trovare $B \in \mathcal{S}$, $B \subset A$ tale che $\mu(B) \geq \mu(A)$ e $\mu(E) > -\varepsilon$ per ogni $E \subset B$, $E \in \mathcal{S}$.

In sostanza, dobbiamo trovare un sottinsieme "quasi positivo" di A, con misura maggiore o uguale a quella di A... Sia $c = \sup\{\mu(C) : C \subset A,\ C \in \mathcal{S}\}$: evidentemente $\mu(A) \leq c < +\infty$ (grazie al Claim I), per cui possiamo trovare un sottinsieme misurabile $B \subset A$ tale che

$$\mu(B) \geq \max\{\mu(A), c - \varepsilon/2\}.$$

Questo insieme ha le proprietà richieste: se esistesse $E \subset B$ con $\mu(E) \leq -\varepsilon$, allora $\mu(B \setminus E) = \mu(B) - \mu(E) \geq c + \varepsilon/2$, contro la definizione di estremo superiore.

Nel terzo passo, il nostro insieme "quasi positivo" diventa positivo:

CLAIM III: se $A \in \mathcal{S}$ esiste un insieme positivo $B \subset A$ tale che $\mu(B) \geq \mu(A)$.

Applichiamo infatti il Claim II con $\varepsilon = 1/n$: troviamo una successione decrescente di insiemi misurabili $A \supset A_1 \supset A_2 \supset A_3 \supset \ldots$ tali che $\mu(A_n) \geq \mu(A)$ e $\mu(E) > -1/n$ per ogni E misurabile, $E \subset A_n$.

Poniamo $B = \bigcap_{n=1}^{\infty} A_n$. Allora $B \subset A$ e, per la continuità della misura sulle successioni decrescenti, $\mu(B) \geq \mu(A)$. Inoltre B è un insieme positivo: se $E \subset B$, allora E è anche un sottinsieme di A_n per ogni n e quindi $\mu(E) > -1/n$.

Dimostrato il Claim III, siamo in grado di trovare la nostra decomposizione di Hahn: sia $s = \sup\{\mu(A) : A \in \mathcal{S}\}$. Prendiamo una successione A_n di insiemi misurabili tale che $\mu(A_n) \to s$: grazie al Claim III posso sostituire ciascuno degli A_n con un suo sottinsieme positivo B_n tale che $\mu(B_n) \geq \mu(A_n)$, per cui $\mu(B_n) \to s$. Definiamo poi

$$P_n = \bigcup_{k=1}^{n} B_k :$$

questi insiemi costituiscono una successione crescente di insiemi positivi tali che $\mu(P_n) \to s$. A questo punto, possiamo porre

$$P = \bigcup_{n=1}^{\infty} P_n, \quad N = X \setminus P.$$

Per quanto detto sopra, $\mu(P) = s$. Inoltre P è positivo: se $E \subset P$, $E \in \mathcal{S}$, allora gli insiemi $E \cap P_n$ formano una successione crescente di insiemi di misura positiva la cui unione è E, per cui $\mu(E) \geq 0$. Infine, N è un insieme negativo: se esistesse $E \subset N$ misurabile con $\mu(E) > 0$, allora $\mu(P \cup E) = \mu(P) + \mu(E) > s$, assurdo! Q.E.D.

DEFINIZIONE (Variazione positiva, negativa, totale e decomposizione di Jordan di μ): Sia μ una misura con segno su X, $X = P \cup N$ una decomposizione di Hahn per μ. Per ogni $E \in \mathcal{S}$ definiamo

$$\begin{array}{rll} \mu^+(E) = & \mu(E \cap P) & \textit{(Variazione positiva di } \mu\textit{)}, \\ \mu^-(E) = & -\mu(E \cap N) & \textit{(Variazione negativa di } \mu\textit{)}, \\ |\mu|(E) = & \mu^+(E) + \mu^-(E) & \textit{(Variazione totale di } \mu\textit{)}. \end{array}$$

Si tratta evidentemente di misure positive, ed inoltre $\mu = \mu^+ - \mu^-$ (decomposizione di Jordan della misura μ).

Quella che segue è una semplice caratterizzazione della misura variazione totale: essa mostra come la variazione totale sia la *più piccola* misura positiva che maggiora il modulo di μ.

PROPOSIZIONE: Se μ è una misura con segno, allora per ogni $A \in \mathcal{S}$ si ha

$$|\mu|(A) = \sup\{\sum_{n=1}^{\infty} |\mu(E_n)| : \ E_n \in \mathcal{S}, \ A = \bigcup_{n=1}^{\infty} E_n, \ E_n \cap E_m = \emptyset \textit{ per } m \neq n\}.$$

DIM. - NON VISTA IN CLASSE: Siano A, E_n come nell'enunciato, $X = P \cup N$ una decomposizione di Hahn di μ, e siano μ^+ e μ^- le variazioni. Allora

$$\sum_{n=1}^{\infty} |\mu(E_n)| = \sum_{n=1}^{\infty} |\mu^+(E_n) - \mu^-(E_n)| \leq$$
$$\sum_{n=1}^{\infty} (\mu^+(E_n) + \mu^-(E_n)) = \sum_{n=1}^{\infty} |\mu|(E_n) = |\mu|(A).$$

L'estremo superiore nell'enunciato è poi un massimo: basta decomporre A nei due insiemi $A \cap P$ e $A \cap N$. Q.E.D.

23 Lezione del 26/11/2009 (2 ore)

Grazie a quanto sappiamo ora sulle misure con segno, possiamo fare il seguente, semplice

ESERCIZIO (Teorema di Radon-Nikodym per le misure con segno): Sia μ una misura positiva finita su X, ν una misura *con segno* con $\nu << \mu$. Allora esiste una funzione $v \in L^1(\mu)$ tale che

$$\nu(E) = \int_E v(x)\, d\mu(x) \quad \forall E \in \mathcal{S}.$$

(SUGG.: Sia $X = P \cup N$ una decomposizione di Hahn. Basta applicare il teorema di Radon-Nikodym alle misure positive ν^+ e ν^-, che sono concentrate su insiemi disgiunti...)

Il teorema di Radon-Nikodym per le misure con segno ci consente finalmente di dimostrare che il duale di $L^p(\mu)$ è $L^q(\mu)$, per $1 \leq p < +\infty$ e a patto che la misura positiva μ sia finita[15]:

TEOREMA: Sia μ una misura positiva finita su X, $1 \leq p < +\infty$. Allora per ogni $T \in (L^p(\mu))'$ esiste un'unica funzione $v \in L^q(\mu)$ (con q esponente coniugato di p) tale che

$$T(u) = \int_X u(x)v(x)\, d\mu(x) \quad \forall u \in L^p(\mu).$$

Inoltre, $\|T\| = \|v\|_{L^q}$.

DIM.: Abbiamo già visto che l'applicazione $\Phi : L^q \to (L^p)'$ che ad ogni $v \in L^q(\mu)$ associa il funzionale $T_v : u \mapsto \int_X uv\, d\mu$ è un'isometria lineare. Ci manca solo da verificare la suriettività di Φ, ed il teorema sarà completamente dimostrato.

Sia dunque $T \in (L^p(\mu))'$. Definiamo $\nu(E) = T(\mathbf{1}_E)$ (si noti che $\mathbf{1}_E \in L^p$ grazie alla finitezza della misura μ): dico che ν è una misura con segno su X, assolutamente continua rispetto a μ.

Infatti, è ovvio che $\nu(E) = 0$ quando $\mu(E) = 0$. Inoltre, se A e B sono misurabili e disgiunti si ha $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$ da cui $\nu(A \cup B) = \nu(A) + \nu(B)$ grazie alla linearità del funzionale.

Verifichiamo la numerabile additività di ν: sia $A = \bigcup_{n=1}^\infty A_n$, con gli A_n misurabili e due a due disgiunti. Si vede subito che

$$\mathbf{1}_A(x) = \sum_{n=1}^\infty \mathbf{1}_{A_n}(x),$$

e la successione delle somme parziali è dominata da $\mathbf{1}_A$: se ne deduce che la serie scritta sopra converge in $L^p(\mu)$.

[15] La dimostrazione si adatta abbastanza facilmente anche al caso delle misure σ-finite

Allora, per la continuità di T si ha $\nu(A) = \sum_{n=1}^{\infty} \nu(A_n)$ e la numerabile additività di ν risulta provata.

Il teorema di Radon-Nikodym garantisce l'esistenza di una funzione $v \in L^1(\mu)$ tale che

$$(A)\ T(\mathbf{1}_E) = \nu(E) = \int_E v(x)\, d\mu(x) \quad \forall E \in S,$$

da cui

$$(B)\ T(s) = \int_X s(x)v(x)\, d\mu(x) \quad \forall s \text{ semplice}.$$

Rimane da far vedere che $v \in L^q$: se questo è vero, al posto della funzione semplice possiamo mettere qualunque funzione $u \in L^p$ perché le funzioni semplici sono dense in quello spazio.

Cominciamo dal caso $p = 1$. Sappiamo da (A) che

$$\left| \int_E v(x)\, d\mu(x) \right| \leq \|T\| \mu(E) \quad \forall E \in \mathcal{S},$$

da cui $\mu(\{x :\ v(x) > \|T\| + 1/n\}) = 0$ per ogni n (altrimenti la disuguaglianza non sarebbe verificata), e analogamente $\mu(\{x :\ v(x) < -\|T\| - 1/n\}) = 0$ da cui $\|v\|_\infty \leq \|T\|$.

Sia poi $1 < p < +\infty$. La relazione (B) vale per tutte le funzioni $s \in L^\infty(\mu)$ (perché ogni funzione limitata si può approssimare uniformemente con funzioni semplici: si veda la dimostrazione del teorema di approssimazione di una funzione misurabile positiva con funzioni semplici!). Per ogni $n \in \mathbf{N}$ sia $E_n = \{x \in X : |v(x)| \leq n\}$ e poniamo $s_n(x) = \mathbf{1}_{E_n}(x)|v(x)|^{q-1}\mathrm{sgn}(v(x))$. Ciascuna di queste funzioni è in $L^\infty(\mu)$, e inoltre $|v(x)|^q = |s_n(x)|^p$ su E_n. Da (B) otteniamo allora

$$\int_{E_n} |v(x)|^q\, d\mu(x) = \int_X s_n(x)v(x)\, d\mu(x) = T(s_n) \leq \|T\| \left(\int_{E_n} |v(x)|^q\, d\mu(x) \right)^{1/p},$$

cioè

$$\left(\int_{E_n} |v(x)|^q\, d\mu(x) \right)^{1/q} \leq \|T\|.$$

Passiamo al limite per $n \to +\infty$: grazie al teorema di Beppo Levi abbiamo finalmente $\|v\|_{L^q} \leq \|T\|$. Q.E.D.

Per studiare problemi che coinvolgono equazioni differenziali (ordinarie o a derivate parziali), abbiamo bisogno di spazi di funzioni *dotate di derivate* e che godano di buone proprietà di compattezza. A questo scopo, sono utilissimi gli *spazi di Sobolev* $W^{1,p}([a,b])$, una famiglia di spazi modellati su L^p. Per definirli, è necessaria seguente nozione:

DEFINIZIONE (Derivata debole): Sia $u \in L^1([a,b])$. Una funzione $v \in L^1([a,b])$ si dice *derivata debole* di u se vale

$$\int_a^b u(x)\phi'(x)\,dx = -\int_a^b v(x)\phi(x)\,dx \quad \forall \phi \in \mathcal{C}_0^1([a,b]).$$

Si noti che se u è di classe $\mathcal{C}^1$, allora la sua derivata è una derivata debole: è la formula di integrazione per parti!

Viceversa vedremo tra poco (è il classico lemma di du Bois-Reymond nel Calcolo delle Variazioni) che se u e v sono continue, allora u è derivabile con derivata v: è infatti a questo risultato che possiamo far risalire l'idea di derivata debole.

Inoltre, si vede facilmente che la derivata debole, se esiste, è anche *unica*. Basta usare il

LEMMA (fondamentale del Calcolo delle Variazioni): Sia $w \in L^1([a,b])$ *tale che* $\int_a^b w\phi\,dx = 0$ *per ogni* $\phi \in \mathcal{C}_C^1([a,b])$. *Allora* $u = 0$ *quasi ovunque.*

DIM.: Si approssimi la funzione $\operatorname{sgn} u(x)$ con una successione ϕ_n di funzioni $\mathcal{C}_C^1$ comprese tra -1 e 1 che vi convergono puntualmente quasi ovunque. Il teorema della convergenza dominata ci dice allora che

$$0 = \int_a^b w\phi_n\,dx \to \int_a^b |u(x)|\,dx,$$

da cui $u = 0$ quasi ovunque. Q.E.D.

Ora, se v e $\tilde{v}$ sono *due* derivate deboli di u, segue dalla definizione di derivata debole che la funzione $w = v - \tilde{v}$ soddisfa le ipotesi del lemma: ne segue che $v = \tilde{v}$ quasi ovunque. Grazie all'unicità della derivata debole, siamo allora legittimati ad indicarla con u'.

Grazie a questa definizione, è facile definire gli spazi di Sobolev:

DEFINIZIONE (Spazi di Sobolev $W^{1,p}$): Se $1 \le p \le +\infty$ definiamo lo spazio di Sobolev

$$W^{1,p}([a,b]) = \{u \in L^p([a,b]) : \text{ esiste } u' \text{ derivata debole di } u, \text{ e } u' \in L^p([a,b])\}.$$

Sullo spazio $W^{1,p}$ si usa mettere una delle seguenti due norme equivalenti:

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|u'\|_{L^p} \text{ oppure } \|u\|_{W^{1,p}} = (\|u\|_{L^p}^p + \|u'\|_{L^p}^p)^{1/p}.$$

Useremo indifferentemente l'una o l'altra. La seconda è vantaggiosa per $p = 2$ perché deriva da un prodotto scalare.

OSSERVAZIONE: E' facile vedere che $W^{1,p}$ è uno spazio normato completo (spazio di Banach): se $\{u_k\} \subset W^{1,p}$ è di Cauchy, allora sia $\{u_k\}$ che $\{u'_k\}$ sono di Cauchy in L^p. Siccome L^p è completo, esistono u, v tali che $u_k \to u$, $u'_k \to v$ in L^p.

Resta da mostrare che $u \in W^{1,p}$ e $v = u'$. Infatti, per definizione di derivata debole abbiamo

$$\int_a^b u_k \phi' \, dx = -\int_a^b u'_k \phi \, dx \quad \forall \phi \in \mathcal{C}_0^1.$$

Passando al limite per $k \to +\infty$ si ottiene

$$\int_a^b u \phi' \, dx = -\int_a^b v \phi \, dx,$$

come volevasi. Si noti che quest'ultimo ragionamento sarebbe corretto anche se avessimo soltanto la convergenza *debole* di u_k a u e di u'_k a v: quest'osservazione ci tornerà utile in seguito!

La nostra definizione degli spazi di Sobolev può essere facilmente generalizzata a funzioni definite su $\mathbf{R}^n$. Tuttavia, nel caso unidimensionale vale una caratterizzazione particolarmente potente degli spazi di Sobolev, che ne semplifica moltissimo l'utilizzo: vedremo tra un attimo che per molti versi le funzioni di Sobolev assomigliano moltissimo alle funzioni $\mathcal{C}^1$ perché sono continue (a meno di modificarle su un insieme di misura nulla) e sono primitive della loro derivata debole!

Questa caratterizzazione purtroppo non vale in dimensione superiore.

Per caratterizzare gli spazi di Sobolev abbiamo però bisogno di una definizione:

DEFINIZIONE (Funzioni assolutamente continue): Definiamo lo spazio delle funzioni assolutamente continue come l'insieme delle funzioni che sono *primitive di funzioni* L^1:

$$AC([a,b]) = \{u : [a,b] \to \mathbf{R} : \exists v \in L^1([a,b]) \; t.c. \; u(x) = u(a) + \int_a^x v(t) \, dt \; \forall x \in [a,b]\}.$$

Un teorema di analisi reale non facilissimo (ma neanche eccessivamente difficile!) mostra che una tale u è derivabile quasi ovunque, e che $u'(x) = v(x)$ per quasi ogni x: vale dunque il teorema fondamentale del calcolo integrale

nel senso che u è primitiva della propria derivata. Inoltre, vale la seguente caratterizzazione delle funzioni AC con gli ε e i δ:
Una funzione $u : [a,b] \to \mathbf{R}$ è assolutamente continua se e soltanto se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che, per ogni collezione finita $[a_1, b_1]$, $[a_2, b_2]$,...,$[a_N, b_N]$ di sottointervalli due a due disgiunti di $[a,b]$ con $\sum_{i=1}^{N}(b_i - a_i) < \delta$, vale

$$\sum_{i=1}^{N} |u(b_i) - u(a_i)| < \varepsilon.$$

Come si vede, questa caratterizzazione è un rafforzamento della definizione di uniforme continuità: in particolare, una funzione assolutamente continua è ovunque continua (come si poteva mostrare anche a partire dalla definizione originale...). Inoltre, la caratterizzazione consente di provare facilmente che il *prodotto di due funzioni assolutamente continue* è una funzione assolutamente continua.

24 Lezione del 27/11/2009 (2 ore)

Proveremo ora che lo spazio delle funzioni assolutamente continue *coincide* in buona sostanza con lo spazio di Sobolev $W^{1,1}([a,b])$. Precisamente, ogni funzione assolutamente continua appartiene allo spazio di Sobolev e, viceversa, data $u \in W^{1,1}$ esiste una funzione assolutamente continua che coincide quasi ovunque con u.

Ci servono due lemmi:

LEMMA 1 (du Bois-Reymond): Se $u \in W^{1,1}([a,b])$ ha derivata debole nulla, allora u coincide quasi ovunque con una costante.

DIM.: Sia $\psi \in \mathcal{C}^0([a,b])$: definiamo $w(x) = \psi(x) - \frac{1}{b-a}\int_a^b \psi(t)\, dt$ e

$$\phi(x) = \int_a^x w(t)\, dt.$$

Allora $\phi \in \mathcal{C}^1_0([a,b])$ e per ipotesi abbiamo

$$0 = \int_a^b u(x)\phi'(x)\, dx =$$
$$\int_a^b [u(x)h(x) - u(x)\frac{1}{b-a}\int_a^b h(s)\, ds]\, dx =$$
$$\int_a^b [u(x) - \frac{1}{b-a}\int_a^b u(s)\, ds]h(x)\, dx$$

dove, nell'ultimo passaggio, abbiamo scambiato l'ordine di integrazione nel secondo termine (e scambiato anche il nome delle variabili di integrazione...).

Per il lemma fondamentale del Calcolo delle Variazioni, questo implica

$$u(x) = \frac{1}{b-a} \int_a^b u(s)\, ds \quad \forall x \in [a,b].$$

Q.E.D.

LEMMA 2: Se $u \in AC([a,b])$, allora $u \in W^{1,1}([a,b])$. Inoltre, la derivata classica di u (che sappiamo definita quasi ovunque) è anche la derivata debole di u.

DIM.: Intanto, dalla definizione di AC segue che sia u che u' sono in L^1. Sia poi $\phi \in \mathcal{C}_0^1([a,b])$: evidentemente, anche ϕ è una funzione assolutamente continua.

Allora la funzione prodotto $u\phi$ è assolutamente continua e si ha $(u\phi)' = u'\phi + u\phi'$ quasi ovunque. Integrando:

$$0 = \int_a^b (u\phi)'\, dx = \int_a^b (u'\phi + u\phi')\, dx,$$

che è esattamente come dire che u' è la derivata debole di u. Q.E.D.

TEOREMA: Sia $u \in W^{1,1}([a,b])$. Allora esiste $\tilde{u} \in AC([a,b])$ tale che $u(x) = \tilde{u}(x)$ per quasi ogni x. Quindi, a meno di cambiare u su un insieme di misura nulla per renderla AC, la derivata debole di u coincide con la sua derivata classica[16].

DIM.: Definiamo $w(x) = \int_a^x u'(t)\, dt$. Questa è una funzione assolutamente continua che, per il LEMMA 2, è anche in $W^{1,1}$ ed ha derivata debole u'. Allora la funzione $u - w$ ha derivata debole nulla, da cui per il LEMMA 1 $u(x) - w(x) = c$ quasi ovunque. Basta allora porre $\tilde{u}(x) = c + w(x)$ per avere la tesi. Q.E.D.

OSSERVAZIONE: Se $1 < p \leq +\infty$, le funzioni di $W^{1,p}$ sono anche in $W^{1,1}$. Se $u \in W^{1,p}$ possiamo dunque supporre (scegliendo il giusto rappresentante della classe di equivalenza di u in L^p) che la nostra funzione sia assolutamente continua. In questo senso, lo spazio $W^{1,p}$ coincide con lo spazio delle funzioni assolutamente continue *la cui derivata appartiene a L^p*. Queste funzioni non sono soltanto assolutamente continue, sono anche Hölderiane di esponente $1 - 1/p$: se $x, y \in [a,b]$ si ha infatti, usando la disuguaglianza di Hölder:

$$|u(x) - u(y)| = |\int_x^y u'(t)\, dt| \leq \|u'\|_{L^p} |x-y|^{1-1/p}.$$

[16] Si tenga presente il LEMMA 2.

Le funzioni di $W^{1,\infty}$ sono poi lipschitziane.

In realtà (sempre a patto di prendere il rappresentante assolutamente continuo di ogni classe di equivalenza), lo spazio $W^{1,\infty}$ *coincide* con lo spazio delle funzioni lipschitziane. Infatti, ogni funzione lipschitziana è assolutamente continua, e se u ha costante di Lipschitz L, i suoi rapporti incrementali sono limitati da L...e quindi $|u'(x)| \leq L$in tutti i punti di derivabilità: una funzione lipschitziana ha derivata in L^∞.

L'osservazione appena fatta (il fatto che le funzioni di $W^{1,p}$ con $p > 1$ sono hölderiane, con costante controllata dalla norma L^p di u') è la chiave del seguente importante risultato di compattezza:

TEOREMA (di compattezza debole in $W^{1,p}$): Sia $\{u_n\} \subset W^{1,p}([a,b])$, $1 < p < +\infty$ (e supponiamo che ognuna delle u_n sia AC). Se esiste una costante $C > 0$ tale che $\|u_n'\|_{L^p} \leq C$ per ogni n, e vale anche una delle seguenti due condizioni:

(i) $|u_n(a)| \leq C$

(ii) $\|u_n\|_{L^p} \leq C$

allora esiste $u \in W^{1,p}$ e una sottosuccessione $\{u_{n_k}\}$ tali che $u_{n_k} \to u$ uniformemente, $u'_{n_k} \rightharpoonup u'$ debolmente in L^p per $k \to +\infty$.

Dimostriamo il teorema di compattezza in $W^{1,p}$. Grazie all'osservazione precedente e all'equilimitatezza delle derivate in L^p, tutte le nostre funzioni soddisfano la disuguaglianza di hölderianità

$$(*) |u_n(x) - u_n(y)| \leq C|x-y|^{1-1/p} \quad \forall x, y \in [a,b].$$

In particolare, le funzioni u_n sono equicontinue.

Supponiamo poi che valga (i). Allora, usando $(*)$ si ha per ogni x e per ogni n

$$|u_n(x)| \leq |u_n(a)| + |u_n(x) - u_n(a)| \leq C + C(b-a)^{1-1/p}$$

e quindi le u_n sono anche equilimitate.

Usando il teorema di Ascoli-Arzelà e il teorema di compattezza debole in L^p troviamo $u \in C^0$, $v \in L^p$ e una sottosuccessione $\{u_{n_k}\}$ tali che $u_{n_k} \to u$ uniformemente, $u'_{n_k} \rightharpoonup v$ debolmente in L^p. Come abbiamo già osservato (si veda la dimostrazione della completezza degli spazi $W^{1,p}$), questo implica che $u \in W^{1,p}$ e $v = u'$.

Rimane da dimostrare il risultato se sostituiamo la (i) con la (ii). In realtà, facciamo vedere che $(ii) \Rightarrow (i)$: abbiamo infatti, per ogni $x \in [a,b]$, $u(a) = u(x) - \int_a^x u'(x)\, dx$.

Allora $|u(a)| \leq |u(x)| + \int_a^b |u'(x)|\, dx$. Integrando su $[a, b]$ (e dividendo ambo i membri per $b - a$):

$$|u(a)| \leq \frac{1}{b-a} \int_a^b |u(x)|\, dx + \int_a^b |u'(x)|\, dx,$$

e abbiamo già visto come maggiorare le norme L^1 a destra con norme L^p (usando la disuguaglianza di Hölder). Q.E.D.

OSSERVAZIONE: Un risultato simile vale anche per $p = +\infty$: se $\{u_n\} \subset W^{1,\infty}$ possiamo utilizzare il teorema per ogni $p < +\infty$ per trovare la sottosuccessione convergente. Inoltre, la successione è formata da funzioni equilipschitziane (perché le derivate sono equilimitate in L^∞), per cui anche il limite sarà una funzione lipschitziana. Le derivate convergeranno debolmente in L^p per tutti i p finiti (e, in realtà, anche in L^∞ debole*...se vi avessi detto cosa vuol dire!).

Il teorema è invece irrimediabilmente falso per $p = 1$: è facile costruire delle successioni equilimitate in norma $W^{1,1}$ che convergono a funzioni discontinue. Per esempio, si considerino le seguenti funzioni su $[-1, 1]$:

$$u_n(x) = \begin{cases} -1 & se\ -1 \leq x \leq -1/n \\ nx & se\ -1/n < x < 1/n \\ 1 & se\ 1/n \leq x \leq 1 \end{cases}$$