

Appunti del corso di
Metodi matematici per l'Ingegneria

Simona Fornaro e Diego Pallara

a.a. 2008/09

Introduzione

In queste dispense sono esposti gli argomenti che costituiscono la prima parte del corso di Metodi matematici per l'ingegneria rivolto agli studenti del primo anno della Laurea specialistica in Ingegneria delle telecomunicazioni. Malgrado gli argomenti siano classici e trovino posto in numerosi libri, alcuni dei quali sono consigliati per la preparazione dell'esame, ci auguriamo che la presente sintetica esposizione possa essere utile come guida alla lettura di testi più approfonditi. Il fascicolo contiene probabilmente più di quanto sia poi esposto in aula (e richiesto all'esame), perché, essendo rivolto a studenti maturi, si è voluta dare la possibilità di qualche approfondimento individuale.

Questo testo non ha carattere definitivo, ed anzi contiamo di poterlo migliorare, giovandoci anche delle osservazioni di tutti i lettori, ed in particolare degli studenti. Ringraziamo in anticipo tutti coloro che ci forniranno suggerimenti utili.

Indice

1	Teoria della misura e dell'integrazione	7
1.1	Misure positive	8
1.2	Funzioni misurabili	12
1.3	Integrazione	15
1.4	La misura di Lebesgue in $\mathbb{R}^n$	18
1.5	Passaggio al limite sotto il segno di integrale	22
1.6	Misure reali e vettoriali	29
1.7	Singolarità e assoluta continuità di misure	32
1.8	Misura prodotto	35
1.9	Formule di riduzione e cambiamento di variabili in $\mathbb{R}^n$	40
1.10	Integrali dipendenti da un parametro	44
1.11	Prodotto di convoluzione	47
2	Teoria delle distribuzioni	51
2.1	Definizioni ed esempi	51
2.2	Successioni e serie di distribuzioni	55
2.3	Derivazione	56
2.4	Supporto e convoluzione	59
2.5	Distribuzioni temperate	62
3	Misure di Lebesgue-Stieltjes e funzioni a variazione limitata	65
3.1	Misura di Lebesgue-Stieltjes	65
3.2	Misura immagine e applicazioni nella teoria delle probabilità	68
3.3	Funzioni a variazione limitata	70
3.4	Funzioni assolutamente continue	76
4	Spazi di Hilbert e serie di Fourier	79

A Appendice – Richiami	93
A.1 Massimo e minimo limite	93
A.2 Insiemi numerabili	94
A.3 Il teorema di Riemann sui riordinamenti	95

Capitolo 1

Teoria della misura e dell'integrazione

In questo capitolo svilupperemo una teoria dell'integrazione più moderna di quella di Riemann, già studiata, che consente di ovviare ad alcuni inconvenienti. I principali difetti della teoria di Riemann sono infatti:

- (i) La non completezza dello spazio $\mathcal{R}([a, b])$ delle funzioni integrabili in $[a, b]$, munito della metrica

$$d(f, g) = \int_a^b |f(x) - g(x)| dx.$$

Questo vuol dire che esistono successioni di funzioni (f_k) integrabili su $[a, b]$ tali che

$$\lim_{k, j \rightarrow +\infty} \int_a^b |f_k - f_j| dx = 0,$$

ma non esiste alcuna funzione integrabile f tale che

$$\lim_{k \rightarrow +\infty} \int_a^b |f_k - f| dx = 0.$$

- (ii) Il fatto che non tutte le funzioni caratteristiche di aperti o chiusi siano integrabili, a meno che l'aperto o il chiuso non siano costituiti da un numero finito di intervalli. Inoltre, in generale le funzioni definite mediante serie non sono integrabili secondo Riemann, a meno che la somma della serie non sia continua.

(iii) La possibilità di passare al limite sotto il segno di integrale

$$f_h \rightarrow f \quad \implies \quad \int_a^b f_h(t) dt \rightarrow \int_a^b f(t) dt$$

solo se le funzioni convergono uniformemente. Vedremo invece che nell'ambito della teoria di Lebesgue la classe delle funzioni integrabili è molto più ricca e comprende molte funzioni che sono limiti di successioni anche non uniformemente convergenti. Inoltre, vi sono formule molto generali di passaggio al limite sotto il segno di integrale.

L'idea di base della teoria di Lebesgue consiste nel suddividere non il *dominio* della funzione da integrare, come nella teoria di Riemann, ma il suo *codominio*. Per esempio, come vedremo, se una funzione assume solo un numero finito di valori, per esempio

$$(1.1) \quad f(x) = \sum_{i=1}^k c_i \chi_{E_i}(x)$$

con $c_i \in \mathbb{R}$ e gli E_i a due a due disgiunti, è naturale definire il suo integrale ponendo

$$\int f = \sum_{i=1}^k c_i \text{mis}(E_i).$$

Per avere una sufficiente flessibilità nell'approssimare funzioni generiche con funzioni del tipo (1.1) non basta (in $\mathbb{R}$) prendere come insiemi di livello E_i degli intervalli, ma bisogna considerare insiemi più generali. Questo pone il problema di definire opportunamente una misura su classi molto ampie di insiemi. Ecco perché trattiamo prima la misura e poi l'integrale (e non viceversa, come nella teoria di Riemann). La trattazione che presentiamo è generale (non si limita alla misura "standard" su $\mathbb{R}^n$) per poter essere applicata fra l'altro nell'ambito della teoria delle distribuzioni, che vedremo in un capitolo successivo, e della teoria delle probabilità.

1.1 Misure positive

Definizione 1.1 Sia X un insieme non vuoto. Una famiglia $\mathcal{E}$ di sottoinsiemi di X si chiama algebra, se

(a) $\emptyset \in \mathcal{E}$;

(b) $\mathcal{E}$ è chiusa rispetto all'unione: per ogni $A, B \in \mathcal{E}$, risulta $A \cup B \in \mathcal{E}$;

(c) $\mathcal{E}$ è chiusa rispetto alla complementazione: per ogni $A \in \mathcal{E}$, risulta $X \setminus A \in \mathcal{E}$.

Osservazione 1.2 Equivalentemente, $\mathcal{E}$ è un'algebra se sono soddisfatte le proprietà (a), (b') e (c), dove

(b') $\mathcal{E}$ è chiusa rispetto all'intersezione: per ogni $A, B \in \mathcal{E}$, risulta $A \cap B \in \mathcal{E}$.

In virtù della proprietà associativa di unione e intersezione, si ha che l'unione e l'intersezione di un numero finito di elementi di $\mathcal{E}$ è ancora in $\mathcal{E}$.

Definizione 1.3 Una famiglia $\mathcal{E}$ di sottoinsiemi di X si chiama σ -algebra se è chiusa rispetto alle operazioni insiemistiche numerabili, cioè se presi $E_n \in \mathcal{E}$ risulta che $\bigcap_{n \in \mathbb{N}} E_n, \bigcup_{n \in \mathbb{N}} E_n \in \mathcal{E}$.

È facile vedere che $\{\emptyset, X\}$ e $\mathcal{P}(X)$ sono σ -algre, rispettivamente la più piccola e la più grande (rispetto all'inclusione) tra tutte quelle che si possono formare con sottoinsiemi di X .

A partire da $\mathcal{G} \subset \mathcal{P}(X)$, è possibile considerare la più piccola σ -algebra contenente $\mathcal{G}$ e definita da

$$\sigma(\mathcal{G}) = \bigcap \{ \mathcal{E} \mid \mathcal{E} \text{ } \sigma\text{-algebra, } \mathcal{G} \subset \mathcal{E} \subset \mathcal{P}(X) \}.$$

La definizione è ben posta, in quanto esiste almeno una σ -algebra, $\mathcal{P}(X)$, che gode delle proprietà richieste. Inoltre, è facile verificare che l'intersezione di una qualunque famiglia di σ -algre è ancora una σ -algebra. In $\mathbb{R}^N$, la σ -algebra di Borel $\mathcal{B}(\mathbb{R}^N)$ è quella generata dalla famiglia di tutti i sottoinsiemi aperti (o, equivalentemente, generata da tutti i sottoinsiemi chiusi).

La coppia $(X, \mathcal{E})$, dove $\mathcal{E}$ è una σ -algebra di sottoinsiemi di X , si chiama *spazio misurabile* e gli elementi di $\mathcal{E}$ prendono il nome di *insiemi misurabili*. Uno spazio misurabile rappresenta l'ambiente adatto per l'introduzione del concetto di *misura*. Data una funzione d'insieme $\mu : \mathcal{E} \rightarrow [0, +\infty]$, diremo che

- μ è *additiva* se $\mu(A \cup B) = \mu(A) + \mu(B)$, per ogni scelta degli insiemi A, B in $\mathcal{E}$ con $A \cap B = \emptyset$;
- μ è σ -*subadditiva* (o equivalentemente *numerabilmente subadditiva*) se per ogni $\{E_h\} \subset \mathcal{E}$ e per ogni $E \in \mathcal{E}$ tali che $E \subset \bigcup_{h=1}^{\infty} E_h$ risulta

$$\mu(E) \leq \sum_{h=1}^{\infty} \mu(E_h);$$

- μ è σ -additiva (o equivalentemente *numerabilmente additiva*) se, posto $E = \bigcup_{h=1}^{\infty} E_h$, con $\{E_h\} \subset E$ e $E_h \cap E_k = \emptyset$ per ogni $h \neq k$, risulta

$$\mu(E) = \sum_{h=1}^{\infty} \mu(E_h);$$

- μ è *finita* se $\mu(X) < +\infty$.

Definizione 1.4 Chiamiamo misura positiva su uno spazio misurabile $(X, \mathcal{E})$ ogni funzione d'insieme σ -additiva $\mu : \mathcal{E} \rightarrow [0, +\infty]$ tale che $\mu(\emptyset) = 0$. La terna $(X, \mathcal{E}, \mu)$ si chiama spazio di misura.

Un insieme $E \in \mathcal{E}$ si dice σ -finito se si può decomporre nell'unione numerabile di insiemi aventi misura finita, cioè $E = \bigcup_{h=1}^{\infty} E_h$ con $\mu(E_h) < +\infty$. Se X è σ -finito, allora la misura μ si dice σ -finita.

Se μ è una misura positiva finita su X con $\mu(X) = 1$, allora μ è una misura di probabilità e la terna $(X, \mathcal{E}, \mu)$ si dice, in questo caso, uno spazio di probabilità.

Osservazione 1.5 Ogni misura positiva è

- (a) crescente rispetto all'inclusione insiemistica

$$E \subset F \implies \mu(E) \leq \mu(F),$$

in quanto $\mu(F) = \mu(E) + \mu(F \setminus E)$. In realtà, per questa proprietà è sufficiente che μ sia una funzione d'insieme positiva e additiva.

- (b) continua rispetto alle successioni crescenti di insiemi misurabili

$$E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots \implies \mu\left(\bigcup_n E_n\right) = \lim_{n \rightarrow \infty} \mu(E_n).$$

Infatti, posto $E_0 = \emptyset$ e $F_n = E_n \setminus E_{n-1}$, per $n \geq 1$, abbiamo una nuova successione di insiemi misurabili *disgiunti*, la cui unione è uguale a quella degli E_n . Siccome μ è σ -additiva, possiamo scrivere

$$\begin{aligned} \mu\left(\bigcup_n F_n\right) &= \sum_n \mu(F_n) = \sum_n (\mu(E_n) - \mu(E_{n-1})) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n (\mu(E_k) - \mu(E_{k-1})) = \lim_{n \rightarrow \infty} \mu(E_n), \end{aligned}$$

avendo ottenuto una somma telescopica nel penultimo passaggio.

(c) continua anche rispetto a successioni decrescenti di insiemi misurabili

$$E_1 \supset E_2 \supset \cdots \supset E_n \supset \cdots \implies \mu\left(\bigcap_n E_n\right) = \lim_{n \rightarrow \infty} \mu(E_n)$$

se vale, in aggiunta, che $\mu(E_1) < +\infty$. Infatti, applicando la (b) alla successione crescente $E_1 \setminus E_n$, la cui unione è $E_1 \setminus E$, otteniamo

$$\begin{aligned} \mu(E_1) - \mu(E) &= \mu(E_1 \setminus E) = \lim_{n \rightarrow +\infty} \mu(E_1 \setminus E_n) \\ &= \mu(E_1) - \lim_{n \rightarrow +\infty} \mu(E_n). \end{aligned}$$

Sottraendo $\mu(E_1)$ ad ambo i membri si ha la tesi.

Osserviamo che la richiesta che almeno un elemento di una successione decrescente abbia misura finita per passare al limite non può essere rimossa. Infatti, se $X = \mathbb{R}$, $E_n = [n, +\infty[$ e μ è la misura sulla retta, allora $\bigcap_n E_n = \emptyset$, mentre $\lim_{n \rightarrow \infty} \mu(E_n) = +\infty$.

Nelle applicazioni, risulta utile il seguente criterio per stabilire se una data funzione d'insieme è σ -additiva.

Proposizione 1.6 *Sia $(X, \mathcal{E})$ uno spazio misurabile e sia $\mu : \mathcal{E} \rightarrow [0, +\infty]$ una funzione d'insieme additiva e σ -subadditiva. Allora μ è σ -additiva.*

DIM. Consideriamo una successione di insiemi in $\mathcal{E}$ a due a due disgiunti, $\{E_h\}$, e sia E la loro unione. In virtù delle ipotesi abbiamo

$$\mu(E) \leq \sum_{h=1}^{\infty} \mu(E_h) = \lim_{n \rightarrow \infty} \sum_{h=1}^n \mu(E_h) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{h=1}^n E_h\right) \leq \mu(E).$$

□

Esempi 1.7 Vediamo degli esempi classici di misure positive.

- (1) Sia X un insieme non vuoto e sia $\mathcal{E} = \mathcal{P}(X)$. Definiamo la cosiddetta *misura del contare* nel modo seguente: $\mu(\emptyset) = 0$, $\mu(E) =$ cardinalità di E , se E è finito, $\mu(E) = \infty$, altrimenti.
- (2) Per ogni $x \in X$, poniamo $\delta_x(E) = 1$, se $x \in E$, $\delta_x(E) = 0$, se $x \notin E$, per ogni $E \subset X$. δ_x prende il nome di *misura di Dirac* concentrata in x .

- (3) La costruzione di Riemann per l'integrale induce la “misura di Peano-Jordan” μ_{PJ} in $\mathbb{R}^n$. Sulla retta $\mathbb{R}$ un insieme E si dice misurabile secondo Peano-Jordan se per ogni $\varepsilon > 0$ esistono intervalli a due a due disgiunti $I_1, \dots, I_k$ e $J_1, \dots, J_n$ tali che

$$\bigcup_{i=1}^k I_i \subset E \subset \bigcup_{j=1}^n J_j, \quad \text{e} \quad \sum_{j=1}^n \ell(J_j) - \sum_{i=1}^k \ell(I_i) < \varepsilon,$$

dove se $I = (a, b)$ abbiamo posto $\ell(I) = b - a$. La famiglia degli insiemi μ_{PJ} -misurabili non è per'ò una σ -algebra: per esempio, $E = \mathbb{Q} \cap [0, 1]$ non è μ_{PJ} -misurabile. Infatti, per la numerabilità E è l'unione di una successione crescente di insiemi finiti (perciò misurabili, e di misura nulla), ma gli unici intervalli I_i contenuti in E sono costituiti da un solo punto, mentre il più piccolo intervallo J_j contenente E è $[0, 1]$. Segue che per ogni scelta degli (I_i) e dei (J_j) come sopra risulta $\sum_{j=1}^n \ell(J_j) - \sum_{i=1}^k \ell(I_i) \geq 1$ ed E non può essere μ_{PJ} -misurabile.

Definizione 1.8 Chiameremo insiemi di misura nulla, o trascurabili, gli insiemi $E \in \mathcal{E}$ tali che $\mu(E) = 0$. Diremo che una certa proprietà $P(x)$ è vera per μ -quasi ogni x se l'insieme

$$\{x \in X : P(x) \text{ è falsa}\}$$

è trascurabile.

1.2 Funzioni misurabili

Introduciamo in questa sezione il concetto di funzione misurabile, che è quello naturale da considerare quando si ha a che fare con spazi misurabili. D'ora in poi, indicheremo per brevità l'insieme $f^{-1}((t, +\infty))$ con $\{f > t\}$. Con significato analogo scriveremo $\{f \geq t\}$, $\{f < t\}$, $\{f \leq t\}$.

Definizione 1.9 Sia $(X, \mathcal{E})$ uno spazio misurabile e sia $f : X \rightarrow \mathbb{R}$. Diremo che f è misurabile se

$$\{f > t\} \in \mathcal{E}, \quad \text{per ogni } t \in \mathbb{R}.$$

Se $f : X \rightarrow \mathbb{R}^m$, diremo che f è misurabile se lo sono tutte le sue componenti scalari.

Condizioni equivalenti alla misurabilità si hanno imponendo la misurabilità di tutti gli insiemi $\{f \geq t\}$ o, passando ai complementari, la misurabilità di tutti gli insiemi $\{f < t\}$ oppure di tutti gli insiemi $\{f \leq t\}$. Si ha infatti

$$\{f \geq t\} = \bigcap_{h=1}^{\infty} \{f > t - 1/h\} \quad \{f > t\} = \bigcup_{h=1}^{\infty} \{f \geq t + 1/h\}$$

quindi dalla misurabilità degli insiemi con la disuguaglianza stretta si deduce la misurabilità di quelli con la disuguaglianza non stretta e viceversa. Le funzioni misurabili sono una classe stabile rispetto alle principali operazioni algebriche.

Proposizione 1.10 *Siano f, g funzioni misurabili. Allora per $\alpha, \beta \in \mathbb{R}$ sono funzioni misurabili le funzioni $\alpha f + \beta g$, la funzione $f \vee g = \max\{f, g\}$, la funzione $f \wedge g = \min\{f, g\}$. Inoltre, se (f_h) è una successione di funzioni misurabili convergenti puntualmente a f , anche f è misurabile. Infine, se f è misurabile e $\{f \neq g\}$ è trascurabile, anche g è misurabile.*

Dim. Se $\alpha \neq 0$ allora $\{\alpha f > t\}$ è uguale a $\{f > t/\alpha\}$ o $\{f < t/\alpha\}$ a seconda che $\alpha > 0$ o $\alpha < 0$. Se $\alpha = 0$ l'insieme è vuoto oppure coincide con X . In ogni caso $\{\alpha f > t\}$ è misurabile e quindi αf è misurabile. Analogamente, βg è misurabile. Per la somma, si ha

$$\{\alpha f + \beta g > t\} = \bigcup_{q \in \mathbb{Q}} \{\alpha f > t - q\} \cap \{\beta g > q\}$$

quindi anche la somma è misurabile, dal momento che i soprallivelli sono esprimibili come unione numerabile di insiemi misurabili. La stabilità rispetto a massimi e minimi segue subito dalle identità

$$\begin{aligned} \{f \vee g > t\} &= \{f > t\} \cup \{g > t\} \\ \{f \wedge g > t\} &= \{f > t\} \cap \{g > t\}. \end{aligned}$$

Siano f_h funzioni misurabili convergenti puntualmente a f . Si ha

$$\begin{aligned} \{f > t\} &= \bigcup_{k=1}^{\infty} \{x \in X : f_h(x) > t + 1/k \text{ definitivamente}\} \\ &= \bigcup_{k=1}^{\infty} \bigcup_{h=1}^{\infty} \bigcap_{i=h}^{\infty} \{f_i > t + 1/k\}. \end{aligned}$$

Infine, se f è misurabile e $E = \{f \neq g\}$ allora

$$\{f > t\} \Delta \{g > t\} \subset E$$

quindi essendo E trascurabile e $\{f > t\}$ misurabile anche $\{g > t\}$ è misurabile. $\square$

Se $f : X \rightarrow \mathbb{R}$ è misurabile, poniamo

$$(1.2) \quad \|f\|_\infty = \inf\{t \geq 0 : \mu(\{x \in X : |f(x)| \geq t\}) = 0\}$$

e diciamo che f è *essenzialmente limitata* se $\|f\|_\infty < +\infty$. Lo spazio delle funzioni essenzialmente limitate in E si denota con $L^\infty(E)$.

Tra tutte le funzioni misurabili godono di particolare interesse le *funzioni caratteristiche* di insiemi misurabili e le loro combinazioni lineari. Ricordiamo che, per ogni $E \subset X$

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

Si verifica immediatamente che χ_E è misurabile se e solo se E è misurabile: infatti,

$$\{\chi_E > t\} = \begin{cases} \emptyset & t \geq 1 \\ E & 0 \leq t < 1 \\ \mathbb{R} & t < 0 \end{cases}$$

e quindi χ_E è misurabile se e solo se E è misurabile.

Definizione 1.11 *Indicheremo con $\mathcal{S}_+(X)$ l'insieme delle funzioni semplici positive, cioè l'insieme delle funzioni esprimibili nella forma*

$$f(x) = \sum_{i=1}^k z_i \chi_{E_i}(x),$$

con $z_1, \dots, z_N$ positivi e $E_1, \dots, E_N \in \mathcal{E}$.

Osserviamo che tutte le funzioni semplici sono misurabili. La rappresentazione come combinazione lineare di funzioni caratteristiche non è certo unica: ad esempio (con $X = \mathbb{R}$)

$$\chi_{[-1,1]} + \chi_{[0,2]} + \chi_{(1,3]} = \chi_{[-1,0)} + 2\chi_{[0,2]} + \chi_{(2,3]}.$$

È facile vedere che, aumentando se necessario il numero degli addendi, ogni funzione semplice è rappresentabile nel seguente modo

$$(1.3) \quad f(x) = \sum_{i=1}^k z_i \chi_{E_i}(x) \quad \text{con } E_i \cap E_j = \emptyset \text{ per } i \neq j.$$

Vale il seguente importante risultato di approssimazione.

Teorema 1.12 Sia $f : X \rightarrow [0, +\infty]$ una funzione misurabile. Allora esiste una successione di funzioni semplici (s_n) , tali che $0 \leq s_n(x) \leq s_{n+1}(x) \leq \dots$ e

$$\lim_{n \rightarrow \infty} s_n(x) = f(x), \quad \text{per ogni } x \in X.$$

Se f è limitata, allora la convergenza è uniforme.

DIM. Supponiamo $f \geq 0$. Per ogni $n \in \mathbb{N}$ e $k = 0, 1, \dots, n2^n - 1$, definiamo

$$E_{k,n} = f^{-1} \left(\left[\frac{k}{2^n}, \frac{k+1}{2^n} \right) \right).$$

Tali insiemi sono misurabili, dato che f lo è. Poniamo poi

$$s_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \chi_{E_{k,n}} + (n+1) \chi_{E_n},$$

dove $E_n = \{f \geq n\}$. Chiaramente, s_n sono funzioni semplici e, si può provare che esse convergono puntualmente e in modo crescente ad f . Se f è limitata si vede che la convergenza è uniforme. Infatti, fissato n , per ogni k gli $E_{k,n}$ sono a due a due disgiunti e, siccome $x \in E_{k,n}$ implica $k/2^n \leq f(x) < (k+1)/2^n$ e $s_n(x) = k/2^n$, si ha $|f(x) - s_n(x)| \leq 1/2^n$ per ogni $x \in X$.

Il caso di f con segno qualunque segue applicando la prima parte a $f^+ = \max\{f, 0\}$ ed $f^- = \max\{-f, 0\}$, tenendo conto che $f = f^+ - f^-$. $\square$

1.3 Integrazione

Dopo le premesse fatte, possiamo finalmente definire l'integrale rispetto ad una generica misura positiva. La costruzione avviene per passi, considerando dapprima le funzioni semplici positive, quindi quelle misurabili positive, per arrivare, infine, ad integrare funzioni misurabili di segno qualunque.

Siano $(X, \mathcal{E})$ uno spazio misurabile e μ una misura positiva in tale spazio. Supponiamo anche che μ sia *completa*, cioè che ogni sottoinsieme di un insieme misurabile con misura nulla sia ancora misurabile (ed abbia misura nulla). Di fatto, la costruzione dell'integrale è possibile anche senza questa ulteriore condizione. Tuttavia, per una misura completa valgono alcune proprietà in più, che preferiamo mettere in evidenza.

Indichiamo con $\mathcal{S}_+(X)$ l'insieme delle funzioni semplici positive in X . Se $s \in \mathcal{S}_+(X)$, $s(x) = \sum_{i=1}^k z_i \chi_{E_i}(x)$, con $E_i \cap E_j = \emptyset$, per $i \neq j$, definiamo

$$\int_X s \, d\mu = \sum_{i=1}^k z_i \mu(E_i),$$

adottando la convenzione che $z_i \mu(E_i) = 0$ se $z_i = 0$ e $\mu(E_i) = \infty$. La definizione non dipende dalla rappresentazione di s ed in particolare risulta

$$\int_X s d\mu = \sum_{z \in Im(s)} z \mu(s^{-1}(z)).$$

Inoltre, l'applicazione che ad una funzione semplice s associa il suo integrale è lineare, pur di moltiplicare per numeri positivi, e gode della seguente proprietà di monotonia:

$$s, t \in \mathcal{S}_+(X), s \geq t \implies \int_X s d\mu \geq \int_X t d\mu.$$

Estendiamo ora l'operazione di integrazione ad una classe più ampia di funzioni.

Definizione 1.13 *Sia f misurabile e positiva. Allora poniamo*

$$\int_X f d\mu = \sup \left\{ \int_X s d\mu : s \in \mathcal{S}_+(X), s \leq f \right\}.$$

Notiamo che l'integrale di una funzione misurabile positiva è ben definito, ma può essere $+\infty$. Inoltre, si vede facilmente che se $f \in \mathcal{S}_+(X)$, il suo integrale coincide con quello già definito.

In generale, se f è misurabile (e di segno qualunque), consideriamo la parte positiva $f^+ = f \vee 0$ e la parte negativa $f^- = (-f) \vee 0$ di f , il cui integrale è definito nella Definizione 1.13. Se f^+ ed f^- hanno integrale finito poniamo

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

Diremo che f è *sommabile* in X se è misurabile e la parte positiva e la parte negativa hanno entrambe integrale finito.

Notiamo che tutte le funzioni limitate, misurabili, nulle fuori di un insieme di misura finita sono sommassibili. L'insieme delle funzioni sommassibili in X verrà indicato con $L^1(X)$.

Teorema 1.14 *L'insieme $L^1(X)$ delle funzioni sommassibili in X è uno spazio vettoriale. Inoltre valgono le proprietà*

(i) *(Linearità)*

$$\int_X [\alpha f + \beta g] d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$$

(ii) *(Monotonia)*

$$f \geq g \implies \int_X f d\mu \geq \int_X g d\mu$$

(iii) (Indipendenza da insiemi di misura nulla) Se $f \in L^1(X)$ e $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$ allora

$$g \in L^1(X) \quad e \quad \int_E f d\mu = \int_X g d\mu.$$

(iv)

$$f \text{ misurabile, } |f| \leq g \in L^1(X) \quad \mu\text{-q.o.} \implies f \in L^1(X).$$

Per quanto riguarda la proprietà (iv), basta osservare che $|f| \leq g$ implica $0 \leq f^+ \leq g$ e $0 \leq f^- \leq g$ e quindi $\int_X f^+ d\mu \leq \int_X g d\mu$ per ogni funzione semplice minorante f^+ (resp. f^-).

Una conseguenza della linearità dell'integrale è la disuguaglianza tra il modulo dell'integrale e l'integrale del modulo:

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu.$$

Infatti si ha

$$\begin{aligned} \left| \int_X f d\mu \right| &= \left| \int_X f^+ d\mu - \int_X f^- d\mu \right| \\ &\leq \int_X f^+ d\mu + \int_X f^- d\mu = \int_X |f| d\mu. \end{aligned}$$

La nozione di integrale si estende immediatamente alle funzioni a valori vettoriali $f = (f_1, \dots, f_m) : X \rightarrow \mathbb{R}^m$ ponendo

$$\int_X f d\mu = \left(\int_X f_1 d\mu, \dots, \int_X f_m d\mu \right)$$

purché le componenti f_i siano integrabili. In particolare, se $m = 2$ con la consueta identificazione del piano con il campo complesso resta definito anche l'integrale di una funzione complessa $f = u + iv : X \rightarrow \mathbb{C}$,

$$\int_X f d\mu = \int_X u d\mu + i \int_X v d\mu.$$

Dato un insieme misurabile $E \subset X$ e $f : X \rightarrow \mathbb{R}$ misurabile diremo che f è *sommabile in E* se $f\chi_E$ è sommabile in X . In tal caso porremo

$$\int_E f d\mu = \int_X f\chi_E d\mu.$$

Avremmo ottenuto lo stesso risultato se, per definire l'integrale in E avessimo usato la restrizione della misura μ ad E . L'insieme delle funzioni sommabili in E verrà indicato con $L^1(E, \mu)$.

1.4 La misura di Lebesgue in $\mathbb{R}^n$

In questo paragrafo vediamo come sia possibile definire una nozione di “misura n -dimensionale” per un sottoinsieme E di $\mathbb{R}^n$. L’idea è quella di assegnare ai “rettangoli”

$$R = [a_1, b_1[\times \cdots \times [a_n, b_n[$$

una misura (lunghezza per $n = 1$, area per $n = 2$, volume per $n = 3, \dots$) uguale a

$$m_n(R) = (b_1 - a_1) \cdot (b_2 - a_2) \cdots (b_n - a_n)$$

e poi di definire la misura di un insieme generico mediante ricoprimenti fatti con rettangoli. Per $n = 1$ si tratta evidentemente di intervalli, per $n = 2$ di rettangoli veri e propri, per $n = 3$ di parallelepipedi. Useremo la parola rettangoli, senza introdurre ulteriori termini. Chiameremo invece “plurirettangolo” un’unione finita di rettangoli del tipo detto.

Definizione 1.15 *Sia $E \subset \mathbb{R}^n$ un insieme. Definiamo $\overline{m}_n(E)$, misura esterna n -dimensionale di Lebesgue di E , nel seguente modo:*

$$\overline{m}_n(E) = \inf \left\{ \sum_{i=1}^{\infty} m_n(R_i) : E \subset \bigcup_{i=1}^{\infty} R_i \right\}.$$

Quindi la misura esterna di E si ottiene cercando, tra tutti i ricoprimenti di E mediante successioni di rettangoli n -dimensionali, quello che minimizza la somma dei volumi. Osserviamo che i rettangoli del ricoprimento non sono necessariamente disgiunti. Vediamo alcune proprietà della misura esterna che seguono direttamente dalla definizione:

- (a) La misura esterna di E è sempre compresa tra 0 e $+\infty$. Inoltre $\overline{m}_n(E) < +\infty$ se E è un insieme limitato e $\overline{m}_n(\emptyset) = 0$. D’altra parte, come vedremo, esistono insiemi non limitati di misura finita ed insiemi non vuoti di misura nulla.
- (b) La misura esterna è invariante per traslazioni

$$\overline{m}_n(x + E) = \overline{m}_n(E) \quad \forall E \subset \mathbb{R}^n, x \in \mathbb{R}^n$$

ed ha il seguente comportamento rispetto ad omotetie

$$(1.4) \quad \overline{m}_n(\lambda E) = \lambda^n \overline{m}_n(E) \quad \forall E \subset \mathbb{R}^n, \lambda > 0.$$

Infatti, se E si ricopre con rettangoli R_i allora $x + E$ si ricopre con rettangoli $x + R_i$, che hanno lo stesso volume, mentre l’insieme λE si ricopre con rettangoli λR_i , il cui volume è $\lambda^n m_n(R_i)$.

(c) La misura esterna è una *funzione crescente d'insieme*:

$$E_1 \subset E_2 \quad \implies \quad \overline{m}_n(E_1) \leq \overline{m}_n(E_2).$$

Inoltre è *numerabilmente subadditiva*:

$$(1.5) \quad E \subset \bigcup_{h=1}^{\infty} E_h \quad \implies \quad \overline{m}_n(E) \leq \sum_{h=1}^{\infty} \overline{m}_n(E_h).$$

La disuguaglianza è infatti ovvia se la sommatoria di $\overline{m}_n(E_h)$ vale $+\infty$. In caso contrario, per ogni $\varepsilon > 0$ ed ogni $h \geq 1$ esiste una successione di rettangoli $(R_{h,i})_{i \in \mathbb{N}}$ tale che

$$E_h \subset \bigcup_{i=1}^{\infty} R_{h,i} \quad \sum_{i=1}^{\infty} m_n(R_{h,i}) < \overline{m}_n(E_h) + \varepsilon 2^{-h}.$$

Sommando in h si ha

$$\sum_{h=1}^{\infty} \sum_{i=1}^{\infty} m_n(R_{h,i}) < \sum_{h=1}^{\infty} \overline{m}_n(E_h) + \varepsilon$$

e dato che i rettangoli $R_{h,i}$ ricoprono E si ottiene

$$\overline{m}_n(E) < \sum_{h=1}^{\infty} \overline{m}_n(E_h) + \varepsilon.$$

Basta allora mandare ε a zero per ottenere la (1.5). □

La misura esterna di $\{x\}$ (il singoletto costituito dal solo x) è zero. Questo perché $\{x\}$ è contenuto in cubi di volume arbitrariamente piccolo. Dalla numerabile subaddittività di $\overline{m}_n$ si ottiene che ogni insieme numerabile ha misura esterna nulla. Infatti, data $x : \mathbb{N} \rightarrow E$ surgettiva, si ha

$$E = \bigcup_{h=0}^{\infty} \{x_h\}.$$

In particolare l'insieme dei punti a coordinate razionali $\mathbb{Q}^n$, essendo numerabile, ha misura nulla.

Anche insiemi “continui” e quindi non numerabili possono avere misura esterna nulla. Ad esempio il segmento $E = (0, 1) \times \{0\}$ in $\mathbb{R}^2$ ha misura esterna nulla: dato infatti $\varepsilon > 0$ ed un intero h tale che $h\varepsilon > 4$, basta ricoprire E con gli h cubi Q_i centrati in i/h e di lato $1/h$ per avere $\overline{m}_n(E) \leq h \cdot (2/h)^2 = 4/h < \varepsilon$.

Vediamo alcune proprietà degli insiemi trascurabili.

Proposizione 1.16 *Ogni sottoinsieme di un insieme trascurabile è trascurabile. L'unione di una successione di insiemi trascurabili è trascurabile. Un insieme E è trascurabile se e solo se per ogni $\varepsilon > 0$ esiste un aperto $A_\varepsilon \supset E$ tale che $\overline{m}_n(A_\varepsilon) < \varepsilon$.*

Dim. I primi due enunciati seguono facilmente dalle proprietà (c) della misura esterna. Se E è contenuto in aperti di misura esterna arbitrariamente piccola, per la monotonia della misura esterna deve essere $\overline{m}_n(E) = 0$. Viceversa, se $\overline{m}_n(E) = 0$ allora per ogni $\varepsilon > 0$ esiste una successione di rettangoli aperti R_i tale che

$$E \subset \bigcup_{i=1}^{\infty} R_i \quad \text{e} \quad \sum_{i=1}^{\infty} m_n(R_i) < \varepsilon.$$

Detta A_ε l'unione dei rettangoli R_i si ha, per la definizione di misura esterna, $\overline{m}_n(A_\varepsilon) < \varepsilon$. $\square$

La misura esterna di Lebesgue ha il vantaggio di essere definita per ogni sottoinsieme di $\mathbb{R}^n$. Il prezzo di questo è però una perdita di additività: esistono esempi di insiemi limitati E, F di $\mathbb{R}^n$ disgiunti e tali che

$$(1.6) \quad \overline{m}_n(E \cup F) < \overline{m}_n(E) + \overline{m}_n(F).$$

Individueremo ora una σ -algebra di insiemi, detti insiemi misurabili, sui quali la misura esterna di Lebesgue ha un comportamento migliore, cioè una σ -algebra $\mathcal{M}(\mathbb{R}^n)$ tale che la restrizione di $\overline{m}_n$ ad $\mathcal{M}(\mathbb{R}^n)$ risulti σ -additiva, definendo una misura.

Definizione 1.17 *Posto $E \triangle F = (E \setminus F) \cup (F \setminus E)$, diremo che un insieme limitato $E \subset \mathbb{R}^n$ è misurabile se per ogni $\varepsilon > 0$ esiste un plurirettangolo P tale che*

$$\overline{m}_n(E \triangle P) < \varepsilon.$$

Se E non è limitato diremo che E è misurabile se $E \cap B_R(0)$ è misurabile per ogni $R > 0$. Indicheremo con $\mathcal{M}(\mathbb{R}^n)$ la classe dei sottoinsiemi misurabili di $\mathbb{R}^n$.

La proposizione precedente implica che tutti gli insiemi trascurabili sono misurabili (si prende $P = \emptyset$).

Vediamo ora alcune proprietà di stabilità degli insiemi misurabili:

Teorema 1.18

(i) *Gli aperti e i chiusi di $\mathbb{R}^n$ sono misurabili.*

(ii) La classe $\mathcal{M}(\mathbb{R}^n)$ è una σ -algebra.

(iii) La classe $\mathcal{M}(\mathbb{R}^n)$ è stabile per omotetie e traslazioni:

$$E \in \mathcal{M}(\mathbb{R}^n) \quad \implies \quad x + E, \lambda E \in \mathcal{M}(\mathbb{R}^n) \quad \forall x \in \mathbb{R}^n, \lambda > 0.$$

(iv) Se E è misurabile e F è tale che $\overline{m}_n(E \triangle F) = 0$ allora anche F è misurabile.

In base al teorema precedente sono misurabili le unioni numerabili di chiusi, le intersezioni numerabili di aperti e così via. La classe degli insiemi misurabili è quindi estremamente ricca.

A confermare questo c'è anche il fatto che i soli esempi di sottoinsiemi di $\mathbb{R}^n$ non misurabili per i quali vale (1.6) si ottengono in modo non costruttivo. Tutti gli insiemi “operativamente” costruiti a partire dai chiusi e dagli aperti con operazioni finite o numerabili sono misurabili.

Possiamo ora definire mediante restrizione ai misurabili la *misura di Lebesgue* in $\mathbb{R}^n$:

Definizione 1.19 Dato $E \subset \mathbb{R}^n$ misurabile poniamo $m_n(E) = \overline{m}_n(E)$. Il numero $m_n(E)$ è detto *misura di Lebesgue* di E .

Teorema 1.20 La misura di Lebesgue è numerabilmente additiva su $\mathcal{M}(\mathbb{R}^n)$: per ogni E misurabile ed ogni sua partizione in insiemi misurabili E_h si ha

$$m_n(E) = \sum_{h=1}^{\infty} m_n(E_h).$$

Inoltre m_n è invariante per traslazioni, soddisfa la (1.4) e $m_n([0, 1]^n) = 1$.

Dal teorema precedente possiamo dedurre altre utili proprietà della misura di Lebesgue, semplici riformulazioni delle proprietà viste nell'Osservazione 1.5.

Corollario 1.21 Per ogni coppia di insiemi $E, F \in \mathcal{M}(\mathbb{R}^n)$ si ha

$$m_n(E \cup F) + m_n(E \cap F) = m_n(E) + m_n(F).$$

Data una successione crescente di insiemi $(E_h) \subset \mathcal{M}(\mathbb{R}^n)$ si ha

$$m_n\left(\bigcup_{h=1}^{\infty} E_h\right) = \lim_{h \rightarrow +\infty} m_n(E_h).$$

Data una successione decrescente di insiemi $(E_h) \subset \mathcal{M}(\mathbb{R}^n)$, si ha

$$m_n(E_1) < +\infty \quad \implies \quad m_n\left(\bigcap_{h=1}^{\infty} E_h\right) = \lim_{h \rightarrow +\infty} m_n(E_h).$$

Osserviamo che, com'è naturale aspettarsi, se una funzione è integrabile nel senso di Riemann allora è integrabile anche nel senso di Lebesgue e i due integrali hanno lo stesso valore.

È naturale chiedersi se in un insieme X qualunque è sempre lecito introdurre una misura positiva. La risposta è affermativa. Come nel caso della misura di Lebesgue, si può dimostrare che tramite una *misura esterna* è possibile costruire una classe di misurabili e una misura positiva. Non vedremo questa costruzione nella massima generalità, bensì in un caso particolare che sarà quello della misura di Stieltjes.

Una funzione d'insieme definita su tutte le parti di X , $\bar{\mu} : \mathcal{P}(X) \rightarrow [0, +\infty]$, si chiama *misura esterna* se

- (i) $\bar{\mu}(\emptyset) = 0$;
- (ii) $\bar{\mu}(A) \leq \bar{\mu}(B)$ se $A \subset B$;
- (iii) $\bar{\mu}(\bigcup_k A_k) \leq \sum_k \bar{\mu}(A_k)$ per ogni collezione numerabile di insiemi $\{A_k\}$.

Diremo che $E \subset X$ è misurabile se

$$\bar{\mu}(A) = \bar{\mu}(A \cap E) + \bar{\mu}(A \setminus E), \quad \forall A \subset X.$$

Si dimostra che la classe dei misurabili è una σ -algebra e che $\bar{\mu}$ ristretta a questa famiglia è σ -additiva, quindi è una misura.

1.5 Passaggio al limite sotto il segno di integrale

In questo paragrafo vedremo sotto quali condizioni vale l'implicazione

$$(1.7) \quad f_h \rightarrow f \quad \implies \quad \int_X f_h d\mu \rightarrow \int_X f d\mu.$$

Vedremo che in molti casi la sola convergenza puntuale della successione implica la convergenza degli integrali. Incominceremo col considerare il caso in cui f_h è una successione monotona crescente. Per poter dare l'enunciato nella sua forma più generale ci sarà utile estendere la nozione di integrale al caso di funzioni a valori in $\overline{\mathbb{R}} = [-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$.

Definizione 1.22 Sia $f : X \rightarrow \overline{\mathbb{R}}$. Diremo che f è misurabile se per ogni $t \in \mathbb{R}$ l'insieme $\{f > t\}$ è misurabile. Se f è misurabile gli insiemi

$$\{f = +\infty\} = \bigcap_{h=1}^{\infty} \{f > h\} \quad \{f = -\infty\} = \bigcap_{h=1}^{\infty} \{f \leq -h\}$$

sono misurabili. Se $f \geq 0$ poniamo

$$\int_X f d\mu = \lim_{R \rightarrow +\infty} \int_X f \wedge R d\mu$$

e nel caso generale poniamo

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu$$

purché almeno uno degli integrali sia finito.

Restano vere le proprietà di monotonia, invarianza e linearità dell'integrale, con l'eccezione di indeterminazioni del tipo $\infty - \infty$. È importante osservare che vale l'implicazione

$$\int_X |f| d\mu < +\infty \quad \implies \quad \mu(X \cap \{|f| = +\infty\}) = 0.$$

Quindi se $f \in L^1(X)$ (cioè il suo modulo ha integrale finito su X) allora gli insiemi $\{f = \pm\infty\}$ sono trascurabili. Usando la disuguaglianza $|f| \geq R\chi_{\{|f|>R\}}$ si ha infatti

$$\int_X |f| d\mu \geq R\mu(\{|f| > R\}) \geq R\mu(\{|f| = +\infty\}).$$

Dividendo ambo i membri per R e passando al limite per $R \rightarrow +\infty$ si ottiene $\mu(\{|f| = \infty\}) = 0$.

Teorema 1.23 (di Beppo Levi della convergenza monotona) *Sia (f_h) una successione crescente di funzioni misurabili a valori in $[0, +\infty]$. Posto, per ogni $x \in X$,*

$$f(x) = \sup_{h \geq 1} f_h(x) = \lim_{h \rightarrow +\infty} f_h(x),$$

si ha

$$\int_X f d\mu = \lim_{h \rightarrow +\infty} \int_X f_h d\mu.$$

Dim. Notiamo che la convergenza puntuale delle f_h ad f non va dimostrata, perché è ovvia per monotonia, così come la convergenza degli integrali,

$$\lim_{h \rightarrow +\infty} \int_X f_h d\mu = \alpha.$$

Inoltre, dalla convergenza puntuale segue anche che f è integrabile, e per la monotonia dell'integrale, che $\alpha \leq \int_X f d\mu$. Pertanto, se $\alpha = +\infty$ non c'è nulla da dimostrare. Altrimenti, fissati $s \in \mathcal{S}_+(X)$ tale che $s \leq f$ e $0 < c < 1$, poniamo

$$E_h = \{x \in X : f_h(x) > cs(x)\}$$

e osserviamo che $E_h \in \mathcal{E}$ per ogni h , e che $\bigcup_h E_h = X$. Per ogni $x \in X$, o $f(x) = 0$, e allora $x \in E_1$, oppure $f(x) > 0$ e $cs(x) < f(x)$; ne segue:

$$\alpha = \lim_{h \rightarrow +\infty} \int_X f_h d\mu \geq \lim_{h \rightarrow +\infty} \int_{E_h} f_h d\mu \geq \lim_{h \rightarrow +\infty} c \int_{E_h} s d\mu = c \int_X s d\mu,$$

da cui, per $c \rightarrow 1$, $\alpha \geq \int_X s d\mu$. Per l'arbitrarietà di $s \leq f$, si ha $\alpha \geq \int_X f d\mu$ e la tesi è provata. $\square$

Osservazione 1.24 L'ipotesi di non negatività sulle funzioni f_h può essere indebolita richiedendo che esista una funzione $g \in L^1(X)$ tale che

$$f_h(x) \geq g(x) \quad \forall x \in X, \quad h \geq 1.$$

Basta infatti applicare il teorema sopra a $f_h - g$. Senza alcuna ipotesi il teorema di convergenza monotona può essere falso: ad esempio se $X = \mathbb{R}$ e $f_h = -\chi_{\mathbb{R} \setminus [-h, h]}$ allora $f \equiv 0$ e

$$\int_{\mathbb{R}} f_h(x) dx = -\infty \quad \forall h, \quad \text{mentre} \quad \int_{\mathbb{R}} f(x) dx = 0.$$

Una importante conseguenza del teorema di convergenza monotona è il fatto che le operazioni di serie e di integrale commutano

$$(1.8) \quad \int_X \sum_{h=1}^{\infty} f_h d\mu = \sum_{h=1}^{\infty} \int_X f_h d\mu$$

purché tutte le funzioni f_h siano non negative. Infatti basta passare al limite per $N \rightarrow +\infty$ nell'uguaglianza

$$\int_X \sum_{h=1}^N f_h d\mu = \sum_{h=1}^N \int_X f_h d\mu$$

ed il teorema di convergenza monotona garantisce che

$$\lim_{N \rightarrow +\infty} \int_X \sum_{h=1}^N f_h d\mu = \int_X \lim_{N \rightarrow +\infty} \sum_{h=1}^N f_h d\mu = \int_X \sum_{h=1}^{\infty} f_h d\mu.$$

È importante osservare che quando l'implicazione (1.7) non vale, c'è comunque una relazione tra il limite degli integrali e l'integrale del limite.

Lemma 1.25 (Lemma di Fatou) *Sia $(f_h) \subset L^1(X)$ una successione di funzioni misurabili a valori in $[0, +\infty]$. Allora*

$$\int_X \liminf_{h \rightarrow +\infty} f_h d\mu \leq \liminf_{h \rightarrow +\infty} \int_X f_h d\mu.$$

Dim. Posto per ogni $h \in \mathbb{N}$, $f = \liminf_{h \rightarrow \infty} f_h$ e $g_h = \inf_{k \geq h} f_k$, si ha che g_h è una successione crescente che converge puntualmente a f in X , ed inoltre ovviamente vale $g_h \leq f_h$ per ogni h . Applicando il teorema della convergenza monotona alla successione (g_h) , risulta:

$$\int_X f d\mu = \lim_{h \rightarrow \infty} \int_X g_h d\mu = \liminf_{h \rightarrow \infty} \int_X g_h d\mu \leq \liminf_{h \rightarrow \infty} \int_X f_h d\mu.$$

□

Osserviamo che, come per il teorema della convergenza monotona, si può sostituire l'ipotesi che le f_h siano tutte positive con l'ipotesi che esista una funzione $g \in L^1(X)$ tale che $f_h \geq g$ per ogni h .

Per trattare limiti non monotoni di successioni è utile il

Teorema 1.26 (Teorema di Lebesgue della convergenza dominata) *Se la successione $(f_h) \subset L^1(X)$ converge puntualmente a $f : X \rightarrow \mathbb{R}$ per μ -quasi ogni $x \in X$ ed esiste una funzione $g \in L^1(X)$ tale che*

$$(1.9) \quad |f_h(x)| \leq g(x) \quad \text{per q.o. } x \in X, \quad h \geq 1,$$

allora $f \in L^1(X)$ e

$$\lim_{h \rightarrow +\infty} \int_X f_h d\mu = \int_X f d\mu.$$

Dim. Notiamo che $f(x) \leq g(x)$ q.o. in X , e che $|f - f_h| \rightarrow 0$, sicché $0 \leq 2g - |f - f_h| \rightarrow 2g$. Ricordiamo anche che vale la seguente relazione tra $\liminf$ e $\limsup$, che segue subito dalle proprietà dell'estremo superiore e dell'estremo inferiore:

$$\liminf_{h \rightarrow \infty} -a_h = -\limsup_{h \rightarrow \infty} a_h$$

per ogni successione reale (a_h) (vedi l'Appendice). Applicando il Lemma di Fatou alla successione $2g - |f - f_h|$ si ha:

$$\begin{aligned} 0 &\leq \int_X 2g = \int_X \lim_{h \rightarrow \infty} (2g - |f - f_h|) \leq \liminf_{h \rightarrow \infty} \int_X (2g - |f - f_h|) \\ &= \int_X 2g + \liminf_{h \rightarrow \infty} \left(- \int_X |f - f_h| \right) = \int_X 2g - \limsup_{h \rightarrow \infty} \int_X |f - f_h| \end{aligned}$$

da cui $\limsup_{h \rightarrow \infty} \int_X |f - f_h| = 0$ e quindi

$$\limsup_{h \rightarrow \infty} \left| \int_X f - \int_X f_h \right| \leq \limsup_{h \rightarrow \infty} \int_X |f - f_h| = 0$$

e il teorema è dimostrato. $\square$

Osservazione 1.27

- (a) Il teorema della convergenza dominata può essere falso se non vale la (1.9):
sia ad esempio $X = \mathbb{R}$ e

$$f_h(x) = \begin{cases} h^3 x(1/h - x) & \text{se } x \in [0, 1/h]; \\ 0 & \text{altrimenti.} \end{cases}$$

Allora f_h converge puntualmente a zero in $\mathbb{R}$ ma

$$\int_{\mathbb{R}} f_h(x) dx = \frac{1}{6}.$$

- (b) Con lo stesso ragionamento usato per dedurre la (1.8) si dimostra che

$$\sum_{h=1}^{\infty} \int_X f_h d\mu = \int_X \sum_{h=1}^{\infty} f_h d\mu$$

purché esista una funzione $g \in L^1(X)$ tale che

$$\left| \sum_{i=1}^N f_i(x) \right| \leq g(x) \quad \forall x \in X, \quad N \geq 1.$$

Per descrivere meglio la struttura dello spazio $L^1(X)$ introduciamo le nozioni contenute nella seguente

Definizione 1.28 (Spazi di Banach) *Se E è uno spazio vettoriale (reale o complesso), si dice norma una funzione $\|\cdot\|$ su E tale che*

- $\|x\| \geq 0 \quad \forall x \in E, \quad \|x\| = 0 \iff x = 0;$
- $\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R} \text{ (}\mathbb{C}\text{)}, \quad \forall x \in E;$
- $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in E$ (diseguaglianza triangolare).

La norma induce una distanza tra i punti di E , $d(x, y) = \|x - y\|$, e quindi una nozione di convergenza di successioni, $x_n \rightarrow x$ per $n \rightarrow \infty$ se e solo se $d(x_n, x) = \|x_n - x\| \rightarrow 0$ per $n \rightarrow \infty$. Una successione $(x_n) \subset E$ si dice *successione di Cauchy* se per ogni $\varepsilon > 0$ esiste $\nu > 0$ tale che $\|x_n - x_k\| < \varepsilon$ per ogni $n, k > \nu$. Uno spazio normato E si dice *completo*, o *spazio di Banach*, se ogni successione di Cauchy in E converge ad un punto di E .

Consideriamo ora per $f \in L^1(X)$ la funzione

$$\|f\|_1 = \int_X |f| d\mu.$$

Si osservi che a rigore questa non è una norma in quanto $\|f\|_1 = 0$ non implica $f = 0$ ma solo $f(x) = 0$ per μ -quasi ogni x . A tutti gli effetti si può lavorare con $\|f\|_1$ come se fosse una norma, considerando equivalenti due funzioni che differiscono in un insieme trascurabile.

Sia ora (f_h) come nel teorema della convergenza dominata. Passando al limite nella (1.9) otteniamo che $|f| \leq g$ quasi ovunque in E . Cambiando se necessario f in un insieme di misura nulla possiamo supporre che la disuguaglianza valga ovunque; dato che $|f_h - f| \leq 2g$ otteniamo

$$\lim_{h \rightarrow +\infty} \int_X |f_h(x) - f(x)| dx = 0.$$

Quindi convergenza puntuale e dominata implica la convergenza nella metrica di $L^1(X)$. In generale la convergenza in $L^1(X)$ non implica la convergenza quasi ovunque: ad esempio le funzioni

$$\begin{cases} f_1 = \chi_{[0,1]} \\ f_2 = \chi_{[0,1/2]}, \quad f_3 = \chi_{[1/2,1]} \\ f_4 = \chi_{[0,1/3]}, \quad f_5 = \chi_{[1/3,2/3]}, \quad f_6 = \chi_{[2/3,1]} \\ \dots \end{cases}$$

convergono a 0 in $L^1([0,1])$ ma non convergono quasi ovunque. Si ha tuttavia convergenza quasi ovunque di sottosuccessioni:

Proposizione 1.29 *Sia $(f_h) \subset L^1(X)$ convergente a $f \in L^1(X)$. Allora esiste una sottosuccessione $g_k = f_{h_k}$ convergente μ -quasi ovunque in X a f .*

Dim. Siano, per $k \geq 1$, h_k indici scelti in modo tale che $\|f_{h_k} - f\|_1 < 2^{-k}$. Posto $g_k = f_{h_k}$ abbiamo

$$\sum_{k=1}^{\infty} \int_X |g_k(x) - f(x)| d\mu \leq 1.$$

Commutando l'integrale con la serie otteniamo

$$\int_X \sum_{k=1}^{\infty} |g_k(x) - f(x)| d\mu \leq 1 < +\infty$$

quindi la (1.8) implica

$$\sum_{k=1}^{\infty} |g_k(x) - f(x)| < +\infty$$

per μ -quasi ogni $x \in X$. Per ogni x con la proprietà sopra i numeri $g_k(x)$ convergono a $f(x)$. $\square$

Con un ragionamento simile a quello usato nella proposizione precedente si verifica la proprietà di completezza di $L^1(X)$:

Proposizione 1.30 *Sia $(f_h) \subset L^1(X)$ una successione di Cauchy. Allora esiste una funzione $f \in L^1(X)$ tale che $\|f_h - f\|_1$ tende a zero.*

Dim. Ragionando come nella dimostrazione precedente, estraiamo dalla (f_h) una sottosuccessione (f_{h_k}) tale che $\|f_{h_{k+1}} - f_{h_k}\| \leq 2^{-k}$. Posto

$$g_n = \sum_{k=1}^n |f_{h_{k+1}} - f_{h_k}|, \quad g = \sum_{k=1}^{\infty} |f_{h_{k+1}} - f_{h_k}|,$$

dalla disuguaglianza triangolare segue che $\|g_k\| \leq 1$ per ogni k , e quindi dal Lemma di Fatou $\|g\| \leq 1$. In particolare, $g(x) < \infty$ q.o., sicché la serie

$$f_{h_1} + \sum_{k=1}^{\infty} f_{h_{k+1}} - f_{h_k}$$

converge assolutamente ad una funzione che possiamo denotare f e che, essendo la serie precedente telescopica, è anche il limite puntuale delle f_{h_k} . A questo punto, fissato $\varepsilon > 0$, sia $\nu > 0$ tale che $\|f_n - f_k\| > \varepsilon$ per $n, k > \nu$. Allora, dal Lemma di Fatou

$$\int_X |f - f_k| d\mu \leq \liminf_{n \rightarrow \infty} \int_X |f_n - f_k| d\mu > \varepsilon$$

e quindi $f \in L^1(X)$ e, per l'arbitrarietà di ε , $\|f - f_k\| \rightarrow 0$. $\square$

La proposizione precedente si può riformulare dicendo che $L^1(X)$ è uno spazio di Banach. È di estremo interesse nelle applicazioni anche lo spazio

$$(1.10) \quad L^2(E) = \left\{ u : E \rightarrow \mathbb{R} : \int_E |u|^2 d\mu < \infty \right\}$$

dove μ è una misura su uno spazio misurabile $(X, \mathcal{E})$ ed $E \in \mathcal{E}$. Lo spazio $L^2(E)$ è uno spazio normato completo con norma

$$\|u\|_2 = \left(\int_E |u|^2 d\mu \right)^{1/2}.$$

Ricordiamo che è stato introdotto anche lo spazio $L^\infty(E)$ delle funzioni misurabili f *essenzialmente limitate*, cioè tali che esiste una costante C tale che

$$\mu(\{x \in E : |f(x)| > C\}) = 0$$

o, equivalentemente

$$|f(x)| \leq C \quad \text{per } \mu\text{-q.o. } x \in E.$$

La minima costante C con questa proprietà (esiste per la numerabile additività della misura) è indicata con $\|f\|_\infty$. Per la norma $\|\cdot\|_\infty$ in $L^\infty(E)$ valgono le stesse considerazioni già fatte per gli spazi $L^1(E)$, $L^2(E)$, ed inoltre anche $L^\infty(E)$ è completo.

Introduciamo anche la seguente notazione. Diciamo che $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è in $L^1_{\text{loc}}(\mathbb{R}^n)$ se $f \in L^1(E)$ per ogni insieme misurabile e limitato $E \subset \mathbb{R}^n$. In modo analogo si definisce $L^2_{\text{loc}}(\mathbb{R}^n)$.

1.6 Misure reali e vettoriali

In questo paragrafo introduciamo le nozioni di misura reale e vettoriale, che sono utili in varie applicazioni e ritroveremo nell'ambito delle distribuzioni.

Definizione 1.31 Se $(X, \mathcal{E})$ è uno spazio misurabile, $m \in \mathbb{N}$, definiamo *misura vettoriale* ogni funzione d'insieme $\mu : \mathcal{E} \rightarrow \mathbb{R}^m$, tale che $\mu(\emptyset) = 0$ e

$$(1.11) \quad \mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j),$$

per ogni successione $\{E_j\} \subset \mathcal{E}$ di insiemi a due a due disgiunti.

In particolare, se nella definizione precedente $m = 1$, allora si parla di *misura reale*. Per $m = 2$, abbiamo una *misura complessa*. Siccome le componenti di una misura vettoriale sono misure reali, spesso, in ciò che segue, sarà sufficiente restringersi al caso reale.

Osservazione 1.32 Una misura positiva *non* è un caso particolare di misura reale. Infatti, direttamente dalla definizione discende che una misura reale è finita, mentre una misura positiva generalmente non lo è. Inoltre, se μ è una misura reale, allora in (1.11) abbiamo una serie convergente, la cui somma non dipende dall'ordine degli addendi, in quanto l'unione insiemistica a primo membro è indipendente dall'ordine. Pertanto la serie in questione è assolutamente convergente, dal teorema del riordinamento di Riemann, vedi in Appendice.

Esempio 1.33 Siano $\{x_k\} \subset X$ e $\{c_k\} \subset \mathbb{R}^m$ due successioni con $\sum_k |c_k| < +\infty$. Definiamo $\mu = \sum_{k=1}^{\infty} c_k \delta_{x_k}$, ossia

$$\mu(E) = \sum_{k: x_k \in E} c_k, \quad E \subset X.$$

Otteniamo così una misura vettoriale definita in $\mathcal{P}(X)$.

Vogliamo affrontare ora il problema di costruire a partire da una misura reale (o vettoriale), μ , una misura positiva λ che domini μ , nel senso che $|\mu(E)| \leq \lambda(E)$, per ogni misurabile E , e che sia, in qualche modo, *la più piccola*. A tal proposito, notiamo che, se $\{E_n\}$ è una partizione misurabile di E , allora deve essere necessariamente

$$(1.12) \quad \lambda(E) = \sum_{n=1}^{\infty} \lambda(E_n) \geq \sum_{n=1}^{\infty} |\mu(E_n)|.$$

Dato che la partizione è arbitraria, $\lambda(E)$ sarà almeno uguale all'estremo superiore delle somme nell'ultimo membro di (1.12). Ciò giustifica la seguente definizione.

Definizione 1.34 Sia μ una misura vettoriale su $(X, \mathcal{E})$. Per ogni $E \in \mathcal{E}$, poniamo

$$|\mu|(E) = \sup \left\{ \sum_{j=1}^{\infty} |\mu(E_j)| \mid E = \bigcup_{j=1}^{\infty} E_j, \ E_i \in \mathcal{E}, \ E_i \cap E_j = \emptyset, i \neq j \right\}.$$

$|\mu|$ si chiama *variazione totale* di μ .

È immediato vedere che $|\mu(E)| \leq |\mu|(E)$ (basta prendere la partizione costituita dal solo E). Il risultato più importante riguardante la variazione totale di una misura è stabilito nel teorema che segue, di cui è omessa la dimostrazione.

Teorema 1.35 $|\mu|$ è una misura positiva finita.

La variazione totale $|\mu|$ è minimale, nel senso che se λ è un'altra misura positiva tale che $|\mu(E)| \leq \lambda(E)$ per ogni $E \in \mathcal{E}$, allora $|\mu|(E) \leq \lambda(E)$.

Esempio 1.36 Sia $(X, \mathcal{E}, \mu)$ uno spazio di misura, prendiamo $f \in L^1(X)$ e definiamo $\nu(E) = \int_E f d\mu$ per ogni $E \in \mathcal{E}$. Allora $|\mu|(E) = \int_E |f| d\mu$.

Infatti, è facile vedere che $|\mu|(E) \leq \int_E |f|$. D'altra parte, se decomponiamo E nell'unione disgiunta di $E^+ = E \cap \{f \geq 0\}$ ed $E^- = E \cap \{f < 0\}$ e consideriamo $\{E_h^+\}$, $\{E_h^-\}$ partizioni misurabili di E^+ ed E^- , rispettivamente, allora

$$\begin{aligned} \sum_{h=1}^{\infty} |\mu(E_h^+)| &= \int_{E^+} f = \int_{E^+} |f|, \\ \sum_{h=1}^{\infty} |\mu(E_h^-)| &= - \int_{E^-} f = \int_{E^-} |f|. \end{aligned}$$

Siccome $\{E_h^+\} \cup \{E_h^-\}$ è una partizione di E , troviamo che $|\mu|(E) \geq \sum_h |\mu(E_h^+)| + \sum_h |\mu(E_h^-)| = \int_{E^+} |f| + \int_{E^-} |f| = \int_E |f|$, che è l'altra disuguaglianza.

Osservazione 1.37 [Decomposizioni di una misura reale]

(a) Per ogni misura reale μ le funzioni

$$\mu^+ = \frac{\mu + |\mu|}{2}, \quad \mu^- = \frac{|\mu| - \mu}{2},$$

sono misure positive finite, dette rispettivamente *parte positiva* e *parte negativa* di μ . Valgono dunque le relazioni

$$\mu = \mu^+ - \mu^-, \quad |\mu| = \mu^+ + \mu^-.$$

La prima delle due identità di sopra è nota come *decomposizione di Jordan*.

(b) Se μ è una misura reale su $(X, \mathcal{E})$, posto

$$\lambda(E) = \sup\{\mu(A) : A \subset E\}, \quad \nu(E) = \sup\{-\mu(A) : A \subset E\}$$

per ogni $E \in \mathcal{E}$, si può provare che esistono due insiemi disgiunti X^+ , $X^- \in \mathcal{E}$ tali che $\lambda(E) = \mu(E \cap X^+)$ e $\nu(E) = -\mu(E \cap X^-)$ per ogni $E \in \mathcal{E}$, e che $\lambda = \mu^+$, $\nu = \mu^-$. Questa proprietà è nota come *decomposizione di Hahn*.

1.7 Singolarità e assoluta continuità di misure

Definizione 1.38 Siano $(X, \mathcal{E})$ uno spazio misurabile, μ una misura positiva e ν una misura vettoriale in $(X, \mathcal{E})$. Diremo che ν è assolutamente continua rispetto a μ e scriveremo

$$\nu \ll \mu$$

se

$$\forall B \in \mathcal{E}, \quad \mu(B) = 0 \implies |\nu|(B) = 0.$$

Diremo che due misure positive μ_1, μ_2 in $(X, \mathcal{E})$ sono (mutuamente) singolari e scriveremo

$$\mu_1 \perp \mu_2,$$

se esiste un insieme misurabile E tale che

$$\mu_1(E) = 0 \quad \text{e} \quad \mu_2(X \setminus E) = 0.$$

Se μ_1 o μ_2 o entrambe sono vettoriali, allora si diranno singolari se lo sono le rispettive variazioni totali.

Se μ è una misura positiva in $(X, \mathcal{E})$, diremo che μ è concentrata in $E \in \mathcal{E}$, se $\mu(X \setminus E) = 0$. Se μ è vettoriale, come al solito la precedente condizione va intesa per $|\mu|$.

Direttamente dalle definizioni discende che due misure sono singolari se e solo se sono concentrate in insiemi disgiunti, per cui due misure singolari vivono su porzioni diverse di X .

Esempio 1.39 Consideriamo $A = \{r_n\}$ un sottoinsieme numerabile di $\mathbb{R}$ e sia λ la misura di Lebesgue. Prendiamo una successione di numeri reali positivi $\{p_n\}$ tale che $\sum_{n=1}^{\infty} p_n < \infty$ e definiamo $\mu = \sum_{n=1}^{\infty} p_n \delta_{r_n}$. Si verifica facilmente che $\mu \perp \lambda$.

Un'utile caratterizzazione dell'assoluta continuità è stabilita nella seguente proposizione. ■

Proposizione 1.40 Siano ν una misura vettoriale e μ una misura positiva in $(X, \mathcal{E})$. Sono equivalenti

(a) $\nu \ll \mu$,

(b) $\forall \varepsilon > 0 \exists \delta > 0$ t.c. se $E \in \mathcal{E}$ e $\mu(E) < \delta$ allora $|\nu|(E) < \varepsilon$.

DIM. (b) $\Rightarrow$ (a): Sia $B \in \mathcal{E}$ tale che $\mu(B) = 0$. Se $|\nu|(B) > 0$, allora esisterebbe $\bar{\varepsilon} > 0$ tale che $|\nu|(B) \geq \bar{\varepsilon}$. Per ipotesi, in corrispondenza di $\bar{\varepsilon}$ troveremmo $\delta > 0$ con la proprietà che $|\nu|(E) < \bar{\varepsilon}$ non appena $\mu(E) < \delta$. La scelta $E = B$ porta ad una contraddizione.

(a) $\Rightarrow$ (b): Supponiamo per assurdo che esistano $\bar{\varepsilon} > 0$ ed una successione di insiemi misurabili $\{E_n\}$ tali che $\mu(E_n) < 2^{-n}$ e $|\nu|(E_n) \geq \bar{\varepsilon}$. Posto $E = \bigcap_{n=0}^{\infty} \bigcup_{j=n}^{\infty} E_j$, risulta che $E \in \mathcal{E}$ e

$$\mu(E) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{j=n}^{\infty} E_j\right) \leq \lim_{n \rightarrow \infty} \sum_{j=n}^{\infty} \mu(E_j) \leq \lim_{n \rightarrow \infty} \sum_{j=n}^{\infty} 2^{-j} = 0,$$

$$\text{mentre } |\nu|(E) = \lim_{n \rightarrow \infty} |\nu|\left(\bigcup_{j=n}^{\infty} E_j\right) \geq \lim_{n \rightarrow \infty} |\nu|(E_n) \geq \bar{\varepsilon}. \quad \square$$

Osserviamo che se la misura ν non è finita, allora nella proposizione precedente l'implicazione (b) $\Rightarrow$ (a) continua a valere, mentre quella inversa non è garantita. Ad esempio, se prendiamo al posto di μ la misura di Lebesgue in $(0, 1)$ e definiamo

$$\nu(E) = \int_E \frac{1}{t} dt,$$

per ogni insieme $E \subset (0, 1)$ misurabile secondo Lebesgue, allora $\nu \ll \mu$, nel senso che $\nu(E) = 0$ se $\mu(E) = 0$. Tuttavia, per ogni $n \in \mathbb{N}$ l'insieme $E_n = (0, 1/n)$ è tale che $\mu(E_n) = 1/n$ mentre $\nu(E_n) = +\infty$.

Vediamo adesso una classe importante di misure assolutamente continue.

Proposizione 1.41 *Sia μ una misura positiva in $(X, \mathcal{E})$ spazio misurabile e sia $f \in [L^1(X, \mu)]^m$ (cioè f ha m componenti in $L^1(X, \mu)$). Poniamo*

$$\mu_f(E) = \int_E f d\mu, \quad E \in \mathcal{E}.$$

Allora μ_f è una misura vettoriale e $|\mu_f|(E) = \int_E |f| d\mu$. In particolare, μ_f è assolutamente continua rispetto a μ .

DIM. La prima parte dell'enunciato è conseguenza del teorema di convergenza dominata. Verifichiamo la formula relativa alla variazione totale. Sia $\{z_n\}$ una successione densa in $\mathbb{S}^{m-1} = \{x \in \mathbb{R}^m : |x| = 1\}$. Per ogni $\varepsilon > 0$ e $x \in X$, poniamo

$$s(x) = \min\{h \in \mathbb{N} \mid \langle f(x), z_h \rangle \geq (1 - \varepsilon)|f(x)|\}.$$

La definizione di $s(x)$ è ben posta perché la densità della successione $\{z_n\}$ in $\mathbb{S}^{m-1}$ assicura che l'insieme in questione sia non vuoto. Poniamo

$$B_n = B \cap s^{-1}(\{z_n\}).$$

Si vede che $B_n \in \mathcal{E}$, $B_n \cap B_k = \emptyset$ se $n \neq k$ e $\bigcup_{n \in \mathbb{N}} B_n = B$. Pertanto

$$\begin{aligned} (1 - \varepsilon) \int_B |f| d\mu &= (1 - \varepsilon) \sum_{n=1}^{\infty} \int_{B_n} |f| d\mu \leq \sum_{n=1}^{\infty} \int_{B_n} \langle f(x), z_n \rangle d\mu \\ &= \sum_{n=1}^{\infty} \langle \int_{B_n} f(x) d\mu, z_n \rangle \leq \sum_{n=1}^{\infty} \left| \int_{B_n} f(x) d\mu \right| \\ &= \sum_{n=1}^{\infty} |\mu_f(B_n)| \leq |\mu_f|(B). \end{aligned}$$

Abbiamo così stabilito che $\int_B |f| d\mu \leq |\mu_f|(B)$. Dato che l'altra disuguaglianza è evidente, l'asserto risulta provato. $\square$

È importante sapere che questa rappresenta di fatto l'unico esempio di misura assolutamente continua, come mostra il seguente risultato concernente la continuità assoluta, ma anche uno dei più importanti in teoria della misura.

Teorema 1.42 *Siano μ una misura positiva e ν una misura vettoriale, a valori in $\mathbb{R}^m$, definite in uno spazio misurabile $(X, \mathcal{E})$. Assumiamo che μ sia σ -finita. Allora esiste un'unica coppia (ν^a, ν^s) di misure vettoriali, a valori in $\mathbb{R}^m$, tali che*

$$(1.13) \quad \begin{aligned} \nu^a &<< \mu, \\ \nu^s &\perp \mu, \\ \nu &= \nu^a + \nu^s. \end{aligned}$$

Inoltre, esiste un'unica $f \in [L^1(X, \mu)]^m$ tale che $\nu^a = \mu_f$.

La coppia formata da ν^a e ν^s viene chiamata la *decomposizione di Lebesgue* di ν rispetto a μ . L'unicità della decomposizione si verifica facilmente, poiché se ν_1^a e ν_1^s è un'altra coppia che soddisfa (1.13), si ha $\nu_1^a - \nu^a = \nu_1^s - \nu^s$. Siccome $\nu_1^a - \nu^a << \mu$ e $\nu_1^s - \nu^s \perp \mu$, necessariamente $\nu_1^a - \nu^a = \nu_1^s - \nu^s = 0$. L'esistenza della decomposizione (1.13) è la parte significativa del primo punto. La seconda parte del Teorema 1.42 è nota sotto il nome di *Teorema di Radon-Nikodym*. L'unicità di f è ancora una volta immediata, mentre il punto essenziale è l'esistenza. Di fatto, si afferma che *ogni* misura assolutamente continua è del tipo μ_f , per qualche f .

La funzione f si chiama *derivata di Radon-Nikodym* di ν^a rispetto a μ e si indica a volte con $\frac{d\nu^a}{d\mu}$.

Il teorema può essere esteso al caso in cui ν è una misura positiva σ -finita.

Si può far vedere che l'ipotesi che le misure siano σ -finite non può essere rimossa. Per esempio, se μ è la misura di Lebesgue in $(0, 1)$ e ν la misura del contare in $(0, 1)$, allora non esiste una decomposizione di Lebesgue di ν rispetto a μ .

1.8 Misura prodotto

Siano $(X, \mathcal{E})$ e $(Y, \mathcal{F})$ due spazi misurabili. Per ogni $E \in \mathcal{E}$ ed $F \in \mathcal{F}$, chiamiamo $E \times F$ *rettangolo misurabile*. Se μ e λ sono due misure positive su $(X, \mathcal{E})$ e $(Y, \mathcal{F})$, rispettivamente, ci poniamo il problema di costruire una misura ν su una σ -algebra opportuna di $X \times Y$ con la proprietà che

$$\nu(E \times F) = \mu(E)\lambda(F), \quad \forall E \in \mathcal{E}, F \in \mathcal{F}.$$

Individuiamo innanzitutto la σ -algebra adatta. Deve necessariamente contenere i rettangoli misurabili, ma questi da soli non costituiscono una σ -algebra, quindi consideriamo

$$\mathcal{E} \times \mathcal{F} := \sigma(\{E \times F \mid E \in \mathcal{E}, F \in \mathcal{F}\}).$$

La scrittura $\mathcal{E} \times \mathcal{F}$ *non* indica il prodotto cartesiano tra $\mathcal{E}$ ed $\mathcal{F}$, ma la σ -algebra generata da tale prodotto. Per ogni $Q \in \mathcal{E} \times \mathcal{F}$, definiamo le *sezioni* di Q tramite le seguenti espressioni

$$\begin{aligned} Q_x &= \{y \in Y \mid (x, y) \in Q\}, & x \in X, \\ Q_y &= \{x \in X \mid (x, y) \in Q\}, & y \in Y. \end{aligned}$$

Lemma 1.43 *Per ogni $x \in X$, $y \in Y$, $Q \in \mathcal{E} \times \mathcal{F}$ risulta $Q_x \in \mathcal{F}$ e $Q_y \in \mathcal{E}$.*

DIM. Poniamo $\mathcal{G} = \{Q \in \mathcal{E} \times \mathcal{F} \mid Q_x \in \mathcal{F}, \forall x \in X\}$. Si vede facilmente che $\mathcal{G}$ è una σ -algebra. Inoltre, se $E \times F$ è un rettangolo misurabile, allora $(E \times F)_x = F$, se $x \in E$, $(E \times F)_x = \emptyset$ se $x \notin E$. In ogni caso, $(E \times F)_x$ è misurabile in Y , per cui $\mathcal{G}$ contiene i rettangoli misurabili. Allora contiene anche la σ -algebra generata da questi, ossia $\mathcal{E} \times \mathcal{F}$. Pertanto $Q_x \in \mathcal{F}$, per ogni $x \in X$ e $Q \in \mathcal{E} \times \mathcal{F}$. Analogamente si dimostra che $Q_y \in \mathcal{E}$, per ogni $y \in Y$ e $Q \in \mathcal{E} \times \mathcal{F}$. $\square$

Data una funzione $f : X \times Y \rightarrow \overline{\mathbb{R}}$ e presi $x \in X$ e $y \in Y$, definiamo le funzioni

$$\begin{aligned} f_x : Y &\rightarrow \overline{\mathbb{R}} & f_y : X &\rightarrow \overline{\mathbb{R}} \\ y &\mapsto f(x, y) & x &\mapsto f(x, y) \end{aligned}$$

Lemma 1.44 *Nella notazione introdotta, se f è $\mathcal{E} \times \mathcal{F}$ -misurabile, allora f_x è $\mathcal{F}$ -misurabile e f_y è $\mathcal{E}$ -misurabile, per ogni $x \in X$, $y \in Y$.*

DIM. Sia $S_t = [t, +\infty]$. Per ipotesi, $f^{-1}(S_t) \in \mathcal{E} \times \mathcal{F}$. Per il Lemma 1.43, $(f^{-1}(S_t))_x \in \mathcal{F}$ e $(f^{-1}(S_t))_y \in \mathcal{E}$, per ogni x e y . Siccome $(f^{-1}(S_t))_x = f_x^{-1}(S_t)$ e $(f^{-1}(S_t))_y = f_y^{-1}(S_t)$, la tesi è provata. $\square$

Teorema 1.45 *Siano $(X, \mathcal{E}, \mu)$ e $(Y, \mathcal{F}, \lambda)$ due spazi dotati di misura positiva σ -finita e sia $Q \in \mathcal{E} \times \mathcal{F}$. Allora*

$$\begin{aligned} x \mapsto \lambda(Q_x) & \quad \text{è } \mathcal{E} - \text{misurabile,} \\ y \mapsto \mu(Q_y) & \quad \text{è } \mathcal{F} - \text{misurabile} \end{aligned}$$

e risulta

$$(1.14) \quad \int_X \lambda(Q_x) d\mu = \int_Y \mu(Q_y) d\lambda.$$

Osserviamo che $\lambda(Q_x)$ e $\mu(Q_y)$ sono ben poste grazie al Lemma 1.43. Inoltre, siccome

$$\lambda(Q_x) = \int_Y \chi_{Q_x}(y) d\lambda,$$

e $\chi_{Q_x}(y) = (\chi_Q)_x(y) = \chi_Q(x, y)$ e analogamente per $\mu(Q_y)$, possiamo riscrivere l'identità (1.14) nella forma

$$\int_X \left(\int_Y \chi_Q(x, y) d\lambda \right) d\mu = \int_Y \left(\int_X \chi_Q(x, y) d\mu \right) d\lambda,$$

così da riconoscere una *formula di inversione dell'ordine di integrazione* per le funzioni caratteristiche. Il Teorema 1.45 rappresenta il passo più importante nella costruzione della misura prodotto. Infatti, grazie a quanto stabilito nel suddetto teorema, se $(X, \mathcal{E}, \mu)$ e $(Y, \mathcal{F}, \lambda)$ sono spazi di misura positiva, σ -finita, per ogni $Q \in \mathcal{E} \times \mathcal{F}$ possiamo definire

$$(1.15) \quad \nu(Q) = \int_X \lambda(Q_x) d\mu \left(= \int_Y \mu(Q_y) d\lambda \right).$$

Proposizione 1.46 *La funzione d'insieme ν definita in (1.15) è una misura positiva σ -finita in $\mathcal{E} \times \mathcal{F}$ e soddisfa $\nu(E \times F) = \mu(E)\lambda(F)$. Inoltre, ν è l'unica misura con queste proprietà.*

DIM. Evidentemente $\nu(\emptyset) = 0$. Se $Q = \bigcup Q_n$, con $Q_n \in \mathcal{E} \times \mathcal{F}$, disgiunti, allora $Q_x = \bigcup (Q_n)_x$ e, dato che λ è σ -additiva, risulta $\lambda(Q_x) = \sum_n \lambda((Q_n)_x)$. Per la (1.8), possiamo scrivere

$$\nu(Q) = \int_X \lambda(Q_x) d\mu = \int_X \sum_n \lambda((Q_n)_x) d\mu = \sum_n \int_X \lambda((Q_n)_x) d\mu = \sum_n \nu(Q_n),$$

da cui la σ -additività di ν . La dimostrazione del fatto che $\nu(E \times F) = \mu(E)\lambda(F)$ segue immediatamente dalla formula (1.15). L'unicità di ν tiene conto del fatto che ν è univocamente determinata sui rettangoli misurabili, che costituiscono un sistema di generatori per $\mathcal{E} \times \mathcal{F}$. $\square$

La misura ν prende il nome di *misura prodotto*. A volte è indicata con $\mu \times \lambda$. Il seguente teorema ora è facile conseguenza della costruzione fatta.

Teorema 1.47 (Fubini) *Siano $(X, \mathcal{E}, \mu)$ e $(Y, \mathcal{F}, \lambda)$ due spazi dotati di misura positiva σ -finita. Sia $f : \mathcal{E} \times \mathcal{F} \rightarrow \overline{\mathbb{R}}$ una funzione $\mathcal{E} \times \mathcal{F}$ -misurabile.*

(a) *Se $f \geq 0$ allora le funzioni*

$$\varphi : x \mapsto \int_Y f_x(y) d\lambda, \quad \psi : y \mapsto \int_X f_y(x) d\mu$$

sono $\mathcal{E}$ -misurabile e $\mathcal{F}$ -misurabile, rispettivamente, e vale

$$\int_X \left(\int_Y f(x, y) d\lambda \right) d\mu = \int_{X \times Y} f(x, y) d\nu = \int_Y \left(\int_X f(x, y) d\mu \right) d\lambda.$$

(formula di riduzione)

(b) *Se f è a valori reali e se*

$$\varphi^*(x) = \int_Y |f|_x d\lambda \quad \text{e} \quad \int_X \varphi^* d\mu < \infty,$$

allora $f \in L^1(X \times Y, \nu)$.

(c) *Se $f \in L^1(X \times Y, \nu)$ allora $f_x \in L^1(Y, \lambda)$ per μ -q.o. x e $f_y \in L^1(X, \mu)$ per λ -q.o. y . Le funzioni definite in (a) sono sommabili nei rispettivi spazi e vale ancora la formula di riduzione per f .*

DIM. (a): Se $f = \chi_Q$, con $Q \in \mathcal{E} \times \mathcal{F}$, allora la tesi è già verificata grazie al Teorema 1.45. Per linearità, essa si estende al caso delle funzioni semplici positive. Supponiamo ora che f sia come in (a). Allora, per il teorema di approssimazione

(Teorema 1.12), esiste una successione di funzioni semplici positive s_n che convergono a f puntualmente e in modo crescente. Dato che $(s_n)_x$ convergono nello stesso modo a f_x , per il teorema di convergenza monotona risulta che

$$\int_Y (s_n)_x(y) d\lambda \longrightarrow \int_Y f_x(y) d\lambda = \varphi(x), \quad \text{per } n \rightarrow \infty,$$

da cui segue che φ è $\mathcal{E}$ -misurabile, in quanto limite di funzioni che sono a loro volta $\mathcal{E}$ -misurabili grazie al passo precedente. Analogamente, ψ è $\mathcal{F}$ -misurabile. Infine, la formula di riduzione per f è conseguenza del fatto che le approssimanti s_n la verificano e si può passare al limite, grazie ancora al teorema di convergenza monotona.

(b): Segue subito applicando (a) a $|f|$.

(c): Sia f una funzione sommabile nello spazio prodotto. Indichiamo con φ_1 e φ_2 le funzioni che corrispondono a f^+ e f^- così come φ corrisponde a f . Siccome (a) è verificata da f^+ , possiamo scrivere

$$\int_X \varphi_1(x) d\mu = \int_{X \times Y} f^+(x, y) d\nu \leq \int_{X \times Y} |f(x, y)| d\nu,$$

da cui segue che $\varphi_1 \in L^1(X, \mu)$. Allo stesso modo $\varphi_2 \in L^1(X, \mu)$. Ne segue che per μ -q.o. $x \in X$, $\varphi_1(x) < \infty$ e $\varphi_2(x) < \infty$. Per tali x , risulta $f_x \in L^1(Y, \lambda)$, essendo $f_x = (f^+)_x - (f^-)_x$. Inoltre, per gli stessi x , $\varphi(x) = \varphi_1(x) - \varphi_2(x)$, per cui $\varphi \in L^1(X, \mu)$. Scambiando il ruolo di x e y si ottiene l'asserto relativo a f_y e ψ . Infine, la formula di riduzione per f si ottiene da quelle per f^+ ed f^- (che verificano (a)), sottraendo membro a membro. $\square$

Notiamo che (b) e (c) insieme implicano il seguente risultato: se f è $\mathcal{E} \times \mathcal{F}$ -misurabile e se

$$\int_X \left(\int_Y |f(x, y)| d\lambda \right) d\mu < \infty,$$

allora i due integrali iterati della formula di riduzione sono finiti e coincidono.

Esempio 1.48 Siano $X = Y = [0, 1]$ e prendiamo al posto di λ la misura del contare e al posto di μ la misura di Lebesgue, definite nelle rispettive σ -algebre. Consideriamo la funzione $f = \chi_\Delta$ dove $\Delta = \{(x, y) \in [0, 1]^2 : x = y\}$ è la parte di diagonale nel quadrato unitario. Per vedere che f è misurabile, basta provare che Δ è misurabile. Per ogni $n \in \mathbb{N}$, suddividiamo l'intervallo $[0, 1]$ in n parti uguali $\{I_j\}_{1 \leq j \leq n}$ e consideriamo $Q_n = I_1^2 \cup \dots \cup I_n^2$. Q_n è unione finita di rettangoli misurabili e come tale è misurabile nello spazio prodotto. Siccome $\Delta = \bigcap_n Q_n$,

abbiamo che anche Δ è misurabile. Per la funzione f considerata non vale la formula di riduzione giacché

$$\int_0^1 f_y(x) d\mu = 0, \quad \text{per ogni } y \in [0, 1],$$

e

$$\int_0^1 f_x(y) d\lambda = 1, \quad \text{per ogni } x \in [0, 1],$$

per cui

$$\int_Y \left(\int_X f(x, y) d\mu \right) d\lambda = 0 \neq 1 = \int_X \left(\int_Y f(x, y) d\lambda \right) d\mu.$$

Il problema è dovuto al fatto che la misura λ non è σ -finita.

Osservazione 1.49 La misura prodotto di due misure complete non è necessariamente completa.

Come applicazione del Teorema di Fubini, proviamo il seguente risultato dovuto a Choquet.

Teorema 1.50 Sia $f : X \rightarrow [0, +\infty]$ una funzione misurabile in $(X, \mathcal{E}, \mu)$, spazio di misura positiva e σ -finita. Allora

$$\int_X f d\mu = \int_0^{+\infty} \mu(\{x \in X : f(x) > t\}) dt,$$

dove, al secondo membro, compare l'integrale di Lebesgue in dimensione uno.

DIM. Applichiamo il Teorema di Fubini al prodotto della misura μ assegnata con la misura di Lebesgue $\mathcal{L}^1$ ristretta alla semiretta $[0, +\infty[$. Posto $E_t = \{f > t\}$, otteniamo

$$\begin{aligned} \int_X f(x) d\mu &= \int_X \left(\int_0^{f(x)} dt \right) d\mu = \int_X \left(\int_0^{+\infty} \chi_{E_t}(x) dt \right) d\mu \\ &= \int_0^{+\infty} \left(\int_X \chi_{E_t}(x) d\mu \right) dt = \int_0^{+\infty} \mu(E_t) dt \\ &= \int_{X \times [0, +\infty[} \chi_{E_t}(x) d(\mu \times \mathcal{L}^1) = \\ &= (\mu \times \mathcal{L}^1)(\{(x, t) \mid t < f(x)\}). \end{aligned}$$

Notiamo anche che l'uguaglianza tra il primo e l'ultimo membro esprime il fatto che l'integrale di una funzione misurabile positiva non è altro che la misura del suo sottografico. $\square$

1.9 Formule di riduzione e cambiamento di variabili in $\mathbb{R}^n$

In questo paragrafo specializziamo la teoria delle misure prodotto allo spazio $\mathbb{R}^n$, pensato come prodotto cartesiano $\mathbb{R}^p \times \mathbb{R}^k$, con $p + k = n$ e otteniamo, come già visto per l'integrale di Riemann, le *formule di riduzione per gli integrali multipli*. Queste sono una conseguenza immediata del Teorema di Fubini, e sono lo strumento essenziale per il calcolo effettivo degli integrali multipli. Naturalmente, in ciascuno spazio euclideo consideriamo la misura di Lebesgue della dimensione corrispondente.

Teorema 1.51 (Formula di riduzione) *Siano $p, k \geq 1$ interi, $n = p + k$ e scriviamo $x = (y, z)$ con $y \in \mathbb{R}^p$ e $z \in \mathbb{R}^k$. Data $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sommabile, per m_p -quasi ogni $y \in \mathbb{R}^p$ la funzione $z \mapsto f(y, z)$ è sommabile in $\mathbb{R}^k$ e*

$$y \mapsto \int_{\mathbb{R}^k} f(y, z) dz$$

è sommabile in $\mathbb{R}^p$. Si ha inoltre

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^k} f(y, z) dz \right) dy.$$

Analogamente, per m_k -quasi ogni $z \in \mathbb{R}^k$ la funzione $y \mapsto f(y, z)$ è sommabile in $\mathbb{R}^p$ e

$$z \mapsto \int_{\mathbb{R}^p} f(y, z) dy$$

è sommabile in $\mathbb{R}^k$. Si ha inoltre

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^p} f(y, z) dy \right) dz.$$

Notiamo che in particolare si ha

$$\int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^k} f(y, z) dz \right) dy = \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^p} f(y, z) dy \right) dz.$$

L'uguaglianza sopra è detta *formula di inversione dell'ordine di integrazione*. Formule simili a quella di riduzione valgono naturalmente anche nel caso in cui la variabile “ y ” non sia data dalle prime p componenti di x e la variabile “ z ” non sia data dalle ultime k componenti di x .

Dato $E \subset \mathbb{R}^n$ misurabile, applicando la formula di riduzione alla funzione $f = \chi_E$ otteniamo che gli insiemi

$$E_y = \{z \in \mathbb{R}^k : (y, z) \in E\} \quad \left(E_z = \{y \in \mathbb{R}^p : (y, z) \in E\} \right)$$

sono misurabili in $\mathbb{R}^k$ ($\mathbb{R}^p$) per m_p -quasi ogni $y \in \mathbb{R}^p$ (m_k -quasi ogni $z \in \mathbb{R}^k$) e

$$\int_{\mathbb{R}^p} m_k(E_y) dy = m_n(E) \quad \left(\int_{\mathbb{R}^k} m_p(E_z) dz = m_n(E) \right).$$

Nel caso dell'integrale su un insieme $E \subset \mathbb{R}^n$, applicando il teorema precedente alla funzione $f\chi_E$ troviamo

$$\int_{\mathbb{R}^p} \left(\int_{E_y} f(y, z) dz \right) = \int_E f(x) dx = \int_{\mathbb{R}^k} \left(\int_{E_z} f(y, z) dy \right) dz.$$

Infine, vediamo una relazione notevole tra l'integrale di una funzione misurabile $f \geq 0$ e la misura $(n+1)$ -dimensionale del sottografico di f :

$$S_f = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : 0 \leq t \leq f(x)\}.$$

Teorema 1.52 (Teorema del sottografico) *Sia $f : \mathbb{R}^n \rightarrow [0, +\infty)$ misurabile. Allora l'insieme S_f è misurabile in $\mathbb{R}^{n+1}$ e*

$$m_{n+1}(S_f) = \int_{\mathbb{R}^n} f(x) dx.$$

Dim. Non dimostreremo la misurabilità di S_f . Usando la formula di riduzione verifichiamo l'uguaglianza sopra:

$$m_{n+1}(S_f) = \int_{\mathbb{R}^n} m_1((S_f)_x) dx = \int_{\mathbb{R}^n} f(x) dx$$

dato che $(S_f)_x = [0, f(x)]$ per ogni $x \in \mathbb{R}^n$. □

Per funzioni di segno qualunque si ha

$$\begin{aligned} \int_{\mathbb{R}^n} f(x) dx &= \int_{\mathbb{R}^n} f^+(x) dx - \int_{\mathbb{R}^n} f^-(x) dx \\ &= m_{n+1}(S_{f^+}) - m_{n+1}(S_{f^-}) \\ &= m_{n+1}(\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : 0 \leq t \leq f(x)\}) \\ &\quad - m_{n+1}(\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq t \leq 0\}). \end{aligned}$$

purché la parte positiva o la parte negativa di f abbiano integrale finito.

Ricordiamo anche la formula di cambiamento di variabili negli integrali multipli, che vale per l'integrale di Lebesgue nella stessa forma in cui s'è visto per l'integrale di Riemann.

Teorema 1.53 *Siano $D, E \subset \mathbb{R}^n$ misurabili e $\Phi : D \rightarrow E$ bigettiva. Supponiamo che Φ sia di classe C^1 in un insieme aperto contenente D . Data $f : \mathbb{R}^n \rightarrow \mathbb{R}$ si ha che f è sommabile in E se e solo se*

$$f(\Phi(x))|\det J\Phi(x)|$$

è sommabile in D e vale l'uguaglianza

$$(1.16) \quad \int_D f(\Phi(x))|\det J\Phi(x)| dx = \int_E f(y) dy.$$

Nel caso particolare $f = \chi_E$ si ha

$$m_n(E) = m_n(\Phi(D)) = \int_D |\det J\Phi(x)| dx.$$

Intuitivamente, la comparsa del determinante di $J\Phi$ nel passaggio da un integrale all'altro tiene conto di come l'applicazione Φ “dilata” o “contrae” gli insiemi. Consideriamo ad esempio $E = B_1((-2, 0)) \cup B_1((2, 0)) \subset \mathbb{R}^2$, $D = B_{1/2}((-2, 0)) \cup B_2((2, 0))$ e una funzione Φ tale che $J\Phi = 2Id$ su $B_{1/2}((-2, 0))$ e $J\Phi = Id/2$ su $B_2((2, 0))$. Allora

$$\begin{aligned} m_2(E) &= \pi + \pi = 4m_2(B_{1/2}((-2, 0))) + \frac{1}{4}m_2(B_2((2, 0))) \\ &= \int_D |\det J\Phi| dx. \end{aligned}$$

In generale, si può verificare che la formula di cambiamento di variabili vale per funzioni lineari Φ ; nel caso generale, dividendo un insieme D in parti molto piccole sulle quali Φ è prossimo ad una funzione lineare (data dal differenziale di Φ) ed usando l'additività della misura si perviene alla formula nel caso generale.

Non dimostreremo la formula di cambiamento di variabili ma la verificheremo in due casi particolari:

(a) Supponiamo $n = 1$, $D = [a, b]$, $E = [c, d]$, Φ di classe C^1 in D e monotona, f continua in E . Dalla formula di cambiamento di variabili vista ad Analisi I otteniamo

$$(1.17) \quad \int_a^b f(\Phi(x))\Phi'(x) dx = \int_{\Phi(a)}^{\Phi(b)} f(y) dy.$$

Se $\Phi' \geq 0$ allora Φ è crescente e $\Phi(a) = c$, $\Phi(b) = d$; si ottiene quindi

$$(1.18) \quad \int_{[a,b]} f(\Phi(x))|\Phi'(x)| dx = \int_{[c,d]} f(y) dy.$$

Se invece $\Phi' \leq 0$ in $[a, b]$ allora $\Phi(a) = d$ e $\Phi(b) = c$; cambiando i segni ad ambo i membri nella (1.17) otteniamo di nuovo (1.18).

(b) Supponiamo che tutte le componenti di Φ tranne una siano l'identità. Per fissare le idee poniamo $x = (z, y)$ con $z \in \mathbb{R}$ e $y \in \mathbb{R}^{n-1}$ e supponiamo che

$$(1.19) \quad \Phi(x) = \Phi(z, y) = (\varphi(z, y), y_1, \dots, y_{n-1}).$$

Si ha allora $\det J\Phi = \partial\varphi/\partial z(z, y) = \varphi'_y(z)$, ove $\varphi_y(z) = \varphi(z, y)$. Posto

$$D_y = \{z \in \mathbb{R} : (z, y) \in D\} \quad E_y = \{z \in \mathbb{R} : (z, y) \in E\}.$$

abbiamo anche

$$E_y = \{\varphi(z, y) : (z, y) \in D\} = \varphi_y(D_y).$$

Usando allora la formula di riduzione e le relazioni scritte sopra otteniamo

$$\begin{aligned} \int_E f(z, y) dz dy &= \int_{\mathbb{R}^{n-1}} \left(\int_{E_y} f(z, y) dz \right) dy \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{D_y} f(\varphi_y(z), y) |\varphi'_y(z)| dz \right) dy \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{D_y} f(\Phi(y, z)) |\det J\Phi(y, z)| dz \right) dy \\ &= \int_D f(\Phi(z, y)) |\det J\Phi(z, y)| dz dy. \end{aligned}$$

In generale, si dimostra che localmente ogni applicazione Φ avente matrice jacobiana non singolare è composizione di applicazioni $\Phi_1, \dots, \Phi_n$ del tipo considerato in (1.19), quindi la (1.16) può essere dedotta, almeno localmente, usando n volte l'argomento nel punto 2. Si passa poi alla formula globale usando l'additività della misura. La formula di cambiamento di variabili è così ricondotta a quella di una variabile.

Tra i vari cambiamenti di variabili ricordiamo le *coordinate polari* in $\mathbb{R}^2$:

$$\Phi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta) \quad \rho > 0, \ 0 \leq \theta < 2\pi$$

con $|\det J\Phi(\rho, \theta)| = \rho$, le *coordinate cilindriche* in $\mathbb{R}^3$:

$$\Phi(\rho, \theta, z) = (\rho \cos \theta, \rho \sin \theta, z) \quad \rho > 0, \ 0 \leq \theta < 2\pi, \ z \in \mathbb{R}$$

con $|\det J\Phi(\rho, \theta, z)| = \rho$ e le *coordinate sferiche* in $\mathbb{R}^3$:

$$\begin{aligned} \Phi(\rho, \theta, \varphi) &= (\rho \cos \theta \sin \varphi, \rho \sin \theta \sin \varphi, \rho \cos \varphi) \\ \rho &> 0, \ 0 \leq \theta < 2\pi, \ 0 < \varphi < \pi \end{aligned}$$

con $|\det J\Phi(\rho, \theta, \varphi)| = \rho^2 \sin \varphi$.

Più in generale, si possono definire *coordinate sferiche in $\mathbb{R}^n$* , usando una variabile lineare $\rho = |x|$ come in $\mathbb{R}^3$ ed $n - 1$ variabili angolari $\omega_1, \dots, \omega_{n-1}$. Non esponiamo i dettagli, limitandoci a segnalare che in questo caso il determinante del cambio di variabili è $\varphi(\omega_1, \dots, \omega_{n-1})\rho^{n-1}$, dove φ è una funzione limitata delle variabili angolari.

1.10 Integrali dipendenti da un parametro

Sia $A \subset \mathbb{R}^m$ un aperto, $E \subset \mathbb{R}^n$ un insieme misurabile e sia $f(t, x) : A \times E \rightarrow \mathbb{R}$. Supponiamo che per ogni $t \in A$ la funzione $x \mapsto f(t, x)$ sia sommabile in E . È allora definita la funzione

$$F(t) = \int_E f(t, x) dx \quad t \in A.$$

Il problema che affronteremo in questo paragrafo è quello della regolarità di F in funzione di quella di f . Incominciamo dalla continuità:

Teorema 1.54 *Supponiamo che $t \mapsto f(t, x)$ sia continua in A per m_n -quasi ogni $x \in E$ ed esista una funzione $g \in L^1(E)$ tale che*

$$(1.20) \quad |f(t, x)| \leq g(x) \quad \forall t \in A, x \in E.$$

Allora F è continua in A .

Dim. Grazie alla caratterizzazione del limite di funzioni tramite limiti di successioni, basta verificare la continuità per successioni. Sia $t \in A$ e $(t_h) \subset A$ convergente a t . Essendo le funzioni $t \mapsto f(t, x)$ continue per m_n -quasi ogni $x \in E$, abbiamo

$$\lim_{h \rightarrow +\infty} f(t_h, x) = f(t, x)$$

per m_n -quasi ogni $x \in E$. Dall'ipotesi (1.20) segue che si può applicare il teorema della convergenza dominata:

$$\lim_{h \rightarrow +\infty} F(t_h) = \lim_{h \rightarrow +\infty} \int_E f(t_h, x) dx = \int_E f(t, x) dx = F(t).$$

□

Esempio. Il teorema precedente può essere falso se non vale la (1.6): sia $A = E = \mathbb{R}$ e

$$f(t, x) = \begin{cases} \frac{|t|-|x|}{t^2} & \text{se } |x| < |t|; \\ 0 & \text{se } |x| \geq |t|. \end{cases}$$

Si trova allora $F(t) = 1$ per $t \neq 0$ e $F(0) = 0$, quindi F non è continua.

Passiamo ora allo studio della regolarità C^k . Premettiamo la seguente notazione, che sarà utile anche nel seguito. Chiamiamo *multiindice* un vettore di numeri naturali, 0 incluso, cioè $\alpha \in \mathbb{N}_0^n$, e per $\alpha = (\alpha_1, \dots, \alpha_n)$ definiamo la *lunghezza* $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$ e poniamo per $x \in \mathbb{R}^n$ $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, e per $f \in C^k$, $D^\alpha f = D_{x_1}^{\alpha_1} D_{x_2}^{\alpha_2} \cdots D_{x_n}^{\alpha_n} f$.

Teorema 1.55 *Supponiamo che $t \mapsto f(t, x)$ sia di classe C^k in A per ogni $x \in E$ ed esista una funzione $g \in L^1(E)$ tale che*

$$(1.21) \quad \sum_{|\alpha| \leq k} |D^\alpha f(t, x)| \leq g(x) \quad \forall t \in A, x \in E.$$

Allora F è di classe C^k in A e

$$D^\alpha F(t) = \int_E D^\alpha f(t, x) dx \quad \forall t \in A, |\alpha| \leq k.$$

Dim. La dimostrazione si può fare per induzione su k . Limitiamoci al caso $k = 1$: sia $i \in \{1, \dots, m\}$ e verifichiamo che

$$(1.22) \quad D_i F(t) = \int_E D_i f(t, x) dx \quad \forall t \in A.$$

Si osservi che se vale la (1.22) allora $D_i F$ è continua in A per il teorema precedente. Fissato $t \in A$ ed una successione $(r_h) \subset \mathbb{R} \setminus \{0\}$ tendente a zero osserviamo che

$$\frac{F(t + r_h e_i) - F(t)}{r_h} = \int_E \frac{f(t + r_h e_i, x) - f(t, x)}{r_h} dx.$$

Fissato un $x \in E$, per il teorema di Lagrange esistono $s_h(x)$ compresi tra 0 e r_h tali che

$$\frac{f(t + r_h e_i, x) - f(t, x)}{r_h} = D_i f(t + s_h(x), x).$$

Passando al limite in h si ha

$$\lim_{h \rightarrow +\infty} \frac{f(t + r_h e_i, x) - f(t, x)}{r_h} = D_i f(t, x).$$

Inoltre, usando la (1.21) otteniamo

$$\left| \frac{f(t + r_h e_i, x) - f(t, x)}{r_h} \right| = |D_i f(t + s_h(x), x)| \leq g(x),$$

quindi il teorema della convergenza dominata implica

$$\begin{aligned} \lim_{h \rightarrow +\infty} \frac{F(t + r_h e_i) - F(t)}{r_h} &= \lim_{h \rightarrow +\infty} \int_E \frac{f(t + r_h e_i, x) - f(t, x)}{r_h} dx \\ &= \int_E D_i f(t, x) dx. \end{aligned}$$

Essendo la successione r_h arbitraria, la funzione F ha derivata parziale i -esima in t e vale la (1.22). $\square$

Il teorema sopra continua a valere se si suppone solamente che $t \mapsto f(t, x)$ è di classe C^k in A per m_n -quasi ogni $x \in E$.

In tal caso la funzione $D^\alpha f(t, x)$, il cui integrale su E dà $D^\alpha F(t)$, è definita in modo arbitrario nell'insieme trascurabile degli $x \in E$ tali che $t \mapsto f(t, x)$ non è C^k .

Nel caso $m = 1$ possiamo anche supporre che l'insieme di integrazione dipenda da t :

$$F(t) = \int_{\alpha(t)}^{\beta(t)} f(t, x) dx.$$

Allora $F(t) = G(t, \alpha(t), \beta(t))$ con

$$G(t, u, v) = \int_u^v f(t, x) dx.$$

Se f è continua nelle variabili (t, x) e C^1 nella variabile t abbiamo (si usa il teorema fondamentale del calcolo integrale)

$$\begin{aligned} G_t(t, u, v) &= \int_u^v f_t(t, x) dx \\ G_u(t, u, v) &= -f(t, u) \\ G_v(t, u, v) &= f(t, v) \end{aligned}$$

quindi dal teorema di derivazione della funzione composta otteniamo

$$\begin{aligned} F'(t) &= G_t(t, \alpha(t), \beta(t)) + G_u(t, \alpha(t), \beta(t))\alpha'(t) + G_v(t, \alpha(t), \beta(t))\beta'(t) \\ &= \int_{\alpha(t)}^{\beta(t)} f_t(t, x) dx + \beta'(t)f(t, \beta(t)) - \alpha'(t)f(t, \alpha(t)). \end{aligned}$$

1.11 Prodotto di convoluzione

Siano f e g funzioni misurabili in $\mathbb{R}^n$. Il *prodotto di convoluzione* $f * g$ è definito dalla formula

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

Si osservi che, anche se f e g sono integrabili, non è detto che $f * g$ sia definito per tutti gli x (ad esempio se $n = 1$ e $f(t) = g(t) = |t|^{-1/2}$ il prodotto $f * g$ è definito solo per $t \neq 0$). Col cambiamento di variabile $x - y = y'$ si vede facilmente che il prodotto di convoluzione è commutativo.

Definizione 1.56 (Supporto) Per $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, definiamo il supporto di f , denotato $\text{supp } f$, come la chiusura dell'insieme $\{x \in \Omega : f(x) \neq 0\}$.

Introduciamo ora un altro spazio funzionale di grande utilità; per $\Omega \subset \mathbb{R}^n$ aperto, poniamo

$$(1.23) \quad C_c^\infty(\Omega) = \left\{ f \in C^\infty(\Omega) : \text{supp } f \subset\subset \Omega \right\},$$

dove $C \subset\subset \Omega$ significa che C è un insieme chiuso e limitato contenuto in Ω ; in particolare, questo vuol dire che c'è una distanza non nulla tra C e il complementare di Ω , ossia, in termini delle funzioni di $C_c^\infty(\Omega)$, che esse si annullano in un intorno della frontiera di Ω . Dimostriamo ora il

Teorema 1.57 Se f appartiene a $L^p(\mathbb{R}^n)$ con $p = 1$ o $p = 2$ e $g \in L^1(\mathbb{R}^n)$ allora il prodotto di convoluzione è definito per m_n -quasi ogni $x \in \mathbb{R}^n$, appartiene a $L^p(\mathbb{R}^n)$ e

$$(1.24) \quad \|f * g\|_2 \leq \|f\|_p \|g\|_1.$$

Inoltre, se $f \in C_c^\infty(\mathbb{R}^n)$ allora $f * g \in C^\infty(\mathbb{R}^n)$ e

$$D^\alpha(f * g) = (D^\alpha f) * g \quad \text{per ogni multiindice } \alpha.$$

Dim. Per quanto riguarda l'appartenenza di $f * g$ a L^p e la (1.24) nel caso $p = 1$, basta osservare che

$$\begin{aligned} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)g(y)| dy dx &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)g(y)| dx dy \\ &= \|f\|_1 \int_{\mathbb{R}^n} |g(y)| dy = \|f\|_1 \|g\|_1. \end{aligned}$$

Nel caso $p = 2$ il discorso è simile: si può porre g nella forma $g = g^{1/2}g^{1/2}$, ottenendo:

$$\begin{aligned} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)| dy \right)^2 dx &\leq \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-y)|^2 |g(y)| dy \left(\int_{\mathbb{R}^n} |g(y)| dy \right) \right] dx \\ &= \|g\|_1 \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)|^2 |g(y)| dx dy \\ &= \|g\|_1^2 \|f\|_2^2 \end{aligned}$$

e la (1.24) segue prendendo le radici quadrate di ambo i membri ed osservando che

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

è finito per m_n -quasi ogni x ed è maggiore di $f * g(x)$.

Infine, la derivabilità di $f * g$ e la commutazione della derivata col prodotto di convoluzione seguono direttamente differenziando sotto il segno di integrale e applicando il Teorema 1.55. $\square$

Osservazione 1.58 Ragionando come nel teorema precedente si verifica che il prodotto di convoluzione è anche *associativo*:

$$f * (g * h) = (f * g) * h \quad f, g, h \in L^1(\mathbb{R}^n).$$

Infatti, risulta

$$\begin{aligned} f * (g * h)(x) &= \int_{\mathbb{R}^n} f(x-y) \left(\int_{\mathbb{R}^n} g(y-z)h(z)dz \right) dy \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y)g(y-z)dy \right) h(z)dz \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(t)g(x-z-t)dt \right) h(z)dz \\ &= \int_{\mathbb{R}^n} (f * g)(x-z)h(z)dz = (f * g) * h(x). \end{aligned}$$

Come prima applicazione del concetto di convoluzione vediamo un metodo che consente di approssimare funzioni L^p ($p = 1, 2$) con funzioni regolari. Diremo che una funzione $\rho \in C^\infty(\mathbb{R}^n)$ è un *nucleo di convoluzione* se ρ è non negativa, pari, il suo integrale vale 1 ed il suo supporto è contenuto in $\overline{B}_1(0)$, vale a dire, $\rho(x) = 0$ per $|x| > 1$. Dato un nucleo di convoluzione ρ , le funzioni

$$\rho_\varepsilon(x) := \varepsilon^{-n} \rho\left(\frac{x}{\varepsilon}\right) \quad \varepsilon > 0$$

sono una *famiglia di funzioni regolarizzanti* (o anche *mollificatori*). Osserviamo che le funzioni ρ_ε sono non negative, pari, hanno supporto contenuto in $\overline{B}_\varepsilon(0)$ e integrale uguale a 1. Un esempio di nucleo di convoluzione è la funzione radiale

$$(1.25) \quad \rho(x) := \begin{cases} ce^{\frac{1}{|x|^2-1}} & \text{se } |x| < 1; \\ 0 & \text{se } |x| \geq 1. \end{cases}, \quad c \text{ tale che } \int_{\mathbb{R}^n} \rho(x) dx = 1$$

Dato un nucleo di convoluzione ρ le funzioni $f * \rho_\varepsilon$ sono dette funzioni *regolarizzate* di f . Il nome viene dal fatto che $f * \rho_\varepsilon \in C^\infty(\mathbb{R}^n)$ per il teorema precedente. Vale anche il

Teorema 1.59 *Sia $p = 1$ o 2 . Per ogni funzione $f \in L^p(\mathbb{R}^n)$ le funzioni regolarizzate $f_\varepsilon := f * \rho_\varepsilon$ convergono a f in $L^p(\mathbb{R}^n)$ e m_n -quasi ovunque in E per $\varepsilon \rightarrow 0$. In particolare, per ogni insieme aperto E di $\mathbb{R}^n$, lo spazio $C^\infty(E)$ è denso in $L^p(E)$, vale a dire, ogni funzione di $L^p(E)$ può essere approssimata in $L^p(E)$ da funzioni in $C^\infty(E)$ con precisione arbitraria.*

Notiamo che con un cambiamento di variabili si ottiene

$$f * \rho_\varepsilon(x) = \int_{\mathbb{R}^n} f(x-y)\rho_\varepsilon(y) dy = \int_{\mathbb{R}^n} f(x-\varepsilon y)\rho(y) dy = \int_{B_1(0)} f(x-\varepsilon y)\rho(y) dy.$$

Possiamo quindi rappresentare il valore $f * \rho_\varepsilon(x)$ come una media (rispetto alla misura di probabilità indotta da ρ) dei valori di f nella palla $B_\varepsilon(x)$. Usando questa caratterizzazione è facile vedere che $f * \rho_\varepsilon(x)$ converge a $f(x)$ in ogni punto di continuità di f .

Capitolo 2

Teoria delle distribuzioni

La teoria delle distribuzioni inizia con S.L. Sobolev nel 1936, ma riceve piena formulazione ad opera di L.Schwartz (1954); oggi costituisce un capitolo fondamentale dell'analisi funzionale. Il concetto di *distribuzione* generalizza quello di funzione in molte situazioni in cui viene meno la regolarità della funzione in questione, soprattutto, come vedremo, nella derivazione. Esempi di distribuzioni sono le *funzioni impulsive*, molto usate nell'elettromagnetismo, nella meccanica quantistica, ancor prima che ne venisse data un'opportuna interpretazione matematica.

2.1 Definizioni ed esempi

Sia Ω un aperto di $\mathbb{R}^n$. Ricordiamo che $C_c^\infty(\Omega)$ è l'insieme delle funzioni f , reali o complesse, di classe C^∞ e a supporto¹ compatto contenuto in Ω . Nell'ambito della teoria delle distribuzioni, si conviene di porre

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega).$$

Chiameremo gli elementi di $\mathcal{D}(\Omega)$ *funzioni test*. Con le usuali operazioni di somma e prodotto per uno scalare, $\mathcal{D}(\Omega)$ è uno spazio vettoriale. Osserviamo che esso è chiuso rispetto all'operazione di derivazione. Se $f \in \mathcal{D}(\Omega)$, poniamo come in (1.2)

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

Un esempio classico di funzione test è dato dalla funzione in (1.25) che ricordiamo

$$(2.1) \quad \varrho(x) = \begin{cases} e^{\frac{1}{|x|^2-1}} & \text{se } |x| < 1, \\ 0 & \text{se } |x| \geq 1. \end{cases}$$

¹ $\text{supp } f = \overline{\{x \in \Omega \mid f(x) \neq 0\}}$.

Assegniamo la definizione di limite in $\mathcal{D}(\Omega)$, tenendo conto delle proprietà peculiari dei suoi elementi.

Definizione 2.1 Siano $\varphi_k, \varphi \in \mathcal{D}(\Omega)$. Diremo che φ_k converge a φ in $\mathcal{D}(\Omega)$ se valgono le seguenti due condizioni:

- a. esiste un compatto K tale che $\text{supp } \varphi_k \subseteq K$, per ogni $k \in \mathbb{N}$;
- b. per ogni multiindice $\alpha \in \mathbb{N}_0^n$, risulta $\lim_{k \rightarrow \infty} \|D^\alpha \varphi_k - D^\alpha \varphi\|_\infty = 0$.

Osserviamo che sulla base della definizione appena data, se $\varphi_k \rightarrow \varphi$ in $\mathcal{D}(\Omega)$, allora $D^\alpha \varphi_k \rightarrow D^\alpha \varphi$ in $\mathcal{D}(\Omega)$, per ogni $\alpha \in \mathbb{N}_0^n$.

Esempio 2.2 Prendiamo $N = 1$, $\Omega = \mathbb{R}$. Sia $\varrho \in \mathcal{D}(\mathbb{R})$ la funzione data in (2.1) e definiamo $\varphi_k(x) = e^{-k} \varrho(x) \sin(kx)$. Proviamo che $\varphi_k \rightarrow 0$ in $\mathcal{D}(\mathbb{R})$. Si vede facilmente che $\text{supp } \varphi_k \subseteq \text{supp } \varrho$. Inoltre, per ogni $k \in \mathbb{N}$, risulta

$$D^n \varphi_k(x) = e^{-k} \sum_{j=0}^n \binom{n}{j} D^j \varrho(x) D^{n-j} \sin(kx),$$

per cui

$$\|D^n \varphi_k\|_\infty \leq e^{-k} \sum_{j=0}^n \binom{n}{j} k^{n-j} \|D^j \varrho\|_\infty \rightarrow 0.$$

Osserviamo che se al posto di e^{-k} avessimo scelto k^{-q} , con q intero positivo, allora $\varphi_k \not\rightarrow 0$ in $\mathcal{D}(\mathbb{R})$, in quanto $\|D^q \varphi_k\|_\infty \not\rightarrow 0$.

Definizione 2.3 Sia T un funzionale lineare su $\mathcal{D}(\Omega)$. Diremo che T è una distribuzione se T è continuo rispetto alla convergenza definita, ossia se per ogni successione $\varphi_k \rightarrow 0$ in $\mathcal{D}(\Omega)$, risulta $T(\varphi_k) \rightarrow 0$.

Siccome T è lineare, è sufficiente considerare il caso in cui la funzione limite è 0. Notiamo che il termine *funzionale* viene usato per il fatto che T è definito in uno spazio di *funzioni*. L'analisi funzionale suggerisce anche la seguente notazione:

$$\begin{aligned} \mathcal{D}'(\Omega) &= \text{lo spazio delle distribuzioni in } \Omega \text{ (ossia il duale topologico di } \mathcal{D}(\Omega)\text{);} \\ \langle T, \varphi \rangle &= \text{il valore di } T \text{ sull'elemento } \varphi \in \mathcal{D}(\Omega). \end{aligned}$$

$\mathcal{D}'(\Omega)$ è uno spazio vettoriale con le operazioni naturali di somma e prodotto per uno scalare. Per verificare se un dato funzionale lineare in $\mathcal{D}(\Omega)$ definisce una distribuzione, può essere utile conoscere la seguente caratterizzazione.

Proposizione 2.4 *Sia T un funzionale lineare in $\mathcal{D}(\Omega)$. Sono equivalenti:*

- (i) $T \in \mathcal{D}'(\Omega)$;
- (ii) *per ogni compatto $K \subset \Omega$, esistono $C > 0$ e $j \in \mathbb{N}_0$ (che dipendono da K), tali che*

$$(2.2) \quad |\langle T, \varphi \rangle| \leq C \sum_{|\alpha| \leq j} \sup_{x \in K} |D^\alpha \varphi(x)|,$$

per ogni $\varphi \in \mathcal{D}(K)$.

La seconda condizione esprime il fatto che una distribuzione, localmente, agisce solo su un numero finito di derivate.

Definizione 2.5 *Se nella (2.2) la scelta di j è indipendente dal compatto K , allora la distribuzione T si dice di ordine finito, e si definisce ordine di T il minimo intero per cui vale (2.2).*

Esempi 2.6

1. Sia $u \in L^1_{\text{loc}}(\Omega)$ una funzione localmente sommabile secondo Lebesgue. Definiamo la distribuzione T_u ponendo

$$(2.3) \quad \langle T_u, \varphi \rangle = \int_{\Omega} u(x) \varphi(x) dx, \quad \varphi \in \mathcal{D}(\Omega).$$

Osserviamo che T_u è ben posta, perché φ ha supporto compatto ed è limitata e u è sommabile su ogni compatto. Inoltre T_u è lineare, dato che l'integrale è lineare, ed è continuo, come si verifica facilmente usando direttamente la definizione o tramite la Proposizione 2.4. Pertanto $T_u \in \mathcal{D}'(\Omega)$. In particolare, T_u è una distribuzione di ordine zero.

Notiamo che la corrispondenza $u \rightarrow T_u$ è iniettiva, cioè se u_1, u_2 sono due funzioni in $L^1_{\text{loc}}(\Omega)$ che individuano, tramite la (2.3), la stessa distribuzione, allora $u_1 = u_2$ q.o. È possibile così identificare u con T_u e, di conseguenza, $L^1_{\text{loc}}(\Omega)$ con un sottospazio di $\mathcal{D}'(\Omega)$. Dunque, le distribuzioni sono *funzioni generalizzate*.

2. Sia $x_0 \in \Omega$. Definiamo δ_{x_0} come segue

$$(2.4) \quad \langle \delta_{x_0}, \varphi \rangle = \varphi(x_0), \quad \varphi \in \mathcal{D}(\Omega).$$

Anche stavolta, è facile vedere che δ_{x_0} è una distribuzione di ordine zero, chiamata *delta di Dirac di polo* x_0 . Se $x_0 = 0$, poniamo semplicemente δ al posto di δ_0 . È significativo vedere che

$$\delta_{x_0} \notin L^1_{\text{loc}}(\Omega),$$

per cui l'inclusione $L^1_{\text{loc}}(\Omega) \subset \mathcal{D}(\Omega)$ è stretta. Infatti, se esistesse $u \in L^1_{\text{loc}}(\Omega)$ tale che $\langle \delta_{x_0}, \varphi \rangle = \int_{\Omega} u(x)\varphi(x)dx$, per ogni $\varphi \in \mathcal{D}(\Omega)$, allora si avrebbe $\int_{\Omega} u(x)\varphi(x)dx = 0$, se $x_0 \notin \text{supp } \varphi$. Ciò implica che $u = 0$ q.o. in $\Omega \setminus \{x_0\}$ e quindi, dato che $\{x_0\}$ ha misura nulla, $u = 0$ q.o. in Ω . Ciò è assurdo, perché $\delta_{x_0} \neq 0$.

2. La (2.4) si può interpretare come *l'integrale di φ rispetto alla misura di Dirac δ_{x_0}* definita nell'Esempio 1.7(2). Più in generale, ad ogni misura μ finita sui compatti si può associare la distribuzione T_{μ} definita da

$$(2.5) \quad \langle T_{\mu}, \varphi \rangle = \int_{\Omega} \varphi d\mu, \quad \varphi \in \mathcal{D}(\Omega).$$

Questa formula contiene entrambi gli esempi precedenti come casi particolari.

3. Prendiamo $\Omega = I = (0, 1)$ e definiamo $\langle T, \varphi \rangle = \sum_{j=2}^{\infty} D^j \varphi \left(\frac{1}{j} \right)$. La definizione

è ben posta, perché per ogni $\varphi \in \mathcal{D}(I)$, la somma che definisce $\langle T, \varphi \rangle$ è finita. T è lineare. Per provare che T è continuo usiamo la Proposizione 2.4. Sia K un compatto contenuto in I . Sia $\nu \in \mathbb{N}$ il minimo intero tale che $1/j \notin K$ per ogni $j \geq \nu$. Se $\varphi \in \mathcal{D}(K)$, allora $|\langle T, \varphi \rangle| \leq \sum_{j=2}^{\nu-1} \|D^j \varphi\|_{\infty}$, da cui la tesi.

Notare che ν dipende da K , per cui, in questo caso, la distribuzione ha ordine infinito.

Definiamo il *prodotto* Tf tra una distribuzione $T \in \mathcal{D}'(\Omega)$ ed una funzione f di classe C^{∞} in Ω mediante la formula

$$(2.6) \quad \langle Tf, \varphi \rangle = \langle T, f\varphi \rangle, \quad \varphi \in \mathcal{D}(\Omega).$$

Siccome $f \in C^{\infty}(\Omega)$, il prodotto $f\varphi \in \mathcal{D}(\Omega)$, per cui la definizione appena data è ben posta. In particolare, se $T = \delta$ allora $\delta f = f(0)\delta$, come è immediato verificare.

2.2 Successioni e serie di distribuzioni

Definizione 2.7 Siano $T_n, T \in \mathcal{D}'(\Omega)$. Diremo che T_n converge a T nel senso delle distribuzioni se $\lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle = \langle T, \varphi \rangle$, per ogni $\varphi \in \mathcal{D}(\Omega)$.

In particolare, una successione di funzioni $(f_n) \subset L^1_{\text{loc}}(\Omega)$ converge ad una distribuzione T , se $\lim_{n \rightarrow \infty} \int_{\Omega} f_n(x) \varphi(x) dx = \langle T, \varphi \rangle$, per ogni $\varphi \in \mathcal{D}(\Omega)$. Il fatto di considerare il limite puntuale della successione non aiuta, giacché in generale il limite nel senso delle distribuzioni non è più una funzione in generale. Tuttavia, se f_n converge ad una funzione f puntualmente ed esiste una funzione $g \in L^1_{\text{loc}}(\Omega)$ tale che $|f_n(x)| \leq g(x)$, per quasi ogni x , allora f_n converge a f anche in $\mathcal{D}'(\Omega)$.

Esempio 2.8 Sia ϱ la funzione test definita in (2.1), normalizzata, e sia $\varrho_n(x) = n\varrho(nx)$. Vogliamo provare che $\lim_{n \rightarrow \infty} \varrho_n = \delta$ in $\mathcal{D}'(\mathbb{R})$. Tenendo conto del fatto che $\int_{\mathbb{R}} \varrho_n = 1$, abbiamo

$$\int_{\mathbb{R}} \varrho_n(x) \varphi(x) dx - \varphi(0) = \int_{\mathbb{R}} \varrho_n(x) (\varphi(x) - \varphi(0)) dx = \int_{-1/n}^{1/n} n\varrho(nx) (\varphi(x) - \varphi(0)) dx.$$

Poniamo $\sigma_n = \max_{x \in [-1/n, 1/n]} |\varphi(x) - \varphi(0)|$. Così, risulta

$$\left| \int_{\mathbb{R}} \varrho_n(x) \varphi(x) dx - \varphi(0) \right| \leq \sigma_n \int_{-1/n}^{1/n} n\varrho(nx) dx = \sigma_n \longrightarrow 0,$$

grazie alla continuità di φ . Questo esempio chiarisce la natura di δ come funzione impulsiva e conferma il fatto che essa non è propriamente una funzione. Infatti, avremmo $\delta(x) = 0$, per ogni $x \neq 0$, $\delta(0) = +\infty$ ma $\int_{\mathbb{R}} \delta = 1$.

Analogamente, si può provare che $\lim_{n \rightarrow \infty} f_n = \delta$ in $\mathcal{D}'(\mathbb{R})$, dove

$$f_n(x) = \begin{cases} n & \text{se } |x| \leq \frac{1}{2n} \\ 0 & \text{se } |x| > \frac{1}{2n} \end{cases}$$

Alla definizione di limite di una successione in $\mathcal{D}'(\Omega)$ si associa, in modo naturale, quella di serie di distribuzioni.

Definizione 2.9 Se $T_n, T \in \mathcal{D}'(\Omega)$, diremo che $\sum_{n=1}^{\infty} T_n = T$ in $\mathcal{D}'(\Omega)$, se accade

che $\lim_{n \rightarrow \infty} \sum_{k=1}^n T_k = T$ in $\mathcal{D}'(\Omega)$.

Osservazione 2.10 La definizione del prodotto tra una distribuzione e una funzione C^∞ data in precedenza non è casuale. Essa è un'operazione continua nel senso che se $T_n \rightarrow T$ in $\mathcal{D}'(\Omega)$, allora $T_n f \rightarrow T f$ in $\mathcal{D}'(\Omega)$.

2.3 Derivazione

Uno dei principali vantaggi offerti dalla teoria delle distribuzioni concerne la derivazione. Vedremo che le distribuzioni sono derivabili infinite volte e che basta sapere che una successione di distribuzioni converge, per avere la convergenza di tutte le derivate. Ciò spiega il motivo per cui le distribuzioni sono largamente impiegate nella risoluzione di equazioni differenziali.

Prima di dare la definizione di derivata di una distribuzione, vediamo un esempio che chiarisce ancora una volta il ruolo delle distribuzioni come funzioni generalizzate. Sia $f \in C^1(\mathbb{R})$. Allora per ogni $\varphi \in \mathcal{D}(\mathbb{R})$ risulta

$$\int_{\mathbb{R}} f(x)\varphi'(x)dx = - \int_{\mathbb{R}} f'(x)\varphi(x)dx,$$

grazie alla formula di integrazione per parti.

Definizione 2.11 Siano $T \in \mathcal{D}'(\Omega)$ e $\beta \in \mathbb{N}_0^n$. Definiamo

$$(2.7) \quad \langle D^\beta T, \varphi \rangle = (-1)^{|\beta|} \langle T, D^\beta \varphi \rangle,$$

derivata di T .

Notiamo che, in effetti, l'identità (2.7) generalizza la formula di integrazione per parti per le funzioni regolari. Pertanto, se $f \in C^1(\Omega)$, derivate classiche e distribuzionali coincidono. In realtà, ciò continua ad essere vero anche sotto condizioni di minore regolarità. Per esempio, se $f \in C^1(\mathbb{R} \setminus \{0\})$ e la derivata prima è localmente sommabile, allora la formula (2.7) è ancora verificata dalla derivata classica.

Dalla definizione discende subito che $D^\beta T$ è ancora una distribuzione. Perciò, è possibile iterare la derivazione e di fatto ogni distribuzione può essere derivata infinite volte. In particolare, una funzione localmente sommabile è derivabile infinite volte *nel senso delle distribuzioni*. Non è detto, però, che le derivate siano ancora funzioni localmente sommabili.

Osservazione 2.12 Osserviamo che l'operazione di derivazione è continua rispetto alla convergenza in $\mathcal{D}'(\Omega)$, cioè se $\lim_{n \rightarrow \infty} T_n = T$, allora per ogni multiindice β

anche $\lim_{n \rightarrow \infty} D^\beta T_n = D^\beta T$ in $\mathcal{D}'(\Omega)$. Un risultato analogo può essere enunciato per la derivata di una serie di distribuzioni.

Queste proprietà non hanno riscontro nell'ambito delle funzioni e costituiscono un valido argomento a favore della teoria delle distribuzioni.

Esempio 2.13 Consideriamo la funzione di Heaviside

$$H(x) = \begin{cases} 1 & \text{se } x > 0, \\ 0 & \text{se } x < 0. \end{cases}$$

Siccome $H \in L^1_{\text{loc}}(\mathbb{R})$, $H \in \mathcal{D}(\mathbb{R})$. Calcoliamo H' . Per ogni $\varphi \in \mathcal{D}(\mathbb{R})$, risulta

$$\langle H, \varphi' \rangle = \int_0^\infty \varphi'(x) dx = -\varphi(0) = -\langle \delta, \varphi \rangle,$$

pertanto $H' = \delta \notin L^1_{\text{loc}}(\mathbb{R})$. Questo risultato si può interpretare dicendo che H è una soluzione (discontinua) dell'equazione differenziale $T' = \delta$.

Esempio 2.14 Sia $u(x) = 1/|x|$, $x \in \mathbb{R}^3 \setminus \{0\}$. Calcoliamo Δu nel senso delle distribuzioni. Sia $\varphi \in \mathcal{D}(\mathbb{R}^3)$ e sia $R > 0$ grande abbastanza affinché $\text{supp } \varphi \subset B_R(0)$. Osserviamo che

$$\langle \Delta u, \varphi \rangle = \langle u, \Delta \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \int_{B_R(0) \setminus B_\varepsilon(0)} u(x) \Delta \varphi(x) dx.$$

Dato che u è di classe C^∞ in $\mathbb{R}^3 \setminus \{0\}$, integrando per parti, risulta

$$\begin{aligned} \int_{B_R(0) \setminus B_\varepsilon(0)} u(x) \Delta \varphi(x) dx &= \int_{B_R(0) \setminus B_\varepsilon(0)} (\Delta u)(x) \varphi(x) dx \\ &\quad + \int_{\partial B_\varepsilon(0)} u(x) \nabla \varphi(x) \cdot \nu(x) d\sigma \\ &\quad - \int_{\partial B_\varepsilon(0)} \nabla u(x) \cdot \nu(x) \varphi(x) d\sigma \end{aligned}$$

dove $\nu(x)$ è la normale esterna unitaria al dominio di integrazione, quindi $\nu(x) = -x/\varepsilon$. Con un semplice calcolo si vede che $\nabla u(x) = -x/|x|^3$ e $\Delta u(x) = 0$, se $x \neq 0$. Quindi, il primo integrale è zero. Poi

$$\left| \int_{\partial B_\varepsilon(0)} u(x) \nabla \varphi(x) \cdot \nu(x) d\sigma \right| \leq \|\nabla \varphi\|_\infty \frac{|\partial B_\varepsilon(0)|}{\varepsilon} = \|\nabla \varphi\|_\infty 4\pi\varepsilon \longrightarrow 0$$

per $\varepsilon \rightarrow 0$, e

$$\begin{aligned} \int_{\partial B_\varepsilon(0)} \nabla u(x) \cdot \nu(x) \varphi(x) d\sigma &= \frac{1}{\varepsilon^2} \int_{\partial B_\varepsilon(0)} \varphi(x) d\sigma \\ &= 4\pi \oint_{\partial B_\varepsilon(0)} \varphi(x) d\sigma \longrightarrow 4\pi \varphi(0), \text{ se } \varepsilon \rightarrow 0. \end{aligned}$$

In definitiva, abbiamo provato che $\Delta \frac{1}{4\pi|x|} = -\delta$.

Calcoliamo una derivata prima del prodotto Tf . Se $\varphi \in \mathcal{D}(\Omega)$, risulta

$$\begin{aligned} \langle D_i(Tf), \varphi \rangle &= -\langle Tf, D_i\varphi \rangle = -\langle T, f D_i\varphi \rangle = -\langle T, D_i(f\varphi) \rangle + \langle T, (D_i f)\varphi \rangle \\ &= \langle (D_i T)f + T D_i f, \varphi \rangle, \end{aligned}$$

cioè

$$D_i(Tf) = (D_i T)f + T D_i f.$$

Ritroviamo così l'usuale regola di derivazione di un prodotto, che può essere estesa alle derivate di ordine superiore dando luogo alla formula di Leibniz in $\mathcal{D}'(\Omega)$.

Esempio 2.15 È noto che la funzione $f(x) = 1/x$, $x \neq 0$, non è sommabile in alcun intorno dell'origine, e non è neanche integrabile in senso generalizzato secondo la teoria dell'integrale di Riemann; infatti, il limite che dovrebbe definire l'integrale generalizzato di f

$$\lim_{\varepsilon \rightarrow 0, \delta \rightarrow 0} \int_{-1}^{-\varepsilon} \frac{1}{x} dx + \int_{\delta}^1 \frac{1}{x} dx$$

non esiste. Pertanto, non si può associare ad f una distribuzione su $\mathbb{R}$ come si è fatto nell'Esempio 2.6 con le funzioni L_{loc}^1 . Si può però associare ad f una distribuzione osservando che seguente limite esiste per ogni $c \in \mathbb{R}$ e per ogni $M > 0$:

$$(2.8) \quad \lim_{\delta \rightarrow 0} \int_{-M}^{-\delta} \frac{c}{x} dx + \int_{\delta}^M \frac{c}{x} dx = 0.$$

Presa allora una qualunque funzione test φ , definiamo la distribuzione $\text{PV } \frac{1}{x}$ ponendo

$$\langle \text{PV } \frac{1}{x}, \varphi \rangle = \lim_{\delta \rightarrow 0} \int_{-\infty}^{-\delta} \frac{\varphi(x)}{x} dx + \int_{\delta}^{\infty} \frac{\varphi(x)}{x} dx.$$

Il limite esiste per ogni $\varphi \in \mathcal{D}(\mathbb{R})$. Infatti si ha

$$\lim_{x \rightarrow 0} \frac{\varphi(x) - \varphi(0)}{x} = \varphi'(0)$$

e quindi, usando la (2.8) con $c = \varphi(0)$ ed M tale che $\text{supp } \varphi \in [-M, M]$, il limite

$$\lim_{\delta \rightarrow 0} \int_{\delta < |x| < M} \frac{\varphi(x)}{x} dx = \lim_{\delta \rightarrow 0} \int_{\delta < |x| < M} \frac{\varphi(x) - \varphi(0)}{x} dx$$

esiste perché in un intorno di 0 la funzione integranda è continua. A differenza dalle funzioni L^1_{loc} però, la distribuzione $\text{PV } \frac{1}{x}$ è di ordine 1 e non di ordine 0.

Se ora consideriamo la funzione $g(x) = \log |x|$, $x \neq 0$, questa è $L^1_{\text{loc}}(\mathbb{R})$ e quindi definisce una distribuzione T_g . Calcoliamone la derivata:

$$\begin{aligned} \langle T'_g, \varphi \rangle &= -\langle T_g, \varphi' \rangle = -\int_{-\infty}^{\infty} \log |x| \varphi'(x) dx \\ &= -\lim_{\delta \rightarrow 0} \int_{\infty}^{-\delta} \log |x| \varphi'(x) dx + \int_{\delta}^{\infty} \log |x| \varphi'(x) dx \\ &= \lim_{\delta \rightarrow 0} \int_{\infty}^{-\delta} \frac{\varphi(x)}{x} dx + \int_{\delta}^{\infty} \frac{\varphi(x)}{x} dx + \log \delta (\varphi(\delta) - \varphi(-\delta)) \\ &= \langle \text{PV } \frac{1}{x}, \varphi \rangle, \end{aligned}$$

dal momento che

$$\lim_{\delta \rightarrow 0} \log \delta (\varphi(\delta) - \varphi(-\delta)) = \lim_{\delta \rightarrow 0} \delta \log \delta \frac{\varphi(\delta) - \varphi(-\delta)}{\delta} = 0.$$

Perciò $T'_g = \text{PV } \frac{1}{x}$.

2.4 Supporto e convoluzione

Per una funzione misurabile u in Ω , si definisce supporto il complementare dell'insieme dei punti x aventi un intorno in cui u è identicamente nulla. In particolare, questa nozione può essere data per le funzioni di $L^1_{\text{loc}}(\Omega)$. La definizione di supporto di una distribuzione estende quella nota per funzioni localmente sommabili.

Definizione 2.16 Sia $T \in \mathcal{D}'(\Omega)$. Diciamo che un punto $x \in \Omega$ non appartiene al supporto $\text{supp } T$ di T quando esiste un aperto $\Omega' \subset \Omega$ contenente x e tale che $T|_{\Omega'} = 0$.

Esempio 2.17 Il supporto di δ è $\{0\}$.

L'insieme $\text{supp } T$ è chiuso in Ω . Inoltre, se $\Omega \setminus \text{supp } T$ è non vuoto e $\varphi \in \mathcal{D}(\Omega)$ ha supporto contenuto in $\Omega \setminus \text{supp } T$ allora $\langle T, \varphi \rangle = 0$. In realtà, $\text{supp } T$ è il più piccolo chiuso tale che $\langle T, \varphi \rangle = 0$, per ogni $\varphi \in \mathcal{D}(\Omega)$, tale che $\text{supp } \varphi \cap \text{supp } T = \emptyset$.

Proposizione 2.18 *Ogni distribuzione a supporto compatto è di ordine finito.*

DIM. Sia $T \in \mathcal{D}'(\Omega)$ e supponiamo che $K = \text{supp } T$ sia compatto in Ω . Consideriamo due aperti limitati Ω_1 e Ω_2 tali che $K \subset \Omega_1 \subset \Omega_2$ e la chiusura $\overline{\Omega_2}$ di Ω_2 sia ancora contenuta in Ω . Sia $\psi \in C^\infty(\Omega)$ tale che $\psi = 1$ in Ω_1 e $\psi = 0$ in $\Omega \setminus \Omega_2$. Se $\varphi \in \mathcal{D}(\Omega)$, la funzione $\varphi(1 - \psi)$ ha supporto in $\Omega \setminus \Omega_1$, dunque disgiunto da K . Pertanto $\langle T, \varphi(1 - \psi) \rangle = 0$, cioè

$$(2.9) \quad \langle T, \varphi \rangle = \langle T, \varphi\psi \rangle.$$

Applicando ora la Proposizione 2.4(ii) relativamente al compatto $\overline{\Omega_2}$, otteniamo che esistono una costante C e un intero $j \in \mathbb{N}_0$ tali che

$$|\langle T, \Phi \rangle| \leq C \sum_{|\alpha| \leq j} \|D^\alpha \Phi\|_\infty, \quad \forall \Phi \in \mathcal{D}(\overline{\Omega_2}).$$

Per ogni $\varphi \in \mathcal{D}(\Omega)$, la funzione $\Phi = \varphi\psi$ soddisfa la stima di sopra da cui, tenendo conto di (2.9), deduciamo

$$|\langle T, \varphi \rangle| \leq C \sum_{|\alpha| \leq j} \|D^\alpha(\varphi\psi)\|_\infty \leq \tilde{C} \sum_{|\alpha| \leq j} \|D^\alpha \varphi\|_\infty,$$

dove, nell'ultimo passaggio, abbiamo utilizzato la regola di Leibniz per esplicitare $D^\alpha(\varphi\psi)$ e abbiamo inglobato nella costante $\tilde{C}$ tutte le norme $\|D^\beta \psi\|_\infty$, con β multiindice di lunghezza compresa tra 0 e j . Dunque l'asserto è provato. $\square$

Le distribuzioni a supporto compatto hanno un certo interesse, dovuto al fatto che esse possono essere estese ad uno spazio più ampio di $\mathcal{D}(\Omega)$. L'idea per costruire tale spazio è la seguente: le funzioni test dovranno essere ancora di classe C^∞ , perché l'altro oggetto è una distribuzione e non una funzione, ma non c'è motivo di imporre che esse abbiano supporto compatto. Infatti, la distribuzione (quando a supporto compatto) non vede ciò che avviene vicino al bordo. Il modo di rendere rigorosa quest'idea intuitiva poggia sul seguente lemma.

Lemma 2.19 *Siano $T \in \mathcal{D}'(\Omega)$ una distribuzione a supporto compatto K , $v \in C^\infty(\Omega)$ e $\varphi_1, \varphi_2 \in \mathcal{D}(\Omega)$ tali che $\varphi_1 = \varphi_2 = 1$ in un intorno aperto di K . Allora $\langle T, v\varphi_1 \rangle = \langle T, v\varphi_2 \rangle$.*

DIM. Basta osservare che i supporti di $(\varphi_1 - \varphi_2)v$ e T sono disgiunti, per cui $\langle T, (\varphi_1 - \varphi_2)v \rangle = 0$, che è la tesi. $\square$

Risulta dunque ben posta la seguente definizione.

Definizione 2.20 *Siano $T \in \mathcal{D}(\Omega)$ una distribuzione a supporto compatto K e $v \in C^\infty(\Omega)$. Allora*

$$\langle T, v \rangle := \langle T, \varphi v \rangle,$$

dove φ è una qualunque funzione di $\mathcal{D}(\Omega)$ tale che $\varphi = 1$ in un intorno aperto di K .

Non è difficile provare che con questa estensione la distribuzione T risulta continua rispetto ad una nuova definizione di limite in $C^\infty(\Omega)$. Più precisamente

se $v_n, v \in C^\infty(\Omega)$ e se la successione v_n converge a v con le derivate di qualunque ordine uniformemente sui compatti di Ω , allora

$$\langle T, v_n \rangle \rightarrow \langle T, v \rangle.$$

La Definizione 2.20 rende possibile, per una distribuzione a supporto compatto T in Ω , il prolungamento a tutto lo spazio $\mathbb{R}^n$. Basta infatti porre

$$\langle \tilde{T}, v \rangle = \langle T, v|_\Omega \rangle, \quad v \in \mathcal{D}(\mathbb{R}^n).$$

Allora $\tilde{T} \in \mathcal{D}'(\mathbb{R}^n)$, $\tilde{T}$ ha lo stesso supporto di T e la sua restrizione a $\mathcal{D}(\Omega)$ coincide con T . Naturalmente, una distribuzione arbitraria potrebbe non ammettere alcun prolungamento.

Alle distribuzioni si può anche estendere l'operazione di convoluzione. Iniziamo col definire la convoluzione tra una distribuzione e una funzione test.

Definizione 2.21 *Per $\varphi \in \mathcal{D}(\mathbb{R}^n)$ e $T \in \mathcal{D}'(\mathbb{R}^n)$ sia $T * \varphi$ la distribuzione definita da $\langle (T * \varphi), \psi \rangle = \langle T, \varphi * \psi \rangle$.*

Osservazione 2.22 Valgono le seguenti proprietà:

- (1) $\text{supp } T * \varphi \subset \text{supp } T + \text{supp } \varphi$
- (2) $(T * \varphi) * \psi = T * (\varphi * \psi)$ per ogni $\psi \in \mathcal{D}(\mathbb{R}^n)$
- (3) $T * \varphi \in C^\infty$
- (4) $D^\alpha(T * \varphi) = (D^\alpha T) * \varphi = T * (D^\alpha \varphi)$ per ogni $\alpha \in \mathbb{N}^n$

- (5) Se almeno una tra le distribuzioni T, S ha supporto compatto, si può definire la convoluzione tra S e T ponendo

$$\langle (S * T), \varphi \rangle = \langle S(x), \langle T(y), \varphi(x+y) \rangle \rangle$$

per ogni φ test, dove la notazione ha il significato che la distribuzione $T(y)$ agisce sulla variabile y producendo una funzione della x a cui si applica poi la distribuzione $S(x)$. Per esempio, $T * \delta = T$ per ogni $T \in \mathcal{D}'$; infatti,

$$\langle \delta(y), \varphi(x+y) \rangle = \varphi(x) \quad \Rightarrow \quad \langle (T * \delta), \varphi \rangle = \langle T(x), \langle \delta(y), \varphi(x+y) \rangle \rangle = \langle T, \varphi \rangle.$$

2.5 Distribuzioni temperate

Le distribuzioni temperate sono particolari distribuzioni in $\mathbb{R}^n$ che all'infinito hanno un comportamento più controllato di quanto possa accadere alla generica distribuzione. In questo modo è consentita maggiore generalità alle funzioni test. Cominciamo proprio a definire queste ultime.

Definizione 2.23 Diremo che una funzione φ appartiene alla classe di Schwartz $\mathcal{S}(\mathbb{R}^n)$ se $\varphi \in C^\infty(\mathbb{R}^n)$ e per ogni $\alpha, \beta \in \mathbb{N}_0^n$ risulta

$$\sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty.$$

Un elemento di $\mathcal{S}(\mathbb{R}^n)$ si chiama anche funzione a decrescita rapida all'infinito.

Osserviamo che $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, ma il viceversa non è vero: basti considerare la funzione $e^{-|x|^2}$.

Definizione 2.24 Date $\varphi_k, \varphi \in \mathcal{S}(\mathbb{R}^n)$, diremo che φ_k converge a φ in $\mathcal{S}(\mathbb{R}^n)$ se

$$\lim_{k \rightarrow \infty} \sup_{x \in \mathbb{R}^n} |x^\alpha (D^\beta \varphi_k - D^\beta \varphi)(x)| = 0, \quad \text{per ogni } \alpha, \beta \in \mathbb{N}_0^n.$$

Chiameremo distribuzione temperata ogni funzionale T su $\mathcal{S}(\mathbb{R}^n)$ lineare e continuo rispetto alla convergenza introdotta. In simboli, $T \in \mathcal{S}'(\mathbb{R}^n)$.

Siccome $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ e la convergenza nel primo spazio implica quella nel secondo, ogni distribuzione temperata, ristretta a $\mathcal{D}(\mathbb{R}^n)$, è una distribuzione classica.

A differenza di quanto avviene per $\mathcal{D}'(\Omega)$, lo spazio $L^1_{\text{loc}}(\mathbb{R}^n)$ non è interamente contenuto in $\mathcal{S}'(\mathbb{R}^n)$. Affinché una funzione $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ dia origine ad una distribuzione è necessario che essa soddisfi condizioni supplementari all'infinito.

Proposizione 2.25 *Ogni polinomio è una distribuzione temperata.*

DIM. Sia $p(x) = \sum_{\alpha \leq k} a_\alpha x^\alpha$ un polinomio. Poniamo $q(x) = p(x)(1 + |x|^{2N})$ e scriviamo $p = \frac{q}{1 + |x|^{2N}}$. Sia $\varphi_n \in \mathcal{S}(\mathbb{R}^n)$ tale che $\varphi_n \rightarrow 0$ in $\mathcal{S}(\mathbb{R}^n)$. Allora

$$\begin{aligned} \left| \int_{\mathbb{R}^n} p(x) \varphi_n(x) dx \right| &= \left| \int_{\mathbb{R}^n} \frac{q(x)}{1 + |x|^{2N}} \varphi_n(x) dx \right| \\ &\leq \left(\int_{\mathbb{R}^n} \frac{dx}{1 + |x|^{2N}} \right) \sup_{x \in \mathbb{R}^n} |q(x) \varphi_n(x)| \\ &\leq \left(\int_{\mathbb{R}^n} \frac{dx}{1 + |x|^{2N}} \right) \sum_{\alpha \leq k} \sup_{x \in \mathbb{R}^n} |x^\alpha (1 + |x|^{2N}) \varphi_n(x)| \rightarrow 0. \end{aligned}$$

□

La dimostrazione della precedente proposizione prova di fatto che ogni *funzione a crescita lenta*, ossia decomponibile nel prodotto di un polinomio e di una funzione sommabile, definisce una distribuzione temperata. Tuttavia, una distribuzione temperata non può crescere più velocemente di un polinomio, come dimostra la seguente osservazione.

Osservazione 2.26 Proviamo che $e^x \notin \mathcal{S}'(\mathbb{R})$. Sia $\varphi \in \mathcal{D}(\mathbb{R})$, con le seguenti proprietà: $\varphi \geq 0$, $\varphi \equiv 1$ in $[0, 1]$ e $\text{supp } \varphi \subset (-2, 2)$. Poniamo $\varphi_k(x) = \frac{1}{2^k} \varphi\left(\frac{x}{k}\right)$ e notiamo che il supporto di φ_k è contenuto in $(-2k, 2k)$. Allora $\varphi_k \rightarrow 0$ in $\mathcal{S}'(\mathbb{R})$, poiché per ogni scelta di interi positivi p, q risulta $x^p D^q \varphi_k(x) = 2^{-k} k^{-q} x^p D^q \varphi\left(\frac{x}{k}\right)$ e quindi

$$\sup_{x \in \mathbb{R}} |x^p D^q \varphi_k(x)| \leq \frac{k^{p-q} 2^p}{2^k} \|D^q \varphi\|_\infty \rightarrow 0.$$

Tuttavia

$$\int_{\mathbb{R}} e^x \varphi_k(x) dx \geq \frac{1}{2^k} \int_0^k e^x dx = \frac{e^k - 1}{2^k} \rightarrow +\infty.$$

Proposizione 2.27 *Se T è una distribuzione a supporto compatto allora T è anche una distribuzione temperata.*

DIM. Per quanto osservato nella sezione §2.4, T verifica la seguente proprietà: se $v_n, v \in C^\infty(\Omega)$ e se la successione v_n converge a v con le derivate di qualunque ordine uniformemente sui compatti di Ω , allora $\langle T, v_n \rangle \rightarrow \langle T, v \rangle$. Sulla base di questo fatto, non è difficile provare che se $\varphi_n \in \mathcal{S}(\mathbb{R}^n)$ e $\varphi_n \rightarrow 0$ in $\mathcal{S}(\mathbb{R}^n)$ allora $\langle T, \varphi_n \rangle \rightarrow 0$. □

La classe delle distribuzioni temperate è chiusa rispetto all'operazione di derivazione. ■
 Così ogni distribuzione temperata può essere derivata infinite volte. La moltiplicazione per una funzione f è lecita se f verifica delle condizioni sufficienti a garantire che $f\varphi \in \mathcal{S}(\mathbb{R}^n)$, per ogni $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Ad esempio, il prodotto di una distribuzione temperata con un polinomio è ancora una distribuzione temperata.

In modo del tutto analogo a quanto visto in $\mathcal{D}'(\mathbb{R}^n)$, è possibile definire limite e serie di una successione di distribuzioni temperate in $\mathcal{S}'(\mathbb{R}^n)$.

Concludiamo con un'osservazione, che motiva l'introduzione delle distribuzioni temperate. Queste sono più idonee di quelle classiche per estendere la trasformata di Fourier. Infatti, se $\varphi \in \mathcal{S}(\mathbb{R}^n)$, allora la sua trasformata di Fourier $\widehat{\varphi}$ è ancora un elemento di $\mathcal{S}(\mathbb{R}^n)$ e quindi è ben posta la seguente definizione

$$\langle \widehat{T}, \varphi \rangle = \langle T, \widehat{\varphi} \rangle, \quad \varphi \in \mathcal{S}(\mathbb{R}^n),$$

dove $T \in \mathcal{S}'(\mathbb{R}^n)$. In questo modo, molte proprietà della trasformata di Fourier già note per le funzioni si trasferiscono automaticamente alle distribuzioni temperate.

Esempio 2.28 Calcoliamo la trasformata di Fourier della delta di Dirac in 0:

$$\langle \widehat{\delta}, \varphi \rangle = \langle \delta, \widehat{\varphi} \rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}^n} \varphi(x) dx,$$

sicché $\widehat{\delta} = 1$, dove l'eguaglianza ha il significato che la distribuzione *delta* coincide con quella associata alla funzione costante di valore 1, che appartiene ad $L^1_{\text{loc}}(\mathbb{R}^n)$. Questo esempio va confrontato con il Teorema 1.59: infatti, ricordando che $\widehat{f * g} = \widehat{f} \widehat{g}$, dire che $f * \rho_\varepsilon \rightarrow f$ equivale formalmente a dire che $\widehat{f * \rho_\varepsilon} = \widehat{f} \widehat{\rho_\varepsilon} \rightarrow \widehat{f}$, cioè che $\widehat{\rho_\varepsilon} \rightarrow 1$, o $\rho_\varepsilon \rightarrow \delta$, che segue appunto dal Teorema 1.59.

Capitolo 3

Misure di Lebesgue-Stieltjes e funzioni a variazione limitata

3.1 Misura di Lebesgue-Stieltjes

Nell'introdurre il concetto di misura di Lebesgue in $\mathbb{R}^n$, si è partiti dalla definizione ordinaria di area o volume di un rettangolo. In particolare, la costruzione per il caso unidimensionale poggia sulla nozione di *lunghezza* di un intervallo. In questa sezione, affrontiamo il problema di introdurre in dimensione uno il concetto di misura in un modo più generale.

Sia $F : \mathbb{R} \rightarrow \mathbb{R}$ una funzione crescente e continua a destra. Per ogni $a \leq b$, poniamo

$$\begin{aligned}m_F(]a, b]) &= F(b) - F(a) \\m_F([a, b]) &= F(b) - F(a^-) \\m_F([a, b[) &= F(b^-) - F(a^-) \\m_F(]a, b[) &= F(b^-) - F(a).\end{aligned}$$

Indichiamo con $\mathcal{I}$ la famiglia di tutti gli intervalli aperti, chiusi o semiaperti di $\mathbb{R}$. La funzione m_F così definita è non negativa e additiva in $\mathcal{I}$, cioè

$$I = \bigcup_{h=1}^n I_h \quad \Longrightarrow \quad m_F(I) = \sum_{h=1}^n m_F(I_h),$$

per ogni $I, I_h \in \mathcal{I}$, con $I_k \cap I_h = \emptyset$.

Al fine di estendere m_F ad una classe più ampia di insiemi, ragioniamo come per la misura di Lebesgue, corrispondente al caso $F(x) = x$ e introduciamo il concetto di misura esterna.

Definizione 3.1 Sia $E \subseteq \mathbb{R}$. Definiamo

$$\overline{m}_F(E) = \inf \left\{ \sum_{h=1}^{\infty} m_F(I_h) \mid E \subseteq \bigcup_{h=1}^{\infty} I_h, I_h \in \mathcal{I} \right\}.$$

Dunque, $\overline{m}_F(E)$ è l'estremo inferiore tra tutte le misure dei ricoprimenti di E con famiglie finite o numerabili di intervalli. Vediamo subito di quali proprietà gode la funzione $\overline{m}_F$.

Proprietà

- a) per ogni $E \subseteq \mathbb{R}$, $\overline{m}_F(E) \in [0, +\infty]$ e $\overline{m}_F(\emptyset) = 0$;
- b) $\overline{m}_F$ è σ -subadditiva, cioè

$$E \subseteq \bigcup_{h=1}^{\infty} E_h \quad \implies \quad \overline{m}_F(E) \leq \sum_{h=1}^{\infty} \overline{m}_F(E_h).$$

La proprietà b) implica, in particolare, che $\overline{m}_F$ è crescente, ossia

$$E_1 \subseteq E_2 \quad \implies \quad \overline{m}_F(E_1) \leq \overline{m}_F(E_2).$$

La verifica di a) è immediata, quella di b) si fa come nel caso della misura esterna di Lebesgue $\overline{m}_n$. Una funzione d'insieme definita su tutte le parti di un insieme che soddisfa a) e b) si chiama, in generale, *misura esterna*.

A differenza di $\overline{m}_n$, la misura esterna $\overline{m}_F$ non è, in generale, invariante per traslazioni, né omogenea.

Si può verificare che se $I \in \mathcal{I}$, allora $\overline{m}_F(I) = m_F(I)$, pertanto $\overline{m}_F$ è un'estensione di m_F . Tuttavia, $\overline{m}_F$ non è additiva. Per rimediare a questo problema, è necessario restringere $\overline{m}_F$ ad un'opportuna classe di insiemi (che chiameremo misurabili).

Definizione 3.2 Se $E \subset \mathbb{R}$ è limitato, diremo che E è misurabile se per ogni $\varepsilon > 0$ esiste un pluriintervallo P , cioè un'unione finita di intervalli, tale che

$$\overline{m}_F(E \Delta P) < \varepsilon.$$

Dunque, E si approssima in modo arbitrariamente preciso mediante un insieme elementare.

Se E non è limitato, diremo che è misurabile se per ogni $R > 0$ l'insieme $E \cap]-R, R]$ è misurabile.

Indicheremo con $\mathcal{M}_F$ la classe dei sottoinsiemi misurabili di $\mathbb{R}$.

Teorema 3.3 *La collezione $\mathcal{M}_F$ costituisce una σ -algebra e la funzione $\overline{m}_F$ ristretta a $\mathcal{M}_F$ è una misura, che denoteremo con m_F , chiamata misura di Lebesgue-Stieltjes generata da F .*

Ogni insieme la cui misura esterna è nulla è misurabile. Basta scegliere $P = \emptyset$ nella Definizione 3.2. Naturalmente, la classe $\mathcal{M}_F$ dipende dalla scelta della funzione F . Tuttavia, si può vedere che essa contiene sempre gli insiemi aperti (e quindi i chiusi) di $\mathbb{R}$.

Osserviamo che la misura così costruita è automaticamente completa.

Esempi 3.4

1. Siano $x_1 < x_2 < \dots < x_n$ n punti in $\mathbb{R}$ e $h_1, \dots, h_n$ numeri positivi. Definiamo

$$F(x) = \sum_{x_j \leq x} h_j.$$

La funzione F è crescente, continua in $\mathbb{R} \setminus \{x_1, \dots, x_n\}$ e continua a destra in ogni x_i ove il salto è pari ad h_i . Allora

$$\mathcal{M}_F = \mathcal{P}(\mathbb{R}) \quad \text{e} \quad m_F = \sum_{i=1}^n h_i \delta_{x_i}.$$

Infatti, $m_F(x_i) = F(x_i) - F(x_i^-) = h_i$, per ogni i , e $m_F(\mathbb{R} \setminus \{x_1, \dots, x_n\}) = 0$. Se $A \subseteq \mathbb{R}$, allora $A \cap (\mathbb{R} \setminus \{x_1, \dots, x_n\})$ è misurabile, perché sottoinsieme di un insieme misurabile di misura nulla, e $A \cap \{x_1, \dots, x_n\}$ è misurabile perché unione di singoletti $\{x_i\} = [x_i, x_i]$, che sono misurabili. Pertanto, A è misurabile e $m_F(A) = m_F(A \cap \{x_1, \dots, x_n\}) = \sum_{x_i \in A} h_i$.

Questo risultato si estende al caso di un'infinità numerabile di punti $\{x_n\}$ e di numeri positivi $\{h_n\}$ tali che $\sum h_n < \infty$.

2. Sia ora $F \in C_c^1(\mathbb{R})$ e poniamo $f = F'$. Allora $\mathcal{M}_F = \mathcal{M}(\mathbb{R})$, classe dei misurabili secondo Lebesgue e

$$m_F(A) = \int_A f(x) dx, \quad \text{per ogni } A \in \mathcal{M}(\mathbb{R}).$$

Infatti, basta osservare che

$$m_F([a, b]) = F(b) - F(a) = \int_a^b f(x) dx = \int_{[a, b]} f(x) dx.$$

L'unicità del prolungamento di m_F assicura la tesi.

Osservazione 3.5 Se μ è una misura positiva finita sui boreliani di $\mathbb{R}$, allora

$$F(x) = \mu([-\infty, x]), \quad x \in \mathbb{R}$$

definisce una funzione crescente, limitata, continua a destra, la cui misura di Lebesgue Stieltjes coincide con μ . Inoltre, $\lim_{x \rightarrow -\infty} F(x) = 0$. Viceversa, se F è una siffatta funzione, allora posto $\tilde{F}(x) = m_F([-\infty, x])$, risulta $F = \tilde{F}$. Pertanto, le misure di Lebesgue-Stieltjes esauriscono tutte le misure finite sulla retta.

Sia m_F la misura di Lebesgue-Stieltjes generata dalla funzione F . Per questa misura si definisce in modo ordinario la classe delle funzioni sommabili e si introduce la nozione di integrale. L'integrale fatto rispetto a m_F si chiama *integrale di Lebesgue-Stieltjes* e si denota con il simbolo

$$\int f dF.$$

Se F è la funzione dell'Esempio 3.4-1. allora

$$(3.1) \quad \int_A f dF = \sum_{x_i \in A} f(x_i) h_i, \quad A \subset \mathbb{R}.$$

Ciò si verifica considerando dapprima funzioni caratteristiche, quindi funzioni semplici, per passare poi alle funzioni misurabili positive e infine alle funzioni sommabili.

Nel caso in cui F soddisfi le condizioni dell'Esempio 3.4-2. si trova che

$$(3.2) \quad \int_A g dF = \int_A g F' dx, \quad A \in \mathcal{M}(\mathbb{R}).$$

Anche qui, la verifica si fa per passi.

3.2 Misura immagine e applicazioni nella teoria delle probabilità

L'integrale di Lebesgue-Stieltjes trova uso in numerose questioni applicative. In particolare, questa nozione si usa largamente nella teoria delle probabilità. Mostriamo un'applicazione che fa uso della nozione di misura immagine.

Siano $(X, \mathcal{E})$ uno spazio misurabile, $f : X \rightarrow \mathbb{R}$ una funzione misurabile e μ una misura qualunque in $(X, \mathcal{E})$. La formula

$$f_{\#}\mu(A) = \mu \circ f^{-1}(A) = \mu(f^{-1}(A)), \quad \forall A \text{ aperto},$$

definisce una misura in $\mathbb{R}$, detta *misura immagine* di μ mediante f . L'importanza di questa nozione è quella di produrre in modo automatico un teorema di cambiamento di variabili.

Teorema 3.6 *Per ogni u funzione sommabile rispetto alla misura $f_{\#}\mu$, risulta che $u \circ f$ è sommabile rispetto a μ e vale*

$$\int_{\mathbb{R}} u df_{\#}\mu = \int_X (u \circ f) d\mu.$$

Ricordiamo che si chiama *spazio di probabilità* uno spazio $(\Omega, \mathcal{E}, P)$ dotato di una misura positiva P con $P(\Omega) = 1$. Una *variabile aleatoria* in $(\Omega, \mathcal{E}, P)$ è una funzione $X : \Omega \rightarrow \mathbb{R}$ misurabile. Si chiama *legge* o *distribuzione* di X la funzione definita da

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega \mid X(\omega) \leq x\}),$$

per ogni $x \in \mathbb{R}$. Dunque, $F(x)$ è la probabilità che la variabile aleatoria X assuma valori più piccoli di x . Osserviamo che $F(x) = (P \circ X^{-1})(]-\infty, x])$, dove $P \circ X^{-1}$ è la misura immagine di P mediante X . È evidente che F è crescente, continua a destra e soddisfa $F(-\infty) = 0$, $F(+\infty) = 1$. Inoltre si può verificare che $P \circ X^{-1}$ coincide con la misura di Lebesgue-Stieltjes generata da F , ossia

$$(3.3) \quad m_F = P \circ X^{-1}.$$

Tra le caratteristiche essenziali di una variabile aleatoria, c'è la sua *media* o *speranza*, definita da

$$E(X) = \int_{\Omega} X dP.$$

In virtù del Teorema 3.6 e della relazione (3.3), otteniamo che

$$E(X) = \int_{\mathbb{R}} (X \circ X^{-1}) d(P \circ X^{-1}) = \int_{\mathbb{R}} x dF.$$

In questo modo, per conoscere $E(X)$ è sufficiente calcolare l'integrale a secondo membro che è in $\mathbb{R}$, quindi sicuramente più semplice da maneggiare rispetto all'integrale in Ω .

Di solito in probabilità si incontrano variabili aleatorie discrete o assolutamente continue. Una variabile aleatoria X si dice discreta se può assumere solo un numero finito o numerabile di valori $x_1, \dots, x_n, \dots$. In questo caso, la distribuzione di X sarà una funzione di salto come quella considerata nell'Esempio 3.4-1. Infatti, posto $p_j = P(X = x_j)$, risulta $F(x) = \sum_{x_j \leq x} p_j$. Grazie alla formula (3.1) troviamo

che $E(X) = \sum_{j=1}^{\infty} x_j p_j$. A titolo esemplificativo, consideriamo la variabile aleatoria X_1 che conta il tempo di attesa per il verificarsi di un certo evento, ad esempio “la pioggia a Lecce”. Se tale evento ha probabilità p di verificarsi e $1 - p$ di non verificarsi in un giorno, allora $P(X_1 = k) = (1 - p)^{k-1} p$, per ogni $k \in \mathbb{N}$, esprime la probabilità che per $k-1$ giorni non piova e che piova al k -simo giorno. Per conoscere il tempo medio di attesa (inteso come numero di giorni) per il primo giorno di pioggia a Lecce bisognerà calcolare la media di X_1 , dunque $\sum_{k=1}^{\infty} k(1 - p)^{k-1} p = \frac{1}{p}$.

Un'altra categoria di variabili aleatorie sono quelle per cui la misura immagine $P \circ X^{-1}$ è assolutamente continua rispetto alla misura di Lebesgue, per cui esiste $f \in L^1(\mathbb{R})$ tale che la distribuzione di X si esprime nella forma

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

Siccome $F'(x) = f(x)$, la formula (3.2) permette di ottenere $E(X) = \int_{\mathbb{R}} x f(x) dx$.

3.3 Funzioni a variazione limitata

Sia $I = (a, b)$ un intervallo di $\mathbb{R}$ con $-\infty \leq a < b \leq +\infty$ e sia $u : I \rightarrow \mathbb{R}$. Definiamo

$$\text{pV}(u, I) = \sup \left\{ \sum_{i=1}^n |u(t_i) - u(t_{i-1})| \mid a < t_0 < t_1 < \dots < t_n < b \right\}$$

variazione puntuale di u in I . Essa dipende fortemente dai valori puntuali di u . Dato che avremo a che fare con funzioni che sono solo in $L^1_{\text{loc}}(I)$ e dato che queste possono essere modificate arbitrariamente su insiemi di misura nulla senza che cambi la loro sommabilità, è utile introdurre la seguente quantità

$$\text{eV}(u, I) = \inf \{ \text{pV}(v, I) \mid v = u \text{ q.o.} \},$$

detta *variazione essenziale* di u in I . Infine, chiamiamo

$$\text{V}(u, I) = \sup \left\{ \int_a^b u(x) \varphi'(x) dx \mid \varphi \in C_c^1(I), |\varphi(x)| \leq 1, \forall x \in I \right\},$$

variazione di u in I .

Vediamo una serie di osservazioni ed esempi relativi alla variazione puntuale.

- Se $pV(u, I) < +\infty$ allora u è limitata in I .

Infatti, fissato $t_0 \in I$, se $t < t_0$ consideriamo la partizione $t < t_0$. Dalla definizione deduciamo che

$$|u(t_0) - u(t)| \leq pV(u, I),$$

da cui $|u(t)| \leq |u(t_0)| + pV(u, I)$. Il caso $t > t_0$ è analogo.

- Se u è monotona e limitata, allora $pV(u, I) < +\infty$ e $pV(u, I) = |u(b^-) - u(a^+)|$.

Supponiamo u crescente per fissare le idee. Per ogni partizione $a < t_0 < t_1 < \dots < t_n < b$, risulta $\sum_{i=1}^n |u(t_i) - u(t_{i-1})| = u(t_n) - u(t_0)$. Al variare della partizione, $t_n \rightarrow b^-$ e $t_0 \rightarrow a^+$ per cui $pV(u, I) = u(b^-) - u(a^+)$.

- Se $u \in C^1[a, b]$, con a, b finiti, allora $pV(u, I) = \int_a^b |u'(x)| dx$.

Se $a \leq t_0 < t_1 < \dots < t_n \leq b$ è una partizione di $[a, b]$, allora per il Teorema di Lagrange esistono $\xi_i \in [t_{i-1}, t_i]$ tali che

$$\sum_{i=1}^n |u(t_i) - u(t_{i-1})| = \sum_{i=1}^n |u'(\xi_i)|(t_i - t_{i-1}).$$

Da qui, l'asserto.

- Se $a < c < b$, allora $pV(u, [a, b]) = pV(u, [a, c]) + pV(u, [c, b])$, cioè la variazione puntuale è additiva.
- Consideriamo la seguente funzione

$$u(x) = \begin{cases} x \cos(\pi/x), & x \in]0, 1] \\ 0, & x = 0. \end{cases}$$

Allora u è continua in $[0, 1]$, ma $pV(u, [0, 1]) = +\infty$. Infatti, osservato che

$$\cos(\pi/x) = 1, \quad \text{se } x = 1/(2k), \quad k \in \mathbb{N},$$

$$\cos(\pi/x) = -1, \quad \text{se } x = 1/(2k+1), \quad k \in \mathbb{N},$$

in corrispondenza della partizione $P_N = \{1/n \mid n = 1, \dots, N\}$ si ha che

$$\sum_{n=1}^N \left| u\left(\frac{1}{n+1}\right) - u\left(\frac{1}{n}\right) \right| = \sum_{n=1}^N \left(\frac{1}{n+1} + \frac{1}{n} \right) \rightarrow +\infty$$

per $N \rightarrow +\infty$.

Proposizione 3.7 *Sia $u : I \rightarrow \mathbb{R}$ una funzione tale che $pV(u, I) < +\infty$. Allora esistono due funzioni $u_1, u_2 : I \rightarrow \mathbb{R}$ limitate e crescenti tali che $u = u_1 - u_2$ e $pV(u, I) = pV(u_1, I) + pV(u_2, I)$.*

Inoltre, se u è continua a destra (risp. a sinistra), anche u_1, u_2 sono continue a destra (risp. a sinistra).

DIM. Per ogni $t \in I = (a, b)$, poniamo

$$g(t) = pV(u, (a, t]) = \sup \left\{ \sum_{i=1}^n |u(t_i) - u(t_{i-1})| \mid a < t_0 < t_1 < \cdots < t_n \leq t \right\}.$$

Siccome $0 \leq g(t) \leq pV(u, I) < +\infty$, g è limitata. Proviamo che g è crescente. Siano $t < s$ e $a < t_0 < t_1 < \cdots < t_n \leq t < s$. Allora

$$|u(t) - u(s)| + \sum_{i=1}^n |u(t_i) - u(t_{i-1})| \leq g(s),$$

da cui, al variare della partizione

$$|u(t) - u(s)| + g(t) \leq g(s),$$

che, in forma equivalente si legge

$$(3.4) \quad |u(s) - u(t)| \leq g(s) - g(t).$$

Definiamo ora

$$u_1 = \frac{g + u}{2}, \quad u_2 = \frac{g - u}{2}.$$

Chiaramente $u = u_1 - u_2$, u_1, u_2 sono limitate e, in virtù di (3.4) u_1, u_2 sono anche crescenti. In particolare, si ha $pV(u_i, I) = u_i(b^-) - u_i(a^+)$, $i = 1, 2$ e quindi $pV(u_1, I) + pV(u_2, I) = g(b^-) - g(a^+) \leq g(b^-) = pV(u, I)$. Per il viceversa, basta osservare che

$$\begin{aligned} \sum_{i=1}^n |u(t_i) - u(t_{i-1})| &\leq \sum_{i=1}^n |u_1(t_i) - u_1(t_{i-1})| + \sum_{i=1}^n |u_2(t_i) - u_2(t_{i-1})| \\ &\leq pV(u_1, I) + pV(u_2, I). \end{aligned}$$

□

Vediamo ora un teorema importante che mette in luce la relazione tra le nozioni introdotte all'inizio del capitolo.

Teorema 3.8 *Sia $u \in L^1_{\text{loc}}(I)$. Allora $V(u, I) = eV(u, I)$ ed esiste una funzione $\tilde{u}$ tale che $\tilde{u} = u$ q.o. e $pV(\tilde{u}, I) = eV(u, I)$ (cioè $\tilde{u}$ realizza il minimo).*

La funzione $\tilde{u}$ la cui esistenza è garantita dal teorema precedente prende il nome di *buon rappresentante* di u . Ogni funzione monotona è buon rappresentante di sé stessa.

Definizione 3.9 *Se $u \in L^1(I)$ e $V(u, I) < +\infty$, allora u si dice a variazione limitata e si scrive $u \in BV(I)$.*

Dalla dimostrazione del Teorema 3.8 si ricavano alcune proprietà fondamentali delle funzioni a variazione limitata che enunciamo nella proposizione e nel teorema seguenti.

Proposizione 3.10 *Se $u \in BV(I)$, allora la derivata distribuzionale di u è una misura reale μ definita sui boreliani contenuti in I , $\mathcal{B}(I)$, cioè*

$$\int_I u(x) \varphi'(x) dx = - \int_I \varphi d\mu, \quad \forall \varphi \in C_c^1(I).$$

Inoltre, $|\mu|(I) = V(u, I)$.

D'ora in avanti, indicheremo la derivata distribuzionale di u con Du .

Teorema 3.11 *Se $u \in BV(I)$ e $\tilde{u}$ è buon rappresentante di u allora $\tilde{u}$ è derivabile q.o. in I e la derivata puntuale di $\tilde{u}$ coincide con la densità (della parte assolutamente continua) di Du rispetto a $\mathcal{L}^1$, ossia*

$$Du = \tilde{u}' \mathcal{L}^1 + (Du)^s,$$

con $(Du)^s \perp \mathcal{L}^1$.

Spesso, in quello che segue, trattando una funzione in $BV(I)$ intenderemo riferirci al suo buon rappresentante.

In virtù di quanto visto sinora possiamo affermare che

- (1) ogni funzione $u \in BV(I)$ è differenza di due funzioni crescenti;
- (2) ogni funzione $u \in BV(I)$ ammette limiti sinistro e destro in $x \in I$ finiti ed esistono al più un'infinità numerabile di punti di discontinuità (di prima specie).

Notiamo che la proprietà (2) discende da (1) e dalle proprietà delle funzioni monotone.

Esempio 3.12 [L'insieme di Cantor] Poniamo $E_0 = [0, 1]$. Prendiamo ora in E_0 i due intervalli $[0, \frac{1}{3}]$ e $[\frac{2}{3}, 1]$, ciascuno di ampiezza $1/3$ e definiamo $E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. In ogni intervallo che compone E_1 ripetiamo lo spesso discorso prendendo i sottointervalli $[0, \frac{1}{9}]$, $[\frac{2}{9}, \frac{1}{3}]$, ecc... e chiamiamo E_2 l'unione dei quattro intervalli, ciascuno di ampiezza $\frac{1}{9}$, così ottenuti.

Al passo n -simo, E_n sarà costituito da 2^n intervalli disgiunti ciascuno di ampiezza $\frac{1}{3^n}$. Notiamo che $E_n \subset E_{n-1} \subset E_0$. Poniamo infine $E = \bigcap_{n \in \mathbb{N}_0} E_n$ e chiamiamo E *insieme di Cantor*. Risulta

$$\mathcal{L}^1(E) = \lim_{n \rightarrow \infty} \mathcal{L}^1(E_n) = \lim_{n \rightarrow \infty} (2/3)^n = 0.$$

Di seguito elenchiamo le proprietà fondamentali di E .

- (i) E è compatto, cioè chiuso e limitato.
- (ii) E è perfetto, cioè ogni suo punto è di accumulazione e non ce ne sono fuori.

Infatti, sia $x \in E$, dunque $x \in E_n$ per ogni n . Fissato $\varepsilon > 0$, prendiamo $n \in \mathbb{N}$ tale che $\frac{1}{3^n} < \varepsilon$. Sia $I_n = [a_n, b_n]$ l'unico intervallo tra quelli di E_n contenente x . Se, ad esempio, $x \neq a_n$ allora $|x - a_n| \leq \frac{1}{3^n} < \varepsilon$.

- (iii) E è in nessuna parte denso, cioè E non contiene intervalli.

Se E contenesse un intervallo di ampiezza $\delta > 0$, allora, scegliendo n abbastanza grande, si avrebbe che E_n è formato da intervalli disgiunti di ampiezza $\frac{1}{3^n} < \delta$, assurdo!

- (iv) E ha la cardinalità del continuo. Infatti, se si scrivono i punti di $[0, 1]$ in forma ternaria $x = 0, a_1 a_2 a_3 \dots a_n \dots$ con $a_i \in \{0, 1, 2\}$, per costruzione i punti di C sono esattamente quelli per cui a_n non è mai 1 (questi punti corrispondono ai "terzi medi" di ogni suddivisione); allora, la funzione che ad $x \in C$ associa

$$f(x) = 0, \frac{a_1}{2} \frac{a_2}{2} \frac{a_3}{2} \dots \frac{a_n}{2} \dots$$

è ben definita (perché per $x \in C$ ogni $a_n/2$ è 0 o 1) e, pensando lo sviluppo di $f(x)$ in base 2, la funzione f risulta bigettiva su $[0, 1]$.

Esempio 3.13 [La funzione di Cantor] Poniamo

$$f_n(x) = \int_0^x \left(\frac{3}{2}\right)^n \chi_{E_n}(t) dt, \quad x \in [0, 1], \quad n = 0, 1, \dots$$

dove E_n è l'insieme costruito nell'esempio precedente. Ogni f_n è continua, crescente e costante a tratti in E_n^c . Inoltre, f_n è derivabile a tratti con $f'_n \in L^1(0, 1)$. La successione $(f_n(x))$ è definitivamente costante se $x \notin E$. Si può anche dimostrare che vale la seguente stima

$$|f_{n+1}(x) - f_n(x)| \leq \frac{1}{2^{n-1}}, \quad n \in \mathbb{N}, \quad x \in [0, 1].$$

Pertanto la successione (f_n) soddisfa il criterio di Cauchy uniformemente in $[0, 1]$ e, di conseguenza, esiste $f \in C([0, 1])$ limite uniforme di (f_n) . Possiamo aggiungere che $f(0) = 0$, $f(1) = 1$ e f è crescente. Alla luce di queste proprietà, si vede che $f \in BV(0, 1)$.

Se $x_0 \notin C$, allora esiste $k \in \mathbb{N}$ tale che $x \notin E_k$. Essendo $E_{n+1} \subset E_n$, passando al complementare risulta $E_n^c \subset E_{n+1}^c$. Dunque $x_0 \notin E_n$, per ogni $n \geq k$. Se I è la componente di E_k^c contenente x , allora $f_n(x) = f_k(x) = f_k(x_0)$, per ogni $x \in I$. Ciò dimostra che f è costante in I , pertanto esiste $f'(x_0) = 0$. Data l'arbitrarietà di x_0 , ricaviamo che f' esiste ed è uguale a zero in E^c , cioè q.o., essendo $\mathcal{L}^1(E) = 0$.

Per la Proposizione 3.10, la derivata distribuzionale di f , Df , è una misura sui boreliani di $[0, 1]$. Dato che f è crescente, essa è buon rappresentante di sé stessa. Quindi, la Proposizione 3.10 insieme al fatto che $f' = 0$ q.o. provano che Df non ha parte assolutamente continua rispetto alla misura di Lebesgue. Ne segue che $Df \perp \mathcal{L}^1$. Proviamo che tale misura è singolare rispetto alla misura di Lebesgue. Risulta $\mathcal{L}^1(E) = 0$, in virtù dell'esempio precedente. Se facciamo vedere che $Df(E^c) = 0$, allora il nostro asserto è provato. Sia I una componente connessa di E^c , ossia uno degli intervalli scartati durante la costruzione di E . Sia $\varphi \in C_c^\infty(I)$. Per definizione di Df risulta allora

$$\int_I f \varphi' dx = - \int_I \varphi d(Df).$$

Ma f è costante in I . Dunque può essere portata fuori dall'integrale a primo membro e ciò che rimane è zero, essendo φ nulla agli estremi di I . Pertanto $\int_I \varphi d(Df) = 0$ per ogni $\varphi \in C_c^\infty(I)$, che implica $Df(I) = 0$. Essendo I arbitrario, ne segue che $Df(E^c) = 0$.

Osserviamo che la funzione di Cantor costituisce un controesempio al Teorema Fondamentale del Calcolo Integrale per l'integrale di Lebesgue. Infatti, risulta

$$f(1) - f(0) = 1 > 0 = \int_0^1 f'.$$

Pertanto, le funzioni a variazione limitata non sono il candidato ideale per ottenere l'estensione del Teorema Fondamentale del Calcolo Integrale all'integrale di Lebesgue. Peraltro, l'esempio della funzione di Cantor dimostra che il problema non è legato né all'esistenza q.o. della derivata puntuale, né alla sua integrabilità. A questo proposito però è opportuno indicare il seguente risultato.

Proposizione 3.14 *Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione tale che esiste $f'(x)$ per ogni $x \in [a, b]$ e $f' \in L^1(a, b)$. Allora*

$$f(x) - f(a) = \int_a^x f', \quad \forall x \in [a, b].$$

Nel paragrafo seguente introduciamo nell'ambito delle funzioni a variazione limitata la più grande sottofamiglia di funzioni che soddisfano il Teorema Fondamentale del Calcolo Integrale.

3.4 Funzioni assolutamente continue

Definizione 3.15 *Sia I un intervallo in $\mathbb{R}$. Diciamo che una funzione $f \in BV(I)$ è assolutamente continua in I se la sua derivata distribuzionale Df è assolutamente continua rispetto alla misura di Lebesgue in I . Scriveremo, in tal caso, che $f \in AC(I)$.*

Se $f \in AC(I)$, ricordando il Teorema 3.11, abbiamo che

$$Df([\alpha, \beta]) = \int_{\alpha}^{\beta} f' dx, \quad [\alpha, \beta] \subset I.$$

Osservazione 3.16 Siccome ogni funzione assolutamente continua è anche a variazione limitata, essa può essere scritta come differenza di due funzioni crescenti limitate. Non è difficile verificare che la funzione g , definita nella dimostrazione della Proposizione 3.7, è assolutamente continua se u lo è. Pertanto, se $u \in AC(I)$ allora $u = u_1 - u_2$ con u_1, u_2 crescenti, limitate e assolutamente continue.

La risposta al nostro problema è stabilita nel seguente importante teorema.

Teorema 3.17 *Se $f : I \rightarrow \mathbb{R}$ è assolutamente continua allora per ogni $[\alpha, \beta] \subset I$ risulta*

$$(3.5) \quad \int_{\alpha}^{\beta} f' dx = f(\beta) - f(\alpha).$$

DIM. Sia $f \in AC(I)$. Per l'Osservazione 3.16 possiamo scrivere $f = f_1 - f_2$, con f_1, f_2 crescenti, limitate e assolutamente continue. Pertanto, non è restrittivo dimostrare l'identità (3.5) per una funzione assolutamente continua crescente. Se $[\alpha, \beta] \subset I$, risulta allora

$$f(\beta) - f(\alpha) = pV(f, [\alpha, \beta]) = V(f, [\alpha, \beta]) = |Df|([\alpha, \beta]) = \int_{\alpha}^{\beta} |f'| dx = \int_{\alpha}^{\beta} f' dx.$$

Per differenza, la tesi è vera nel caso generale. $\square$

Grazie alla formula (3.5) è facile dimostrare la seguente proprietà.

Proposizione 3.18 *Sia $f : I \rightarrow \mathbb{R}$ una funzione assolutamente continua. Allora per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che comunque si scelgano n intervalli disgiunti*

$$(t_0, t_1), \dots, (t_{n-1}, t_n) \subset I \text{ con } \sum_{i=1}^n (t_i - t_{i-1}) < \delta \text{ risulta } \sum_{i=1}^n |f(t_i) - f(t_{i-1})| < \varepsilon.$$

DIM. In virtù della Proposizione 1.40, fissato $\varepsilon > 0$ esiste $\delta > 0$ tale che se E è un boreliano con $\mathcal{L}^1(E) < \delta$ allora $|Df|(E) < \varepsilon$. Scegliamo ora n intervalli disgiunti $(t_0, t_1), \dots, (t_{n-1}, t_n) \subset I$, la cui ampiezza complessiva sia minore di δ . Posto $E = \cup_{i=1}^n (t_{i-1}, t_i)$, si ha allora $|Df|(\cup_{i=1}^n (t_{i-1}, t_i)) < \varepsilon$. D'altra parte, grazie a (3.5) possiamo fare la seguente stima

$$\sum_{i=1}^n |f(t_i) - f(t_{i-1})| \leq \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |f'| dt = \sum_{i=1}^n |Df|(t_{i-1}, t_i) = |Df|(\cup_{i=1}^n (t_{i-1}, t_i)) < \varepsilon,$$

che è quanto si voleva provare. $\square$

Notiamo che la Proposizione precedente implica che ogni funzione assolutamente continua è uniformemente continua e quindi continua. Mentre non vale il viceversa. La funzione di Cantor ne è un controesempio. Inoltre, se I è limitato allora vale anche il viceversa della Proposizione 3.18, sicchè essa rappresenta una caratterizzazione delle funzioni assolutamente continue su intervalli limitati.

Le proprietà dell'integrale di Lebesgue permettono di dimostrare il seguente risultato.

Proposizione 3.19 *Sia $f \in L^1(a, b)$. Allora la funzione $g(x) = \int_a^x f$ è assolutamente continua. Inoltre, g è derivabile per quasi ogni $x \in I$ e risulta $g' = f$ q.o.*

Quest'ultima proposizione è il viceversa del Teorema 3.17. Quindi, possiamo concludere che tutte e sole le funzioni per cui vale la formula del Teorema Fondamentale del Calcolo Integrato con l'integrale di Lebesgue sono le funzioni assolutamente continue.

Capitolo 4

Spazi di Hilbert e serie di Fourier

Nelle applicazioni, soprattutto in Fisica, lo spazio ambiente naturale per numerosissimi modelli è lo spazio L^2 introdotto nella Definizione 1.10. Estenderemo ora ad alcuni spazi di dimensione infinita le nozioni di prodotto scalare, base, proiezione ortogonale. In seguito vedremo l'applicazione di questi concetti generali allo spazio L^2 .

Definizione 4.1 *Sia E uno spazio vettoriale reale. Un'applicazione $\langle u, v \rangle : E \times E \rightarrow \mathbb{R}$ è detta prodotto scalare se soddisfa le seguenti proprietà di linearità, simmetria e positività:*

$$\begin{aligned}\langle \alpha u + \beta v, w \rangle &= \alpha \langle u, w \rangle + \beta \langle v, w \rangle & \forall u, v, w \in E, \quad \forall \alpha, \beta \in \mathbb{R} \\ \langle u, v \rangle &= \langle v, u \rangle & \forall u, v \in E \\ \langle u, u \rangle &> 0 & \forall u \in E \setminus \{0\}.\end{aligned}$$

In particolare si ha linearità rispetto anche alla seconda variabile e

$$\langle 0, v \rangle = \langle v, 0 \rangle = 0 \quad \forall v \in E.$$

Nel caso degli spazi vettoriali complessi si richiede che l'applicazione $\langle \cdot, \cdot \rangle$ sia a valori complessi e l'assioma di simmetria è sostituito da

$$\langle u, v \rangle = \overline{\langle v, u \rangle} \quad \forall u, v \in E.$$

Esempi 4.2

(1) Lo spazio $\mathbb{R}^n$ è in genere munito del prodotto scalare euclideo

$$\langle u, v \rangle = \sum_{i=1}^n u_i v_i.$$

(2) Sia A una matrice simmetrica $n \times n$ definita positiva; allora

$$\langle u, v \rangle_A := \langle Au, v \rangle = \sum_{i,j=1}^n a_{ij} u_j v_i$$

definisce un nuovo prodotto scalare in $\mathbb{R}^n$.

(3) Lo spazio $\mathbb{C}^n$ è in genere munito del prodotto scalare (detto hermitiano)

$$\langle u, v \rangle = \sum_{i=1}^n u_i \bar{v}_i.$$

(4) Sia ℓ^2 lo spazio vettoriale delle successioni reali (a_n) a quadrato sommabile:

$$(a_n) \in \ell^2 \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} a_n^2 < +\infty.$$

Il fatto che ℓ^2 sia uno spazio vettoriale segue dalla disuguaglianza $(a+b)^2 \leq 2a^2 + 2b^2$. Inoltre, la disuguaglianza $2|ab| \leq a^2 + b^2$ consente di definire un prodotto scalare in ℓ^2 :

$$\langle (a_n), (b_n) \rangle := \sum_{n=1}^{\infty} a_n b_n.$$

Sullo spazio vettoriale delle successioni complesse (a_n) a quadrato sommabile:

$$\sum_{n=1}^{\infty} |a_n|^2 < +\infty.$$

si definisce il prodotto scalare:

$$\langle (a_n), (b_n) \rangle := \sum_{n=1}^{\infty} a_n \bar{b}_n.$$

(5) Per $(X, \mathcal{E}, \mu)$ spazio di misura, possiamo definire il prodotto scalare

$$\langle u, v \rangle = \int_X u \bar{v} d\mu$$

per $u, v : X \rightarrow \mathbb{C}$ appartenenti ad $L^2(X)$, e risulta $\|u\|_2 = \sqrt{\langle u, u \rangle}$.

Proposizione 4.3 *Ogni prodotto scalare in E soddisfa la disuguaglianza di Cauchy-Schwartz*

$$|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle \langle v, v \rangle} \quad \forall u, v \in E.$$

Inoltre la funzione $\|u\| := \sqrt{\langle u, u \rangle}$ è una norma in E .

Dim. Fissiamo u e v in E ; per ogni $t \in \mathbb{R}$ si ha

$$0 \leq \langle tu + v, tu + v \rangle = t^2 \langle u, u \rangle + 2t \langle u, v \rangle + \langle v, v \rangle.$$

Il discriminante del trinomio di secondo grado dovrà allora essere negativo; si trova quindi

$$\frac{\Delta}{4} = \langle u, v \rangle^2 - \langle u, u \rangle \langle v, v \rangle \leq 0.$$

Verifichiamo ora che gli proprietà delle norme sono soddisfatte. La positività e la omogeneità sono evidenti. Per verificare la disuguaglianza triangolare osserviamo che

$$\|u + v\| \leq \|u\| + \|v\|$$

equivale, elevando al quadrato ambo i membri, a

$$\langle u + v, u + v \rangle \leq \langle u, u \rangle + \langle v, v \rangle + 2\|u\|\|v\|$$

quindi semplificando si ottiene

$$\langle u, v \rangle \leq \|u\|\|v\|$$

che segue dalla disuguaglianza di Cauchy-Schwartz. □

Osservazione 4.4 La disuguaglianza di Cauchy-Schwartz implica che, a v fissato, le applicazioni $u \mapsto \langle u, v \rangle$ sono Lipschitziane di costante $\|v\|$. Vedremo in seguito che, sotto opportune ipotesi, *ogni* funzionale lineare e continuo è di questo tipo.

Dati $u, v \in E$ non nulli, *l'angolo* tra u e v è a volte definito come l'unico $\theta \in [0, 2\pi)$ tale che

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\|\|v\|}.$$

Tale angolo esiste per la disuguaglianza di Cauchy-Schwartz.

Definizione 4.5 Chiameremo spazio pre-Hilbertiano ogni spazio vettoriale reale munito di un prodotto scalare e della norma corrispondente. La distanza indotta dal prodotto scalare sarà naturalmente

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}.$$

Se E , munito della distanza d , è uno spazio metrico completo allora diremo che E è uno spazio di Hilbert.

Esempio 4.6 L'insieme $C([a, b])$ delle funzioni continue in $[a, b]$, munito del prodotto scalare

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx$$

è uno spazio pre-Hilbertiano.

Osservazione 4.7

(1) Si può verificare che ogni norma derivata da un prodotto scalare soddisfa l'identità del parallelogramma

$$(4.1) \quad \|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Viceversa, se $(E, \|\cdot\|)$ è uno spazio normato e la norma soddisfa l'identità del parallelogramma, allora

$$\langle x, y \rangle := \frac{\|x + y\|^2 - \|x - y\|^2}{4}$$

definisce un prodotto scalare in E che induce la norma di partenza.

(2) Negli spazi di pre-hilbertiani, che sono ovviamente dei particolari spazi normati, sussistono le nozioni di convergenza di successioni e di completezza. Vale il seguente teorema di completamento:

Per ogni spazio pre-Hilbertiano E esistono uno spazio di Hilbert H ed una isometria $j : E \rightarrow H$ tali che l'immagine $J(E)$ è densa in H , vale a dire, ogni $x \in H$ può essere approssimato da una successione $(x_h) \subset j(E)$. Lo spazio H è unico a meno di isometrie ed è detto completamento di E .

(3) Nello spazio $L^2(X)$ ovviamente una successione di funzioni (u_n) converge ad u se e solo se

$$\lim_{n \rightarrow \infty} \int_X |u_n - u|^2 d\mu = 0,$$

e notiamo che, ragionando come nel caso di L^1 , si prova che se $u_n \rightarrow u$ in L^2 allora esiste una sottosuccessione (u_{k_n}) convergente ad u μ -q.o. Notiamo anche che, ragionando ancora come in L^1 , si dimostra che $L^2(X)$ è completo (per $X \subseteq \mathbb{R}^n$ aperto può essere costruito come completamento di $C(X)$).

Esempio 4.8 Il completamento di $C([a, b])$ è lo spazio $L^2(a, b)$ (più precisamente, lo spazio quoziente) munito del prodotto scalare

$$\langle f, g \rangle := \int_a^b f(t)g(t) dt.$$

Il concetto di base in dimensione infinita richiede il passaggio da combinazioni lineari *finite* a combinazioni lineari *infinite* ma con coefficienti a quadrato sommabile.

Definizione 4.9 Sia H uno spazio di Hilbert di dimensione infinita e sia $(v_n) \subset H$ un insieme di vettori. Diremo che (v_n) è una base ortonormale di H se

(1) $\langle v_n, v_m \rangle = \delta_{nm}$ per ogni $n, m \geq 1$;

(2) per ogni $u \in H$ esiste una successione $(a_n) \subset \mathbb{R}^n$ tale che

$$(4.2) \quad u = \sum_{n=1}^{\infty} a_n v_n.$$

Diremo che uno spazio di Hilbert è separabile se ha almeno una base ortonormale.

La proprietà di separabilità è legata al fatto di avere un'infinità al più numerabile di direzioni indipendenti e a due a due ortogonali. Tra gli spazi di Hilbert separabili vi è ℓ^2 : basta scegliere

$$v_n = (0, 0, \dots, 0, 1, 0, \dots)$$

con il numero 1 all' n -mo posto. I vettori v_n sono a due a due ortogonali ed hanno norma 1. Data inoltre $(a_n) \in \ell^2$ definiamo

$$(a_n)_N = \sum_{i=1}^N a_i v_i = (a_1, a_2, \dots, a_N, 0, 0, \dots).$$

Si ha quindi

$$(a_n) - (a_n)_N = (0, 0, \dots, 0, a_{N+1}, a_{N+2}, \dots)$$

$$\lim_{N \rightarrow +\infty} \|(a_n) - (a_n)_N\|^2 = \lim_{N \rightarrow +\infty} \sum_{i=N+1}^{\infty} a_i^2 = 0$$

il che vuol dire

$$(a_n) = \lim_{N \rightarrow +\infty} (a_n)_N = \sum_{n=1}^{\infty} a_n v_n.$$

La seguente proposizione mostra che ℓ^2 è il modello isometrico di tutti gli spazi di Hilbert separabili di dimensione infinita.

Proposizione 4.10 *Sia H uno spazio di Hilbert separabile di dimensione infinita e sia (v_n) una base ortonormale. Allora i coefficienti (a_n) , detti coordinate di u rispetto alla base (v_n) , sono univocamente determinati dalla (4.2) e valgono le relazioni $a_n = \langle u, v_n \rangle$. Vale inoltre l'identità di Parseval:*

$$\|u\|^2 = \sum_{n=1}^{\infty} a_n^2 = \sum_{n=1}^{\infty} \langle u, v_n \rangle^2.$$

Infine, l'applicazione

$$(4.3) \quad u \mapsto (\langle u, v_n \rangle)$$

è una isometria lineare e bigettiva di H con ℓ^2 .

Dim. Supponiamo che valga la (4.2) e prendiamo i prodotti scalari di ambo i membri con v_m , per m fissato. Essendo il prodotto scalare continuo e lineare otteniamo

$$(4.4) \quad \begin{aligned} \langle u, v_m \rangle &= \left\langle \sum_{n=1}^{\infty} a_n v_n, v_m \right\rangle = \lim_{N \rightarrow +\infty} \left\langle \sum_{n=1}^N a_n v_n, v_m \right\rangle \\ &= \lim_{N \rightarrow +\infty} \sum_{n=1}^N a_n \delta_{nm} = a_m. \end{aligned}$$

Posto $u_N = \sum_{i=1}^N a_i v_i$ abbiamo

$$\begin{aligned} \|u_N\|^2 &= \left\langle \sum_{i=1}^N a_i v_i, \sum_{j=1}^N a_j v_j \right\rangle = \sum_{i,j=1}^N a_i a_j \langle v_i, v_j \rangle \\ &= \sum_{i,j=1}^N a_i a_j \delta_{ij} = \sum_{i=1}^N a_i^2. \end{aligned}$$

Passando al limite per $N \rightarrow +\infty$ ed usando la continuità della norma si ottiene l'identità di Parseval.

Infine l'applicazione $\Phi : H \rightarrow \ell^2$ in (4.3) è evidentemente lineare ed è una isometria tra i due spazi per l'identità di Parseval. L'iniettività di Φ segue dalle implicazioni

$$\Phi(u) = \Phi(v) \Rightarrow \|\Phi(u - v)\| = 0 \Rightarrow \|u - v\| = 0 \Rightarrow u = v.$$

Data $(a_n) \in \ell^2$ si verifica facilmente che la successione $u_N = \sum_{n=1}^N a_n v_n$ è di Cauchy:

$$\|u_M - u_N\|^2 = \left\| \sum_{i=N+1}^M a_i v_i \right\|^2 = \sum_{i=N+1}^M a_i^2 \leq \sum_{i=N+1}^{\infty} a_i^2 \quad N < M.$$

Dato che $(a_n) \in \ell^2$ si ha

$$\lim_{N \rightarrow +\infty} \sum_{i=N+1}^{\infty} a_i^2 = 0.$$

Definendo

$$u := \sum_{n=1}^{\infty} a_n v_n$$

il ragionamento usato in (4.4) mostra che $a_n = \langle u, v_n \rangle$, quindi $\Phi(u) = (a_n)$. $\square$

Osservazione 4.11 Si osservi che le serie del tipo

$$\sum_{n=1}^{\infty} a_n v_n \quad (a_n) \in \ell^2$$

non convergono totalmente in generale (si prenda ad esempio $a_n = 1/n$).

Quali spazi di Hilbert sono separabili? È ben noto che il *metodo di ortonormalizzazione di Gram-Schmidt* fornisce una base ortonormale negli spazi di Hilbert di dimensione finita, che sono quindi separabili. Una caratterizzazione degli spazi di Hilbert separabili è fornita dal seguente teorema, che si dimostra anch'esso col metodo di Gram-Schmidt.

Teorema 4.12 *Uno spazio di Hilbert H è separabile se e solo se esiste un insieme al più numerabile D denso in H .*

Osservazione 4.13 Sia E uno spazio pre-Hilbertiano, sia $F \subset E$ un sottospazio di dimensione N e $v_1, \dots, v_N$ una base ortonormale di F . Verificare che per ogni $u \in E$ la funzione

$$\left\| u - \sum_{i=1}^N a_i v_i \right\|^2$$

ha valore minimo per $a_i = \langle u, v_i \rangle$. Il vettore

$$\sum_{i=1}^N \langle u, v_i \rangle v_i$$

è detto *proiezione* di u su F .

Tra le proprietà fondamentali che gli spazi di Hilbert hanno in comune con gli spazi di dimensione finita ricordiamo l'esistenza di proiezioni sui convessi chiusi (anche di dimensione infinita) e la possibilità di rappresentare ogni funzionale lineare e continuo mediante il prodotto scalare per un opportuno vettore (teorema di Riesz).

Teorema 4.14 Sia H uno spazio di Hilbert e $C \subset H$ un insieme convesso, chiuso, non vuoto. Allora per ogni $x \in H$ esiste unico $y \in C$ tale che

$$\|x - y\| = \min_{z \in C} \|x - z\|.$$

Il vettore y è detto *proiezione* di x sul convesso C ed è caratterizzato dalla proprietà:

$$(4.5) \quad \langle x - y, z - y \rangle \leq 0 \quad \forall z \in C.$$

Dim. Sia $(y_h) \subset C$ una successione *minimizzante*, vale a dire

$$\lim_{h \rightarrow +\infty} \|y_h - x\| = m := \inf_{y \in C} \|y - x\|.$$

Usando l'identità del parallelogramma (4.1) otteniamo

$$\begin{aligned} 2 \left\| \frac{y_h - y_k}{2} \right\|^2 &= \|y_h - x\|^2 + \|y_k - x\|^2 - 2 \left\| \frac{y_h + y_k}{2} - x \right\|^2 \\ &\leq \|y_h - x\|^2 + \|y_k - x\|^2 - 2m \end{aligned}$$

quindi

$$\lim_{h, k \rightarrow +\infty} \|y_h - y_k\| = 0$$

e per la completezza di H possiamo concludere che y_h converge a y . Inoltre, dato che C è chiuso, $y \in C$ e

$$\|y - x\| = \lim_{h \rightarrow +\infty} \|y_h - x\| = m$$

quindi y è il minimo cercato. L'unicità del minimo segue ancora dall'identità del parallelogramma: se y, y' sono minimi si ha

$$2\left\|\frac{y - y'}{2}\right\|^2 = \|y - x\|^2 + \|y' - x\|^2 - 2\left\|\frac{y + y'}{2} - x\right\|^2 \leq m + m - 2m = 0$$

quindi $y = y'$. Verifichiamo ora la (4.5): scelto $z \in C$, poniamo $y_t := y + t(z - y) \in C$ ed usiamo la disuguaglianza (per $t \in (0, 1]$)

$$\|x - y\|^2 \leq \|x - y_t\|^2 = \|x - y\|^2 - 2t\langle x - y, z - y \rangle + t^2\|z - y\|^2.$$

Semplificando $\|x - y\|^2$, dividendo ambo i membri per t e facendo tendere t a zero otteniamo la (4.5). Viceversa, se la (4.5) vale per un certo $y \in C$, fissiamo $z \in C$ e consideriamo la funzione

$$\varphi(t) := \|(y + t(z - y)) - x\|^2.$$

Dalla (4.5) deduciamo che φ è crescente (basta sviluppare la norma) quindi

$$\|z - x\|^2 = \varphi(1) \geq \varphi(0) = \|y - x\|^2.$$

Essendo z arbitrario, y è minimo. □

Osservazione 4.15 Elenchiamo alcune proprietà di cui lasciamo la verifica come esercizio.

(1) Se C è uno spazio vettoriale affine la formula sopra equivale

$$\langle x - y, z - y \rangle = 0 \quad \forall z \in C$$

e se C è uno spazio vettoriale, $\langle x - y, z \rangle = 0$ per ogni $z \in C$. Si dice in tal caso che $x - y$ è *ortogonale* a C . L'insieme dei vettori ortogonali a C si indica con $C^\perp$. Possiamo quindi dire che x si decompone in un vettore $(x - y) \in C^\perp$ ed un vettore $y \in C$.

(2) Siano $g_1, g_2 : (a, b) \rightarrow \mathbb{R}$ funzioni misurabili e sia

$$C := \{f \in L^2(a, b) : g_1(x) \leq f(x) \leq g_2(x) \text{ per q.o. } x \in (a, b)\}.$$

Verificare che C è un convesso chiuso di $L^2(a, b)$ e determinare l'operatore di proiezione da $L^2(a, b)$ in C e trovare condizioni su g_1 e g_2 che garantiscano che C è non vuoto.

(3) Siano $(E_1, \|\cdot\|_1)$, $(E_2, \|\cdot\|_2)$ spazi normati completi e $E := E_1 \times E_2$. Verificare che

$$\|(x, y)\| := \sqrt{\|x\|_1^2 + \|y\|_2^2}$$

rende E uno spazio normato completo.

Teorema 4.16 (Teorema di Riesz) *Sia H uno spazio di Hilbert e $L : H \rightarrow \mathbb{R}$ un operatore lineare e continuo, non identicamente nullo. Allora esiste unico $x \in H$ tale che $L(y) = \langle x, y \rangle$ per ogni $y \in H$.*

Dim. (esistenza) Sia C lo spazio vettoriale chiuso $L^{-1}(0)$ e sia $v \in H \setminus C$. Detta $u \in C$ la proiezione ortogonale di v su C , indichiamo con w il vettore $v - u$ normalizzato e verifichiamo che le applicazioni lineari

$$L(y), \quad L(w)\langle w, y \rangle$$

coincidono. Infatti sono entrambe nulle su C , per definizione di C e per l'ortogonalità di w (Osservazione 4.15.1). Inoltre le due applicazioni coincidono per $y = w$, quindi sullo spazio vettoriale $\mathbf{R}_w$ generato da w . La coincidenza delle due funzioni seguirà allora dal fatto che ogni $y \in H$ può essere decomposto nella somma di un vettore di C ed un vettore di $\mathbf{R}_w$, infatti

$$y = \left(y - \frac{L(y)}{L(w)} w \right) + \frac{L(y)}{L(w)} w$$

ed il vettore tra parentesi appartiene a C :

$$L\left(y - \frac{L(y)}{L(w)} w\right) = L(y) - \frac{L(y)}{L(w)} L(w) = 0.$$

Basta allora porre $x := L(w)w$ per avere la tesi.

(unicità) Se $\langle x, y \rangle = \langle x', y \rangle$ per ogni $y \in H$, prendendo $y = x - x'$ otteniamo $\|x - x'\|^2 = 0$ quindi $x = x'$. $\square$

Vediamo ora l'applicazione di questi discorsi astratti ad un particolare spazio di funzioni, lo spazio $L^2(-\pi, \pi)$.

Definizione 4.17 *Diremo che f è regolare a tratti in $[-\pi, \pi]$ se f è di classe C^1 in $[-\pi, \pi]$ privato di al più un numero finito di punti ed in questi punti f e f' hanno limiti destro e sinistro finiti.*

Consideriamo ora una funzione regolare a tratti in $[-\pi, \pi]$ (estesa per periodicità ad $\mathbb{R}$) e la serie di Fourier ad essa associata:

$$(4.6) \quad f(x) = \sum_{n=0}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx$$

con i *coefficienti dello sviluppo* dati da

$$a_n = \begin{cases} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx & \text{se } n = 0; \\ \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx & \text{se } n > 0 \end{cases}$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx.$$

Vale allora il

Teorema 4.18 *Sia f regolare a tratti e 2π periodica in $\mathbb{R}$.*

(a) *la serie di Fourier di f converge puntualmente alla media dei limiti destro e sinistro:*

$$\sum_{n=0}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx = \frac{1}{2} \left(\lim_{y \rightarrow x^+} f(y) + \lim_{y \rightarrow x^-} f(y) \right);$$

(b) *la serie di Fourier di f converge uniformemente in tutti gli intervalli chiusi che non contengono punti di discontinuità di f .*

L'interpretazione della serie di Fourier nel contesto degli spazi di Hilbert è data dal

Teorema 4.19 *Lo spazio $H = L^2(-\pi, \pi)$ è di Hilbert e separabile. Una base ortonormale di H è costituita dalle funzioni*

$$v_0 = \frac{1}{\sqrt{2\pi}} \quad v_n = \frac{1}{\sqrt{\pi}} \cos nx \quad n = 1, 2, \dots \quad w_n = \frac{1}{\sqrt{\pi}} \sin nx \quad n = 1, 2, \dots$$

Dim. Si verifica immediatamente che le funzioni v_n, w_n sono a due a due ortogonali ed hanno norma 1. Consideriamo lo spazio H delle funzioni f del tipo

$$f = \sum_{n=0}^{\infty} c_n v_n + \sum_{n=1}^{\infty} d_n w_n$$

con coefficienti (c_n) , (d_n) in l_2 . L'ipotesi di sommabilità sui coefficienti garantisce che H è un sottospazio di $L^2(-\pi, \pi)$, vale a dire, le somme parziali convergono nella norma $L^2(-\pi, \pi)$. Gli stessi calcoli visti nel caso dell'identità di Parseval mostrano inoltre che

$$\|f\|^2 = \sum_{n=0}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2$$

quindi H è isometrico allo spazio $\ell^2 \times \ell^2$ munito della norma prodotto introdotta nell'esercizio 6. Lo spazio H , essendo isometrico ad uno spazio completo è completo e, in particolare, chiuso.

Osserviamo ora che le funzioni $C^1([-\pi, \pi])$ appartengono ad H . Infatti, data f di questo tipo, integrando per parti abbiamo (per $n \geq 1$)

$$\begin{aligned} c_n &= -\frac{1}{\sqrt{\pi}n} \int_{-\pi}^{\pi} f'(t) \sin nt \, dt, \\ d_n &= \frac{1}{\sqrt{\pi}n} (f(-\pi) - f(\pi)) \cos n\pi + \int_{-\pi}^{\pi} f'(t) \cos nt \, dt \end{aligned}$$

dalle quali è facile dedurre che (c_n) , (d_n) appartengono a ℓ^2 . Quindi la somma della serie è anche una funzione $g \in H$ in quanto i coefficienti sono di quadrato sommabile. Dette ora $s_n(t)$ le somme parziali della serie di Fourier, sappiamo che esse convergono puntualmente (tranne, in generale, i punti $\pm\pi$) a $f(t)$. D'altronde, sappiamo anche, dall'Osservazione 3, che vi è una sottosuccessione $s_{n_k}(t)$ convergente quasi ovunque in (a, b) a $g(t)$. Ne segue che $f = g \in H$.

Infine, sapendo che ogni funzione $f \in L^2(-\pi, \pi)$ è limite, in norma L^2 , di funzioni di classe $C^1([-\pi, \pi])$, concludiamo che $H = L^2(-\pi, \pi)$. $\square$

Per ogni funzione $f \in L^2(-\pi, \pi)$ abbiamo quindi

$$(4.7) \quad f = \sum_{n=0}^{\infty} c_n v_n + \sum_{n=1}^{\infty} d_n w_n$$

con

$$c_n = \langle f, v_n \rangle, \quad d_n = \langle f, w_n \rangle$$

e

$$\int_{-\pi}^{\pi} f^2(x) \, dx = \sum_{n=0}^{\infty} c_n^2 + \sum_{n=1}^{\infty} d_n^2$$

per l'identità di Parseval.

Osservazione 4.20

- (a) È facile vedere che se f è pari allora tutti i coefficienti b_n sono nulli e se f è dispari tutti i coefficienti a_n sono nulli. Di fatto, la prima sommatoria e la seconda sommatoria in (4.7) rappresentano rispettivamente la proiezione di f sullo spazio delle funzioni pari e la proiezione di f sullo spazio delle funzioni dispari.
- (b) Per avere basi ortonormali nei rettangoli di $\mathbb{R}^k$ basta considerare prodotti di seni e coseni dipendenti da una sola delle coordinate ed aventi tutte le possibili frequenze, cioè espressioni del tipo

$$\cos(n_1 x_1) \cos(n_2 x_2) \sin(n_3 x_3) \dots \cos(n_k x_k).$$

In questo caso il formalismo complesso è quasi d'obbligo per semplificare le notazioni. Ad esempio, nel rettangolo $(-\pi, \pi)^k$ una base ortonormale è

$$v_n(x) := \frac{1}{(2\pi)^{k/2}} e^{i\langle n, x \rangle} \quad n \in \mathbb{Z}^k.$$

- (c) Osserviamo che la convergenza della serie in (4.7) ha luogo in $L^2(-\pi, \pi)$, quindi in una metrica di tipo integrale. In base ad un teorema dimostrato da Lennart Carleson nel 1966 la serie di Fourier converge anche quasi ovunque in $(-\pi, \pi)$ per ogni $f \in L^2(-\pi, \pi)$.
- (c) Infine, ricordiamo altre possibili scelte di basi ortonormali in $L^2(a, b)$. Si chiamano *polinomi di Chebyshev* in (a, b) i polinomi ottenuti applicando il metodo di Gram-Schmidt ai vettori

$$(4.8) \quad 1, t, t^2, t^3, \dots$$

Nel caso $(a, b) = (-1, 1)$ i polinomi sono anche detti *polinomi di Legendre*. Nel caso $(a, b) = \mathbb{R}$, per ovviare al fatto che i polinomi non sono a quadrato sommabile si definisce

$$L_*^2(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ misurabile, } \int_{-\infty}^{+\infty} f^2(s) e^{-s^2} ds < +\infty \right\}$$

con prodotto scalare

$$\langle f, g \rangle := \int_{-\infty}^{+\infty} f(s) g(s) e^{-s^2} ds.$$

I polinomi ottenuti ortonormalizzando i vettori (4.8) si chiamano *polinomi di Hermite*.

Appendice A

Appendice – Richiami

A.1 Massimo e minimo limite

Ricordiamo che per una successione reale (a_n) si definiscono il *minimo limite* ed il *massimo limite* ponendo

$$\liminf_{n \rightarrow +\infty} a_n = \sup_{n \in \mathbb{N}} \inf_{k \geq n} a_k, \quad \limsup_{n \rightarrow +\infty} a_n = \inf_{n \in \mathbb{N}} \sup_{k \geq n} a_k.$$

Commentiamo solo la definizione di minimo limite, lasciano per esercizio la riformulazione delle considerazioni che seguono al caso del massimo limite. Data (a_n) , si può costruire la successione (e'_n) ponendo $e'_n = \inf_{k \geq n} a_k$ per ogni $n \in \mathbb{N}$. Siccome ad ogni passo si calcola l'estremo inferiore su un insieme più piccolo, la Successione (e'_n) è crescente e quindi esiste il suo limite, e coincide con $\sup_n \{e'_n\}$. Inoltre, dalle proprietà dell'estremo inferiore e dell'estremo superiore si ricava la seguente caratterizzazione del minimo limite:

$\ell = \liminf_{n \rightarrow +\infty} a_n$ se e solo se valgono le condizioni:

- (a) per ogni $\varepsilon > 0$ esiste $\nu > 0$ tale che $a_n > \ell - \varepsilon$ per ogni $n > \nu$;
- (b) per ogni $\varepsilon > 0$ e per ogni $\nu \in \mathbb{N}$ esiste $n > \nu$ tale che $a_n < \ell + \varepsilon$.

Ne segue in particolare che $\liminf_{n \rightarrow +\infty} a_n$ è il più piccolo tra i numeri reali ℓ che godono della proprietà che esiste una successione estratta da (a_n) convergente ad ℓ . Analogamente,

$\ell = \limsup_{n \rightarrow +\infty} a_n$ se e solo se valgono le condizioni:

- (a) per ogni $\varepsilon > 0$ esiste $\nu > 0$ tale che $a_n < \ell + \varepsilon$ per ogni $n > \nu$;

(b) per ogni $\varepsilon > 0$ e per ogni $\nu \in \mathbb{N}$ esiste $n > \nu$ tale che $a_n > \ell - \varepsilon$.

e $\limsup_{n \rightarrow +\infty} a_n$ è il più grande tra i numeri reali ℓ che godono della proprietà che esiste una successione estratta da (a_n) convergente ad ℓ . Ovvvia conseguenza di quanto detto è che $\liminf_{h \rightarrow \infty} f_h \leq \limsup_{h \rightarrow \infty} f_h$, con uguaglianza se e solo se esiste il limite di a_n .

A.2 Insiemi numerabili

Agli insiemi costituiti da un numero finito di elementi associamo in modo elementare il numero naturale che ci dice quanti elementi contiene, e fra due insiemi finiti esiste un'applicazione biunivoca se e solo se essi contengono lo stesso numero di elementi. Di conseguenza, non può esistere un'applicazione biunivoca tra un insieme finito ed una sua parte propria. Questa proprietà distingue gli insiemi finiti da quelli infiniti. Infatti, un insieme si dice *infinito* se è in corrispondenza biunivoca con una sua parte propria. L'esempio più semplice è probabilmente quello degli insiemi $\mathbb{N}$ e $\mathbb{P}$ dei numeri pari. Ovviamente $\mathbb{P} \subsetneq \mathbb{N}$, e quindi $\mathbb{P}$ è un sottoinsieme proprio di $\mathbb{N}$, ma ciò non ostante l'applicazione $f : \mathbb{N} \rightarrow \mathbb{P}$ definita da $f(n) = 2n$ è biunivoca ($\mathbb{N}$ contiene “solo” il doppio degli elementi di $\mathbb{P}$). Questo però non vuol dire che fra due insiemi infiniti esista *sempre* un'applicazione biunivoca, e quindi bisogna accettare l'idea che esiste una gerarchia fra gli insiemi infiniti, cioè che anche fra gli insiemi infiniti ce ne siano di “più o meno numerosi”. Quest'idea vaga si può formalizzare, ma ciò esula dagli scopi del corso: ci limitiamo ad osservare che se X è un insieme infinito allora non c'è *mai* un'applicazione biunivoca tra X e l'insieme $\mathcal{P}(X)$ di tutti i suoi sottoinsiemi, che risulta “molto più numeroso” di X . L'insieme $\mathbb{N}$ è “il più piccolo” insieme infinito, e si dice che un insieme E è *numerabile* se esiste un'applicazione surgettiva $x = x_n : \mathbb{N} \rightarrow E$. Rivolgendo il nostro interesse agli insiemi numerici, è ovvio che anche l'insieme $\mathbb{Z}$ dei numeri interi è numerabile (contiene “solo” il doppio degli elementi di $\mathbb{N}$), mentre la situazione è di gran lunga meno intuitiva per quanto riguarda $\mathbb{Q}$ ed $\mathbb{R}$. Risulta:

Teorema A.1 *L'insieme $\mathbb{Q}$ è numerabile, mentre $\mathbb{R}$ non lo è.*

Dim. Scriviamo $\mathbb{Q}$ come matrice infinita

$$\begin{array}{ccccccc}
 \frac{1}{1} & \rightarrow & \frac{2}{1} & & \frac{3}{1} & \rightarrow & \frac{4}{1} & \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \cdots \\
 \frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & \cdots \\
 \downarrow & \nearrow & & \swarrow & & \nearrow & & \cdots \\
 \frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & & \frac{4}{3} & \cdots \\
 \vdots & \swarrow & & \nearrow & & \swarrow & & \cdots
 \end{array}$$

dove le frecce hanno il seguente significato: partendo dall'elemento in alto a sinistra (il numero 1) la freccia porta induttivamente il numero razionale $f(n)$ nel termine successivo $f(n+1)$ della successione $(f(n))_{n \in \mathbb{N}}$ che stiamo definendo. Ci si convince immediatamente che la funzione così definita è surgettiva e quindi che $\mathbb{Q}$ è numerabile.

Viceversa, supponiamo che $\mathbb{R}$ sia numerabile. allora lo sarà anche l'intervallo $[0, 1]$, sicché, fissata una convenzione che induca un'unica rappresentazione decimale con parte intera nulla, possiamo scrivere l'intero intervallo $[0, 1]$ come successione (x_n) in modo tale che

$$\begin{aligned}
 x_1 &= 0, a_{11} a_{12} a_{13} \cdots \\
 x_2 &= 0, a_{21} a_{22} a_{23} \cdots \\
 x_3 &= 0, a_{31} a_{32} a_{33} \cdots \\
 &\vdots \\
 x_n &= 0, a_{n1} a_{n2} a_{n3} \cdots a_{nn} \cdots \\
 &\vdots
 \end{aligned}$$

e le cifre decimali a_{nk} sono numeri interi tra 0 e 9. Allora, scelta la successione (b_n) tale che $b_n \neq a_{nn}$ per ogni n , il numero reale $\bar{x} = 0, b_1 b_2 b_3 \cdots b_n \cdots$ appartiene a $[0, 1]$ perché ha parte intera nulla, ma non è nessuno degli x_n perché per ogni n differisce da x_n (almeno) per l' n -esima cifra decimale. Questo prova che $[0, 1]$ (e quindi $\mathbb{R}$) non è numerabile.

A.3 Il teorema di Riemann sui riordinamenti

Richiamiamo alcune proprietà delle serie numeriche usate in relazione alle misure reali. Le considerazioni che seguono dovrebbero far capire quanta distanza ci sia tra le serie e le somme finite. Una delle proprietà più “ovvie” delle somme

finite è la proprietà *commutativa*. È naturale domandarsi se essa valga anche per le serie. Per formulare correttamente il problema bisogna introdurre il concetto di *permutazione dei termini di una serie*. Date la serie $\sum_k a_k$ ed una funzione bigettiva $\pi : \mathbb{N} \rightarrow \mathbb{N}$ (permutazione), si dice *serie ottenuta permutando i termini di $\sum_k a_k$ secondo π* la serie $\sum_k a_{\pi(k)}$. Notiamo che i valori assunti dalla successione $(a_{\pi(k)})_{k \in \mathbb{N}}$ sono *gli stessi di $(a_k)_{k \in \mathbb{N}}$, e vengono assunti lo stesso numero di volte*, per cui se avessimo a che fare con una somma finita passare da $\sum_k a_k$ a $\sum_k a_{\pi(k)}$ si ridurrebbe a “cambiare l’ordine degli addendi”, ed è ben noto che in tal caso “la somma non cambia”. Per le serie infinite le cose vanno in modo completamente diverso, a meno che non si abbia convergenza assoluta.

Teorema A.2 *Sia $\sum_k a_k$ una serie semplicemente convergente. Allora:*

- (i) *se la serie $\sum_{k=0}^{\infty} a_k$ converge assolutamente e la sua somma è S , allora per ogni permutazione π la serie $\sum_{k=0}^{\infty} a_{\pi(k)}$ converge assolutamente ed ha per somma S .*
- (ii) *se la serie $\sum_{k=0}^{\infty} a_k$ non converge assolutamente, allora nessuna serie permutata converge assolutamente, ed inoltre per ogni $S \in \mathbb{R}$ esiste una permutazione π tale che la serie permutata $\sum_{k=0}^{\infty} a_{\pi(k)}$ converga (semplicemente) ad S .*

Non presentiamo la dimostrazione di questo risultato, ma ci limitiamo a sottolineare ancora la differenza tra le somme finite e le serie non assolutamente convergenti: queste, cambiando l’ordine degli addendi, possono dare *qualunque* somma!