

Diario del Corso di Analisi Matematica II - Mod. 2

Corso di Laurea: Matematica Applicata

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 9/12/2013 (2 ore) 4**
Teorema della divergenza. Misura di Lebesgue: motivazione, ripasso sulla misura di Peano Jordan, misura esterna di Lebesgue.
- 2 Lezione del 11/12/2013 (2 ore) 8**
Prime proprietà della misura esterna di Lebesgue. Misure esterne astratte. Insiemi misurabili secondo Caratheodory. Proprietà della misura sui misurabili.
- 3 Lezione del 13/12/2013 (2 ore) 12**
Proprietà degli insiemi misurabili e della misura sugli insiemi misurabili. Regolarità della misura di Lebesgue.
- 4 Lezione del 16/12/2013 (3 ore) 16**
Esistenza di insiemi non misurabili secondo Lebesgue. Funzioni misurabili. Funzioni semplici e loro integrale. Approssimazione di funzioni misurabili non negative con funzioni semplici. Integrale di Lebesgue di funzioni misurabili non negative.
- 5 Lezione del 18/12/2013 (2 ore) 22**
Teorema di Beppo Levi e conseguenze. Lemma di Fatou e teorema della convergenza dominata di Lebesgue.
- 6 Lezione del 19/12/2013 (1 ora) 25**
Qualche conseguenza dei teoremi di convergenza integrale. Proprietà vere "quasi ovunque". Confronto con l'integrale di Riemann.
- 7 Lezione del 20/12/2013 (2 ore) 28**
Teorema di Fubini (senza dimostrazione). Teorema di derivazione sotto il segno di integrale.
- 8 Lezione del 8/1/2014 (2 ore) 31**
Teorema di Cauchy-Lipschitz: esistenza e unicità locale per il problema di Cauchy.
- 9 Lezione del 10/1/2014 (2 ore) 35**
Prolungabilità delle soluzioni locali e soluzioni massimali. Un risultato di esistenza globale. Esistenza e unicità per i sistemi di equazioni ordinarie. Equazioni lineari di ordine n : esistenza e unicità globale per il problema di Cauchy. Equazioni lineari omogenee: struttura dell'insieme delle soluzioni.

- 10 Lezione del 13/1/2014 (2 ore)** **40**
Equazioni lineari complete: struttura dell'insieme delle soluzioni. Metodo della variazione delle costanti. Esponenziale complesso. Soluzione generale di equazioni differenziali lineari omogenee a coefficienti costanti.
- 11 Lezione del 15/1/2014 (2 ore)** **45**
Metodo degli annichilatori. Serie di Fourier: richiamo di alcuni risultati di convergenza (in L^2 , puntuale, uniforme).
- 12 Lezione del 17/1/2014 (2 ore)** **50**
Risoluzione per separazione di variabili dell'equazione del calore e dell'equazione della corda vibrante.

1 Lezione del 9/12/2013 (2 ore)

Cominciamo la seconda parte del corso con l'enunciato e la dimostrazione del *teorema della divergenza*, che il Prof. Orlandi non ha avuto il tempo di trattare.

TEOREMA (della divergenza): Sia $\Omega \subset \mathbf{R}^3$ un aperto limitato regolare¹, $F : A \subset \mathbf{R}^3 \rightarrow \mathbf{R}^3$ una funzione vettoriale di classe C^1 (ove A è un aperto che contiene $\overline{\Omega}$). Allora, se n denota la normale a $\partial\Omega$ diretta verso l'esterno di Ω , si ha

$$\int_{\partial\Omega} F \cdot n \, d\sigma = \int_{\Omega} \operatorname{div} F(x, y, z) \, dx \, dy \, dz.$$

DIMOSTRAZIONE del teorema della divergenza: Per semplicità dimostriamo il teorema in un caso particolare, cioè sotto l'ipotesi più restrittiva che l'aperto Ω sia *normale rispetto ai tre piani coordinati*.

In particolare, essendo il nostro dominio normale rispetto al piano xy (in altre parole, semplice rispetto all'asse z), esistono un aperto $D \subset \mathbf{R}^2$ e due funzioni $f, g : D \rightarrow \mathbf{R}$ di classe C^1 tali che $\Omega = \{(x, y, z) : (x, y) \in D, g(x, y) < z < f(x, y)\}$.

Facciamo vedere che

$$(**) \int_{\partial\Omega} (0, 0, F_3) \cdot n \, d\sigma = \int_{\Omega} \frac{\partial F_3}{\partial z} \, dx \, dy \, dz,$$

dove n rappresenta la normale esterna a $\partial\Omega$. I due "pezzi mancanti" dell'enunciato (quelli che coinvolgono F_2 e F_3) si otterranno allo stesso modo, sfruttando la normalità di Ω rispetto agli altri due piani coordinati: per linearità, avremo così l'enunciato del teorema per campi di vettori del tutto generali.

Calcoliamo dapprima il flusso. Si noti che il contributo all'integrale della "superficie laterale" di $\partial\Omega$ è zero, perchè il campo di vettori che stiamo considerando è diretto lungo l'asse z , mentre la normale è orizzontale: rimangono solo i contributi dovuti all'integrazione sui grafici di f e di g : in conclusione,

¹Questo significa che Ω , in un intorno di ogni punto della sua frontiera, coincide con il *sottografico* di una funzione di classe C^1 definita su uno dei piani coordinati.

ricordando la definizione di integrale di superficie, otteniamo

$$\begin{aligned} & \int_{\partial\Omega} (0, 0, F_3) \cdot n \, d\sigma = \\ & \int_D (0, 0, F_3(x, y, f(x, y))) \cdot (-f_x, -f_y, 1) \, dx \, dy - \\ & \int_D (0, 0, F_3(x, y, g(x, y))) \cdot (-g_x, -g_y, 1) \, dx \, dy = \\ & \int_D F_3(x, y, f(x, y)) \, dx \, dy - \int_D F_3(x, y, g(x, y)) \, dx \, dy. \end{aligned}$$

Si noti che il segno meno è dovuto al fatto che la normale esterna a $\partial\Omega$ punta verso il basso sul grafico di g .

D'altra parte,

$$\begin{aligned} & \int_{\Omega} \frac{\partial F_3}{\partial z} \, dx \, dy \, dz = \int_D \, dx \, dy \int_{g(x,y)}^{f(x,y)} \frac{\partial F_3}{\partial z} \, dz = \\ & \int_D F_3(x, y, f(x, y)) \, dx \, dy - \int_D F_3(x, y, g(x, y)) \, dx \, dy \end{aligned}$$

e la (***) è dimostrata. Q.E.D.

OSSERVAZIONE: Si può reinterpretare il teorema della divergenza con il linguaggio delle forme. Come sapete, al campo di vettori (F_1, F_2, F_3) possiamo far corrispondere la 2-forma

$$\omega = F_1 \, dx_2 \wedge dx_3 - F_2 \, dx_1 \wedge dx_3 + F_3 \, dx_1 \wedge dx_2.$$

Avete già visto con Giandomenico Orlandi che l'integrale di questa forma su una qualunque superficie parametrizzata sufficientemente regolare, coincide con il flusso del campo di vettori attraverso la stessa superficie (orientata con la normale indotta dalla parametrizzazione).

Ora, calcolando formalmente il differenziale esterno della forma ω troviamo la 3-forma

$$d\omega = (\operatorname{div} F) \, dx_1 \wedge dx_2 \wedge dx_3.$$

Quindi, nel linguaggio delle forme il teorema della divergenza si legge

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega.$$

L'enunciato è identico a quello del teorema di Stokes, sono solo aumentate di uno le dimensioni degli oggetti geometrici coinvolti!

La prima parte delle mie lezioni 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.

Nelle lezioni di Giandomenico Orlandi avete (sostanzialmente) 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: $I = (a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n)$. 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 $I_1, \dots, I_N, J_1, \dots, J_K \subset \mathbf{R}^n$ 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

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

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 $g, h : [a, b] \rightarrow \mathbf{R}$ due funzioni di una variabile, integrabili secondo Riemann e con $g(x) \leq h(x)$ per ogni $x \in [a, b]$. Consideriamo l'insieme $A = \{(x, y) \in \mathbf{R}^2 : x \in [a, b], g(x) \leq y \leq h(x)\}$. 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.

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

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

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

2 Lezione del 11/12/2013 (2 ore)

La misura esterna di Lebesgue gode delle seguenti proprietà elementari:

TEOREMA (Proprietà elementari della misura esterna di Lebesgue): Sia $m : \mathcal{P}(\mathbf{R}^n) \rightarrow [0, +\infty]$ la misura esterna di Lebesgue². 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 $A, A_1, A_2, \dots \subset \mathbf{R}^n$, 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).

(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 $\{I_j^i\}_j$ tali che $\bigcup_{j=1}^{\infty} I_j^i \supset A_i$ e

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

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

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

² $\mathcal{P}(\mathbf{R}^n)$ denota l'insieme delle parti di \mathbf{R}^n , ossia l'insieme di tutti i sottinsiemi di \mathbf{R}^n .

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 $j = 1, 2, \dots$ 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.

Sorprendentemente, 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 $J \subset \bigcup_{j=1}^N I_j$. 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 astratto più generale, chiamato misura esterna:

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

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

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

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

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 Carathéodory): Un sottinsieme $A \subset \mathbf{R}^n$ si dice *misurabile secondo Lebesgue* o *m-misurabile* se vale l’ugualianza

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

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

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

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

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

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 m la misura di Lebesgue su \mathbf{R}^n . Valgono i seguenti fatti

(i) Se A è misurabile secondo Lebesgue, allora $A^C = \mathbf{R}^n \setminus A$ è misurabile. Inoltre, se $m(A) = 0$ allora A è misurabile.

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

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

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

(numerabile additività della misura di Lebesgue sui misurabili).

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

$$m(A) = \lim_{i \rightarrow +\infty} m(A_i).$$

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

$$m(A) = \lim_{i \rightarrow +\infty} m(A_i).$$

Identico enunciato vale se la misura di Lebesgue è sostituita con una qualunque misura esterna μ e gli insiemi misurabili secondo Lebesgue con gli insiemi μ -misurabili.

3 Lezione del 13/12/2013 (2 ore)

Dimostriamo il teorema sugli insiemi misurabili enunciato la volta scorsa!

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

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

Che un insieme di misura nulla sia misurabile è immediato. Da questo segue in particolare che \emptyset e \mathbf{R}^n 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 $T \subset \mathbf{R}^n$ si ha

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

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 $\geq m(T \cap (A \cup B))$. Invece, l'insieme nell'ultimo addendo non è altro che $T \setminus (A \cup B)$: si ha allora

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

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 $m(A \cup B) = m((A \cup B) \cap A) + m((A \cup B) \setminus A) = m(A) + m(B)$. 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 $m(A) \leq \sum_{i=1}^{\infty} m(A_i)$, mentre la monotonia assicura che per ogni $N \in \mathbf{N}$

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

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

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

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

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

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

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

e per (iv) si ha

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

da cui $\lim_{i \rightarrow +\infty} m(A_i) \leq m(A)$. 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 $A = \bigcup_{i=1}^{\infty} A_i$, 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 $B_N := \bigcup_{i=1}^N A_i$: essi sono misurabili anche per la misura esterna \tilde{m} data dalla

restrizione di m all'insieme T (cioè la misura definita da $\tilde{m}(A) := m(T \cap A)$ per ogni $A \subset \mathbf{R}^n$). Per la monotonia della misura abbiamo:

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

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

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

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

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

Q.E.D.

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 $m(B \setminus C) < \varepsilon$.

Per dimostrarlo ci servirà il seguente fatterello topologico: qualunque aperto di \mathbf{R}^n , comunque complicato, può 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 $(q_1 - r, q_1 + r) \times (q_2 - r, q_2 + r) \times \dots \times (q_n - r, q_n + r)$, 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.

Dimostriamo il teorema di regolarità della misura di Lebesgue.

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 $\bigcup_{i=1}^{\infty} I_i \supset A$ e $\sum_{i=1}^{\infty} |I_i| \leq m(A) + \varepsilon/2$. Abbiamo già

visto che non è restrittivo supporre che gli I_i siano tutti aperti. Se $B = \bigcup_{i=1}^{\infty} I_i$, 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 $B = \bigcup_{N=1}^{\infty} B_N$.

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

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.

4 Lezione del 16/12/2013 (3 ore)

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 equivalenza³. Mostriamo 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 $B_q = A_q \setminus (0, 1)$, 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

$$1 = m([0, 1)) = \sum_{q \in [0, 1) \cap \mathbf{Q}} m(\tilde{A}_q) = \sum_{q \in [0, 1) \cap \mathbf{Q}} m(A).$$

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 nessun caso essa 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*.

³Per poter definire questo insieme, dobbiamo assumere la validità dell'assioma della scelta!

DEFINIZIONE (funzione misurabile): Sia $A \subset \mathbf{R}^n$ misurabile, $f : A \rightarrow \overline{\mathbf{R}}$. 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

PROPOSIZIONE (Caratterizzazione delle funzioni misurabili): Una funzione $f : A_0 \rightarrow \overline{\mathbf{R}}$ (con $A_0 \subset \mathbf{R}^n$ misurabile) 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 e dimostriamo che la controimmagine di un aperto è sempre 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 “destra”. Ne segue che $f^{-1}(\{-\infty\})$ è misurabile: $f^{-1}(\{-\infty\}) = \bigcap_{N=1}^{\infty} f^{-1}((-\infty, -N])$...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 $U = \bigcup_{i=1}^{\infty} I_i$, con $I_i \subset \mathbf{R}$ intervalli aperti. Allora $f^{-1}(U) = \bigcup_{i=1}^{\infty} f^{-1}(I_i)$ è misurabile. Q.E.D.

Osserviamo che 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 $\phi : \overline{\mathbf{R}} \rightarrow \overline{\mathbf{R}}$ 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

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

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) $\{x \in \mathbf{R}^n : f(x) = +\infty, g(x) > 0\}$, vale 0 sull'insieme $\{x \in \mathbf{R}^n : f(x) = +\infty, g(x) = 0\}$, etc. In conclusione, se ne deduce facilmente che la funzione prodotto è misurabile.

Per quanto riguarda la (iv), sia $f(x) = \sup\{f_n(x) : n = 1, 2, \dots\}$. 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 $\liminf_{n \rightarrow +\infty} f_n(x)$ è misurabile in quanto $\liminf_{n \rightarrow +\infty} f_n(x) = \sup_n \inf\{f_m(x) : m \geq n\}$. Analogamente, $\limsup_{n \rightarrow +\infty} f_n(x)$ 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 $A \subset \mathbf{R}^n$, la sua *funzione caratteristica* è la funzione

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

DEFINIZIONE: Una *funzione semplice* $\phi : \mathbf{R}^n \rightarrow \mathbf{R}$ 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, \dots, A_N e dei numeri reali c_1, c_2, \dots, 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 m come

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

Osserviamo che una funzione a scala è una funzione semplice in cui gli insiemi A_i sono intervalli. Per la misura di Lebesgue e per questo tipo di funzioni, la nuova definizione di integrale coincide con quella di Riemann. Inoltre, non è difficile vedere che l'integrale sulle funzioni semplici gode delle usuali proprietà di monotonia, di additività e di omogeneità rispetto alla funzione integranda.

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: $\int f(x) dx = \sup\{\int \phi(x) dx : \phi \text{ semplice}, \phi \leq f\}$.

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 $f : \mathbf{R}^n \rightarrow [0, +\infty]$ una funzione misurabile. Allora esiste una successione $\phi_k : \mathbf{R}^n \rightarrow [0, +\infty)$ di funzioni semplici tali che $f \geq \phi_{k+1} \geq \phi_k$ ($k = 1, 2, 3, \dots$) e tali che

$$\lim_{k \rightarrow +\infty} \phi_k(x) = f(x) \quad \forall x \in \mathbf{R}^n.$$

DIM.: Per ogni fissato $k = 1, 2, \dots$ e $j = 0, 2, \dots, k2^k - 1$ 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 semplici⁴

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

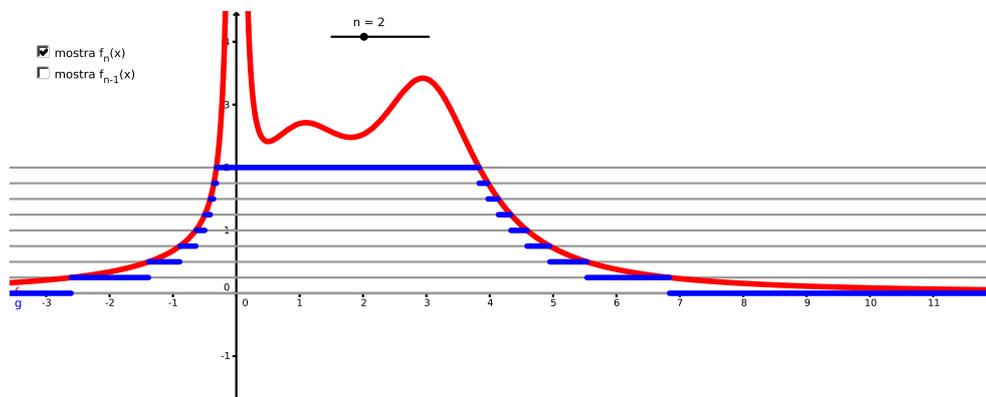
Per costruzione, queste funzioni sono misurabili e sono tutte minori o uguali a f . Inoltre, esse formano una successione crescente: basta osservare che per ogni k e per ogni $j = 1, \dots, k2^k - 1$ si ha $E_{k,j} = E_{k+1,2j} \cup E_{k+1,2j+1}$. Quanto poi a $E_{k,k2^k}$, questo verrà suddiviso al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

È poi facile vedere che $\phi_k(x) \rightarrow f(x)$ 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 $\phi_k(x) = k \rightarrow +\infty$.

Ecco un tentativo di visualizzare la costruzione delle funzioni ϕ_k con un foglio GeoGebra⁵.

⁴Possiamo esprimere queste funzioni anche nel seguente modo più compatto: $\phi_k(x) = \min\{k, 2^{-k} \lfloor 2^k f(x) \rfloor\}$, dove $\lfloor \cdot \rfloor$ denota la parte intera.

⁵http://profs.sci.univr.it/~baldo/aa2010/Funzioni_semplici.html



Q.E.D.

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

DEFINIZIONE: L'integrale di Lebesgue di una funzione misurabile $f : \mathbf{R}^n \rightarrow [0, +\infty]$ si definisce come

$$\int_{\mathbf{R}^n} f(x) dx = \sup \left\{ \int_{\mathbf{R}^n} \phi(x) dx : \phi \text{ semplice, } \phi \leq f \right\}.$$

Se poi $f : A \rightarrow [0, +\infty]$ è misurabile, definiamo $\int_A f(x) dx$ come $\int_{\mathbf{R}^n} \tilde{f}(x) dx$, dove $\tilde{f} : \mathbf{R}^n \rightarrow [0, +\infty]$ 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, $f_k : \mathbf{R}^n \rightarrow [0, +\infty]$, e supponiamo che la successione sia anche crescente: $f_{k+1}(x) \geq f_k(x)$ per ogni $x \in \mathbf{R}^n$ e per ogni $k = 1, 2, 3, \dots$. Allora, se $f(x) = \lim_{k \rightarrow +\infty} f_k(x)$, si ha

$$\int_{\mathbf{R}^n} f(x) dx = \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} f_k(x) dx.$$

5 Lezione del 18/12/2013 (2 ore)

Prima della dimostrazione, vediamo una importante conseguenza del teorema di Beppo Levi:

OSSERVAZIONE (Additività dell'integrale rispetto alla funzione integranda): Siano $f, g : \mathbf{R}^n \rightarrow [0, +\infty]$ funzioni misurabili. Allora

$$\int_{\mathbf{R}^n} (f(x) + g(x)) dx = \int_{\mathbf{R}^n} f(x) dx + \int_{\mathbf{R}^n} g(x) dx.$$

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

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

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

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

$$\int_A f(x) dx := \int_A f^+(x) dx - \int_A f^-(x) dx.$$

Se poi i due integrali della parte positiva e della parte negativa 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.

Grazie all'additività dell'integrale e al fatto evidente che le costanti possono essere portate fuori dal segno di integrale, l'integrale di Lebesgue è *lineare* sullo spazio vettoriale delle funzioni sommabili.

Dimostrazione del teorema di Beppo Levi: Notiamo innanzitutto che la funzione f è misurabile in quanto limite (sup) di funzioni misurabili. Inoltre, la successione $k \mapsto \int_{\mathbf{R}^n} f_k(x) dx$ è crescente: indichiamo con α il suo limite.

Evidentemente, essendo $f \geq f_k$ per ogni k , si ha $\int_{\mathbf{R}^n} f(x) dx \geq \alpha$: in particolare, se $\alpha = +\infty$, il teorema è dimostrato. Se invece $\alpha \in \mathbf{R}$, ci rimane da

dimostrare la disuguaglianza opposta

$$\int_{\mathbf{R}^n} f(x) dx \leq \alpha.$$

A tal fine, fissiamo $c \in (0, 1)$ e una funzione semplice $s : \mathbf{R}^n \rightarrow [0, +\infty)$ 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 $E_k = \{x \in \mathbf{R}^n : f_k(x) \geq cs(x)\}$. Grazie al fatto che le f_k tendono a f e che $c < 1$, abbiamo che $\bigcup_{k=1}^{\infty} E_k = \mathbf{R}^n$, e inoltre la successione di insiemi misurabili E_k è crescente perché lo è $\{f_k\}$. Definiamo poi $A_{j,k} = A_j \cap E_k$: grazie alla continuità della misura sulle successioni crescenti abbiamo $m(A_{j,k}) \rightarrow m(A_j)$ per $k \rightarrow +\infty$. Allora:

$$\begin{aligned} \alpha &= \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} f_k(x) dx \geq \\ &\lim_{k \rightarrow +\infty} \int_{E_k} f_k(x) dx \geq \lim_{k \rightarrow +\infty} \int_{E_k} c s(x) dx = \\ &\lim_{k \rightarrow +\infty} c \sum_{j=1}^N s_j m(A_{j,k}) = c \sum_{j=1}^N s_j m(A_j) = c \int_{\mathbf{R}^n} s(x) dx. \end{aligned}$$

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 un'interessante conseguenza del teorema di Beppo Levi:
ESEMPIO (Integrazione per serie): Se $\{f_k\}$ è una successione di funzioni misurabili non negative definite su A , allora

$$\int_A \sum_{i=1}^{\infty} f_k(x) dx = \sum_{i=1}^{\infty} \int_A f_k(x) dx.$$

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

Si noti che la proprietà di additività rispetto alla funzione integranda continua 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 l'enunciato di un altro celebre risultato: il Lemma di Fatou!

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

$$\int_{\mathbf{R}^n} f(x) dx \leq \liminf_{k \rightarrow +\infty} \int_{\mathbf{R}^n} f_k(x) dx.$$

DIM.: Sappiamo già che f è misurabile non negativa. Si ha $f(x) = \lim_{k \rightarrow +\infty} g_k(x)$, 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

$$\int_{\mathbf{R}^n} f(x) dx = \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} g_k(x) dx.$$

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.

Un paio di osservazioni: il lemma di Fatou è in generale falso per funzioni di segno qualunque. Si prenda per esempio $n = 1$, $f_k(x) = -1/k$ (funzioni costanti). Allora $f_k(x) \rightarrow 0$, ma

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

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:

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

$$\begin{aligned} \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} |f_k(x) - f(x)| dx &= 0, \\ \int_{\mathbf{R}^n} f(x) dx &= \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} f_k(x) dx. \end{aligned}$$

DIM.: 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

$$\liminf_{k \rightarrow +\infty} \int_{\mathbf{R}^n} (2\phi(x) - |f_k(x) - f(x)|) dx \geq \int_{\mathbf{R}^n} 2\phi(x) dx,$$

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

$$\limsup_{k \rightarrow +\infty} \int_{\mathbf{R}^n} |f_k(x) - f(x)| dx \leq 0,$$

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

$$\left| \int_{\mathbf{R}^n} f_k(x) dx - \int_{\mathbf{R}^n} f(x) dx \right| \leq \int_{\mathbf{R}^n} |f_k(x) - f(x)| dx.$$

Q.E.D.

6 Lezione del 19/12/2013 (1 ora)

Vediamo subito qualche altra conseguenza interessante dei teoremi che abbiamo dimostrato.

ESEMPIO (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 $A = \bigcup_{i=1}^{\infty} A_i$, allora

$$\int_A f(x) dx = \sum_{i=1}^{\infty} \int_{A_i} f(x) dx.$$

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

La convergenza puntuale nei teoremi di convergenza integrale non è necessaria in *tutti i punti*: non c'è niente di male se essa viene a mancare in un insieme di misura nulla. A questo proposito 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 di Lebesgue nulla.

Per esempio, date due funzioni $f, g : \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$, diremo che esse sono *quasi ovunque uguali* se $m(\{x : f(x) \neq g(x)\}) = 0$.

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. Di più, nei teoremi di Beppo Levi, Fatou e Lebesgue, basta avere la convergenza *quasi ovunque* delle funzioni coinvolte (e la funzione dominante ϕ nel teorema di Lebesgue basta che domini le f_k quasi ovunque).

Vediamo per esercizio una semplice conseguenza del teorema della convergenza dominata:

Un risultato di integrazione per serie: se $f_k : \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$ sono sommabili e $\sum_{k=1}^{\infty} \int_{\mathbf{R}^n} |f_k(x)| < +\infty$, allora la serie $\sum_{k=1}^{\infty} f_k(x)$ converge puntualmente ad una funzione sommabile $f(x)$ e

$$\int_{\mathbf{R}^n} f(x) dx = \sum_{k=1}^{\infty} \int_{\mathbf{R}^n} f_k(x) dx.$$

Sia infatti $g(x) = \sum_{k=1}^{\infty} |f_k(x)|$: grazie al teorema di integrazione per serie che abbiamo già visto (come conseguenza del teorema di Beppo Levi), questa funzione è sommabile. Ne segue subito che $g(x) < +\infty$ per quasi ogni x : dunque la serie $\sum_{k=1}^{\infty} f_k(x)$ converge assolutamente per quasi ogni x ad una funzione che battezziamo $f(x)$.

Per concludere, basta applicare il teorema della convergenza dominata alle somme parziali della serie: esse sono dominate dalla funzione sommabile $g(x)$.

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 complicata⁶).

⁶Si 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 Beppo 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 $f : [a, b] \rightarrow \mathbf{R}$ 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 $f : [a, b] \rightarrow \mathbf{R}$, 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

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

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

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

da cui passando al limite

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

e analogamente

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

Siccome $\bar{\psi} \geq \underline{\phi}$, se ne deduce che $\int_a^b (\bar{\psi} - \underline{\phi}) dx = 0$. Ora, è semplice mostrare che una funzione non negativa ha integrale nullo se e soltanto se è quasi ovunque nulla (vedremo domani la dimostrazione). Quindi $\bar{\psi} - \underline{\phi} = 0$ quasi ovunque, ossia $\bar{\psi} = \underline{\phi} = f$ 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.

7 Lezione del 20/12/2013 (2 ore)

Mostriamo che una funzione non negativa con integrale 0 è nulla quasi ovunque:

PROPOSIZIONE: Se $f : A \rightarrow [0, +\infty]$ è misurabile e $\int_A f(x) dx = 0$, 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 $m(E_{\bar{n}}) > 0$, con $E_{\bar{n}} = \{x \in A : f(x) > \frac{1}{\bar{n}}\}$ avremmo

$$\int_A f(x) dx \geq \int_{E_{\bar{n}}} f(x) dx \geq m(E_{\bar{n}})/\bar{n} > 0,$$

contro l'ipotesi. Q.E.D.

Continuamo con un simpatico esercizio:

Assoluta continuità dell'integrale di Lebesgue: se $f : \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$ è sommabile, allora l'integrale non si può "concentrare su insiemi piccoli": per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni insieme misurabile $A \subset \mathbf{R}^n$ con $m(A) < \delta$ si ha

$$\int_A |f(x)| dx < \varepsilon.$$

Dimostriamolo: se per assurdo non fosse vero, troveremmo $\varepsilon_0 > 0$ e per ogni $n \in \mathbf{N}$ un insieme misurabile A_n con $m(A_n) < 2^{-n}$ e $\int_{A_n} |f(x)| dx > \varepsilon_0$.

Poniamo poi $B_n = \bigcup_{k=n}^{\infty} A_k$: questa è una successione decrescente di insiemi misurabili, con $m(B_n) \leq 2^{-n+1}$ (numerabile subadditività della misura): in particolare, l'intersezione dei B_n ha misura nulla. Ma allora la successione di funzioni $g_k(x) = |f(x)| \cdot \mathbf{1}_{B_k}(x)$ tende puntualmente a 0, a meno di un insieme di misura nulla. La successione è poi dominata da $|f(x)|$. Per il teorema della convergenza dominata, possiamo concludere che gli integrali delle g_k tendono a 0. Questo è assurdo perché

$$\int_{B_k} |f(x)| dx \geq \int_{A_k} |f(x)| dx \geq \varepsilon_0.$$

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 $f : \mathbf{R}^2 \rightarrow \overline{\mathbf{R}}$ 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

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

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).

La dimostrazione è un po' complicata, per cui la ometteremo.

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 semplici⁷. 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 $\Phi : A \rightarrow B$ un diffeomorfismo, con A e B aperti di \mathbf{R}^n , $f : B \rightarrow \overline{\mathbf{R}}$ una funzione integrabile. Allora si ha

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

I due teoremi precedenti sono ovviamente utilissimi nelle applicazioni. . . e hanno anche delle suggestive conseguenze "geometriche": dal teorema di

⁷Si consideri per esempio la funzione $f(x, y) = (x - y)/(x + y)^3$: i suoi integrali iterati sono finiti e diversi sul quadrato $[0, 1] \times [0, 1]$.

Fubini segue subito che l'integrale di Lebesgue di una funzione misurabile $f : \mathbf{R}^n \rightarrow [0, +\infty]$ non è altro che la *misura del sottografico di f* in \mathbf{R}^{n+1} . Il teorema del cambiamento di variabili garantisce che la misura di Lebesgue è *invariante per isometrie*: fino ad ora, sapevamo solo che lo era per traslazioni!

Concludiamo la parte sulla misura di Lebesgue con un teorema di derivazione sotto il segno di integrale:

TEOREMA (di derivazione sotto il segno di integrale): Sia $A \subset \mathbf{R}^n$ un insieme misurabile, $f : A \times (a, b) \rightarrow \overline{\mathbf{R}}$ una funzione tale che $x \mapsto f(x, t)$ è sommabile per ogni $t \in (a, b)$ e $t \mapsto f(x, t)$ è di classe C^1 per quasi ogni $x \in A$. Supponiamo inoltre che esista una funzione sommabile $g : A \rightarrow \overline{\mathbf{R}}$ tale che $|\frac{\partial f}{\partial t}(x, t)| \leq g(x)$ per quasi ogni $x \in A$ e per ogni $t \in (a, b)$. Allora la funzione $F(t) = \int_A f(x, t) dx$ è di classe $C^1((a, b))$ e vale

$$\frac{d}{dt} \int_A f(x, t) dx = \int_A \frac{\partial f}{\partial t}(x, t) dx$$

DIM.: Sia $\bar{t} \in (a, b)$, δ_k una qualunque successione che tende a 0 per $k \rightarrow +\infty$. Consideriamo i rapporti incrementali

$$\frac{F(\bar{t} + \delta_k) - F(\bar{t})}{\delta_k} = \int_A \frac{f(x, \bar{t} + \delta_k) - f(x, \bar{t})}{\delta_k} dx = \int_A \frac{\partial f}{\partial t}(x, \bar{t} + \theta \delta_k) dx,$$

dove nell'ultimo passaggio $\theta = \theta(x, k) \in (0, 1)$ è fornito dal teorema di Lagrange.

Per $k \rightarrow +\infty$, l'integranda nell'ultima espressione tende puntualmente a $\frac{\partial f}{\partial t}(x, \bar{t})$ e la convergenza è dominata da $g(x)$: il teorema di Lebesgue garantisce quindi che

$$\lim_{k \rightarrow +\infty} \frac{F(\bar{t} + \delta_k) - F(\bar{t})}{\delta_k} = \int_A \frac{\partial f}{\partial t}(x, \bar{t}) dx.$$

Grazie all'arbitrarietà della successione $\delta_k \rightarrow 0$, questo dice che $F'(\bar{t}) = \int_A \frac{\partial f}{\partial t}(x, \bar{t}) dx$.

Rimane da dimostrare che $F'(t)$ è una funzione continua: se, come prima, $\delta_k \rightarrow 0$ è un'arbitraria successione infinitesima, abbiamo

$$F'(\bar{t} + \delta_k) = \int_A \frac{\partial f}{\partial t}(x, \bar{t} + \delta_k) dx.$$

L'integranda dell'ultima espressione tende puntualmente quasi ovunque a $\frac{\partial f}{\partial t}(x, \bar{t})$ e la convergenza è dominata da $g(x)$: il teorema di Lebesgue consente allora di concludere che $F'(\bar{t} + \delta_k) \rightarrow F'(\bar{t})$ e la funzione F è di classe C^1 . Q.E.D.

Il teorema ammette un'ovvia generalizzazione, con dimostrazione identica, a funzioni $f(x, t)$ con $x \in A \subset \mathbf{R}^n$, $t \in \Omega \subset \mathbf{R}^m$, A misurabile e Ω aperto: se $x \mapsto f(x, t)$ è sommabile per ogni $t \in \Omega$ e $t \mapsto f(x, t)$ è di classe C^1 per quasi ogni $x \in A$, e se inoltre esistono funzioni sommabili $g_1(x), \dots, g_m(x)$ tali che

$$\left| \frac{\partial f}{\partial t_i}(x, t) \right| \leq g_i(x) \quad \forall t \in \Omega, \text{ q.o. } x \in A, \quad i = 1, \dots, m,$$

allora la funzione $F(t) = \int_A f(x, t) dx$ è di classe $C^1(\Omega)$ e vale

$$\frac{\partial}{\partial t_i} \int_A f(x, t) dx = \int_A \frac{\partial f}{\partial t_i}(x, t) dx.$$

8 Lezione del 8/1/2014 (2 ore)

Nella parte seguente del corso, vogliamo studiare meglio la teoria delle equazioni differenziali ordinarie, già incontrata al primo anno. Vogliamo cominciare con un teorema che assicura, sotto certe condizioni, che la soluzione del problema di Cauchy *esiste localmente* ed è *unica*.

Consideriamo dunque il problema di Cauchy del primo ordine

$$(*) \quad \begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}$$

dove f è una data funzione continua di due variabili definita in un intorno di (t_0, x_0) .

Supponiamo di avere una soluzione $x(t)$ del problema: integriamo ambo i membri dell'equazione tra t_0 e t , tenendo conto della condizione iniziale. Otteniamo allora l'identità

$$(**) \quad x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds,$$

valida per tutti i t in cui è definita la soluzione.

Vale anche il viceversa:

PROPOSIZIONE: Sia $x \in \mathcal{C}^0([t_0 - \delta, t_0 + \delta])$ una funzione per cui vale l'identità integrale (**). Allora $x(t)$ è derivabile, ed è una soluzione del problema di Cauchy (*) nell'intervallo $[t_0 - \delta, t_0 + \delta]$.

DIM.: Sostituendo $t = t_0$ in (**) si ottiene la condizione iniziale $x(t_0) = x_0$. Inoltre, a secondo membro abbiamo l'integrale tra t_0 e t di una funzione continua: per il Teorema fondamentale del calcolo integrale, la derivata di questa quantità è $f(t, x(t))$, e quindi $x(t)$ è derivabile e soddisfa l'equazione differenziale (*). Q.E.D.

Siamo finalmente in grado di dimostrare un Teorema di esistenza e unicità locale per il nostro problema di Cauchy:

TEOREMA: Sia $f : [t_0 - a, t_0 + a] \times [x_0 - R, x_0 + R] \rightarrow \mathbf{R}$ una funzione continua soddisfacente alla seguente condizione di Lipschitz: esiste $L > 0$ tale che

$$|f(t, x_1) - f(t, x_2)| \leq L|x_1 - x_2| \quad \forall t \in [t_0 - a, t_0 + a], \forall x_1, x_2 \in [x_0 - R, x_0 + R].$$

Allora è possibile trovare un $\delta \in (0, a]$, tale che nell'intervallo $[t_0 - \delta, t_0 + \delta]$ esiste una soluzione del problema di Cauchy (*). Inoltre tale soluzione è unica, nel senso che due soluzioni di (*) definite in uno stesso intorno di t_0 coincidono.

DIM.: Grazie alla Proposizione, basta in realtà cercare soluzioni continue dell'equazione integrale (**).

Per il teorema di Weierstrass, la funzione continua $|f(x, t)|$ ammette massimo sul rettangolo $[t_0 - a, t_0 + a] \times [x_0 - R, x_0 + R]$: chiamiamo M tale massimo. Scegliamo poi

$$\delta = \min\left\{a, \frac{1}{2L}, \frac{R}{M}\right\}.$$

Definiamo un'applicazione T che ad ogni elemento dell'insieme di funzioni

$$X = \left\{ x \in \mathcal{C}^0([t_0 - \delta, t_0 + \delta]) : x(t) \in [x_0 - R, x_0 + R] \forall t \in [t_0 - \delta, t_0 + \delta] \right\}$$

ne associa un'altra che chiamiamo $T(x)$:

$$(T(x))(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds.$$

Chiaramente, per ogni $x \in X$ la funzione $T(x)$ è continua (ed anzi derivabile). Verifichiamo che, in realtà, $T : X \rightarrow X$ ed è una contrazione.

Intanto, se $t \in [t_0 - \delta, t_0 + \delta]$ abbiamo

$$|(T(x))(t) - x_0| \leq \left| \int_{t_0}^t |f(s, x(s))| ds \right| \leq M \cdot \delta \leq R$$

grazie alla condizione di Lipschitz. Dunque $(T(x))(t) \in [x_0 - R, x_0 + R]$, e effettivamente $T : X \rightarrow X$.

Inoltre, se $x_1(t), x_2(t) \in X$:

$$\begin{aligned} |(T(x_1))(t) - (T(x_2))(t)| &\leq \left| \int_{t_0}^t |f(s, x_1(s)) - f(s, x_2(s))| ds \right| \leq \\ &\left| \int_{t_0}^t L|x_1(s) - x_2(s)| ds \right| \leq L \cdot \delta \|x_1 - x_2\|_\infty \leq \frac{1}{2} \|x_1 - x_2\|_\infty. \end{aligned}$$

Passando al sup per $t \in [t_0 - \delta, t_0 + \delta]$ abbiamo allora

$$\|T(x_1) - T(x_2)\|_\infty \leq \frac{1}{2} \|x_1 - x_2\|_\infty,$$

e T è una contrazione di X in X .

L'insieme di funzioni X , dotato della metrica uniforme, è evidentemente un sottospazio chiuso di $\mathcal{C}^0([t_0 - \delta, t_0 + \delta])$: X è dunque uno *spazio metrico completo*.

Allora, grazie al teorema delle contrazioni, T ha un unico punto fisso: esso è evidentemente l'unica soluzione dell'equazione integrale (**), e quindi del problema di Cauchy (**). Q.E.D.

La dipendenza di δ dalla costante di Lipschitz L può essere facilmente eliminata: mostriamo che l'(unica) soluzione locale esiste su $[t_0 - \delta, t_0 + \delta]$ con $\delta = \min\{a, R/M\}$.

A questo scopo, dimostriamo il seguente, immediato corollario del teorema delle contrazioni:

PROPOSIZIONE: Sia X uno spazio metrico completo, $T : X \rightarrow X$ una funzione per la quale esiste $k \in \mathbf{N}$, $k \geq 1$ tale che $T^{(k)}$ (iterata k volte di T) è una contrazione. Allora T ha un unico punto fisso in X .

DIM.: Il teorema delle contrazioni ci assicura che $T^{(k)}$ ha un unico punto fisso \bar{x} . Ma allora anche $T(\bar{x})$ è un punto fisso di $T^{(k)}$: $T^{(k)}(T(\bar{x})) = T(T^{(k)}(\bar{x})) = T(\bar{x})$. Ma allora $T(\bar{x}) = \bar{x}$, grazie all'unicità del punto fisso di $T^{(k)}$: anche T possiede un punto fisso! Tale punto fisso è poi unico, perché ogni punto fisso di T è tale anche per $T^{(k)}$. Q.E.D.

⁸ Cioè $T^{(k)} = T \circ T \circ \dots \circ T$ (k volte).

COROLLARIO: Con le notazioni della dimostrazione del Teorema di Cauchy-Lipschitz, la soluzione (unica) del problema di Cauchy esiste anche scegliendo $\delta = \min\{a, R/M\}$.

DIM.: La condizione su δ è sufficiente a garantire che $T : X \rightarrow X$. Se mostriamo che esiste un'iterata di T che è una contrazione, possiamo applicare la proposizione appena dimostrata e concludere che esiste un unico punto fisso.

A questo scopo, proviamo per induzione che per ogni $x, y \in X$ si ha

$$|T^{(k)}(x)(t) - T^{(k)}(y)(t)| \leq \frac{L^k}{k!} |t - t_0|^k \|x - y\|_\infty.$$

Per $k = 1$ abbiamo già provato questa disuguaglianza dimostrando il teorema. Supponiamola vera per k . Allora, usando la definizione dell'operatore T e la lipschitzianità:

$$|T^{(k+1)}(x)(t) - T^{(k+1)}(y)(t)| \leq L \left| \int_{t_0}^t |T^{(k)}(x)(s) - T^{(k)}(y)(s)| ds \right|.$$

Usando allora l'ipotesi induttiva ed integrando si ottiene la tesi con $k + 1$ al posto di k .

Passando al sup:

$$\|T^{(k)}(x) - T^{(k)}(y)\|_\infty \leq \frac{L^k}{k!} \delta^k \|x - y\|_\infty.$$

Siccome $(L\delta)^k/k! \rightarrow 0$, l'iterata è una contrazione per k abbastanza grande. Q.E.D.

OSSERVAZIONE: Si noti che, nel teorema di esistenza e unicità, in generale è lecito sperare soltanto nell'esistenza *locale*: per esempio, il problema di Cauchy

$$\begin{cases} x'(t) = x^2(t), \\ x(0) = 1. \end{cases}$$

ha soluzione unica, che però è definita soltanto su una semiretta.

Analogamente, se rimuoviamo l'ipotesi di lipschitzianità in generale la soluzione non è più unica: il problema

$$\begin{cases} x'(t) = \sqrt{|x(t)|}, \\ x(0) = 0. \end{cases}$$

ammette infinite soluzioni. Vedremo invece che l'*esistenza locale* è garantita nella sola ipotesi di continuità della funzione f .

OSSERVAZIONE (Lipschitzianità delle funzioni regolari): Una condizione sufficiente per avere la lipschitzianità di f nella seconda variabile è la richiesta che sia $f_x \in \mathcal{C}^0([t_0 - a, t_0 + a] \times [x_0 - R, x_0 + R])$. Sia infatti $L = \max\{|f_x(t, x)| : (t, x) \in [t_0 - a, t_0 + a] \times [x_0 - R, x_0 + R]\}$. Dati $t \in [t_0 - a, t_0 + a]$ e $x, y \in [x_0 - R, x_0 + R]$ possiamo scrivere

$$|f(t, y) - f(t, x)| = \left| \int_x^y f_x(t, u) du \right| \leq L|y - x|.$$

Si vede facilmente che la costante L che abbiamo scelto è anche *la più piccola* per cui la disuguaglianza di Lipschitz è vera in tutto il rettangolo.

9 Lezione del 10/1/2014 (2 ore)

OSSERVAZIONE (Esistenza globale per f globalmente lipschitziana su una striscia): La versione migliorata del teorema ci garantisce che se f è definita sulla *striscia* $[t_0 - a, t_0 + a] \times \mathbf{R}$ ed è lipschitziana in x uniformemente rispetto a t , allora la soluzione è *globale*, nel senso che esiste sull'intero intervallo $[t_0 - a, t_0 + a]$. Infatti, la richiesta che $\delta \leq R/M$ è necessaria solo per evitare che $T(x)(t)$ esca dall'intervallo $[x_0 - R, x_0 + R]$: questa richiesta può essere tolta se la funzione è definita sulla striscia infinita.

L'ipotesi di lipschitzianità sull'intera striscia è però piuttosto forte (anche se è soddisfatta nel caso importantissimo delle equazioni - e dei sistemi - lineari): vedremo tra breve come indebolirla per avere un risultato di esistenza globale più generale.

OSSERVAZIONE (Prolungamento di soluzioni, soluzione massimale del problema di Cauchy): Sia data l'equazione differenziale $x'(t) = f(t, x(t))$, con f definita su un aperto $A \subset \mathbf{R}^2$ e soddisfacente le ipotesi del teorema di Cauchy-Lipschitz *localmente*, cioè in un intorno rettangolare di ogni punto di A . Questo garantisce l'esistenza di un'(unica) soluzione locale *per ogni* dato iniziale scelto in A .

Una semplice osservazione è allora che *due soluzioni locali dell'equazione che coincidono in un punto t_0 , devono coincidere nell'intersezione degli intervalli su cui sono definite*: supponiamo per assurdo che ci siano punti a destra di t_0 in cui le soluzioni differiscono, e sia t_* il loro estremo inferiore. Le due soluzioni devono coincidere in t_* (per continuità): il teorema di Cauchy-Lipschitz assicura che allora coincidono anche in un intorno destro di t_* (basta risolvere il problema di Cauchy con punto iniziale t_*), contro la definizione di inf. In modo analogo, le due soluzioni devono coincidere a sinistra di t_0 .

Questa osservazione permette allora di dedurre che tra tutte le soluzioni di un fissato problema di Cauchy, ve n'è una definita su un *intervallo massimale*: questo è semplicemente l'unione dei domini di tutte le soluzioni.

OSSERVAZIONE (Teorema di fuga dai compatti): Nelle stesse ipotesi dell'osservazione precedente, è possibile dimostrare che una soluzione massimale $u_* : I_* \rightarrow \mathbf{R}$ deve necessariamente *uscire da ogni sottinsieme compatto* K di A : possiamo esprimere informalmente questo fatto dicendo che una soluzione massimale *scappa verso il bordo dell'insieme* A su cui è definito il secondo membro dell'equazione, quando t tende agli estremi dell'intervallo in cui tale soluzione esiste.

OSSERVAZIONE (Esistenza e unicità per sistemi di equazioni del primo ordine): Il problema di Cauchy per un sistema di n equazioni differenziali ordinarie del primo ordine (in forma normale) è formalmente identico a quello per un'equazione singola:

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}$$

se assumiamo che la funzione f e l'incognita x siano *vettoriali*:

$$\begin{aligned} f(t, x) &= (f_1(t, x_1, x_2, \dots, x_n), \dots, f_n(t, x_1, \dots, x_n)) : \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n, \\ x(t) &= (x_1(t), \dots, x_n(t)). \end{aligned}$$

Il teorema di Cauchy-Lipschitz vale allora, *con dimostrazione assolutamente identica*, anche per i sistemi!

Inoltre, anche in questo caso la lipschitzianità è implicata dalla continuità delle derivate parziali di f rispetto alle ultime n variabili: la dimostrazione è un semplice adattamento di quella che abbiamo visto nel caso scalare.

Vediamo subito un risultato di esistenza globale: anche in questo caso, lo enunciamo e dimostriamo per un'equazione scalare, ma il risultato vale (con identica dimostrazione) per i sistemi.

TEOREMA (di esistenza globale): Sia $f : [a, b] \times \mathbf{R} \rightarrow \mathbf{R}$ una funzione continua soddisfacente alla seguente condizione di Lipschitz locale: per ogni $R > 0$ esiste $L = L(R) > 0$ tale che

$$|f(t, x_1) - f(t, x_2)| \leq L|x_1 - x_2| \quad \forall t \in [a, b], \quad \forall x_1, x_2 \in [-R, R].$$

Supponiamo inoltre che f abbia crescita lineare nella variabile x : esistono due costanti $A, B > 0$ tali che

$$|f(t, x)| \leq A + B|x| \quad \forall t \in [a, b], \quad \forall x \in \mathbf{R}.$$

Allora esiste un'(unica) soluzione del problema di Cauchy (*) (con $t_0 \in [a, b]$), definita sull'intero intervallo $[a, b]$.

DIM.: Poiché abbiamo già dimostrato l'unicità della soluzione massimale, rimane solo da provare che essa è definita sull'intero intervallo!

Mostreremo solo l'esistenza della soluzione su $[t_0, b]$: procedendo allo stesso modo si prova che la soluzione esiste anche su $[a, t_0]$.

Cominciamo col mostrare che per ogni fissato $\bar{t} \in [a, b]$ e per ogni fissato $\bar{x} \in \mathbf{R}$, la soluzione del problema di Cauchy con dato iniziale $x(\bar{t}) = \bar{x}$ è definita *almeno* sull'intervallo $[\bar{t}, \bar{t} + \delta]$ con

$$(*) \delta = \min\left\{b - \bar{t}, \frac{1}{1 + B}\right\}.$$

Sappiamo infatti, dalla stima "migliorata" di δ nel teorema di Cauchy-Lipschitz, che se f soddisfa la condizione di Lipschitz su $[\bar{t}, b] \times [\bar{x} - R, \bar{x} + R]$, allora la soluzione esiste su $[\bar{t}, \bar{t} + \delta]$ con $\delta = \min\{b - \bar{t}, R/M\}$ ove

$$M = \max\{|f(t, x)| : (t, x) \in [\bar{t}, b] \times [\bar{x} - R, \bar{x} + R]\}.$$

Per ottenere (*) basta scegliere $R = A + B|\bar{x}|$: in tal caso, per ogni $(t, x) \in [\bar{t}, b] \times [\bar{x} - R, \bar{x} + R]$ abbiamo

$$|f(t, x)| \leq A + B|x| \leq A + B|\bar{x}| + BR = (1 + B)R,$$

da cui $M \leq (1 + B)R$ e $R/M \geq 1/(1 + B)$.

In soldoni abbiamo provato che risolvendo il problema di Cauchy *con un punto iniziale arbitrario*, la soluzione o esiste fino a b , oppure è definita su un intervallino di ampiezza almeno $1/(1 + B)$ a destra del punto iniziale.

Partendo da t_0 , la soluzione sarà quindi definita almeno su $[t_0, t_0 + 1/(1 + B)]$. A questo punto, possiamo prendere il *valore finale* di tale soluzione nel punto $t_0 + 1/(1 + B)$ come dato iniziale di un nuovo problema di Cauchy...e potremo prolungare la soluzione fino a $t_0 + 2/(1 + B)$. Procedendo in questo modo, in un numero finito di passi avremo prolungato la soluzione sino al punto b , estremo finale del nostro intervallo. Q.E.D.

La condizione di Lipschitz serve a garantire l'unicità delle soluzioni: essa non è invece necessaria a dimostrare l'esistenza locale. Vale infatti un teorema, dovuto a Peano, che assicura l'esistenza di una soluzione nella sola ipotesi di continuità della funzione $f(t, x)$:

⁹Siccome vogliamo la soluzione soltanto a destra di \bar{t} , le richieste su f sono necessarie solo a destra di tale punto!

TEOREMA (Peano): Sia $f : [t_0 - a, t_0 + a] \times [x_0 - R, x_0 + R] \rightarrow \mathbf{R}$ una funzione continua. E' possibile trovare un $\delta \in (0, a]$, tale che nell'intervallo $[t_0 - \delta, t_0 + \delta]$ esiste una soluzione del problema di Cauchy

$$(P) \quad \begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}$$

Tale soluzione non è necessariamente unica.

Omettiamo la dimostrazione di questo risultato.

Passiamo ora a studiare in modo un po' sistematico le soluzioni di un'equazione lineare di ordine n . Per prima cosa, definiamo un operatore differenziale lineare di ordine n : si tratta di un'applicazione lineare $L : \mathcal{C}^n([a, b]) \rightarrow \mathcal{C}^0([a, b])$ del tipo

$$Lx(t) = D^n x(t) + a_{n-1}(t)D^{n-1}x(t) + a_{n-2}D^{n-2}x(t) + \dots + a_1 Dx(t) + a_0 x(t),$$

ove $D^k x(t)$ indica la derivata k -esima della funzione $x(t)$, mentre $a_0, a_1, \dots, a_{n-1} \in \mathcal{C}^0([a, b])$ sono fissate funzioni continue.

Un'equazione differenziale lineare di ordine n associata a questo operatore è del tipo $Lx(t) = f(t)$, con $f \in \mathcal{C}^0([a, b])$. Il problema di Cauchy relativo consiste nel trovare la soluzione dell'equazione che soddisfa un' n -upla di condizioni iniziali del tipo $x(t_0) = x_0, x'(t_0) = x_1, \dots, x^{(n-1)}(t_0) = x_{n-1}$, con $t_0 \in [a, b]$ e x_0, x_1, \dots, x_{n-1} fissati numeri reali.

Osserviamo per prima cosa che questo problema di Cauchy ha un'unica soluzione globale definita sull'intero intervallo $[a, b]$: a questo scopo, basta osservare che tale problema di Cauchy, indichiamolo con (P), è equivalente al seguente, che coinvolge un sistema di n equazioni del primo ordine:

$$(P') \quad \begin{cases} x'_{n-1}(t) = -a_{n-1}(t)x_{n-1}(t) - a_{n-2}x_{n-2}(t) + \dots - a_1 x_1(t) - a_0 x_0(t) + f(t) \\ x'_{n-2}(t) = x_{n-1}(t) \\ x'_{n-3}(t) = x_{n-2}(t) \\ \dots \\ x'_0(t) = x_1(t) \\ x_{n-1}(t_0) = x_{n-1} \\ x_{n-2}(t_0) = x_{n-2} \\ \dots \\ x_1(t_0) = x_1 \\ x_0(t_0) = x_0 \end{cases}$$

Infatti, la funzione $x_0(t)$ data dal problema (P') è evidentemente una soluzione di (P). Viceversa, una soluzione $x(t)$ del problema (P) assieme alle sue derivate fino alla $(n-1)$ -esima forma l' n -upla soluzione del sistema (P').

Evidentemente, al sistema (P') è applicabile il teorema di esistenza e unicità globale: per ogni n -upla di dati iniziali esiste una ed una sola soluzione definita sull'intero intervallo $[a, b]$. Lo stesso vale dunque per il problema di Cauchy originale (P) per l'equazione di ordine n ¹⁰.

Stabilito che vale il risultato di esistenza e unicità, passiamo a studiare la *struttura dell'insieme delle soluzioni* dell'equazione lineare $Lx(t) = f(t)$. Per farlo, partiamo dal caso dell'equazione *omogenea* $Lx(t) = 0$:

TEOREMA (Spazio delle soluzioni di un'equazione lineare omogenea di ordine n): Dato l'operatore lineare L di cui sopra, l'insieme delle soluzioni dell'equazione omogenea $Lx(t) = 0$

$$\mathcal{V} = \{x(t) : Lx(t) = 0\}$$

è uno spazio vettoriale reale di dimensione n . In particolare, n soluzioni linearmente indipendenti dell'equazione omogenea generano l'intero insieme \mathcal{V} .

DIM.: È immediato verificare che una combinazione lineare di soluzioni dell'equazione omogenea è ancora una soluzione (è una conseguenza della linearità della derivata): l'insieme \mathcal{V} delle soluzioni è dunque uno spazio vettoriale reale.

Fissiamo allora $t_0 \in [a, b]$ e consideriamo l'applicazione

$$\begin{aligned} \Psi : \mathcal{V} &\rightarrow \mathbf{R}^n \\ x(t) &\mapsto (x(t_0), x'(t_0), x''(t_0), \dots, x^{(n-1)}(t_0)) \end{aligned}$$

che ad una soluzione $x(t)$ associa il vettore dei suoi *dati iniziali di Cauchy*.

L'applicazione Ψ è evidentemente lineare: dico che si tratta di un isomorfismo di spazi vettoriali.

La suriettività di Ψ viene dal teorema di esistenza globale: qualunque dato di Cauchy si scelga, esiste una soluzione.

L'iniettività è invece conseguenza dell'unicità: $x(t) \in \ker \Psi$ se e soltanto se il vettore dei dati iniziali è nullo. Ora, la funzione identicamente nulla è chiaramente una soluzione del problema di Cauchy per l'equazione omogenea con dati nulli: grazie al risultato di unicità non ve ne sono altre. In altre parole, $\ker \Psi = \{0\}$ e Ψ è un isomorfismo. Q.E.D.

¹⁰Un analogo risultato di esistenza e unicità locale vale anche per il problema di Cauchy relativo ad un'equazione *non lineare* di ordine n : anche questa è equivalente ad un sistema di n equazioni del primo ordine, ottenuto con il trucco sopra.

10 Lezione del 13/1/2014 (2 ore)

Vediamo subito come è fatto l'insieme delle soluzioni di un'equazione lineare di ordine n completa:

*COROLLARIO (Soluzioni di un'equazione lineare di ordine n completa):
L'insieme delle soluzioni di un'equazione lineare di ordine n*

$$\mathcal{W} = \{x(t) : Lx(t) = f(t)\}$$

è uno spazio affine modellato sullo spazio vettoriale \mathcal{V} delle soluzioni dell'equazione omogenea associata. Precisamente, se $x_p(t)$ è una qualunque soluzione particolare dell'equazione, allora

$$\mathcal{W} = \{x_p(t) + x_0(t) : Lx_0(t) = 0\}.$$

DIM.: Siano $x_1(t), x_2(t)$ due soluzioni dell'equazione completa $Lx(t) = f(t)$. È immediato verificare che $x_1(t) - x_2(t)$ è una soluzione dell'equazione omogenea $Lx(t) = 0$. Q.E.D.

Il teorema appena visto ci dice che per risolvere un'equazione lineare omogenea di ordine n è sufficiente trovarne n soluzioni *linearmente indipendenti*.

Un criterio per stabilire se n soluzioni sono, appunto, linearmente indipendenti, coinvolge il *determinante wronskiano*:

PROPOSIZIONE (Wronskiano): Siano $x_1(t), \dots, x_n(t)$ soluzioni dell'equazione lineare omogenea di ordine n $Lx(t) = 0$. Queste soluzioni sono linearmente indipendenti se e soltanto se il loro determinante wronskiano

$$W(t) = \det \begin{pmatrix} x_1(t) & x_2(t) & \dots & x_n(t) \\ x_1'(t) & x_2'(t) & \dots & x_n'(t) \\ x_1''(t) & x_2''(t) & \dots & x_n''(t) \\ \dots & \dots & \dots & \dots \\ x_1^{(n-1)}(t) & x_2^{(n-1)}(t) & \dots & x_n^{(n-1)}(t) \end{pmatrix}$$

non si annulla mai nell'intervallo $[a, b]$.

DIM.: Per ogni fissato $t_0 \in [a, b]$, $W(t_0)$ è il determinante di una matrice che ha come colonne i vettori dei dati iniziali dei problemi di Cauchy di cui sono soluzioni $x_1(t), x_2(t), \dots, x_n(t)$. Ricordando l'isomorfismo lineare $\Psi : \mathcal{V} \rightarrow \mathbf{R}^n$ definito nella dimostrazione del teorema si ha facilmente la tesi: $x_1(t), \dots, x_n(t)$ formano una base di \mathcal{V} se e solo se le loro immagini secondo

Ψ (che sono le colonne della matrice) formano una base di \mathbf{R}^n , cioè se e solo se la matrice wronskiana è non singolare. Q.E.D.

Vedremo tra poco come trovare n soluzioni indipendenti di un'equazione lineare a coefficienti costanti. Il caso generale, purtroppo, non è così roseo. . .

In compenso, se siamo in grado di risolvere l'equazione omogenea, c'è anche un metodo per trovare una soluzione particolare di un'equazione completa $Lx(t) = f(t)$:

TEOREMA (Metodo della variazione delle costanti): Data l'equazione lineare di ordine n $Lx(t) = f(t)$, siano $x_1(t), x_2(t), \dots, x_n(t)$ soluzioni linearmente indipendenti dell'equazione omogenea associata. Allora esistono n funzioni di classe $C^1([a, b])$, $C_1(t), C_2(t), \dots, C_n(t)$ che soddisfano il sistema

$$(*) \begin{cases} C_1'(t)x_1(t) + C_2'(t)x_2(t) + \dots + C_n'(t)x_n(t) = 0 \\ C_1'(t)x_1'(t) + C_2'(t)x_2'(t) + \dots + C_n'(t)x_n'(t) = 0 \\ C_1'(t)x_1''(t) + C_2'(t)x_2''(t) + \dots + C_n'(t)x_n''(t) = 0 \\ \dots \\ C_1'(t)x_1^{(n-1)}(t) + C_2'(t)x_2^{(n-1)}(t) + \dots + C_n'(t)x_n^{(n-1)}(t) = f(t) \end{cases}$$

Si ha allora che la funzione $x_p(t) := C_1(t)x_1(t) + C_2(t)x_2(t) + \dots + C_n(t)x_n(t)$ è una soluzione particolare dell'equazione completa.

DIM.: Mostriamo innanzitutto che il sistema (*) è risolubile. Innanzitutto, per ogni fissato t è certamente possibile trovare una n -upla di numeri reali $C_1'(t), \dots, C_n'(t)$ che soddisfano il sistema (che è un onesto sistema *algebrico* di equazioni lineari, di quelli studiati nel corso di algebra lineare!): la matrice incompleta del sistema non è altro che la matrice wronskiana, che sappiamo essere non singolare!

Inoltre, le soluzioni $C_1'(t), \dots, C_n'(t)$ sono funzioni continue di t : per rendersene conto, basta ricordarsi come è fatta l'espressione dell'inversa di una matrice non singolare $n \times n$ (i coefficienti della matrice da invertire sono funzioni continue di t ...ed il determinante, che finisce a denominatore nell'espressione dell'inversa, non si annulla mai per la proposizione precedente). Siccome anche il vettore dei termini noti è continuo in t , l'asserto è provato.

A questo punto, per trovare $C_1(t), C_2(t), \dots, C_n(t)$ basta prendere delle primitive (arbitrarie) delle funzioni che abbiamo trovato risolvendo il sistema.

Non ci resta allora che far vedere che $x_p(t) = C_1(t)x_1(t) + C_2(t)x_2(t) + \dots + C_n(t)x_n(t)$ è soluzione dell'equazione completa.

Calcoliamoci le derivate: usando la prima equazione del sistema si vede subito che $x_p'(t) = C_1(t)x_1'(t) + C_2(t)x_2'(t) + \dots + C_n(t)x_n'(t)$. Analogamente avremo $x_p''(t) = C_1(t)x_1''(t) + C_2(t)x_2''(t) + \dots + C_n(t)x_n''(t)$ e più in generale $x_p^{(k)}(t) = C_1(t)x_1^{(k)}(t) + C_2(t)x_2^{(k)}(t) + \dots + C_n(t)x_n^{(k)}(t)$ per $k = 1, 2, \dots, n-1$.

Invece

$$x_p^{(n)}(t) = C_1(t)x_1^{(n)}(t) + C_2(t)x_2^{(n)}(t) + \dots + C_n(t)x_n^{(n)}(t) + f(t).$$

Sostituendo nell'espressione dell'operatore L si ottiene

$$Lx_p(t) = C_1(t)Lx_1(t) + C_2(t)Lx_2(t) + \dots + C_n(t)Lx_n(t) + f(t) = f(t),$$

dove si è usato il fatto che $x_1(t), x_2(t), \dots, x_n(t)$ sono soluzioni dell'omogenea. Q.E.D.

Vedremo ora come fare a trovare n soluzioni linearmente indipendenti di un'equazione lineare omogenea *a coefficienti costanti*.

Per ragioni che saranno chiare tra un attimo, ci tornerà utile cercare *soluzioni a valori complessi* della nostra equazione differenziale. Una funzione (di una variabile reale) a valori complessi si deriva nel modo ovvio: poniamo per definizione $D(a(t) + ib(t)) := Da(t) + iDb(t)$. A questo punto, ha perfettamente senso cercare soluzioni a valori complessi dell'equazione omogenea $Lx(t) = 0$. Si noti poi che la parte reale e la parte immaginaria di una soluzione complessa saranno soluzioni reali.

Altra cosa che ci servirà sarà saper fare l'esponenziale di un *numero complesso*.

Ricordando lo sviluppo in serie di e^x , è abbastanza naturale dare la definizione seguente: se $z \in \mathbf{C}$, poniamo

$$e^z := \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

Con ragionamenti non difficili si può verificare che la serie esponenziale converge per ogni $z \in \mathbf{C}$.

Si può anche far vedere, ma è un po' più complicato (c'è bisogno del prodotto alla Cauchy di serie...), che $e^{z_1+z_2} = e^{z_1}e^{z_2}$, per cui il nome "esponenziale complesso" è legittimo.

Cosa succede se calcoliamo l'esponenziale complesso di un numero immaginario puro? Usando la definizione e ricordando le serie di Taylor di seno e coseno, si vede subito che vale la *formula di Eulero*

$$e^{i\beta} = \cos \beta + i \sin \beta.$$

In conclusione, avremo dunque

$$e^{\alpha+i\beta} = e^\alpha(\cos \beta + i \sin \beta).$$

[In alternativa, si potrebbe prendere questa formula come definizione, per la verità un po' misteriosa, di esponenziale complesso. Si verifichi per esercizio che vale la proprietà dell'esponenziale richiamata sopra, cioè che si ha

$$e^{(\alpha_1+i\beta_1)+(\alpha_2+i\beta_2)} = e^{\alpha_1+i\beta_1} \cdot e^{\alpha_2+i\beta_2}.]$$

Una cosa che possiamo verificare senza difficoltà è che $D(e^{\lambda t}) = \lambda e^{\lambda t}$ anche per λ complesso. Questo ci suggerisce come trovare soluzioni di equazioni lineari omogenee a coefficienti costanti: sia dato l'operatore

$$Lx(t) := D^n x(t) + a_{n-1}D^{n-1}x(t) + a_{n-2}D^{n-2}x(t) + \dots + a_1Dx(t) + a_0x(t),$$

con a_0, a_1, \dots, a_{n-1} costanti reali fissate.

L'equazione caratteristica associata all'operatore L è, per definizione, l'equazione polinomiale nell'incognita $\lambda \in \mathbf{C}$

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0.$$

Per il teorema fondamentale dell'algebra, tale equazione possiede n radici complesse $\lambda_1, \lambda_2, \dots, \lambda_n$ (non necessariamente tutte distinte). È immediato verificare che le radici caratteristiche permettono di fattorizzare l'operatore L nel prodotto di n operatori lineari a coefficienti costanti del primo ordine:

$$Lx(t) = \left(\prod_{i=1}^n (D - \lambda_i I) \right) x(t),$$

dove I è l'identità. In particolare, si noti che gli operatori del primo ordine che compaiono in questa fattorizzazione *commutano tra loro*.

Usando questa fattorizzazione dell'operatore si vede subito che

PROPOSIZIONE (Soluzioni esponenziali di un'equazione lineare omogenea a coefficienti costanti): Una funzione del tipo $e^{\lambda t}$ è soluzione (a valori complessi) dell'equazione omogenea a coefficienti costanti $Lx(t) = 0$ se e solo se λ è una delle radici caratteristiche dell'operatore L . Inoltre, se $\lambda_1, \lambda_2, \dots, \lambda_k$ sono radici caratteristiche distinte, allora le soluzioni $e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_k t}$ sono linearmente indipendenti.

DIM.: Sia λ_i una delle radici dell'equazione caratteristica. Allora $(D - \lambda_i I)e^{\lambda t} = (\lambda - \lambda_i)e^{\lambda t}$: il risultato è nullo se e solo se $\lambda = \lambda_i$. Usando la fattorizzazione dell'operatore deduciamo allora che

$$Le^{\lambda t} = e^{\lambda t} \prod_{i=1}^n (\lambda - \lambda_i) :$$

$e^{\lambda t}$ è soluzione dell'equazione omogenea se e soltanto se λ è una delle radici dell'equazione caratteristica.

Mostriamo che k soluzioni di questo tipo $e^{\lambda_1 t}, \dots, e^{\lambda_k t}$, con $\lambda_1, \dots, \lambda_k$ tutti distinti, sono linearmente indipendenti¹¹. Supponiamo infatti che valga l'identità

$$C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + \dots + C_k e^{\lambda_k t} = 0.$$

Moltiplicando per $e^{-\lambda_k t}$ e derivando si ottiene

$$C_1(\lambda_1 - \lambda_k)e^{(\lambda_1 - \lambda_k)t} + C_2(\lambda_2 - \lambda_k)e^{(\lambda_2 - \lambda_k)t} + \dots + C_{k-1}(\lambda_{k-1} - \lambda_k)e^{(\lambda_{k-1} - \lambda_k)t} = 0.$$

Moltiplichiamo poi per $e^{-(\lambda_{k-1} - \lambda_k)t}$ e deriviamo nuovamente:

$$C_1(\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1})e^{(\lambda_1 - \lambda_{k-1})t} + \dots + C_{k-2}(\lambda_{k-2} - \lambda_k)(\lambda_{k-2} - \lambda_{k-1})e^{(\lambda_{k-2} - \lambda_{k-1})t} = 0.$$

Ripetendo il giochetto per k volte arriviamo all'espressione

$$C_1(\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1})(\lambda_1 - \lambda_{k-2}) \dots (\lambda_1 - \lambda_2)e^{(\lambda_1 - \lambda_2)t} = 0$$

da cui $C_1 = 0$. In maniera del tutto analoga (basta permutare gli indici...) si dimostra che $C_2 = C_3 = \dots = C_k = 0$: le soluzioni sono indipendenti. Q.E.D.

OSSERVAZIONE: Che fare se ci sono radici complesse coniugate dell'equazione caratteristica.

Grazie alla proposizione appena dimostrata, siamo in grado di trovare n soluzioni indipendenti (a valori complessi) di un'equazione lineare omogenea di ordine n la cui equazione caratteristica abbia radici distinte. Queste si trasformano poi facilmente in n soluzioni *reali* linearmente indipendenti, che come sappiamo generano l'intero spazio delle soluzioni.

Se infatti l'equazione caratteristica ha una radice complessa $\lambda = \alpha + i\beta$, vi è anche la coniugata $\bar{\lambda} = \alpha - i\beta$ (perché l'equazione caratteristica ha coefficienti reali). Ma allora

$$\begin{aligned} \frac{e^{\lambda t} + e^{\bar{\lambda} t}}{2} &= e^{\alpha t} \cos \beta t \\ \frac{e^{\lambda t} - e^{\bar{\lambda} t}}{2i} &= e^{\alpha t} \sin \beta t \end{aligned}$$

sono due soluzioni reali che generano lo stesso sottospazio...

¹¹Questa parte della dimostrazione è stata omessa in classe.

Vediamo come comportarci nel caso che il polinomio caratteristico abbia radici multiple, cioè sia del tipo

$$(\lambda - \lambda_1)^{\mu_1} (\lambda - \lambda_1)^{\mu_2} \dots (\lambda - \lambda_r)^{\mu_r},$$

ove evidentemente si avrà $\mu_1 + \mu_2 + \dots + \mu_r = n$. La fattorizzazione dell'operatore ha ora dei fattori ripetuti:

$$Lx(t) = \left(\prod_{i=1}^r (D - \lambda_i I)^{\mu_i} \right) x(t).$$

Mostriamo che sotto queste condizioni le funzioni $t^k e^{\lambda_i t}$ sono soluzioni dell'equazione per ogni $i = 1, \dots, r$ e per ogni $k = 0, 1, \dots, \mu_i - 1$: si ha infatti $(D - \lambda_i I)(t^k e^{\lambda_i t}) = k t^{k-1} e^{\lambda_i t}$ da cui iterando $(D - \lambda_i I)^k (t^k e^{\lambda_i t}) = k! e^{\lambda_i t}$ e $(D - \lambda_i I)^{k+1} (t^k e^{\lambda_i t}) = 0$.

In questo modo, ci siamo fabbricati esattamente n soluzioni dell'equazione omogenea: se dimostrarimo che sono linearmente indipendenti, abbiamo finito¹²!

Una relazione di dipendenza lineare tra di esse ha la forma $P_1(t)e^{\lambda_1 t} + P_2(t)e^{\lambda_2 t} + \dots + P_r(t)e^{\lambda_r t} = 0$, dove P_1, P_2, \dots, P_r sono polinomi a coefficienti complessi e il grado di P_i è strettamente minore di μ_i . Supponiamo per assurdo che, per esempio, P_1 non sia identicamente nullo. Allora moltiplichiamo per $e^{-\lambda_r t}$ e deriviamo μ_r volte: troviamo un'identità del tipo

$$Q_1(t)e^{(\lambda_1 - \lambda_r)t} + Q_2(t)e^{(\lambda_2 - \lambda_r)t} + \dots + Q_{r-1}(t)e^{(\lambda_{r-1} - \lambda_r)t} = 0,$$

ove i Q_i sono ancora polinomi di grado minore di μ_i , ed il grado di Q_1 è esattamente uguale a quello di P_1 . Ripetendo il trucco, arriviamo all'espressione assurda $R_1(t)e^{(\lambda_1 - \lambda_2)t} = 0$, ove R_1 ha lo stesso grado di P_1 . Le soluzioni trovate sono dunque indipendenti!

11 Lezione del 15/1/2014 (2 ore)

Cominciamo con un esempio di applicazione del risultato della volta scorsa sulle soluzioni di un'equazione lineare omogenea a coefficienti costanti.

ESEMPIO: Si consideri l'equazione omogenea $y^{(11)} + 2y^{(7)} + y''' = 0$. L'equazione caratteristica è $\lambda^{11} + 2\lambda^7 + \lambda^3 = \lambda^3(\lambda^4 + 1)^2 = 0$, che ha come radici

¹²L'indipendenza lineare in classe l'abbiamo data per buona: riporto la dimostrazione per completezza.

$\frac{\sqrt{2}}{2} \pm i\frac{\sqrt{2}}{2}$ e $-\frac{\sqrt{2}}{2} \pm i\frac{\sqrt{2}}{2}$, ciascuna con molteplicità 2, e poi 0 con molteplicità 3. La soluzione generale è quindi

$$y(t) = e^{\frac{\sqrt{2}}{2}t} \left(C_1 \cos \frac{\sqrt{2}}{2}t + C_2 t \cos \frac{\sqrt{2}}{2}t + C_3 \sin \frac{\sqrt{2}}{2}t + C_4 t \sin \frac{\sqrt{2}}{2}t \right) + e^{-\frac{\sqrt{2}}{2}t} \left(C_5 \cos \frac{\sqrt{2}}{2}t + C_6 t \cos \frac{\sqrt{2}}{2}t + C_7 \sin \frac{\sqrt{2}}{2}t + C_8 t \sin \frac{\sqrt{2}}{2}t \right) + C_9 + C_{10}t + C_{11}t^2.$$

OSSERVAZIONE: Equazioni complete con $f(t) = P(t)e^{\lambda t}$, P polinomio di grado qualunque. $\lambda \in \mathbf{C}$.

Ora abbiamo una ricetta generale per risolvere un'equazione lineare omogenea a coefficienti costanti. Se l'equazione è poi completa, il metodo della variazione delle costanti ci permette di trovarne una soluzione particolare: anche in questo caso, in linea di principio, siamo capaci di trovare tutte le soluzioni!

Purtroppo, in pratica, il metodo della variazione delle costanti conduce spesso a calcoli complicati ed antipatici. In molti casi di interesse, esistono modi più veloci per trovare una soluzione particolare di un'equazione lineare non omogenea.

Uno di questi casi è quello in cui *il secondo membro dell'equazione è della forma $f(t) = P(t)e^{\lambda t}$, dove P è un polinomio e λ un numero complesso*. In questo caso, abbiamo a disposizione il cosiddetto *metodo degli annihilatori o di somiglianza*: a parte alcuni casi "degeneri" che vedremo tra un attimo, esiste una soluzione particolare *dello stesso tipo* del secondo membro, ossia del tipo $x_P(t) = Q(t)e^{\lambda t}$, ove Q è un polinomio dello stesso grado di P . Per trovare la soluzione particolare è allora sufficiente sostituire nell'equazione per determinare i coefficienti del polinomio $Q(t)$.

I soli casi in cui non esiste una soluzione di questo tipo sono quelli in cui $f(t)$ è una *soluzione dell'equazione omogenea*, cioè se $Lf(t) = 0$... ma come vedremo anche in questi casi il problema si aggira molto facilmente!

L'osservazione fondamentale per trovare soluzioni particolari di questo tipo è che il secondo membro $f(t)$ è soluzione di un'opportuna equazione lineare omogenea $Mx(t) = 0$: basta prendere un qualunque operatore lineare a coefficienti costanti M di cui λ sia una *radice caratteristica di molteplicità maggiore del grado del polinomio $P(t)$* .

Da questo si ricava che una qualunque soluzione particolare dell'equazione $Lx(t) = f(t)$ è anche soluzione dell'equazione omogenea a coefficienti costanti $(LM)x(t) = 0$. Le radici caratteristiche dell'operatore LM sono semplicemente l'unione di quelle di L e di quelle di M (e la molteplicità di una di esse è la somma delle molteplicità con cui compare in L ed in M).

Ne ricaviamo la seguente semplice ricetta:

Si voglia trovare una soluzione particolare dell'equazione lineare a coefficienti costanti $Lx(t) = P(t)e^{\lambda t}$, con $\lambda \in \mathbf{C}$ e $P(t)$ polinomio di grado n . Allora

- *Se λ non è una radice caratteristica di L , cerchiamo una soluzione del tipo*

$$x_P(t) = Q(t)e^{\lambda t},$$

con Q polinomio di grado n .

- *Se λ è una radice caratteristica di L di molteplicità μ , cerchiamo una soluzione del tipo*

$$x_P(t) = t^\mu Q(t)e^{\lambda t},$$

con Q polinomio di grado n .

Come argomento finale del corso, vogliamo mettere a frutto le serie di Fourier per risolvere alcuni problemi che coinvolgono equazioni alle derivate parziali.

Siccome useremo, appunto, le serie di Fourier, vediamo di ricapitolare (e completare) alcuni dei risultati già visti con Giandomenico Orlandi. Uno spazio naturale in cui studiare le serie di Fourier è lo spazio $L^2(2\pi)$ costituito dalle funzioni $f : \mathbf{R} \rightarrow \mathbf{R}$ che siano 2π -periodiche e *a quadrato sommabile* sul periodo $[-\pi, \pi]$. Tale spazio è dotato del prodotto scalare $\langle f, g \rangle := \int_{-\pi}^{\pi} f(x)g(x) dx$, che induce la norma

$$\|f\|_{L^2} = \left(\int_{-\pi}^{\pi} f^2(x) dx \right)^{1/2}.$$

In realtà, occorre una piccola accortezza nella definizione dello spazio: se $L^2(2\pi)$ fosse costituito dalle *funzioni* periodiche a quadrato sommabile sul periodo, il prodotto scalare e la norma sarebbero degeneri: la norma di una funzione *quasi ovunque nulla* è 0.

Questo problema scompare definendo $L^2(2\pi)$ come l'insieme delle *classi di equivalenza di funzioni quasi ovunque uguali* (che siano, ovviamente, 2π -periodiche e a quadrato sommabile nel periodo): decidiamo cioè di *identificare* due funzioni la cui differenza è quasi ovunque uguale a 0. Con questa definizione, $L^2(2\pi)$ diventa un'onesto spazio vettoriale dotato di prodotto scalare e di norma.

Questo spazio è inoltre *completo*:

TEOREMA (Riesz-Fischer): Lo spazio $L^2(2\pi)$ con la norma indotta dal prodotto scalare L^2 è completo.

Omettiamo la dimostrazione di questo teorema.

Tornando alle serie di Fourier in $L^2(2\pi)$, Giandomenico vi ha fatto vedere che le funzioni

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(nx), \frac{1}{\sqrt{\pi}} \sin(nx) : n = 1, 2, 3, \dots \right\}$$

sono una famiglia ortonormale in $L^2(2\pi)$. Questo ha un'importante conseguenza: se $u \in L^2(2\pi)$, consideriamo la somma parziale N -esima della serie di Fourier di u , cioè

$$S_N(x) = a_0/2 + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx),$$

ove al solito

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} u(x) \cos nx \, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} u(x) \sin nx \, dx$$

sono i coefficienti di Fourier. Ebbene, $S_N(x)$ non è altro che la *proiezione ortogonale* di $u(x)$ sul sottospazio (di dimensione finita) generato dalle funzioni

$$\left\{ 1/\sqrt{2\pi}, \frac{1}{\sqrt{\pi}} \cos(nx), \frac{1}{\sqrt{\pi}} \sin(nx) : n = 1, 2, 3, \dots, N \right\}.$$

In altre parole, è il punto di tale sottospazio *più vicino* ad $u(x)$ (rispetto alla distanza L^2): tra tutti i polinomi trigonometrici di grado N , $S_N(x)$ è quello che meglio approssima $u(x)$.

Esiste poi un teorema (noto come *Teorema di Stone-Weierstrass trigonometrico* e di dimostrazione non facilissima...), che assicura che *ogni funzione* $u \in L^2(2\pi)$ può essere approssimata in norma L^2 con una successione di polinomi trigonometrici (cioè combinazioni lineari finite delle funzioni trigonometriche di cui sopra). Grazie all'ottimalità di $S_N(x)$ tra tutti i polinomi trigonometrici di grado N , questo garantisce che la serie di Fourier converge a u in norma L^2 :

TEOREMA: La serie di Fourier di una funzione $u \in L^2(2\pi)$ converge a u in norma L^2 .

Si noti che, grazie alla disuguaglianza di Bessel (vedi appunti di Giandomenico Orlandi), questo implica in particolare che *i coefficienti di Fourier* a_n, b_n di una funzione L^2 tendono a zero per $n \rightarrow +\infty$.

Che dire della convergenza puntuale ed uniforme delle serie di Fourier?
Un tipico risultato è il seguente:

TEOREMA: Sia $f : \mathbf{R} \rightarrow \mathbf{R}$ una funzione 2π -periodica. Supponiamo che f sia regolare a tratti: assumiamo cioè che essa sia continua tranne eventualmente che in un numero finito di punti nel suo periodo, e che in questi abbia discontinuità di salto con limiti destro e sinistro finiti. Supponiamo inoltre che la funzione sia derivabile, tranne eventualmente che in un numero finito di punti nel suo periodo, che la derivata sia continua nel suo dominio e che nei punti di non derivabilità esistano finiti i limiti destro e sinistro della derivata.

Allora la serie di Fourier di f converge puntualmente a $f(x)$ in ogni punto di continuità x . Se invece x è un punto di salto per f , la serie di Fourier converge alla media aritmetica tra il limite destro e il limite sinistro.

Infine, la convergenza è uniforme in ogni compatto che non intersechi l'insieme dei punti di salto per f .

OSSERVAZIONE: Le ipotesi che abbiamo fatto su f sono piuttosto forti: infatti che per poter scrivere i coefficienti di Fourier di f , è sufficiente che la funzione sia limitata ed integrabile secondo Riemann in $[-\pi, \pi]$. D'altra parte, il teorema comprende un'ipotesi - molto forte - sulla derivata di f ...

In effetti, neanche la continuità di f sarebbe sufficiente a garantire la convergenza: esistono funzioni continue periodiche la cui serie di Fourier non converge in moltissimi punti, anche se è possibile (ma MOLTO difficile) dimostrare che vi è sempre convergenza in quasi tutti i punti (cioè al di fuori da un insieme di misura nulla).

Per contro, le ipotesi che abbiamo dato non sono ottimali e si possono rilassare parecchio.

OSSERVAZIONE: Per ragioni di tempo, non dimostreremo il teorema di convergenza puntuale. Per i nostri discorsi sulle equazioni alle derivate parziali, però, è utile vedere come si dimostra che vi è convergenza uniforme sull'intera retta reale, nel caso particolarmente semplice in cui $f \in C^1$. Diamo per buono il risultato di convergenza puntuale: supponiamo di sapere già che la serie di Fourier di f converge a $f(x)$ in ogni punto $x \in \mathbf{R}$.

La derivata f' è una funzione continua e 2π -periodica: in particolare, essa è a quadrato sommabile sul periodo e vale la disuguaglianza di Bessel. Quest significa che, se denotiamo con α_n, β_n i coefficienti di Fourier di f' , abbiamo

$$\sum_{n=1}^{\infty} (\alpha_n^2 + \beta_n^2) < +\infty.$$

Inoltre, con una integrazione per parti scopriamo che c'è una semplice relazione tra i coefficienti di Fourier di f' ed i coefficienti di Fourier a_n, b_n di f : precisamente, abbiamo $\alpha_n = nb_n, \beta_n = -na_n$ da cui

$$\sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2) < +\infty.$$

Se ne deduce facilmente che la serie di Fourier

$$a_0/2 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

converge *totalmente*: infatti, la serie della norme è maggiorata da

$$\begin{aligned} \sum_{n=1}^{\infty} (|a_n| + |b_n|) &= \sum_{n=1}^{\infty} \frac{1}{n} (n|a_n| + n|b_n|) \leq \\ &\left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{1/2} \left(\left(\sum_{n=1}^{\infty} n^2 a_n^2 \right)^{1/2} + \left(\sum_{n=1}^{\infty} n^2 b_n^2 \right)^{1/2} \right) \end{aligned}$$

e le serie che appaiono nell'ultima espressione sono convergenti.

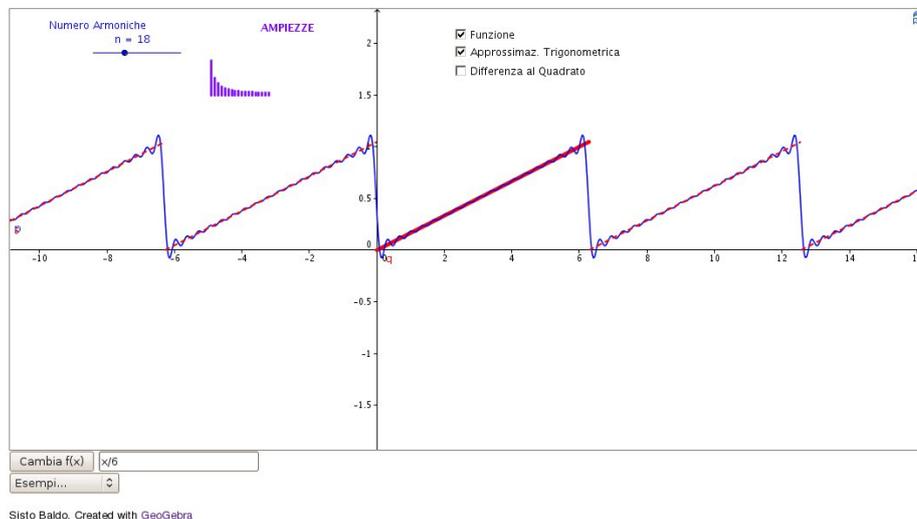
Dunque la serie di Fourier converge uniformemente ad una certa funzione continua, che è ovviamente f perché abbiamo convergenza puntuale.

12 Lezione del 17/1/2014 (2 ore)

Per sperimentare un po' sulla convergenza delle serie di Fourier, potete utilizzare un foglio GeoGebra¹³ che permette di visualizzare le somme parziali della serie di Fourier di una funzione data: ovviamente, i coefficienti di Fourier sono approssimati (con un metodo noto come DFT - discrete Fourier transform)! Il programma interpreta la funzione data dall'utente come la definizione della funzione periodica tra 0 e 2π (nel senso che la estende periodicamente fuori da tale intervallo).

¹³http://profs.sci.univr.it/~baldo/divulgazione/dft/dft_1.html

Serie di Fourier



Torniamo finalmente alle equazioni alle derivate parziali.

Il primo esempio che vogliamo considerare è quello dell'*equazione del calore*: vogliamo trovare una funzione $u(t, x)$ che soddisfi il problema

$$(*) \quad \begin{cases} u_t = ku_{xx} & \text{se } 0 < x < \pi, t > 0 \\ u(0, t) = u(\pi, t) = 0 & \text{se } t > 0 \\ u(x, 0) = f(x) & \text{se } 0 < x < \pi \end{cases}$$

La funzione $u(x, t)$ rappresenta la *temperatura*, nella posizione x ed al tempo t , di una sbarra metallica di lunghezza π che giace sull'asse delle x tra $x = 0$ e $x = \pi$. La seconda equazione ci dice che gli estremi della sbarra sono tenuti a temperatura 0 (valori al contorno), la terza equazione che la temperatura iniziale in posizione x è data da una funzione nota $f(x)$ (condizione iniziale). Infine, la prima equazione è un'equazione alle derivate parziali nota, appunto, come *equazione del calore*: essa descrive l'evoluzione temporale della temperatura in un corpo unidimensionale conduttore di calore.

Per risolvere il nostro *problema ai valori iniziali ed al contorno*, usiamo il *metodo della separazione delle variabili*. Cominciamo a cercare soluzioni non banali dell'equazione differenziale che siano a variabili separate, ossia del tipo $u(x, t) = X(x) \cdot T(t)$ (prodotto di una funzione della sola x e di una della sola t). Visto che abbiamo condizioni al contorno nulle, è naturale imporre anche che $X(0) = X(\pi) = 0$.

Sostituendo nell'equazione differenziale otteniamo $XT' = kX''T$, da cui $\frac{X''}{X} = \frac{T'}{kT}$. In quest'ultima equazione, a primo membro abbiamo una funzione

della sola x , a destra una funzione della sola t : i due membri dell'equazione devono allora essere costanti, ossia deve essere

$$\frac{X''(x)}{X(x)} = \frac{T'(t)}{kT(t)} = -\lambda, \quad \lambda \in \mathbf{R}.$$

In particolare abbiamo $X''(x) + \lambda X(x) = 0$, equazione lineare omogenea che si risolve immediatamente. Se $\lambda < 0$, la soluzione generale è $X(x) = Ae^{\sqrt{-\lambda}x} + Be^{-\sqrt{-\lambda}x}$: imponendo le condizioni al contorno troviamo $A = B = 0$, per cui c'è solo la soluzione nulla. Se $\lambda = 0$ abbiamo $X(x) = A + Bx$ e imponendo le condizioni al contorno troviamo ancora $A = B = 0$.

Infine, se $\lambda > 0$, la soluzione generale dell'equazione è $X(x) = A \cos \sqrt{\lambda}x + B \sin \sqrt{\lambda}x$. La condizione $X(0) = 0$ implica $A = 0$, mentre la seconda condizione diventa $B \sin \sqrt{\lambda}\pi = 0$. Possiamo prendere $B \neq 0$ se e solo se $\sin \sqrt{\lambda}\pi = 0$, cioè se $\lambda = n^2$ con $n = 1, 2, 3, \dots$. In questi casi, troviamo le seguenti soluzioni non nulle:

$$X(x) = B \sin(nx).$$

Le corrispondenti equazioni per $T(t)$ sono $T'(t) = -kn^2T(t)$, che hanno soluzione $T_n(t) = Ce^{-kn^2t}$.

In conclusione, le sole soluzioni non banali a variabili separate sono date da

$$u_n(x, t) = Be^{-kn^2t} \sin nx, \quad n = 1, 2, 3, \dots$$

Vogliamo ora provare a trovare una soluzione del nostro problema in forma di *serie* di queste soluzioni a variabili separate: vogliamo scrivere cioè

$$(**) \quad u(x, t) = \sum_{n=1}^{\infty} b_n e^{-kn^2t} \sin nx,$$

ove ovviamente i coefficienti b_n devono essere scelti in modo da soddisfare la condizione iniziale e *speriamo che la serie converga, e converga proprio ad una soluzione del problema (*)!*

Per trovare i coefficienti b_n , osserviamo che

$$(***) \quad f(x) = u(x, 0) = \sum_{n=1}^{\infty} b_n \sin nx \dots$$

e il secondo membro ha tutta l'aria di una serie di Fourier di una funzione dispari! Estendiamo allora f ad una funzione 2π -periodica e dispari: i suoi coefficienti di Fourier sono dati da

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx.$$

e con questa scelta dei coefficienti la condizione iniziale (***) è soddisfatta purché f sia abbastanza regolare: per esempio, regolare a tratti e continua in modo da avere convergenza uniforme della serie con $t = 0$ al dato iniziale. In effetti, con questa ipotesi la serie di due variabili (**) converge totalmente sulla striscia $[0, \pi] \times [0, +\infty)$: la serie delle norme è maggiorata da $\sum_{k=1}^{\infty} |b_k|$, che è convergente (si riveda la dimostrazione del teorema di convergenza uniforme!). In particolare, la somma $u(x, t)$ della serie è continua sulla striscia e la condizione iniziale è genuinamente soddisfatta!

Si noti, in particolare, che i coefficienti di Fourier b_n tendono a 0 e sono quindi limitati da una certa costante M .

Le condizioni al contorno sono evidentemente soddisfatte: ci manca “solo” da mostrare che la funzione $u(x, t)$ definita dalla serie (**) è regolare per $t > 0$ e soddisfa effettivamente l’equazione del calore!

Consideriamo, sul rettangolo $[0, \pi] \times [\tau, T]$ con $0 < \tau < T$, la convergenza totale delle serie che si ottengono derivando termine a termine la (**) una volta rispetto a t o due volte rispetto a x (vale a dire, le due serie che si ottengono calcolando formalmente u_t e u_{xx}). Le relative serie delle norme sono dominate (a meno di una costante moltiplicativa k nel caso di u_t) dalla serie convergente

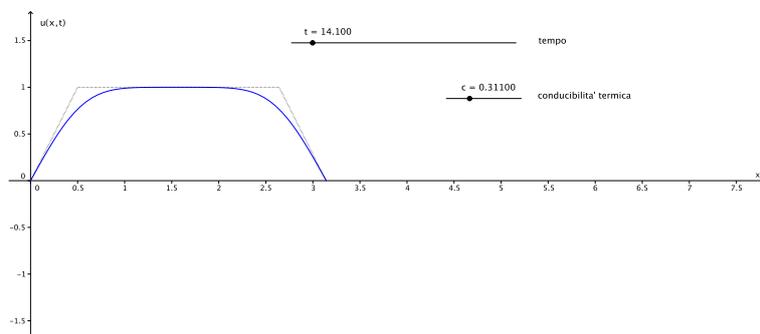
$$M \sum_{n=1}^{\infty} n^2 e^{-kn^2\tau} < +\infty,$$

per cui si ha convergenza totale. La derivazione termine a termine è allora lecita, e (**) fornisce effettivamente una soluzione del nostro problema.

Il ragionamento appena fatto mostra la proprietà *regolarizzante* dell’equazione del calore: anche se il dato iniziale $f(x)$ è non derivabile, o persino discontinuo in qualche punto, la soluzione (**) diventa immediatamente regolare per tempi positivi¹⁴.

Ecco un foglio GeoGebra che mostra un’animazione delle soluzioni dell’equazione del calore:

¹⁴In realtà, con ragionamenti del tutto analoghi a quelli fatti sopra, si vede che le serie delle derivate di qualunque ordine convergono totalmente sul rettangolo $[0, \pi] \times [\tau, T]$: la soluzione è di classe C^∞



Tecniche molto simili si possono usare per risolvere l'equazione della corda vibrante (o equazione delle onde unidimensionale), cioè il problema

$$(V) \quad \begin{cases} u_{tt} = c^2 u_{xx} & \text{se } 0 < x < \pi, t > 0 \\ u(0, t) = u(\pi, t) = 0 & \text{se } t > 0 \\ u(x, 0) = f(x) & \text{se } 0 < x < \pi \\ u_t(x, 0) = 0 & \text{se } 0 < x < \pi \end{cases}$$

Questo problema descrive il movimento di una corda vibrante che a riposo giace sull'asse delle x tra 0 e π : $u(x, t)$ rappresenta lo spostamento verticale (rispetto a 0) del punto della corda di ascissa x , al tempo t . Le condizioni al contorno $u(0, t) = u(\pi, t) = 0$ dicono che gli estremi della corda sono fissati, $f(x)$ rappresenta la *posizione iniziale* della corda, mentre l'ultima condizione dice che la corda viene lasciata andare con velocità iniziale nulla (problema della *corda pizzicata*).

Procediamo ancora per separazione di variabili, cercando soluzioni non banali dell'equazione alle derivate parziali della forma $u(x, t) = X(x)T(t)$: otteniamo le equazioni

$$\frac{X''}{X} = \frac{T''}{c^2 T} = -\lambda.$$

La prima equazione è esattamente quella di prima: abbiamo soluzioni non banali del tipo $X(x) = B \sin nx$ se e solo se $\lambda = n^2$. La corrispondente equazione per T è $T'' + n^2 c^2 T = 0$ con la condizione $T'(0) = 0$ (proveniente dalla condizione su u_t): le soluzioni a variabili separate sono allora $u_n(x, t) = B \sin nx \cos cnt$ e cerchiamo di scrivere la soluzione del problema (V) come serie

$$(VI) \quad u(x, t) = \sum_{n=1}^{\infty} b_n \sin nx \cos cnt.$$

Ancora una volta, affinché sia soddisfatta la condizione iniziale $f(x)$ dovremo scegliere $b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx$.

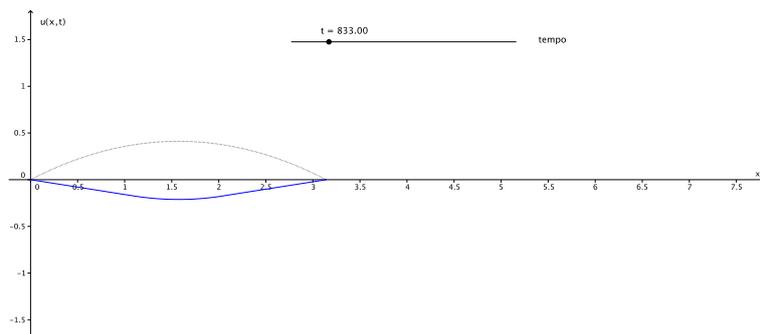
Mostriamo che, sotto opportune ipotesi su f , la serie (VI) fornisce effettivamente una soluzione del nostro problema. Per non avere problemi, supponiamo che f sia estendibile ad una funzione 2π -periodica e dispari di classe almeno C^3 . La disuguaglianza di Bessel applicata alla derivata terza ci garantisce che $\sum_{k=1}^{\infty} k^6 |b_k|^2 < +\infty$ (si integri per parti 3 volte...).

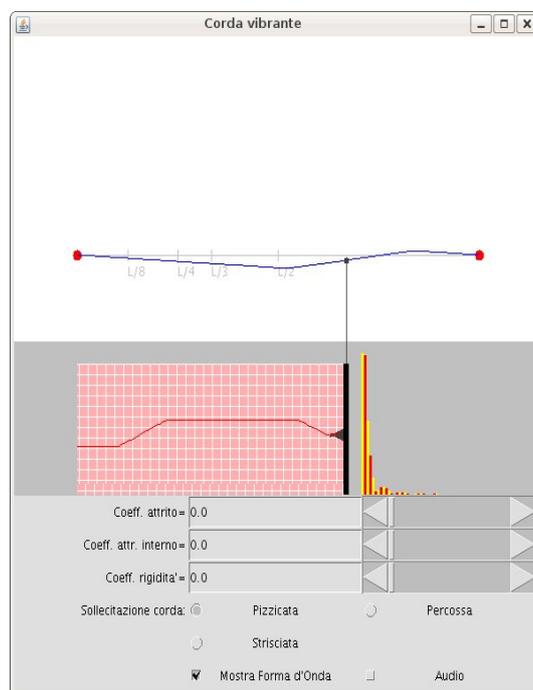
Si ha ovviamente convergenza totale della serie (VI) su tutta la striscia $[0, \pi] \times [0, +\infty]$: la somma è una funzione continua che soddisfa la condizione iniziale e le condizioni al contorno. Per verificare che vale l'equazione, calcoliamo formalmente u_{tt} e u_{xx} derivando termine a termine la serie e verifichiamo che vi è convergenza totale: le serie derivate sono dominate da

$$\sum_{k=1}^{\infty} k^2 |b_k| = \sum_{k=1}^{\infty} \frac{1}{k} k^3 |b_k| \leq \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \right)^{1/2} \left(\sum_{k=1}^{\infty} k^6 |b_k|^2 \right)^{1/2} < +\infty,$$

dove abbiamo usato la disuguaglianza di Cauchy-Schwarz.

Ecco un foglio GeoGebra che mostra un'animazione delle soluzioni dell'equazione delle onde:





Concludo mostrandovi una simulazione di corda vibrante, scritta in java, che permette di osservare il moto della corda pizzicata con un modello leggermente più completo, che tiene conto anche di effetti di smorzamento e di rigidità. Selezionando le apposite caselle, è possibile anche vedere ed ascoltare la forma d'onda generata.