

Introduzione alla teoria della misura e dell'integrale di Lebesgue per il Corso di Metodi Matematici per l'Ingegneria

Marco Bramanti
Politecnico di Milano

2 maggio 2012

Indice

1	Motivazioni per studiare la teoria della misura e dell'integrazione di Lebesgue	2
1.1	Inadeguatezza dell'integrale di Riemann per gli scopi dell'analisi funzionale	2
1.2	L'idea dell'integrale di Lebesgue	6
1.3	Il problema di definire una "buona" misura	8
1.4	Sviluppo e impatto della teoria della misura e dell'integrale di Lebesgue	11
2	Teoria della misura e dell'integrale di Lebesgue	12
2.1	Spazi misurabili	12
2.2	Misura	13
2.3	Costruzione della misura di Lebesgue in \mathbb{R}^n	14
2.4	Restrizione di una misura	19
2.5	Funzioni misurabili	20
2.6	Integrazione astratta	24
2.7	Relazione tra integrale di Riemann e integrale di Lebesgue	29
2.8	I teoremi di convergenza per l'integrale di Lebesgue	30
2.9	Lo spazio L^1	34
2.10	Il teorema fondamentale del calcolo integrale	37
3	La teoria della misura astratta in probabilità	41

4	Misure di Hausdorff	47
4.1	Misura e dimensione di Hausdorff	47
4.2	La misura $(n - 1)$ -dimensionale e il problema isoperimetrico . . .	52
4.3	Misure e distribuzioni di carica	54
5	Spazi L^p	56
5.1	Gli spazi $L^p(\Omega)$ per $1 < p < \infty$	56
5.2	Lo spazio L^∞	60
5.3	Inclusioni tra spazi L^p	61
5.4	Approssimazione di funzioni L^p	63
5.5	Convoluzione in \mathbb{R}^n	63

1 Motivazioni per studiare la teoria della misura e dell'integrazione di Lebesgue

1.1 Inadeguatezza dell'integrale di Riemann per gli scopi dell'analisi funzionale

Abbiamo visto come gli spazi di Lagrange $C^k[a, b]$ costituiscano un ambiente funzionale adatto allo studio delle funzioni *derivabili*: lo spazio $C^1[a, b]$, ad esempio, è uno spazio vettoriale normato completo, costituito dalle funzioni derivabili con continuità; la norma dello spazio esprime un controllo uniforme della “grandezza” sia della funzione che della sua derivata; proprio questo fa sì che la convergenza in questa norma conservi le proprietà caratteristiche delle funzioni di questo spazio, cioè appunto l’essere derivabili con continuità, ed esprima l’approssimazione della funzione limite e della sua derivata con le funzioni approssimanti e le loro derivate. Inoltre la completezza dello spazio si può esprimere intuitivamente dicendo che la convergenza in questa norma¹ garantisce l’appartenenza del limite allo spazio stesso.

Un’esigenza altrettanto naturale è quella di definire uno spazio vettoriale normato di funzioni *integrabili*, ad esempio su un intervallo $[a, b]$. Non solo lo spazio dovrà contenere soltanto funzioni integrabili, ma la norma dovrà esprimere un controllo sulla “grandezza” dell’integrale della funzione, in modo che la convergenza in questa norma conservi l’integrabilità e al tempo stesso esprima l’approssimazione dell’integrale della funzione limite con l’integrale delle funzioni approssimanti. Verosimilmente se tutte queste proprietà risultano soddisfatte lo spazio sarà completo rispetto a questa norma.

Proviamo a definire uno spazio di questo tipo utilizzando l’integrale di Riemann, cioè quello che si studia nei corsi di analisi di base, che è anche quello che è stato utilizzato fino alla fine del 19° secolo.

¹Rigorosamente si dovrebbe dire che *il fatto che la successione sia di Cauchy* garantisce l’appartenenza del limite allo spazio. Tuttavia per gli spazi di funzioni di solito è vero che una successione di Cauchy nello spazio X è una successione convergente “in qualche spazio Y ”, magari più grande di X . Perciò intuitivamente pensiamo la completezza come il fatto che il limite di una successione convergente appartenga allo spazio stesso.

La candidata naturale come norma (usiamo il simbolo provvisorio $\| \cdot \|_I$ con I come integrale per distinguerla da altre norme viste fin qui) è:

$$\|f\|_I = \int_a^b |f(x)| dx,$$

che soddisfa sicuramente la proprietà di omogeneità,

$$\begin{aligned} \|\lambda f\|_I &= |\lambda| \|f\|_I \quad \forall \lambda \in \mathbb{R} \text{ cioè} \\ \int_a^b |\lambda f(x)| dx &= |\lambda| \int_a^b |f(x)| dx \end{aligned}$$

e la disuguaglianza triangolare,

$$\begin{aligned} \|f + g\|_I &\leq \|f\|_I + \|g\|_I \text{ cioè} \\ \int_a^b |f(x) + g(x)| dx &\leq \int_a^b |f(x)| dx + \int_a^b |g(x)| dx. \end{aligned}$$

Lasciamo un momento in sospeso la verifica della prima proprietà della norma (quella di annullamento) e chiediamoci su quale spazio di funzioni mettere questa norma.

Una prima possibilità è $C^0[a, b]$, dal momento che le funzioni continue su $[a, b]$ sono certamente Riemann integrabili. Su questo spazio la norma soddisfa anche la proprietà di annullamento: per una funzione continua su $[a, b]$ è vero che

$$\int_a^b |f(x)| dx = 0 \implies f(x) = 0 \quad \forall x \in [a, b].$$

Perciò $(C^0[a, b], \|\cdot\|_I)$ è uno spazio vettoriale normato; la norma misura la grandezza dell'integrale, e la convergenza in norma esprime un'approssimazione dell'integrale. Il problema è che questa convergenza non conserva l'appartenenza allo spazio stesso, in altre parole questo spazio non è completo. Infatti:

Esempio 1 Sia

$$f_n(x) = \operatorname{sgn}(x) \sqrt[n]{|x|} \text{ in } [-1, 1].$$

Si vede che

$$\|f_n - f\|_I \rightarrow 0 \text{ con } f(x) = \begin{cases} 1 & \text{per } x > 0 \\ -1 & \text{per } x < 0. \end{cases}$$

In particolare $f \notin C^0[-1, 1]$, anche se la successione è di Cauchy.

In un certo senso, abbiamo scelto un insieme di funzioni troppo piccolo: le funzioni continue sono solo *una parte* delle funzioni Riemann integrabili, e l'approssimazione in norma integrale può fare uscire dallo spazio delle funzioni continue. Lo spazio $(C^0[a, b], \|\cdot\|_I)$ non è completo, quindi non è un ambiente funzionale in cui sia facile dimostrare teoremi di esistenza. Non è adeguato per gli scopi dell'analisi funzionale.

Pensiamo allora di considerare lo spazio $R[a, b]$ di *tutte* le funzioni limitate e Riemann integrabili su $[a, b]$, e di munirlo della norma $\|\cdot\|_I$. Una prima complicazione è che questa non soddisfa la proprietà di annullamento: per una generica funzione integrabile,

$$\int_a^b |f(x)| dx = 0 \not\Rightarrow f(x) = 0 \quad \forall x \in [a, b].$$

Ad esempio, una funzione nulla tranne in un numero finito di punti ha integrale del modulo nullo, eppure non è identicamente nulla. Si esce da questo problema decidendo di identificare con 0 le funzioni per cui $\|f\|_I = 0$ e identificare tra loro funzioni la cui differenza ha norma nulla. Poiché questo problema non si presenta solo in quest'esempio, ma si presenterà nella teoria di Lebesgue di cui vorremo parlare, ne approfittiamo per spiegare questo punto rigorosamente.

Nell'insieme $R[a, b]$ introduciamo una *relazione di equivalenza*, ponendo

$$f \sim g \Leftrightarrow \int_a^b |f(x) - g(x)| dx = 0.$$

La relazione \sim soddisfa le proprietà riflessiva, simmetrica e transitiva, perciò è una relazione di equivalenza. L'insieme $R[a, b]$ risulta così ripartito in *classi di equivalenza*, cioè insiemi a due a due disgiunti di funzioni tali che per ogni coppia f, g di funzioni nella stessa classe risulta $\int_a^b |f(x) - g(x)| dx = 0$. Indichiamo con $\{f\}$ la classe di equivalenza a cui appartiene f (si dice che f è un rappresentante della classe), e consideriamo ora l'insieme $R[a, b] / \sim$ (che si dice insieme quoziente di $R[a, b]$ rispetto alla relazione di equivalenza \sim) che ha per elementi le classi di equivalenza. E' uno spazio vettoriale, con le operazioni

$$\begin{aligned} \{f\} + \{g\} &= \{f + g\} \\ c\{f\} &= \{cf\}. \end{aligned}$$

La classe $\{0\}$ che contiene la funzione identicamente nulla e tutte quelle che hanno integrale del modulo nullo, è lo zero dello spazio vettoriale: $\{f\} + \{0\} = \{f\}$.

Poniamo ora:

$$\|\{f\}\|_I = \|f\|_I.$$

Notare che la definizione è coerente perché se g è un'altra funzione della classe $\{f\}$, si avrà

$$\|g\|_I = \|f\|_I.$$

Infatti per la disuguaglianza triangolare si ha

$$\begin{aligned} \|g\|_I &\leq \|g - f\|_I + \|f\|_I \\ \|f\|_I &\leq \|f - g\|_I + \|g\|_I \end{aligned}$$

e poiché, essendo $f \sim g$, risulta $\|g - f\|_I = 0$, concludiamo $\|g\|_I \leq \|f\|_I$ e viceversa.

Il vantaggio di essere passati all'insieme quoziente è che ora se

$$\|\{f\}\|_I = 0,$$

questo implica ora che $\|f\|_I = 0$ cioè $f \sim 0$, quindi $\{f\} = \{0\}$, ossia: se una classe di equivalenza ha norma nulla, questa è la classe che contiene lo zero, cioè la classe $\{0\}$. Quindi vale anche la proprietà di annullamento, ossia lo spazio quoziente $R[a, b] / \sim$ con la norma $\|\{\cdot\}\|_I$ risulta uno *spazio vettoriale normato*.

Introdurre uno spazio di “classi di equivalenza di funzioni” è il modo, rigoroso ma molto pesante, in cui la teoria degli insiemi formalizza questa situazione. L'idea intuitiva con cui noi possiamo pensare questo spazio, più agilmente, è che i suoi elementi siano le solite funzioni, con l'avvertenza però che due funzioni la cui differenza ha norma nulla sono pensate come la stessa funzione; in particolare una funzione a norma nulla si indentifica con la funzione identicamente nulla. Si comincia a vedere all'orizzonte come la riflessione sull'integrale dal punto di vista dell'analisi funzionale spinga a modificare un po' il modo tradizionale di pensare le funzioni. Questo spunto sarà approfondito nel seguito, parlando di integrale di Lebesgue.

Proseguiamo con l'analisi dello spazio $(R[a, b], \|\cdot\|_I)$ (seguendo l'idea intuitiva appena enunciata, continuiamo a indicarlo senza far uso delle notazioni relative alle classi di equivalenza). Si constata che, purtroppo,

lo spazio $(R[a, b], \|\cdot\|_I)$ non è completo.

Un esempio di successione che in questo spazio è di Cauchy ma non converge si potrà dare più facilmente quando avremo introdotto una parte della teoria dell'integrale di Lebesgue. Ma segnaliamo da subito questo risultato negativo.

Prima di proseguire commentando la non completezza di questo spazio, segnaliamo qualche altra proprietà insoddisfacente di questo spazio. Si consideri il seguente

Esempio 2 Sia $\{x_n\}_{n=1}^\infty$ la successione dei razionali dell'intervallo $[0, 1]$ (poiché sono un'infinità numerabile, possono essere messi in successione, anche se questa non sarà una successione monotona) e sia

$$\begin{aligned} f_n &: [0, 1] \rightarrow \mathbb{R} \\ f_n(x) &= \begin{cases} 1 & \text{per } x = x_1, x_2, \dots, x_n \\ 0 & \text{altrimenti.} \end{cases} \end{aligned}$$

Si ha

$$f_n(x) \rightarrow f(x) = \begin{cases} 1 & \text{per } x \in \mathbb{Q} \\ 0 & \text{altrimenti.} \end{cases}$$

Notiamo che ciascuna f_n è Riemann-integrabile, con $\|f_n\|_I = 0$, mentre f (che è la famosa “funzione di Dirichlet”) non è Riemann integrabile. Nello spazio delle classi di equivalenza, ciascuna f_n è la funzione identicamente nulla (!!!), quindi nello spazio considerato questa è una successione costante (nulla), che ovviamente tende a zero, eppure dal punto di vista della convergenza puntuale $f_n \rightarrow f$ non integrabile.

La situazione è strana: il passaggio alle classi di equivalenza non è coerente col concetto di convergenza puntuale di una successione di funzioni. Questo accade perché *la classe delle funzioni Riemann-integrabili non è chiusa rispetto al limite puntuale*. (Se la funzione di Dirichlet fosse Riemann integrabile -e avesse integrale nullo-, avremmo una successione nulla che tende a zero: niente di interessante, ma niente di strano).

Se nello spazio delle funzioni Riemann integrabili, infine, mettessimo la norma C^0 (della convergenza uniforme) otterremmo sì uno spazio completo (limite uniforme di funzioni Riemann integrabili è Riemann integrabile), ma a prezzo di perdere una delle motivazioni da cui siamo partiti: usare una norma *integrale*, in modo che la convergenza in norma traduca l'approssimazione dell'integrale della funzione limite con l'integrale delle funzioni approssimanti. La convergenza uniforme è una condizione troppo forte, che *implica* la convergenza degli integrali, ma non è soddisfatta in tutte le situazioni in cui gli integrali convergono.

1.2 L'idea dell'integrale di Lebegue

Torniamo ora a commentare la non completezza dello spazio $(R[a, b], \|\cdot\|_I)$. Così come la non completezza dello spazio $(C^0[a, b], \|\cdot\|_I)$ è stata interpretata dicendo che le funzioni continue sono solo una parte delle funzioni integrabili, e perciò la convergenza in norma integrale può far uscire dallo spazio delle funzioni continue, così potremmo immaginare che anche le funzioni Riemann-integrabili siano solo una parte di quelle per cui è possibile (in qualche altro modo) definire un integrale, che naturalmente coincida col solito integrale quando applicato a funzioni Riemann-integrabili. Nasce quindi l'idea che si possa allargare la classe delle funzioni integrabili, per ottenere uno spazio completo. Questo però potrà essere fatto solo *cambiando la definizione di integrale* (finché l'integrale è quello, le funzioni integrabili sono quelle!), e più precisamente cambiandolo in un senso che renda *meno restrittiva* la richiesta di integrabilità. Ricordiamo ora come è stato definito l'integrale di Riemann come limite di somme:

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{b-a}{n} f(\xi_k^{(n)})$$

dove al passo n -esimo l'intervallo $[a, b]$ è stato suddiviso in n intervalli uguali mediante i punti $a = x_0^{(n)}, x_1^{(n)}, x_2^{(n)}, \dots, x_n^{(n)} = b$ e ad ogni passo si sono scelti, arbitrariamente, gli n punti $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$, $k = 1, 2, \dots, n$. La funzione f si dice Riemann-integrabile se il limite esiste finito e non dipende da come si sono scelti, ad ogni passo, i punti $\xi_k^{(n)}$. Si capisce dalla definizione che ciò che “mette a rischio” l'integrabilità è il fatto che f abbia numerosi punti di discontinuità, e quindi nel corso della costruzione succeda “tante volte” che scegliendo il punto $\xi_k^{(n)}$ in modi diversi all'interno dell'intervallino $[x_{k-1}^{(n)}, x_k^{(n)}]$ si ottengano delle variazioni importanti nel valore di s_n . Ad esempio, la funzione di Dirichlet non è integrabile perché, ad ogni passo della costruzione iterativa, è possibile scegliere i punti $\xi_k^{(n)}$ tutti razionali, e allora $s_n = 1$, o tutti irrazionali, e allora

$s_n = 0$; perciò certamente successioni di Cauchy-Riemann diverse hanno limiti diversi. *Se vogliamo dare una definizione di integrale diversa, che risulti meno restrittiva, dobbiamo trovare il modo di neutralizzare, nell'algoritmo di calcolo dell'integrale, gli effetti di instabilità dovuti alle (eventuali) grandi oscillazioni o discontinuità della funzione.*

Un'idea di questo tipo si trova nella nozione di integrale introdotta da Henri Lebesgue nel 1902 nella sua tesi di dottorato presentata alla Facoltà di Scienze di Parigi.

L'idea, apparentemente banale, è: invece di fare una costruzione iterativa suddividendo l'intervallo $[a, b]$ sull'asse x in parti uguali, facciamo una costruzione iterativa suddividendo in parti uguali, sull'asse y , un intervallo che contenga l'insieme dei valori assunti dalla funzione (che per il momento supponiamo limitata). Ad esempio, se la funzione ha valori in $[0, 1]$, al passo n -esimo suddivideremo $[0, 1]$ in parti uguali e considereremo gli insiemi

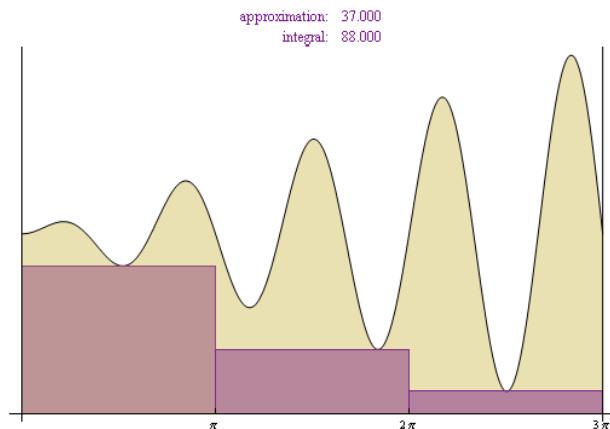
$$E_k^{(n)} = \left\{ x \in [a, b] : \frac{k-1}{n} \leq f(x) < \frac{k}{n} \right\}, \text{ con } k = 1, 2, \dots, n.$$

Ovviamente un'approssimazione, per difetto o per eccesso, dell'area sotto il grafico di f si ottiene, rispettivamente, con le somme:

$$s_n^- = \sum_{k=1}^n |E_k^{(n)}| \frac{k-1}{n};$$

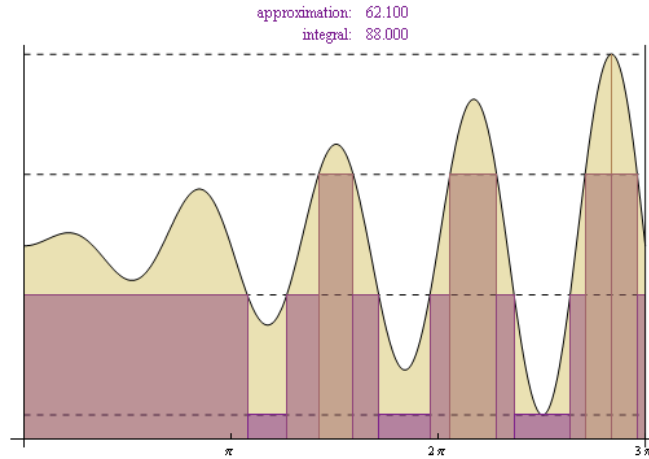
$$s_n^+ = \sum_{k=1}^n |E_k^{(n)}| \frac{k}{n},$$

dove $|E_k|$ indica la misura dell'insieme E_k , cioè se ad es. è un intervallo la sua lunghezza, se è l'unione di più intervalli la somma delle lunghezze degli intervalli, e così via².



Una somma di Riemann di passo 3

²Come vedremo, buona parte del problema da affrontare sta in questo semplicistico “e così via”.



Una somma di Lebesgue di passo 3 della stessa funzione

Ciò che dimostra quanto sia buona quest'idea, e quanto sia diversa da quella dell'integrale di Riemann (non ostante l'apparente simmetria: suddividiamo l'asse x / suddividiamo l'asse y) è che per come sono costruite le somme risulta sempre

$$0 \leq s_n^+ - s_n^- = \sum_{k=1}^n \left| E_k^{(n)} \right| \frac{1}{n} = \frac{1}{n}$$

per cui lo scarto tra le approssimazioni per eccesso e per difetto tende necessariamente a zero! Significa che questo algoritmo di approssimazione restituisce “sempre” un limite. Apparentemente, ogni funzione risulta integrabile a questo modo. In realtà, abbiamo in un certo senso spostato il problema: data una funzione f e interi n, k qualsiasi, come sarà fatto l'insieme $E_k^{(n)}$? Se f ha molte oscillazioni e discontinuità potrà essere un insieme molto diverso da un intervallo o l'unione di un numero finito di intervalli. Ad esempio, per la funzione di Dirichlet tra gli insiemi $E_k^{(n)}$ ci saranno

$$\{x \in [0, 1] : x \in \mathbb{Q}\}, \{x \in [0, 1] : x \notin \mathbb{Q}\},$$

che non sono esprimibili come unioni finite di intervalli.

Dunque: *per costruire una teoria dell'integrazione di questo nuovo tipo dobbiamo prima impegnarci a costruire una “teoria della misura” che sia in grado di assegnare una “lunghezza” anche a sottoinsiemi molto irregolari della retta.*

1.3 Il problema di definire una “buona” misura

Dobbiamo dunque trovare il modo di definire la “lunghezza” (d'ora in poi la chiameremo “misura”) di un sottoinsieme della retta il più possibile generico. Idealmente, ad ogni sottoinsieme E di \mathbb{R} vorremmo poter assegnare un numero $\mu(E)$, non negativo ed eventualmente infinito (la retta intera, o una semiretta,

hanno lunghezza infinita), in modo che questa “funzione d’insieme”³

$$\mu : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$$

abbia alcune proprietà ragionevoli. Certamente vorremo che sia $\mu(\emptyset) = 0$. Non sarà vero il viceversa, ossia: ci saranno insiemi non vuoti di misura nulla (un punto ha lunghezza nulla). La misura elementare è additiva (la misura del tutto è somma della misura delle parti), quindi vorremo che risulti

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ ogni volta che } A \cap B = \emptyset.$$

Se teniamo presente che uno dei nostri obiettivi finali è costruire uno spazio di funzioni integrabili che risulti completo, sarà fondamentale che la nostra teoria permetta di trattare agevolmente i limiti di successioni di funzioni; sarà prudente allora richiedere che già al livello della misura le successioni (in questo caso di insiemi) abbiano un posto di riguardo. Chiediamo perciò che la misura μ risulti additiva anche per *successioni di insiemi*:

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

per ogni successione $\{E_n\}_{n=1}^{\infty}$ di insiemi a due a due disgiunti. Questa proprietà si chiama *numerabile additività*. Infine, stiamo costruendo una misura che generalizzi l’idea elementare di lunghezza, quindi vogliamo che per un intervallo $[a, b]$ risulti

$$\mu([a, b]) = b - a$$

(lunghezza elementare). Riassumiamo: vorremmo definire una funzione

$$\mu : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty] \tag{1}$$

tale che

$$\mu(\emptyset) = 0 \tag{2}$$

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n) \tag{3}$$

per ogni successione $\{E_n\}_{n=1}^{\infty}$ di insiemi a due a due disgiunti, e

$$\mu([a, b]) = b - a \quad \forall a, b \in \mathbb{R}, a \leq b \tag{4}$$

Purtroppo *questo programma si è rivelato impossibile*. Infatti un profondo teorema di S. Ulam (1909-1984) mostra che *non esiste alcuna funzione μ che soddisfi simultaneamente (1), (2), (3), (4)*. Precisamente, il teorema di Ulam afferma che:

³ Il simbolo $\mathcal{P}(\mathbb{R})$ indica l’insieme delle parti di \mathbb{R} , cioè l’insieme di tutti i sottoinsiemi di \mathbb{R} .

“L’unica funzione μ soddisfacente (1), (2), (3) e tale che per ogni $x \in \mathbb{R}$ sia $\mu(\{x\}) = 0$ (cioè l’insieme costituito da un punto solo ha sempre misura nulla) è quella identicamente nulla”.

Poiché la proprietà (4) implica che per ogni $x \in \mathbb{R}$ è $\mu(\{x\}) = 0$ (prendere $a = b = x$) e implica anche che μ non è identicamente nulla (prendere $a < b$), ne segue quanto abbiamo affermato.

A quale dell 4 proprietà possiamo rinunciare? L’unica scelta possibile, in vista del nostro obiettivo, è rinunciare a (1): costruiremo una misura che abbia proprietà ragionevoli ed estenda la misura elementare, ma *non tutti* i sottoinsiemi di \mathbb{R} risulteranno misurabili. La misura μ sarà definita solo su una certa famiglia \mathcal{L} di sottoinsiemi di \mathbb{R} . Come scegliere \mathcal{L} ? Che proprietà dovrà avere? Per coerenza con le richieste (2) e (3), \mathcal{L} dovrà contenere l’insieme vuoto, e inoltre dovrà essere chiuso per unioni numerabili:

$$E_n \in \mathcal{L} \text{ per } n = 1, 2, 3, \dots \implies \bigcup_{n=1}^{\infty} E_n \in \mathcal{L}. \quad (5)$$

Anche l’insieme \mathbb{R} (cioè lo spazio intero) dovrà essere misurabile (con misura infinita, ci aspettiamo); inoltre, \mathcal{L} dovrà essere chiuso per qualche altra operazione insiemistica elementare, ad esempio la complementazione: se sappiamo misurare E , dovremmo saper misurare anche il complementare E^c di E :

$$E \in \mathcal{L} \implies E^c \in \mathcal{L}.$$

Infine, per coerenza con la richiesta (4), \mathcal{L} dovrà contenere gli intervalli. Siamo così arrivati a capire che per costruire una buona misura μ dobbiamo anzitutto occuparci di individuare un’opportuna famiglia di sottoinsiemi di \mathbb{R} su cui definirla, famiglia che deve anch’essa godere di opportune proprietà algebriche, rispetto alle operazioni insiemistiche. Provando a guardare in avanti, immaginiamo ora che cosa succederà quando, avendo costruito una buona misura sulla retta, che assegna una “lunghezza” appropriata ad ogni insieme $E \in \mathcal{L}$, dovremo definire l’integrale di una funzione $f(x)$; affinché l’algoritmo di approssimazione ideato da Lebesgue sia utilizzabile, occorrerà che gli insiemi

$$E_k^{(n)} = \left\{ x \in [a, b] : \frac{k-1}{n} \leq f(x) < \frac{k}{n} \right\}$$

appartengano tutti alla famiglia \mathcal{L} ; questo si traduce in una richiesta sulla funzione f , che non potrà essere totalmente arbitraria, ma dovrà essere anch’essa “misurabile”, il che per definizione significherà appunto che tutti gli insiemi di livello $E_k^{(n)}$ sono misurabili. Ragionevolmente, quindi, non tutte le funzioni risulteranno integrabili secondo Lebesgue; tuttavia, dovrebbero risultare integrabili molte più funzioni di quelle Riemann-integrabili; soprattutto, data la “intrinseca stabilità” della definizione di integrale di Lebesgue e dato che tra le proprietà della misura abbiamo inserito la numerabile additività, che dovrebbe essere d’aiuto nel dimostrare proprietà delle successioni di funzioni, possiamo sperare che questo spazio risulti completo rispetto alla norma integrale.

Il percorso logico che ci attende è ora sostanzialmente capovolto rispetto a quello seguito in questa introduzione. Dovremo:

1. Definire e studiare particolari famiglie di sottoinsiemi dell'insieme universo (che nel discorso fatto fin qui era \mathbb{R} , ma potrebbe essere \mathbb{R}^n o anche un insieme diverso), le cosiddette “ σ -algebre”;
2. definire il concetto di misura e studiarne le proprietà; definire una particolare misura, quella di Lebesgue, che estende la misura elementare in \mathbb{R}^n ;
3. definire cosa si intenda per “funzioni misurabili” e studiarne le proprietà;
4. definire l'integrale di una funzione misurabile e dimostrarne le proprietà;
5. definire uno spazio vettoriale normato di funzioni Lebesgue integrabili e dimostrarne la completezza.

1.4 Sviluppo e impatto della teoria della misura e dell'integrale di Lebesgue

Nel percorrere i passi appena descritti di questa teoria, ci si è resi conto che buona parte di essa può essere sviluppata in un contesto molto astratto, cioè senza supporre che lo spazio ambiente sia \mathbb{R} o \mathbb{R}^n , e qualora anche lo spazio ambiente sia \mathbb{R}^n , senza supporre che la misura estenda la misura elementare. In altre parole, risulta importante sia costruire una misura e un integrale sui consueti sottoinsiemi di \mathbb{R}^n che generalizzi la misura elementare e l'integrale di Riemann, permettendo di costruire spazi di funzioni integrabili che siano di Banach, sia costruire una teoria della misura e dell'integrazione astratta, (cioè rispetto ad una misura “qualsiasi” in uno spazio ambiente “qualsiasi”), che sia uno strumento teorico flessibile e applicabile a situazioni diverse.

La teoria di Lebesgue ha rivoluzionato l'analisi matematica del 20° secolo in molti modi. L'idea di identificare due funzioni f, g ogni volta che $\int |f - g| = 0$ (si dirà che “ f e g sono uguali quasi ovunque”) ha cambiato il modo stesso di pensare al concetto di funzione, rispetto alla definizione “punto per punto” data da Dirichlet negli anni 1830; gli spazi di funzioni integrabili secondo Lebesgue, uniti a un altro sviluppo dell'analisi funzionale astratta di cui parleremo in seguito, cioè la *teoria degli spazi di Hilbert*, ha gettato nuova luce sullo studio delle *serie di Fourier*, rivoluzionando l'*analisi armonica*; il teorema fondamentale del calcolo integrale nel contesto della teoria dell'integrazione di Lebesgue è stato il punto di partenza per una generalizzazione del concetto di derivata, la cosiddetta “derivata debole”, di cui parleremo, che entra nella definizione degli spazi di funzioni (“*spazi di Sobolev*”) che costituiscono dagli anni 1950 il contesto standard in cui studiare le *equazioni alle derivate parziali*.

La versione astratta della teoria dell'integrazione ha avuto altri risvolti fondamentali: negli anni 1930 Kolmogorov ha rifondato assiomaticamente il *calcolo delle probabilità* dandogli la veste che ha ancor oggi, interamente fondata sulla teoria della misura astratta; la teoria astratta della misura ha anche dato la possibilità di definire le misure di dimensione inferiore a quella dello spazio ambiente (misura di superficie nello spazio tridimensionale, ad es.), creando una teoria adatta a studiare anche oggetti geometrici molto irregolari; ne è nata la *teoria geometrica della misura*, che ha saputo risolvere questioni che erano

aperte da secoli, come il problema isoperimetrico classico, e ha dato strumenti potenti al *calcolo delle variazioni*. L'idea di derivata debole, già citata, e la teoria della misura, infine, sono stati tra gli spunti per la nascita della *teoria delle distribuzioni*, che negli anni 1950 ha ampiamente generalizzato il concetto di funzione, derivata, trasformata di Fourier, con conseguenze vaste sia sullo studio delle equazioni a derivate parziali che sull'analisi armonica.

2 Teoria della misura e dell'integrale di Lebesgue

Per approfondimenti sul contenuto di questa sezione si rimanda a: [1], [3], [2], [6].

2.1 Spazi misurabili

Definizione 3 Sia Ω un insieme. Si dice σ -algebra (su Ω) una famiglia \mathcal{M} di sottoinsiemi di Ω (cioè⁴ $\mathcal{M} \subseteq \mathcal{P}(\Omega)$) tale che:

$\Omega \in \mathcal{M}$;

$E \in \mathcal{M} \implies E^c \in \mathcal{M}$ (dove E^c indica il complementare di E in Ω);

se $\{E_n\}_{n=1}^\infty$ è una successione di insiemi di \mathcal{M} , allora $\bigcup_{n=1}^\infty E_n \in \mathcal{M}$.

Gli insiemi di \mathcal{M} si dicono insiemi misurabili, (Ω, \mathcal{M}) si dice spazio misurabile.

Ogni insieme Ω ha due σ -algebre banali: la più piccola è quella costituita solo da Ω e \emptyset ; la più grande è tutto $\mathcal{P}(\Omega)$. Dalla definizione segue facilmente la:

Proposizione 4 Se \mathcal{M} è una σ -algebra, \mathcal{M} è chiusa anche rispetto alle seguenti operazioni insiemistiche: unione finita, intersezione finita o numerabile; differenza insiemistica⁵.

Proposizione 5 Se $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ è una qualsiasi famiglia di sottoinsiemi di Ω , è ben definita la più piccola σ -algebra contenente \mathcal{E} , detta " σ -algebra generata da \mathcal{E} " e si indica con $\sigma(\mathcal{E})$.

Infatti, di σ -algebre contenenti \mathcal{E} ce n'è almeno una: $\mathcal{P}(\Omega)$; consideriamo allora l'intersezione di tutte le σ -algebre contenenti \mathcal{E} ; si tratta di dimostrare che quest'intersezione è a sua volta una σ -algebra (facile esercizio); questa sarà per costruzione la più piccola σ -algebra contenente \mathcal{E} .

Esempio 6 (Relazione tra topologia e teoria della misura) Sia Ω uno spazio metrico qualsiasi, e \mathcal{A} la famiglia di tutti gli insiemi aperti di Ω . La σ -algebra $\sigma(\mathcal{A})$ (generata da \mathcal{A} nel senso della definizione precedente) si dice σ -algebra

⁴dove $\mathcal{P}(\Omega)$ indica l'insieme delle parti di Ω , cioè l'insieme di tutti i sottoinsiemi di Ω , compresi Ω stesso e l'insieme vuoto \emptyset

⁵La differenza insiemistica è definita così:

$$A \setminus B = A \cap B^c.$$

di Borel in Ω , si indica con $\mathcal{B}(\Omega)$, e i suoi insiemi si chiamano boreliani di Ω . Tra i boreliani troviamo tutti gli insiemi aperti, tutti i chiusi, ma anche tutti gli insiemi (generalmente né aperti né chiusi) che si ottengono facendo unione numerabile di chiusi o intersezione numerabile di aperti. Se ad esempio $\Omega = \mathbb{R}^n$ (con la distanza euclidea), diciamo pure che è molto difficile costruire un insieme di che non sia un boreliano. Questa σ -algebra è quindi una classe molto ampia e generale di sottoinsiemi di Ω .

2.2 Misura

Definizione 7 Sia (Ω, \mathcal{M}) uno spazio misurabile. Si dice misura (su questo spazio) una qualunque funzione (d'insieme)

$$\mu : \mathcal{M} \rightarrow [0, +\infty]$$

che sia numerabilmente additiva, ossia tale che per ogni successione $\{E_n\}_{n=1}^{\infty}$ di insiemi di \mathcal{M} a due a due disgiunti, si abbia

$$\mu \left(\bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n).$$

(Dove ambo i membri possono essere finiti o infiniti). In tal caso $(\Omega, \mathcal{M}, \mu)$ si dice spazio di misura.

Da questa sola richiesta di numerabile additività seguono in realtà molte altre proprietà della misura:

Teorema 8 Sia $(\Omega, \mathcal{M}, \mu)$ uno spazio di misura. Allora:

1. $\mu(\emptyset) = 0$;
2. μ è finitamente additiva, cioè per ogni famiglia finita E_1, E_2, \dots, E_n di insiemi di \mathcal{M} a due a due disgiunti si ha

$$\mu \left(\bigcup_{i=1}^n E_i \right) = \sum_{i=1}^n \mu(E_i);$$

3. μ è monotona, cioè $\forall A, B \in \mathcal{M}, A \subseteq B \implies \mu(A) \leq \mu(B)$;
4. μ è condizionatamente sottrattiva, cioè $\forall A, B \in \mathcal{M}, A \subseteq B$ e $\mu(B) < \infty \implies \mu(B \setminus A) = \mu(B) - \mu(A)$;
5. μ è continua da sotto: se $\{E_n\}_{n=1}^{\infty}$ è una successione di insiemi di \mathcal{M} tale che $E_n \nearrow E$ (cioè: $E_n \subseteq E_{n+1} \forall n$ e $\bigcup_{n=1}^{\infty} E_n = E$) allora $\mu(E_n) \rightarrow \mu(E)$;
6. μ è condizionatamente continua da sopra: se $\{E_n\}_{n=1}^{\infty}$ è una successione di insiemi di \mathcal{M} tale che $E_n \searrow E$ (cioè: $E_n \supseteq E_{n+1} \forall n$ e $\bigcap_{n=1}^{\infty} E_n = E$) e inoltre $\mu(E_1) < \infty$, allora $\mu(E_n) \rightarrow \mu(E)$;
7. μ è numerabilmente subadditiva, cioè se $\{E_n\}_{n=1}^{\infty}$ è una successione di insiemi di \mathcal{M} (non necessariamente a due a due disgiunti), allora

$$\mu \left(\bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

Lo studente può provare per esercizio a dimostrare qualcuna di queste proprietà.

Nel seguito, come vedremo, saranno molto importanti gli insiemi di misura nulla. Notiamo che se $E, E_0 \in \mathcal{M}$, $E_0 \subset E$ e $\mu(E) = 0$, allora (per la monotonia della misura) anche $\mu(E_0) = 0$. Talvolta siamo in una situazione leggermente diversa: abbiamo un insieme $E \in \mathcal{M}$ tale che $\mu(E) = 0$, e un altro insieme $E_0 \subset E$ (di cui a priori non sappiamo che sia misurabile); ci piacerebbe poter comunque concludere che $\mu(E_0) = 0$ (cioè: che $E_0 \in \mathcal{M}$, e quindi $\mu(E_0) = 0$). Questo non è vero per tutte le misure, ma solo per quelle per cui è noto che *i sottoinsiemi degli insiemi di misura nulla sono tutti misurabili* (e quindi di misura nulla).

Definizione 9 Sia $(\Omega, \mathcal{M}, \mu)$ uno spazio di misura. Si dice che la misura μ è completa se i sottoinsiemi degli insiemi di misura nulla sono tutti misurabili (e quindi hanno misura nulla).

Esempi di misure. Non è così semplice costruire misure significative. Cominciamo a introdurre un paio di esempi semplici (ma comunque importanti).

La misura del conteggio. Sia Ω un insieme qualsiasi, $\mathcal{M} = \mathcal{P}(\Omega)$ e $\mu : \mathcal{M} \rightarrow [0, +\infty]$ tale che

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

Si verifica che μ è una misura detta misura del conteggio. Come vedremo, l'integrale rispetto a questa misura risulterà una serie numerica, quindi la teoria astratta dell'integrazione assorbirà al suo interno la teoria delle serie numeriche.

La misura atomica o di Dirac. Sia Ω un insieme qualsiasi, $x_0 \in \Omega$ un suo elemento fissato, e $\mathcal{M} = \mathcal{P}(\Omega)$. Definiamo $\mu : \mathcal{M} \rightarrow [0, +\infty]$ tale che

$$\mu(A) = \begin{cases} 1 & \text{se } x_0 \in A \\ 0 & \text{altrimenti.} \end{cases}$$

Si verifica che μ è una misura detta misura di Dirac concentrata in x_0 e si indica talvolta col simbolo δ_{x_0} .

Vediamo ora come costruire la misura fondamentale che ci interessa negli spazi \mathbb{R}^n , che permetterà di definire un integrale che estenda l'integrale classico di Riemann.

2.3 Costruzione della misura di Lebesgue in \mathbb{R}^n

Il nostro obiettivo ora è costruire una σ -algebra di sottoinsiemi di \mathbb{R}^n che contenga gli insiemi elementari (di conseguenza, conterrà i boreliani), e definire su di essa una misura che estenda la misura elementare. La costruzione che faremo non è quella originale di Lebesgue ma è dovuta a Carathéodory, e procede in due passi. In un primo tempo si definisce una funzione d'insieme non negativa

μ^* che è definita su *tutti* i sottoinsiemi di \mathbb{R}^n , estende la misura elementare, ma non è una misura (*non può esserlo*, in base al teorema di Ulam). Questa μ^* si chiamerà *misura esterna*. In un secondo tempo individueremo una σ -algebra \mathcal{L} di sottoinsiemi di \mathbb{R}^n tale che μ^* , ristretta a \mathcal{L} , sia numerabilmente additiva, quindi una misura: \mathcal{L} sarà la σ -algebra degli insiemi Lebesgue misurabili (e conterrà la σ -algebra di Borel), e la restrizione di μ^* a \mathcal{L} sarà la misura di Lebesgue μ .

Consideriamo gli insiemi più semplici che sono l'analogo in \mathbb{R}^n degli intervalli sulla retta.

Definiamo n -cella un insieme del tipo:

$$I = [a_1, b_1] \times [a_2, b_2] \times \dots [a_n, b_n], \text{ con } a_i, b_i \in \mathbb{R}.$$

Per ogni n -cella I come sopra poniamo

$$\mu^*(I) = |b_1 - a_1| \cdot |b_2 - a_2| \cdot \dots \cdot |b_n - a_n|.$$

(Cioè sulle n -celle μ^* è definito come il volume elementare). Ora poniamo, per ogni $E \subset \mathbb{R}^n$,

$$\mu^*(E) = \inf \left\{ \sum_{k=1}^{\infty} \mu^*(I_k) : E \subset \bigcup_{k=1}^{\infty} I_k \right\}.$$

In altre parole: si considerano tutte le successioni di n -celle la cui unione ricopre E (poiché le n -celle sono una successione e non un numero finito, si può ricoprire anche un insieme E illimitato); per ciascuna di queste successioni di n -celle si considera la somma (serie) delle misure elementari; quindi si prende l'estremo inferiore al variare delle coperture di E con n -celle. Notare che se ogni possibile copertura di E con n celle è tale che la serie delle misure sia infinita, si avrà $\mu^*(E) = \infty$. Se E è un insieme limitato, certamente $\mu^*(E) < \infty$ (ma può essere $\mu^*(E) < \infty$ anche per certi E illimitati).

Come anticipato nell'introduzione di questo paragrafo, la funzione μ^* non è una misura. La proprietà importante che si può dimostrare per μ^* è espressa dal seguente:

Teorema 10 *La funzione μ^* è numerabilmente subadditiva, cioè per ogni successione $\{E_k\}_{k=1}^{\infty}$ di sottoinsiemi di \mathbb{R}^n (non necessariamente a due a due disgiunti), si ha*

$$\mu^* \left(\bigcup_{k=1}^{\infty} E_k \right) \leq \sum_{k=1}^{\infty} \mu^*(E_k).$$

(Dove i due membri possono essere eventualmente infiniti).

Il punto fondamentale è che anche se gli insiemi E_k si assumono a due a due disgiunti, in generale non si può garantire che nella disuguaglianza precedente valga il segno di $=$. Per questo μ^* non è una misura. Del resto, per il teorema di Ulam, non potrebbe esistere una misura definita su tutti i sottoinsiemi di \mathbb{R}^n che estenda la misura elementare.

Apriamo una parentesi, diamo la seguente definizione astratta, che ci sarà utile più avanti:

Definizione 11 *Sia Ω un insieme qualsiasi. Si dice misura esterna su Ω una funzione d'insieme*

$$\nu : \mathcal{P}(\Omega) \rightarrow [0, +\infty]$$

che sia numerabilmente subadditiva, ossia tale che per ogni successione $\{E_k\}_{k=1}^{\infty}$ di sottoinsiemi di Ω si abbia

$$\nu \left(\bigcup_{k=1}^{\infty} E_k \right) \leq \sum_{k=1}^{\infty} \nu(E_k).$$

Torniamo al contesto di \mathbb{R}^n e alla misura esterna μ^* . L'idea ingegnosa di Carathéodory consiste nell'individuare una condizione che permetta di selezionare, mediante μ^* , una σ -algebra di insiemi su cui μ^* sia numerabilmente additiva.

Definizione 12 *Si dice che $E \subset \mathbb{R}^n$ ha la proprietà di Carathéodory se*

$$\mu^*(E) = \mu^*(E \cap T) + \mu^*(E \cap T^c) \text{ per ogni } T \subset \mathbb{R}^n.$$

Detto a parole: un insieme ha la proprietà di Carathéodory se suddivide additivamente ogni sottoinsieme di \mathbb{R}^n . Vale il seguente

Teorema 13 (di Carathéodory) *Sia \mathcal{L} la famiglia dei sottoinsiemi di \mathbb{R}^n aventi la proprietà di Carathéodory. Allora:*

- \mathcal{L} è una σ -algebra;*
- \mathcal{L} contiene la σ -algebra di Borel;*
- la funzione μ^* ristretta a \mathcal{L} è una misura μ , detta misura di Lebesgue;*
- la misura di Lebesgue è completa;*
- la misura di Lebesgue è invariante per traslazioni ed estende la misura elementare.*

Qualche commento sul teorema precedente. Dal punto di vista pratico, la misura di Lebesgue è definita esattamente come la misura esterna μ^* , soltanto che la classe di insiemi misurabili non coincide con $\mathcal{P}(\mathbb{R}^n)$ ma è più ristretta. In realtà la classe degli insiemi Lebesgue misurabili è veramente molto ampia e comprende non solo tutti gli insiemi di Borel, come affermato dal teorema, ma in pratica tutti gli insiemi di cui si possa dare un procedimento costruttivo esplicito di definizione. Per costruire esempi di insiemi non misurabili bisogna utilizzare l'assioma della scelta, un assioma della teoria degli insiemi che asserisce la possibilità di definire insiemi con un certo procedimento non costruttivo che richiede di effettuare infinite scelte arbitrarie. Un esempio di questo tipo si trova in [1], §2.3.

Il fatto che operativamente $\mu(E)$ si calcoli come $\mu^*(E)$ fa capire che μ è invariante per traslazioni ed estende la misura elementare: infatti μ^* è definita a partire dalla misura elementare delle n -celle, che è invariante per traslazioni;

inoltre μ è (numerabilmente e quindi anche finitamente) additiva; dalla proprietà di additività finita e il fatto che la misura delle n -celle sia il volume elementare segue che la misura degli insiemi elementari (ad esempio, se siamo in \mathbb{R}^2 o \mathbb{R}^3 i poligoni e i poliedri, rispettivamente) coincide con la loro misura elementare. (Si pensi a come, in geometria elementare, si arriva a stabilire le formule per il calcolo dell'area dei poligoni: si basano tutte sulla formula per l'area del rettangolo -che è la "misura elementare di una 2-cella"- e l'additività dell'area).

Esempio 14 Per visualizzare meglio come agisce la definizione di misura di Lebesgue, facciamo qualche esempio.

1. Poiché la misura di Lebesgue estende la misura elementare, un punto ha misura nulla (in \mathbb{R}^n , qualunque sia n).

2. D'altro canto la misura è numerabilmente additiva perciò, ogni insieme numerabile (è misurabile e) ha misura nulla.

3. In \mathbb{R} quanto appena affermato è particolarmente significativo, perché mostra che \mathbb{Q} è un sottoinsieme di \mathbb{R} misurabile e di misura nulla, non ostante il fatto che \mathbb{Q} sia denso in \mathbb{R} . Si capisce che la valutazione di quanto sia "grande" un insieme dal punto di vista della teoria della misura e dal punto di vista della topologia può essere molto diverso. La misurabilità dell'insieme \mathbb{Q} è un fatto che deve far riflettere sulla potenza della teoria: utilizzando le nozioni elementari di misura dei segmenti, un insieme come \mathbb{Q} sarebbe "intrattabile".

4. Si rifletta sul fatto che, ad esempio, una retta (o una curva regolare) nel piano o nello spazio ha misura nulla, così come un piano (o una superficie regolare) nello spazio ha misura nulla. Questo si può vedere riflettendo sulla definizione di μ^* mediante coperture dell'insieme da misurare con successioni di n -celle. A titolo di esempio, mostriamo che una retta, ad esempio l'asse x , ha misura nulla nel piano. Per cominciare sia E il semiasse $x \geq 0, y = 0$ nel piano, e proviamo che $\mu(E) = 0$. Possiamo coprire E con l'unione delle 2 celle così fatte:

$$E \subset \bigcup_{k=0}^{\infty} [k, k+1] \times \left[-\frac{\varepsilon}{2^k}, \frac{\varepsilon}{2^k}\right], \text{ con}$$

$$\sum_{k=0}^{\infty} \mu^* \left([k, k+1] \times \left[-\frac{\varepsilon}{2^k}, \frac{\varepsilon}{2^k}\right] \right) = \sum_{k=0}^{\infty} 1 \times \frac{2\varepsilon}{2^k} = 4\varepsilon.$$

Perciò $\mu^*(E) = \inf \{4\varepsilon : \varepsilon > 0\} = 0$. Analogamente si prova che il semiasse $x \leq 0$ ha misura nulla, da cui essendo unione di due insiemi di misura nulla l'asse x ha misura nulla in \mathbb{R}^2 . Naturalmente, in \mathbb{R} l'asse x avrebbe invece misura infinita.

5. Tornando al caso unidimensionale di \mathbb{R} , ci si può chiedere se solo gli insiemi numerabili hanno misura nulla. In altre parole, esistono in \mathbb{R} insiemi non numerabili di misura nulla? La domanda non è facile, perché i sottoinsiemi non numerabili di \mathbb{R} che vengono subito in mente sono gli intervalli (che sicuramente hanno misura positiva, pari alla loro lunghezza) o, volendo fare qualche esempio più complicato, l'insieme degli irrazionali, che però è misurabile e ha misura

infinita, essendo il complementare dell'insieme dei razionali, che è misurabile e ha misura nulla. Analogamente, l'insieme di tutti gli irrazionali compresi in un intervallo ha misura pari alla lunghezza dell'intervallo. Un esempio interessante di sottoinsieme non numerabile di \mathbb{R} che ha misura nulla è il ternario di Cantor, che ora definiamo.

Si esegue la seguente costruzione iterativa. Consideriamo l'intervallo $[0, 1]$. Al passo 1, rimuoviamo l'intervallo aperto centrale di lunghezza $1/3$, cioè l'intervallo $(1/3, 2/3)$, e chiamiamo T_1 l'insieme residuo, cioè $[0, 1/3] \cup [2/3, 1]$, che ha misura $2/3$.

Al passo 2, da ciascuno dei due intervallini che costituiscono T_1 si rimuove l'intervallo aperto centrale di lunghezza $1/3$ dell'intervallino da cui lo si sta togliendo, cioè di lunghezza $1/9$, e chiamiamo T_2 l'insieme residuo, che ora ha misura $4/9$.

Si prosegue iterativamente a questo modo. L'insieme T_n costruito al passo n -esimo ha misura $(2/3)^n$, perché è l'unione di 2^n intervalli ognuno di lunghezza $1/3^n$.

L'insieme ternario di Cantor è definito come $T = \bigcap_{n=1}^{\infty} T_n$. Poiché T_n è una successione decrescente di insiemi e T_1 ha misura finita, per la condizionata continuità da sopra della misura di Lebesgue si ha

$$\mu(T) = \lim_{n \rightarrow \infty} \mu(T_n) = \lim_{n \rightarrow \infty} (2/3)^n = 0.$$

Si potrebbe pensare che questo insieme T , ottenuto continuando a rimuovere intervalli da $[0, 1]$, sia un insieme vuoto o comunque molto “magro”, quindi non avremmo trovato niente di interessante. Invece T non è affatto vuoto; è un insieme chiuso (intersezione di chiusi), che contiene almeno tutti gli estremi degli intervalli residui, cioè $0, 1/3, 2/3, 1, 1/9, 2/9, \dots$, ma in realtà contiene molto di più; si può dimostrare che T è un insieme non numerabile. Per rendercene conto, ragioniamo così. Come potremmo indicare univocamente un punto di T ? Assegniamo una sequenza di istruzioni con cui individuarlo, al modo seguente: al passo 1 della costruzione, il punto si trova nell'intervallino di lunghezza $1/3$ di destra o di sinistra? Nel primo caso scriviamo 1, altrimenti scriviamo 0; al passo 2 della costruzione, il punto si trova dentro l'intervallino di lunghezza $1/3$ individuato al passo 1, e precisamente nell'intervallino di destra o di sinistra di lunghezza $1/9$? Nel primo caso scriviamo 1, altrimenti scriviamo 0. Così facendo associamo ad ogni punto del ternario di Cantor una successione di cifre 0 o 1; ma la totalità di queste successioni è un insieme non numerabile (ad esempio, perché possiamo vedere queste successioni come le rappresentazioni in base 2 dei numeri reali dell'intervallo $[0, 1]$). Ciò dimostra che T è non numerabile, pur avendo misura nulla.

6. Una generalizzazione della costruzione precedente ci sarà utile in seguito per costruire un esempio importante riguardo all'integrale di Lebesgue.

Si esegue la seguente costruzione iterativa. Consideriamo l'intervallo $[0, 1]$ e fissiamo un numero λ positivo e $\leq 1/3$. Al passo 1, rimuoviamo l'intervallo

aperto centrale di lunghezza λ , cioè l'intervallo $(\frac{1}{2} - \frac{\lambda}{2}, \frac{1}{2} + \frac{\lambda}{2})$, e chiamiamo T_1^λ l'insieme residuo, cioè $[0, \frac{1}{2} - \frac{\lambda}{2}] \cup [\frac{1}{2} + \frac{\lambda}{2}, 1]$, che ha misura $1 - \lambda$.

Al passo 2, da ciascuno dei due intervallini che costituiscono T_1^λ si rimuove l'intervallo aperto centrale di lunghezza λ volte l'intervallino da cui lo si sta togliendo, cioè di lunghezza λ^2 , e chiamiamo T_2^λ l'insieme residuo, che ora ha misura $1 - \lambda - 2\lambda^2$.

Si prosegue iterativamente a questo modo. L'insieme T_n^λ costruito al passo n -esimo ha misura

$$1 - [\lambda + 2\lambda^2 + 2^2\lambda^3 + \dots + 2^{n-1}\lambda^n] = 1 - \lambda \sum_{k=0}^{n-1} (2\lambda)^k.$$

L'insieme ternario di Cantor generalizzato è definito come $T^\lambda = \bigcap_{n=1}^{\infty} T_n^\lambda$.

Poiché T_n^λ è una successione decrescente di insiemi e T_1^λ ha misura finita, per la condizionata continuità da sopra della misura di Lebesgue si ha

$$\begin{aligned} \mu(T^\lambda) &= \lim_{n \rightarrow \infty} \mu(T_n^\lambda) = \lim_{n \rightarrow \infty} \left[1 - \lambda \sum_{k=0}^{n-1} (2\lambda)^k \right] = \\ &= 1 - \lambda \sum_{k=0}^{\infty} (2\lambda)^k = 1 - \frac{\lambda}{1 - 2\lambda} = \frac{1 - 3\lambda}{1 - 2\lambda}. \end{aligned}$$

In particolare, per ogni $\lambda \in (0, \frac{1}{3})$ il ternario di Cantor generalizzato T^λ ha misura positiva. L'insieme $T^{1/3}$ è il ternario di Cantor "classico" costruito al punto precedente, e ha misura nulla.

2.4 Restrizione di una misura

Abbiamo costruito la misura di Lebesgue su \mathbb{R}^n . Naturalmente in molte questioni non consideriamo lo spazio intero, ma un sottoinsieme Ω di \mathbb{R}^n , e vorremmo avere una misura sui sottoinsiemi di Ω . Più in generale, dato uno spazio di misura astratto $(\Omega, \mathcal{M}, \mu)$ a volte ci interessa considerare un sottoinsieme $\Omega_0 \subset \Omega$ e vorremmo essere sicuri di avere uno spazio di misura su Ω_0 che sia la restrizione di quello già considerato. Questo in effetti si può fare in generale, grazie al prossimo risultato, semplice da dimostrare ma che vale la pena di citare esplicitamente.

Teorema 15 (Restrizione di una misura) *Sia $(\Omega, \mathcal{M}, \mu)$ uno spazio di misura astratto (qualsiasi), e sia $\Omega_0 \subset \Omega, \Omega_0 \in \mathcal{M}$ (cioè Ω_0 è un sottoinsieme misurabile di Ω). Allora:*

(a) *La famiglia di insiemi $\mathcal{M}_0 = \{E_0 = E \cap \Omega_0 : E \in \mathcal{M}\}$ è una σ -algebra di sottoinsiemi di Ω_0 .*

(b) *La funzione d'insieme μ ristretta a \mathcal{M}_0 è una misura su $(\Omega_0, \mathcal{M}_0)$, quindi dà luogo ad un nuovo spazio di misura $(\Omega_0, \mathcal{M}_0, \mu_0)$ dove $\mu_0 = \mu|_{\mathcal{M}_0}$ è detta restrizione della misura μ a Ω_0 .*

La costruzione precedente è significativa quando $\mu(\Omega_0) > 0$, altrimenti la misura restrizione è identicamente nulla.

In particolare, in \mathbb{R}^n , dato un qualsiasi sottoinsieme misurabile Ω_0 di misura positiva, parleremo della misura di Lebesgue in Ω_0 per indicare la restrizione di μ a Ω_0 . Naturalmente ha interesse anche definire una misura “naturale” su insiemi come una curva o una superficie nello spazio \mathbb{R}^3 , che estenda il concetto elementare di misura di lunghezza e misura di superficie che si possono definire mediante il calcolo differenziale e integrale su curve e superfici regolari. Questo è però un altro problema, che considereremo brevemente in seguito e porterà al concetto di misura di Hausdorff k -dimensionale.

2.5 Funzioni misurabili

Torniamo ora alla teoria generale astratta. Abbiamo uno spazio di misura qualsiasi $(\Omega, \mathcal{M}, \mu)$ e vogliamo definire l'integrale rispetto a questa misura. Ricordando quanto spiegato nell'introduzione, ci attende ancora un passo preliminare, quello di definire cosa sono le funzioni misurabili, che saranno quelle per cui la costruzione dell'integrale di Lebesgue è possibile. Cominciamo dal seguente semplice

Teorema 16 *Sia (Ω, \mathcal{M}) uno spazio misurabile qualsiasi e sia $f : \Omega \rightarrow \mathbb{R}$. Sono equivalenti le seguenti 4 condizioni:*

1. $\{x \in \Omega : f(x) > a\} \in \mathcal{M}$ per ogni $a \in \mathbb{R}$;
2. $\{x \in \Omega : f(x) \geq a\} \in \mathcal{M}$ per ogni $a \in \mathbb{R}$;
3. $\{x \in \Omega : f(x) < a\} \in \mathcal{M}$ per ogni $a \in \mathbb{R}$;
4. $\{x \in \Omega : f(x) \leq a\} \in \mathcal{M}$ per ogni $a \in \mathbb{R}$.

Dimostrazione. L'equivalenza di 1 e 4 è ovvia perché i due insiemi sono l'uno il complementare dell'altro; analogamente 2 e 3 sono ovviamente equivalenti. L'equivalenza tra 1 e 2, ad esempio, si vede scrivendo

$$\{x \in \Omega : f(x) \geq a\} = \bigcap_{n=1}^{\infty} \left\{ x \in \Omega : f(x) > a - \frac{1}{n} \right\},$$

il che mostra che $2 \implies 1$, perché se vale 2 ogni insieme a 2° membro è misurabile, quindi lo è anche la loro intersezione, perciò vale 1; d'altro canto

$$\{x \in \Omega : f(x) > a\} = \bigcup_{n=1}^{\infty} \left\{ x \in \Omega : f(x) \geq a + \frac{1}{n} \right\},$$

il che mostra che $1 \implies 2$, perché se vale 1 ogni insieme a 2° membro è misurabile, quindi lo è anche la loro unione, perciò vale 2. ■

Definizione 17 *Sia (Ω, \mathcal{M}) uno spazio misurabile qualsiasi e sia $f : \Omega \rightarrow \mathbb{R}$. Si dice che f è misurabile (su (Ω, \mathcal{M})) se vale una delle condizioni equivalenti espresse dal teorema precedente.*

Esempio 18 Sia (Ω, \mathcal{M}) uno spazio misurabile qualsiasi ed $E \subset \Omega$. Definiamo la funzione caratteristica dell'insieme E (un concetto che ci servirà spesso in seguito),

$$\chi_E(x) = \begin{cases} 1 & \text{se } x \in E \\ 0 & \text{se } x \notin E. \end{cases}$$

Allora si vede facilmente che

la funzione χ_E è misurabile se e solo se l'insieme E è misurabile.

Quindi costruire esempi di funzioni non misurabili è tanto difficile quanto costruire esempi di insiemi non misurabili. Questo suggerisce che la richiesta di misurabilità di una funzione sia una richiesta generalmente molto debole.

Esempio 19 Sia Ω uno spazio metrico e sia \mathcal{M} la σ -algebra di Borel in Ω . Allora: se $f : \Omega \rightarrow \mathbb{R}$ è continua, f è misurabile. Infatti, per una funzione continua, gli insiemi ai punti 1 e 3 del teorema sopra sono aperti, mentre gli insiemi ai punti 2 e 4 del teorema sono chiusi; in tutti e 4 i casi, sono insiemi di Borel.

In particolare se $\Omega \subset \mathbb{R}^n$ e \mathcal{M} è la σ -algebra degli insiemi Lebesgue misurabili (che come sappiamo contiene la σ -algebra di Borel), se $f : \Omega \rightarrow \mathbb{R}$ è continua, f è misurabile.

Sappiamo quindi che, nel caso basilare che ci interessa, cioè quello della misura di Lebesgue in \mathbb{R}^n , le funzioni continue sono misurabili. Vediamo una serie di risultati che dicono come, a partire da funzioni misurabili, costruire altre funzioni misurabili. Prima però facciamo la seguente osservazione. Abbiamo visto come per dare la definizione di funzione misurabile non occorra avere uno spazio di misura $(\Omega, \mathcal{M}, \mu)$, ma solo uno spazio misurabile (Ω, \mathcal{M}) . Tuttavia, nel seguito della teoria ci sono alcune proprietà delle funzioni misurabili che si discutono meglio supponendo di avere già introdotto una misura. Ad esempio, se f è misurabile e g è uguale a f tranne che su un insieme “trascurabile”, vorrei poter concludere che anche g è misurabile. Questo però presuppone la presenza di una misura rispetto alla quale valutare la piccolezza dell'insieme in cui g è diversa da f .

Supponiamo quindi d'ora in poi che $(\Omega, \mathcal{M}, \mu)$ sia uno spazio di misura e la misura μ inoltre sia completa.

Illustriamo subito l'importanza di quest'ipotesi con la prossima proprietà:

Proposizione 20 Siano $f, g : \Omega \rightarrow \mathbb{R}$ con f misurabile e $g = f$ tranne che su un insieme di misura nulla. Allora g è misurabile.

Dimostrazione. Sia $E_0 = \{x \in \Omega : g(x) \neq f(x)\}$, che per ipotesi ha misura nulla. Allora, per ogni $a \in \mathbb{R}$,

$$\begin{aligned} & \{x \in \Omega : g(x) > a\} = \\ &= \{x \in \Omega : f(x) > a\} \cup \{x \in \Omega : g(x) > a \text{ e } g(x) \neq f(x)\} \\ & \quad \setminus \{x \in \Omega : f(x) > a \text{ e } g(x) \neq f(x)\} \\ &= E_1 \cup E_2 \setminus E_3 \end{aligned}$$

dove E_1 è misurabile perché f è misurabile, mentre E_2, E_3 sono misurabili perché sono sottoinsiemi dell'insieme di misura nulla E_0 , e la misura è completa. Pertanto $\{x \in \Omega : g(x) > a\}$ è misurabile per ogni $a \in \mathbb{R}$, quindi g è misurabile. ■

In particolare il teorema precedente implica che nel seguito della teoria possiamo considerare funzioni definite in Ω *salvo al più un insieme di misura nulla* o, come si dice comunemente, *definite quasi ovunque* (abbreviato in q.o.). Se f è definita q.o. in Ω , possiamo pensare di definirla in un qualsiasi modo anche nell'insieme di misura nulla residuo, e la sua misurabilità o meno in Ω non dipende da come l'abbiamo definita (quindi, in definitiva, possiamo non definirla proprio).

Teorema 21 (Operazioni sulle funzioni misurabili) *Siano $f, g : \Omega \rightarrow \mathbb{R}$ misurabili. Allora:*

$f \pm g$ è misurabile; $f \cdot g$ è misurabile; cf è misurabile (se c è una costante reale);

f/g è misurabile purché l'insieme in cui g si annulla abbia misura nulla;

$f^+ = \max(f, 0)$; $f^- = -\min(f, 0)$; $|f|$ sono misurabili;

se $\phi : \mathbb{R} \rightarrow \mathbb{R}$ è continua, $\phi(f)$ è misurabile;

se $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ è continua, $\phi(f, g)$ è misurabile.

La dimostrazione è un noioso esercizio. Il teorema precedente mostra sostanzialmente che ogni sequenza finita di operazioni su funzioni misurabili produce funzioni misurabili. Ci interessano però anche operazioni infinite, prima fra tutte il passaggio al limite: è vero che se $f_n : \Omega \rightarrow \mathbb{R}$ è una successione di funzioni misurabili convergente puntualmente a f in Ω , anche f è misurabile? Per provare questo non è facile operare direttamente sull'operazione di limite; più facile operare su quella di estremo superiore e inferiore.

Proposizione 22 *Sia $f_n : \Omega \rightarrow [-\infty, +\infty]$ (per $n = 1, 2, 3, \dots$) una successione di funzioni misurabili, allora*

$$f(x) = \sup_n f_n(x) \quad \text{e} \quad g(x) = \inf_n f_n(x)$$

(dove f, g sono anch'esse finite o infinite) sono misurabili.

La dimostrazione segue subito dal fatto che

$$\begin{aligned} \left\{x \in \Omega : \sup_n f_n(x) > a\right\} &= \bigcup_{n=1}^{\infty} \{x \in \Omega : f_n(x) > a\} \\ \left\{x \in \Omega : \inf_n f_n(x) < a\right\} &= \bigcup_{n=1}^{\infty} \{x \in \Omega : f_n(x) < a\}. \end{aligned}$$

Vogliamo ricondurre il calcolo del limite a un calcolo di sup / inf, per sfruttare il risultato precedente. Occorre per questo introdurre il concetto di *limite superiore* e *limite inferiore* di una successione, che ci servirà anche per altri motivi.

Definizione 23 Sia $\{a_n\}_{n=1}^{\infty}$ una successione di numeri reali (non necessariamente convergente). Allora la successione

$$b_n = \sup_{k \geq n} a_k$$

è monotona decrescente, quindi ammette limite (finito o $-\infty$). Si pone

$$\limsup_{n \rightarrow \infty} a_n = \inf_n \left(\sup_{k \geq n} a_k \right) = \lim_{n \rightarrow +\infty} \left(\sup_{k \geq n} a_k \right).$$

Analogamente, la successione

$$c_n = \inf_{k \geq n} a_k$$

è monotona crescente, quindi ammette limite (finito o $+\infty$). Si pone

$$\liminf_{n \rightarrow \infty} a_n = \sup_n \left(\inf_{k \geq n} a_k \right) = \lim_{n \rightarrow +\infty} \left(\inf_{k \geq n} a_k \right).$$

L'importanza dei concetti di \liminf e \limsup è che per *ogni* successione di numeri reali (limitata o illimitata, convergente o no) \liminf e \limsup esistono sempre (finiti o infiniti); tra i due numeri vale la relazione $\liminf a_n \leq \limsup a_n$; essi coincidono tra loro se e solo se la successione è convergente o divergente, e in quel caso coincidono col limite.

Esempio 24 1. Sia $a_n = (-1)^n + \frac{1}{n}$. Si ha:

$$b_n = \sup_{k \geq n} \left[(-1)^k + \frac{1}{k} \right] = \begin{cases} 1 + \frac{1}{n} & \text{se } n \text{ è pari} \\ 1 + \frac{1}{n+1} & \text{se } n \text{ è dispari} \end{cases}$$

quindi

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow +\infty} b_n = 1.$$

$$c_n = \inf_{k \geq n} \left[(-1)^k + \frac{1}{k} \right] = -1$$

quindi

$$\liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow +\infty} c_n = -1.$$

2. Sia $a_n = n \sin^2 \left(n \frac{\pi}{3} \right)$. Si ha:

$$\liminf_{n \rightarrow \infty} a_n = 0; \limsup_{n \rightarrow \infty} a_n = +\infty.$$

Se ora è f_n è una successione di funzioni misurabili in Ω , dalla misurabilità di $\sup_n f_n$ e $\inf_n f_n$ segue subito la misurabilità di $\limsup f_n$ e $\liminf f_n$, e quindi la misurabilità del limite di f_n , quando questo esiste. La completezza della misura consente di provare la seguente versione più flessibile del risultato appena annunciato:

Teorema 25 Sia $f_n : \Omega \rightarrow [-\infty, +\infty]$ (per $n = 1, 2, 3, \dots$) una successione di funzioni, ciascuna definita q.o. in Ω e misurabile, allora le funzioni

$$\liminf_{n \rightarrow \infty} f_n, \limsup_{n \rightarrow \infty} f_n$$

sono definite q.o. e misurabili. In particolare se esiste $f(x)$, limite puntuale q.o. delle $f_n(x)$, f è misurabile.

2.6 Integrazione astratta

Abbiamo ora tutti gli ingredienti per definire l'integrale rispetto ad una misura (astratta) qualsiasi. L'integrale di una qualsiasi funzione (misurabile) f sarà definito come estremo superiore o limite di opportune "somme di Lebesgue di f " (anziché somme di Cauchy-Riemann), dove queste somme (la cui costruzione è stata anticipata intuitivamente nell'introduzione) si possono vedere come integrali di opportune funzioni approssimanti, le funzioni semplici, che ora introduciamo. Il titolo "integrazione astratta" ricorda che tutta la costruzione che descriveremo in questo paragrafo vale in *qualsiasi* spazio di misura, ma naturalmente vale anche in particolare nel caso della misura di Lebesgue in \mathbb{R}^n , e questo è uno dei casi più interessanti.

Nel seguito supporremo sempre che $(\Omega, \mathcal{M}, \mu)$ sia uno spazio di misura e la misura μ sia completa.

Definizione 26 Si dice che una funzione $s : \Omega \rightarrow \mathbb{R}$ è semplice se è misurabile e assume un numero finito di valori.

Una funzione semplice si può sempre scrivere nella forma

$$s(x) = \sum_{j=1}^n c_j \chi_{E_j}(x)$$

con E_1, E_2, \dots, E_n insiemi misurabili e $c_1, c_2, \dots, c_n \in \mathbb{R}$. (Ricordare che χ_{E_j} è la funzione caratteristica di E_j , che vale 1 in E_j e 0 altrove). Gli insiemi si possono scegliere a due a due disgiunti (e allora i numeri c_j sono esattamente i possibili valori assunti da $s(x)$) ma se anche non lo sono la funzione rimane semplice. E' naturale definire l'integrale di s rispetto alla misura μ come

$$\int_{\Omega} s(x) d\mu(x) = \sum_{j=1}^n c_j \mu(E_j)$$

(pensare al caso della misura di Lebesgue sulla retta; se gli insiemi E_j sono intervalli l'integrale risulta l'area sotto il grafico della poligonale $s(x)$; naturalmente la teoria è stata fatta proprio per poter considerare le situazioni in cui gli insiemi E_j non sono intervalli ma insiemi molto complicati). L'idea è allora definire l'integrale di una funzione misurabile e positiva, per cominciare, come l'estremo

superiore degli integrali delle funzioni semplici $s(x) \leq f(x)$. Il problema è se c'è un modo standard di definire funzioni semplici che approssimano tanto bene quanto si vuole la funzione f . Questa è esattamente l'idea, che è stata anticipata nell'introduzione, di suddividere in parti uguali l'insieme dei valori assunti da f , anziché il dominio di f . La costruzione è contenuta nel prossimo

Teorema 27 (Approssimazione con funzioni semplici) *Sia $f : \Omega \rightarrow [0, +\infty]$ misurabile. Esiste una successione monotona crescente di funzioni semplici s_k che converge puntualmente a f in Ω . Se inoltre f è limitata la convergenza è uniforme.*

Dimostrazione. Fissato un intero $k = 1, 2, 3, \dots$, sia

$$\begin{aligned} E_k &= \{x \in \Omega : f(x) > k\} \text{ e siano} \\ E_k^j &= \left\{x \in \Omega : \frac{j-1}{2^k} < f(x) \leq \frac{j}{2^k}\right\} \text{ per } j = 1, 2, \dots, k2^k. \end{aligned}$$

Poniamo

$$s_k(x) = k\chi_{E_k}(x) + \sum_{j=1}^{k2^k} \frac{j-1}{2^k} \chi_{E_k^j}(x).$$

Si verifica che le s_k hanno le proprietà richieste. ■

Si può ora dare la seguente

Definizione 28 (Integrale di una funzione positiva) *Sia $f : \Omega \rightarrow [0, +\infty]$ misurabile. Si pone*

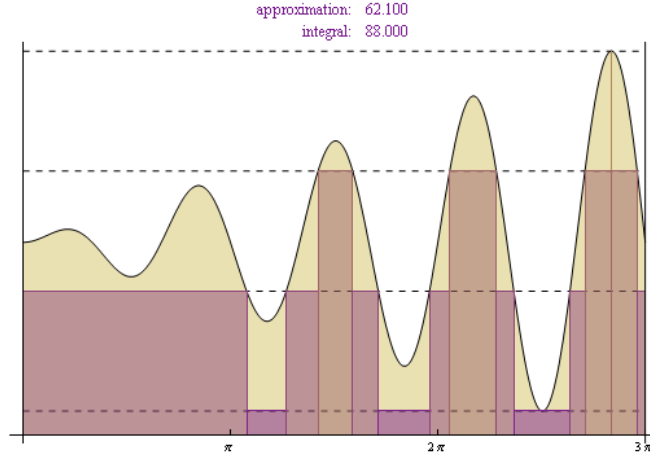
$$\int_{\Omega} f(x) d\mu(x) = \sup \left\{ \int_{\Omega} s(x) d\mu(x) : s \text{ semplice, } s(x) \leq f(x) \right\}$$

(dove l'estremo superiore può essere finito o $+\infty$).

Ci si convince facilmente che questo estremo superiore si può realizzare in particolare mediante le s_k costruite nel teorema precedente, perciò si può anche scrivere

$$\int_{\Omega} f(x) d\mu(x) = \lim_{k \rightarrow \infty} \left(k\mu(E_k) + \sum_{j=1}^{k2^k} \frac{j-1}{2^k} \mu(E_k^j) \right)$$

vedendo quindi l'integrale come un limite di somme "alla Lebesgue":



Dunque per ogni funzione misurabile e non negativa è ben definito (finito o $+\infty$) l'integrale di Lebesgue. Si noti in particolare che in questa teoria il caso in cui la funzione o il dominio sono illimitati vengono trattati direttamente e non, come accadeva per la teoria di Riemann, in un secondo tempo facendo il limite di integrali di funzioni limitate su domini limitati.

Per comprendere meglio il significato geometrico della costruzione dell'integrale di Lebesgue, consideriamo il caso particolare in cui f è misurabile, positiva e limitata, $0 \leq f(x) \leq M$. La costruzione del teorema precedente si può allora ritoccare ponendo

$$E_k^j = \left\{ x \in \Omega : \frac{j-1}{2^k} M < f(x) \leq \frac{j}{2^k} M \right\} \text{ per } j = 1, 2, \dots, 2^k$$

$$s_k(x) = M \sum_{j=1}^{2^k} \frac{j-1}{2^k} \chi_{E_k^j}(x)$$

$$\int_{\Omega} f(x) d\mu(x) = \lim_{k \rightarrow \infty} \left(M \sum_{j=1}^{2^k} \frac{j-1}{2^k} \mu(E_k^j) \right)$$

e si ha, per ogni k ,

$$M \sum_{j=1}^{2^k} \frac{j-1}{2^k} \mu(E_k^j) \leq \int_{\Omega} f(x) d\mu(x) \leq M \sum_{j=1}^{2^k} \frac{j}{2^k} \mu(E_k^j),$$

dove lo scarto tra l'approssimazione per eccesso e per difetto al passo k non supera, se Ω ha misura finita,

$$M \sum_{j=1}^{2^k} \frac{1}{2^k} \mu(E_k^j) \leq M \frac{1}{2^k} \sum_{j=1}^{2^k} \mu(E_k^j) = \frac{M}{2^k} \mu(\Omega)$$

e quindi può essere resa piccola a piacere. In particolare l'integrale in questo caso è certamente finito. Per esempio, in un intervallo di \mathbb{R} , la funzione di Dirichlet (che non è Riemann integrabile),

$$f(x) = \begin{cases} 1 & \text{per } x \in \mathbb{Q} \\ 0 & \text{per } x \notin \mathbb{Q} \end{cases}$$

è misurabile, q.o. nulla, quindi è Lebesgue integrabile con integrale nullo. Nella teoria di Lebesgue dunque, in particolare, tutte le funzioni misurabili e limitate hanno integrale finito sugli insiemi di misura finita.

Arriviamo ora alla definizione di integrale di Lebesgue per una funzione di segno qualsiasi.

Definizione 29 Sia $f : \Omega \rightarrow [-\infty, +\infty]$ misurabile. Si dice che f è Lebesgue integrabile, o sommabile, se

$$\int_{\Omega} |f(x)| d\mu(x) < \infty$$

e in tal caso si pone

$$\int_{\Omega} f(x) d\mu(x) = \int_{\Omega} f^+(x) d\mu(x) - \int_{\Omega} f^-(x) d\mu(x)$$

e risulta ovviamente

$$\left| \int_{\Omega} f(x) d\mu(x) \right| \leq \int_{\Omega} |f(x)| d\mu(x) < \infty.$$

Notare che se f è misurabile allora anche $|f|$ lo è (questa è una delle proprietà delle funzioni misurabili che abbiamo elencato), ma il viceversa non è vero: se $E \subset \Omega$ è un insieme non misurabile e definiamo

$$f(x) = \begin{cases} 1 & \text{per } x \in E \\ -1 & \text{per } x \notin E \end{cases}$$

allora f non è misurabile, mentre $|f|$ (funzione costante uguale a 1!) ovviamente lo è. Quindi nella definizione precedente è necessario richiedere la misurabilità di f e la finitezza dell'integrale di $|f|$, non è possibile esprimere le ipotesi unicamente su $|f|$.

Si confrontino le due definizioni di integrale introdotte (per funzioni positive o di segno qualsiasi): per una funzione misurabile e positiva l'integrale è sempre definito (finito o $+\infty$) ma per dar senso all'integrale di una funzione di segno variabile richiediamo la finitezza dell'integrale del modulo, che implica quella della parte positiva e negativa.

Definizione 30 Indichiamo con $L^1(\Omega, \mathcal{M}, \mu)$, o più brevemente con $L^1(\Omega)$ quando \mathcal{M} e μ si possono sottointendere, l'insieme delle funzioni sommabili nel senso della definizione precedente.

Si verifica facilmente che $L^1(\Omega, \mathcal{M}, \mu)$ è uno spazio vettoriale. Studieremo in seguito questo spazio come spazio vettoriale normato.

Occorre naturalmente provare che l'integrale di Lebesgue soddisfa "le solite" proprietà elementari dell'integrale. Premettiamo la seguente definizione, che ci serve per dar senso all'integrale di una funzione su un sottoinsieme (misurabile) di Ω :

Definizione 31 Se $E \in \mathcal{M}$, poniamo

$$\int_E f(x) d\mu(x) = \int_{\Omega} (f\chi_E)(x) d\mu(x)$$

se f è una funzione misurabile in Ω oppure è una funzione misurabile in E e noi la definiamo zero (o qualunque altro valore!) fuori da E . La definizione va intesa nel senso che se $f\chi_E \in L^1(\Omega, \mathcal{M}, \mu)$ allora diciamo che $f \in L^1(E, \mathcal{M}_E, \mu_E)$ e questa uguaglianza ne assegna l'integrale.

Teorema 32 Siano $f, g \in L^1(\Omega, \mathcal{M}, \mu)$ allora

1. *Linearità dell'integrale:* per ogni $c_1, c_2 \in \mathbb{R}$,

$$\int_{\Omega} [c_1 f(x) + c_2 g(x)] d\mu(x) = c_1 \int_{\Omega} f(x) d\mu(x) + c_2 \int_{\Omega} g(x) d\mu(x);$$

2. *Monotonia dell'integrale*

$$\begin{aligned} f(x) \leq g(x) \text{ q.o. in } \Omega &\implies \int_{\Omega} f(x) d\mu(x) \leq \int_{\Omega} g(x) d\mu(x); \\ E, F \in \mathcal{M}, E \subseteq F, f \geq 0 &\implies \int_E f(x) d\mu(x) \leq \int_F f(x) d\mu(x) \end{aligned}$$

3. *Proprietà di annullamento*

$$\begin{aligned} \text{se } \mu(E) = 0 &\quad \text{allora } \int_E f(x) d\mu(x) = 0; \\ \text{se } f(x) = 0 \text{ q.o. in } \Omega &\quad \text{allora } \int_{\Omega} f(x) d\mu(x) = 0; \\ \text{se } \int_{\Omega} |f(x)| d\mu(x) = 0 &\quad \text{allora } f(x) = 0 \text{ q.o. in } \Omega. \end{aligned}$$

Il teorema precedente si dimostra abbastanza facilmente in base alla definizione di integrale, riconducendosi al caso delle funzioni semplici.

La proprietà di *additività rispetto all'insieme di integrazione*, che non è compresa nel teorema precedente, vale in una forma più forte di quella che conosciamo per l'integrale di Riemann; precisamente, vale una *numerabile additività*, che però enunciamo solo per funzioni positive:

Teorema 33 Sia $f : \Omega \rightarrow [0, +\infty]$ misurabile e $\{E_n\}_{n=1}^\infty$ una successione di sottoinsiemi misurabili di Ω a due a due essenzialmente disgiunti⁶. Allora

$$\int_{\bigcup_{n=1}^\infty E_n} f(x) d\mu(x) = \sum_{n=1}^\infty \int_{E_n} f(x) d\mu(x).$$

Si osservi che questo teorema si può rileggere anche dicendo che una funzione misurabile $f : \Omega \rightarrow [0, +\infty]$ induce una nuova misura μ_f su (Ω, \mathcal{M}) , definita da

$$\mu_f(E) = \int_E f(x) d\mu(x). \quad (6)$$

Si dice che f è la *densità* di μ_f (rispetto alla misura originaria μ).

Serie numeriche come integrali. Consideriamo lo spazio di misura $(\Omega, \mathcal{M}, \mu)$ dove $\Omega = \mathbb{N}$, $\mathcal{M} = \mathcal{P}(\mathbb{N})$, μ è la misura del conteggio. Qualsiasi successione $\{a_n\}_{n=1}^\infty$ a valori reali si può quindi vedere come una funzione misurabile $a(n)$ su Ω ; se la serie $\sum_{n=1}^\infty a_n$ risulta assolutamente convergente, la funzione a sarà sommabile e si avrà

$$\int_{\Omega} a(n) d\mu(n) = \sum_{n=1}^\infty a_n.$$

In altre parole, le serie sono particolari integrali astratti. Nella teoria di Lebesgue quindi l'analisi del discreto e del continuo non hanno più solamente “certe analogie”, ma possono vedersi formalmente come due diverse applicazioni concrete della medesima teoria astratta. Questo è fondamentale ad esempio nelle applicazioni al Calcolo delle Probabilità, che difatti nella sua formulazione moderna, dovuta a Kolmogorov, anni 1930, è fondata sulla teoria astratta della misura.

2.7 Relazione tra integrale di Riemann e integrale di Lebesgue

Particolarizziamo ora la definizione di integrale astratto al caso della misura di Lebesgue sulla retta \mathbb{R} . In questo contesto possiamo ora confrontare le nozioni di integrale di Riemann e integrale di Lebesgue. Per poter effettuare il confronto, mettiamoci nell'insieme delle funzioni limitate definite su un intervallo $[a, b]$. Tra queste, le funzioni sommabili secondo Lebesgue sono tutte e solo le funzioni misurabili. Per confronto, si può dimostrare il seguente

Teorema 34 Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione limitata. Allora f è Riemann-integrabile se e solo se è continua quasi ovunque (dove l'espressione “quasi ovunque” ha il solito significato, quindi significa che l'insieme dei punti di discontinuità ha misura di Lebesgue nulla).

⁶ Due insiemi misurabili si dicono essenzialmente disgiunti se la loro intersezione ha misura nulla. In particolare, due insiemi disgiunti sono anche essenzialmente disgiunti.

Poiché si può dimostrare che tutte le funzioni continue quasi ovunque sono misurabili, ne segue che *tutte le funzioni Riemann integrabili⁷ sono anche Lebesgue integrabili*. L'integrale che abbiamo costruito è quindi effettivamente più generale di quello che conoscevamo dalla teoria classica.

Per esempio, come sappiamo la funzione di Dirichlet (uguale a 1 sui razionali e 0 sugli irrazionali) non è Riemann integrabile, e difatti è discontinua ovunque. Dal punto di vista della teoria di Lebesgue la funzione di Dirichlet è indistinguibile dalla funzione identicamente nulla, ovviamente integrabile.

Un esempio più interessante è il seguente:

Esempio 35 Consideriamo il ternario di Cantor generalizzato T^λ costruito nel § 2.3, con $\lambda < 1/3$, che come si ricorderà ha misura positiva, e sia

$$f : [0, 1] \rightarrow \mathbb{R}$$

$$f(x) = \chi_{T^\lambda}(x).$$

Come qualsiasi funzione caratteristica di un insieme, questa funzione è discontinua nei punti di frontiera dell'insieme stesso. D'altro canto, una delle proprietà significative di T^λ , dovuta alla sua particolare definizione iterativa, è quella di non contenere interamente alcun intervallo (a, b) (ce ne si rende facilmente conto se si pensa che ogni intervallino che costituisce T_n^λ viene spezzato in 3 intervallini più brevi al passo $n + 1$, e di questi quello centrale viene rimosso; quindi l'insieme finale T^λ non contiene alcun intervallino, per quanto breve. Ma allora T^λ non possiede alcun punto interno, perciò tutti i suoi punti sono di frontiera. Questo implica che la funzione caratteristica χ_{T^λ} è discontinua in tutti i punti di T^λ , che ha misura positiva. Pertanto, per il teorema precedente, χ_{T^λ} non è Riemann integrabile. D'altro canto χ_{T^λ} è sommabile secondo Lebesgue, con integrale pari alla misura di Lebesgue di T^λ . L'esempio è interessante perché questa funzione, a differenza della funzione di Dirichlet, non può essere resa continua alterandola su un insieme di misura nulla.

2.8 I teoremi di convergenza per l'integrale di Lebesgue

Finora abbiamo costruito l'integrale di Lebesgue, ne abbiamo elencato le proprietà di base, abbiamo constatato (ultimo paragrafo) che le funzioni Lebesgue integrabili sono più di quelle Riemann integrabili, ma non abbiamo realmente illustrato i vantaggi di questo integrale rispetto a quello classico. I vantaggi consistono soprattutto in alcuni importanti teoremi sul passaggio al limite per successioni, che qui presenteremo, e che hanno importanti conseguenze.

Torniamo ancora nel contesto astratto di un qualsiasi spazio di misura $(\Omega, \mathcal{M}, \mu)$, in cui supponiamo come in precedenza che la misura μ sia anche completa.

Teorema 36 (della convergenza monotona) Sia $f_n : \Omega \rightarrow [0, +\infty]$ una successione di funzioni misurabili, monotona crescente, cioè $f_n(x) \leq f_{n+1}(x)$

⁷In senso proprio, non generalizzato. La relazione tra integrale di Lebesgue e integrale di Riemann generalizzato è meno banale, ma non ce ne occuperemo.

per ogni intero n e $x \in \Omega$. Allora

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) d\mu(x) = \int_{\Omega} \lim_{n \rightarrow +\infty} f_n(x) d\mu(x)$$

dove i due membri dell'uguaglianza possono essere finiti o infiniti. (L'esistenza dei due limiti è parte della tesi).

Dimostrazione. Poiché una successione monotona crescente di numeri reali ha sempre limite (finito o $+\infty$), è ben definita la funzione $f : \Omega \rightarrow [0, +\infty]$

$$f(x) = \lim_{n \rightarrow +\infty} f_n(x) \text{ per ogni } x \in \Omega.$$

La funzione f è misurabile e non negativa perché limite di funzioni misurabili e non negative, quindi esiste, finito o $+\infty$, $\int_{\Omega} f(x) d\mu(x)$. Inoltre per la monotonia dell'integrale,

$$f_n(x) \leq f(x) \implies \int_{\Omega} f_n(x) d\mu(x) \leq \int_{\Omega} f(x) d\mu(x)$$

e per il teorema di permanenza del segno

$$\alpha \equiv \lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) d\mu(x) \leq \int_{\Omega} f(x) d\mu(x). \quad (7)$$

Rimane da provare la disuguaglianza opposta. Ricordando la definizione di integrale di una funzione positiva, sia $s(x)$ una qualsiasi funzione semplice $\leq f(x)$ in Ω e, fissato un qualsiasi numero $c \in (0, 1)$, consideriamo gli insiemi

$$E_n = \{x \in \Omega : f_n(x) \geq cs(x)\}.$$

Abbiamo allora per ogni intero n :

$$\alpha \geq \int_{\Omega} f_n(x) d\mu(x) \geq \int_{E_n} f_n(x) d\mu(x) \geq \int_{E_n} cs(x) d\mu(x) \quad (8)$$

D'altro canto: poiché per ogni $x \in \Omega$ è $f_n(x) \rightarrow f(x) \geq s(x)$,

$$\bigcup_{n=1}^{\infty} E_n = \Omega;$$

inoltre per la monotonia delle f_n è $E_n \subset E_{n+1}$, quindi $E_n \nearrow \Omega$ e per la continuità dal basso della misura, $\nu(E_n) \rightarrow \nu(E)$ per ogni misura su (Ω, \mathcal{M}) , in particolare per la misura di densità $cs(x)$ (v. (6)). Quindi se nella (8) passiamo al limite per $n \rightarrow \infty$ abbiamo

$$\alpha \geq c \int_{\Omega} s(x) d\mu(x).$$

Questo vale per ogni numero $c \in (0, 1)$ quindi per $c \rightarrow 1$ abbiamo

$$\alpha \geq \int_{\Omega} s(x) d\mu(x).$$

Infine, questo vale per ogni funzione semplice $s(x) \leq f(x)$, perciò passando al sup su tutte queste $s(x)$, per definizione di integrale di f abbiamo

$$\alpha \geq \int_{\Omega} f(x) d\mu(x),$$

che, ricordando (7), è quanto restava da provare. ■

Teorema 37 (di Fatou) Sia $f_n : \Omega \rightarrow [0, +\infty]$ una successione di funzioni misurabili, allora

$$\int_{\Omega} \liminf_{n \rightarrow \infty} f_n(x) d\mu(x) \leq \liminf_{n \rightarrow \infty} \int_{\Omega} f_n(x) d\mu(x)$$

dove, come nel teorema precedente, i due membri possono essere finiti o infiniti.

Si noti che in questo teorema non si suppone a priori l'esistenza di alcun limite; tuttavia, ricordiamo che *il liminf di una successione esiste sempre*. Conviene ricordarne la definizione:

$$\liminf_{n \rightarrow \infty} a_n = \sup_n \left(\inf_{k \geq n} a_k \right) = \lim_{n \rightarrow +\infty} \left(\inf_{k \geq n} a_k \right).$$

Dimostrazione. La successione di funzioni

$$\left\{ \inf_{k \geq n} f_k(x) \right\}_{n=1}^{\infty}$$

per definizione è monotona crescente, quindi possiamo applicare ad essa il teorema della convergenza monotona, e affermare che

$$\begin{aligned} \int_{\Omega} \liminf_{n \rightarrow \infty} f_n(x) d\mu(x) &= \int_{\Omega} \lim_{n \rightarrow +\infty} \left(\inf_{k \geq n} f_k(x) \right) d\mu(x) \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega} \left(\inf_{k \geq n} f_k(x) \right) d\mu(x). \end{aligned}$$

Ora è sufficiente maggiorare nell'ultimo integrale:

$$\left(\inf_{k \geq n} f_k(x) \right) \leq f_n(x).$$

Tuttavia così facendo non possiamo più essere certi che la successione degli integrali che ne risulta abbia limite; ne esiste però certamente il liminf. La catena di passaggi è quindi:

$$\begin{aligned} \int_{\Omega} \liminf_{n \rightarrow \infty} f_n(x) d\mu(x) &= \lim_{n \rightarrow +\infty} \int_{\Omega} \left(\inf_{k \geq n} f_k(x) \right) d\mu(x) = \\ &= \liminf_{n \rightarrow \infty} \int_{\Omega} \left(\inf_{k \geq n} f_k(x) \right) d\mu(x) \leq \liminf_{n \rightarrow \infty} \int_{\Omega} f_n(x) d\mu(x). \end{aligned}$$

■

I due teoremi precedenti riguardano una successione di funzioni misurabili e positive, e considerano integrali finiti o infiniti. Il prossimo teorema invece considera successioni di funzioni di segno qualsiasi, ma sommabili. Questo è probabilmente il più importante teorema della teoria di Lebesgue:

Teorema 38 (della convergenza dominata, o “teorema di Lebesgue”)

Sia $f_n : \Omega \rightarrow [-\infty, +\infty]$ una successione di funzioni misurabili, convergente puntualmente (quasi ovunque) a una certa funzione f . Supponiamo che esista una funzione g sommabile in Ω tale che per ogni intero n sia

$$|f_n(x)| \leq g(x) \text{ per q.o. } x \in \Omega.$$

Allora

$$\int_{\Omega} |f_n(x) - f(x)| d\mu(x) \rightarrow 0 \text{ per } n \rightarrow \infty.$$

In particolare,

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n(x) d\mu = \int_{\Omega} f(x) d\mu$$

cioè il limite si scambia con l'integrale.

La funzione g che compare nell'ipotesi del teorema si chiama funzione *dominante integrabile* (perché domina, cioè maggiore, il modulo delle f_n), da cui il nome del teorema.

Dimostrazione. Dall'ipotesi $|f_n(x)| \leq g(x)$ e $f_n(x) \rightarrow f(x)$ (q.o.) ricaviamo che anche $|f(x)| \leq g(x)$ (q.o.), quindi la successione

$$\phi_n = 2g - |f_n - f|$$

è non negativa (q.o.); inoltre $\phi_n(x) \rightarrow 2g(x)$ q.o.; possiamo applicare a ϕ_n il Teorema di Fatou, ottenendo

$$\begin{aligned} \int_{\Omega} 2gd\mu &= \int_{\Omega} \liminf_{n \rightarrow \infty} \phi_n d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} \phi_n d\mu \\ &= \liminf_{n \rightarrow \infty} \int_{\Omega} [2g - |f_n - f|] d\mu = \int_{\Omega} 2gd\mu + \liminf_{n \rightarrow \infty} \int_{\Omega} -|f_n - f| d\mu \end{aligned}$$

e semplificando ai due membri la quantità (finita per ipotesi) $\int_{\Omega} 2gd\mu$ abbiamo

$$0 \leq \liminf_{n \rightarrow \infty} \int_{\Omega} -|f_n - f| d\mu = -\limsup_{n \rightarrow \infty} \int_{\Omega} |f_n - f| d\mu$$

cioè

$$\limsup_{n \rightarrow \infty} \int_{\Omega} |f_n - f| d\mu = 0,$$

quindi $\int_{\Omega} |f_n(x) - f(x)| d\mu(x) \rightarrow 0$. Infine,

$$\left| \int_{\Omega} f_n(x) d\mu - \int_{\Omega} f(x) d\mu \right| = \left| \int_{\Omega} [f_n(x) - f(x)] d\mu \right| \leq \int_{\Omega} |f_n(x) - f(x)| d\mu(x)$$

perciò è anche $\int_{\Omega} f_n(x) d\mu \rightarrow \int_{\Omega} f(x) d\mu$, e la dimostrazione è completa. ■

2.9 Lo spazio L^1

Abbiamo introdotto lo studio della teoria di Lebesgue con l'esigenza di costruire uno spazio vettoriale di funzioni integrabili che, normato con una norma integrale, risulti completo. Lo scopo principale di questo paragrafo è mostrare che lo spazio $L^1(\Omega, \mathcal{M}, \mu)$ è uno spazio vettoriale normato completo. Occorre anzitutto una precisazione. Se definiamo in $L^1(\Omega, \mathcal{M}, \mu)$

$$\|f\|_{L^1(\Omega, \mathcal{M}, \mu)} = \int_{\Omega} |f(x)| d\mu(x)$$

questa risulta soddisfare le proprietà della norma, tranne quella di annullamento. Infatti, come visto nel Teorema 32,

$$\|f\|_{L^1(\Omega, \mathcal{M}, \mu)} = 0 \implies f(x) = 0 \text{ q.o. in } \Omega.$$

Per ottenere una norma è quindi necessario identificare funzioni uguali q.o. tra loro, ossia introdurre in L^1 la relazione di equivalenza $f \sim g$ se $f = g$ q.o. e considerare lo spazio delle classi di equivalenza di funzioni, che può essere reso uno spazio vettoriale normato a questo modo, come spiegato nel § 1.1. Questa operazione è coerente anche perché la misura è completa, perciò se f è misurabile e $g = f$ q.o., anche g è misurabile. Naturalmente sul piano intuitivo continueremo a pensare gli elementi di L^1 come funzioni, che sono tra loro indistinguibili quando sono uguali quasi ovunque; sul piano formale, invece, gli elementi di L^1 sono classi d'equivalenza di funzioni.

Possiamo ora, sfruttando vari fatti dimostrati fin qui, provare anzitutto un'affermazione fatta nell'introduzione, ossia:

Proposizione 39 *Lo spazio vettoriale delle funzioni limitate e Riemann-integrabili su un intervallo $[a, b]$, normato con la norma integrale*

$$\|f\|_{L^1} = \int_a^b |f(x)| dx$$

(dopo aver identificato due funzioni f, g ogni volta che $\int_a^b |f(x) - g(x)| dx = 0$) non è completo.

Dimostrazione. Consideriamo il ternario di Cantor T^λ corrispondente a un $\lambda < 1/3$, che ha misura di Lebesgue positiva, gli insiemi T_n^λ che compaiono nella sua costruzione iterativa, e poniamo

$$\begin{aligned} f_n &: [0, 1] \rightarrow \mathbb{R} \\ f_n &= \chi_{T_n^\lambda}. \end{aligned}$$

Ogni insieme T_n^λ è l'unione di un numero finito di intervalli chiusi, quindi la sua frontiera è costituita da un numero finito di punti (gli estremi di tutti questi intervalli); pertanto f_n è discontinua in un numero finito di punti, perciò è Riemann integrabile. Questa è dunque una successione di funzioni in $R[0, 1]$.

Per $n \rightarrow \infty$, $f_n(x) \rightarrow \chi_{T^\lambda}(x)$. Inoltre $|f_n(x)| \leq 1$, e la costante è integrabile in $[0, 1]$, quindi per il teorema della convergenza dominata

$$\|f_n - \chi_{T^\lambda}\|_{L^1} \rightarrow 0.$$

Perciò la successione è di Cauchy in $L^1[0, 1]$, e quindi anche in $R[0, 1]$. D'altro canto abbiamo dimostrato in precedenza che χ_{T^λ} non è Riemann integrabile. Quindi f_n è una successione di Cauchy in $R[0, 1]$ che non converge in $R[0, 1]$, e questo spazio non è completo. ■

Passiamo ora al risultato positivo che ci interessa:

Teorema 40 *Lo spazio vettoriale normato $L^1(\Omega, \mathcal{M}, \mu)$ (definito indentificando funzioni uguali q.o.) è completo.*

Per provare il teorema ci servirà il seguente risultato astratto:

Proposizione 41 *Sia X uno spazio vettoriale normato. Allora X è completo se e solo se, per ogni successione $\{x_n\}_{n=1}^\infty$ di elementi di X si ha che:*

se la serie numerica $\sum_{n=1}^\infty \|x_n\|$ converge in \mathbb{R} , allora

la serie $\sum_{n=1}^\infty x_n$ converge in X ,

ossia esiste $x \in X$ tale che

$$\left\| \sum_{k=1}^n x_k - x \right\| \rightarrow 0 \text{ per } n \rightarrow \infty.$$

Dimostrazione. Il fatto che se X è completo allora vale l'implicazione segnalata si dimostra facilmente, e lo lasciamo per esercizio anche perché questa è la parte del teorema che non ci servirà. Dimostriamo invece la parte del teorema che sfrutteremo per provare la completezza di L^1 ossia: se la convergenza della serie delle norme implica la convergenza della serie in X , allora X è completo.

Consideriamo quindi una successione $\{x_n\}_{n=1}^\infty$ di Cauchy in X , e proviamo che converge.

Per definizione di successione di Cauchy, possiamo estrarre da $\{x_n\}$ una successione $\{x_{n_k}\}$ tale che per ogni $k = 1, 2, 3, \dots$ sia

$$\|x_{n_{k+1}} - x_{n_k}\| \leq \frac{1}{k^2}.$$

Dunque $\sum_{k=1}^\infty \|x_{n_{k+1}} - x_{n_k}\|$ converge in \mathbb{R} e per la proprietà che vale per ipotesi, la serie $\sum_{k=1}^\infty (x_{n_{k+1}} - x_{n_k})$ converge in X ad un certo elemento x . Ossia:

$$\sum_{k=1}^N (x_{n_{k+1}} - x_{n_k}) \rightarrow x \text{ per } N \rightarrow \infty.$$

D'altro canto

$$\sum_{k=1}^N (x_{n_{k+1}} - x_{n_k}) = x_{n_{N+1}} - x_{n_1}$$

(somma telescopica), quindi $x_{n_{N+1}} - x_{n_1} \rightarrow x$ e $x_{n_{N+1}} \rightarrow x + x_{n_1}$ per $N \rightarrow \infty$. Abbiamo quindi provato che la successione $\{x_n\}$ contiene una sottosuccessione convergente, il che unito alla condizione di Cauchy implica⁸ la convergenza di tutta la successione $\{x_n\}$. ■

Veniamo ora alla

Dimostrazione del teorema di completezza. Proveremo che $L^1(\Omega, \mathcal{M}, \mu)$ è completo mostrando che vale la proprietà enunciata nella Proposizione precedente. Sia quindi $\{f_n\}_{n=1}^\infty$ una successione di funzioni in $L^1(\Omega, \mathcal{M}, \mu)$ di cui sappiamo che

$$\sum_{n=1}^\infty \|f_n\|_{L^1} < \infty$$

e proviamo che la serie $\sum_{n=1}^\infty f_n$ converge in $L^1(\Omega, \mathcal{M}, \mu)$. Introduciamo la successione delle somme parziali

$$g_k(x) = \sum_{n=1}^k |f_n(x)|.$$

Le funzioni g_k sono non negative e la loro successione è monotona crescente; sia

$$g(x) = \lim_{k \rightarrow \infty} g_k(x).$$

Per il teorema della convergenza monotona si ha

$$\int_{\Omega} g d\mu = \lim_{k \rightarrow \infty} \int \sum_{n=1}^k |f_n| d\mu = \lim_{k \rightarrow \infty} \sum_{n=1}^k \int |f_n| d\mu = \sum_{n=1}^\infty \int |f_n| d\mu < \infty \text{ per ipotesi.}$$

Si noti che lo scambio tra integrale e sommatoria è stato fatto su una sommatoria finita, quindi è senz'altro lecito; inoltre, il limite della somma parziale è per definizione la serie. Abbiamo dimostrato quindi che g è sommabile, e in particolare $g(x)$ è finita q.o. Ciò significa, essendo per la monotonia $\sum_{n=1}^k |f_n(x)| = g_k(x) \leq g(x)$, che la serie $\sum_{n=1}^\infty f_n(x)$ converge (per q.o. x) assolutamente e quindi puntualmente. Sia dunque

$$f(x) = \sum_{n=1}^\infty f_n(x).$$

Per la permanenza del segno si ha anche $|f(x)| \leq g(x)$, quindi anche f è sommabile.

⁸ come nella dimostrazione della completezza di \mathbb{R} .

Introduciamo ora le somme parziali della serie senza i moduli,

$$s_k(x) = \sum_{n=1}^k f_n(x).$$

Avremo $|s_k(x) - f(x)| \leq 2g(x)$, dominante integrabile, e poiché $s_k(x) \rightarrow f(x)$ puntualmente, per il teorema della convergenza dominata $\|s_k - f\| \rightarrow 0$, cioè la serie $\sum_{n=1}^{\infty} f_n$ converge in L^1 , e per il criterio della Proposizione precedente L^1 è completo. ■

Enunciamo infine il seguente fatto, che si può dimostrare combinando le tecniche delle due dimostrazioni precedenti:

Proposizione 42 *Da ogni successione $\{f_n\}$ di Cauchy in L^1 , e quindi convergente a una certa f in L^1 , si può estrarre una sottosuccessione che converge a f anche in senso puntuale quasi ovunque.*

2.10 Il teorema fondamentale del calcolo integrale

Tra le proprietà di base dell'integrale di Riemann che ancora non abbiamo rivisitato all'interno della teoria di Lebesgue ci sono i due teoremi fondamentali del calcolo integrale, espressi rispettivamente dalle uguaglianze

$$\int_a^b F'(x) dx = F(b) - F(a) \quad (9)$$

$$\frac{d}{dx} \int_a^x f(t) dt = f(x). \quad (10)$$

Cominciamo dal secondo (che in effetti viene spesso dimostrato per primo). Per l'integrale di Riemann sappiamo che l'ipotesi naturale sotto cui vale è che f sia continua in $[a, b]$; allora per ogni $x \in [a, b]$ vale la (10). Ciò significa che questa fondamentale proprietà dell'integrale di Riemann viene stabilita quando la funzione integranda appartiene ad una classe ben più ristretta di quella delle funzioni Riemann integrabili. Questa è certamente un'altra caratteristica insoddisfacente della teoria classica dell'integrazione ed in effetti, anche se noi abbiamo motivato l'introduzione dell'integrale di Lebesgue con l'esigenza di definire uno spazio di Banach di funzioni integrabili (munito di una norma integrale), storicamente la principale motivazione di Lebesgue fu proprio quella di definire un tipo di integrale per cui l'identità (10) potesse essere stabilita senza eccezioni per *tutte* le funzioni integrabili⁹. Enunciamo quindi, senza dimostrazione, il teorema che si può dimostrare nella teoria di Lebesgue:

Teorema 43 *Sia $f \in L^1[a, b]$, $x_0 \in [a, b]$ e sia*

$$F(x) = \int_{x_0}^x f(t) dt.$$

⁹Questo è proprio quanto scrive Lebesgue nella prima pagina dell'Introduzione alla sua tesi del 1902 in cui introduce la nuova teoria della misura e dell'integrazione.

Allora la funzione F è continua; inoltre per q.o. $x \in [a, b]$ la funzione F è derivabile in x e per q.o. x vale $F'(x) = f(x)$.

Qualche osservazione. Ricordiamo che una funzione $f \in L^1[a, b]$ è in realtà una classe d'equivalenza di funzioni, o se vogliamo è definita solo quasi ovunque. Tuttavia, qualunque sia la particolare funzione f che noi scegliamo come rappresentante della classe d'equivalenza, il valore dell'integrale $\int_{x_0}^x f(t) dt$ non cambia, per x_0 e x fissati. Questo significa che la funzione $F(x)$ è definita in modo univoco in ogni punto $x \in [a, b]$ (coerentemente al fatto che il teorema dimostra che F è continua: alterando una funzione continua in un insieme di misura nulla si trova in generale una funzione discontinua); invece la f può essere alterata su insiemi di misura nulla. Il teorema afferma anzitutto che questa ben definita funzione F è derivabile in quasi ogni punto di $[a, b]$; meglio di questo non si può sperare, cioè non possiamo aspettarci che F sia derivabile in ogni punto, come mostra il prossimo semplice

Esempio 44

$$f(t) = \begin{cases} 1 & \text{per } t \geq 0 \\ 0 & \text{per } t < 0 \end{cases}$$

e

$$F(x) = \int_0^x f(t) dt = \begin{cases} x & \text{per } x \geq 0 \\ 0 & \text{per } x < 0. \end{cases}$$

La funzione F non è derivabile nel punto $x = 0$.

La seconda affermazione del teorema consiste nell'uguaglianza $F'(x) = f(x)$ quasi ovunque. Di nuovo, meglio di questo non si può sperare, perché se anche $F'(x)$ esistesse ovunque, non potremmo garantire a priori che $F'(x)$ sia uguale a $f(x)$ ovunque, dal momento che il valore $f(x)$ può essere alterato in un insieme di misura nulla senza cambiare la funzione $F(x)$:

Esempio 45 Le funzioni

$$f_1(t) = 0 \text{ in } [0, 1]$$

$$f_2(t) = \begin{cases} 1 & \text{per } t \in \mathbb{Q} \\ 0 & \text{per } t \notin \mathbb{Q} \end{cases}$$

sono rappresentanti diversi della stessa classe d'equivalenza, ossia sono la stessa funzione $L^1[0, 1]$. Per questa funzione si ha

$$\begin{aligned} F(x) &= \int_0^x f_i(t) dt = 0 \quad (i = 1, 2), \text{ e quindi} \\ F'(x) &= 0. \end{aligned}$$

Come si vede $F'(x)$ è uguale quasi ovunque a $f_2(x)$ (mentre è uguale ovunque a $f_1(x)$).

Le osservazioni precedenti mostrano che il risultato contenuto nel teorema è ottimale. Del resto l'uguaglianza tra due funzioni L^1 è per definizione una uguaglianza quasi ovunque.

Veniamo ora all'altro teorema fondamentale del calcolo integrale, cioè la (9). Nella teoria di Riemann si richiede che F' sia una funzione continua, quindi che $F \in C^1[a, b]$; di nuovo, si chiede qualcosa che è più forte delle ipotesi minime sotto le quali l'integrale di F' ha senso. D'altro canto la continuità di F' non si può far cadere impunemente:

Esempio 46 *Sia*

$$F(x) = \begin{cases} 1 & \text{per } x \geq 0 \\ 0 & \text{per } x < 0. \end{cases}$$

Ovviamente esiste $F'(x) = 0$ per ogni $x \neq 0$ e la funzione F' (costante al di fuori di un unico punto in cui non è definita) è Riemann integrabile, ma l'uguaglianza (9) è falsa perché

$$\begin{aligned} \int_{-1}^1 F'(x) dx &= 0, \text{ mentre} \\ F(1) - F(-1) &= 1. \end{aligned}$$

Si osservi, per confronto, il prossimo esempio:

Esempio 47 *Sia*

$$F(x) = |x|$$

Anche in questo caso esiste $F'(x) = \operatorname{sgn}(x)$ per ogni $x \neq 0$ e la funzione F' è Riemann integrabile, anche se non è definita in $x = 0$ ed ha una discontinuità a salto. Le ipotesi del teorema fondamentale del calcolo integrale (nella teoria di Riemann) non sono verificate, quindi. Tuttavia l'uguaglianza (9) sussiste in questo caso perché

$$\begin{aligned} \int_{-1}^1 F'(x) dx &= \int_{-1}^0 -1 dx + \int_0^1 1 dx = -1 + 1 = 0, \text{ e} \\ F(1) - F(-1) &= 1 - 1 = 0. \end{aligned}$$

Questo esempio, che nel contesto della teoria di Riemann è anomalo (pur non contraddicendo alcun teorema), nella teoria di Lebesgue, come vedremo, è inquadrato in modo naturale. Per poter enunciare il primo teorema fondamentale del calcolo integrale nella teoria di Lebesgue occorre prima introdurre una nuova nozione.

Definizione 48 *Sia $f : [a, b] \rightarrow \mathbb{R}$. Si dice che f è assolutamente continua in $[a, b]$ (scriveremo $f \in AC[a, b]$) se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni successione di intervalli $[a_n, b_n]$ ($n = 1, 2, 3, \dots$), contenuti in $[a, b]$, disgiunti tra loro, e aventi lunghezza complessiva $< \delta$, ossia*

$$\sum_{n=1}^{\infty} |b_n - a_n| < \delta,$$

si ha che

$$\sum_{n=1}^{\infty} |f(b_n) - f(a_n)| < \varepsilon.$$

Questa nozione naturalmente implica la continuità, ma è più forte di essa.

Esempio 49 1. Se una funzione è Lipschitziana, cioè

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

allora è assolutamente continua (fissato ε basta scegliere $\delta = \varepsilon/L$).

2. In particolare, se $f \in C^1[a, b]$ allora $f \in AC[a, b]$ (perché ogni funzione C^1 è lipschitziana).

3. Le condizioni 1 o 2 non sono però necessarie affinché f sia assolutamente continua. Ad esempio, si dimostri per esercizio che

$$f(x) = \sqrt{x}$$

è $AC[0, 1]$ (mentre non è $Lip[0, 1]$).

3. Un esempio di funzione continua ma non assolutamente continua è il seguente, in $[0, 1]$:

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{se } x > 0 \\ 0 & \text{per } x = 0. \end{cases}$$

Infatti: sia

$$a_k = \frac{1}{\frac{3}{2}\pi + 2k\pi}; b_k = \frac{1}{\frac{1}{2}\pi + 2k\pi} \quad \text{per } k = 1, 2, 3, \dots$$

Si ha:

$$\sum_{k=1}^{\infty} |b_k - a_k| = \sum_{k=1}^{\infty} \left| \frac{1}{\frac{1}{2}\pi + 2k\pi} - \frac{1}{\frac{3}{2}\pi + 2k\pi} \right| = \sum_{k=1}^{\infty} \frac{\pi}{\left(\frac{1}{2}\pi + 2k\pi\right)\left(\frac{3}{2}\pi + 2k\pi\right)},$$

serie convergente, mentre

$$\sum_{k=1}^{\infty} |f(b_k) - f(a_k)| = \sum_{k=1}^{\infty} \left(\frac{1}{\frac{1}{2}\pi + 2k\pi} + \frac{1}{\frac{3}{2}\pi + 2k\pi} \right),$$

serie divergente. Questo significa che fissato $\varepsilon = 1$ e fissato qualsiasi $\delta > 0$, possiamo rendere

$$\sum_{k=N}^{\infty} |b_k - a_k| < \delta$$

pur di scegliere un N opportunamente grande (il resto di una serie convergente tende a zero), e tuttavia

$$\sum_{k=N}^{\infty} |f(b_k) - f(a_k)| = \infty.$$

La rilevanza del concetto di assoluta continuità nella teoria di Lebesgue è dovuta anzitutto al seguente:

Teorema 50 *Sia $f \in L^1[a, b]$, $x_0 \in [a, b]$ e*

$$F(x) = \int_{x_0}^x f(t) dt.$$

Allora $F \in AC[a, b]$.

Il Teorema 43 già affermava che F è continua e q.o. derivabile; apprendiamo ora che F è anche assolutamente continua. Poiché in base al teorema precedente è anche $F' = f$ q.o., possiamo riscrivere

$$F(x) = \int_{x_0}^x F'(t) dt$$

e scopriamo così che se, a partire da una qualsiasi funzione $f \in L^1[a, b]$, costruiamo la funzione integrale F , questa risulta essere assolutamente continua e soddisfa il primo teorema fondamentale del calcolo. Il fatto interessante è che vale anche la proprietà inversa, ossia:

Teorema 51 *Sia $f \in AC[a, b]$. Allora f è derivabile per q.o. $x \in [a, b]$, la funzione f' è $L^1[a, b]$ e inoltre per ogni $x_0, x \in [a, b]$ è*

$$f(x) - f(x_0) = \int_{x_0}^x f'(t) dt.$$

Possiamo quindi concludere che le funzioni assolutamente continue sono tutte e sole quelle che soddisfano il primo teorema fondamentale del calcolo integrale, nella teoria di Lebesgue.

3 La teoria della misura astratta in probabilità

Dal nascere della teoria della misura e dell'integrazione moderna (Lebesgue 1902, Carathéodory 1914), una delle prime applicazioni rilevanti della teoria nella sua versione astratta (in cui cioè lo spazio Ω non è necessariamente un sottoinsieme di \mathbb{R}^n e la misura è una misura "qualsiasi") è stata quella al calcolo delle probabilità. Nel 1933 N. Kolmogorov diede al calcolo delle probabilità un fondamento assiomatico che da allora è stato sistematicamente utilizzato. Lo scopo di questo paragrafo è introdurre la terminologia probabilistica comunemente utilizzata ed esemplificare, sui primi concetti e proprietà delle variabili aleatorie e del valore atteso, i vantaggi dell'utilizzo della teoria della misura astratta. Naturalmente l'importanza dell'impostazione di Kolmogorov va ben oltre le osservazioni elementari che faremo qui. Ulteriori approfondimenti in questa direzione si possono trovare in [6], cap.9.

Uno spazio di probabilità è definito come uno spazio di misura (Ω, \mathcal{M}, P) dove l'insieme Ω è lo *spazio campionario*, detto anche spazio degli *eventi elementari*, \mathcal{M} è la σ -algebra degli *eventi*, P la *misura di probabilità*, e si assume la condizione $P(\Omega) = 1$.

Una *variabile aleatoria* (abbreviato v.a.) sullo spazio Ω è una qualsiasi funzione $X : \Omega \rightarrow \mathbb{R}$, \mathcal{M} -misurabile; si usano correntemente le seguenti scritte abbreviate:

$$\{X < t\} \text{ indica l'insieme } \{\omega \in \Omega : X(\omega) < t\}, \text{ per } t \in \mathbb{R},$$

e analoghi significati hanno le scritte $\{X \leq t\}, \{X = t\}, \{X \in I\}$ per $I \subset \mathbb{R}$. Una variabile aleatoria doppia sarà una funzione $(X, Y) : \Omega \rightarrow \mathbb{R}^2$, \mathcal{M} -misurabile, nel senso che per ogni aperto A di \mathbb{R}^2 risulti $\{(X, Y) \in A\} \in \mathcal{M}$. Analogamente si definisce una n -upla di variabili aleatorie (X_1, X_2, \dots, X_n) , detta anche vettore aleatorio.

Esempio 52 Lanciamo due dadi e guardiamo quanto fa la somma dei punti. In questo caso

$$\begin{aligned} \Omega &= \{(h, k) : h, k = 1, 2, 3, 4, 5, 6\} \\ \mathcal{M} &= \mathcal{P}(\Omega) \\ P\{(h, k)\} &= \frac{1}{36} \text{ per ogni } (h, k) \end{aligned}$$

(cioè P è la probabilità uniforme su uno spazio campionario finito). La somma dei punti dei due dati è la variabile aleatoria

$$X : \Omega \rightarrow \mathbb{R}$$

$$X(h, k) = h + k.$$

La X è ovviamente misurabile perché la σ -algebra è tutto $\mathcal{P}(\Omega)$. Questo è ciò che capita sempre nel discreto: si può scegliere $\mathcal{P}(\Omega)$ come σ -algebra, per cui ogni funzione $X : \Omega \rightarrow \mathbb{R}$ è una variabile aleatoria; la necessità di usare σ -algebre non banali nasce dal continuo.

Vediamo come il formalismo della teoria astratta della misura e dell'integrazione interviene in modo importante nel concetto di *valore atteso*.

Definizione 53 Sia (Ω, \mathcal{M}, P) uno spazio di probabilità e $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria. Si definisce *valore atteso di X* il numero

$$EX = \int_{\Omega} X dP,$$

se questo è finito.

Dunque il valore atteso è definito semplicemente come integrale, ma attenzione: come integrale astratto sullo spazio campionario Ω , non in generale come integrale su \mathbb{R} .

Esempio 54 Calcoliamo in base alla definizione il valore atteso della somma del lancio dei due dadi (v. esempio precedente). Poiché Ω è finito (con 36 elementi, ciascuno di probabilità $1/36$) la X è una funzione semplice, e per definizione il suo integrale si calcola così:

$$\begin{aligned} EX &= \int_{\Omega} X dP = \sum_{h,k=1}^6 (h+k) \cdot \frac{1}{36} = \frac{1}{36} \sum_{h=1}^6 \left(\sum_{k=1}^6 (h+k) \right) = \\ &= \frac{1}{36} \sum_{h=1}^6 (6h+21) = \frac{1}{36} (6 \cdot 21 + 6 \cdot 21) = 7. \end{aligned}$$

Nelle presentazioni elementari del calcolo delle probabilità il valore atteso di una v.a. non è definito a questo modo ma, dopo aver dato separatamente le definizioni di v.a. discreta e v.a. continua, è definita in due modi diversi per questi due tipi di variabili, usando rispettivamente una serie numerica o un integrale su \mathbb{R} . Vediamo come si recuperano le definizioni elementari in questo quadro più generale. Premettiamo una semplice

Proposizione 55 Sia (Ω, \mathcal{M}, P) uno spazio di probabilità e $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria. Se, per ogni insieme di Borel $E \in \mathcal{B}(\mathbb{R})$ poniamo

$$P_X(E) = P(X \in E)$$

allora:

1. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ è uno spazio di probabilità.
2. Per ogni funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ Borel misurabile si ha:

$$\int_{\Omega} f(X) dP = \int_{\mathbb{R}} f dP_X. \quad (11)$$

3. In particolare,

$$EX = \int_{\mathbb{R}} t dP_X(t).$$

Dimostrazione. Per provare 1, osserviamo che se $\{E_n\}$ è una successione di insiemi di Borel in \mathbb{R} a due a due disgiunti, si ha:

$$\begin{aligned} P_X \left(\bigcup_{n=1}^{\infty} E_n \right) &= P \left(\left\{ X \in \bigcup_{n=1}^{\infty} E_n \right\} \right) = P \left(\bigcup_{n=1}^{\infty} \{X \in E_n\} \right) \\ &= \sum_{n=1}^{\infty} P \{X \in E_n\} = \sum_{n=1}^{\infty} P_X(E_n). \end{aligned}$$

Proviamo la 2 anzitutto per $f = \chi_E$ dove $E \in \mathcal{B}(\mathbb{R})$. Si ha:

$$\int_{\Omega} \chi_E(X) dP = P(X \in E) = P_X(E) = \int_{\mathbb{R}} \chi_E dP_X.$$

Per linearità, la (11) risulta allora verificata per ogni funzione semplice f ; per continuità, risulta verificata per ogni funzione misurabile f .

La 3 è semplicemente la 2 applicata alla funzione $f(t) = t$, perché $EX = \int_{\Omega} X dP$. ■

La misura di probabilità P_X si dice *distribuzione di X* . Chiediamoci ora che aspetto concreto ha la distribuzione di una variabile aleatoria. E' qui che si ritrova il legame con i concetti introdotti nelle presentazioni elementari del calcolo delle probabilità. Infatti: se X è una v.a. *discreta*, ossia Ω è un insieme finito o numerabile, e quindi X può assumere solo un numero finito o numerabile di valori $\{x_k\}_{k=1}^{\infty}$ (spesso in questi casi X è una variabile che *conta* oggetti di qualche tipo, e i valori x_k sono semplicemente $0, 1, 2, 3, \dots$), si ha:

$$P_X(E) = \sum_{x_k \in E} P(X = x_k).$$

I numeri $p_X(x_k) = P(X = x_k)$ definiscono completamente la distribuzione P_X . Si dice anche che la successione $\{p_X(x_k)\}_{k=1}^{\infty}$ è la *densità discreta* di X . Si ha naturalmente:

$$\begin{aligned} p_X(x_k) &\geq 0 \quad \forall k \\ \sum_{k=1}^{\infty} p_X(x_k) &= 1. \end{aligned}$$

In questo caso il valore atteso di X si calcola, per il punto 3 delle Proposizione precedente,

$$EX = \int_{\mathbb{R}} t dP_X(t) = \sum_{k=1}^{\infty} x_k p_X(x_k),$$

se questa serie converge.

Esempio 56 *Calcoliamo di nuovo il valore atteso della somma di due dadi, usando ora questa proprietà. La v.a. $X(h, k) = h + k$ assume i valori*

$$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,$$

con probabilità, rispettivamente, pari a

$$\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}.$$

Perciò

$$\begin{aligned} EX &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} \\ &\quad + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7 \end{aligned}$$

ma si osservi che in questo caso è più semplice calcolare il valore atteso come integrale astratto di X su Ω piuttosto che usando la densità discreta di X .

Una v.a. continua X su (Ω, \mathcal{M}, P) è per definizione una v.a. la cui distribuzione P_X è una misura su \mathbb{R} dotata di densità rispetto alla misura di Lebesgue in \mathbb{R} . In altre parole, esiste una funzione $f_X : \mathbb{R} \rightarrow [0, +\infty)$ Borel misurabile, detta densità di X , tale che

$$P_X(E) = \int_E f_X(t) dt \text{ per ogni } E \in \mathcal{B}(\mathbb{R}).$$

Dev'essere, naturalmente,

$$\int_{\mathbb{R}} f_X(t) dt = 1.$$

In questo caso il valore atteso si calcola, per il punto 3 della proposizione precedente:

$$EX = \int_{\mathbb{R}} t dP_X(t) = \int_{\mathbb{R}} t f_X(t) dt,$$

in quanto $dP_X(t) = f_X(t) dt$.

Nella presentazione elementare del calcolo delle probabilità, solitamente le v.a. continue sono introdotte direttamente mediante la loro densità, spesso senza neppure precisare quale sia lo spazio di probabilità soggiacente. Ad esempio, si dice: “Sia T il tempo di vita di una lampadina, e supponiamo che T abbia legge esponenziale di parametro ν ”, intendendo con questo il fatto che

$$f_T(t) = \begin{cases} \nu e^{-\nu t} & \text{per } t > 0 \\ 0 & \text{per } t \leq 0. \end{cases}$$

Si noti che con questo abbiamo in sostanza specificato come si fa a calcolare $P(T \in E)$ senza definire esplicitamente lo spazio Ω e la funzione $T : \Omega \rightarrow \mathbb{R}$. Potremmo ad esempio pensare che Ω sia l'insieme di tutte le lampadine esistenti, esistite e che esisteranno, e T sia la funzione che associa a ciascuna di queste il suo tempo di vita. Si tratterebbe di definire in qualche modo la probabilità su Ω , ma comunque $f_X(t)$ risulterebbe come sopra. Come si vede, ai fini pratici specificare lo spazio Ω è in questi casi inutile, perciò non lo si fa.

Vediamo ora come l'utilizzo della teoria dell'integrazione astratta intervenga nella dimostrazione delle proprietà del valore atteso. Ad esempio: *il valore atteso è lineare*, cioè se X, Y sono due v.a. dotate di valore atteso finito, allora

$$E(X + Y) = EX + EY.$$

Nelle presentazioni elementari della probabilità, questa proprietà viene enunciata separatamente per v.a. discrete e continue, e dimostrata separatamente (utilizzando serie numeriche o integrali). Si noti anzitutto che nell'impostazione assiomatica della probabilità il concetto di v.a. è più generale rispetto a quella di v.a. discreta o continua (la distribuzione di X potrebbe essere una misura né discreta né dotata di densità, ma ad esempio un misto delle due cose). In

secondo luogo, la *linearità del valore atteso segue semplicemente dalla linearità dell'integrale astratto*:

$$E(X + Y) = \int_{\Omega} (X + Y) dP = \int_{\Omega} X dP + \int_{\Omega} Y dP = EX + EY.$$

La dimostrazione di questa proprietà usando le densità sarebbe invece più tortuosa: nel caso continuo, ad esempio, si tratta di provare che

$$E(X + Y) = \int_{\mathbb{R}} t f_{X+Y}(t) dt = \int_{\mathbb{R}} t f_X(t) dt + \int_{\mathbb{R}} t f_Y(t) dt = EX + EY \quad (12)$$

cosa non ovvia perché $f_{X+Y}(t) \neq f_X(t) + f_Y(t)$: la densità della somma non è uguale alla somma delle densità, ma alla *convoluzione* delle densità:

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x - t) f_Y(t) dt. \quad (13)$$

Quindi bisogna prima provare la (13), e poi il fatto che da questo segua la (12).

Veniamo all'altra proprietà importante del valore atteso, legato al concetto di indipendenza.

Ricordiamo che due v.a. X_1, X_2 si dicono *indipendenti* se

$$P(X_1 \in E_1 \text{ e } X_2 \in E_2) = P(X_1 \in E_1) \cdot P(X_2 \in E_2) \quad (14)$$

per ogni coppia di insiemi di Borel E_1, E_2 in \mathbb{R} . Si può introdurre la distribuzione della v.a. doppia (X_1, X_2) come la misura di probabilità su \mathbb{R}^2 definita da

$$P_{(X_1, X_2)}(E) = P((X_1, X_2) \in E)$$

per ogni insieme di Borel $E \subset \mathbb{R}^2$.

Se X_1, X_2 sono indipendenti allora la distribuzione di (X_1, X_2) è la misura prodotto delle distribuzioni P_{X_1}, P_{X_2} , in simboli:

$$P_{(X_1, X_2)} = P_{X_1} \cdot P_{X_2}$$

che significa proprio la (14) e implica che negli integrali doppi risulta:

$$dP_{(X_1, X_2)}(t_1, t_2) = dP_{X_1}(t_1) dP_{X_2}(t_2) \quad (15)$$

Si può provare per le variabili aleatorie doppie una proposizione analoga a quella prima dimostrata, ossia: per ogni funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ Borel misurabile si ha:

$$\int_{\Omega} f(X_1, X_2) dP = \int_{\mathbb{R}^2} f dP_{(X_1, X_2)}.$$

In particolare,

$$E(X_1 X_2) = \int_{\mathbb{R}^2} t_1 t_2 dP_{(X_1, X_2)}(t_1, t_2). \quad (16)$$

Si ottiene subito allora la proprietà fondamentale sul *valore atteso del prodotto di variabili aleatorie*: se due v.a. X_1, X_2 sono indipendenti (e hanno valore atteso finito), allora

$$E(X_1 X_2) = EX_1 \cdot EX_2.$$

Infatti per (15), (16) si ha

$$\begin{aligned} E(X_1 X_2) &= \int_{\mathbb{R}^2} t_1 t_2 dP_{(X_1, X_2)}(t_1, t_2) = \int_{\mathbb{R}^2} t_1 t_2 dP_{X_1}(t_1) dP_{X_2}(t_2) \\ &= \int_{\mathbb{R}} t_1 dP_{X_1}(t_1) \int_{\mathbb{R}} t_2 dP_{X_2}(t_2) = EX_1 \cdot EX_2. \end{aligned}$$

Di nuovo, la dimostrazione di questa proprietà utilizzando le densità sarebbe meno immediata, richiederebbe tecniche diverse nel caso discreto e continuo, e non coprirebbe il caso generale di variabili né discrete né continue.

4 Misure di Hausdorff

4.1 Misura e dimensione di Hausdorff

Abbiamo visto che la misura di Lebesgue in \mathbb{R}^n estende la misura elementare, cioè ad esempio la lunghezza in \mathbb{R} , l'area in \mathbb{R}^2 e il volume in \mathbb{R}^3 . Ma, anche dal punto di vista elementare, noi siamo talvolta interessati a calcolare la lunghezza non di un sottoinsieme della retta, ma di un arco di curva nel piano o nello spazio, oppure l'area non di un sottoinsieme del piano, ma di una superficie nello spazio. In altre parole, un problema naturale che non abbiamo ancora affrontato è quello di misurare sottoinsiemi di \mathbb{R}^n aventi dimensione minore di n (come sono ad esempio le curve o le superfici regolari in \mathbb{R}^3), usando una misura appropriata alla loro dimensione (ad es. la lunghezza, cioè una *misura unidimensionale*, per una curva, anche se è immersa in uno spazio ambiente di dimensione 2 o 3). A questo problema risponde il concetto di misura di Hausdorff s -dimensionale in \mathbb{R}^n , che ora introduciamo.

Cominciamo dalla seguente precisazione: se $C \subset \mathbb{R}^n$ è un qualsiasi insieme limitato, si dice *diametro di E* il numero

$$\text{diam}(C) = \sup \{|x - y| : x, y \in E\}.$$

Definizione 57 Sia $A \subseteq \mathbb{R}^n$ un sottoinsieme qualsiasi, e s un numero reale non negativo fissato. Per ciascun $\delta \geq 0$ definiamo

$$H_\delta^s(A) = \inf \left\{ \sum_{j=1}^{\infty} \alpha(s) \left(\frac{\text{diam}(C_j)}{2} \right)^s : A \subseteq \bigcup_{j=1}^{\infty} C_j, \text{diam}(C_j) \leq \delta \right\} \quad (17)$$

dove

$$\alpha(s) = \frac{\pi^{s/2}}{\Gamma\left(\frac{s}{2} + 1\right)}$$

e Γ è la funzione di Eulero definita da $\Gamma(t) = \int_0^\infty e^{-x} x^{t-1} dx$.

Qualche spiegazione. La costante $\alpha(s)$ ha il seguente significato geometrico. Se $s = n$, la misura di Lebesgue di una sfera di raggio r in \mathbb{R}^n è uguale¹⁰ a $\alpha(n) r^n$. Se fosse $s = n$ e uno degli insiemi C_j che compaiono nella (17) fosse una sfera $B(x, r)$ di raggio r , si avrebbe

$$\alpha(n) \left(\frac{\text{diam}(C_j)}{2} \right)^n = \alpha(n) r^n = \text{vol}(B(x, r)).$$

L'idea della definizione precedente è allora la seguente. Vogliamo definire la misura s -dimensionale di un insieme A in \mathbb{R}^n . Ricopriamo per prima cosa l'insieme con unioni numerabili di insiemi generici, di cui ciò che conta è soltanto il diametro; di ognuno di questi insiemi, e quindi della loro unione, calcoliamo non il volume, ma una misura s -dimensionale, espressa dalla quantità $\alpha(s) \left(\frac{\text{diam}(C_j)}{2} \right)^s$ e dalla somma di questi numeri. Prendiamo infine l'estremo inferiore delle misure così ottenute. Per costringere però la copertura $\bigcup_{j=1}^{\infty} C_j$ a “seguire la forma” dell'insieme A , imponiamo che i diametri degli insiemi C_j siano tutti minori di un δ fissato, che immaginiamo piccolo. Spieghiamo il significato di questa richiesta col prossimo

Esempio 58 Sia $A \subset \mathbb{R}^2$ la poligonale di 2 lati aventi vertici $(1, 0)$, $(0, 0)$, $(0, 1)$ (fig.1).

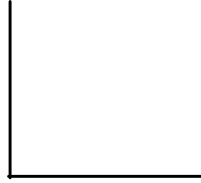


fig.1

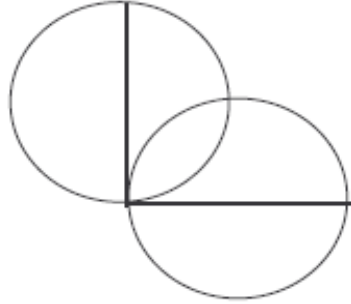


fig.2

Proviamo a misurare la lunghezza di A (che è uguale a 2) usando la precedente definizione di H_δ^1 . Se $\delta = 1$ possiamo coprire ciascun segmento con un cerchio di diametro 1 (fig. 2), e troviamo

$$A \subset C_1 \cup C_2$$

con $\text{diam}(C_1) = \text{diam}(C_2) = 1$, $\alpha(1) = 2$, perciò

$$\sum_{j=1}^2 \alpha(1) \left(\frac{\text{diam}(C_j)}{2} \right)^1 = 2 \left(\frac{1}{2} + \frac{1}{2} \right) = 2.$$

¹⁰Per la spiegazione di questo fatto e qualche dettaglio sulla funzione Gamma di Eulero, si veda [5], §§ 5.1-5-2.

Se decidessimo di ricoprire la poligonale con 4 cerchi di diametro $1/2$ ciascuno (fig.3), otterremmo

$$\sum_{j=1}^4 \alpha(1) \left(\frac{\text{diam}(C_j)}{2} \right)^1 = 2 \left(\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \right) = 2.$$

Si capisce perciò che 2 l'estremo inferiore dei valori che possiamo ottenere con questo tipo di coperture. Se non avessimo però imposto la condizione $\text{diam}(C_j) \leq \delta$ avremmo potuto ricoprire l'intera poligonale con un unico cerchio C di centro $(\frac{1}{2}, \frac{1}{2})$ e diametro $\sqrt{2}$ (fig.4), trovando

$$\alpha(1) \left(\frac{\text{diam}(C)}{2} \right)^1 = 2 \cdot \frac{\sqrt{2}}{2} = \sqrt{2},$$

un risultato inferiore, che non esprime la vera lunghezza della poligonale.

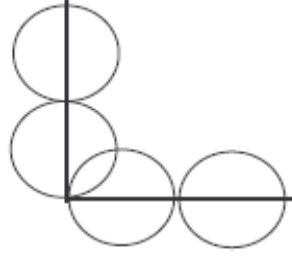


fig.3

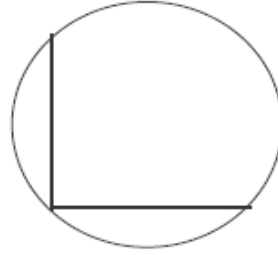


fig.4

Immaginiamo ora di ripetere la costruzione per una poligonale A i cui vertici sono $(\delta, 0), (0, 0), (0, \delta)$, e δ è molto piccolo. Per trovare la corretta lunghezza della poligonale dovremmo imporre il vincolo $\text{diam}(C_j) \leq \delta$ per questo δ . Si capisce quindi che per avere una definizione di misura s -dimensionale che funzioni per ogni insieme occorre far tendere a zero il numero δ nella definizione di H_δ^s . Questo è esattamente quello che si fa:

Definizione 59 Sia $A \subseteq \mathbb{R}^n$ un sottoinsieme qualsiasi, e s un numero reale non negativo fissato. Si pone:

$$H^s(A) = \lim_{\delta \rightarrow 0} H_\delta^s(A).$$

Il limite precedente esiste (finito o $+\infty$) perché la funzione $\delta \mapsto H_\delta^s(A)$ è decrescente: diminuendo δ diminuiscono le famiglie di coperture ammissibili, quindi l'estremo inferiore aumenta.

La quantità $H^s(A)$ viene detta *misura di Hausdorff s -dimensionale di A in \mathbb{R}^n* . A dispetto del nome, non è esattamente una misura, come ora chiariremo. Notiamo esplicitamente che questo numero è stato definito per ogni sottoinsieme A di \mathbb{R}^n e per ogni numero reale $s \geq 0$, anche non intero.

Enunciamo senza alcuna dimostrazione le proprietà fondamentali di H^s , rimandando per approfondimenti a [7], Chap.1.

Teorema 60 Per ogni $s \geq 0$, la funzione d'insieme H^s è una misura esterna su \mathbb{R}^n , in altre parole

$$H^s : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, +\infty]$$

è numerabilmente subadditiva: per ogni successione di sottoinsiemi $A_k \subset \mathbb{R}^n$ ($k = 1, 2, 3, \dots$) si ha:

$$H^s \left(\bigcup_{k=1}^{\infty} A_k \right) \leq \sum_{k=1}^{\infty} H^s(A_k).$$

Si può utilizzare la costruzione di Carathéodory già descritta nella definizione della misura di Lebesgue in \mathbb{R}^n per ottenere a partire da questa misura esterna una vera misura: è sufficiente selezionare tra tutti i sottoinsiemi di \mathbb{R}^n quelli per cui vale la condizione

$$H^s(E) = H^s(E \cap T) + H^s(E \cap T^c) \quad \forall T \subset \mathbb{R}^n.$$

Questi insiemi si diranno H^s -misurabili.

Teorema 61 Per ogni $s \geq 0$, la famiglia degli insiemi H^s -misurabili in \mathbb{R}^n è una σ -algebra, contenente la σ -algebra di Borel. La funzione H^s ristretta a questa σ -algebra è una misura, che continueremo a indicare col simbolo H^s .

Siamo quindi riusciti a definire un concetto di misura s -dimensionale per una vasta classe di sottoinsiemi di \mathbb{R}^n (in particolare, tutti i boreliani) e per ogni $s \geq 0$ (anche non intero!). Che aspetto avrà questa misura per i diversi valori di s ? Anzitutto si può provare quanto segue:

Teorema 62 Per $s = 0$ la misura H^s è la misura del conteggio.

Per $s = n$ la misura H^s è la misura di Lebesgue in \mathbb{R}^n .

Per $s > n$ la funzione H^s è identicamente nulla.

Per ogni s , la misura H^s è invariante per rototraslazioni, e soddisfa l'omogeneità:

$$H^s(\lambda E) = \lambda^s H^s(E) \quad \text{per ogni } \lambda > 0.$$

Si può anche dimostrare che per valori interi la misura H^s coincide con quanto ci aspettiamo, ad esempio: se γ è un arco di curva di classe C^1 nello spazio \mathbb{R}^3 , $H^1(\gamma)$ coincide con la lunghezza d'arco. Una delle motivazioni per introdurre la misura di Hausdorff, tuttavia, è proprio quella di poter misurare anche oggetti geometrici molto meno regolari.

Poniamoci ora dal seguente punto di vista. Immaginiamo un sottoinsieme $A \subset \mathbb{R}^n$ molto irregolare, per il quale non sia evidente qual è la dimensione naturale. In altre parole, per quale s è naturale misurare $H^s(A)$? Il prossimo teorema chiarisce la situazione che si può verificare:

Teorema 63 Sia $A \subset \mathbb{R}^n$ e $0 \leq s < t \leq n$.

1. Se $H^s(A) < \infty$, allora $H^t(A) = 0$.

2. Se $H^t(A) < \infty$, allora $H^s(A) = \infty$.

Intuitivamente: se un oggetto geometrico ha misura unidimensionale finita, ad esempio è un segmento, la sua area sarà zero; se ha area finita, ad esempio è un cerchio, la sua lunghezza (pensandola come linea) sarà infinita.

La precedente proprietà offre la possibilità di definire la dimensione di A :

Definizione 64 Sia $A \subset \mathbb{R}^n$. Si definisce *dimensione di Hausdorff* di A il numero

$$H_{\dim}(A) = \inf \{s \geq 0 : H^s(A) = 0\}.$$

Notiamo che per ogni $A \subset \mathbb{R}^n$ risulta $H_{\dim}(A) \leq n$ e che:

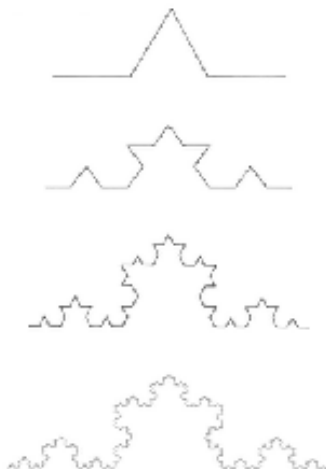
per ogni $s > H_{\dim}(A)$ è $H^s(A) = 0$, e

per ogni $s < H_{\dim}(A)$ è $H^s(A) = \infty$.

Notiamo esplicitamente che $H_{\dim}(A) = s$ non implica che $H^s(A) < \infty$, ad esempio una retta ha dimensione 1 ma misura unidimensionale infinita.

Riflettendo sugli esempi più comuni di sottoinsiemi dello spazio è naturale chiedersi: esistono insiemi la cui dimensione di Hausdorff non è intera? La risposta è affermativa, e questi insiemi si dicono *frattali*.

Esempio 65 (La curva di Koch) Consideriamo la seguente curva nel piano, costruita con un procedimento iterativo. Al passo 0 consideriamo il segmento $[0, 1]$ contenuto nell'asse x , nel piano xy . Al passo 1, dividiamo il segmento in 3 parti uguali, rimuoviamo il terzo centrale, e al suo posto inseriamo due segmenti di lunghezza $1/3$, ottenendo complessivamente una poligonale γ_1 fatta da 4 segmenti di lunghezza $1/3$ ciascuno, quindi la poligonale ha lunghezza $4/3$. Al passo 2 suddividiamo ognuno dei 4 intervallini in 3 parti uguali, rimuoviamo il terzo centrale e lo sostituiamo con 2 intervallini di lunghezza $1/9$ ciascuno, ottenendo complessivamente una poligonale γ_2 fatta da 4^2 segmenti di lunghezza $1/3^2$ ciascuno, quindi la poligonale ha lunghezza $(4/3)^2$. Procediamo iterativamente a questo modo. Al passo n la poligonale γ_n è costituita da 4^n intervallini di lunghezza $1/3^n$ ciascuno, quindi ha lunghezza totale $(4/3)^n$.



Si può dimostrare che $\{\gamma_n\}$ è una successione di curve continue nel piano, che cioè si possono parametrizzare nella forma $\{(x_n(t), y_n(t))\}_{n=1}^\infty$ con $x_n, y_n : [0, 1] \rightarrow \mathbb{R}$ continue, e la successione converge uniformemente a una curva limite $\gamma(t) = (x(t), y(t))$, continua, che si dice “curva di Koch”, o “curva fiocco di neve”. Il procedimento iterativo con cui è stata costruita γ dà alla curva limite quella caratteristica di “autosimilarità” tipica degli insiemi frattali: un particolare della curva, messo sotto il microscopio, assomiglia alla curva tutta intera. Poiché γ_n ha lunghezza $(4/3)^n$, si capisce che γ avrà lunghezza infinita. In effetti γ non è un insieme di dimensione 1. Calcoliamone la dimensione di Hausdorff. Al passo n -esimo della costruzione, il modo naturale di ricoprire γ_n in modo “accurato” consiste nel coprire ogni segmentino di lunghezza $1/3^n$ con un cerchio di diametro $1/3^n$; occorreranno in tutto 4^n cerchi C_j di questo tipo, e si otterrà, per qualsiasi s fissato,

$$\sum_{j=1}^{4^n} \alpha(s) \left(\frac{\text{diam}(C_j)}{2} \right)^s = \alpha(s) 4^n \left(\frac{1}{2 \cdot 3^n} \right)^s = \frac{\alpha(s)}{2^s} \left(\frac{4}{3^s} \right)^n.$$

Al crescere di n la copertura richiesta è sempre più fine, e occorre fare il limite per $n \rightarrow \infty$ della quantità appena scritta. La domanda ora è: per quale valore di s si ha che

$$\lim_{n \rightarrow \infty} \frac{\alpha(s)}{2^s} \left(\frac{4}{3^s} \right)^n$$

è finito e diverso da zero? Dovrà essere $\frac{4}{3^s} = 1$, cioè $s = \log_3 4$. Questo numero è la dimensione di Hausdorff della curva di Koch:

$$H_{\dim}(\gamma) = \log_3 4 \simeq 1.2619.$$

Questa dimensione è maggiore di 1, a riprova del fatto che questa curva è “talmente intricata” da essere un oggetto di dimensione maggiore della dimensione naturale di una curva; la dimensione è anche minore di 2, dimensione dello spazio ambiente. L’insieme è un frattale.

4.2 La misura $(n - 1)$ -dimensionale e il problema isoperimetrico

Tra le misure di Hausdorff, una molto usata in \mathbb{R}^n è la H^{n-1} , che è l’analogo della misura di superficie in \mathbb{R}^3 . Se ad esempio $A \subset \mathbb{R}^n$ è un aperto limitato la cui frontiera ∂A è regolare, sarà $H_{\dim}(\partial A) = n-1$, perciò H^{n-1} è la misura naturale sul bordo dei sottoinsiemi “ragionevoli” di \mathbb{R}^n . La misura H^{n-1} compare ad esempio negli integrali di bordo che sono coinvolti nel teorema della divergenza in \mathbb{R}^n (su cui ritorneremo quando studieremo gli spazi di Sobolev):

$$\int_A \text{div} \mathbf{F}(x) dx = \int_{\partial A} \mathbf{F} \cdot \mathbf{n}_e dH^{n-1}(x).$$

Un altro fatto importante che coinvolge la misura H^{n-1} è la *disuguaglianza isoperimetrica*, su cui ora spenderemo qualche parola. Un problema molto

antico¹¹ è il seguente: qual è, tra tutte le figure geometriche piane, quella che, a pari perimetro, racchiude l'area massima? La risposta, piuttosto naturale, è: il cerchio. Analogamente, tra tutti i solidi, quello che, a parità di area della superficie laterale, racchiude il volume massimo, è la sfera. Questi enunciati sono tanto intuitivi quanto difficili da dimostrare rigorosamente. Anzitutto infatti bisognerebbe chiarire chi sono “tutte le figure piane”, ossia qual è l'insieme di oggetti geometrici tra cui affermiamo che il cerchio ha la proprietà di massimo che abbiamo espresso. Di ciascuna “figura ammissibile” dobbiamo essere in grado di misurare area e perimetro, quindi la definizione della classe delle figure ammissibili dipende da quale teoria della misura si adotta. Si intuisce che la risoluzione rigorosa di questo problema non poteva essere alla portata dei greci che se lo posero 2300 anni fa. La teoria della misura del 20° secolo ha permesso di affrontare rigorosamente questo problema, anche nel caso più generale dei sottoinsiemi di \mathbb{R}^n . Come “volume” di A considereremo la misura di Lebesgue di A in \mathbb{R}^n , che indichiamo con $|A|$. Come “area della superficie laterale” di A considereremo $H^{n-1}(\partial A)$. Una formulazione quantitativa della disuguaglianza isoperimetrica si può ottenere a questo modo. Consideriamo il *quoziente adimensionato*:

$$J(A) = \frac{|A|^{(n-1)/n}}{H^{n-1}(\partial A)}.$$

Diciamo che è adimensionato in quanto $J(\lambda A) = J(A)$ per ogni dilatazione $\lambda > 0$. Il quoziente, in altre parole, dipende solo dalla *forma* dell'insieme A ma non dalle sue dimensioni assolute. La disuguaglianza isoperimetrica afferma che $J(A)$ è massimo quando A è una sfera. Calcoliamo il valore di $J(A)$ quando $A = B(0, r)$ è una sfera di raggio r . Si ha:

$$\begin{aligned} |B(0, r)| &= \alpha(n) r^n \\ H^{n-1}(\partial B(0, r)) &= \alpha(n) n r^{n-1} \end{aligned}$$

(la misura del bordo della sfera è la derivata del volume rispetto al raggio). Quindi

$$J(B(0, r)) = \frac{|B(0, r)|^{(n-1)/n}}{H^{n-1}(\partial B(0, r))} = \frac{[\alpha(n) r^n]^{(n-1)/n}}{\alpha(n) n r^{n-1}} = \frac{1}{\alpha(n)^{1/n} n}.$$

Pertanto la disuguaglianza isoperimetrica afferma:

$$\frac{|A|^{(n-1)/n}}{H^{n-1}(\partial A)} \leq \frac{1}{\alpha(n)^{1/n} n}.$$

Enunciamo un risultato preciso:

Teorema 66 *Per ogni insieme A di Borel in \mathbb{R}^n , di misura di Lebesgue $|A|$ finita, vale la disuguaglianza:*

$$H^{n-1}(\partial A) \geq \alpha(n)^{1/n} n |A|^{(n-1)/n}.$$

¹¹Risale almeno a Zenodoro, matematico greco del 3° secolo a.C.

Questo si può esprimere anche dicendo che il funzionale di insieme

$$J : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, \infty)$$

ammette massimo $1/\alpha(n)^{1/n} n$, e il massimo è assunto in corrispondenza delle sfere.

La forma in cui abbiamo enunciato questo teorema non è la migliore possibile. Il risultato preciso in questa direzione è il teorema dimostrato da Ennio De Giorgi nel 1958, che stabilisce la proprietà isoperimetrica della sfera in una classe più ampia di insiemi, precisamente quelli di *perimetro finito*, una nozione introdotta dallo stesso De Giorgi e oggi comunemente utilizzata nella teoria geometrica della misura. Per un insieme A di perimetro $P(A)$ finito si ha sempre $H^{n-1}(\partial A) \geq P(A)$, perciò la disuguaglianza di De Giorgi,

$$P(A) \geq \alpha(n)^{1/n} n |A|^{(n-1)/n}$$

implica quella che abbiamo sopra enunciato. Si osservi che il risultato di De Giorgi contiene quindi la soluzione generale di un problema che era aperto da 2300 anni, una soluzione resa possibile dagli sviluppi della teoria della misura, e successivamente della *teoria geometrica della misura*, nella prima metà del '900.

4.3 Misure e distribuzioni di carica

Abbiamo studiato il concetto astratto di misura, ma abbiamo anche incontrato, in \mathbb{R}^n , esempi concreti significativi di misura:

- la misura di Lebesgue n -dimensionale;
- le misure di Hausdorff s -dimensionali;
- le misure definite da una densità rispetto alla misura di Lebesgue, cioè $d\mu(x) = f(x) dx$ con $f \geq 0$ misurabile;
- le misure concentrate, tipo “massa di Dirac”.

Vediamo un esempio di situazione fisica in cui queste varie misure sono coinvolte. Supponiamo di studiare il problema elettrostatico di determinare il potenziale del campo generato da una certa distribuzione di cariche elettriche nello spazio. Dal punto di vista fisico, le cariche elettriche possono essere:

- cariche puntiformi;
- cariche distribuite lungo fili materiali;
- cariche distribuite su superfici materiali;
- cariche distribuite nello spazio tridimensionale (come in un gas ionizzato, o in una soluzione elettrolitica).

Nelle trattazioni matematicamente elementari queste situazioni vengono affrontate separatamente, sfruttando concetti diversi: per la carica puntiforme è comodo trattare direttamente la carica, senza introdurre una funzione densità; per una carica distribuita nello spazio si introduce una funzione densità volumica, il cui integrale triplo in una certa regione assegna la carica totale contenuta in quella regione; per fili e superfici regolari bisogna introdurre una densità lineare e superficiale, specificando che di quella funzione andrà calcolato

un integrale di linea o di superficie. Il concetto matematico di misura consente di fare unità tra tutte queste situazioni:

la distribuzione di una carica è sempre espressa da una misura μ , in modo tale che la quantità totale di carica contenuta in una regione A qualsiasi dello spazio sia pari a $\mu(A)$. Questa misura potrà essere a sua volta somma di vari addendi:

-una carica distribuita in una regione tridimensionale sarà espressa da una misura

$$f(x) dx$$

con f densità volumica, e dx misura di Lebesgue in \mathbb{R}^3 ;

-una carica distribuita lungo una superficie Σ sarà espressa da una misura di Hausdorff bidimensionale concentrata lungo Σ , quindi

$$\chi_{\Sigma}(x) \sigma(x) dH^2(x)$$

con σ densità superficiale (e χ_{Σ} è la funzione caratteristica di Σ);

-una carica distribuita lungo una curva γ sarà espressa da una misura di Hausdorff unidimensionale concentrata lungo γ , quindi

$$\chi_{\gamma}(x) \rho(x) dH^1(x)$$

con ρ densità lineare (e χ_{γ} è la funzione caratteristica di γ);

-un sistema di cariche puntiformi q_1, q_2, \dots, q_n collocate nei punti x_1, x_2, \dots, x_n sarà espresso da una combinazione lineare di misure di Dirac, del tipo

$$\sum_{i=1}^n q_i \delta_{x_i}(x).$$

In definitiva, potremo esprimere la densità di carica in tutto lo spazio con un'unica misura

$$d\mu(x) = f(x) dx + \chi_{\Sigma}(x) \sigma(x) dH^2(x) + \chi_{\gamma}(x) \rho(x) dH^1(x) + \sum_{i=1}^n q_i \delta_{x_i}(x).$$

In ogni regione A dello spazio, di forma e dimensione qualsiasi, la quantità totale di carica sarà data da $\mu(A)$.

Ricordando ora che l'equazione soddisfatta dal potenziale elettrostatico u in presenza di una densità di carica (continua) f è l'equazione di Poisson

$$-\Delta u = f,$$

possiamo pensare di formalizzare il problema della ricerca del potenziale elettrostatico dovuto ad una distribuzione di carica di tipo generale mediante un'equazione di Poisson con termine noto misura,

$$-\Delta u = \mu,$$

ma come andrà intesa quest'equazione? Cosa significherà affermare che u risolve l'equazione? Questi ed altri problemi spinsero il matematico francese Laurent Schwarz, verso il 1950, a creare la *teoria delle distribuzioni* che, prendendo a prestito il nome proprio da problemi fisici di questo tipo, introdusse una significativa generalizzazione del concetto di funzione, derivata, soluzione di un'equazione differenziale. In questo quadro funzioni e misure sono particolari distribuzioni, un concetto più generale che racchiude questi ed altri oggetti matematici.

5 Spazi L^p

5.1 Gli spazi $L^p(\Omega)$ per $1 < p < \infty$

Sia $(\Omega, \mathcal{M}, \mu)$ uno spazio di misura, con μ misura completa. Oltre allo spazio $L^1(\Omega)$ delle funzioni sommabili, è utile in molte questioni introdurre spazi di funzioni misurabili una cui opportuna potenza è sommabile. Diamo anzitutto la seguente

Definizione 67 Per $p \in [1, \infty)$, poniamo

$$L^p(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \text{ misurabile} : \int_{\Omega} |f(x)|^p d\mu(x) < \infty \right\}.$$

Per $p = 1$ ritroviamo lo spazio $L^1(\Omega)$ già introdotto, delle funzioni Lebesgue-sommabili.

Chiediamoci: la richiesta che risulti $|f|^p$ sommabile per qualche $p > 1$ è più forte o più debole rispetto a $|f|$ sommabile? In generale le due richieste non sono confrontabili:

Esempio 68 La funzione $f(x) = 1/x$ sta in $L^p(1, +\infty)$ per $p > 1$ ma non per $p = 1$; invece $f(x) = 1/x \log^2 x$ sta in $L^1(0, 1/2)$ ma non sta in $L^p(0, 1/2)$ per nessun $p > 1$.

Dunque questi spazi non soddisfano reciproche inclusioni banali, in generale. Approfondiremo in seguito questo punto.

Gli spazi $L^p(\Omega)$ sono vettoriali anche per $p > 1$ (per $p = 1$ lo sappiamo già). Infatti la combinazione lineare di funzioni misurabili è misurabile, e vale la disuguaglianza

$$|f + g|^p \leq 2^p (|f|^p + |g|^p),$$

perciò se $f, g \in L^p(\Omega)$ anche $f + g \in L^p$, con

$$\int_{\Omega} |f + g|^p d\mu \leq 2^p \left(\int_{\Omega} |f|^p d\mu + \int_{\Omega} |g|^p d\mu \right) < \infty. \quad (18)$$

Vorremmo renderli spazi vettoriali normati. E' chiaro che la norma naturale è (se vogliamo soddisfare l'omogeneità):

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

(Nel seguito scriveremo spesso $\|f\|_p$ per indicare $\|f\|_{L^p(\Omega)}$, quando non c'è pericolo di confusioni). Questa “candidata norma” soddisfa la proprietà di positività (pur di identificare come di consueto funzioni tra loro uguali quasi ovunque) e quella di omogeneità della norma. La disuguaglianza triangolare non è però immediata da dimostrare, perché dalla (18) segue

$$\begin{aligned} \left(\int_{\Omega} |f+g|^p d\mu \right)^{1/p} &\leq 2 \left(\int_{\Omega} |f|^p d\mu + \int_{\Omega} |g|^p d\mu \right)^{1/p} \\ &\leq 2 \left[\left(\int_{\Omega} |f|^p d\mu \right)^{1/p} + \left(\int_{\Omega} |g|^p d\mu \right)^{1/p} \right], \end{aligned}$$

col che avremmo ottenuto una disuguaglianza di tipo triangolare ma con una costante 2 a secondo membro, il che non va bene. Per dimostrare una disuguaglianza triangolare “pulita” occorre un argomento più raffinato. Cominciamo a dimostrare la seguente disuguaglianza, che ha molte applicazioni:

Teorema 69 (Disuguaglianza di Hölder) *Siano $p, q \in (1, \infty)$ tali che*

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

(due numeri con queste proprietà si dicono esponenti coniugati). Se $f \in L^p(\Omega)$ e $g \in L^q(\Omega)$, allora $fg \in L^1(\Omega)$ e vale la disuguaglianza (di Hölder)

$$\int_{\Omega} |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Prima di proseguire, prendiamo confidenza col concetto di *esponenti coniugati*. Affinché valga la relazione $\frac{1}{p} + \frac{1}{q} = 1$, più grande è p più piccolo dev'essere q ; in particolare, 2 è il coniugato di se stesso; se $1 < p < 2$ il suo coniugato q soddisfa $2 < q < \infty$. Ad esempio, il coniugato di $p = 3 > 2$ è $q = \frac{3}{2} < 2$. Per $p \rightarrow 1^+$, $q \rightarrow +\infty$. Dalla relazione $\frac{1}{q} = 1 - \frac{1}{p}$ ricaviamo:

$$q = \frac{p}{p-1}$$

e anche

$$pq = p + q.$$

Per provare la disuguaglianza di Hölder si fa uso della seguente disuguaglianza numerica

Lemma 70 (Disuguaglianza di Young) *Per ogni $a, b \geq 0$, si ha*

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

se p, q sono esponenti coniugati.

Dimostrazione della disuguaglianza di Young. Consideriamo la funzione di due variabili

$$f(a, b) = \frac{\frac{a^p}{p} + \frac{b^q}{q}}{ab}.$$

Nel primo quadrante $a > 0, b > 0$ la funzione è positiva, cerchiamo il suo minimo. Si trova

$$\nabla f(a, b) = \left(\frac{a^p - b^q}{qa^2b}, \frac{b^q - a^p}{pab^2} \right)$$

che si annulla dove $a^p = b^q$, punti in cui è

$$f\left(a, a^{p/q}\right) = \frac{a^p \left(\frac{1}{p} + \frac{1}{q}\right)}{aa^{p/q}} = a^{p-1-p/q} = a^0 = 1.$$

Si verifica che questo è effettivamente un punto di minimo, perciò $f(a, b) \geq 1$, che è quanto volevamo provare. ■

Dimostrazione della disuguaglianza di Hölder. Applicando la disuguaglianza di Young ai numeri:

$$a = \frac{|f(x)|}{\|f\|_p}, b = \frac{|g(x)|}{\|g\|_q}$$

per ogni $x \in \Omega$ e integrando in Ω abbiamo

$$\begin{aligned} \int_{\Omega} \frac{|fg(x)|}{\|f\|_p \|g\|_q} d\mu &\leq \int_{\Omega} \left(\frac{|f(x)|^p}{p \|f\|_p^p} + \frac{|g(x)|^q}{q \|g\|_q^q} \right) d\mu \\ &= \frac{1}{p \|f\|_p^p} \int_{\Omega} |f(x)|^p d\mu + \frac{1}{q \|g\|_q^q} \int_{\Omega} |g(x)|^q d\mu \\ &= \frac{1}{p \|f\|_p^p} \|f\|_p^p + \frac{1}{q \|g\|_q^q} \|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 \end{aligned}$$

ossia

$$\int_{\Omega} \frac{|fg(x)|}{\|f\|_p \|g\|_q} d\mu \leq 1$$

e quindi la tesi. ■

Utilizzando il risultato precedente possiamo ora dimostrare la disuguaglianza triangolare della norma L^p :

Teorema 71 (Disuguaglianza di Minkowski) Per ogni $p \in [1, +\infty)$, $f, g \in L^p(\Omega)$ vale la disuguaglianza

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Dimostrazione. Per $p = 1$ lo sappiamo già. Per $p \in (1, \infty)$ scriviamo:

$$|f + g|^p = |f + g| |f + g|^{p-1} \leq (|f| + |g|) |f + g|^{p-1}$$

e quindi

$$\|f + g\|_p^p = \int_{\Omega} |f + g|^p d\mu \leq \int_{\Omega} |f| |f + g|^{p-1} d\mu + \int_{\Omega} |g| |f + g|^{p-1} d\mu.$$

Ora in ciascuno degli integrali a secondo membro applichiamo la disuguaglianza di Hölder, sfruttando il fatto che $f, g \in L^p$ e $|f + g|^{p-1} \in L^q$ con q coniugato di p , perché

$$\begin{aligned} |f + g|^{(p-1)q} &= |f + g|^{pq-q} = |f + g|^p, \text{ perciò} \\ \left\| |f + g|^{p-1} \right\|_q &= \left(\int_{\Omega} |f + g|^{(p-1)q} d\mu \right)^{1/q} = \left(\int_{\Omega} |f + g|^p d\mu \right)^{1/q} = \|f + g\|_p^{p/q} \end{aligned}$$

Dunque:

$$\begin{aligned} \|f + g\|_p^p &\leq \|f\|_p \left\| |f + g|^{p-1} \right\|_q + \|g\|_p \left\| |f + g|^{p-1} \right\|_q \\ &= \|f + g\|_p^{p/q} (\|f\|_p + \|g\|_p) \end{aligned}$$

da cui

$$\|f + g\|_p^{p-p/q} \leq \|f\|_p + \|g\|_p$$

ossia

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

■

Dunque $L^p(\Omega)$ è uno spazio vettoriale normato per tutti i $p \in [1, \infty)$. Chiediamoci ora se c'è una ragione particolare per cui gli spazi L^p non sono stati definiti anche per $0 < p < 1$. In effetti per $0 < p < 1$ la “candidata norma” $\|f\|_p$ non soddisfa la disuguaglianza triangolare, come mostra il seguente

Esempio 72 Consideriamo $L^p[0, 2]$ per $0 < p < 1$, siano $f = \chi_{[0,1]}, g = \chi_{[1,2]}$. Si ha:

$$\begin{aligned} \left(\int_0^2 |f + g|^p dx \right)^{1/p} &= \left(\int_0^2 1 dx \right)^{1/p} = 2^{1/p}; \\ \left(\int_0^2 |f|^p dx \right)^{1/p} &= \left(\int_0^1 1 dx \right)^{1/p} = 1; \\ \left(\int_0^2 |g|^p dx \right)^{1/p} &= \left(\int_1^2 1 dx \right)^{1/p} = 1; \end{aligned}$$

dunque

$$\|f + g\|_p = 2^{1/p} > 2 = \|f\|_p + \|g\|_p$$

e la disuguaglianza triangolare viene a cadere.

Vale anche il prossimo fondamentale risultato:

Teorema 73 *Gli spazi $L^p(\Omega)$ per $1 \leq p < \infty$ sono completi.*

Per $p = 1$ lo sappiamo già. Per $1 < p < \infty$ la dimostrazione è molto simile a quella vista per $p = 1$ nel §2.9. Lo studente può provare per esercizio a modificare quella dimostrazione per adattarla al caso $p > 1$ (si tratta solo di introdurre opportunamente qualche esponente p).

5.2 Lo spazio L^∞

Introduciamo ora uno spazio che in qualche senso completa la scala degli spazi L^p .

Definizione 74 *Poniamo:*

$$L^\infty(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \text{ misurabile} : \exists K > 0 \text{ per cui } |f(x)| \leq K \text{ per q.o. } x \in \Omega\}.$$

Le funzioni di questo spazio si dicono anche essenzialmente limitate. Esse sono limitate salvo un insieme di misura nulla. Se $f \in L^\infty(\Omega)$ si definisce

$$\|f\|_{L^\infty(\Omega)} = \inf \{K > 0 : |f(x)| \leq K \text{ per q.o. } x \in \Omega\}.$$

Quindi in particolare risulta

$$|f(x)| \leq \|f\|_{L^\infty(\Omega)} \text{ per q.o. } x \in \Omega.$$

Si presti attenzione a non equivocare la definizione di “limitatezza essenziale”. I prossimi esempi dovrebbero chiarire le idee.

Esempio 75 1. *Se f è limitata in Ω in particolare è essenzialmente limitata.*

2. *La funzione $1/x$ non è essenzialmente limitata in \mathbb{R} . Attenzione: f non è “limitata per $x \neq 0$ ”, in quanto per nessun $K > 0$ è vero che $\frac{1}{|x|} \leq K$ per ogni $x \neq 0$, o per ogni x appartenente a qualche altro insieme di misura nulla.*

3. *Un esempio di funzione essenzialmente limitata senza essere limitata è:*

$$f(x) = \begin{cases} \sin x & \text{per } x \in \mathbb{R} \setminus \mathbb{Q} \\ x & \text{per } x \in \mathbb{Q} \end{cases}$$

E' chiaro infatti che

$$|f(x)| \leq 1 \text{ per q.o. } x \in \mathbb{R},$$

in quanto \mathbb{Q} è un insieme di misura nulla. D'altro canto ovviamente f è illimitata. D'altro canto ovviamente f è illimitata.

4. *f è essenzialmente limitata se e solo se f è uguale quasi ovunque ad una funzione limitata. Ad esempio, la f dell'esempio precedente è uguale quasi ovunque a $\sin x$.*

5. *La funzione $f(x) = \log x$ appartiene in $L^p(0,1)$ per ogni $p \in [1, \infty)$ ma non appartiene a $L^\infty(0,1)$; la funzione $f(x) = \sin x$ appartiene a $L^\infty(\mathbb{R})$ ma non appartiene a $L^p(\mathbb{R})$ per nessun $p \in [1, \infty)$, quindi non c'è in generale alcuna inclusione banale tra lo spazio L^∞ e gli altri spazi L^p .*

Si verifica facilmente che $L^\infty(\Omega)$ è uno spazio vettoriale normato (con la solita identificazione tra funzioni uguali quasi ovunque). Inoltre:

Teorema 76 *Lo spazio vettoriale normato $L^\infty(\Omega)$ è completo.*

Dimostrazione. Sia $\{f_n\}_{n=1}^\infty$ una successione di Cauchy in $L^\infty(\Omega)$. Questo significa che $\|f_n - f_m\|_{L^\infty(\Omega)} \rightarrow 0$ per $n, m \rightarrow \infty$. D'altro canto

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_{L^\infty(\Omega)} \text{ per ogni } x \in \Omega \setminus E_{n,m}$$

dove $E_{n,m}$ è un insieme di misura nulla. Poiché l'unione numerabile di insiemi di misura nulla ha misura nulla, fuori da un certo insieme E di misura nulla (unione di *tutti* gli insiemi $E_{n,m}$) la successione $\{f_n\}_{n=1}^\infty$ soddisfa la condizione di Cauchy per la convergenza uniforme, dunque converge uniformemente a una certa funzione f , che risulta definita e misurabile in $\Omega \setminus E$, quindi definita q.o. e misurabile in Ω . Inoltre $\|f_n - f\|_{L^\infty(\Omega)} \rightarrow 0$. ■

Notiamo che il concetto di coppia di esponenti coniugati e la disuguaglianza di Hölder si estendono a comprendere anche lo spazio L^∞ , al seguente modo:

Definizione 77 *Diciamo che l'esponente coniugato di $p = 1$ è $q = \infty$, e l'esponente coniugato di $p = \infty$ è $q = 1$.*

Ora, se $f \in L^1(\Omega)$ e $g \in L^\infty(\Omega)$ si ha, essendo $|g(x)| \leq \|g\|_{L^\infty(\Omega)}$ per q.o. $x \in \Omega$,

$$\int_{\Omega} |fg(x)| d\mu(x) \leq \int_{\Omega} \|g\|_{L^\infty(\Omega)} |f(x)| d\mu(x) = \|g\|_{L^\infty(\Omega)} \|f\|_{L^1(\Omega)}.$$

Perciò la disuguaglianza di Hölder si estende anche ai casi estremi in cui $p = 1$ o ∞ .

5.3 Inclusioni tra spazi L^p

Cominciamo a dimostrare il seguente risultato positivo di inclusione tra spazi L^p su un dominio di misura finita:

Teorema 78 *Sia $(\Omega, \mathcal{M}, \mu)$ uno spazio di misura (con μ misura completa), di misura finita, cioè $\mu(\Omega) < \infty$. Se $1 \leq p < r \leq \infty$ allora $L^r(\Omega) \subset L^p(\Omega)$ e vale la disuguaglianza:*

$$\|f\|_p \leq \mu(\Omega)^{\frac{1}{p} - \frac{1}{r}} \|f\|_r \text{ per ogni } f \in L^r(\Omega)$$

dove nel caso $r = \infty$ a secondo membro abbiamo $\mu(\Omega)^{\frac{1}{p}} \|f\|_\infty$.

Dimostrazione. Supponiamo prima $r < \infty$. Applichiamo la disuguaglianza di Hölder, scrivendo:

$$\|f\|_p^p = \int_{\Omega} |f(x)|^p \cdot 1 d\mu(x)$$

e osservando che se $f \in L^r(\Omega)$, allora $|f|^p \in L^{r/p}(\Omega)$ e la costante 1 appartiene allo spazio $L^s(\Omega)$ con s coniugato di r/p , quindi $1/s = 1 - p/r$. Perciò:

$$\|f\|_p^p \leq \left(\int_{\Omega} |f(x)|^{p \cdot \frac{r}{p}} d\mu(x) \right)^{\frac{p}{r}} \left(\int_{\Omega} 1^s d\mu \right)^{1 - \frac{p}{r}} = \|f\|_r^p \mu(\Omega)^{1 - p/r}$$

ed elevando ambo i membri alla $1/p$ si ha la tesi. Nel caso $r = \infty$ si ha subito

$$\left(\int_{\Omega} |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} \leq \|f\|_{\infty} \left(\int_{\Omega} 1 d\mu(x) \right)^{\frac{1}{p}} = \mu(\Omega)^{\frac{1}{p}} \|f\|_{\infty}.$$

■

Se $\mu(\Omega) = \infty$ abbiamo già visto con semplici esempi che non valgono inclusioni banali tra gli spazi L^p . Dati due esponenti $p, r \in [1, \infty]$ diversi tra loro, esistono $f \in L^p(\Omega) \setminus L^r(\Omega)$ e $g \in L^r(\Omega) \setminus L^p(\Omega)$. Qui sotto si riporta qualche esempio in tal senso. Si può dimostrare che, ad ogni modo, se $f \in L^p(\Omega) \cap L^r(\Omega)$ per due esponenti p, r con $1 \leq p < r \leq \infty$, allora $f \in L^s(\Omega)$ per ogni $s \in (p, r)$. Ad esempio, se $f \in L^1(\Omega) \cap L^{\infty}(\Omega)$ allora $f \in L^s(\Omega)$ per ogni $s \in [1, \infty]$.

Esempio 79 1. $f(x) = \frac{1}{\sqrt{x}} \in L^p(0, 1)$ per ogni $p \in [1, 2)$ ma non per $p \geq 2$.

2. $f(x) = \frac{1}{\sqrt{x}} \in L^p(1, +\infty)$ per ogni $p \in (2, +\infty]$ ma non per $p \leq 2$.

3. $f(x) = \frac{1}{\sqrt{x} \log x} \in L^p(0, 1/2)$ per ogni $p \in [1, 2]$ ma non per $p > 2$.

4. $f(x) = \frac{1}{\sqrt{x} \log x} \in L^p(2, \infty)$ per ogni $p \in [2, +\infty]$ ma non per $p < 2$.

5. $f(x) = \frac{1}{\sqrt{x}(1+|\log x|)} \in L^p(0, \infty)$ se e solo se $p = 2$.

6. Se $r \in [1, \infty)$, $f(x) = \frac{1}{x^{1/r}(1+|\log x|^2)} \in L^p(0, \infty)$ se e solo se $p = r$.

7. $f(x) = \sin x \in L^p(0, \infty)$ se e solo se $p = \infty$.

8. Passando da \mathbb{R} a \mathbb{R}^n , chiediamoci ora, per $\alpha > 0$ fissato, per quali p è $f(x) = \frac{1}{|x|^{\alpha}} \in L^p(B_1)$, con $B_1 = \{x \in \mathbb{R}^n : |x| < 1\}$. Con cambio di variabili polari si ha:

$$\begin{aligned} \int_{|x|<1} \frac{dx}{|x|^{\alpha p}} &= \int_0^1 \frac{1}{\rho^{\alpha p}} c_n \rho^{n-1} d\rho = c_n \int_0^1 \frac{d\rho}{\rho^{\alpha p - n + 1}} < 1 \text{ per:} \\ \alpha p - n + 1 &< 1, \text{ cioè} \\ 1 &\leq p < \frac{n}{\alpha}, \text{ purché } \alpha < n. \end{aligned}$$

Per $\alpha \geq n$, $f \notin L^p(B_1)$ per nessun p .

9. Analogamente chiediamoci per quali p è $f(x) = \frac{1}{|x|^{\alpha}} \in L^p(\mathbb{R}^n \setminus B_1)$. Si ha ora:

$$\begin{aligned} \int_{|x|\geq 1} \frac{dx}{|x|^{\alpha p}} &= c_n \int_1^{+\infty} \frac{d\rho}{\rho^{\alpha p - n + 1}} < 1 \text{ per:} \\ \alpha p - n + 1 &> 1, \text{ cioè} \\ p &> \frac{n}{\alpha} \text{ se } \alpha \leq n \\ p &\geq 1 \text{ se } \alpha > n. \end{aligned}$$

5.4 Approssimazione di funzioni L^p

Tutto quanto abbiamo detto fin qui per gli spazi L^p vale su un generico spazio di misura $(\Omega, \mathcal{M}, \mu)$ (con μ completa). Ora ci poniamo un problema che tipicamente ha senso quando Ω è un sottoinsieme misurabile di \mathbb{R}^n e μ la misura di Lebesgue in \mathbb{R}^n : è possibile approssimare in norma $L^p(\Omega)$ una funzione $f \in L^p(\Omega)$ con funzioni continue, oppure regolari (cioè derivabili un certo numero di volte)? Vale in proposito un risultato molto forte, che ci limitiamo ad enunciare. Sia Ω un aperto di \mathbb{R}^n . Consideriamo lo spazio $C_0^\infty(\Omega)$ delle funzioni f derivabili infinite volte in Ω e aventi supporto compatto in Ω , ossia tali che esiste un insieme chiuso e limitato $K \subset \Omega$ per cui $f(x) = 0$ fuori da K (l'insieme K dipende dalla funzione f). Si osservi che $C_0^\infty(\Omega)$ è uno spazio vettoriale, anche se non possiede una norma naturale. E' ovviamente contenuto in $L^p(\Omega)$ per ogni $p \in [1, \infty]$; infatti per il teorema di Weierstrass f è limitata e si annulla fuori da un insieme limitato, quindi di misura finita. Con queste premesse, si ha allora:

Teorema 80 *Lo spazio $C_0^\infty(\Omega)$ è denso in $L^p(\Omega)$ per ogni $p \in [1, \infty)$.*

Esplicitamente, la densità significa che per ogni $f \in L^p(\Omega)$ con $p \in [1, \infty)$ esiste una successione $\{\varphi_n\}_{n=1}^\infty \subset C_0^\infty(\Omega)$ tale che $\|\varphi_n - f\|_p \rightarrow 0$ per $n \rightarrow \infty$. A maggior ragione, quindi, ogni funzione $L^p(\Omega)$ con $p \in [1, \infty)$ può essere approssimata bene quanto vogliamo in norma L^p mediante funzioni continue in $\bar{\Omega}$ o con funzioni $C^\infty(\bar{\Omega})$. E' utile talvolta anche un risultato di approssimazione di tipo diverso, non mediante funzioni regolari ma mediante funzioni semplici:

Teorema 81 *Lo spazio delle funzioni semplici in Ω (cioè misurabili e che assumono un numero finito di valori) è denso in $L^p(\Omega)$ per ogni $p \in [1, \infty)$.*

Usando i teoremi precedenti e qualche altro argomento si può dimostrare in particolare che:

Teorema 82 *Lo spazio $L^p(\Omega)$ è separabile per tutti i $p \in [1, \infty)$.*

Si osservi esplicitamente che i teoremi precedenti escludono lo spazio $L^\infty(\Omega)$. Infatti, si può dimostrare che lo spazio $L^\infty(\Omega)$ non è separabile, e non è possibile approssimare in norma L^∞ una qualunque funzione L^∞ mediante funzioni continue o regolari.

5.5 Convoluzione in \mathbb{R}^n

Un'operazione che riguarda gli spazi $L^p(\mathbb{R}^n)$ è la *convoluzione*, di cui ci limitiamo a segnalare la definizione e un paio di proprietà.

Teorema 83 *Se $f, g \in L^1(\mathbb{R}^n)$, allora l'integrale di convoluzione*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y) g(y) dy$$

è ben definito per q.o. $x \in \mathbb{R}^n$, e risulta:

$$\|f * g\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$

Dimostrazione. Tralasciamo di dimostrare il fatto che la funzione di due variabili $f(x-y)g(y)$ è misurabile. Posto questo, proviamo la disuguaglianza delle norme, da cui in particolare segue che la convoluzione è definita per quasi ogni x .

$$\|f * g\|_{L^1(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y)g(y)dy \right| dx \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)g(y)| dy dx$$

scambiando l'ordine di integrazione nell'integrale iterato

$$= \int_{\mathbb{R}^n} |g(y)| \left(\int_{\mathbb{R}^n} |f(x-y)| dx \right) dy.$$

Ora l'integrale interno è

$$\begin{aligned} \int_{\mathbb{R}^n} |f(x-y)| dx &= \text{(traslando } x-y=t) \\ &= \int_{\mathbb{R}^n} |f(t)| dt = \|f\|_{L^1(\mathbb{R}^n)}, \end{aligned}$$

perciò

$$\|f * g\|_{L^1(\mathbb{R}^n)} \leq \int_{\mathbb{R}^n} |g(y)| \|f\|_{L^1(\mathbb{R}^n)} dy = \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$

■

Qualche osservazione. Si noti anzitutto che la definizione di convoluzione ha senso tra due funzioni definite su *tutto lo spazio* \mathbb{R}^n . Si verifica con un cambio di variabile che la convoluzione è commutativa, ossia $(f * g) = (g * f)$. Si noti che l'integranda $f(x-y)g(y)$, per ogni x fissato si può vedere come una funzione di y che è il prodotto di due funzioni L^1 , quindi *in generale* non è L^1 . Tuttavia il teorema appena dimostrato afferma che *per quasi ogni* x in effetti è L^1 . Ad esempio, se

$$f(x) = \frac{1}{\sqrt{|x|}} \chi_{(-1,1)}(x),$$

la funzione $f \in L^1(\mathbb{R})$, quindi $f * f \in L^1(\mathbb{R})$ e in particolare l'integrale

$$(f * f)(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{|x-y|}} \chi_{(-1,1)}(x-y) \frac{1}{\sqrt{|y|}} \chi_{(-1,1)}(y) dy$$

è finito per q.o. x . Però per $x = 0$ si ha

$$(f * f)(0) = \int_{\mathbb{R}} \frac{1}{\sqrt{|y|}} \chi_{(-1,1)}(-y) \frac{1}{\sqrt{|y|}} \chi_{(-1,1)}(y) dy = \int_{-1}^1 \frac{dy}{|y|} = \infty.$$

E' facile verificare direttamente che per ogni $x \neq 0$ invece l'integrale converge.

Segnaliamo anche la seguente estensione del teorema precedente:

Teorema 84 (Disuguaglianza di Young) *Siano $f \in L^1(\mathbb{R}^n)$ e $g \in L^p(\mathbb{R}^n)$ per qualche $p \in [1, \infty]$. Allora è ben definita per q.o. x la convoluzione $f * g$ e si ha:*

$$\|f * g\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^p(\mathbb{R}^n)}.$$

Riferimenti bibliografici

- [1] V. Pata: Analisi reale e funzionale. Dispensa in rete:
http://web.mate.polimi.it/viste/studenti/send.phptipo=a&id=121&id_insegnamento=419&filename=ARF.pdf
- [2] K. Saxe. Beginning functional analysis. Springer 2010.
- [3] W. Rudin: Real and complex analysis. McGraw-Hill Book Co., New York, 1987.
- [4] M. Bramanti: Esercitazioni di Analisi 3. CUSL, 1993.
- [5] M. Bramanti, C. D. Pagani, S. Salsa: Analisi Matematica 2. Zanichelli, 2009.
- [6] G. Folland: Real Analysis. Modern applications and techniques. Wiley, 1984.
- [7] L. Fanghua, Y. Xiaoping: Geometric measure theory. An introduction. Science Press, 2002.