Introduzione

Sono qui raccolti alcuni appunti sulla teoria della misura e dell'integrazione. Si consiglia in ogni caso di consultare i testi di Wheeden e Zygmund [6], Rudin [5] e Lieb e Loss [3], da cui il materiale qui esposto è stato preso. Segnalatemi gli errori, sicuramente presenti, all'indirizzo `zelati@unina.it`

Napoli, 9 novembre 2007

Capitolo 1

Preliminari

1.1. Insiemi

Se $\mathcal{F} = \{E\}$ è una collezione di sottoinsiemi di un insieme X (usualmente $X = \mathbb{R}$ o $X = \mathbb{R}^n$), indicheremo con

$$\bigcup_{E\in\mathcal{F}} E := \{\, x \mid x \in E \text{ per qualche } E \in \mathcal{F} \,\}$$
$$\bigcap_{E\in\mathcal{F}} E := \{\, x \mid x \in E \text{ per ogni } E \in \mathcal{F} \,\}.$$

(La notazione $\{\, a \mid b \,\}$ indica l'insieme di tutte le cose di tipo a che soddisfano alla condizione b e i simboli $a := b$ e $b =: a$ significano che a è definito da b.) Se la famiglia $\mathcal{F}$ è numerabile (diremo che è numerabile anche nel caso in cui sia finita) la chiamiamo una successione di insiemi, e la indichiamo con $\mathcal{F} = \{\, E_k \mid k = 1, 2, \dots \,\}$. In tal caso indichiamo l'unione e l'intersezione di $\mathcal{F}$ con

$$\bigcup_{k\in\mathbb{N}} E_k \qquad \bigcap_{k\in\mathbb{N}} E_k$$

o anche semplicemente con

$$\bigcup_k E_k \qquad \bigcap_k E_k.$$

Diremo che la successione $E_1, E_2, \dots$ è crescente se $E_k \subset E_{k+1}$, e che è decrescente se $E_k \supset E_{k+1}$. In questi casi scriveremo

$$E_k \nearrow \bigcup_k E_k \qquad E_k \searrow \bigcap_k E_k.$$

Definiamo poi il *massimo limite* e il *minimo limite* della successione di insiemi $E_1, E_2, \dots$ ponendo

$$\limsup_k E_k := \bigcap_{j=1}^{\infty} \Big(\bigcup_{k=j}^{\infty} E_k \Big)$$

$$\liminf_k E_k := \bigcup_{j=1}^{\infty} \Big(\bigcap_{k=j}^{\infty} E_k \Big).$$

Posto $U_j = \bigcup_{k=j}^{\infty} E_k$ e $V_j = \bigcap_{k=j}^{\infty} E_k$ ne segue che

$$U_j \searrow \limsup_k E_k \qquad V_j \nearrow \liminf_k E_k.$$

Si può dimostrare che $\limsup_k E_k$ è l'insieme dei punti che appartengono a infiniti E_k, mentre $\liminf_k E_k$ è l'insieme dei punti che appartengono a tutti gli E_k tranne al piú un numero finito. Ne segue che

$$\liminf_k E_k \subset \limsup_k E_k.$$

Ricordiamo anche che

$$\Big(\bigcup_{E \in \mathcal{F}} E \Big)^c = \bigcap_{E \in \mathcal{F}} E^c \qquad \Big(\bigcap_{E \in \mathcal{F}} E \Big)^c = \bigcup_{E \in \mathcal{F}} E^c.$$

1.2. Lo spazio euclideo $\mathbb{R}^n$

Ricordiamo qui alcuni fatti elementari e delle notazioni che ci serviranno in seguito. I numeri reali sono indicati con $\mathbb{R}$, i complessi con $\mathbb{C}$ e $\bar{z}$ indica il complesso coniugato del numero complesso z. Indicheremo con $[a, b]$ l'*intervallo chiuso* di $\mathbb{R}$, definito da $a \le x \le b$, e con (a, b) l'*intervallo aperto*, definito da $a < x < b$. Assumeremo noti i fatti fondamentali del calcolo negli *spazi euclidei n-dimensionali*

$$\mathbb{R}^n = \big\{ (x_1, \dots, x_n) \bigm| \text{ciascun } x_i \text{ sta in } \mathbb{R} \big\}.$$

La *distanza euclidea* tra due punti x e y in $\mathbb{R}^n$ è definita da $|x - y|$, dove, per ogni $x \in \mathbb{R}^n$,

$$|x| := \Big(\sum_{i=1}^{n} x_i^2 \Big)^{1/2}.$$

Assumiamo che il lettore conosca alcune diseguaglianze elementari come la *diseguaglianza triangolare*

$$|x| + |y| \ge |x - y|.$$

Assumeremo nota anche la definizione di *limite*, di *integrale secondo Riemann* e di *funzione differenziabile*. Ricordiamo poi che $\mathbb{R}^n$ è uno spazio metrico completo, che $\mathbb{Q}^n$ è denso in $\mathbb{R}^n$, cosí che $\mathbb{R}^n$ è uno spazio metrico separabile.

Data una successione $a_1, a_2, \ldots \in \mathbb{R}$, poniamo

$$b_j = \sup_{k \geq j} a_k \qquad c_j = \inf_{k \geq j} a_k.$$

Ne segue che

$$-\infty \leq c_j \leq a_j \leq b_j \leq +\infty,$$

e che b_j è una successione monotona decrescente, mentre c_j è monotona crescente. Poniamo

$$\limsup_{k \to +\infty} a_k := \lim_{j \to +\infty} b_j = \lim_{j \to +\infty} \sup_{k \geq j} a_k = \inf_j \sup_{k \geq j} a_k$$
$$\liminf_{k \to +\infty} a_k := \lim_{j \to +\infty} c_j = \lim_{j \to +\infty} \inf_{k \geq j} a_k = \sup_j \inf_{k \geq j} a_k$$

Un'utile caratterizzazione del $\liminf a_k$ e $\limsup a_k$ (detti rispettivamente minimo limite e massimo limite della successione a_n) è data dal seguente teorema

Teorema 1.2.1. *Sia $a_1, a_2, \ldots \in \mathbb{R}$ una successione di numeri reali. Allora*

(a) $L = \limsup_k a_k$ *se e solo se (i) esiste una sottosuccessione a_{k_j} convergente a L; (ii) per ogni $L' > L$ esiste $K \in \mathbb{N}$ tale che $a_k < L'$ per tutti i $k \geq K$;*

(b) $\ell = \liminf_k a_k$ *se e solo se (i) esiste una sottosuccessione a_{k_j} convergente a ℓ; (ii) per ogni $\ell' < \ell$ esiste $K \in \mathbb{N}$ tale che $a_k > \ell'$ per tutti i $k \geq K$;*

(c) $a = \lim_k a_k$ *se e solo se* $a = \liminf_k a_k = \limsup_k a_k$.

(d) *Se a_{k_j} è una sottosuccessione della successione a_k, e $\lim_j a_{k_j} = a$, allora*

$$\liminf_k a_k \leq a \leq \limsup_k a_k.$$

Il teorema sopra giustifica il nome di massimo e minimo limite.

Se $E \subset \mathbb{R}^n$, il suo *diametro* è

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

Se il diametro di E è finito, E si dice *limitato*. Se E_1, E_2 sono sottoinsiemi di $\mathbb{R}^n$, la *distanza di E_1 da E_2* è il numero reale

$$\operatorname{dist}(E_1, E_2) := \inf\{\, |x - y| \mid x \in E_1, y \in E_2 \,\}.$$

1.3. Aperti, chiusi e insiemi speciali di $\mathbb{R}^n$

In questa sezione ricordiamo alcune proprietà di $\mathbb{R}^n$.

Dati $x \in \mathbb{R}^n$ e $\delta > 0$, denotiamo la *palla aperta* di centro x e raggio δ con

$$B(x, \delta) := \{\, y \in \mathbb{R}^n \mid \operatorname{dist}(x, y) < \delta \,\}.$$

Un punto $x \in E$ si dice *punto interno* di E se, per qualche $\delta > 0$ la palla $B(x, \delta)$ è contenuta in E. E è *aperto* se tutti i suoi punti sono interni. La collezione dei punti interni ad E si chiama la parte interna di E e sarà indicata con $\mathring{E}$. Un punto $x \in \mathbb{R}^n$ si dice *punto di accumulazione* per E se tutte le palle $B(x, \delta)$ contengono punti di E diversi da x stesso.

Ricordiamo che un insieme E si dice *chiuso* se il suo complementare E^c è aperto, che la *chiusura di un insieme* $A \subset \mathbb{R}^n$ è l'unione di A e dei suoi punti di accumulazione. Denotiamo la chiusura di A con $\overline{A}$. Ricordiamo i seguenti teoremi:

Teorema 1.3.1. *Valgono le seguenti proprietà:*

(a) $\overline{B(x, \delta)} = \left\{ y \in \mathbb{R}^n \;\middle|\; \mathrm{dist}(x, y) \leq \delta \right\}$;

(b) $\overline{E}$ *è chiuso; è il piú piccolo chiuso contenente* E*;*

(c) $\overline{\overline{A}} = \overline{A}$*;*

(d) E *è chiuso se e solo se* $E = \overline{E}$.

Teorema 1.3.2. *Valgono le seguenti proprietà:*

(a) *L'unione di aperti è aperta;*

(b) *L'intersezione di un numero finito di aperti è aperta;*

(c) *L'intersezione di chiusi è chiusa;*

(d) *L'unione di un numero finito di chiusi è chiusa;*

Ricordiamo che $E_1 \subset E \subset \mathbb{R}^n$ è *aperto relativamente ad* E se $E_1 = E \cap G$ con G aperto in $\mathbb{R}^n$, ed è *chiuso relativamente ad* E se $E_1 = E \cap F$, F chiuso in $\mathbb{R}^n$.

Teorema 1.3.3. $E_1 \subset E$ *è chiuso relativamente ad* E *se e solo se* $E_1 = E \cap \overline{E_1}$.

Non tutti i sottoinsiemi di $\mathbb{R}^n$ sono aperti o chiusi. Una classe importante di sottoinsiemi (che incontreremo in seguito) è la seguente

Definizione 1.3.4. Diciamo che $B \subset \mathbb{R}^n$ è un $\mathcal{G}_\delta$ (scriveremo anche $B \in \mathcal{G}_\delta$) se B è l'intersezione di una successione (numerabile) di aperti, e che B è un $\mathcal{F}_\sigma$ ($B \in \mathcal{F}_\sigma$) se B è l'unione di una successione (numerabile) di chiusi.

Osservazione 1.3.5. È facile vedere che

(a) Se B è un $\mathcal{G}_\delta$, allora B^c è un $\mathcal{F}_\sigma$ (ed analogamente se B è un $\mathcal{F}_\sigma$ allora B^c è un $\mathcal{G}_\delta$);

(b) Se $B \subset \mathbb{R}^n$ è aperto o chiuso, allora B è un $\mathcal{G}_\delta$ ed anche un $\mathcal{F}_\sigma$.

Tutto (o quasi, si veda l'esercizio 6) quanto detto sopra vale in realtà anche sostituendo $\mathbb{R}^n$ con uno spazio metrico. Passiamo ora ad analizzare proprietà specifiche di $\mathbb{R}^n$. Il seguente teorema ci mostra come sono fatti gli aperti in $\mathbb{R}$.

Teorema 1.3.6 (Georg Cantor, 1882, [2])**.** *Sia $E \subset \mathbb{R}$ aperto. Allora E è l'unione di una famiglia numerabile di intervalli aperti a due a due disgiunti.*

Dimostrazione. Per ogni $x \in E$, sia I_x il piú grande intervallo aperto tale che (i) $x \in I_x$; (ii) $I_x \subset E$. Un tale intervallo esiste; basta considerare l'unione della famiglia di intervalli I che soddisfano (i) e (ii) (tale famiglia è non vuota in quanto E è aperto). Tale unione è ovviamente un intervallo aperto.

Dati x e y in E, mostriamo che I_x e I_y sono disgiunti oppure coincidenti. In effetti, se $z \in I_x \cap I_y$, $I_x \cup I_y$ è un intervallo aperto non vuoto contenente x e y in E. Ma allora $I_x \cup I_y \subset I_x$, da cui $I_x = I_x \cup I_y$.

Infine mostriamo che la collezione $\mathcal{F} = \{I_x\}$ è numerabile. Per far ciò basta osservare che ciascun $I \in \mathcal{F}$ contiene un razionale. $\square$

Un teorema analogo non vale per gli aperti di $\mathbb{R}^n$, cioè gli aperti di $\mathbb{R}^n$ non si possono caratterizzare come unioni numerabili di aperti "semplici". Si veda in ogni caso il teorema 1.3.8.

Definizione 1.3.7. Un rettangolo $R \subset \mathbb{R}^n$ è un insieme della forma

$$R = \prod_{i=1}^{n} [a_i, b_i] = \left\{ \xi \in \mathbb{R}^n \;\middle|\; \xi_i \in [a_i, b_i], i = 1, \ldots, n \right\}$$

per qualche scelta di numeri reali $a_i \leq b_i$. Nel caso in cui tutti i lati siano uguali (cioè $b_i - a_i = b_j - a_j$ per ogni i, j), il rettangolo si dice cubo. Dato un rettangolo, il suo diametro è dato da

$$\operatorname{diam}(R) = \left(\sum_{i=1}^{n} (b_i - a_i)^2 \right)^{1/2}$$

mentre definiamo il suo volume ponendo

$$\operatorname{vol}(R) = \prod_{i=1}^{n} (b_i - a_i).$$

Diremo inoltre che due rettangoli R_1 e R_2 sono non sovrapposti (o che non si sovrappongono) se le loro parti interne non si intersecano, cioè se $\mathring{R}_1 \cap \mathring{R}_2 = \varnothing$. Analogamente diremo che una collezione di rettangoli è una collezione di rettangoli non sovrapposti se questi sono a due a due non sovrapposti.

Dati $x_0, e_1, e_2, \dots, e_n \in \mathbb{R}^n$, il parallelepipedo chiuso P è l'insieme

$$P := \Big\{ x \ \Big|\ x = x_0 + \sum_{k=1}^{n} t_k e_k,\ 0 \le t_k \le 1 \Big\}.$$

Il volume di P è, per definizione,

$$\mathrm{vol}(P) = |\det[e_1, \dots, e_n]|.$$

Si può verificare che tale definizione coincide con quella data sopra per i rettangoli nel caso in cui P si riduce ad un rettangolo.

Se $\mathsf{T}\colon \mathbb{R}^n \to \mathbb{R}^n$ è un'applicazione lineare, il parallelepipedo P viene trasformato nel parallelepipedo P' e vale

$$\mathrm{vol}(P') = |\det(\mathsf{T})|\,\mathrm{vol}(P).$$

In particolare, se T è una rotazione, $\mathrm{vol}(P') = \mathrm{vol}(P)$.

Abbiamo detto che gli aperti di $\mathbb{R}^n$ non si possono caratterizzare come unione numerabile di aperti "semplici". Vale però il seguente teorema:

Teorema 1.3.8. *Sia $G \subset \mathbb{R}^n$ un aperto. Allora G è l'unione di una famiglia numerabile di cubi chiusi non sovrapposti.*

Dimostrazione. Sia $\mathcal{K}_0$ l'insieme dei cubi di lato 1 e vertici nei punti di coordinate intere. Dividendo a metà ciascun lato otteniamo da ciascun cubo in $\mathcal{K}_0$ 2^n cubi di lato $1/2$. Sia $\mathcal{K}_1$ la collezione di tali cubi. Procedendo analogamente, definiamo per ogni $j \in \mathbb{N}$ una famiglia $\mathcal{K}_j$ di cubi di lato 2^{-j} non sovrapposti e tali che ciascuno di essi è l'unione di 2^n cubi non sovrapposti in $\mathcal{K}_{j+1}$.

Sia G un aperto di $\mathbb{R}^n$. Sia $\mathcal{S}_0$ la collezione dei cubi in $\mathcal{K}_0$ contenuti in G, $\mathcal{S}_1$ la collezione dei cubi in $\mathcal{K}_1$ che sono contenuti in G ma non nei cubi di $\mathcal{S}_0$, e cosí via.

Posto $\mathcal{S} = \cup \mathcal{S}_j$, abbiamo che $\mathcal{S}$ è una collezione numerabile (in quanto unione numerabile di insiemi numerabili) di cubi non sovrapposti.

Mostriamo che $G = \cup_{Q \in \mathcal{S}} Q$. Ciò è una semplice conseguenza del fatto che, essendo G aperto, per ogni $x \in G$ esiste un cubo $Q \in \cup \mathcal{K}_j$ tale che $x \in Q \subset G$. $\square$

La famiglia di cubi $\{ Q \mid Q \in \mathcal{K}_j \text{ per qualche } j \}$ costruita sopra si dice una famiglia diadica di cubi.

Nel lemma seguente si enuncia un fatto intuitivamente ovvio la dimostrazione del quale è però piuttosto laboriosa.

Lemma 1.3.9. *Sia $\mathcal{C}$ una collezione finita di rettangoli. Allora*

(a) *Se $\mathcal{C}$ è una collezione di rettangoli non sovrapposti, tutti contenuti nel rettangolo J vale*

$$\mathrm{vol}(J) \geq \sum_{R \in \mathcal{C}} \mathrm{vol}(R)$$

(b) *Se $J \subset \bigcup_{R \in \mathcal{C}} R$, allora*

$$\mathrm{vol}(J) \leq \sum_{R \in \mathcal{C}} \mathrm{vol}(R)$$

Dimostrazione. Osserviamo che possiamo sempre assumere, anche nella dimostrazione del punto **(b)**, che ciascun rettangolo $R \in \mathcal{C}$ sia contenuto in J.

Infatti, sia **(b)** vera sotto l'ipotesi ulteriore $R \subset J$ per ogni $R \in \mathcal{C}$. Mostriamo che allora **(b)** vale in generale. Poniamo

$$\mathcal{C}' = \left\{ R \cap J \;\middle|\; R \in \mathcal{C} \right\}.$$

Allora $\mathcal{C}'$ è una collezione di rettangoli tutti contenuti in J, e

$$\bigcup_{R \in \mathcal{C}'} R = \bigcup_{R \in \mathcal{C}} R \cap J \supset J.$$

Ne segue che

$$\mathrm{vol}(J) \leq \sum_{R \in \mathcal{C}'} \mathrm{vol}(R) = \sum_{R \in \mathcal{C}} \mathrm{vol}(R \cap J) \leq \sum_{R \in \mathcal{C}} \mathrm{vol}(R).$$

Osserviamo poi che, se $J = \prod_1^N [a_k, b_k]$ e se $a_1 \leq c \leq b_1$, chiamati $J^+ = [a_1, c] \times \prod_2^N [a_k, b_k]$ e $J^- = [c, b_1] \times \prod_2^N [a_k, b_k]$, vale $\mathrm{vol}(J) = \mathrm{vol}(J^+) + \mathrm{vol}(J^-)$. Più in generale, se per ogni $1 \leq k \leq N$ $a_k = c_{k,0} \leq c_{k,1} \leq \cdots \leq c_{k,n_k} = b_k$, denotata con $\mathcal{S}$ la collezione dei rettangoli della forma

$$R(m_1, \ldots, m_N) = \prod_{k=1}^{N} [c_{k,m_k - 1}, c_{k,m_k}]$$

vale

$$\mathrm{vol}(J) = \sum_{R \in \mathcal{S}} \mathrm{vol}(R).$$

Sia ora $\mathcal{C}$ una collezione finita di rettangoli I contenuti nel rettangolo $J = \prod_1^N [a_k, b_k]$. Per ciascun $1 \leq k \leq N$ scegliamo $a_k = c_{k,0} \leq \cdots \leq c_{k,n_k} = b_k$ in modo che ciscun $I \in \mathcal{C}$ abbia la forma

$$I = \prod_{k=1}^{N} \left[c_{k,m_k}, c_{k,m'_k} \right]$$

per una qualche scelta di $0 \leq m_k \leq m'_k \leq n_k$. Sia $\mathcal{S}$ come prima la collezione di tutti i rettangoli della forma $R(m_1, \ldots, m_N)$. Allora si ha che

$$\mathrm{vol}(J) = \sum_{R \in \mathcal{S}} \mathrm{vol}(R),$$

e, analogamente, se $I \in \mathcal{C}$, indicata con $\mathcal{S}(I)$ la collezione di $R \in \mathcal{S}$ tali che $R \subset I$, si ha che

$$\mathrm{vol}(I) = \sum_{R \in \mathcal{S}(I)} \mathrm{vol}(R).$$

Dimostriamo ora il punto **(a)**. Siano I e $I' \in \mathcal{C}$. Se $R = R(m_1, \ldots, m_N) \in \mathcal{S}$ appartiene sia a $\mathcal{S}(I)$ che a $\mathcal{S}(I')$, necessariamente $c_{k,m_k-1} = c_{k,m_k}$ per almeno un $1 \leq k \leq N$. Allora $\mathrm{vol}(R) = 0$ e quindi

$$\mathrm{vol}(J) = \sum_{R \in \mathcal{S}} \mathrm{vol}(R) \geq \sum_{I \in \mathcal{C}} \sum_{R \in \mathcal{S}(I)} \mathrm{vol}(R) = \sum_{I \in \mathcal{C}} \mathrm{vol}(I).$$

Mostriamo infine **(b)**. Se $J \subset \bigcup_{I \in \mathcal{C}} I$, allora ogni $R \in \mathcal{S}$ appartiene a $\mathcal{S}(I)$ per almeno un $I \in \mathcal{C}$. Quindi

$$\mathrm{vol}(J) = \sum_{R \in \mathcal{S}} \mathrm{vol}(R) \leq \sum_{I \in \mathcal{C}} \sum_{R \in \mathcal{S}(I)} \mathrm{vol}(R) = \sum_{I \in \mathcal{C}} \mathrm{vol}(I).$$

.

$\square$

Definizione 1.3.10. Un insieme A si dice perfetto se A è chiuso e se ogni punto $x \in A$ è punto limite di A.

Una proprietà importante degli insiemi perfetti è la seguente:

Teorema 1.3.11. *Sia $E \subset \mathbb{R}^n$ un insieme perfetto non vuoto. Allora E non è numerabile.*

Dimostrazione. Supponiamo per assurdo che E sia numerabile. Allora $E = \{c_1, c_2, \ldots\}$. Definiamo $E_k = E \setminus \{c_k\}$. Dato $x_1 \in E_1$, esiste $0 < r_1 \leq \frac{1}{1}$ tale che $c_1 \notin \overline{B(x_1, r_1)}$. Poiché $x_1 \in E$, E perfetto, x_1 è un punto limite per E, e quindi anche per E_2. Ne segue che $E_2 \cap B(x_1, r_1)$ non è vuoto. Sia $x_2 \in E_2 \cap B(x_1, r_1)$, e scegliamo $0 < r_2 < \frac{1}{2}$ in modo tale che $\overline{B(x_2, r_2)} \subset B(x_1, r_1)$, $c_2 \notin \overline{B(x_2, r_2)}$. Continuando cosí otteniamo una successione $\{x_n\} \subset E$ e una successione $r_n \to 0$ tali che $x_n \in E_n \cap B(x_{n-1}, r_{n-1})$, $\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1})$, $c_n \notin \overline{B(x_n, r_n)}$. La successione $\{x_n\}$ è di Cauchy (poichè $x_m \in \overline{B(x_n, r_n)}$ per ogni $m \geq n$), e quindi converge a $\overline{x} \in E$. Ma $\overline{x} \in \overline{B(x_n, r_n)}$ per ogni n implica $\overline{x} \neq c_n$ per ogni n, contraddizione che dimostra il teorema. $\square$

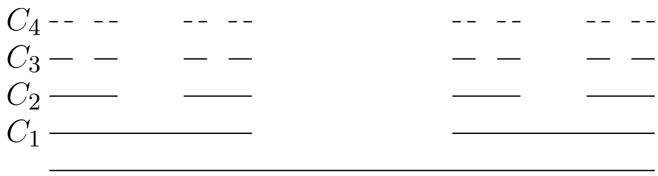

Figura 1. La costruzione dell'insieme di Cantor

Sono ovviamente perfetti gli intervalli chiusi di $\mathbb{R}$. Un altro esempio classico di insieme perfetto è dato dall'insieme di Cantor.

Esempio 1.3.12 (L'insieme di Cantor)**.** Consideriamo l'intervallo compatto $I_0 = [0,1]$. Rimuoviamo da questo intervallo il terzo centrale, cioè l'intervallo $\left(\frac{1}{3}, \frac{2}{3}\right)$. Otteniamo cosí l'insieme C_1, unione di 2^1 intervalli chiusi di lunghezza 3^{-1}. Ai passi successivi togliamo a ciascuno degli intervalli rimasti il terzo centrale. Abbiamo quindi che l'insieme C_n è l'unione di 2^n intervalli I_n^k, chiusi, disgiunti e di lunghezza $(1/3)^n$ (si veda la figura 1). L'insieme di Cantor è l'insieme

$$C = \bigcap_{n=1}^{\infty} C_n.$$

C è ovviamente limitato, e chiuso in quanto intersezione di chiusi. Ne segue che C è compatto.

Mostriamo ora che C è non numerabile. Osserviamo innanzi tutto che gli estremi di ciascun I_n^k appartengono a C. Se $x \in C$, $x \in C_n$ per ogni n, e quindi per ogni n x appartiene ad un intervallo $I_n^k \subset C_n$. È immediato allora verificare che x è il limite della successione formata dagli estremi inferiori (o superiori) di tali intervalli. Poiché tale successione è in C, ne deduciamo che ogni punto di C è il limite di una successione di punti di C. Quindi C è un insieme perfetto. La tesi segue allora dal teorema 1.3.11.

Una dimostrazione alternativa della non numerabilità di C si può fare costruendo direttamente una applicazione iniettiva da $[0,1]$ in C. Dato $x \in [0,1]$, costruiamo una successione $I_n^{k_n} \subset C_n$ di intervalli come segue:

(a) Poniamo $J_1^1 = [0,1/2]$ e $J_1^2 = (1/2,1]$. Definiamo $k_1 = 1$ se $x \in J_1^1$ e $k_1 = 2$ se $x \in J_1^2$. Quindi $I_1^{k_1}$ è il primo dei due intervalli che compongono C_1 se $x \in [0,1/2]$ e il secondo in caso contrario.

(b) Poniamo $J_1^{k_1} = J_2^1 \cup J_2^2$, dove J_2^i sono intervalli di lunghezza $2^{-2} = 1/4$, disgiunti e con J_2^1 chiuso a destra. Se $x \in J_2^1$ poniamo $I_2^{k_2}$ essere il primo dei due intervalli in $I_1^{k_1}$, in caso contrario il secondo.

(c) Proseguiamo in modo analogo, prendendo $I_n^{k_n}$ essere il primo dei due subintervalli che si ottengono da $I_{n-1}^{k_{n-1}}$ togliendo il terzo centrale se x sta nella prima metà di $J_{n-1}^{k_{n-1}}$ e il secondo in caso contrario.

È immediato vedere che $\{y\} = \bigcap_n I_n^{k_n} \subset C$. Quindi abbiamo definito un'applicazione $x \in [0,1] \mapsto y \in C$. Poiché la iniettività di tale applicazione è ovvia, la non numerabilità di C segue. Notiamo infine che gli elementi dell'insieme di Cantor si possono caratterizzare nel modo seguente:

$$C = \left\{ y \in [0,1] \;\middle|\; y = \sum_{n=1}^{\infty} \frac{c_n}{3^n} \text{ dove } c_n \in \{0,2\} \right\}$$

sono cioè i numeri reali la cui rappresentazione in base 3 non contiene la cifra 1. La corrispondenza fra C e $[0,1]$ si può allora descrivere (usando la rappresentazione in base 2 dei numeri in $[0,1]$) nel modo seguente

$$x = \sum_{n=1}^{\infty} \frac{c_n}{2^n} \mapsto y = \sum_{n=1}^{\infty} \frac{2c_n}{3^n}$$

dove $c_n \in \{0,1\}$. Si noti però che tale definizione presenta delle ambiguità: infatti vi sono numeri che hanno piú di una rappresentazione in base 2: $\frac{1}{2} = \sum_{n=2}^{\infty} \frac{1}{2^n}$.

Osservazione 1.3.13. Ogni chiuso di $\mathbb{R}$ si ottiene togliendo da $\mathbb{R}$ stesso una famiglia numerabile di intervalli aperti. Se E è perfetto, si ottiene da $\mathbb{R}$ togliendo una famiglia numerabile di intervalli in modo che nessuna coppia di tali intervalli abbia un estremo in comune.

1.4. Connessione

Ricordiamo qui la definizione e le principali proprietà degli insiemi connessi.

Definizione 1.4.1. Uno spazio topologico X si dice disconnesso se X può essere espresso come unione di due insiemi aperti, non vuoti e disgiunti. In caso contrario X si dice connesso.

Ricordiamo che i soli sottoinsiemi di $\mathbb{R}$ connessi (rispetto alla topologia ususale) sono gli intervalli (aperti, chiusi o semiaperti).

Vale il seguente teorema:

Teorema 1.4.2. *Sia X uno spazio topologico connesso, e $f\colon X \to Y$ un'applicazione continua.*

Allora $f(X) \subset Y$ è uno spazio topologico connesso.

Diamo altre caratterizzazioni della connessione:

Proposizione 1.4.3. *Sia X uno spazio topologico. Allora sono equivalenti le seguenti affermazioni*

(a) *X è connesso;*

(b) *X non può essere espresso come unione di due chiusi non vuoti e disgiunti;*

(c) *I soli sottoinsiemi di X aperti e chiusi sono X stesso e l'insieme vuoto $\varnothing$;*

(d) *Se $A \subset X$, A diverso da X e da $\varnothing$, allora $\partial A \neq \varnothing$;*

(e) *Consideriamo $Y = \{0, 1\}$ con la topologia discreta. Allora non vi è alcuna funzione $f\colon X \to Y$ continua e suriettiva.*

Vediamo ora il concetto di connessione per archi, e il suo legame con la connessione.

Definizione 1.4.4. Sia X uno spazio topologico. Un arco o un cammino in X è un'applicazione $\gamma\colon [0,1] \to X$ continua.

Uno spazio topologico X si dice connesso per archi se per ogni coppia di punti $x, y \in X$ esiste un arco γ tale che $\gamma(0) = x$, $\gamma(1) = y$.

Teorema 1.4.5. *Sia X uno spazio topologico connesso per archi. Allora X è connesso.*

Dimostrazione. Supponiamo, per assurdo che $X = U \cup V$, con U e V aperti, non vuoti e disgiunti. Prendiamo $x \in U$, $y \in V$ e consideriamo un cammino γ che parte da x e arriva in y. Allora, per il teorema 1.4.2, l'immagine $\gamma^* = \gamma([0,1])$ è connesso in X. Ma $\gamma^* = (\gamma^* \cap U) \cup (\gamma^* \cap V)$ con $(\gamma^* \cap U)$ e $(\gamma^* \cap V)$ aperti, non vuoti e disgiunti, contraddizione che termina la dimostrazione. □

La nozione di connessione per archi è equivalente a quella di connessione in $\mathbb{R}$, ma non in spazi topologici generali: si consideri, ad esempio, in $\mathbb{R}^2$ l'insieme

$$\Big\{ (x, y) \;\Big|\; y = \sin\frac{1}{x}, x > 0 \Big\} \cup \{(0,0)\}$$

che si può mostrare essere connesso ma non connesso per archi. Nonostante ciò, vi è un'importante classe di spazi topologici per cui le due nozioni coincidono: i sottoinsiemi aperti di $\mathbb{R}^n$. Vale infatti il seguente

Teorema 1.4.6. *Sia U un sottonsieme aperto e connesso di $\mathbb{R}^n$. Allora U è connesso per archi.*

Dimostrazione. Sia $x \in U$ e consideriamo l'insieme

$$A = \Big\{ a \in U \;\Big|\; \text{esiste un cammino } \gamma \text{ che parte da } x \text{ e arriva a } a \Big\}.$$

Poniamo $B = U \setminus A$. È facile mostrare che A è aperto: sia $a \in A$, e sia γ un cammino che connette x e a. Prendiamo una palla $B(a, r) \subset U$. Ogni punto $y \in B(a, r)$ può essere connesso a x seguendo prima il cammino γ e poi il raggio della palla da a fino a y. Mostriamo che A è anche chiuso: infatti, data una successione a_n di punti di A, $a_n \to a \in U$, essendo U aperto esiste una palla $B(a, r) \subset U$. Ma allora possiamo connettere a ad x seguendo il cammino γ che connette x ad $a_n \in B(a, r)$ e poi il raggio che connette a_n ad a. Essendo A chiuso, B è aperto. Ma $U = A \cup B$ con A e B aperti e U connesso implica che A o $B = \varnothing$. Poiché $x \in A$, $B = \varnothing$ e $U = A$, ovvero U è connesso per archi. $\square$

1.5. Funzioni

In generale se f è una funzione definita in A a valori in B, scriviamo $f\colon A \to B$. Se $x \in A$, scriviamo $x \mapsto f(x)$ per distinguere l'immagine di un singolo punto x dall'immagine dell'intero insieme A.

Sia f una funzione definita in $E \subset \mathbb{R}^n$ a valori reali e sia x_0 un punto di accumulazione per E. Poniamo

$$M(x_0, \delta) := \sup_{\substack{x \in B(x_0,\delta)\cap E \\ x \neq x_0}} f(x) \qquad m(x_0, \delta) := \inf_{\substack{x \in B(x_0,\delta)\cap E \\ x \neq x_0}} f(x).$$

M è una funzione crescente e m una funzione decrescente di δ. Poniamo

$$\limsup_{\substack{x \to x_0 \\ x \in E}} f(x) := \lim_{\delta \to 0} M(x_0, \delta) \qquad \liminf_{\substack{x \to x_0 \\ x \in E}} f(x) := \lim_{\delta \to 0} m(x_0, \delta).$$

Vale un teorema analogo a quello sui $\liminf$ e $\limsup$ di successioni

Teorema 1.5.1. *Sia $f\colon E \to \mathbb{R}$. Allora*

(a) $m = \limsup_{x \to x_0} f(x)$ *se e solo se: (i) esiste una successione $x_1, x_2, \dots \in E \setminus \{x_0\}$ convergente a x_0 tale che $\lim_k f(x_k) = m$; (ii) per ogni $m' > m$ esiste $\delta > 0$ tale che $f(x) < m'$ per tutti gli $x \in B(x_0, \delta) \cap E$, $x \neq x_0$;*

(b) $\ell = \liminf_{x \to x_0} f(x)$ *se e solo se: (i) esiste una successione $x_1, x_2, \dots \in E \setminus \{x_0\}$ convergente a x_0 tale che $\lim_k f(x_k) = \ell$; (ii) per ogni $\ell' < \ell$ esiste $\delta > 0$ tale che $f(x) > \ell'$ per tutti gli $x \in B(x_0, \delta) \cap E$, $x \neq x_0$;*

(c) $\lim_{x \to x_0} f(x) = a$ *se e solo se* $a = \liminf_{x \to x_0} f(x) = \limsup_{x \to x_0} f(x)$;

(d) $a = \limsup_{x \to x_0} f(x)$ *se e solo se: (i) esiste una successione $x_1, x_2, \dots \in E \setminus \{x_0\}$ convergente a x_0 tale che $\limsup_{k \to \infty} f(x_k) = a$; (ii) per ogni successione $x_1, x_2, \dots \in E \setminus \{x_0\}$ convergente a x_0 abbiamo che $\limsup_{k \to \infty} f(x_k) \leq a$;*

(e) $\ell = \liminf_{x \to x_0} f(x)$ *se e solo se: (i) esiste una successione $x_k \in E \setminus \{x_0\}$ convergente a x_0 tale che $\liminf_{k \to \infty} f(x_k) = \ell$; (ii) per ogni successione $x_k \in E \setminus \{x_0\}$ convergente a x_0 abbiamo che $\liminf_{k \to +\infty} f(x_k) \geq \ell$;*

Definizione 1.5.2. Sia X uno spazio topologico. Una funzione $f\colon X \to \mathbb{R}$ si dice semicontinua inferiormente in X se $\{f > a\} := f^{-1}((a, +\infty))$ è aperto (in X) per ogni $a \in \mathbb{R}$, e che f è semicontinua superiormente in X se e solo se $\{f < a\} := f^{-1}((-\infty, a))$ è aperto (in X) per ogni $a \in \mathbb{R}$.

Si può dimostrare che

Teorema 1.5.3. *Sia $E \subset \mathbb{R}^n$. Allora $f\colon E \to \mathbb{R}$ è semicontinua inferiormente in E se e solo se per ogni $x_0 \in E$*

$$f(x_0) \leq \liminf_{x \to x_0} f(x)$$

ed è semicontinua superiormente E se e solo se

$$f(x_0) \geq \limsup_{x \to x_0} f(x).$$

Alcune delle proprietà delle funzioni semicontinue:

Proposizione 1.5.4. *Sia X uno spazio topologico.*

(a) *f è semicontinua inferiormente se e solo se $g = -f$ è semicontinua superiormente;*

(b) *χ_E è semicontinua inferiormente se e solo se E è aperto ed è semicontinua superiormente se e solo se E è chiuso;*

(c) *Se è f semicontinua inferiormente allora αf è semicontinua inferiormente per ogni $\alpha \geq 0$;*

(d) *Se f e g sono semicontinue inferiormente allora anche $f + g$ è semicontinua inferiormente;*

(e) *Se f_α, $\alpha \in I$ è una famiglia di funzioni semicontinue inferiormente, allora $\sup_{\alpha \in I} f_\alpha$ è una funzione semicontinua inferiormente;*

(f) *Se f è semicontinua inferiormente e X è compatto, allora esiste $x_0 \in X$ tale che $f(x_0) = \inf_{x \in X} f(x)$.*

Introduciamo la notazione

$$C^k(\Omega)$$

per indicare l'insieme delle funzioni a valori complessi, definite nell'aperto $\Omega \subset \mathbb{R}^n$, che siano k volte differenziabili con continuità (cioè tali che le derivate parziali $\partial f / \partial x_{i_1} \dots \partial x_{i_k}$ esistono in tutti i punti $x \in \Omega$ e sono funzioni continue su Ω). Se una funzione f sta in $C^k(\Omega)$ per tutti i $k \in \mathbb{N}$, scriviamo $f \in C^\infty(\Omega)$.

Una classe importante di funzioni è quella delle funzioni caratteristiche di insiemi. Se A è un insieme, definiamo

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases} \tag{1.5.1}$$

Queste funzioni verranno utilizzate come blocchi fondamentali per costruire funzioni piú generali. Si noti che $\chi_A \chi_B = \chi_{A \cap B}$.

Il *supporto di una funzione continua* $f\colon \mathbb{R}^n \to \mathbb{C}$ si indica con $\operatorname{supp}\{f\}$ ed è definito come la chiusura dell'insieme dei punti $x \in \mathbb{R}^n$ in cui $f(x) \neq 0$, cioè

$$\operatorname{supp}\{f\} = \overline{\left\{ x \in \mathbb{R}^n \;\middle|\; f(x) \neq 0 \right\}}. \tag{1.5.2}$$

È importante ricordare che questa definizione è di tipo topologico. Piú avanti daremo la definizione di *supporto essenziale* di una funzione misurabile (si veda l'osservazione 4.1.9). Denotiamo l'insieme delle funzioni $C^\infty(\Omega)$ il cui supporto è limitato e contenuto in Ω con $C_c^\infty(\Omega)$. L'indice c sta per "compatto", poiché un sottoinsieme è chiuso e limitato se e solo se è compatto.

Il seguente è l'esempio classico di una funzione in $C_c^\infty(\mathbb{R}^n)$; il suo supporto è la palla unità $\left\{ x \in \mathbb{R}^n \;\middle|\; |x| \leq 1 \right\}$:

$$j(x) = \begin{cases} \exp\left[-\frac{1}{1-|x|^2}\right] & \text{se } |x| < 1 \\ 0 & \text{se } |x| \geq 1. \end{cases} \tag{1.5.3}$$

La verifica che j è effettivamente di classe $C^\infty(\mathbb{R}^n)$ è lasciata come esercizio.

1.6. Convessità

Ricordiamo alcune definizioni e risultati sulle funzioni convesse.

Definizione 1.6.1. Sia V uno spazio vettoriale, e $X \subset V$. Diciamo che X è un insieme convesso se per ogni coppia $x, y \in X$ e per ogni $\lambda \in [0,1]$ si ha che $\lambda x + (1-\lambda)y \in X$. Se $A \subset V$ è un generico sottoinsieme, l'inviluppo convesso di A è il piú piccolo (rispetto all'inclusione) insieme convesso che contiene A, ovvero l'intersezione di tutti i sottoinsiemi convessi di V che contengono A (è immediato dimostrare che una tale intersezione è un insieme convesso).

Una funzione $f\colon X \to \mathbb{R}$, dove X è un insieme convesso, si dice funzione convessa se

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \qquad \text{per ogni } \lambda \in [0,1]\,, x, y \in K.$$

Se la diseguaglianza è stretta per ogni $\lambda \in (0,1)$ e $x \neq y$, la funzione si dice strettamente convessa.

Osserviamo che i convessi in $\mathbb{R}$ sono gli intervalli. Alcune proprietà delle funzioni convesse di cui faremo uso nel seguito sono le seguenti.

Proposizione 1.6.2. *Sia* $f\colon I \subset \mathbb{R} \to \mathbb{R}$. *Allora*

(a) *Se* f *è convessa, allora per ogni* $x < z < y$ *si ha*

$$\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(z)}{y-z}. \tag{1.6.1}$$

(b) *Se per ogni $x < z < y$ vale*

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(z)}{y - z}, \tag{1.6.2}$$

allora f è convessa.

(c) *Se f è convessa, in ogni $x_0 \in \mathring{I}$ f ammette derivata destra $D^+ f(x_0)$ e sinistra $D^- f(x_0)$. Inoltre se $x_1 < x_2 \in \mathring{I}$, si ha che*

$$D^- f(x_1) \leq D^+ f(x_1) \leq D^- f(x_2) \leq D^+ f(x_2). \tag{1.6.3}$$

(d) *Se f è convessa, allora per ogni $x \in \mathring{I}$ vale*

$$f(y) \geq f(x) + \beta(y - x) \qquad \text{per ogni } y \in I \tag{1.6.4}$$

per ogni $\beta \in [D^- f(y), D^+ f(y)]$, e quindi esiste (almeno) una retta passante per $(x, f(x))$ che tocca il grafico della funzione solo in questo punto.

(e) *Se f è convessa, f è continua in ogni $x \in \mathring{I}$.*

(f) *Se f è derivabile in (a, b), allora è convessa se e solo se la sua derivata è una funzione monotona crescente.*

Dimostrazione. Se $x < z < y$ allora $z = \lambda x + (1 - \lambda) y$ ove

$$\lambda = \frac{y - z}{y - x} \qquad 1 - \lambda = \frac{z - x}{y - x}. \tag{1.6.5}$$

(a) Dalla diseguaglianza di convessità $f(z) \leq \lambda f(x) + (1 - \lambda) f(y)$, si ha

$$\begin{aligned} &\lambda(f(y) - f(x)) \leq f(y) - f(z) \\ &f(z) - f(x) \leq (1 - \lambda)(f(y) - f(x)) \end{aligned}$$

e, utilizzando le espressioni per λ e $1 - \lambda$ si ottiene subito la (1.6.1).

(b) Fissiamo x, y e $\lambda \in (0, 1)$. Poniamo $z = \lambda x + (1 - \lambda) y$, in modo che continui a valere la (1.6.5). Dalla (1.6.2) segue allora che

$$\begin{aligned} f(z) - f(x) &\leq \frac{z - x}{y - z}(f(y) - f(z)) \\ &= \frac{1 - \lambda}{\lambda}(f(y) - f(z)) \end{aligned}$$

da cui la diseguaglianza di convessità

$$f(\lambda x + (1 - \lambda) y) = f(z) \leq \lambda f(x) + (1 - \lambda) f(y)$$

segue immediatamente.

(c) Segue immediatamente dal punto **(a)** che, i rapporti incrementali $\frac{f(x) - f(x_0)}{x - x_0}$ dipendono in modo monotono non decrescente dalla variabile x.

Quindi è immediato dedurre l'esistenza delle derivate destra e sinistra, cosí come la (1.6.3). Sempre dalla monotonia, segue anche che

$$D^- f(x_0) = \sup_{x<x_0} \frac{f(x) - f(x_0)}{x - x_0} \qquad D^+ f(x_0) = \inf_{x>x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1.6.6}$$

(d) Dalla (1.6.6) segue immediatamente la tesi.

(e) Dalla (1.6.4) segue subito che

$$\begin{aligned} f(x) &\geq f(y) + \alpha(x - y) \\ f(y) &\geq f(x) + \beta(y - x) \end{aligned}$$

con $\alpha \in [D^- f(y), D^+ f(y)]$, $\beta \in [D^- f(x), D^+ f(x)]$, da cui

$$f(y) + \alpha(x - y) \leq f(x) \leq f(y) + \beta(x - y).$$

Possiamo ora calcolare il limite per $x \to y$, osservando che α non dipende da x, mentre β, che dipende da x, è una quantità limitata al variare di $x \in [y - \delta, y + \delta]$ (ricordiamo la (1.6.3)). Ne deduciamo subito la continuità di f in $y \in \mathring{I}$.

(f) f è derivabile se e solo se $D^- f(x) = D^+ f(x)$. Dalla (1.6.3) segue allora che le funzioni convesse e derivabili hanno derivata monotona crescente. Nel caso in cui f è derivabile, allora abbiamo, per ogni terna $x < z < y$,

$$\frac{f(z) - f(x)}{z - x} = f'(s), \qquad \frac{f(y) - f(z)}{y - z} = f'(t),$$

con $s \in (x, z)$, $t \in (z, y)$. La tesi segue allora dalla monotonia di f' e dalla **(b)**. □

Osservazione 1.6.3. La funzione $f\colon [0, 1] \to \mathbb{R}$ definita come

$$f(x) = \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1 \end{cases}$$

è convessa ma non continua e mostra che la tesi della Proposizione 1.6.2**(e)** non può essere in generale rafforzata.

1.7. Compattezza

Ricordiamo alcuni teoremi sugli insiemi compatti in spazi topologici e metrici.

Definizione 1.7.1. Se X è un insieme, un ricoprimento di X è una famiglia $\mathcal{F}$ di insiemi tali che $X \subset \cup_{A\in\mathcal{F}} A$.

Un sottoricoprimento del ricoprimento $\mathcal{F}$ di X è un ricoprimento $\mathcal{F}_1$ di X tale che $A \in \mathcal{F}_1$ implica $A \in \mathcal{F}$.

Un ricoprimento aperto è un ricoprimento $\mathcal{F}$ formato da insiemi aperti.

L'insieme X è compatto se da ogni ricoprimento aperto di X possiamo estrarre un sottoricoprimento finito.

Una famiglia $\mathcal{F}$ di insiemi si dice famiglia centrata se l'intersezione di ogni sua parte finita è non vuota.

Teorema 1.7.2. *Sia X uno spazio topologico. Sono equivalenti:*

(a) *X è compatto;*

(b) *Ogni famiglia centrata di chiusi in X ha intersezione non vuota.*

(a) implica (b). Sia $\mathcal{F} = \{F\}$ una famiglia centrata di chiusi in X, e sia X compatto. Allora $\mathcal{G} = \left\{ G \mid G = F^c,\ F \in \mathcal{F} \right\}$ è una famiglia di aperti in X. Se $\bigcap_{F\in\mathcal{F}} F = \varnothing$, abbiamo che $\bigcup_{G\in\mathcal{G}} G = \bigcup_{F\in\mathcal{F}} F^c = \left(\bigcap_{F\in\mathcal{F}} F\right)^c = \varnothing^c = X$, cioè $\mathcal{G}$ è un ricoprimento aperto del compatto X. Ne deduciamo che esiste un sottoricoprimento finito $\{G_1, \dots, G_\ell\}$. Ma allora $\cap_{i=1}^{\ell} F_i = \cap_{i=1}^{\ell} G_i^c = X^c = \varnothing$, contraddizione. □

(b) implica (a). Sia $\mathcal{G} = \{G\}$ un ricoprimento aperto di X. Allora, posto $\mathcal{F} = \left\{ F \mid F = G^c,\ G \in \mathcal{G} \right\}$, abbiamo che $\bigcap_{F\in\mathcal{F}} F = \varnothing$ e quindi $\mathcal{F}$ è una famiglia di chiusi che non può essere centrata (se no avrebbe intersezione non vuota). Quindi esiste una parte finita $\{F_1, \dots, F_\ell\}$ di $\mathcal{F}$ con intersezione vuota. Ma ciò implica che la famiglia $G_1 = F_1^c, \dots, G_\ell = F_\ell^c$ è un sottoricoprimento del ricoprimento aperto $\mathcal{G}$ di X. □

Corollario 1.7.3. *Sia X uno spazio topologico compatto, $F \subset X$ chiuso. Allora F è compatto.*

Dimostrazione. Sia $\mathcal{F} = \{F_\alpha\}$ una famiglia centrata di chiusi in F. Ma se F_α è chiuso in F, e F è chiuso in X, allora F_α è chiuso in X. Ne segue che la famiglia $\mathcal{F}$ è una famiglia centrata di chiusi in X. La tesi segue allora dalla compattezza di X. □

Corollario 1.7.4. *Sia X uno spazio topologico compatto. Allora ogni sottoinsieme infinito di X ha un punto di accumulazione.*

Dimostrazione. Sia $A \subset X$ un insieme infinito. In particolare A contiene un sottoinsieme numerabile $A_1 = \{x_1, x_2, \dots\}$. Supponiamo che A, e quindi anche A_1, sia senza punti di accumulazione. Allora la famiglia $A_n = \{x_n, x_{n+1}, \dots\}$ è una famiglia centrata di chiusi (gli A_n sono chiusi poiché non hanno punti di accumulazione). Segue allora che $\bigcap_n A_n \neq \varnothing$, contraddizione. □

Ricordiamo che uno spazio topologico X si dice di *Hausdorff* se, dati x e $y \in X$ con $x \neq y$ esistono due aperti U e V disgiunti tali che $x \in U$, $y \in V$.

Teorema 1.7.5. *Sia X uno spazio topologico di Hausdorff, e K un compatto in X. Allora K è chiuso.*

Dimostrazione. Mostriamo che il complementare è aperto. Sia $x_0 \in X \setminus K$. Per ogni $x \in K$ possiamo trovare una coppia di aperti disgiunti U_x e V_x tali che $x \in U_x$, $x_0 \in V_x$. Poiché $K \subset \bigcup_{x \in K}(U_x \cap K)$ e $U_x \cap K$ è aperto in K, segue dalla compattezza di K che esistono $x_1, \dots, x_n \in K$ tali che $K \subset (U_{x_1} \cup \dots \cup U_{x_n})$. Ma allora $V_{x_1} \cap \dots \cap V_{x_n}$ è un aperto che contiene x_0 e che non interseca K. Ne segue che x_0 è interno al complementare di K. Dall'arbitrarietà di x_0 segue che il complementare di K è aperto. □

Teorema 1.7.6. *Sia X uno spazio topologico di Hausdorff, e K_α una famiglia centrata di compatti in X. Allora $\bigcap_\alpha K_\alpha \neq \varnothing$.*

Dimostrazione. Scegliamo un elemento K_{α_0} della famiglia. Per il teorema precedente è (un compatto) chiuso in X. Quindi la famiglia $\{K_\alpha \cap K_{\alpha_0}\}$ è una famiglia centrata di chiusi nel compatto K_α. Ne segue che

$$\varnothing \neq \bigcap_\alpha (K_\alpha \cap K_{\alpha_0}) = \bigcap_\alpha K_\alpha.$$

□

Passiamo ora ad esaminare la situazione nel caso di spazi metrici. Ricordiamo le seguenti definizioni

Definizione 1.7.7. Sia X uno spazio metrico, con distanza $\mathrm{dist}(x, y)$. Diciamo che X è uno spazio metrico completo se ogni successione di Cauchy è convergente. Diciamo che X è totalmente limitato se per ogni $\epsilon > 0$ esiste un ricoprimento finito di X mediante palle aperte di raggio ϵ.

Teorema 1.7.8. *Sia X uno spazio metrico. Sono equivalenti:*

(a) *X è compatto;*

(b) *Ogni sottoinsieme infinito di X ha un punto di accumulazione in X;*

(c) *Ogni successione in X ammette una sottosuccessione convergente ad un punto $x \in X$;*

(d) *X è totalmente limitato e completo.*

(a) implica (b). Si veda il teorema 1.7.4. □

(b) implica (c). Sia $x_1, x_2, \dots$ una successione in X. Se $x_1, x_2, \dots$ ha solo un numero finito di termini distinti, possiamo costruire una sottosuccessione a termini costanti, e l'implicazione segue. Se l'insieme $\{x_n\}$ è infinito ha per **(b)** un punto di accumulazione $x \in X$. Sia $x_{n_1} = x_1$. Supponiamo di aver scelto $x_{n_1}, \dots, x_{n_k}$. Scegliamo $x_{n_{k+1}} \in B(x, \frac{1}{k+1})$ in modo tale che $n_{k+1} > n_k$. □

(c) implica (d). È ovvio che X è completo. Supponiamo non sia totalmente limitato. Allora esiste ϵ_0 tale che X non può essere ricoperto con un numero finito di palle di raggio ϵ_0. Prendiamo $x_1 \in X$ arbitrariamente. Poiché $B(x_1, \epsilon_0)$ non ricopre X, possiamo trovare $x_2 \in X \setminus B(x_1, \epsilon_0)$. Ma anche $B(x_1, \epsilon_0) \cup B(x_2, \epsilon_0)$ non ricopre X. Continuando cosí otteniamo una successione $x_n \in X$ tale che $\operatorname{dist}(x_i, x_j) \geq \epsilon_0$ per ogni $i \neq j$. Ma allora tale successione non ammette sottosuccessioni convergenti. □

(d) implica (b). Prendiamo un sottoinsieme infinito di X. Sia $\{x_n\}$ una sua parte numerabile. X è ricoperto da un numero finito di palle di raggio 1. Almeno una di esse, che chiameremo B_1, contiene infiniti elementi $\{x_n^1\}$ della successione $x_1, x_2, \ldots$. Scegliamo poi un ricoprimento di $\overline{B}_1$ con palle di raggio 1/2. Almeno una di esse, che chiameremo B_2, contiene infiniti elementi $\{x_1^2\}$ della successione $\{x_n^1\}$. Costruiamo cosí una successione B_n di palle di raggio 2^{-n+1}. Sia, per ogni n, A_n una palla chiusa con centro uguale a quello di B_n e raggio doppio. Abbiamo che $A_1, A_2, \ldots$ è una successione decrescente di chiusi il cui diametro tende a zero. Quindi la successione dei centri è di Cauchy e, essendo X completo, converge a $x_0 = \bigcap A_n$. Ne segue che x_0 è un punto di accumulazione per la successione $x_1, x_2, \ldots$. □

(d) implica (a). Osserviamo innanzi tutto che dalla **(d)** segue che esiste una base numerabile di intorni per X. Infatti per ogni n X è ricoperto da un numero finito, che indicheremo con $N(n)$ di palle di centro x_i^n e raggio $1/n$. Ne segue che l'insieme $\left\{ x_i^n \;\middle|\; n \in \mathbb{N},\ 1 \leq i \leq N(n) \right\}$ è numerabile e denso in X e che la famiglia $\left\{ B(x_i^n, 1/k) \;\middle|\; n, k \in \mathbb{N},\ 1 \leq i \leq N(n) \right\}$ è una base numerabile per gli intorni di X.

Sia ora $X \subset \bigcup_\alpha G_\alpha$. Per ogni α esistono i, n, k tali che $B(x_i^n, 1/k) \subset G_\alpha$. Consideriamo allora la famiglia delle palle della forma $\mathcal{B} = \left\{ B(x_i^n, 1/k) \;\middle|\; B(x_i^n, 1/k) \subset G_\alpha \text{ per qualche } \alpha \right\}$. È ovviamente numerabile, indichiamola con $\mathcal{B} = \left\{ B_j \;\middle|\; j \in \mathbb{N} \right\}$. Per ciascuna palla B_j in $\mathcal{B}$ scegliamo poi un G_{α_j} che la contiene. Otteniamo cosí una famiglia numerabile $\{G_{\alpha_j}\}$ tale che $X \subset \bigcup_{j \in \mathbb{N}} G_{\alpha_j}$. Abbiamo quindi mostrato che è sufficiente mostrare che ogni ricoprimento *numerabile* di X ammette un sottoricoprimento finito.

Supponiamo allora per assurdo che dal ricoprimento $\mathcal{G} = \left\{ G_{\alpha_k} \;\middle|\; k \in \mathbb{N} \right\}$ non sia possibile estrarre alcun sottoricoprimento finito. Sia $F_k = G_{\alpha_k}^c$, e consideriamo $\Phi_n = \bigcap_{k=1}^n F_k$. Poiché i G_{α_k} formano un ricoprimento di X, abbiamo che $\Phi_n \searrow \varnothing$. Gli insiemi Φ_n sono una successione decrescente di chiusi non vuoti (se $\Phi_n = \varnothing$, abbiamo che $X = \left(\bigcap_{k=1}^n F_k\right)^c = \bigcup_{k=1}^n G_{\alpha_k}$, contraddizione). Due casi sono possibili: (i) $\Phi_n = \Phi_{n+1}$ per tutti gli $n \geq n_0$. In tal caso $\bigcap_{n=1}^\infty \Phi_n = \Phi_{n_0} \neq \varnothing$, contraddizione. (ii) per infiniti n $\Phi_n \supsetneq \Phi_{n+1}$. Possiamo allora definire una successione $x_1, x_2, \ldots$ tale che $x_n \in \Phi_n \setminus \Phi_{n+1}$ per gli n per cui tale differenza è non vuota. Tale successione avrà almeno un punto di accumulazione x_0 (abbiamo già mostrato che **(d)** implica **(b)**). Ma $x_0 \in \Phi_n$ per ogni n poiché Φ_n è chiuso e $\{x_n, x_{n+1}, \ldots\} \subset \Phi_n$. Quindi $x_0 \in \bigcap_n \Phi_n$, di nuovo una contraddizione. $\square$

Un'immediato corollario del teorema precedente è la ben nota caratterizzazione dei compatti di $\mathbb{R}^n$:

Teorema 1.7.9. *Sono equivalenti:*

(a) *$E \subset \mathbb{R}^n$ è compatto;*

(b) *ogni successione in E ammette una sottosuccessione convergente;*

(c) *E è chiuso e limitato;*

(d) *Ogni famiglia centrata di chiusi in E ha intersezione non vuota.*

Infatti ogni insieme limitato di $\mathbb{R}^n$ è totalmente limitato, e i chiusi di $\mathbb{R}^n$ sono completi (proprietà vera per ogni sottoinsieme chiuso di uno spazio metrico completo).

Un'altro concetto di cui faremo uso in seguito è il seguente:

Definizione 1.7.10. Sia X uno spazio topologico. L'insieme $E \subset X$ si dice relativamente compatto se la sua chiusura $\overline{E}$ è compatta.

Segue immediatamente dalla definizione e dal teorema 1.7.8 che

Teorema 1.7.11. *Sia X uno spazio metrico completo, e $E \subset X$. Sono equivalenti:*

(a) *E è relativamente compatto;*

(b) *E è totalmente limitato;*

(c) *Ogni successione in E ammette una sottosuccessione di Cauchy.*

Esercizi

(1) Si dimostri che $\limsup_k E_k$ è l'insieme dei punti che appartengono a infiniti E_k, mentre $\liminf_k E_k$ è l'insieme dei punti che appartengono a tutti gli E_k tranne al piú un numero finito.

(2) Si dimostri il teorema 1.2.1.

(3) Dimostrare che $\liminf_n a_n + \liminf_n b_n \leq \liminf_n (a_n + b_n) \leq \liminf_n a_n + \limsup_n b_n$. Si dia un esempio in cui le diseguaglianza sono strette. Dimostrare poi che $\liminf_k (a_k + b_k) = \lim_k a_k + \liminf_k b_k$ nel caso in cui il limite di a_k esista e le somme abbiano senso.

(4) Dimostrare che $\liminf_k a_k = -\limsup_k (-a_k)$.

(5) Si mostri che ogni aperto di $\mathbb{R}^n$ è un $\mathcal{F}_\sigma$ e che ogni chiuso è un $\mathcal{G}_\delta$. Si mostri anche che ogni intervallo (chiuso, aperto o semiaperto) è sia un $\mathcal{F}_\sigma$ che un $\mathcal{G}_\delta$.

(6) Sia X uno spazio metrico con distanza dist. Si mostri che $\overline{B(x,\delta)} \subset \{\, y \in X \mid \operatorname{dist}(x,y) \le \delta \,\}$. Si mostri con un esempio che tale inclusione può essere stretta.

(7) Si mostri che la definizione di volume per un parallelepipedo coincide con quella di un rettangolo nel caso in cui il parallelepipedo sia un rettangolo.

(8) Si dimostri che i connessi di $\mathbb{R}$ sono tutti e solo gli intervalli.

(9) Si dimostri il teorema 1.4.2.

(10) Si dimostri la proposizione 1.4.3.

(11) Si dimostri il teorema 1.5.1.

(12) Sia $f\colon \mathbb{R} \to \mathbb{R}$ una funzione continua. Si mostri che per ogni valore $c \in \big[\liminf_{x\to+\infty} f(x), \limsup_{x\to+\infty} f(x)\big]$ esiste una successione $x_k \to +\infty$ tale che $\lim_k f(x_k) = c$.

(13) Si dimostri il teorema 1.5.3.

(14) Sia $f\colon E \to (-\infty,+\infty]$ una funzione semicontinua inferiormente nell'insieme compatto E. Mostrare che f è inferiormente limitata e che assume il minimo in E.

(15) Si dimostri che la funzione $j(x)$ definita nella (1.5.3) è infinitamente differenziabile.

Capitolo 2

Prime nozioni di teoria della misura

2.1. Introduzione

Il primo incontro con l'integrazione è nei primi corsi di analisi dove l'integrale viene definito come integrale secondo Riemann. Tale concetto, che ha una motivazione storica ed è sufficiente per affrontare molti problemi matematici, non è adeguato per le necessità dell'analisi moderna. La difficoltà che presenta l'integrale di Riemann è che la classe di funzioni per cui può essere definito non è chiusa per passaggio al limite puntuale di funzioni in questa classe, nemmeno nel caso in cui la successione sia monotona. La limitazione di dover trattare con una teoria dell'integrazione che non permette di passare al limite è simile a quella di dover fare della matematica con i soli numeri razionali.

Se pensiamo al grafico di una funzione reale in n variabili, l'integrale di una tale funzione è il volume $(n+1)$-dimensionale sotteso dal grafico. Il problema è di come definire un tale volume. L'integrale di Riemann cerca di definirlo come "base per altezza" di parallelepipedi che hanno come base dei piccoli cubi n-dimensionali e altezza un valore "ragionevole" della funzione al variare della variabile indipendente nel cubo. La difficoltà risiede nel fatto che può essere impossibile definire propriamente una tale altezza se la funzione è molto discontinua.

L'idea innovatrice e feconda di risultati di Lebesgue è stata quella di calcolare il volume $(n+1)$-dimensionale "nella direzione opposta", cioè calcolando prima il volume n-dimensionale dell'insieme dove la funzione è piú grande di un dato numero y. Questa volume è una buona funzione monotona noncrescente del numero y che può essere integrata secondo Riemann.

Per poter attuare questo programma è però necessario definire in modo opportuno "il volume n-dimensionale dell'insieme dove la funzione è piú grande di un dato numero y", e piú in generale il volume dei sottoinsiemi di $\mathbb{R}^n$ (o perlomeno di una classe abbastanza ampia di tali sottoinsiemi). Ciò richiede una buona nozione di *volume* per una classe (la piú ampia possibile) di sottoinsiemi di $\mathbb{R}^n$.

Vedremo in seguito che la proprietà principale per una tale nozione di *volume* è quella della numerabile additività. Per motivare tale proprietà ricordiamo l'essenza della teoria dell'integrazione. Dato uno spazio X e una famiglia $\mathcal{B}$ di sottoinsiemi di X per cui è definita una nozione ragionevole di misura, cioè un'applicazione $A \in \mathcal{B} \mapsto \mu(A) \in [0, +\infty]$, per calcolare l'integrale di una funzione $f\colon X \to \mathbb{R}$ procediamo scegliendo una qualche partizione $\mathcal{P}$ di X in sottoinsiemi $A \in \mathcal{B}$. Data $\mathcal{P}$, scegliamo poi per ciascun $A \in \mathcal{P}$ un valore a_A *tipico* di f su A. Formiamo allora le somme

$$\sum_{A\in\mathcal{P}} a_A \mu(A). \tag{2.1.1}$$

Infine, usando eventualmente un procedimento di limite, si eliminano le ambiguità scegliendo partizioni $\mathcal{P}$ in modo che le restrizioni di f a ciascun A siano sempre piú vicine a essere delle costanti.

Per dare senso alla (2.1.1), anche ignorando i problemi di convergenza, dobbiamo ovviamente restringerci a considerare partizioni finite o al piú numerabili di X, e il procedimento di limite sarà essenziale. Nel caso particolare in cui X sia numerabile, e che $\{x\} \in \mathcal{B}$ per ogni $x \in X$, vi è un modo *ovvio* di evitare il procedimento di limite: basta prendere $\mathcal{P} = \bigl\{ \{x\} \bigm| x \in X \bigr\}$ e definire l'integrale come (non ci occupiamo di problemi di convergenza-notiamo però che questa definizione è ben posta ogni volta che $f(x) \geq 0$ su tutto X)

$$\sum_{x\in X} f(x)\mu(\{x\}). \tag{2.1.2}$$

Questa è l'idea seguita da Riemann nella sua teoria dell'integrazione. Ma quello di Riemann non è l'unico modo *ovvio* in cui procedere per definire l'integrale nel caso in cui X sia numerabile. Infatti, se $f\colon X \to \mathbb{R}$, con X numerabile, allora anche $\mathrm{Range}(f)$ è numerabile e, assumendo che $A(a) := \bigl\{ x \in X \bigm| f(x) = a \bigr\} \in \mathcal{B}$ per ogni $a \in \mathbb{R}$, Lebesgue affermerebbe che un altro candidato ovvio per l'integrale è

$$\sum_{a\in\mathrm{Range}(f)} a\mu(A(a)). \tag{2.1.3}$$

Affinché queste due definizioni ovvie coincidano è necessario esaminare l'applicazione $A \in \mathcal{B} \mapsto \mu(A) \in [0, +\infty]$. Infatti, anche se X è numerabile e $\mathcal{B}$ contiene tutti i sottoinsiemi di X, le formule (2.1.2) e (2.1.3) danno lo stesso risultato solo se uno sa che, per ogni collezione numerabile $A_1, A_2, \ldots \subset \mathcal{B}$,

$$\mu(\bigcup_n A_n) = \sum_n \mu(A_n) \qquad \text{se } A_n \cap A_m = \varnothing \text{ per } m \neq n. \tag{2.1.4}$$

Nel caso in cui X sia numerabile, la (2.1.4) equivale ad assumere che

$$\mu(A) = \sum_{x \in A} \mu(\{x\}), \qquad A \subset X.$$

Ma nel caso in cui X non sia numerabile o finito, la (2.1.4) non è affatto banale. Si deve a Lebesgue aver compreso che tale nozione è comunque fondamentale e di aver mostrato che esistono misure non banali che godono di una tale proprietà.

2.2. Nozioni base di teoria della misura

Prima di cercare di definire una misura di un insieme bisogna studiare la struttura degli insiemi misurabili, cioè degli insiemi a cui si potrà associare un valore numerico in un modo non ambiguo. Non necessariamente tutti gli insiemi saranno misurabili.

Consideriamo un insieme X i cui elementi saranno chiamati *punti*. Possiamo pensare ad X come ad un sottoinsieme di $\mathbb{R}^n$, ma potrebbe essere un insieme molto piú generale, ad esempio l'insieme dei cammini in un spazio su cui cerchiamo di definire un "integrale funzionale".

Definizione 2.2.1. Un collezione Σ di sottoinsiemi di X è chiamata una sigma-algebra se sono soddisfatti i seguenti assiomi:

(a) Se $A \in \Sigma$, allora anche $A^c \in \Sigma$.

(b) Se $A_1, A_2, \ldots$ è una famiglia numerabile di insiemi in Σ, allora la loro unione $\bigcup_{i=1}^{\infty} A_i$ è pure in Σ.

(c) $X \in \Sigma$.

Si osservi che queste ipotesi implicano che l'insieme vuoto $\varnothing$ sta in Σ, che Σ è chiuso anche per intersezioni numerabili, cioè $A_1, A_2, \ldots \in \Sigma$ implica $\bigcap_{i=1}^{\infty} A_i \in \Sigma$ e che $A_1 \setminus A_2 \in \Sigma$.

È ovvio che ogni famiglia $\mathcal{F}$ di sottoinsiemi di X può essere estesa ad una sigma-algebra (si prenda, ad esempio, la sigma-algebra di tutti i sottoinsiemi di X). Ma tra tutte le estensioni ve ne è una speciale. Consideriamo tutte le sigma-algebre che contengono $\mathcal{F}$ e facciamone l'intersezione, che chiamiamo Σ. Quindi un sottoinsieme $A \subset X$ è in Σ se sta in tutte le sigma-algebre che contengono $\mathcal{F}$. È facile vedere che Σ è una sigma-algebra. Infatti è la piú piccola sigma-algebra che contiene $\mathcal{F}$, detta anche la sigma-algebra generata da $\mathcal{F}$. Un esempio importante è fornito dalla sigma-algebra $\mathcal{B}_n = \mathcal{B}(\mathbb{R}^n)$ dei Boreliani in $\mathbb{R}^n$ che è generata dagli aperti di $\mathbb{R}^n$. Alternativamente $\mathcal{B}(\mathbb{R}^n)$ può essere definita come la sigma-algebra generata dalle palle di $\mathbb{R}^n$, cioè dagli insiemi della forma

$$B_{x,R} = \left\{\, y \in \mathbb{R}^n \;\middle|\; |x-y| < R \,\right\}.$$

Segue da **(a)** che i chiusi sono Boreliani. Usando l'assioma della scelta è possibile mostrare che $\mathcal{B}(\mathbb{R}^n)$ non contiene tutti i sottoinsiemi di $\mathbb{R}^n$ (si veda l'osservazione 4.1.6), ma vogliamo enfatizzare che non faremo mai uso di ciò nel seguito. Piú in generale indicheremo con $\mathcal{B}(X)$ la sigma-algebra dei Boreliani nello spazio topologico X, cioè la sigma-algebra generata dagli aperti in X.

Definizione 2.2.2. Una misura (a volte chiamata misura positiva) μ definita su di una sigma-algebra Σ è una funzione da Σ in $[0,+\infty]$ tale che $\mu(\varnothing) = 0$ e che soddisfi alla seguente proprietà di additività numerabile (detta anche di sigma-additività). Se $A_1, A_2, \dots$ è una successione di insiemi *disgiunti* di Σ, allora

$$\mu\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} \mu(A_i). \tag{2.2.1}$$

Osservazione 2.2.3. Si noti che la somma della serie che appare nella definizione di sigma-additività è sempre definita (eventualmente vale $+\infty$), e che sarebbe contrario alla nostra intuizione di misura richiede l'additività piú che numerabile. Infatti ne seguirebbe, supponendo che ciascun punto di $x \in \mathbb{R}$ abbia misura $\mu(\{x\}) = 0$, che $\mu([0,1]) = \sum_{x\in[0,1]} \mu(\{x\}) = 0$.

Storicamente il grosso passo avanti è stato quello di rendersi conto che l'additività numerabile è un requisito fondamentale. È infatti facile costruire misure finitamente additive (in cui la (2.2.1) vale con $+\infty$ rimpiazzato da un qualunque intero finito), ma non è possibile sviluppare in tal modo una teoria dell'integrazione soddisfacente. Poiché $\mu(\varnothing) = 0$, la (2.2.1) include, come caso particolare, l'additività finita. Altre conseguenze importanti della (2.2.1) sono:

Proposizione 2.2.4. *Data una misura μ definita in una sigma-algebra Σ in un insieme X, valgono le seguenti proprietà:*

(a) $\mu(\bigcup_{k=1}^N A_k) = \sum_{k=1}^N \mu(A_k)$ *se $A_1, \dots, A_N$ è una famiglia finita di insiemi disgiunti;*

(b) $\mu(A) \le \mu(B)$ *se $A \subset B$;*

(c) $\lim_{j\to+\infty} \mu(A_j) = \mu\Big(\bigcup_{i=1}^{+\infty} A_i\Big)$ *se $A_1 \subset A_2 \subset A_3 \subset \dots;$*

(d) $\lim_{j\to+\infty}\mu(A_j) = \mu\Big(\bigcap_{i=1}^{+\infty} A_i\Big)$ *se* $A_1 \supset A_2 \supset A_3 \supset \dots$ *e almeno uno degli insiemi* A_j *ha misura finita.*

Dimostrazione. (a): segue immediatamente dall'additività numerabile e dal fatto che $\mu(\varnothing) = 0$;

(b): $B = A \cup (B \setminus A)$ implica che $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

(c): $\bigcup_k A_k = A_1 \cup \bigcup_k (A_{k+1} \setminus A_k)$ implica che

$$\mu(\bigcup_k A_k) = \mu(A_1) + \sum_k \mu(A_{k+1} \setminus A_k)$$
$$= \lim_{n\to\infty}\big(\mu(A_1) + \sum_{k=1}^{n}\mu(A_{k+1}\setminus A_k)\big) = \lim_{n\to\infty}\mu(A_{n+1}).$$

(d): Possiamo supporre $\mu(A_1) < +\infty$. Abbiamo che $A_1 = \bigcap_k A_k \cup \bigcup_k (A_k \setminus A_{k+1})$, e quindi

$$\mu(A_1) = \mu(\bigcap_k A_k) + \sum_{k=1}^{\infty}\mu(A_k \setminus A_{k+1})$$
$$= \mu(\bigcap_k A_k) + \lim_{n\to\infty}\sum_{k=1}^{n}\mu(A_k \setminus A_{k+1})$$
$$= \mu(\bigcap_k A_k) + \lim_{n\to\infty}\mu(A_1 \setminus A_{n+1}).$$

Da cui si ricava che

$$\mu(\bigcap_k A_k) = \lim_{n\to\infty}\big(\mu(A_1) - \mu(A_1 \setminus A_{n+1})\big) = \lim_{n\to\infty}\mu(A_{n+1}).$$

□

Uno spazio di misura è quindi composto da tre oggetti: un insieme X, una sigma-algebra Σ e una misura μ. Se $X = \mathbb{R}^n$ (o, piú in generale, se X contiene degli aperti, cosí che sia possibile definire la sigma-algebra dei boreliani $\mathcal{B}(X)$) e se $\Sigma \supset \mathcal{B}(X)$, allora μ si chiama una misura di Borel. Gli elementi di Σ si chiamano insiemi misurabili. Si noti anche che ogni qualvolta X' è un sottoinsieme misurabile di X possiamo definire il sottospazio di misura (X', Σ', μ) in cui Σ' è l'insieme dei sottoinsiemi misurabili di X'. Questo sottospazio di misura è chiamato la restrizione di μ a X'.

Un esempio semplice ma importante è la misura delta di Dirac δ_y centrata in un punto $y \in \mathbb{R}^n$ arbitrario ma fissato di $\mathbb{R}^n$:

$$\delta_y(A) = \begin{cases} 1 & \text{se } y \in A \\ 0 & \text{se } y \notin A. \end{cases} \tag{2.2.2}$$

In altre parole, usando la definizione delle funzioni caratteristiche (1.5.1)

$$\delta_y(A) = \chi_A(y). \tag{2.2.3}$$

Qui come sigma-algebra si può prendere sia la sigma algebra $\mathcal{B}(\mathbb{R}^n)$ dei Boreliani che la sigma-algebra di tutti i sottoinsiemi di $\mathbb{R}^n$.

Una misura μ si dice una misura sigma-finita se esiste una successione $A_1, A_2, \ldots \subset X$ di insiemi misurabili tali che (i) $X = \bigcup_n A_n$ e (ii) $\mu(A_n) < +\infty$ per ogni $n \in \mathbb{N}$.

2.3. Classi monotone

Il prossimo teorema 2.3.3 è importante per la costruzione delle misure prodotto e per la dimostrazione del teorema di Fubini 10.2.1.

Definizione 2.3.1. Una classe monotona $\mathcal{M}$ è una collezione di insiemi con due proprietà:

(1) se $A_i \in \mathcal{M}$ per $i = 1, 2, \ldots$, e se $A_1 \subset A_2 \subset \cdots$, allora $\bigcup_i A_i \in \mathcal{M}$;
(2) se $A_i \in \mathcal{M}$ per $i = 1, 2, \ldots$, e se $A_1 \supset A_2 \supset \cdots$, allora $\bigcap_i A_i \in \mathcal{M}$.

Una collezione di insiemi, $\mathcal{A}$, forma un'algebra di insiemi se per ogni A e B in $\mathcal{A}$ le differenze $A \setminus B$, $B \setminus A$ e le unioni $A \cup B$ sono in $\mathcal{A}$.

Osservazione 2.3.2. Ovviamente ogni sigma-algebra è una classe monotona, e l'insieme delle parti di un insieme X è una classe monotona. Quindi ogni collezione di sottoinsiemi di X è contenuta in una classe monotona (l'intersezione di tutte le classi monotone che contengono la collezione).

Si noti che il passaggio da un'algebra $\mathcal{A}$ che contiene X ad una sigma-algebra si riduce ad aggiungere le unioni numerabili di sottoinsiemi di $\mathcal{A}$, creando una collezione $\mathcal{A}_1$ di insiemi che non è piú chiusa per intersezioni. Al passo successivo aggiungiamo le intersezioni numerabili di sottoinsiemi di $\mathcal{A}_1$. Otteniamo una nuova collezione $\mathcal{A}_2$ che non è chiusa per unioni. Procedendo cosí si arriva ad una sigma-algebra mediante un'"induzione transfinita". Il teorema seguente permette di evitare ciò e semplicemente afferma che le sigma-algebre sono "limiti" monotoni di algebre.

Teorema 2.3.3 (Teorema sulle classi monotone)**.** *Sia X un insieme e sia $\mathcal{A}$ un'algebra di insiemi di X tale che X stesso appartiene ad $\mathcal{A}$.*

Allora esiste la piú piccola classe monotona $\mathcal{S}$ che contiene $\mathcal{A}$. Tale classe $\mathcal{S}$ è anche la piú piccola sigma-algebra che contiene $\mathcal{A}$.

Dimostrazione. Sia $\mathcal{S}$ l'intersezione di tutte le classi monotone che contengono $\mathcal{A}$, cioè tali che $Y \in \mathcal{S}$ se e solo se Y sta in tutte le classi monotone che contengono $\mathcal{A}$. Lasciamo come esercizio al lettore la verifica che $\mathcal{S}$ è una classe monotona che contiene $\mathcal{A}$. Per definizione è allora la piú piccola classe monotona che contiene $\mathcal{A}$.

Notiamo che è sufficiente mostrare che $\mathcal{S}$ è chiusa per complementazione e unioni finite. Infatti, supponendo ciò vero per il momento, abbiamo che, se $A_1, A_2, \dots$ stanno in $\mathcal{S}$, allora $B_n := \bigcup_{i=1}^n A_i$ è una successione crescente di insiemi in $\mathcal{S}$. Poiché $\mathcal{S}$ è una classe monotona $\bigcup_{n=1}^\infty B_n = \bigcup_{i=1}^\infty A_i$ sta in $\mathcal{S}$. Quindi $\mathcal{S}$ è chiusa per unioni numerabili. Quindi $\mathcal{S}$ è una sigma-algebra e poiché ogni sigma-algebra è una classe monotona, $\mathcal{S}$ è la piú piccola sigma-algebra che contiene $\mathcal{A}$.

Mostriamo ora che $\mathcal{S}$ è chiusa per unioni finite. Fissiamo $A \in \mathcal{A}$ e consideriamo la collezione $\mathcal{C}(A) = \{ B \in \mathcal{S} \mid B \cup A \in \mathcal{S} \}$. Poiché $\mathcal{A}$ è un'algebra, $\mathcal{C}(A)$ contiene $\mathcal{A}$. Per ogni successione crescente di insiemi B_n in $\mathcal{C}(A)$, $A \cup B_n$ è una successione crescente di insiemi in $\mathcal{S}$. Poiché $\mathcal{S}$ è una classe monotona,

$$A \cup \Big(\bigcup_{i=1}^\infty B_i \Big) = \bigcup_{i=1}^\infty A \cup B_i$$

sta in $\mathcal{S}$ e quindi $\bigcup_{i=1}^\infty B_i$ sta in $\mathcal{C}(A)$. Il lettore può verificare che $\mathcal{C}(A)$ è chiusa per intersezioni di successioni numerabili e decrescenti di insiemi, e quindi concludere che $\mathcal{C}(A)$ è una classe monotona che contiene $\mathcal{A}$. Poiché sta in $\mathcal{S}$ e $\mathcal{S}$ è la piú piccola classe monotona che contiene $\mathcal{A}$, $\mathcal{C}(A) = \mathcal{S}$.

Prendiamo ora un insieme arbitrario A in $\mathcal{S}$, e consideriamo la collezione $\mathcal{C}(A) = \{ B \in \mathcal{S} \mid B \cup A \in \mathcal{S} \}$. Dagli argomenti precedenti sappiamo che $\mathcal{A}$ è un sottoinsieme di $\mathcal{C}(A)$. Ripetendo il ragionamento fatto prima di vede che anche in questo caso $\mathcal{C}(A)$ è una classe monotona e quindi di nuovo $\mathcal{C}(A) = \mathcal{S}$. Quindi $\mathcal{S}$ è chiuso per unioni finite, come affermato.

Infine, consideriamo il comportamento per complementazione. Sia $\mathcal{C} = \{ B \in \mathcal{S} \mid B^c \in \mathcal{S} \}$. Questo insieme contiene $\mathcal{A}$ in quanto $\mathcal{A}$ è un'algebra che contiene X. Data una successione crescente di insiemi $B_i \in \mathcal{C}$, $i = 1, 2, \dots$, B_i^c è una successione decrescente di insiemi in $\mathcal{S}$. Poiché $\mathcal{S}$ è una classe monotona,

$$\Big(\bigcup_{i=1}^\infty B_i \Big)^c = \bigcap_{i=1}^\infty B_i^c$$

sta in $\mathcal{S}$. In modo simile data una successione decrescente di insiemi $B_i \in \mathcal{C}$, $i = 1, 2, \dots$, B_i^c è una successione crescente di insiemi in $\mathcal{S}$ e quindi

$$\Big(\bigcap_{i=1}^\infty B_i \Big)^c = \bigcup_{i=1}^\infty B_i^c$$

sta in $\mathcal{S}$ e $\mathcal{C} = \mathcal{S}$.

Quindi $\mathcal{S}$ è chiuso per unioni finite e complementazione. $\square$

2.4. Unicità di misure

Come applicazione del teorema sulle classi monotone presentiamo un teorema di unicità per misure. Illustra un modo tipico di utilizzare il teorema sulle classi monotone e sarà utile nella sezione 10.1 sulle misure prodotto.

Teorema 2.4.1 (Unicità delle misure)**.** *Sia X un insieme, $\mathcal{A}$ un'algebra di sottoinsiemi di X contenente X stesso e Σ la piú piccola sigma-algebra che* *contiene $\mathcal{A}$. Sia μ_1 una misura sigma-finita nel senso piú forte seguente: esiste una successione di insiemi $A_i \in \mathcal{A}$ (e non semplicemente $A_i \in \Sigma$), $i = 1, 2, \dots$, ciascuno di misura μ_1 finita, tali che $\bigcup_{i=1}^{\infty} A_i = X$. Se μ_2 è una misura che coincide con μ_1 su $\mathcal{A}$, allora $\mu_1 = \mu_2$ su tutto Σ.*

Dimostrazione. Dall'ipotesi di sigma-finitezza segue che $X = \bigcup_i A_i$ dove possiamo assumere, oltre a (i) $A_i \in \mathcal{A}$, (ii) $\mu(A_i) < +\infty$ anche che (iii) $A_i \cap A_j = \varnothing$ per ogni $i \neq j$ (infatti basta eventualmente considerare la nuova famiglia $\tilde{A}_1 = A_1, \tilde{A}_2 = A_2 \setminus A_1, \tilde{A}_3 = A_3 \setminus (A_1 \cup A_2), \dots$).

Consideriamo la collezione di insiemi

$$\mathcal{M} = \left\{ B \in \Sigma \;\middle|\; \mu_1(B \cap A_i) = \mu_2(B \cap A_i) \text{ per ogni } i \in \mathbb{N} \right\}.$$

Ovviamente $\mathcal{M}$ contiene l'algebra $\mathcal{A}$. Mostriamo che è una classe monotona. Sia $B_1 \subset B_2 \subset \dots$ una successione crescente di insiemi in $\mathcal{M}$. Poiché, per ogni i fissato, $B_n \cap A_i \subset B_{n+1} \cap A_i$, anche la successione di insiemi $B_1 \cap A_i, B_2 \cap A_i, \dots$ è crescente. Poiché sia μ_1 che μ_2 sono misure, ne segue che $\lim_{n\to+\infty} \mu_1(B_n \cap A_i) = \mu_1(A_i \cap \bigcup B_n)$ e $\lim_{n\to+\infty} \mu_2(B_n \cap A_i) = \mu_2(A_i \cap \bigcup B_n)$, e quindi $\mu_1(A_i \cap \bigcup B_n) = \mu_2(A_i \cap \bigcup B_n)$. Ne deduciamo che $\bigcup B_n \in \mathcal{M}$. Prendendo poi una successione $B_1 \supset B_2 \supset \dots$, si deduce dal fatto che ciascun A_i ha misura finita che $\lim_{n\to+\infty} \mu_1(B_n \cap A_i) = \mu_1(A_i \cap \bigcap B_n)$ e $\lim_{n\to+\infty} \mu_2(B_n \cap A_i) = \mu_2(A_i \cap \bigcap B_n)$. Ne segue che $\mu_1(A_i \cap \bigcap B_n) = \mu_2(A_i \cap \bigcap B_n)$ per tutti gli $i \in \mathbb{N}$, e quindi $\bigcap B_n \in \mathcal{M}$. Quindi $\mathcal{M}$ è una classe monotona, che quindi coincide con Σ per il teorema 2.3.3. Quindi, per ogni $B \in \Sigma$ si ha che $\mu_1(B \cap A_i) = \mu_2(B \cap A_i)$ per tutti gli i. Ne deduciamo che

$$\mu_1(B) = \sum_{i=1}^{\infty} \mu_1(B \cap A_i) = \sum_{i=1}^{\infty} \mu_2(B \cap A_i) = \mu_2(B).$$

$\square$

Esercizi Capitolo 2

(1) Sia X un insieme, e $\mathcal{F}$ una classe di sottoinsiemi di X. Si dimostri che l'intersezione di tutte le sigma-algebre che contengono $\mathcal{F}$ è una sigma-algebra.

(2) Può una sigma algebra avere un'infinità numerabile di elementi?

(3) Mostrare con un esempio che la **(d)** della proposizione 2.2.4 può essere falsa se $\mu(A_k) = +\infty$ per tutti i k.

(4) Sia X un insieme qualunque. Mostrare che la funzione che ad ogni sottoinsieme $A \subset X$ associa la sua cardinalità $\#(A)$, definita sulla sigma-algebra dell'insieme delle parti di X è una misura.

(5) Sia X un insieme non numerabile e Σ la famiglia dei sottoinsiemi $A \subset X$ tali che A o A^c è al piú numerabile. Si ponga $\mu(A) = 0$ se A è al piú

numerabile e $\mu(A) = 1$ se A non è numerabile. Si dimostri che Σ è una sigma-algebra e che μ è una misura.

(6) Dimostrare che la σ-algebra $\mathcal{B}_1$ dei boreliani in $\mathbb{R}$ è la piú piccola σ-algebra generata da
  (a) gli intervalli aperti;
  (b) le semirette della forma $(-\infty, a)$, $a \in \mathbb{R}$;
  (c) le semirette della forma $(a, +\infty)$, $a \in \mathbb{R}$;
  (d) gli intervalli della forma $(a, b]$.

(7) Sia (X, Σ, μ) uno spazio di misura, e siano E_1 ed E_2 misurabili. Mostrare che $\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2) - \mu(E_1 \cap E_2)$.

(8) Sia (X, Σ, μ) uno spazio di misura, e $E_1, E_2, \ldots$ una successione di insiemi misurabili. Si dimostri che
  (a) $\mu(\liminf_k E_k) \leq \liminf_k \mu(E_k)$;
  (b) se esiste $k_0 \in \mathbb{N}$ tale che $\mu(\bigcup_{k_0}^{\infty} E_k) < +\infty$ allora

$$\limsup_{k\to\infty} \mu(E_k) \leq \mu(\limsup_k E_k).$$

(9) Sia (X, Σ, μ) uno spazio di misura, e $E_1, E_2, \ldots$ una successione di insiemi misurabili. Diremo che $E_k \to E$ se

$$E = \liminf_{k\to+\infty} E_k = \limsup_{k\to+\infty} E_k$$

Si dimostri che se $E_k \to E$ e se esiste $k_0 \in \mathbb{N}$ tale che $\mu(\bigcup_{k_0}^{\infty} E_k) < +\infty$ allora $\mu(E_k) \to \mu(E)$.

(10) Sia (X, Σ, μ) uno spazio di misura, e $E_1, E_2, \ldots$ una successione di insiemi misurabili. Supponiamo che $\sum_k \mu(E_k) < +\infty$. Si dimostri che

$$\mu(\limsup_k E_k) = 0.$$

(11) Sia (X, Σ, μ) uno spazio di misura, e $E_1, E_2, \ldots$ una successione di insiemi misurabili tali che $\mu(E_k \cap E_j) = 0$. Mostrare che

$$\mu(\bigcup_k E_k) = \sum_k \mu(E_k).$$

(12) Dare un esempio di una classe monotona che non sia una σ-algebra.

Capitolo 3

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

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

In questa e nelle prossime sezioni costruiremo la misura di Lebesgue in $\mathbb{R}^n$, il piú importante esempio di misura con cui avremo a che fare nel seguito, e nei capitoli successivi svilupperemo la teoria dell'integrazione. La costruzione non è semplice, ma ha la proprietà di dare in modo corretto il volume euclideo degli insiemi "ben fatti". Prima di addentrarci nella sua costruzione, enunciamo qui le sue proprietà piú importanti, con l'intento di rendere piú chiare le costruzioni successive.

La misura di Lebesgue è ovviamente una *misura*, cioè una funzione *numerabilmente additiva*, definita su una sigma-algebra Σ. Gli elementi di Σ sono i sottoinsiemi *misurabili* (secondo Lebesgue). La misura di Lebesgue (o volume) di un insieme misurabile si indica con $\mathcal{L}^n(E)$ o con il simbolo

$$|E| := \mathcal{L}^n(E).$$

La sigma-algebra Σ contiene la sigma-algebra $\mathcal{B}(\mathbb{R}^n)$ dei Boreliani, ma è strettamente piú grande di questa. Facendo uso dell'assioma della scelta si può mostrare che, necessariamente, esistono insiemi non misurabili. La misura di Lebesgue è l'*unica* misura in $\mathbb{R}^n$ che gode delle seguenti proprietà:

Invarianza per traslazione. Per ogni fissato $y \in \mathbb{R}^n$ si ha $\mathcal{L}^n(E) = \mathcal{L}^n(E+y) = \mathcal{L}^n(\{\, x+y \mid x \in E \,\})$;

Normalizzazione. $\mathcal{L}^n([0,1]^n) = 1$.

Segue da queste due proprietà che tale misura generalizza l'usuale concetto di area e di volume[1].

Sempre utilizzando l'assioma della scelta si dimostra che non possiamo misurare tutti i sottoinsiemi di $\mathbb{R}^n$, se $n > 1$, anche se ci accontentiamo di una misura invariante per rototraslazioni, normalizzata e solo *finitamente* additiva.

Il caso di $\mathbb{R}$ fa eccezione, ed esiste una misura (diversa da quella di Lebesgue) solo *finitamente additiva*, normalizzata, invariante per traslazioni e definita per *tutti* i sottoinsiemi di $\mathbb{R}$.

Abbandonando invece l'assioma della scelta si può mostrare l'esistenza di un modello di matematica in cui tutti i sottoinsiemi di $\mathbb{R}^n$ sono misurabili.

La misura di Lebesgue di un palla è

$$\mathcal{L}^n(B_{x,R}) = |B_{0,1}|r^n = \frac{2\pi^{n/2}r^n}{n\Gamma(n/2)} = \frac{1}{n}|\mathbb{S}^{n-1}|r^n, \tag{3.1.1}$$

dove

$$|\mathbb{S}^{n-1}| = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

è l'area di $\mathbb{S}^{n-1}$, la sfera di raggio 1 in $\mathbb{R}^n$.

La misura di Lebesgue gode di due proprietà importanti (si veda il teorema 3.4.3), la regolarità interna:

$$\mathcal{L}^n(E) = \sup\big\{\,\mathcal{L}^n(C)\;\big|\; C \subset E \text{ e } C \text{ compatto}\,\big\}, \qquad \forall E \in \Sigma \tag{3.1.2}$$

e la regolarità esterna:

$$\mathcal{L}^n(E) = \inf\big\{\,\mathcal{L}^n(O)\;\big|\; E \subset O \text{ e } O \text{ aperto}\,\big\}, \qquad \forall E \in \Sigma \tag{3.1.3}$$

L'importanza di tali proprietà risiede nel fatto che legano misura e topologia.

Un'altra proprietà importante della misura di Lebesgue è il fatto che è *sigma-finita*. Una misura (X, Σ, μ) è sigma-finita se esiste una successione numerabile di insiemi $E_1, E_2, \dots$ tale che $\mu(E_i) < +\infty$ per ogni $i = 1, 2, \dots$ e tale che $X = \bigcup_{i=1}^{+\infty} E_i$. Se vale la sigma-finitezza è facile vedere che gli E_i si possono prendere disgiunti. Nel caso della misura di Lebesgue $\mathcal{L}^n$ si possono prendere come E_i i cubi di lato unitario con i vertici a coordinate intere.

[1] L'approccio elementare alla definizione di area di un poligono è basato sui concetti di congruenza (ovvero uguaglianza a meno di rototraslazioni) e di equiscomponibilità. Ricordiamo che due poligoni si dicono *equiscomponibili*, se si possono scomporre in parti poligonali a due a due congruenti. Segue allora dalle proprietà della misura di Lebesgue che due poligoni equiscomponibili hanno la stessa misura. Si può anche dimostrare che due *poligoni* di uguale misura sono equiscomponibili. Per una discussione di questo approccio si veda il libro [4].

3.2. La misura esterna in $\mathbb{R}^n$

Cominciamo ora a costruire la misura di Lebesgue in $\mathbb{R}^n$. Per fare ciò costruiamo innanzi tutto una misura esterna, cioè

Definizione 3.2.1. Una funzione $\mu_e \colon \mathcal{P}(X) \to [0, +\infty]$ si dice misura esterna se soddisfa alla seguente proprietà

Subadditività numerabile. Se $E_1, E_2, \ldots$ è una successione numerabile di sottoinsiemi di X, si ha che

$$\mu_e\Big(\bigcup_{i=1}^{\infty} E_i\Big) \le \sum_{i=1}^{+\infty} \mu_e(E_i).$$

La misura di Lebesgue sarà ottenuta restringendo il dominio di definizione di una misura esterna ad un'opportuna sottoclasse di insiemi $\Sigma \subset \mathcal{P}(\mathbb{R}^n)$. Vedremo che tale restrizione $\mathcal{L}^n \colon \Sigma \to [0, +\infty]$ sarà una misura normalizzata e invariante per rototraslazioni.

Definizione 3.2.2. Sia $E \subset \mathbb{R}^n$. La misura esterna di Lebesgue di E è il numero reale non negativo

$$\mathcal{L}^n_e(E) = \inf\Big\{ \sum_{i=1}^{+\infty} \mathrm{vol}(R_i) \;\Big|\; E \subset \bigcup_{i=1}^{+\infty} R_i,\ R_i \text{ rettangolo} \Big\}. \tag{3.2.1}$$

Si osservi che la famiglia di rettangoli R_i è numerabile.

Osservazione 3.2.3. E' chiaro che, per ogni $E \subset \mathbb{R}^n$, $0 \le \mathcal{L}^n_e(E) \le +\infty$, in quanto la serie è a termini positivi.

Tale definizione è consistente, in quanto

Teorema 3.2.4. *Per ogni rettangolo R*

$$\mathcal{L}^n_e(R) = \mathrm{vol}(R).$$

Dimostrazione. Poiché R ricopre se stesso, $\mathcal{L}^n_e(R) \le \mathrm{vol}(R)$.

Dimostriamo ora l'altra diseguaglianza. Sia R_i una successione (numerabile) di rettangoli che ricopre R. Consideriamo la famiglia di aperti $\mathring{R}_i$. Dato $\epsilon > 0$, consideriamo, per ogni i, un rettangolo S_i tale che (a) $R_i \subset \mathring{S}_i$ e (b) $\mathrm{vol}\,(S_i) \le \mathrm{vol}(R_i) + \epsilon 2^{-i}$. La famiglia $\mathring{S}_i$ di aperti ricopre il rettangolo R. Poiché R è compatto, esiste esiste k_0 tale che $R \subset \bigcup_{i=1}^{k_0} \mathring{S}_i \subset \bigcup_{i=1}^{k_0} S_i$. Usando il lemma 1.3.9 otteniamo allora che $\mathrm{vol}(R) \le \sum_{i=1}^{k_0} \mathrm{vol}(S_i) \le \sum_{i=1}^{k_0} \mathrm{vol}(R_i) + \epsilon \le \sum_{i=1}^{\infty} \mathrm{vol}(R_i) + \epsilon$, da cui segue, prendendo l'estremo inferiore su tutti i ricoprimenti, $\mathrm{vol}(R) \le \mathcal{L}^n_e(R) + \epsilon$. Il teorema segue dall'arbitrarietà di ϵ. $\square$

Vediamo ora le proprietà principali della misura esterna:

Teorema 3.2.5 (Monotonia)**.** *Se $E_1 \subset E_2$ allora $\mathcal{L}_e^n(E_1) \leq \mathcal{L}_e^n(E_2)$.*

Teorema 3.2.6 (Subadditività numerabile)**.** *Se $E = \bigcup E_k$ è l'unione di una famiglia numerabile di insiemi, allora $\mathcal{L}_e^n(E) \leq \sum \mathcal{L}_e^n(E_k)$.*

Dimostrazione. Possiamo assumere $\mathcal{L}_e^n(E_k) < +\infty$ per ogni k (in caso contrario la dimostrazione è ovvia). Sia $\epsilon > 0$. Per ogni k esiste una famiglia R_i^k di rettangoli tale che

$$E_k \subset \bigcup_{i=1}^{\infty} R_i^k \qquad \text{e} \qquad \sum_{i=1}^{\infty} \operatorname{vol}(R_i^k) < \mathcal{L}_e^n(E_k) + \tfrac{\epsilon}{2^k}.$$

Poiché $E \subset \bigcup_{i,k} R_i^k$, ne segue che $\mathcal{L}_e^n(E) \leq \sum_{i,k} \operatorname{vol}(R_i^k) = \sum_k \sum_i \operatorname{vol}(R_i^k)$. Quindi

$$\mathcal{L}_e^n(E) \leq \sum_k (\mathcal{L}_e^n(E_k) + \tfrac{\epsilon}{2^k}) = \sum_k \mathcal{L}_e^n(E_k) + \epsilon$$

e il risultato segue dall'arbitrarietà di ϵ. $\square$

Segue immediatamente dalla definizione e dai teoremi precedenti che

Proposizione 3.2.7. *Valgono le seguenti proprietà:*

(a) $\mathcal{L}_e^n(\{x\}) = 0$;

(b) *Sia $E = \{x_n\}_{n=1}^{\infty} \subset \mathbb{R}^n$ un sottoinsieme numerabile. Allora*

$$\mathcal{L}_e^n(E) = 0.$$

(c) $\mathcal{L}_e^n(\partial R) = 0$ *se R è un rettangolo.*

Esempio 3.2.8. I sottoinsiemi numerabili di $\mathbb{R}^n$ sono di misura nulla. In particolare $\mathbb{Q}$ è denso in $\mathbb{R}$ e di misura nulla. Ma vi sono anche insiemi non numerabili di misura nulla. L'esempio canonico è dato dall'insieme di Cantor (si veda l'esempio 1.3.12). Ricordando che $C = \bigcap_k C_k$, e che C_k è l'unione di 2^k intervalli chiusi di lunghezza 3^{-k}, abbiamo che $\mathcal{L}_e^1(C) \leq (2/3)^k \to 0$.

Mostriamo ora che insiemi arbitrari sono contenuti in insiemi "semplici" di misura esterna vicina.

Teorema 3.2.9. *Sia $E \subset \mathbb{R}^n$. Allora*

(a) *Per ogni $\epsilon > 0$ esiste un insieme G aperto tale che $E \subset G$ e $\mathcal{L}_e^n(G) \leq \mathcal{L}_e^n(E) + \epsilon$. Quindi*

$$\mathcal{L}_e^n(E) = \inf\{\, \mathcal{L}_e^n(G) \mid E \subset G, G \text{ aperto} \,\}.$$

(b) *Esiste un insieme H di tipo $\mathcal{G}_\delta$ tale che $E \subset H$ e $\mathcal{L}_e^n(E) = \mathcal{L}_e^n(H)$.*

Dimostrazione. (a): Sia $\epsilon > 0$ assegnato. Scegliamo dei rettangoli R_k in modo che $E \subset \bigcup_k R_k$ e che $\sum_k \mathrm{vol}(R_k) \leq \mathcal{L}_e^n(E) + \epsilon/2$. Per ogni k consideriamo un altro rettangolo S_k che contenga R_k nel suo interno e tale che $\mathrm{vol}(S_k) \leq \mathrm{vol}(R_k) + 2^{-k-1}\epsilon$. Ponendo $G = \bigcup_k \mathring{S}_k$, si ha che G è un aperto contenente E di misura esterna $\mathcal{L}_e^n(G) \leq \sum_k \mathrm{vol}(S_k) \leq \sum_k \mathrm{vol}(R_k) + \epsilon/2 \leq \mathcal{L}_e^n(E) + \epsilon$.

(b): Per **(a)** abbiamo che, per ogni k, esiste un aperto G_k in cui E è contenuto e tale che $\mathcal{L}_e^n(G_k) \leq \mathcal{L}_e^n(E) + 1/k$. Posto $H = \bigcap_k G_k$, abbiamo che H è un $\mathcal{G}_\delta$, che $E \subset H$ e che $\mathcal{L}_e^n(E) \leq \mathcal{L}_e^n(H) \leq \mathcal{L}_e^n(G_k) \leq \mathcal{L}_e^n(E) + 1/k$. Il risultato segue dall'arbitrarietà di k. □

Abbiamo visto che la misura esterna è numerabilmente subadditiva. Vedremo piú avanti che non è numerabilmente additiva. In alcuni casi, però vale l'additività:

Teorema 3.2.10. *Siano $E_1, E_2 \subset \mathbb{R}^n$ due insiemi tali che* $\mathrm{dist}(E_1, E_2) > 0$. *Allora*

$$\mathcal{L}_e^n(E_1 \cup E_2) = \mathcal{L}_e^n(E_1) + \mathcal{L}_e^n(E_2).$$

Dimostrazione. La diseguaglianza $\mathcal{L}_e^n(E_1 \cup E_2) \leq \mathcal{L}_e^n(E_1) + \mathcal{L}_e^n(E_2)$ segue immediatamente dalla subadditività.

Dimostriamo $\mathcal{L}_e^n(E_1) + \mathcal{L}_e^n(E_2) \leq \mathcal{L}_e^n(E_1 \cup E_2)$. Consideriamo una successione di rettangoli $\{R_k\}$ tale che $E_1 \cup E_2 \subset \bigcup_k R_k$ e $\sum_k \mathrm{vol}(R_k) \leq \mathcal{L}_e^n(E_1 \cup E_2) + \epsilon$ e tale che ciascun rettangolo R_k abbia diametro $\mathrm{diam}(R_k) < \frac{1}{2}\,\mathrm{dist}(E_1, E_2)$ (si può sempre fare, suddividendo eventualmente i rettangoli di diametro grande, senza cambiare il valore della serie). Poniamo $\mathcal{N} = \{\, k \mid R_k \cap E_1 \neq \varnothing \,\}$ e $\mathcal{M} = \{\, k \mid R_k \cap E_2 \neq \varnothing \,\}$. Ne segue che $\mathcal{N} \cap \mathcal{M} = \varnothing$, che $E_1 \subset \bigcup_{k \in \mathcal{N}} R_k$ e che $E_2 \subset \bigcup_{k \in \mathcal{M}} R_k$. Allora

$$\begin{aligned} \mathcal{L}_e^n(E_1) + \mathcal{L}_e^n(E_2) &\leq \sum_{k \in \mathcal{N}} \mathrm{vol}(R_k) + \sum_{k \in \mathcal{M}} \mathrm{vol}(R_k) \\ &\leq \sum_{k \in \mathbb{N}} \mathrm{vol}(R_k) \leq \mathcal{L}_e^n(E_1 \cup E_2) + \epsilon \end{aligned}$$

da cui il risultato segue. □

Usando il teorema appena dimostrato, possiamo provare il seguente fatto, che utilizzeremo nel seguito.

Teorema 3.2.11. *Sia $\{R_k\}$ una famiglia finita di rettangoli non sovrapposti. Allora*

$$\mathcal{L}_e^n\Big(\bigcup_{k=1}^N R_k\Big) = \sum_{k=1}^N \mathrm{vol}(R_k) = \sum_{k=1}^N \mathcal{L}_e^n(R_k).$$

Dimostrazione. Sia $\epsilon > 0$. Per ogni k tale che $\mathring{R}_k \neq \varnothing$ (indicheremo con $\mathcal{M}$ l'insieme di tali k) consideriamo $S_k \subset \mathring{R}_k$ tale che $\mathrm{vol}(R_k) \leq \mathrm{vol}(S_k) + \epsilon/N$. Allora abbiamo che i rettangoli $S_1, S_2, \ldots$ sono compatti e a due a due disgiunti. Ne segue che $\mathrm{dist}(S_k, S_j) > 0$ e possiamo utilizzare il teorema 3.2.10 per dedurre che

$$
\begin{aligned}
\mathcal{L}^n_e\Big(\bigcup_{k\in\mathcal{M}} S_k\Big) &\leq \mathcal{L}^n_e\Big(\bigcup_{k=1}^{N} R_k\Big) \leq \sum_{k=1}^{N} \mathcal{L}^n_e(R_k) \\
&= \sum_{k\in\mathcal{M}} \mathcal{L}^n_e(R_k) \leq \sum_{k\in\mathcal{M}} \mathcal{L}^n_e(S_k) + \epsilon \\
&= \mathcal{L}^n_e\Big(\bigcup_{k\in\mathcal{M}} S_k\Big) + \epsilon.
\end{aligned}
$$

Da cui segue la tesi per l'arbitrarietà di ϵ. □

Potrebbe sembrare, dalla nostra definizione, che la misura esterna possa dipendere dalla scelta degli assi coordinati (abbiamo usato, nella definizione, rettangoli con i lati paralleli agli assi). Ciò non accade, ma rimandiamo alla sezione 3.6 la dimostrazione di tale fatto.

3.3. Insiemi misurabili secondo Lebesgue

Definiamo ora la classe degli insiemi misurabili secondo Lebesgue, cioè la classe di insiemi sulla quale la misura esterna di Lebesgue risulta essere una misura. Ovviamente, una volta definita, dovremo mostrare che tale classe è una sigma-algebra.

Definizione 3.3.1. Un sottoinsieme $E \subset \mathbb{R}^n$ si dice misurabile secondo Lebesgue, o semplicemente misurabile se per ogni $\epsilon > 0$ esiste un insieme aperto G tale che (i) $E \subset G$; (ii) $\mathcal{L}^n_e(G \setminus E) < \epsilon$.

Se E è un insieme misurabile, la sua misura di Lebesgue, o semplicemente la sua misura, è il numero reale nonnegativo $\mathcal{L}^n(E) := \mathcal{L}^n_e(E)$.

Osservazione 3.3.2. Non si confonda la condizione $\mathcal{L}^n_e(G \setminus E) < \epsilon$ con l'affermazione dimostrata nel 3.2.9. In effetti, poiché $G = E \cup (G \setminus E)$, abbiamo solo che $\mathcal{L}^n_e(G) \leq \mathcal{L}^n_e(E) + \mathcal{L}^n_e(G \setminus E)$, e non possiamo concludere da $\mathcal{L}^n_e(G) \leq \mathcal{L}^n_e(E) + \epsilon$ che $\mathcal{L}^n_e(G \setminus E) < \epsilon$.

Segue immediatamente dalla definizione

Proposizione 3.3.3. *Sia $E \subset \mathbb{R}^n$ un insieme aperto. Allora E è misurabile.*

Proposizione 3.3.4. *Sia $E \subset \mathbb{R}^n$ un insieme di misura esterna nulla. Allora E è misurabile.*

Dimostrazione. Sappiamo che, per ogni $\epsilon > 0$, esiste $G \supset E$ tale che $\mathcal{L}^n_e(G) \leq \mathcal{L}^n_e(E) + \epsilon = \epsilon$. Allora $\mathcal{L}^n_e(G \setminus E) \leq \mathcal{L}^n_e(G) \leq \epsilon$. $\square$

Teorema 3.3.5. *L'unione $E = \bigcup_k E_k$ di una famiglia numerabile $\{E_k\}$ di insiemi misurabili è misurabile.*

Dimostrazione. Sia $\epsilon > 0$. Per ogni k esiste un aperto G_k che contiene E_k tale che $\mathcal{L}^n_e(G_k \setminus E_k) \leq \epsilon 2^{-k}$. Allora $G = \bigcup_k G_k$ è aperto, contiene E e si ha

$$\mathcal{L}^n_e(G \setminus E) \leq \mathcal{L}^n_e(\bigcup_k (G_k \setminus E_k)) \leq \sum_k \mathcal{L}^n_e(G_k \setminus E_k) \leq \epsilon.$$

$\square$

Teorema 3.3.6. *Sia R un rettangolo. Allora R è misurabile.*

Dimostrazione. Basta osservare che $R = \mathring{R} \cup \partial R$. $\mathring{R}$ è misurabile in quanto aperto, ∂R è misurabile in quanto di misura nulla, e quindi R è misurabile in quanto unione di misurabili. $\square$

Possiamo ora dimostrare che anche i chiusi sono misurabili

Teorema 3.3.7. *Sia $E \subset \mathbb{R}^n$ chiuso. Allora E è misurabile.*

Dimostrazione. Supponiamo per il momento E compatto, e sia $\epsilon > 0$ assegnato. Sappiamo che esiste un aperto G tale che $E \subset G$ e $\mathcal{L}^n_e(G) \leq \mathcal{L}^n_e(E) + \epsilon$. L'insieme $G \setminus E$ è aperto, e quindi esiste una famiglia numerabile di cubi chiusi non sovrapposti Q_k tale che $G \setminus E = \bigcup_k Q_k$ (teorema 1.3.8). Quindi $\mathcal{L}^n_e(G \setminus E) \leq \sum_k \mathcal{L}^n_e(Q_k)$. Basta allora mostrare che $\sum_k \mathcal{L}^n_e(Q_k) \leq \epsilon$. Poiché, per ogni N

$$G = E \cup \bigcup_k^{\infty} Q_k \supset E \cup \bigcup_k^{N} Q_k,$$

abbiamo che

$$\mathcal{L}^n(G) \geq \mathcal{L}^n_e(E \cup \bigcup_k^{N} Q_k) = \mathcal{L}^n_e(E) + \mathcal{L}^n_e(\bigcup_k^{N} Q_k).$$

(Infatti gli insiemi compatti e disgiunti E e $\bigcup_k^N Q_k$ hanno distanza positiva, e quindi si può utilizzare il teorema 3.2.10.) D'altra parte, usando il teorema 3.2.11 abbiamo che $\mathcal{L}^n(\bigcup_k^N Q_k) = \sum_{k=1}^N \mathcal{L}^n(Q_k)$, e quindi otteniamo che

$$\sum_{k=1}^{N} \mathcal{L}^n(Q_k) \leq \mathcal{L}^n(G) - \mathcal{L}^n_e(E) \leq \epsilon.$$

Dall'arbitrarietà di N segue allora che $\sum_{k=1}^{\infty} \mathcal{L}^n(Q_k) \leq \epsilon$, e il teorema è dimostrato se E è compatto.

Sia ora E chiuso in $\mathbb{R}^n$. Ma allora $E = \bigcup_k E_k$, ove $E_k = \{\, x \in E \mid |x| \leq k \,\}$. Poiché E_k è compatto, è misurabile e quindi E, unione di misurabili, è misurabile. □

Teorema 3.3.8. *Sia E misurabile. Allora E^c è misurabile.*

Dimostrazione. Per ogni k esiste un aperto $G_k \supset E$ tale che $\mathcal{L}^n_e(G_k \setminus E) \leq 1/k$. G_k^c è chiuso, e quindi misurabile, da cui segue la misurabilità anche dell'insieme $H = \bigcup_k G_k^c \subset E^c$. Sia $Z = E^c \setminus H$. Poiché $Z \subset E^c \setminus G_k^c = G_k \setminus E$, abbiamo che $\mathcal{L}^n_e(Z) \leq \mathcal{L}^n_e(G_k \setminus E) \leq 1/k$. Poiché k è arbitrario, abbiamo che $\mathcal{L}^n_e(Z) = 0$ e quindi Z è misurabile. Quindi $E^c = H \cup Z$ è l'unione di due insiemi misurabili ed è pertanto misurabile. □

Dai risultati sopra otteniamo pertanto che

Corollario 3.3.9. *I sottoinsiemi di $\mathbb{R}^n$ misurabili formano una σ-algebra che contiene gli aperti. In particolare sono misurabili tutti i Boreliani.*

La definizione di misurabilità dell'insieme E è stata data in termini di insiemi aperti che contengono E. Come ci si può aspettare, tale definizione è equivalente ad una condizione data mediante i *chiusi contenuti* in E, e la misurabilità può anche essere caratterizzata mediante gli insiemi di tipo $\mathcal{F}_\sigma$ e $\mathcal{G}_\delta$.

Teorema 3.3.10. *Sia $E \subset \mathbb{R}^n$. Sono equivalenti*

(a) *E è misurabile;*

(b) *Per ogni $\epsilon > 0$ esiste un chiuso $F \subset E$ tale che $\mathcal{L}^n_e(E - F) < \epsilon$;*

(c) *Esistono un insieme H di tipo $\mathcal{G}_\delta$ e un insieme Z di misura nulla tali che $E = H \setminus Z$;*

(d) *Esistono un insieme F di tipo $\mathcal{F}_\sigma$ e un insieme Z di misura nulla tali che $E = F \cup Z$.*

Dimostrazione. **(a)** equivalente **(b)**: Sia E misurabile, $\epsilon > 0$ assegnato. Poiché E^c è misurabile, esiste un aperto $G \supset E^c$, $\mathcal{L}^n_e(G \setminus E^c) \leq \epsilon$. Ma allora $E \supset G^c$, G^c è chiuso e $\mathcal{L}^n_e(E \setminus G^c) = \mathcal{L}^n_e(G \setminus E^c) \leq \epsilon$. Per dimostrare il viceversa, sia $\epsilon > 0$ assegnato. Allora esiste $C \subset E$, C chiuso, $\mathcal{L}^n_e(E \setminus C) \leq \epsilon$. Ne segue che $E^c \subset C^c$, con C^c aperto e $\mathcal{L}^n_e(C^c \setminus E^c) = \mathcal{L}^n_e(E \setminus C) \leq \epsilon$, da cui segue la misurabilità di E^c e quindi di E.

(a) equivalente **(c)**: È chiaro che $E = H \setminus Z$ con H di tipo $\mathcal{G}_\delta$ (e quindi misurabile, poiché gli insiemi misurabili formano una sigma algebra) e Z di misura nulla (e quindi misurabile) implica E misurabile. Per dimostrare il viceversa basta osservare che, per ogni $k \in \mathbb{N}$, esiste un aperto $G_k \supset E$ tale che $\mathcal{L}^n(G_k - E) \leq 1/k$. Posto allora $H = \cap_k G_k$ e $Z = H \setminus E$ il risultato segue.

(a) equivalente **(d)**: Si può dimostrare sia per passaggio al complementare in **(c)**, sia utilizzando il fatto che per ogni $k \in \mathbb{N}$ esiste un chiuso $F_k \subset E$ tale che $\mathcal{L}^n(E \setminus F_k) \leq \epsilon$. □

3.4. Additività della misura di Lebesgue

Abbiamo visto nella sezione precedente che la misura di Lebesgue è definita per tutti gli insiemi appartenenti ad una sigma-algebra che contiene gli aperti. Vogliamo ora dimostrare che la misura di Lebesgue è effettivamente una misura, cioè che è numerabilmente additiva.

Teorema 3.4.1. *Sia $E_k \subset \mathbb{R}^n$ una successione numerabile di insiemi a due a due disgiunti e misurabili in $\mathbb{R}^n$. Allora*

$$\mathcal{L}^n(\bigcup_k E_k) = \sum_k \mathcal{L}^n(E_k).$$

Dimostrazione. Supponiamo innanzi tutto che ciascun E_k sia limitato. Sia $\epsilon > 0$ assegnato. Sappiamo che per ogni k esiste un chiuso $F_k \subset E_k$ tale che $\mathcal{L}^n(E_k \setminus F_k) \leq 2^{-k}\epsilon$. Essendo F_k chiuso e limitato, F_k è compatto; ne segue, utilizzando il teorema 3.2.10, che per ogni N

$$\mathcal{L}^n(\bigcup_{k=1}^N F_k) = \sum_{k=1}^N \mathcal{L}^n(F_k).$$

Abbiamo allora che

$$\begin{aligned}\mathcal{L}^n(\bigcup_{k=1}^\infty E_k) \geq \mathcal{L}^n(\bigcup_{k=1}^N F_k) &= \sum_{k=1}^N \mathcal{L}^n(F_k)\\ &\geq \sum_{k=1}^N (\mathcal{L}^n(E_k) - \mathcal{L}^n(E_k \setminus F_k)) \geq \sum_{k=1}^N \mathcal{L}^n(E_k) - \epsilon.\end{aligned}$$

Dall'arbitrarietà di N e ϵ segue allora che $\mathcal{L}^n(\bigcup_{k=1}^\infty E_k) \geq \sum_{k=1}^\infty \mathcal{L}^n(E_k)$. Poiché la diseguaglianza opposta segue dalla subadditività della misura esterna, il teorema è dimostrato nel caso in cui E_k è limitato per tutti i k.

Nel caso generale, consideriamo, per ogni k e $j \in \mathbb{N}$, gli insiemi $E_{k,j} = \{ x \in E_k \mid j-1 \leq |x| < j \}$. Ovviamente $E_k = \bigcup_j E_{k,j}$, e gli $E_{k,j}$ sono limitati e a due a due disgiunti. Ne segue che

$$\mathcal{L}^n(\bigcup_k E_k) = \mathcal{L}^n(\bigcup_{k,j} E_{k,j}) = \sum_{k,j} \mathcal{L}^n(E_{k,j}) = \sum_k \mathcal{L}^n(E_k).$$

□

Un (utile) corollario è il seguente

Corollario 3.4.2. *Sia* $\{R_k\}$ *una collezione numerabile di rettangoli non sovrapposti. Allora*
$$\mathcal{L}^n(\bigcup_k R_k) = \sum_k \mathrm{vol}(R_k).$$

Dimostrazione. Abbiamo che $\{\mathring{R}_k\}$ è una famiglia numerabile di insiemi a due a due disgiunti misurabili. Quindi
$$\mathcal{L}^n(\bigcup_k R_k) \geq \mathcal{L}^n(\bigcup_k \mathring{R}_k) = \sum_k \mathcal{L}^n(\mathring{R}_k) = \sum_k \mathrm{vol}(R_k).$$
Poiché la diseguaglianza opposta segue dalla subadditività, il corollario è dimostrato. □

Dimostriamo ora un'importante proprietà della misura di Lebesgue:

Corollario 3.4.3 (Regolarità della misura di Lebesgue)**.** *La misura di Lebesgue è* **regolare***, cioè gode delle seguenti proprietà:*

Regolarità esterna. $\mathcal{L}^n(E) = \inf\{\, \mathcal{L}^n(O) \mid E \subset O,\ O \text{ aperto} \,\}$*;*

Regolarità interna. $\mathcal{L}^n(E) = \sup\{\, \mathcal{L}^n(C) \mid C \subset E,\ C \text{ compatto} \,\}$*;*

Dimostrazione. La regolarità esterna vale addirittura per la misura esterna di un sottoinsieme qualunque di $\mathbb{R}^n$ (si veda il teorema 3.2.9**(a)**).

Per dimostrare la regolarità interna, osserviamo che dal teorema 3.3.10**(b)** segue che per ogni ϵ esiste un chiuso $F \subset E$ tale che
$$\mathcal{L}^n(E) = \mathcal{L}^n(F) + \mathcal{L}^n(E \setminus F) < \mathcal{L}^n(F) + \epsilon.$$
e quindi $\mathcal{L}^n(F) > \mathcal{L}^n(E) - \epsilon$. Se F è limitato, abbiamo terminato. In caso contrario, poniamo
$$F_k = \{\, x \in F \mid |x| \leq k \,\}.$$
Allora F_k è compatto per ogni k, $F_k \subset F_{k+1}$ e $\bigcup_{k=1}^{\infty} F_k = F$. Segue allora dalla Proposizione 2.2.4**(c)** che $\lim_{k\to\infty} \mathcal{L}^n(F_k) \to \mathcal{L}^n(F)$ e il corollario segue. □

Possiamo ora dimostrare un'altra caratterizzazione dei misurabili, dovuta a Carathéodory.

Teorema 3.4.4 (Condizione di Carathéodory)**.** $E \subset \mathbb{R}^n$ *è misurabile se, e solo se, per ogni* $A \subset \mathbb{R}^n$ *si ha che*
$$\mathcal{L}_e^n(A) = \mathcal{L}_e^n(A \cap E) + \mathcal{L}_e^n(A \setminus E).$$

Osservazione 3.4.5. Un insieme è quindi misurabile se spezza ogni insieme A (non necessariamente misurabile) in due parti in modo tale che la somma delle misure esterne delle parti è la misura esterna di A. L'interesse della caratterizzazione di Carathéodory sta nel fatto che non dipende da proprietà specifiche di $\mathbb{R}^n$ o della misura esterna considerata. Tale condizione permette, data una qualunque misura esterna definita sui sottoinsiemi di un insieme X, di definire una classe di insiemi (quelli che soddisfano alla condizione di Carathéodory) che si può dimostrare essere una sigma-algebra, su cui la misura esterna risulta essere una misura. Fornisce quindi uno strumento generale per costruire misure a partire da misure esterne.

Dimostrazione. Supponiamo E misurabile. Sia A un sottoinsieme qualunque di $\mathbb{R}^n$. Ovviamente $\mathcal{L}^n_e(A) \leq \mathcal{L}^n_e(A \cap E) + \mathcal{L}^n_e(A \setminus E)$. Sappiamo inoltre che esiste un insieme H di tipo $\mathcal{G}_\delta$ tale che $A \subset H$, $\mathcal{L}^n_e(A) = \mathcal{L}^n(H)$. Abbiamo che $H = (H \cap E) \cup (H \setminus E)$. Poiché H è misurabile e $(H \cap E) \cap (H \setminus E) = \varnothing$, $\mathcal{L}^n_e(A) = \mathcal{L}^n(H) = \mathcal{L}^n(H \cap E) + \mathcal{L}^n(H \setminus E) \geq \mathcal{L}^n_e(A \cap E) + \mathcal{L}^n_e(A \setminus E)$.

Dimostriamo ora il viceversa. Supponiamo che E soddisfi alla condizione di Carathéodory per ogni A. Assumiamo, per il momento, $\mathcal{L}^n_e(E) < +\infty$. Allora esiste un insieme H di tipo $\mathcal{G}_\delta$ (in particolare misurabile), tale che $E \subset H$ e $\mathcal{L}^n_e(E) = \mathcal{L}^n(H)$. La condizione di Carathéodory, con $A = H$, implica che

$$\mathcal{L}^n(H) = \mathcal{L}^n_e(H \cap E) + \mathcal{L}^n_e(H \setminus E) = \mathcal{L}^n_e(E) + \mathcal{L}^n_e(H \setminus E).$$

Ne deduciamo che $\mathcal{L}^n_e(H \setminus E) = \mathcal{L}^n(H) - \mathcal{L}^n(E) = 0$. Ma allora l'insieme $H \setminus E$ ha misura esterna nulla; quindi è misurabile. Ma allora anche $E = H \setminus (H \setminus E)$ è misurabile.

Se E non è di misura finita, consideriamo $E_k = \left\{ x \in E \;\middle|\; |x| \leq k \right\}$. Gli E_k sono di misura finita e $E = \bigcup_k E_k$. Sia, per ogni k, H_k un insieme di tipo $\mathcal{G}_\delta$ che contiene E_k e tale che $\mathcal{L}^n(H_k) = \mathcal{L}^n_e(E_k)$. Per ipotesi

$$\mathcal{L}^n(H_k) = \mathcal{L}^n_e(H_k \cap E) + \mathcal{L}^n_e(H_k \setminus E) = \mathcal{L}^n_e(E_k) + \mathcal{L}^n_e(H_k \setminus E).$$

Ne segue che $\mathcal{L}^n_e(H_k \setminus E) = 0$ per ogni k. Ma allora tutti gli insiemi $H_k \setminus E$ sono misurabili e quindi lo è anche $E = \bigcup_k (H_k \setminus (H_k \setminus E))$. $\square$

Un risultato che si ottiene facilmente da questo teorema è il seguente

Teorema 3.4.6. *Sia E un insieme misurabile, ed A un sottoinsieme qualunque. Se $E \subset A$ si ha che*

$$\mathcal{L}^n_e(A) = \mathcal{L}^n(E) + \mathcal{L}^n_e(A \setminus E).$$

Nel caso in cui $\mathcal{L}^n(E) < +\infty$, ne deduciamo che $\mathcal{L}^n_e(A \setminus E) = \mathcal{L}^n_e(A) - \mathcal{L}^n(E)$.

Possiamo anche dimostrare una versione piú forte del teorema 3.2.9**(b)**.

Teorema 3.4.7. *Sia E un sottoinsieme di $\mathbb{R}^n$. Allora esiste un insieme H di tipo $\mathcal{G}_\delta$ tale che $E \subset H$ e $\mathcal{L}^n_e(E \cap M) = \mathcal{L}^n(H \cap M)$ per tutti gli insiemi M misurabili.*

Dimostrazione. Consideriamo prima il caso in cui $\mathcal{L}^n_e(E) < \infty$. Sia H un insieme di tipo $\mathcal{G}_\delta$ tale che $E \subset H$, $\mathcal{L}^n_e(E) = \mathcal{L}^n(H)$. Se M è un qualunque misurabile, abbiamo che $\mathcal{L}^n_e(E) = \mathcal{L}^n_e(E \cap M) + \mathcal{L}^n_e(E \setminus M)$ e che $\mathcal{L}^n(H) = \mathcal{L}^n_e(H \cap M) + \mathcal{L}^n_e(H \setminus M)$. Quindi

$$\mathcal{L}^n_e(E \cap M) + \mathcal{L}^n_e(E \setminus M) = \mathcal{L}^n(H \cap M) + \mathcal{L}^n(H \setminus M).$$

Poiché $E \setminus M \subset H \setminus M$, ne deduciamo $\mathcal{L}^n_e(E \cap M) \geq \mathcal{L}^n(H \cap M)$; e l'inclusione $E \cap M \subset H \cap M$ permette di concludere.

Nel caso in cui $\mathcal{L}^n_e(E) = +\infty$, sia $E_k = E \cap B(0,k)$. Allora $E_k \nearrow E$, e $\mathcal{L}^n_e(E_k) < +\infty$. Per quanto visto, per ogni k esiste un insieme H_k di tipo $\mathcal{G}_\delta$ tale che $E_k \subset H_k$ e $\mathcal{L}^n_e(E_k \cap M) = \mathcal{L}^n(H_k \cap M)$ per tutti gli M misurabili. Poniamo $\hat{H}_m = \bigcap_{k=m}^\infty H_k$. Abbiamo che $E_m \subset \hat{H}_m \subset H_m$, $\hat{H}_m \nearrow H = \bigcup_m \hat{H}_m$ e quindi $\mathcal{L}^n(\hat{H}_m \cap M) = \mathcal{L}^n_e(E_m \cap M)$ per ogni M misurabile. Utilizzando la proposizione 2.2.4 e il teorema 3.4.8 otteniamo che $\mathcal{L}^n(\hat{H}_m \cap M) \to \mathcal{L}^n(H \cap M)$, che $\mathcal{L}^n_e(E_m \cap M) \to \mathcal{L}^n_e(E \cap M)$ e quindi $\mathcal{L}^n(H \cap M) = \mathcal{L}^n(E \cap M)$. L'insieme H ha la proprietà voluta, ma non è di tipo $\mathcal{G}_\delta$. (È di tipo $\mathcal{G}_{\delta\sigma}$.) Per trovare un insieme di tipo $\mathcal{G}_\delta$ con la stessa proprietà, osserviamo che $H = H_1 \setminus Z$, ove H_1 è un $\mathcal{G}_\delta$ e $\mathcal{L}^n(Z) = 0$. Quindi $E \subset H \subset H_1$, e $H_1 \cap M = (H \cap M) \cup (Z \cap M)$. Il risultato allora segue dal fatto che $\mathcal{L}^n(Z \cap M) = 0$. □

Sappiamo che una delle conseguenze della numerabile additività è che la misura dell'unione di una successione crescente di misurabili è il limite delle misure degli elementi della successione (si veda la proposizione 2.2.4). La stessa proprietà vale per la misura esterna di Lebesgue.

Teorema 3.4.8. *Sia E_k una successione crescente di insiemi qualunque in $\mathbb{R}^n$. Allora $\mathcal{L}^n_e(E_k) \to \mathcal{L}^n_e(\bigcup_k E_k)$.*

Dimostrazione. Consideriamo, per ogni k, un insieme H_k di tipo $\mathcal{G}_\delta$ tale che $E_k \subset H_k$, $\mathcal{L}^n_e(E_k) = \mathcal{L}^n(H_k)$. Poniamo poi $\hat{H}_k = \bigcap_{m=k}^\infty H_m$. Ovviamente vale $E_k \subset \hat{H}_k \subset H_k$, da cui otteniamo che $\mathcal{L}^n_e(E_k) = \mathcal{L}^n(\hat{H}_k)$, e la successione di insiemi (misurabili) $\hat{H}_k$ è crescente. Allora, per la **(c)** della proposizione 2.2.4, $\mathcal{L}^n_e(\bigcup_k E_k) \geq \mathcal{L}^n_e(E_k) = \mathcal{L}^n(\hat{H}_k) \nearrow \mathcal{L}^n(\bigcup_k \hat{H}_k) \geq \mathcal{L}^n_e(\bigcup_k E_k)$, e il teorema segue. □

3.5. Insiemi di misura nulla

Definizione 3.5.1. Sia (X, Σ, μ) uno spazio di misura, e $P(x)$ una proprietà dipendente dal punto $x \in X$. Diciamo che tale proprietà vale quasi ovunque se esiste un insieme $Z \in \Sigma$ tale che $P(x)$ è vera per ogni $x \in X \setminus Z$ e $\mu(Z) = 0$.

Ad esempio la funzione di Dirichlet $\chi_{\mathbb{Q}}$ è ≤ 0 per quasi ogni $x \in \mathbb{R}$, se consideriamo in $\mathbb{R}$ la misura di Lebesgue (ma $\chi_{\mathbb{Q}}(x) = 1$ per quasi ogni $x \in \mathbb{R}$ se in $\mathbb{R}$ consideriamo la misura delta di Dirac).

Abbiamo visto che tutti gli insiemi di misura di Lebesgue esterna nulla sono misurabili. In particolare da $E \subset A$, A misurabile secondo Lebesgue e $\mathcal{L}^n(A) = 0$ deduciamo che E è misurabile (secondo Lebesgue). Questa proprietà non vale per una misura arbitraria, ad esempio mostreremo che non vale per lo spazio di misura $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathcal{L}^n)$.

Definizione 3.5.2. Diciamo che la misura μ definita nello spazio di misura (X, Σ, μ) è una misura completa se Σ contiene tutti i sottoinsiemi degli insiemi di misura nulla.

Il fatto che non tutte le misure siano complete è un fatto del tutto naturale (sono in generale non complete le misure prodotto) e può creare alcuni inconvenienti. Vale però il seguente teorema:

Teorema 3.5.3. *Sia (X, Σ, μ) uno spazio di misura. Allora esiste una sigma algebra $\overline{\Sigma} \supset \Sigma$ e una misura completa μ' definita in $\overline{\Sigma}$, che estende μ. Tale misura μ' si chiama il completamento della misura μ.*

Dimostrazione. Basta considerare $\overline{\Sigma} = \{ E \subset X \mid A \subset E \subset B, \mu(B \setminus A) = 0 \text{ per qualche } A, B \in \Sigma \}$. Si mostra facilmente che $\overline{\Sigma}$ è una sigma algebra. Si pone poi $\mu'(E) = \mu(A)$ se $A \subset E \subset B$, $A, B \in \Sigma$, $\mu(B \setminus A) = 0$. Tale μ' è la misura completa cercata. □

Dalla caratterizzazione dei misurabili secondo Lebesgue è chiaro che la misura di Lebesgue è il completamento della misura di Lebesgue ristretta ai Boreliani.

3.6. Invarianza per rototraslazioni

In questa sezione vogliamo mostrare che la misura di Lebesgue è effettivamente invariante per rototraslazioni. L'invarianza per traslazioni è insita nel fatto che la misura esterna è stata costruita a partire dal volume dei rettangoli, quantità invariante per traslazioni. Il problema con le rotazioni sorge dal fatto che i rettangoli non vengono mandati in rettangoli da rotazioni arbitrarie. Esaminiamo il problema da un punto di vista piú generale, analizzando il modo con cui gli insiemi misurabili si comportano per trasformazioni generali.

Lemma 3.6.1. *Sia $F \subset \mathbb{R}^n$ un insieme chiuso e $\Phi\colon F \to \mathbb{R}^m$ un'applicazione continua. Allora, per ogni insieme Γ di tipo $\mathcal{F}_\sigma$, abbiamo che l'insieme $\Phi(F \cap \Gamma)$ è di tipo $\mathcal{F}_\sigma$.*

Inoltre, se sappiamo che $\mathcal{L}_e^m(\Phi(F \cap A)) = 0$ per tutti gli insiemi A di misura esterna nulla, allora $\Phi(F \cap E)$ è misurabile ogni qualvolta E è misurabile.

In particolare, se $n \le m$ e Φ è un'applicazione lipschitziana con costante di Lipschitz L, cioè se $|\Phi(x) - \Phi(y)| \le L|x - y|$ per tutti gli $x, y \in F$, allora $\mathcal{L}_e^m(\Phi(E \cap F)) \le (2L\sqrt{n})^m \mathcal{L}_e^n(E)$ e quindi Φ manda sottoinsiemi misurabili di F in sottoinsiemi misurabili.

Dimostrazione. Poiché $\Phi(\bigcup_k E_k) = \bigcup_k \Phi(E_k)$ la classe $\mathcal{A}$ dei sottoinsiemi Γ tali che $\Phi(\Gamma \cap F)$ è di tipo $\mathcal{F}_\sigma$ è chiusa per unioni numerabili.

Se Γ è compatto, allora $\Phi(\Gamma \cap F)$ è compatto, e quindi di tipo $\mathcal{F}_\sigma$. Ne segue che $\mathcal{A}$ contiene tutti i compatti, e quindi i chiusi, che sono unioni numerabili di compatti, e quindi tutti gli insiemi di tipo $\mathcal{F}_\sigma$, che sono unioni numerabili di chiusi. La prima affermazione è quindi dimostrata.

Supponiamo ora che $\mathcal{L}^m_e(\Phi(F \cap A)) = 0$ per tutti gli insiemi A di misura esterna nulla, e sia E misurabile. Allora $E = H \cup Z$, con H di tipo $\mathcal{F}_\sigma$ e Z di misura esterna nulla. Ne segue che $\Phi(E \cap F) = \Phi((H \cap F) \cup (Z \cap F)) = \Phi(H \cap F) \cup \Phi(Z \cap F)$ è misurabile in quanto unione dei misurabili $\Phi(H \cap F)$ (misurabile poiché di tipo $\mathcal{F}_\sigma$) e $\Phi(Z \cap F)$ (misurabile poiché di misura esterna nulla).

Sia ora Φ un'applicazione lipschitziana con costante di Lipschitz L. Mostriamo innanzi tutto che $\mathcal{L}^m_e(\Phi(Q \cap F)) \le (2L\sqrt{n})^m\, \mathcal{L}^n_e(Q)$ per ogni cubo Q di lato $r \le 1$. Se $Q \cap F = \varnothing$, non vi è nulla da dimostrare. Supponiamo che $x_0 \in Q \cap F$. Allora $\Phi(Q \cap F)$ contiene $\Phi(x_0)$ ed è contenuto nella palla di centro $\Phi(x_0)$ e raggio $Lr\sqrt{n}$. Quindi $\Phi(Q \cap F)$ è contenuto nel cubo $\prod_{i=1}^m [\Phi(x_0)_i - \sqrt{n}rL, \Phi(x_0)_i + \sqrt{n}rL]$. Ne segue che $\mathcal{L}^m_e(\Phi(Q \cap F)) \le (2rL\sqrt{n})^m = (2L\sqrt{n})^m r^{m-n}\, \mathcal{L}^n_e(Q) \le (2L\sqrt{n})^m\, \mathcal{L}^n_e(Q)$. Sia ora $E \subset \mathbb{R}^n$. Per ogni $\epsilon > 0$ esiste un aperto G tale che $E \cap F \subset G$ e $\mathcal{L}^n(G) \le \mathcal{L}^n_e(E \cap F) + \epsilon$. Poiché G è aperto, $G = \bigcup_k Q_k$, dove i Q_k sono cubi non sovrapposti che possiamo assumere di lato $r_k \le 1$. Allora

$$
\begin{aligned}
\mathcal{L}^m_e(\Phi(E \cap F)) &\le \mathcal{L}^m_e(\Phi(G \cap F)) = \mathcal{L}^m_e(\Phi(F \cap \bigcup_k Q_k)) \\
&= \mathcal{L}^m_e(\bigcup_k \Phi(F \cap Q_k)) \le \sum_k \mathcal{L}^m_e(\Phi(F \cap Q_k)) \\
&\le (2L\sqrt{n})^m \sum_k \mathcal{L}^n_e(Q_k) = (2L\sqrt{n})^m\, \mathcal{L}^n_e(G) \\
&\le (2L\sqrt{n})^m\, \mathcal{L}^n_e(E \cap F) + (2L\sqrt{n})^m \epsilon,
\end{aligned}
$$

da cui il risultato. □

Teorema 3.6.2. *Sia* $\mathsf{T} \colon \mathbb{R}^n \to \mathbb{R}^n$ *un'applicazione lineare, che identificheremo con la matrice* $\mathsf{T} = \{a_{ij}\}$. *Allora* $\mathsf{T}(E)$ *è misurabile per ogni* $E \subset \mathbb{R}^n$ *misurabile e vale*

$$\mathcal{L}^n_e(\mathsf{T}(E)) = |\det(\mathsf{T})|\, \mathcal{L}^n_e(E) \qquad \textit{per ogni } E \subset \mathbb{R}^n.$$

Dimostrazione. Ovviamente T, essendo lineare, è lipschitziana, e quindi manda misurabili in misurabili. Dividiamo ora la dimostrazione in vari passi.

Passo 1. $\mathcal{L}^n_e(c + \lambda E) = |\lambda|^n\, \mathcal{L}^n_e(E)$ per tutti i $c \in \mathbb{R}^n$, per tutti i $\lambda \in \mathbb{R}$ e per tutti gli $E \subset \mathbb{R}^n$.

Per l'invarianza per traslazioni della misura esterna possiamo assumere $c = 0$. Se $\lambda = 0$ è ovvio. Assumiamo allora $\lambda > 0$. Abbiamo che per ogni rettangolo R vale $\mathrm{vol}(\lambda R) = |\lambda|^n \mathrm{vol}(R)$. Poiché $\{R_k\}$ ricopre E se e solo se $\{\lambda R_k\}$ ricopre λE, il primo passo segue dalla definizione di misura esterna.

Passo 2. Se $\mathsf{T}\colon \mathbb{R}^n \to \mathbb{R}^n$ è un'applicazione lineare e Q un cubo, si ha che $\mathcal{L}^n(\mathsf{T}(Q)) = \alpha(\mathsf{T})\,\mathcal{L}^n(Q)$, dove $\alpha(\mathsf{T}) = \mathcal{L}^n(\mathsf{T}(Q_0))$, $Q_0 = \prod_{i=1}^n [0,1]$.

Per ogni cubo Q esistono $c \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ tali che $Q = c + \lambda Q_0$ e $\mathrm{vol}(Q) = |\lambda|^n$. Allora, usando il passo 1 e la linearità di T si ottiene

$$\mathcal{L}^n(\mathsf{T}(Q)) = \mathcal{L}^n(c + \lambda \mathsf{T}(Q_0)) = |\lambda|^n\,\mathcal{L}^n(\mathsf{T}(Q_0)),$$

da cui il risultato.

Passo 3. $\mathcal{L}^n(\mathsf{T}(G)) \le \alpha(\mathsf{T})\,\mathcal{L}^n(G)$ per tutti i G aperti e per tutte le applicazioni lineari T. Vale l'uguaglianza se T è non singolare.

Sappiamo che $G = \bigcup_k Q_k$, dove i Q_k sono cubi non sovrapposti. Ne segue che

$$\begin{aligned}\mathcal{L}^n(\mathsf{T}(G)) &= \mathcal{L}^n(\bigcup_k \mathsf{T}(Q_k)) \le \sum_k \mathcal{L}^n(\mathsf{T}(Q_k))\\ &= \alpha(\mathsf{T}) \sum_k \mathcal{L}^n(Q_k) = \alpha(\mathsf{T})\,\mathcal{L}^n(G).\end{aligned}$$

Nel caso in cui T è non singolare, abbiamo che $\mathsf{T}(\mathring{Q}_k) \cap \mathsf{T}(\mathring{Q}_j) = \varnothing$ se $k \neq j$, e che $\mathcal{L}^n(\mathsf{T}(\partial Q_j)) = 0$ in quando $\mathcal{L}^n(\partial Q_j) = 0$ e T è lipschitziana. Ne segue subito che

$$\begin{aligned}\mathcal{L}^n(\mathsf{T}(G)) \ge \mathcal{L}^n(\bigcup_k \mathsf{T}(\mathring{Q}_k)) &= \sum_k \mathcal{L}^n(\mathsf{T}(\mathring{Q}_k)) = \sum_k \mathcal{L}^n(\mathsf{T}(\mathring{Q}_k) \cup \mathsf{T}(\partial Q_k))\\ &= \sum_k \mathcal{L}^n(\mathsf{T}(Q_k)) = \alpha(\mathsf{T}) \sum_k \mathcal{L}^n(Q_k) = \alpha(\mathsf{T})\,\mathcal{L}^n(G)\end{aligned}$$

e il risultato segue.

Passo 4. $\mathcal{L}^n_e(\mathsf{T}(E)) = \alpha(\mathsf{T})\,\mathcal{L}^n_e(E)$ per tutte le applicazioni lineari non singolari T e per tutti i sottoinsiemi $E \subset \mathbb{R}^n$.

Questo segue da: (i) $E \subset G$ con G aperto se e solo se $\mathsf{T}(E) \subset \mathsf{T}(G)$ con $T(G)$ aperto (dalla continuità di T e di T^{-1}); (ii) $\mathcal{L}^n_e(E) = \inf\big\{\,\mathcal{L}^n(G) \;\big|\; E \subset G,\ G \text{ aperto}\,\big\}$; (iii) $\mathcal{L}^n(G) = \alpha(\mathsf{T})\,\mathcal{L}^n(\mathsf{T}(G))$ per tutti gli aperti G (passo 3) e per tutte le applicazioni lineari non singolari T.

Passo 5. Se S e T sono applicazioni lineari, e S è non singolare, allora $\alpha(S \circ \mathsf{T}) = \alpha(S)\alpha(\mathsf{T})$.

Immediato da

$$\alpha(S \circ \mathsf{T}) = \mathcal{L}^n(S \circ \mathsf{T}(Q_0)) = \mathcal{L}^n(S(\mathsf{T}(Q_0))) = \alpha(S)\,\mathcal{L}^n(\mathsf{T}(Q_0)) = \alpha(S)\alpha(\mathsf{T}).$$

Passo 6. Se T è una matrice rotazione $\alpha(\mathsf{T}) = 1$.

Segue dal fatto che $B(0,1) = \mathsf{T}(B(0,1))$. Infatti ciò implica

$$\mathcal{L}^n(B(0,1)) = \alpha(\mathsf{T})\,\mathcal{L}^n(B(0,1)),$$

da cui $\alpha(\mathsf{T}) = 1$.

Passo 7. $\alpha(\mathsf{T}) = |\det(\mathsf{T})|$ se T è non singolare e simmetrica.

Se T è già una matrice diagonale, è chiaro che $\alpha(\mathsf{T}) = \mathcal{L}^n(\mathsf{T}(Q_0)) = |\lambda_1 \cdots \lambda_n| = |\det(\mathsf{T})|$ se i λ_k sono gli elementi sulla diagonale. Nel caso generale esiste una matrice S ortogonale che diagonalizza T, cioè tale che $\mathsf{A} = \mathsf{S}\mathsf{T}\mathsf{S}^{\mathrm{T}}$, con A diagonale e tale che $\det(\mathsf{A}) = \det(\mathsf{T})$. Allora $|\det(\mathsf{A})| = \alpha(\mathsf{A}) = \alpha(\mathsf{S})\alpha(\mathsf{T})\alpha(\mathsf{S}^{\mathrm{T}})$, e il risultato segue in quanto S e S^{T} sono rotazioni.

Passo 8. $\alpha(\mathsf{T}) = |\det(\mathsf{T})|$ per ogni T nonsingolare.

Consideriamo la matrice simmetrica e non singolare $\mathsf{T}^{\mathrm{T}}\mathsf{T}$. Esiste una rotazione R tale che $\mathsf{R}^{\mathrm{T}}\mathsf{T}^{\mathrm{T}}\mathsf{T}\mathsf{R} = \mathsf{D}$, ove D è una matrice diagonale, di elementi λ_i. Poiché $\langle e_i, \mathsf{D}e_i\rangle = \lambda_i = |\mathsf{T}\mathsf{R}e_i|^2$, abbiamo subito che $0 < \lambda_i$ per ogni i. Inoltre $\lambda_1 \cdot \lambda_2 \cdots \lambda_n = \det(\mathsf{D}) = |\det(T)|^2$. Definiamo la matrice $\mathsf{D}^{1/2}$ essere la matrice diagonale di elementi $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}$. Notiamo che tale matrice verifica l'equazione $\mathsf{D}^{1/2}\mathsf{D}^{1/2} = D$. Definiamo poi la matrice $\mathsf{B} = \mathsf{R}\mathsf{D}^{1/2}\mathsf{R}^{\mathrm{T}}$. Poichè verifica l'equazione $\mathsf{B}\mathsf{B} = \mathsf{R}\mathsf{D}\mathsf{R}^{\mathrm{T}} = \mathsf{T}^{\mathrm{T}}\mathsf{T}$, tale matrice viene solitamente indicata come $\mathsf{B} = (\mathsf{T}^{\mathrm{T}}\mathsf{T})^{1/2}$.

Inoltre è immediato verificare che B è simmetrica, nonsingolare e che $\det(\mathsf{B}) = |\det(\mathsf{T})|$. Poniamo poi $\mathsf{S} = \mathsf{B}^{-1}\mathsf{T}^{\mathrm{T}}$ e osserviamo che $\mathsf{S}^{\mathrm{T}} = \mathsf{T}\mathsf{B}^{-1}$ (B è simmetrica), cosí che $\mathsf{S}\mathsf{S}^{\mathrm{T}} = \mathsf{B}^{-1}\mathsf{T}^{\mathrm{T}}\mathsf{T}\mathsf{B}^{-1} = \mathsf{B}^{-1}\mathsf{B}\mathsf{B}\mathsf{B}^{-1}\mathbb{I}$, la matrice identità. Quindi S è ortogonale. Poiché $\mathsf{T} = \mathsf{S}^{\mathrm{T}}\mathsf{B}$, abbiamo che $\alpha(\mathsf{T}) = \alpha(\mathsf{B}) = |\det(\mathsf{B})| = |\det(\mathsf{T})|$.

Passo 9. $\alpha(\mathsf{T}) = 0$ se T è singolare.

Scegliamo $y \in \mathbb{R}^n$ ortogonale al codominio di T e di norma 1. Prendiamo poi una matrice ortogonale S tale che $e_1 = \mathsf{S}y$. Allora e_1 è ortogonale al codominio di $\mathsf{S} \circ \mathsf{T}$, e quindi esiste un rettangolo R in $\mathbb{R}^{n-1}$ tale che $\mathsf{S} \circ \mathsf{T}(Q_0) \subset \{0\} \times R$. Quindi $\alpha(\mathsf{T}) = \alpha(\mathsf{S} \circ \mathsf{T}) = 0$. □

3.7. Insiemi non misurabili

Costruiamo ora un insieme non misurabile. Per fare ciò dovremo, come annunciato, fare uso dell'assioma della scelta. Cominciamo dimostrando un risultato preliminare, interessante di per se e che non fa uso dell'assioma della scelta.

Lemma 3.7.1. *Sia $E \subset \mathbb{R}$ è un insieme misurabile, di misura $\mathcal{L}^1(E) > 0$.*

Allora l'insieme $E - E = \{ x - y \mid x, y \in E \}$ contiene, per qualche $\delta > 0$, l'intervallo $(-\delta, \delta)$.

Dimostrazione. Esiste un insieme aperto G che contiene E e tale che $\mathcal{L}^1(G \setminus E) < \frac{1}{3}\mathcal{L}^1(E)$. Sappiamo che $G = \bigcup_k I_k$, dove $I_k = (a_k, b_k)$ è una famiglia numerabile di intervalli aperti a due a due disgiunti. Ne segue che $\sum_k \mathcal{L}^1(I_k) = \mathcal{L}^1(G) \leq \mathcal{L}^1(E) + \mathcal{L}^1(G \setminus E) \leq \frac{4}{3}\mathcal{L}^1(E) = \frac{4}{3}\sum_k \mathcal{L}^1(E \cap I_k)$. Quindi esiste k_0 tale che $\mathcal{L}^1(I_{k_0}) \leq \frac{4}{3}\mathcal{L}^1(E \cap I_{k_0})$.

Sia $A = E \cap I_{k_0}$. Supponiamo che $d \in \mathbb{R}$ sia tale che $(d + A) \cap A = \varnothing$. Allora

$$2\,\mathcal{L}^1(A) = \mathcal{L}^1(d + A) + \mathcal{L}^1(A) = \mathcal{L}^1((d + A) \cup A) \leq \mathcal{L}^1((d + I_{k_0}) \cup I_{k_0}).$$

Poiché $(d + I_{k_0}) \cup I_{k_0} \subset (a_{k_0}, b_{k_0} + d)$ se $d \geq 0$ o $(d + I_{k_0}) \cup I_{k_0} \subset (a_{k_0} + d, b_{k_0})$ se $d \leq 0$, abbiamo che $\mathcal{L}^1((d + I_{k_0}) \cup I_{k_0}) \leq |d| + \mathcal{L}^1(I_{k_0})$. Ne deduciamo allora che

$$2\,\mathcal{L}^1(A) \leq |d| + \frac{4}{3}\mathcal{L}^1(A),$$

ovvero $|d| \geq \frac{2}{3}\mathcal{L}^1(A)$. Ma questo equivale a dire che $(d + A) \cap A \neq \varnothing$ per tutti i $d \in (-\delta, \delta)$ se $\delta = \frac{2}{3}\mathcal{L}^1(A)$, ed implica $(-\delta, \delta) \subset E - E$. □

La dimostrazione del seguente lemma (e quindi quella del teorema successivo) fa uso dell'assioma della scelta.

Lemma 3.7.2 (Vitali)**.** *Esiste un sottoinsieme $E \subset \mathbb{R}$ tale che*

(a) $(E - E) \cap \mathbb{Q} = \{0\}$*;*

(b) $\mathbb{R} = \bigcup_{q \in \mathbb{Q}}(q + E)$*.*

Tale insieme E è non misurabile.

Dimostrazione. Introduciamo in $\mathbb{R}$ la seguente relazione d'equivalenza: $a \sim b$ se $b - a \in \mathbb{Q}$. La classe d'equivalenza di $x \in \mathbb{R}$ è data da $x + \mathbb{Q}$. Costruiamo, usando l'assioma della scelta, un insieme E che contenga uno e un solo elemento per ciascuna classe d'equivalenza. Allora le proprietà **(a)** e **(b)** seguono immediatamente. Per dimostrare che E è non misurabile, è sufficiente osservare che: (i) $\mathcal{L}^1(\mathbb{R}) \leq \sum_{q \in \mathbb{Q}} \mathcal{L}^1_e(q + E)$; (ii) $\mathcal{L}^1_e(q + E) = \mathcal{L}^1_e(E)$ per tutti i $q \in \mathbb{Q}$. Ne segue che $\mathcal{L}^1_e(E) > 0$. Ma allora, se E è misurabile, $E - E$ deve contenere un intervallo centrato nell'origine, contraddicendo la **(a)**. □

Teorema 3.7.3. *Ogni insieme $A \subset \mathbb{R}$ tale che $\mathcal{L}^1_e(A) > 0$ contiene un sottoinsieme non misurabile.*

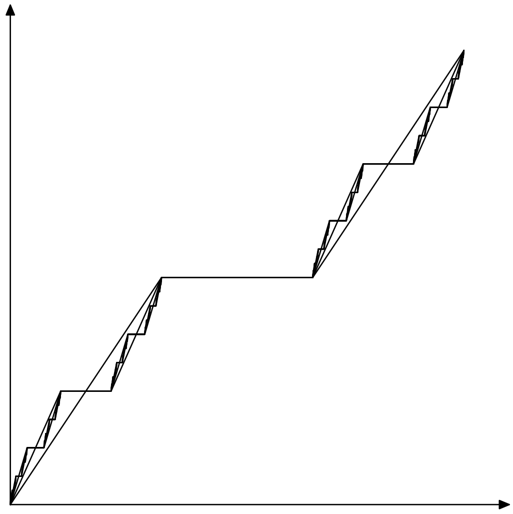

Figura 1. Alcune approssimanti della funzione di Cantor-Lebesgue

Dimostrazione. Consideriamo l'insieme E costruito nel lemma precedente. Allora, posto $A_q = A \cap (q + E)$, si ha che $A = \bigcup_{q\in\mathbb{Q}} A_q$, e quindi $0 < \mathcal{L}^1_e(A) \le \sum_{q\in\mathbb{Q}} \mathcal{L}^1_e(A_q)$. Ne segue che per qualche $q \in \mathbb{Q}$, $\mathcal{L}^1_e(A_q) > 0$. Se A_q è misurabile, $A_q - A_q$ contiene un intervallo centrato nell'origine, contraddicendo il lemma 3.7.2**(a)**. □

Esempio 3.7.4. Mostriamo ora un esempio di una funzione $f\colon [0,1] \to [0,1]$ continua, monotona crescente, tale che $f(E)$ non è misurabile per un E misurabile. La costruzione di tale funzione è strettamente legata alla costruzione dell'insieme di Cantor.

Ricordiamo che l'insieme di Cantor $C = \bigcap_n C_n$, ove $C_n = \bigcup_{k=1}^{2^n} I^k_n$, ove gli I^k_n sono intervalli chiusi di lunghezza 3^{-n}. Poniamo $D_n = [0,1] \setminus C_n = \bigcup_{k=1}^{2^n-1} J^k_n$. Gli insiemi J^k_n sono intervalli aperti. Ad esempio $J^1_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$, $J^1_2 = \left(\frac{1}{9}, \frac{2}{9}\right)$, $J^2_2 = \left(\frac{1}{3}, \frac{2}{3}\right)$, $J^3_2 = \left(\frac{7}{9}, \frac{8}{9}\right)$. Si noti che $J^{2j}_{n+1} = J^j_n$. Definiamo ora una successione di funzioni continue e monotone f_n ponendo

(1) $f_n(0) = 0$ e $f_n(1) = 1$ per ogni n;

(2) $f_n(x) = \frac{j}{2^n}$ nell'intervallo J^j_n, per ogni n e ogni $1 \le j \le 2^n - 1$;

(3) $f_n(x)$ è continua su $[0,1]$ e lineare in $C_n = [0,1] \setminus D_n$.

Si veda in Figura 1 alcune delle funzioni f_n. Notiamo che, per ogni $x \in J^{2k}_{n+1} = J^k_n$, $f_{n+1}(x) = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = f_n(x)$. Quindi $f_m(x) = f_n(x)$ per ogni $m \ge n$ e per ogni $x \in J^k_n$, qualunque sia $k = 1, \dots, 2^n - 1$, cioè per ogni $x \in D_n$. Abbiamo quindi convergenza della successione $f_n(x)$ per tutti gli $x \in \bigcup_n D_n$. Sia ora $x \in I^k_n = \left[a^k_n, b^k_n\right]$. Ovviamente, essendo f_m una funzione monotona, abbiamo che $f_m(a_n^k) \leq f_m(x) \leq f_m(b_n^k)$, qualunque sia m. Supponiamo ora che $m \geq n$. Allora $f_m(a_n^k) = f_n(a_n^k)$ (a_n^k sta in $\overline{D}_n$), e analogamente $f_m(b_n^k) = f_n(b_n^k)$. Quindi $|f_m(x) - f_n(x)| \leq f_n(b_n^k) - f_n(a_n^k) \leq \frac{1}{2^n}$. Riepilogando abbiamo che per tutti gli $m \geq n$

$$\begin{cases} |f_m(x) - f_n(x)| \leq \frac{1}{2^n} & x \in C_n \\ |f_m(x) - f_n(x)| = 0 & x \in D_n. \end{cases}$$

Quindi f_m converge uniformemente in $[0,1]$ ad una funzione f che chiamiamo la funzione di Cantor-Lebesgue. Abbiamo che f è ovviamente continua (è limite uniforme di funzioni continue) e monotona crescente, tale che $f(0) = 0$, $f(1) = 1$. Si osservi poi che, per ogni $x \in J_n^k$, $f(x) = \frac{k}{2^n}$. Quindi $f(D_n) \subset \left\{ x = \frac{k}{2^n} \mid 1 \leq k < 2^n \right\}$. Ne segue che, posto $D = \bigcup_n D_n$, $f(D) = \left\{ \frac{k}{2^n} \mid n \in \mathbb{N}, 1 \leq k < 2^n \right\}$. In particolare $f(D)$ è numerabile e quindi di misura nulla. Poiché $f([0,1]) = [0,1]$, ne deduciamo che $f(C)$ è un insieme compatto, di misura 1 (si ha in effetti che $f(C) = [0,1]$). Quindi f manda l'insieme di misura nulla C in un insieme di misura positiva. Sia $A \subset f(C) \setminus f(D)$ un insieme non misurabile, e poniamo $B = f^{-1}(A)$. Allora $B \subset C$ è ovviamente misurabile (C è di misura nulla), ma $f(B) = A$ non è misurabile. Un'altra particolarità della funzione di Cantor-Lebesgue è che ha derivata nulla in tutti i punti di D, come si può facilmente verificare. Quindi è una funzione che ha derivata nulla in tutti i punti tranne che in un insieme di misura nulla (l'insieme C).

La funzione $f(x)$ di Cantor-Lebesgue si può modificare nel modo seguente: si consideri la funzione

$$\phi(x) = x + f(x).$$

Tale funzione ϕ risulta, come si può facilmente verificare, strettamente crescente e continua. Ne segue che ϕ non solo è continua, ma anche invertibile, con inversa ψ continua. Si ha poi che $\phi([0,1]) = [0,2]$, $\mathcal{L}^n(\phi(D)) = 1$, $\mathcal{L}^n(\phi(C)) = 1$. Quindi anche ϕ manda insiemi di misura nulla in insiemi di misura positiva, e come prima si vede che manda insiemi misurabili (sottoinsiemi dell'insieme C) in insiemi non misurabili. Basta prendere un sottoinsieme A non misurabile di $\phi(C) \setminus \phi(D)$; la controimmagine $B = \psi(A) \subset C$ è chiaramente misurabile.

3.8. Il Paradosso di Banach-Tarski

Vogliamo mostrare in questa sezione un noto paradosso di teoria della misura, il paradosso di Banach-Tarski. Tale paradosso afferma che è possibile dividere la palla di $\mathbb{R}^n$, se $n \geq 3$, in un numero finito di parti disgiunte: $B(0,1) = A_1 \cup \cdots \cup A_N$ tali che $B(0,1) = \mathsf{T}_1 A_1 \cup \cdots \cup \mathsf{T}_m A_m$ e $B(0,1) = \mathsf{T}_{m+1} A_{m+1} \cup \cdots \cup \mathsf{T}_N A_N$ dove $A_i \cap A_j = \varnothing$ e le T_i sono rototraslazioni. Tale paradosso mostra in particolare la non esistenza di una misura *finitamente additiva*, normalizzata e invariante per rototraslazioni in $\mathbb{R}^n$ (se $n \geq 3$).

Definizione 3.8.1. Sia G un gruppo, X un insieme. G agisce su X se per ogni $g \in G$ esiste una funzione $g\colon X \to X$ tale che (i) per tutti i $g, h \in G$ vale $g(h(x)) = (gh)(x)$ per ogni $x \in X$; (ii) $1(x) = x$ per ogni $x \in X$.

Definizione 3.8.2. Supponiamo che il gruppo G agisca su X, e sia E un sottoinsieme in X. E è G-paradossale se esistono $m+n$ sottoinsiemi $A_1, \ldots, A_n$, $B_1, \ldots, B_m$ di E a due a due disgiunti e $m+n$ elementi $g_1, \ldots, g_n$, $h_1, \ldots, h_m$ di G tali che

$$E = \bigcup_{i=1}^{n} g_i(A_i) \qquad E = \bigcup_{j=1}^{m} h_j(B_j).$$

Osservazione 3.8.3. Ovviamente $\bigcup A_i$ e $\bigcup B_j$ sono sottoinsiemi disgiunti di E. È possibile fare in modo che anche $\{g_i(A_i)\}$ e $\{h_j(B_j)\}$ siano collezioni disgiunte.

Teorema 3.8.4 (Paradosso di Banach-Tarski)**.** *Ogni palla di $\mathbb{R}^3$ è paradossale rispetto al gruppo delle isometrie di $\mathbb{R}^3$.*

Osservazione 3.8.5. Sia G un gruppo. Allora G agisce su G per moltiplicazione a sinistra.

Dato un insieme M, ricordiamo che il gruppo libero F generato dall'insieme M è l'insieme delle parole finite formate dalle lettere $\{\sigma, \sigma^{-1} \mid \sigma \in M\}$. Due parole sono equivalenti se si ottengono una dall'altra per cancellazione o addizione di un numero finito di coppie della forma $\sigma\sigma^{-1}$ o $\sigma^{-1}\sigma$. Una parola si dice ridotta se non contiene coppie di questo tipo. Possiamo pensare ad F come all'insieme delle parole ridotte. L'operazione di gruppo è semplicemente la concatenazione di parole. L'identità del gruppo è la parola vuota. Si dice rango del gruppo libero F la cardinalità di M. Due gruppi liberi dello stesso rango sono isomorfi.

Teorema 3.8.6. *Ogni gruppo libero F di rango due è F-paradossale se F agisce su se stesso per moltiplicazione a sinistra.*

Dimostrazione. Siano σ e τ i generatori del gruppo. Se $\rho \in \{\sigma, \sigma^{-1}, \tau, \tau^{-1}\}$ poniamo $w(\rho) = \{h \in F \mid h = \rho h'\}$. Cioè $w(\rho)$ è l'insieme delle parole (ridotte) che iniziano con la lettera ρ. Ovviamente vale che

$$F = \{1\} \cup w(\sigma) \cup w(\sigma^{-1}) \cup w(\tau) \cup w(\tau^{-1}),$$

e la decomposizione è in sottoinsiemi disgiunti. Inoltre abbiamo che

$$F = w(\sigma) \cup \sigma w(\sigma^{-1})$$
$$F = w(\tau) \cup \tau w(\tau^{-1})$$

(infatti, data una parola ridotta v, se non inizia con σ, la parola $\sigma^{-1}v$ è una parola ridotta in $w(\sigma^{-1})$, e quindi appartiene a $\sigma w(\sigma^{-1})$). □

Proposizione 3.8.7. *Se G è un gruppo G-paradossale che agisce su X senza punti fissi (cioè se per tutti i $g \in G$, $g \neq 1$ si ha $g(x) \neq x$), allora X è G-paradossale.*

Per quanto visto nel precedente teorema 3.8.6, ne segue che X è F-paradossale ogni volta che un gruppo libero F di rango due agisce su X senza punti fissi.

Dimostrazione. Supponiamo $A_1, \dots, A_n$, $B_1, \dots, B_m \subset G$ siano insiemi disgiunti, e $g_1, \dots, g_n$, $h_1, \dots, h_m \in G$ siano tali che $G = \bigcup_{i=1}^n g_i(A_i) = \bigcup_{j=1}^m h_j(B_j)$. Consideriamo le orbite di G in X, $G_x = \{ g(x) \mid g \in G \}$. Sono delle classi di equivalenza per la relazione $x \approx y$ se $y = g(x)$ per qualche $g \in G$. Per l'assioma della scelta esiste un insieme M che interseca ciascuna orbita in un solo punto. Ne segue che $X = \bigcup_{g \in G} g(M)$ e che $g(M) \cap g'(M) = \varnothing$ se $g \neq g'$. Sia $A_i^* = \bigcup_{g \in A_i} g(M)$, $B_j^* = \bigcup_{g \in B_j} g(M)$. Si verifica facilmente che $\{A_i^*\} \cup \{B_j^*\}$ è una collezione di sottoinsiemi di X a due a due disgiunti. Poiché ovviamente $X = \bigcup_i g_i(A_i^*) = \bigcup_j h_j(B_j^*)$, la proposizione segue. $\square$

Teorema 3.8.8. *$SO(3)$ ha un sottogruppo libero di rango 2, cioè esistono due rotazioni ρ e ϕ attorno ad assi passanti per l'origine che sono un insieme di generatori liberi del sottogruppo generato da ϕ e ρ.*

Dimostrazione. Sia ϕ la rotazione attorno all'asse z di un angolo di $\arccos 1/3$ e ρ la rotazione attorno all'asse x di un angolo di $\arccos 1/3$. Allora

$$\phi^{\pm 1} = \begin{pmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \rho^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}.$$

Prendiamo una parola ridotta non banale w di lettere $\{\phi, \phi^{-1}, \rho, \rho^{-1}\}$ e mostriamo che non è mai l'identità. Possiamo sempre supporre che w termini con la lettera ϕ o con la lettera ϕ^{-1} (se non è cosí, osserviamo che $w \neq 1$ se e solo se $\phi w \phi^{-1} \neq 1$ e sostituiamo a w la parola $\phi w \phi^{-1}$).

Mostriamo che $w(1,0,0) = (a, b\sqrt{2}, c)/3^k$ se w è di lunghezza k ove $a, b, c \in \mathbb{Z}$ e b non è divisibile per 3. Vedremo che ciò implica che $w(1,0,0) \neq (1,0,0)$ per ogni parola $w \neq 1$.

Dimostriamo l'affermazione fatta per induzione sulla lunghezza della parola w.

Se la lunghezza di w è uno, $w = \phi^{\pm 1}$, e $w(1,0,0) = (1, \pm 2\sqrt{2}, 0)/3$ e l'affermazione è ovvia.

Supponiamo ora l'affermazione vera per le parole di lunghezza $k \leq n$ e dimostriamolo per quelle di lunghezza $n+1$. Consideriamo i casi $w = \phi^{\pm 1} w'$ e $w = \rho^{\pm 1} w'$. Sappiamo, per l'ipotesi di induzione, che $w'(1,0,0) = (a', b'\sqrt{2}, c')/3^n$.

Allora $\phi^{\pm 1} w'(1,0,0) = (a' \mp 4b', (b' \pm 2a')\sqrt{2}, 3c')/3^{n+1}$ e $\rho^{\pm 1} w'(1,0,0) = (3a', (b' \mp 2c')\sqrt{2}, c' \pm 4b')/3^{n+1}$. Ne segue subito che $a, b, c \in \mathbb{Z}$. Mostriamo ora che b non è divisibile per 3. Per fare ciò dobbiamo considerare quattro casi: $w = \phi^{\pm 1} \rho^{\pm 1} v$, $w = \rho^{\pm 1} \phi^{\pm 1} v$, $w = \phi^{\pm 1} \phi^{\pm 1} v$, $w = \rho^{\pm 1} \rho^{\pm 1} v$. Posto $v = (a'', b''\sqrt{2}, c'')$ abbiamo, nel primo caso, $b = b' \pm 2a' = b' \pm 6a''$ e nel secondo caso $b = b' \mp 2c' = b' \mp 6c''$, non divisibili per 3 se b' non lo è. Nel terzo caso $b = b' \pm 2a' = b' \pm 2(a'' \mp 4b'') = b' \pm 2a'' - 8b'' = b' + b'' \pm 2a'' - 9b'' = 2b' - 9b''$, non divisibile per 3 se b' non lo è. Infine, nell'ultimo caso, $b = b' \mp 2c' = b' \mp 2(c'' \pm 4b'') = b' \mp 2c'' - 8b'' = b' + b'' \mp 2c'' - 9b'' = 2b' - 9b''$, di nuovo non divisibile per 3 se b' non lo è. Quindi abbiamo dimostrato l'ipotesi d'induzione.

Mostriamo ora che ciò implica $w(1,0,0) \neq (1,0,0)$ per ogni $w \neq 1$. Se fosse $w(1,0,0) = (1,0,0)$, avremmo che $3^k = a$, $b = 0$, $c = 0$. Allora analizziamo i casi possibili: (i) $0 = b' \pm 2a'$, $3^k = a' \mp 4b'$, da cui $b' = \mp 2a' = \mp 2(3^k \pm 4b') = \mp 2 \cdot 3^k - 8b'$, ovvero $b' = \mp 2 \cdot 3^{k-2}$ e b' è divisibile per 3. Nel caso (ii): $3^k = 3a'$, $0 = b' \mp 2c'$, $0 = c' \pm 4b'$, da cui $b' = \pm 2c' = -8b'$, ovvero $b' = c' = 0$ e anche $w'(1,0,0) = (1,0,0)$, assurdo. $\square$

Esercizi Capitolo 3

(1) Sia $X = \mathbb{Q}$. Dato un intervallo $I \subset \mathbb{Q}$ di estremi razionali $p \leq q$, si definisca la lunghezza di un tale intervallo come $\ell(I) = q - p$. Si definisca poi, per ogni insieme $E \subset \mathbb{Q}$, $\mu^*(E) = \inf\{\sum_1^\infty \ell(I_i) \mid I_i = [p_i, q_i],\ p_i,\ q_i \in \mathbb{Q},\ E \subset \cup I_i\}$. Si dimostri che $\mu^*(E) = 0$ per ogni $E \subset \mathbb{Q}$. In particolare $0 = \mu^*(I) < 1 = \ell(I)$ se I è l'insieme dei razionali fra 0 e 1. Ciò mostra che il fatto che la misura esterna di un intervallo sia uguale alla sua lunghezza è una proprietà di $\mathbb{R}$.

(2) Sia $b > 1$ un intero e $0 < x < 1$. Mostrare che esiste una successione di interi c_k, $0 \leq c_k < b$ tali che $x = \sum_{k=1}^{+\infty} c_k b^{-k}$. Mostrare che tale espansione è unica a meno che $x = cb^{-k}$ per qualche intero $0 \leq c < b$, nel qual caso esistono due espansioni per x.

(3) Se $b = 3$ nell'esercizio 2, l'espansione si chiama triadica o ternaria. Mostrare che l'insieme di Cantor C consiste di tutti gli x che hanno almeno un'espansione ternaria in cui tutti i c_k sono 0 o 2.

Sia $f(x)$ la funzione di Cantor-Lebesgue. Si mostri che se $x \in C$ e $x = \sum c_k 3^{-k}$, con tutti i c_k o 0 o 2, allora $f(x) = \sum \frac{1}{2} c_k 2^{-k}$.

(4) Costruire un insieme di Cantor nel quadrato unitario $\{(x,y) \mid 0 \leq x, y \leq 1\}$ nel modo seguente. Suddividere il quadrato in nove parti uguali e tenere solo i quattro quadrati d'angolo chiusi, rimuovendo la regione rimanente (a forma di croce). Ripetere la procedura in ciascun quadrato rimasto, infinite volte. Mostrare che l'insieme che si ottiene è perfetto, ha misura bidimensionale 0 ed è uguale a $C \times C$.

(5) Costruire un sottoinsieme di $[0,1]$ nella stessa maniera con cui si costruisce il Cantor, ma rimuovendo a ciascun passo dagli intervalli un sottointervallo di lunghezza relativa θ, $0 < \theta < 1$. Mostrare che anche in questo caso si ottiene un insieme perfetto di misura 0.

(6) Costruire un sottoinsieme di $[0,1]$ nella stessa maniera con cui si costruisce il Cantor, ma rimuovendo al passo k da ciascun intervallo rimasto un sottointervallo di lunghezza relativa $\delta 3^{-k}$, $0 < \delta < 1$. Mostrare che in questo caso si ottiene un insieme perfetto di misura $1 - \delta$ che non contiene alcun intervallo.

(7) Costruire un insieme di Borel $A \subset \mathbb{R}$ tale che $0 < \mathcal{L}^1(A \cap I) < \mathcal{L}^1(I)$ per ogni intervallo non vuoto I.

(8) Sia $Z \subset \mathbb{R}$, $\mathcal{L}^n(Z) = 0$. Mostrare che $\mathcal{L}^n(\{\, x^2 \mid x \in Z \,\}) = 0$.

(9) Definiamo la misura *interna* di un sottoinsieme $E \subset \mathbb{R}^n$ come

$$\mu(E)_i = \sup\{\, \mathcal{L}^n(F) \mid F \text{ chiuso } \subset E \,\}$$

Dimostrare che

(a) $\mu(E)_i \leq \mathcal{L}^n_e(E)$;

(b) se $\mathcal{L}^n_e(E) < +\infty$, E è misurabile se e solo se $\mu(E)_i = \mathcal{L}^n_e(E)$.

(c) mostrare che 9b è in generale falso se $\mathcal{L}^n_e(E) = +\infty$.

(10) Abbiamo visto che $E - E$ contiene un intervallo se E è un sottoinsieme di $\mathbb{R}$ di misura positiva. Si mostri che $C - C = [0, 1]$ se C è l'insieme di Cantor.

(11) Mostrare che $\mathcal{L}^n(P) = \mathrm{vol}(P)$ se P è un parallelepipede.

Capitolo 4

Funzioni Misurabili

4.1. Funzioni misurabili

Avremo a che fare, nel seguito, con funzioni a valori nella retta reale estesa $[-\infty, \infty] = \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. In $\overline{\mathbb{R}}$ assumeremo l'usuale topologia. Ricordiamo che per tale topologia gli intorni di $a \in \mathbb{R}$ sono gli intervalli aperti contenenti a, mentre gli intorni di $+\infty$ sono gli insiemi della forma $(a, +\infty]$, $a \in \mathbb{R}$ e quelli di $-\infty$ sono gli insiemi $[-\infty, a)$, $a \in \mathbb{R}$. Ricordiamo che la somma $\infty - \infty$ non è definita, mentre assumeremo che $0 \cdot \pm\infty = 0$. Ne segue che il prodotto di due elementi di $\overline{\mathbb{R}}$ è sempre ben definito (anche se non risulta piú continuo). Tale assunzione ci permetterà di semplificare l'enunciato di molti teoremi.

Definizione 4.1.1. Sia (X, Σ) uno spazio misurabile e $f\colon X \to [-\infty, +\infty]$. Diciamo che f è una funzione misurabile (rispetto alla sigma-algebra Σ) se per ogni numero reale t l'*insieme di livello*

$$S_f(t) := \left\{ x \in X \;\middle|\; f(x) > t \right\} \tag{4.1.1}$$

è misurabile, cioè $S_f(t) \in \Sigma$. Si noti che la definizione di misurabilità non richiede l'esistenza di una misura.

Nel caso in cui X è uno spazio topologico e $\Sigma = \mathcal{B}$ è la sigma-algebra dei boreliani diciamo che f è una funzione misurabile secondo Borel o anche Borel-misurabile.

Piú generalmente, se $f\colon X \to \mathbb{C}$ è una funzione in X a valori complessi, diciamo che f è misurabile (Borel-misurabile) se le sue parti reale e immaginaria, $\Re f$ e $\Im f$, sono misurabili (Borel misurabili).

Osservazione 4.1.2. Poiché la sigma algebra dei misurabili secondo Lebesgue in $\mathbb{R}^n$ contiene strettamente la sigma algebra dei boreliani, vi sono funzioni misurabili da $\mathbb{R}$ in $\overline{\mathbb{R}}$ che non sono Borel misurabili (si veda 4.1.6).

Dimostriamo innanzi tutto alcune semplici proprietà delle funzioni misurabili:

Teorema 4.1.3. *Sia* (X, Σ) *uno spazio misurabile, e* $f \colon X \to \overline{\mathbb{R}}$.
Allora

(a) *La classe di insiemi* $E \subset \overline{\mathbb{R}}$ *tali che* $f^{-1}(E) \in \Sigma$ *è una sigma-algebra;*

(b) *Sia* $\mathcal{C}$ *una classe di sottoinsiemi di* $\overline{\mathbb{R}}$. *Assumiamo che (i) la sigma-algebra* $\mathcal{S}$ *generata da* $\mathcal{C}$ *contiene i boreliani; (ii)* $f^{-1}(E) \in \Sigma$ *per ogni* $E \in \mathcal{C}$. *Allora* f *è misurabile. In particolare* f *è misurabile se e solo se* $f^{-1}(G) \in \Sigma$ *per ogni* $G \subset \overline{\mathbb{R}}$ *aperto.*

(c) *Se* f *è misurabile,* $f^{-1}(E) \in \Sigma$ *per ogni boreliano* E*; quindi* f *è misurabile se e solo se* $f^{-1}(E)$ *è misurabile per* ogni *boreliano* $E \subset \overline{\mathbb{R}}$

Dimostrazione. (a): Sia $\mathcal{A} = \left\{ E \subset \overline{\mathbb{R}} \;\middle|\; f^{-1}(E) \in \Sigma \right\}$. Allora, da $f^{-1}(E^c) = (f^{-1}(E))^c$ segue subito che $E \in \mathcal{A}$ implica $E^c \in \mathcal{A}$, e da $f^{-1}(\bigcup_k E_k) = \bigcup_k f^{-1}(E_k)$ segue subito che $E_k \in \mathcal{A}$ implica $\bigcup_k E_k \in \mathcal{A}$. Quindi $\mathcal{A}$ è una sigma algebra.

(b): Basta osservare che, per il punto **(a)**, la classe $\mathcal{A}$ dei sottoinsiemi E di $\overline{\mathbb{R}}$ per cui $f^{-1}(E) \in \Sigma$ forma una sigma-algebra, che contiene $\mathcal{C}$, e quindi $\mathcal{S}$ e quindi tutti i boreliani. Poiché in particolare sono boreliani le semirette $(t, +\infty]$, abbiamo che $f^{-1}((t, \infty]) = S_f(t)$ è misurabile per ogni $t \in \mathbb{R}$, ovvero è misurabile f.

(c): Basta osservare che, per il punto **(a)**, la classe $\mathcal{A}$ dei sottoinsiemi E di $\overline{\mathbb{R}}$ per cui $f^{-1}(E) \in \Sigma$ forma una sigma-algebra, che contiene, per definizione di funzione misurabile, le semirette della forma $(t, +\infty]$. Ma la piú piccola sigma-algebra con tale proprietà è quella dei boreliani (si veda l'esercizio 6 a pagina 31), che quindi è contenuta in $\mathcal{A}$. $\square$

Introduciamo la seguente notazione: $\{a \le f \le b\} := \left\{ x \in X \;\middle|\; a \le f(x) \le b \right\}$, $\{f > a\} := \left\{ x \in X \;\middle|\; f(x) > a \right\}$, e analoghe. Si noti che, con questa notazione, l'insieme di livello $S_f(t) = \{f > t\}$). Se f è una funzione misurabile, segue dalla definizione che $\{f > t\}$ è misurabile. In realtà molti altri insiemi risultano essere misurabili, come segue dal corollario seguente.

Corollario 4.1.4. *Sia* (X, Σ) *uno spazio misurabile, e* $f \colon X \to \overline{\mathbb{R}}$. *Allora*

(a) f *è misurabile, se e solo se è misurabile, per ogni* $t \in \mathbb{Q}$*, l'insieme* $\{f \ge t\}$ *(oppure l'insieme* $\{f < t\}$ *oppure l'insieme* $\{f \le t\}$ *etc.);*

(b) *Se f è misurabile, sono misurabili i seguenti sottoinsiemi di X: $\{f > -\infty\}$, $\{f < +\infty\}$, $\{a \le f \le b\}$, $\{f = a\}$, etc.*

Dimostrazione. (a): Per il teorema 4.1.3**(b)**, basta verificare che i sottoinsiemi $\mathcal{C} = \{ (t, +\infty] \mid t \in \mathbb{Q} \}$ generano la sigma algebra dei Boreliani, fatto di facile verifica.

(b) segue dal teorema 4.1.3**(c)**, in quanto tutti questi insiemi sono Boreliani. □

Ovviamente, avremo a che fare soprattutto con funzioni $f\colon X \to \overline{\mathbb{R}}$, dove $X \subset \mathbb{R}^n$ nel caso in cui Σ è la sigma-algebra dei misurabili secondo Lebesgue. In questo caso la misura è una misura di Borel, e la sigma-algebra contiene quella dei Boreliani. Per tali sigma-algebre sono misurabili le funzioni continue. Infatti vale:

Teorema 4.1.5. *Supponiamo che la sigma-algebra Σ in X contenga gli aperti di X. Sia $f\colon X \to [-\infty, +\infty]$. Allora:*

(a) *f continua implica f misurabile secondo Borel;*

(b) *Se $X \subset \mathbb{R}^n$ e f è semicontinua superiormente o inferiormente allora f è misurabile secondo Borel (per la definizione di funzione semicontinua superiormente e di funzione semicontinua inferiormente si veda la definizione 1.5.2 a pagina 12).*

Ricordiamo che f misurabile secondo Borel implica f misurabile.

Dimostrazione. Il teorema segue immediatamente dalle osservazioni seguenti: (i) se Σ contiene gli aperti allora contiene i Boreliani; (ii) se f è continua o semicontinua inferiormente, allora $S_f(t) = f^{-1}((t, +\infty))$ è aperto (si veda la definizione 1.5.2); (iii) se f è semicontinua superiormente, allora $\{f < t\}$ è aperto. □

Osservazione 4.1.6. Come segue chiaramente dalla **(b)** del teorema 4.1.3, la definizione di funzione misurabile $f\colon X \to \overline{\mathbb{R}}$ è stata data usando una sigma-algebra in X e la topologia di $\overline{\mathbb{R}}$. Sempre nello stesso teorema 4.1.3 abbiamo visto che f è misurabile se e solo se $f^{-1}(E)$ è misurabile per ogni *boreliano* E. Mostriamo ora che esiste una funzione misurabile $f\colon \mathbb{R} \to \mathbb{R}$ (addirittura continua) ed un sottoinsieme $E \subset \overline{\mathbb{R}}$ misurabile (ma non boreliano) tale che $f^{-1}(E)$ non è misurabile.

Consideriamo la funzione $f\colon [0,1] \to [0,1]$ di Cantor-Lebesgue (costruita nell'esempio 3.7.4) e la funzione $\phi(x) = x + f(x)$. Tale ϕ, come abbiamo già detto quando abbiamo costruito la funzione di Cantor-Lebesgue, è continua, strettamente crescente e manda un insieme misurabile $A \subset C$ (C è l'insieme di Cantor) in un sottoinsieme non misurabile B (basta prendere prendere B non misurabile in $\phi(C) \setminus \phi(D)$ e porre $A = \phi^{-1}(B)$). Ma allora la funzione

ψ, inversa della ϕ, (che esiste continua e strettamente monotona in quanto ϕ è continua e strettamente monotona), è tale che $\psi^{-1}(A) = B$, mostrando che la funzione continua (e quindi Borel-misurabile e misurabile) ψ ha la proprietà cercata. Poiché sappiamo che $\psi^{-1}(E)$ è un boreliano per ogni E boreliano, ne deduciamo che A è misurabile ma non boreliano.

Notiamo che a questo punto è semplice costruire una funzione misurabile ma non Borel misurabile: basta prendere la funzione caratteristica χ_A di un insieme A misurabile ma non boreliano.

Vedremo fra poco che due funzioni che differiscono su di un insieme di misura nulla sono indistinguibili dal punto di vista dell'integrazione. Ciò porta a dare una definizione piú generale di funzione misurabile.

Definizione 4.1.7. Sia (X, Σ, μ) uno spazio di misura, e $X' \subset X$. Diciamo che $f \colon X' \to \overline{\mathbb{R}}$ è misurabile in X se: (i) X' è misurabile, e $\mu(X \setminus X') = 0$; (ii) $S_f(t) = \{\, x \in X' \mid f(x) > t \,\} \in \Sigma$ per ogni $t \in \mathbb{R}$.

Osservazione 4.1.8. Si osservi che se $f \colon X' \to \overline{\mathbb{R}}$ è misurabile in X, allora la funzione

$$\tilde{f}(x) := \begin{cases} f(x) & x \in X' \\ 0 & x \in X \setminus X' \end{cases}$$

è misurabile nel senso usuale. Nel caso in cui la misura è completa (come nel caso della misura di Lebesgue in $\mathbb{R}^n$), *ogni* funzione $\tilde{f}$ uguale a f in X' è misurabile.

Osservazione 4.1.9. Il fatto che una funzione misurabile possa essere definita solo quasi ovunque può creare difficoltà con alcuni concetti, per esempio con quello di funzione strettamente positiva. Per rimediare a ciò diciamo che una funzione nonnegativa f è una *funzione misurabile strettamente positiva* su di un insieme misurabile A se l'insieme $\{\, x \in A \mid f(x) = 0 \,\}$ ha misura nulla.

Difficoltà analoghe sorgono nella definizione di supporto di una funzione misurabile. Sia data una misura di Borel μ nello spazio topologico X, e una funzione misurabile f. Ricordiamo che gli aperti di X sono misurabili, cioè fanno parte della sigma-algebra. Consideriamo la collezione $\mathcal{O}$ degli *aperti* G con la proprietà che $f(x) = 0$ per μ-quasi ogni $x \in G$, e sia G^* l'aperto unione di tutti i $G \in \mathcal{O}$. Si noti che $\mathcal{O}$ e G^* possono essere vuoti. Definiamo allora il supporto essenziale di f, $\operatorname{ess\,supp}\{f\}$, essere il complemento di G^*. Quindi $\operatorname{ess\,supp}\{f\}$ è un insieme chiuso e quindi misurabile. Consideriamo, per esempio, la funzione di Dirichlet f in $\mathbb{R}$ definita da $f(x) = 1$ se x è razionale e $f(x) = 0$ se x non è razionale. Ovviamente $f(x) = 0$ per q.o. $x \in \mathbb{R}$ (se $\mu = \mathcal{L}^n$) e quindi $\operatorname{ess\,supp}\{f\} = \varnothing$. Si noti anche che $\operatorname{ess\,supp}\{f\}$ dipende anche dalla misura μ e non solo dalla sigma-algebra. È un semplice esercizio verificare che se μ è la misura di Lebesgue e f una funzione continua, allora $\operatorname{ess\,supp}\{f\}$ coincide con $\operatorname{supp}\{f\}$ come definito nella sezione 1.5.

Nel resto di questo libro useremo, per semplicità, la notazione $\operatorname{supp}\{f\}$ anche per indicare $\operatorname{ess\,supp}\{f\}$.

4.2. L'insieme delle funzioni misurabili

In questo paragrafo analizziamo alcune delle piú importanti proprietà dell'insieme delle funzioni misurabili. In tutta questa sezione assumeremo, senza specificarlo sempre, che X è uno spazio misurabile, con sigma-algebra Σ.

Teorema 4.2.1. *Sia f una funzione misurabile in X, e $\phi\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ una funzione Borel-misurabile. Se $h = \phi \circ f$ è ben definita (almeno quasi ovunque), allora è misurabile in X.*

Dimostrazione. Osserviamo innanzi tutto che, poiché f e ϕ potrebbero essere definite solo quasi ovunque, la richiesta "h ben definita" è necessaria (ad esempio, se $f(x) = 0$ in tutti i punti $x \in X$, e ϕ non è definita in 0, h non è mai definita). Assumiamo h definita in tutto X. Sia $E \subset \overline{\mathbb{R}}$ un sottoinsieme Borel-misurabile. Allora l'insieme $\phi^{-1}(E)$ è un boreliano (ϕ è *Borel*-misurabile) e quindi, per il teorema 4.1.3**(c)** è misurabile l'insieme $h^{-1}(E) = f^{-1}(\phi^{-1}(E))$. Il caso generale segue in modo analogo. $\square$

Una conseguenza immediata è la seguente (si noti che qui le funzioni $|f(x)|$, $f^+(x)$ e $f^-(x)$ sono definite in tutti i punti in cui è definita $f(x)$).

Corollario 4.2.2. *Se f è misurabile in X, sono anche misurabili:*

(1) *la funzione $x \mapsto |f(x)|$;*
(2) *le funzioni $x \mapsto f^+(x) := \max\{f(x), 0\}$ e $x \mapsto f^-(x) := (-f)^+(x) = \max\{-f(x), 0\}$. Si noti che $f^-(x) \geq 0$, e che vale $f(x) = f^+(x) - f^-(x)$, $|f(x)| = f^+(x) + f^-(x)$.*

Teorema 4.2.3. *Siano f e g due funzioni misurabili in X. Allora è misurabile l'insieme $\{f > g\} = \{\, x \in X \mid f(x) > g(x) \,\}$.*

Dimostrazione. Basta osservare che

$$\{f > g\} = \bigcup_{q \in \mathbb{Q}} \{f > q > g\} = \bigcup_{q \in \mathbb{Q}} \{f > q\} \cap \{g < q\}.$$

$\square$

Teorema 4.2.4. *Siano f e g due funzioni misurabili in X, α, $\beta \in \mathbb{R}$. Supponiamo che $h = \alpha f + \beta g$ sia ben definita (almeno quasi ovunque). Allora h è misurabile.*

Dimostrazione. Supponiamo che h sia definita per ogni $x \in X$, e dividiamo la dimostrazione in vari passi:

Passo 1: $f + \lambda$ è misurabile. Basta osservare che $S_{f+\lambda}(t) = \{f + \lambda > t\} = \{f > t - \lambda\}$.

Passo 2: λf è misurabile. Se $\lambda = 0$ non vi è nulla da dimostrare; in caso contrario basta osservare che $S_{\lambda f}(t) = \{\lambda f > t\} = \{f > t/\lambda\}$ se $\lambda > 0$ e $S_{\lambda f}(t) = \{f < t/\lambda\}$ se $\lambda < 0$.

Passo 3: $f + g$ è misurabile. Osserviamo che $S_{f+g}(t) = \{f + g > t\} = \{f > t - g\}$. Per i passi 1 e 2 è misurabile la funzione $t - g$. Basta allora applicare il teorema 4.2.3.

Il caso generale è lasciato al lettore. □

Teorema 4.2.5. *Siano f e g due funzioni misurabili in X. Allora $h = fg$ è misurabile.*

Dimostrazione. Ricordiamo innanzi tutto che abbiamo convenuto che $0 \cdot \infty = 0$, e che quindi il prodotto di due numeri in $\overline{\mathbb{R}}$ è sempre ben definito. Sia $Y \subset X$ l'insieme in cui f e g hanno valori finiti. Y è ovviamente misurabile, e $4fg = (f + g)^2 - (f - g)^2$ in Y. Ne segue facilmente che fg è misurabile in Y. In $X \setminus Y$ almeno una delle due funzioni non è finita. I possibili valori di fg in $X \setminus Y$ sono $\{-\infty, 0, +\infty\}$. Poiché $(fg)^{-1}(+\infty) = (\{f = +\infty\} \cap \{g > 0\}) \cup (\{0 < f < +\infty\} \cap \{g = +\infty\}) \cup (\{-\infty < f < 0\} \cap \{g = -\infty\}) \cup (\{f = -\infty\} \cap \{g < 0\})$ si ha subito che $(fg)^{-1}(+\infty)$ è misurabile. Analogamente si mostra la misurabilità di $(fg)^{-1}(-\infty)$. Poiché $(fg)^{-1}(0) \cap (X \setminus Y) = (\{f = 0\} \cup \{g = 0\}) \cap (X \setminus Y)$, si dimostra facilmente che fg è misurabile. □

Quindi l'insieme delle funzioni misurabili è uno spazio vettoriale e un'algebra (cioè è chiuso sia per combinazioni lineari che per prodotto dei suoi elementi). Mostriamo ora che tale insieme è chiuso anche rispetto alla convergenza puntuale.

Teorema 4.2.6. *Sia $f_1, f_2, \ldots$ una successione di funzioni misurabili in X. Allora sono misurabili:*

(a) *Le funzioni $\sup_k f_k$ e $\inf_k f_k$ definite da $x \mapsto (\sup_k f_k)(x) := \sup_k f_k(x)$ e $x \mapsto (\inf_k f_k)(x) := \inf_k f_k(x)$;*

(b) *Le funzioni $\liminf_k f_k$ e $\limsup_k f_k$ definite da $x \mapsto (\liminf_k f_k)(x) := \liminf_k f_k(x)$ e $x \mapsto (\limsup_k f_k)(x) := \limsup_k f_k(x)$.*

Inoltre, se $f_k(x) \to f(x)$ quasi ovunque, allora f è misurabile in X.

Dimostrazione. **(a)**: Basta osservare che $\{(\sup_k f_k) > t\} = \bigcup_k \{f_k > t\}$.

(b): Poiché $\liminf_k f_k = \sup_n(\inf_{k\geq n} f_k)$, basta utilizzare due volte il punto **(a)**.

Nel caso in cui $f_k \to f$ quasi ovunque, la misurabilità di f segue dal punto **(b)** ricordando che in tal caso $f = \liminf_k f_k = \limsup_k f_k$ (nell'insieme in cui vi è convergenza). □

4.3. Approssimazione di funzioni misurabili

Vogliamo esaminare, in questa sezione, il legame fra le funzioni misurabili e altre classi di funzioni, in particolare il legame fra le funzioni misurabili e le funzioni continue.

Iniziamo analizzando una classe molto particolare di funzioni, che, come vedremo, giocherà un ruolo fondamentale nella teoria dell'integrazione.

Definizione 4.3.1. Una funzione $s\colon X \to \mathbb{R}$ si dice una funzione semplice se il codominio di s è un insieme finito.

Se s è una funzione semplice, la rappresentiamo in modo canonico scrivendo

$$s(x) = \sum_{i=1}^{N} s_i \chi_{A_i}(x),$$

dove $s(X) = \{s_1, s_2, \dots, s_N\}$, $s_i \neq s_j$, e $A_i = s^{-1}(s_i)$. Si noti che ne segue $A_i \cap A_j = \varnothing$ se $i \neq j$ e che questo non è l'unico modo di rappresentare s come combinazione lineare di funzioni caratteristiche. Si vede subito che s è misurabile se e solo se gli A_i della rappresentazione canonica sono tutti misurabili.

Vale il seguente teorema

Teorema 4.3.2. *Sia $f\colon X \to \overline{\mathbb{R}}$. Allora esiste una successione $s_1, s_2, \dots$ di funzioni semplici tale che $s_n(x) \to f(x)$ in tutti i punti $x \in X$.*

Inoltre

(a) *se f è misurabile s_n è misurabile;*

(b) *se $f \geq 0$, $s_n(x) \leq s_{n+1}(x)$ per ogni $x \in X$;*

(c) *se $|f(x)| \leq K$ per tutti gli x, $s_n(x) \to f(x)$ uniformemente.*

Dimostrazione. Supponiamo innanzi tutto $f \geq 0$. Fissiamo $n \in \mathbb{N}$ e poniamo $F_n = \{\, x \in X \mid f(x) \geq n \,\}$. Poi, per ogni $j = 1, 2, \dots, n2^n$, poniamo $E_{n,j} = \{\, x \in X \mid (j-1)/2^n \leq f(x) < j/2^n \,\}$. Gli insiemi $E_{n,j}$ e F_n sono ovviamente a due a due disgiunti, e misurabili se f è misurabile.

Definiamo allora

$$s_n(x) = \sum_{j=1}^{n2^n} \frac{(j-1)}{2^n} \chi_{E_{n,j}}(x) + n\chi_{F_n}.$$

Una tale s_n è chiaramente una funzione semplice (già in forma canonica), misurabile se f è misurabile. Mostriamo che converge puntualmente a $f(x)$. Se $f(x) = +\infty$, $s_n(x) = n$ per ogni n e ovviamente $s_n(x) \to f(x)$. Se $f(x)$ è finito, diciamo $f(x) < m$, allora, se $n > m$ e $(j-1)/2^n \leq f(x) < j/2^n$ per un qualche $j \in \{1, \ldots, n2^n\}$, abbiamo che $0 \leq f(x) - s_n(x) \leq 1/2^n$, da cui la convergenza anche nei punti in cui $f(x)$ è finita e la convergenza uniforme se f è limitata. Si verifica poi facilmente che $s_n(x) \leq s_{n+1}(x)$ (in effetti $s_{n+1}(x) - s_n(x)$ vale 0 o $1/2^{n+1}$ se $f(x) \leq n$).

Il caso in cui f non sia a valori nonnegativi si ottiene dal precedente considerando la decomposizione $f = f^+ - f^-$ di f nelle sue parti positive e negative. □

Presentiamo ora un risultato preliminare che ci servirà per legare convergenza puntuale e convergenza uniforme.

Lemma 4.3.3. *Sia $f_1, f_2, \ldots$ una successione di funzioni misurabili in $X \subset \mathbb{R}^n$. Supponiamo che*

1. $f_k(x) \to f(x)$ *per quasi ogni* $x \in X$;

2. $|f(x)| < +\infty$ *per quasi ogni* $x \in X$;

3. $\mathcal{L}^n(X) < +\infty$.

Allora, per ogni $\epsilon, \eta > 0$ esiste un compatto $K_{\epsilon,\eta} \subset X$ e $N_{\epsilon,\eta} \in \mathbb{N}$ tali che

(a) $\mathcal{L}^n(X \setminus K_{\epsilon,\eta}) < \epsilon$;

(b) $|f_k(x) - f(x)| < \eta$ *per ogni* $x \in K_{\epsilon,\eta}$ *e per ogni* $k \geq N_{\epsilon,\eta}$.

Dimostrazione. Fissiamo ϵ e $\eta > 0$, e definiamo

$$X_n = \left\{ x \in X \;\middle|\; |f(x) - f_k(x)| < \eta \text{ per tutti i } k \geq n \right\}.$$

Poiché $X_n = \bigcap_{k \geq n} \left\{ x \in X \;\middle|\; |f(x) - f_k(x)| < \eta \right\}$, è immediato verificare che X_n è misurabile per ogni n. Inoltre $X_1, X_2, \ldots$ è una successione crescente di insiemi, e $x \in \bigcup_n X_n$ se $f_k(x) \to f(x)$ con $f(x)$ finito. Sia X' l'insieme in cui $f_k(x)$ converge ad $f(x)$ finito. Ovviamente $\bigcup X_n = X'$ e dalle ipotesi segue che $\mathcal{L}^n(X \setminus X') = 0$. Abbiamo subito che $\mathcal{L}^n(X_n) \to \mathcal{L}^n(X') = \mathcal{L}^n(X)$; e dalla finitezza della misura di X si deduce che $\mathcal{L}^n(X \setminus X_n) \to 0$. Esiste quindi $N \in \mathbb{N}$ tale che $\mathcal{L}^n(X \setminus X_n) < \frac{\epsilon}{2}$ per ogni $n \geq N$. Dalla regolarità interna della misura di Lebesgue (si veda il corollario 3.4.3) segue poi che per ogni

$n \geq K$ possiamo trovare un compatto $K_n \subset X_n$ tale che $\mathcal{L}^n(X_n \setminus K_n) < \frac{\epsilon}{2}$. È facile allora verificare che K_N è l'insieme cercato. □

Osservazione 4.3.4. Il lemma è falso senza le ipotesi di finitezza della misura di X (si consideri la successione $f_n(x) = \chi_{\{|x| \geq n\}}(x) \to 0$) o di finitezza di $f(x)$ (si consideri la successione $f_n(x) = n \to +\infty$ in $[0,1]$).

Nel caso in cui X non sia un sottoinsieme di $\mathbb{R}^n$, ma uno spazio di misura qualunque, il teorema continua a valere con la sola differenza che non sappiamo se l'insieme $K_{\epsilon,\eta}$ è compatto (non assumiamo nemmeno l'esistenza di una topologia). Il teorema varrà nella stessa forma enunciata sopra in tutti gli spazi con misure di Borel internamente regolari (si veda la definizione nel corollario 3.4.3).

Teorema 4.3.5 (Teorema di Egorov)**.** *Sia $X \subset \mathbb{R}^n$ un sottoinsieme misurabile, e sia $f_1, f_2, \dots$ una successione di funzioni misurabili in X. Supponiamo che*

1. $f_k(x) \to f(x)$ *per quasi ogni* $x \in X$;

2. $|f(x)| < +\infty$ *per quasi ogni* $x \in X$;

3. $\mathcal{L}^n(X) < +\infty$.

Allora, per ogni $\epsilon > 0$ esiste un compatto $K_\epsilon \subset X$ tale che

(a) $\mathcal{L}^n(X \setminus K_\epsilon) < \epsilon$;

(b) $f_k(x) \to f(x)$ *uniformemente in* K_ϵ.

Dimostrazione. Fissiamo $\epsilon > 0$, e scegliamo $m \in \mathbb{N}$. Per il lemma precedente troviamo $K_m := K_{\epsilon/2^m,1/m}$ e $N_m := N_{\epsilon/2^m,1/m}$ tali che $\mathcal{L}^n(X \setminus K_m) < \epsilon/2^n$ e $|f_k(x) - f(x)| < 1/m$ per ogni $x \in K_m$ e $k \geq N_m$. Poniamo infine $K = \bigcap_m K_m$. Ovviamente tale K dipende solo da ϵ. Si ha che

$$\mathcal{L}^n(X \setminus K) \leq \sum_m \mathcal{L}^n(X \setminus K_m) \leq \epsilon.$$

Mostriamo ora che vi è convergenza uniforme in K. Sia $\delta > 0$ assegnato, e consideriamo $x \in K$. Preso $m_0 \in \mathbb{N}$ tale che $1/m_0 < \delta$, da $x \in K_m$ per ogni m abbiamo che $x \in K_{m_0}$, da cui $|f(x) - f_k(x)| < 1/m_0 < \delta$ per ogni $k \geq N_m$. □

Per dimostrare il prossimo teorema abbiamo bisogno di alcuni risultati preliminari, la cui dimostrazione è semplice in $\mathbb{R}^n$.

Lemma 4.3.6. *Sia $K \subset V \subset \mathbb{R}^n$, con K compatto e V aperto.*

Allora esiste un aperto V_1 tale che

(a) $K \subset V_1 \subset \overline{V_1} \subset V$;

(b) $\overline{V_1}$ *è compatto.*

Dimostrazione. Mostriamo che esiste un $\delta > 0$ tale che

$$N_\delta(K) := \left\{ x \in \mathbb{R}^n \;\middle|\; \operatorname{dist}(x, K) < \delta \right\} \subset V.$$

Se non fosse vero, per ogni $k \in \mathbb{N}$ esisterebbe $x_k \in N_{1/k}(K)$, $x_k \notin V$. Ma allora possiamo trovare una successione $y_1, y_2, \ldots \in K$ tale che $\operatorname{dist}(x_k, y_k) < 1/k$. Essendo K compatto, esiste una sottosuccessione $y_{k_j} \to y \in K$. Ma allora anche $\operatorname{dist}(x_{k_j}, y) \leq \operatorname{dist}(x_{k_j}, y_{k_j}) + \operatorname{dist}(y_{k_j}, y) \to 0$, cioè $x_{k_j} \to y$, contraddizione con il fatto che $x_k \notin V$ e $y \in K \subset V$. Si verifica poi immediatamente che $\overline{N_\delta}$ è compatto (è ovviamente chiuso e limitato), e che $\overline{N_\delta} \subset V$ (eventualmente prendendo δ un poco piú piccolo). □

Osservazione 4.3.7. È chiaro dalla dimostrazione che il lemma precedente continua a valere negli spazi metrici in cui le palle hanno chiusura compatta (ciò implica che anche gli insiemi $N_\delta(K)$ hanno chiusura compatta se K è compatto). In realtà il lemma sopra continua a valere anche in spazi topologici di Hausdorff localmente compatti (cioè tali che ogni punto ha un intorno la cui chiusura è compatta).

Lemma 4.3.8 (Lemma di Urysohn)**.** *Sia $K \subset V \subset \mathbb{R}^n$, con K compatto e V aperto.*

Allora esiste una funzione $g \in C_c(\mathbb{R}^n)$ tale che

(a) *$g(x) = 1$ per ogni $x \in K$;*

(b) *$g(x) = 0$ per ogni $x \notin V$.*

Dimostrazione. Per il lemma precedente, esiste un aperto V_1, la cui chiusura è compatta, tale che $K \subset V_1 \subset \overline{V_1} \subset V$. Definiamo allora $g\colon \mathbb{R}^n \to \mathbb{R}$ ponendo

$$g(x) = \frac{\operatorname{dist}(x, V_1^c)}{\operatorname{dist}(x, V_1^c) + \operatorname{dist}(x, K)}.$$

La verifica che g ha le proprietà richieste è immediata (si veda anche l'esercizio 4). □

Osservazione 4.3.9. Il lemma di Urysohn può essere dimostrato anche in spazi topologici di Hausdorff localmente compatti. Si veda, per esempio [5].

Siamo ora in grado di enunciare e dimostrare il teorema di Lusin

Teorema 4.3.10 (Teorema di Lusin)**.** *Sia $f\colon X \to \overline{\mathbb{R}}$, $X \subset \mathbb{R}^n$. Supponiamo che*

1. *f è misurabile in X;*

2. *$|f(x)| < +\infty$ per quasi ogni $x \in X$;*

3. *Esiste $A \subset X$, A misurabile e di misura finita tale che $f(x) = 0$ per ogni $x \notin A$.*

Allora, per ogni $\epsilon > 0$ esiste una funzione $g \in C_c(\mathbb{R}^n)$ tale che

(a) $\mathcal{L}^n(\{\, x \in X \mid f(x) \neq g(x) \,\}) < \epsilon$;

(b) $\max_{\mathbb{R}^n}|g(x)| \leq \sup_X|f(x)|$.

Dimostrazione. Innanzi tutto possiamo supporre f definita in tutto $\mathbb{R}^n$ semplicemente ponendo $f(x) = 0$ per ogni $x \notin X$. È chiaro che il teorema generale allora segue da questo caso. Notiamo anche che, trovata $\tilde{g} \in C_c(\mathbb{R}^n)$ che soddisfa al punto **(a)**, il punto **(b)** segue subito definendo

$$g(x) = \begin{cases} \tilde{g}(x) & \text{se } |\tilde{g}(x)| \leq R \\ \frac{\tilde{g}(x)}{|\tilde{g}(x)|}R & \text{se } |g(x)| > R \end{cases}$$

dove $R = \sup|f(x)|$. Infatti $|g(x)| \leq R$ per tutti gli x, e $g(x)$ differisce da $\tilde{g}(x)$ solo dove $|\tilde{g}(x)| > R$, e per tali x necessariamente $\tilde{g}(x) \neq f(x)$. Dividiamo la dimostrazione dell'esistenza di g che soddisfa al punto **(a)** in vari passi.

Passo 1: Il teorema vale se $f = \chi_A$. Fissiamo $\epsilon > 0$. Sappiamo, dal corollario 3.4.3, usando anche il fatto che $\mathcal{L}^n(A) < +\infty$, che esistono un compatto K e un aperto V tali che (i) $K \subset A \subset V$; (ii) $\mathcal{L}^n(V \setminus K) < \epsilon$. Usiamo allora il lemma di Urysohn: ne deduciamo che esiste $g \in C_c(\mathbb{R}^n)$ con la proprietà che $g(x) = 1$ per ogni $x \in K$ e $g(x) = 0$ per ogni $x \notin V$. Ma allora $\{\, x \in \mathbb{R}^n \mid g(x) \neq f(x) \,\} \subset V \setminus K$, da cui il teorema se $f = \chi_A$.

Passo 2: Il teorema vale se f è una funzione semplice. Questo è un'immediata conseguenza del passo 1.

Passo 3: Il teorema vale se $0 \leq f \leq K$. .

Consideriamo allora la successione crescente di funzioni semplici, misurabili e non negative $s_1, s_2, \ldots$ che converge a f uniformemente costruita nel teorema 4.3.2. La successione $t_1, t_2, \ldots$ definita da $t_1 = s_1$, $t_n = s_n - s_{n-1}$ per $n > 1$ è una funzione semplice, e $t_n(x) \leq 2^{-n}$ se $n > K$ (è anzi facile vedere che $2^n t_n$ è, per ogni $n > K$, la funzione caratteristica dell'insieme $\{\, x \in \mathbb{R}^n \mid f(x) - s_{n-1}(x) > 2^{-n} \,\}$). Per ogni n consideriamo una funzione $g_n \in C_c(\mathbb{R}^n)$ tale che, posto $B_n = \{\, x \in \mathbb{R}^n \mid g_n(x) \neq t_n(x) \,\}$, si abbia: (i) $\mathcal{L}^n(B_n) < \epsilon 2^{-n}$; (ii) $0 \leq \max g_n \leq \sup t_n \leq 2^{-n}$. Allora la serie di funzioni continue $\sum g_n(x) \leq \sum 2^{-n}$ è uniformemente convergente ad una funzione continua $\tilde{g}$. Si noti che $\mathcal{L}^n(\bigcup_n B_n) \leq \sum_n \mathcal{L}^n(B_n) < \epsilon$ e che $t_n(x) = g_n(x)$ per ogni n se $x \notin \bigcup_n B_n$. Quindi, per ogni $x \notin \bigcup_n B_n$, abbiamo che $\tilde{g}(x) = \sum g_n(x) = \sum t_n(x) = f(x)$. Quindi $\tilde{g}$ è una funzione continua, diversa da f al piú nell'insieme $B = \bigcup_n B_n$, insieme di misura $< \epsilon$. Non sappiamo se $\tilde{g}$ è una funzione a supporto compatto. Ma sappiamo che esistono un compatto K e un aperto V tali che $K \subset A \subset V$

con $\mathcal{L}^n(V \setminus A) < \epsilon$. Con il lemma di Urysohn possiamo allora trovare una funzione $h \in C_c(\mathbb{R}^n)$ che vale 1 in K e che si annulla fuori di V. Quindi la funzione $g(x) = h(x)\tilde{g}(x)$ è la funzione cercata. Infatti è a supporto compatto e $\{ x \in \mathbb{R}^n \mid g(x) \neq f(x) \} \subset B \cup (V \setminus K)$, con $\mathcal{L}^n(B \cup (V \setminus K)) \leq \mathcal{L}^n(B) + \mathcal{L}^n(V \setminus K) < 2\epsilon$.

Passo 4: Il caso generale. Per ogni $n \in \mathbb{N}$, sia $B_n := \{ x \mid |f(x)| \leq n \}$. Allora $\bigcup_n B_n = \mathbb{R}^n \setminus Z$, $\mathcal{L}^n(Z) = 0$. Poiché $B_n^c \subset A$, $\mathcal{L}^n(A) < +\infty$, ne segue che $\mathcal{L}^n(B_n^c) \to \mathcal{L}^n(Z) = 0$. Fissiamo $n \in \mathbb{N}$ in modo che $\mathcal{L}^n(B_n^c) < \epsilon$. Poniamo infine $f_n^+(x) = \min\{f^+(x), n\}$ e $f_n^-(x) = \min\{f^-(x), n\}$. Alle funzioni f_n^+ e f_n^- possiamo applicare il passo 3. Ne deduciamo che esiste g_n^+ e $g_n^- \in C_c(\mathbb{R}^n)$ tali che $\mathcal{L}^n(\{ x \mid g_n^\pm(x) \neq f_n^\pm(x) \}) < \epsilon$. Poiché

$$\{ x \mid g_n^\pm(x) \neq f^\pm(x) \} \subset \{ x \mid g_n^\pm(x) \neq f_n^\pm(x) \} \cup \{ x \mid f_n^\pm(x) \neq f^\pm(x) \},$$

abbiamo subito la tesi. □

Osservazione 4.3.11. Come abbiamo detto nell'osservazione precedente, il lemma di Urysohn continua a valere in spazi topologici localmente compatti. Si può quindi dimostrare il teorema di Lusin (seguendo la dimostrazione sopra) anche in spazi topologici localmente compatti su cui sia definita una misura regolare.

Corollario 4.3.12. *Sia $f\colon X \to \overline{\mathbb{R}}$, $X \subset \mathbb{R}^n$. Supponiamo che*

1. *f è misurabile in X;*

2. *$|f(x)| \leq K$ per quasi ogni $x \in X$;*

3. *Esiste $A \subset X$, A misurabile e di misura finita tale che $f(x) = 0$ per ogni $x \notin A$.*

Allora esiste una successione di funzioni $g_1, g_2, \ldots \in C_c(\mathbb{R}^n)$, tali che

(a) *$|g_n(x)| \leq K$;*

(b) *$f(x) = \lim_{n\to\infty} g_n(x)$.*

Dimostrazione. Dal teorema di Lusin segue che per ogni $n \in \mathbb{N}$ esiste $g_n \in C_c(\mathbb{R}^n)$ tale che (i) $|g_n(x)| \leq K$ e (ii) $\mathcal{L}^n(B_n) < 2^{-n}$ se $B_n = \{ x \mid g_n(x) \neq f(x) \}$. Ovviamente $\mathcal{L}^n(B_n) < \infty$. Se $x \in X$ appartiene solo ad un numero finito di B_n, abbiamo immediatamente che $g_n(x) = f(x)$ definitivamente. Mostriamo che ciò è vero per quasi ogni $x \in X$. Sia B l'insieme degli $x \in X$ che appartengono a B_k per infiniti $k \in \mathbb{N}$. Si ha che $B = \bigcap_{n=1}^\infty \bigcup_{k\geq n} B_k$. Poiché $\tilde{B}_n = \bigcup_{k\geq n} B_k \searrow B$ e $\mathcal{L}^n(\tilde{B}_n) \leq \sum_{k\geq n} \mathcal{L}^n(B_k) \to 0$, abbiamo che $\mathcal{L}^n(B) = \lim \mathcal{L}^n(\tilde{B}_n) = 0$, da cui la tesi. □

Esercizi Capitolo 4

(1) Sia $X = \{a, b, c, d\}$ e consideriamo $\mathcal{C} = \{\{a, b\}, \{c\}, \{d\}\} \subset \mathcal{P}(X)$. Qual'è la piú piccola σ-algebra che contiene $\mathcal{C}$? Quali sono le funzioni $f\colon X \to \mathbb{R}$ misurabili?

(2) Sia $f_1, f_2, \ldots\colon X \to \mathbb{R}$ una successione di funzioni misurabili. Si dimostri che $A = \{\, x \in X \mid f_n(x) \text{ converge} \,\}$ è un insieme misurabile.

(3) Sia $f\colon \mathbb{R}^k \to \mathbb{C}$ una funzione misurabile. Si mostri che esiste una funzione Borel misurabile $g\colon \mathbb{R}^k \to \mathbb{C}$ tale che $f(x) = g(x)$ per q.o. x.

(4) Si dimostri che, per ogni $\varnothing \neq F \subset X$, X spazio metrico, la funzione $x \mapsto \operatorname{dist}(x, F)$ è continua.

Capitolo 5

L'Integrazione secondo Lebesgue

In questo capitolo introduciamo l'integrazione secondo Lebesgue. Lo facciamo nel caso generale di funzioni f definite in uno spazio di misura (X, Σ, μ) a valori in $\mathbb{C}$ o nella retta reale estesa $\overline{\mathbb{R}} = [-\infty, +\infty]$. Ricordiamo che in $[-\infty, +\infty]$ assumiamo sempre, oltre alle usuali operazioni coinvolgenti $\pm\infty$, che $0 \cdot \pm\infty = 0$.

5.1. Definizione dell'integrale per funzioni nonnegative

D'ora in poi denoteremo con (X, Σ, μ) uno spazio di misura. Iniziamo definendo l'integrale per le funzioni semplici.

Definizione 5.1.1. Sia $s\colon X \to [0, +\infty)$ una funzione semplice misurabile di forma canonica

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

e $E \in \Sigma$ un insieme misurabile in X. Definiamo allora l'integrale di s in E ponendo (ricordiamo la convenzione $0 \cdot +\infty = 0$)

$$\int_E s\, d\mu = \sum_{i=1}^{n} \alpha_i\, \mu(A_i \cap E).$$

Vediamo subito alcune proprietà di tale integrale.

Proposizione 5.1.2. *Siano s e t due funzioni semplici, misurabili e non negative e A e B due insiemi misurabili in X. Allora:*

(a) *$s(x) \leq t(x)$ per ogni $x \in A$ implica $\int_A s\,d\mu \leq \int_A t\,d\mu$;*

(b) *$A \subset B$ implica $\int_A s\,d\mu \leq \int_B s\,d\mu$;*

(c) *$\int_A cs\,d\mu = c\int_A s\,d\mu$ per ogni $c \geq 0$;*

(d) *$s(x) = 0$ per ogni $x \in A$ implica $\int_A s\,d\mu = 0$;*

(e) *$\mu(A) = 0$ implica $\int_A s\,d\mu = 0$;*

(f) *la funzione $\phi\colon \Sigma \to [0,+\infty]$ definita da $\phi(A) = \int_A s\,d\mu$ è una misura su Σ;*

(g) *$\int_A (s+t)\,d\mu = \int_A s\,d\mu + \int_A t\,d\mu$.*

(h) *$\int_X s\chi_A\,d\mu = \int_A s\,d\mu$.*

Dimostrazione. Sia $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$ e $t = \sum_{j=1}^m \beta_j \chi_{B_j}$, dove $A_i \cap A_j = \varnothing$ e $B_i \cap B_j = \varnothing$ per ogni $i \neq j$.

Definiamo poi $E_{ij} = A_i \cap B_j$. Allora $E_{ij} \cap E_{i'j'} = \varnothing$ e $X = \bigcup_{i,j} E_{i,j}$ e si vede facilmente che

$$\int_A s\,d\mu = \sum_{i=1}^n \alpha_i\,\mu(A \cap A_i) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i\,\mu(A \cap E_{ij}). \tag{5.1.1}$$

Le **(a)** segue allora notando che, per ogni $x \in A \cap E_{ij}$, $\alpha_i = s(x) \leq t(x) = \beta_j$. Se $A \cap E_{ij} = \varnothing$ non è piú vero che $\alpha_i \leq \beta_j$, ma in ogni caso $\alpha_i\,\mu(A \cap E_{ij}) \leq \beta_j\,\mu(A \cap E_{ij})$. Quindi dalla (5.1.1) otteniamo che

$$\int_A s\,d\mu \leq \sum_{i=1}^n \sum_{j=1}^m \beta_j\,\mu(A \cap E_{ij}) = \int_A t\,d\mu.$$

Le **(b)**-**(e)** sono una conseguenza immediata della definizione.

Mostriamo che vale la **(f)**. Ovviamente $\phi(\varnothing) = \int_\varnothing s\,d\mu = 0$. Quindi dobbiamo verificare solo l'additività numerabile. Sia $E_1, E_2, \dots$ una successione di misurabili a due a due disgiunti in X, e poniamo $A = \bigcup_i E_i$. Dalla definizione abbiamo

$$\begin{aligned}
\phi(A) = \int_A s\,d\mu &= \sum_{i=1}^n \alpha_i\,\mu(A \cap A_i) \\
&= \sum_{i=1}^n \alpha_i\,\mu(A_i \cap (\bigcup_j E_j)) \\
&= \sum_{i=1}^n \sum_{j=1}^\infty \alpha_i\,\mu(A_i \cap E_j) \\
&= \sum_{j=1}^\infty \int_{E_j} s\,d\mu = \sum_{j=1}^\infty \phi(E_j).
\end{aligned}$$

Mostriamo ora l'additività **(g)**. Sia ϕ_{s+t} la misura definita da $s+t$. Allora, per l'additività di tale misura abbiamo

$$\begin{aligned}\int_A (s+t)\,d\mu = \phi_{s+t}(A) &= \sum_{i,j} \phi_{s+t}(A\cap E_{ij})\\ &= \sum_{i,j}\int_{A\cap E_{ij}}(s+t)\,d\mu = \sum_{i,j}(\alpha_i+\beta_j)\,\mu(A\cap E_{ij})\\ &= \sum_i \alpha_i \sum_j \mu(A\cap E_{ij}) + \sum_j \beta_j \sum_i \mu(A\cap E_{ij})\\ &= \sum_i \alpha_i\,\mu(A\cap A_i) + \sum_j \beta_j\,\mu(A\cap B_j)\\ &= \int_A s\,d\mu + \int_A t\,d\mu\end{aligned}$$

e la proprietà cercata segue.

Per dimostrare la **(h)** osserviamo che la $s\chi_A$ è una funzione semplice. Quindi, utilizzando la **(a)** (che implica l'uguaglianza $\int_A s\,d\mu = \int_A t\,d\mu$ se $s=t$ in A) e la **(f)**, abbiamo che

$$\int_X s\chi_A\,d\mu = \int_A s\chi_A\,d\mu + \int_{X\setminus A} s\chi_A\,d\mu = \int_A s\,d\mu + \int_{X\setminus A} 0\,d\mu = \int_A s\,d\mu$$

□

Definizione 5.1.3. Sia ora $f\colon X\to[0,+\infty]$ una funzione misurabile nonnegativa e E un insieme misurabile in X. Definiamo allora l'integrale di f in E ponendo

$$\int_E f\,d\mu := \sup\Big\{\int_E s\,d\mu \;\Big|\; 0\le s(x)\le f(x),\ s \text{ semplice e misurabile}\Big\}.$$

Osservazione 5.1.4. Si noti che, se s è una funzione semplice non negativa, allora segue dalla proposizione 5.1.2**(a)** che questa definizione coincide con quella data sopra.

Notazione 5.1.5. Se $\mu=\mathcal{L}^n$, indicheremo l'integrale di f con l'usuale simbolo $\int_E f\,dx$ invece di $\int_E f\,d\mathcal{L}^n$, e nel caso di $X\subset\mathbb{R}$, $E=[a,b]$ utilizzeremo l'usuale notazione $\int_a^b f\,dx$ utilizzata per l'integrale secondo Riemann. Il fatto di utilizzare lo stesso simbolo sia per l'integrale secondo Lebesgue che per quello secondo Riemann sarà giustificato in seguito (si veda il teorema 5.6.1).

Abbiamo subito alcune proprietà che seguono dalle analoghe proprietà dell'integrale di funzioni semplici:

Proposizione 5.1.6. *Siano f e g due funzioni misurabili e non negative e A e B due insiemi misurabili in X. Allora:*

(a) $f(x) \leq g(x)$ *per quasi ogni* $x \in A$ *implica* $\int_A f\, d\mu \leq \int_A g\, d\mu$ *(in particolare* $\int_A f\, d\mu = \int_A g\, d\mu$ *se* $f(x) = g(x)$ *per quasi ogni* $x \in A$*);*

(b) $A \subset B$ *implica* $\int_A f\, d\mu \leq \int_B f\, d\mu$*;*

(c) $\int_A cf\, d\mu = c \int_A f\, d\mu$ *per ogni* $c \geq 0$*;*

(d) $f(x) = 0$ *per quasi ogni* $x \in A$ *implica* $\int_A f\, d\mu = 0$*;*

(e) $\mu(A) = 0$ *implica* $\int_A f\, d\mu = 0$*;*

(f) $\int_A f\, d\mu = \int_X f\chi_A\, d\mu$.

Dimostrazione. Le **(a)**-**(e)** sono un'immediata conseguenza delle analoghe proprietà per le funzioni semplici. Per quanto riguarda la **(f)** osserviamo che, per ogni funzione semplice s tale che $0 \leq s(x) \leq f(x)\chi_A(x)$ per ogni $x \in A$, si ha che $s = s\chi_A$, e quindi $\int_X s\, d\mu = \int_A s\, d\mu$. Ne segue che $\int_X f\chi_A\, d\mu = \int_A f\chi_A\, d\mu = \int_A f\, d\mu$. □

5.2. Teorema della convergenza monotona di Beppo Levi

Il seguente teorema di passaggio al limite sotto al segno di integrale è uno degli strumenti fondamentali nella teoria dell'integrazione secondo Lebesgue.

Teorema 5.2.1 (Teorema della convergenza monotona o di Beppo Levi)**.** *Sia* $f_1, f_2, \ldots : X \to [0, +\infty]$ *una successione monotona di funzioni misurabili, cioè tale che* $0 \leq f_1(x) \leq f_2(x) \leq \ldots$ *per quasi ogni* $x \in X$*.*

Allora

(a) $f_n(x) \to f(x)$ *per quasi ogni* $x \in X$*;*

(b) $f : X \to [0, +\infty]$ *è misurabile;*

(c) $\int_X f_n\, d\mu \to \int_X f\, d\mu$.

Dimostrazione. Sia X' l'insieme degli $x \in X$ tali che $f_n(x) \leq f_{n+1}(x)$ per ogni n. Abbiamo che $\mu(X \setminus X') = 0$. Abbiamo subito che $f_n(x)$ converge per ogni $x \in X'$ (la successione è monotona). Sia $f(x)$ uguale al limite di tale successione per ogni $x \in X'$ e a 0 per ogni $x \notin X'$. Ovviamente f è misurabile (in quanto limite di funzioni misurabili). Dalla proposizione 5.1.6**(a)** si ha che $\int_X f_n\, d\mu \leq \int_X f_{n+1}\, d\mu$, e quindi la successione (numerica) $\int_X f_n\, d\mu \to \alpha$ per qualche $\alpha \in [0, +\infty]$. Poiché $f_n(x) \leq f(x)$ per quasi ogni x, abbiamo anche che $\int_X f_n\, d\mu \leq \int_X f\, d\mu$, da cui $\alpha \leq \int_X f\, d\mu$.

Per dimostrare il teorema è quindi sufficiente mostrare che $\int_X f\, d\mu \leq \alpha$. Prendiamo una funzione s misurabile, semplice, non negativa e tale che $0 \leq s(x) \leq f(x)$ per ogni $x \in X'$, ed una costante $c \in (0, 1)$. Definiamo poi, per ogni $n \in \mathbb{N}$

$$E_n = \left\{\, x \in X' \;\middle|\; cs(x) \leq f_n(x) \,\right\}.$$

Ovviamente E_n è misurabile ed $E_n \subset E_{n+1}$ per la monotonia della successione. Se $x \in X'$ e $f(x) = 0$, allora $s(x) = 0$ e $x \in E_n$ per ogni n. In caso contrario, $f(x) > 0$ e $cs(x) < f(x)$, e dalla convergenza di $f_n(x) \to f(x)$ deduciamo che $x \in E_n$ per n sufficientemente grande. In ogni caso $x \in \bigcup_n E_n$. Quindi

$$\int_X f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} cs \, d\mu = c \int_{E_n} s \, d\mu \to c \int_X s \, d\mu,$$

dove abbiamo usato le proposizione 5.1.6**(b)**, la definizione di E_n, la proposizione 5.1.6**(c)**, la proposizione 2.2.4**(c)** insieme con la proposizione 5.1.6**(f)**.

Ne segue che $\alpha \geq c \int_X s \, d\mu$ per ogni $0 \leq s \leq f$ e per ogni $c \in (0,1)$. Dalla definizione dell'integrale abbiamo allora che $c \int_X f \, d\mu \leq \alpha$ e prendendo l'estremo superiore su $c \in (0,1)$ otteniamo la diseguaglianza cercata. □

Utilizziamo ora il teorema della convergenza monotona per dimostrare ulteriori proprietà dell'integrale.

Teorema 5.2.2. *Siano $f, g \colon X \to [0, +\infty]$ due funzioni misurabili. Allora*

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

Dimostrazione. Sappiamo, dal teorema 4.3.2, che esistono due successioni monotone non decrescenti s_n e t_n di funzioni semplici, non negative e misurabili tali che $s_n(x) \to f(x)$ e $t_n(x) \to g(x)$ in tutti i punti $x \in X$. Ovviamente anche la successione $h_n = s_n + t_n$ è una successione monotona non decrescente di funzioni semplici, non negative e misurabili, e converge puntualmente alla funzione $f + g$.

Possiamo allora applicare a ciascuna delle tre successioni s_n, t_n e h_n il teorema della convergenza monotona, ottenendo

$$\begin{aligned}\int_X (f+g) \, d\mu &= \lim_{n\to\infty} \int_X (s_n + t_n) \, d\mu \\ &= \lim_{n\to\infty} \int_X s_n \, d\mu + \lim_{n\to\infty} \int_X t_n \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu,\end{aligned}$$

(abbiamo utilizzato la proposizione 5.1.2**(g)**). □

Il seguente è un'immediato corollario del teorema appena dimostrato.

Corollario 5.2.3. *Siano $f_1, f_2, \ldots : X \to [0, +\infty]$ funzioni misurabili. Allora*

$$\int_X \sum_{i=1}^n f_i \, d\mu = \sum_{i=1}^n \int_X f_i \, d\mu.$$

Sempre dal teorema della convergenza monotona deduciamo

Teorema 5.2.4. *Sia $f_1, f_2, \ldots : X \to [0, +\infty]$ una successione di funzioni misurabili non negative. Allora*

$$\int_X \sum_{i=1}^{\infty} f_i \, d\mu = \sum_{i=1}^{\infty} \int_X f_i \, d\mu.$$

Dimostrazione. Si osservi che, essendo tutte le serie coinvolte nell'enunciato a termini non negativi, non vi sono problemi nel dare significato all'enunciato (anche se alcune delle serie possono divergere a $+\infty$).

Definiamo $g_n(x) = \sum_{i=1}^{n} f_i(x)$. Allora g_n è una successione monotona non decrescente di funzioni misurabili non negative che converge a $\sum_{i=1}^{\infty} f_i(x)$. Ne deduciamo, usando il teorema della convergenza monotona e il precedente corollario:

$$\begin{aligned} \int_X \sum_{i=1}^{\infty} f_i \, d\mu &= \int_X \lim_{n \to +\infty} g_n \, d\mu = \lim_{n \to +\infty} \int_X g_n \, d\mu \\ &= \lim_{n \to +\infty} \sum_{i=1}^{n} \int_X f_i \, d\mu = \sum_{i=1}^{\infty} \int_X f_i \, d\mu. \end{aligned}$$

□

Dai teoremi precedenti segue

Teorema 5.2.5. *Sia $f \colon X \to [0, +\infty]$ una funzione misurabile. Allora la funzione $\nu \colon \Sigma \to [0, +\infty]$ definita da*

$$\nu(A) = \int_A f \, d\mu$$

è una misura, che indicheremo con $d\nu = f \, d\mu$. Inoltre, per ogni $g \colon X \to [0, +\infty]$ misurabile vale

$$\int_X g \, d\nu = \int_X g f \, d\mu \tag{5.2.1}$$

Dimostrazione. Ovviamente $\nu(\varnothing) = 0$. Verifichiamo che ν è numerabilmente additiva. Sia $A_1, A_2, \ldots$ una successione di insiemi misurabili e disgiunti. Poiché la successione $A_1, A_2, \ldots$ è disgiunta, abbiamo che

$$\chi_{\bigcup_i A_i}(x) = \sum_{i=1}^{\infty} \chi_{A_i}(x) \qquad \forall x \in X.$$

Ne segue, usando il teorema 5.2.4 di integrazione di una serie termine a termine ,

$$\begin{aligned}\nu(\bigcup_i A_i) &= \int_{\bigcup A_i} f\,d\mu = \int_X f\chi_{\bigcup_i A_i}\,d\mu \\ &= \int_X \sum_{i=1}^{\infty} f\chi_{A_i}\,d\mu = \sum_{i=1}^{\infty}\int_X f\chi_{A_i}\,d\mu \\ &= \sum_{i=1}^{\infty} \nu(A_i).\end{aligned}$$

Se $g = \chi_E$ con E misurabile allora vale

$$\int_X g\,d\nu = \nu(E) = \int_E f\,d\mu = \int_X gf\,d\mu.$$

Quindi la (5.2.1) vale per le funzioni caratteristiche e quindi anche per le funzioni semplici. Dalla definizione dell'integrale per funzioni non negative segue allora che la (5.2.1) vale per ogni g misurabile e non negativa. □

Mostriamo infine un'altra caratterizzazione dell'integrale secondo Lebesgue di una funzione nonnegativa.

Proposizione 5.2.6. *Sia $f\colon X \to [0, +\infty]$ una funzione misurabile. Allora*

$$\int_X f\,d\mu = \int_0^{+\infty} \mu(\{\, x \in X \mid f(x) > t \,\})\,dt \tag{5.2.2}$$

(l'integrale a secondo membro può anche essere inteso nel senso di Riemann. Infatti la funzione $t \mapsto \mu(\{\, x \in X \mid f(x) > t \,\})$ è non crescente in t, e quindi Riemann integrabile, anche se il valore di tale integrale potrebbe essere non finito).

Dimostrazione. Innanzi tutto osserviamo che $\{\, x \in X \mid f(x) > t \,\}$ è un insieme misurabile. Sia $0 \leq s(x) \leq f(x)$ una funzione semplice, $s(x) = \sum_{i=1}^n \alpha_i \chi_{A_i}(x)$. Allora è facile vedere che (5.2.2) vale per s. Prendiamo poi una successione di funzioni semplici $0 \leq s_1 \leq s_2 \leq \cdots \leq f$ convergenti a f. È immediato vedere che, posto $g_n(t) = \mu(\{\, x \in X \mid s_n(x) > t \,\})$, $0 \leq g_n(t) \leq g_{n+1}(t)$, e che $g_n(t) \to g(t) = \mu(\{\, x \in X \mid f(x) > t \,\})$. Allora, applicando il teorema della convergenza monotona sia alla successione s_n che alla successione g_n si ottiene il risultato. Se si considera nella (5.2.2) l'integrale secondo Riemann, è necessario dimostrare che anche in questo caso si può passare al limite sotto il segno di integrale (conseguenza del fatto che le funzioni g_n sono tutte non crescenti). □

Osservazione 5.2.7. Sia f una funzione misurabile e non negativa. Poniamo $G = \{\, (x, t) \in X \times \mathbb{R} \mid 0 \leq t < f(x) \,\}$. Allora $\int_0^{+\infty} \chi_G(x, t)\,dt = f(x)$, $\int_X \chi_G(x, t)\,d\mu = \mu(\{\, x \in X \mid f(x) > t \,\})$.

Dalla proposizione precedente possiamo quindi dedurre che, qualunque sia la funzione misurabile e non negativa f, si ha che

$$\int_X \Big(\int_0^{+\infty} \chi_G(x,t)\, dt \Big)\, d\mu = \int_0^{+\infty} \Big(\int_X \chi_G(x,t)\, d\mu \Big)\, dt.$$

Vedremo in seguito (si veda la sezione 10.1) che la misura prodotto $\mu \times \mathcal{L}^1$ nello spazio prodotto $X \times \mathbb{R}$ si definisce esattamente ponendo, per un generico $G \subset X \times \mathbb{R}$ misurabile $(\mu \times \mathcal{L}^1)(G) = \int_X \Big(\int_0^{+\infty} \chi_G\, dt \Big)\, d\mu$. Possiamo quindi affermare che, per le funzioni caratteristiche χ_G di insiemi G della forma vista sopra vale sempre il teorema di Fubini 10.2.1 che dimostreremo nel seguito.

5.3. Lemma di Fatou

La convergenza puntuale di $f_n(x) \to f(x)$ non è sufficiente ad assicurare la convergenza di $\int f_n \to \int f$ (basta considerare la successione $f_n = n\chi_{(0,1/n)}$, per cui vale $\int_{\mathbb{R}} f_n = 1$, $f_n(x) \to 0$ per ogni $x \in \mathbb{R}$). Il seguente teorema ci fornisce però un'utile diseguaglianza per tali successioni.

Teorema 5.3.1 (Lemma di Fatou)**.** *Sia $f_1, f_2, \ldots : X \to [0, +\infty]$ una successione di funzioni misurabili. Allora*

$$\int_X \liminf_n f_n\, d\mu \leq \liminf_n \int_X f_n\, d\mu.$$

Dimostrazione. Ricordiamo che $\liminf_n f_n(x) = \sup_n \inf_{k \geq n} f_k(x)$. Posto $g_n(x) = \inf_{k \geq n} f_k(x)$, la successione $g_1, g_2, \ldots$ è una successione monotona crescente di funzioni misurabili non negative, che converge a $\liminf_n f_n$. Possiamo allora applicare a tale successione il teorema della convergenza monotona, deducendo che

$$\int_X \liminf_n f_n\, d\mu = \int_X \lim_n g_n\, d\mu = \lim_n \int_X g_n\, d\mu,$$

e il teorema segue dall'ovvia diseguaglianza $\int_X g_n\, d\mu \leq \int_X f_n\, d\mu$. □

Osservazione 5.3.2. Nel caso in cui $f_j(x)$ converge a $f(x)$ per q.o. $x \in X$ il lemma afferma che

$$\liminf_j \int_X f_j(x)\, d\mu \geq \int_X f(x)\, d\mu.$$

5.4. Integrali per funzioni a valori complessi

Consideriamo ora una funzione a valori complessi, $f\colon X \to \mathbb{C}$, e definiamone l'integrale.

Definizione 5.4.1. Sia $f\colon X \to \mathbb{C}$ una funzione misurabile. Diciamo che tale funzione è integrabile secondo Lebesgue se

$$\int_X |f|\, d\mu < +\infty.$$

Indicheremo con $L^1(\mu)$ la classe delle funzioni integrabili, cioè

$$L^1(\mu) := \left\{ f\colon X \to \mathbb{C} \;\middle|\; f \text{ è misurabile e } \int_X |f|\, d\mu < +\infty \right\}.$$

Se $f \in L^1(\mu)$, $f = u + iv$, poniamo

$$\int_X f\, d\mu = \int_X u^+\, d\mu - \int_X u^-\, d\mu + i\int_X v^+\, d\mu - i\int_X v^-\, d\mu,$$

(dove $u = u^+ - u^-$; si noti che dalla diseguaglianza $0 \le u^+ \le |u| \le |f|$ segue che $\int_X u^+\, d\mu < +\infty$, e analogamente $\int_X u^-\, d\mu, \int_X v^\pm\, d\mu < +\infty$, cosí che la somma sopra è ben definita).

Mostriamo innanzi tutto che $L^1(\mu)$ è uno spazio vettoriale e che l'integrale è lineare.

Teorema 5.4.2. *Siano f e g due funzioni in $L^1(\mu)$ e $\alpha, \beta \in \mathbb{C}$. Allora $\alpha f + \beta g \in L^1(\mu)$ e vale*

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

Dimostrazione. La misurabilità di $\alpha f + \beta g$ segue dal teorema 4.2.4. La diseguaglianza $|\alpha f + \beta g|(x) \le |\alpha||f|(x) + |\beta||g|(x)$ implica poi che $\alpha f + \beta g \in L^1(\mu)$. Mostriamo poi che valgono le due uguaglianze

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

Supponiamo f e g a valori reali. Allora, posto $h = f + g$, vale l'uguaglianza

$$h^+ - h^- = f^+ - f^- + g^+ - g^-,$$

da cui deduciamo che

$$h^+ + f^- + g^- = f^+ + g^+ + h^-.$$

Usando l'additività dell'integrale per funzioni non negative (teorema 5.2.2) otteniamo

$$\int_X h^+\, d\mu + \int_X f^-\, d\mu + \int_X g^-\, d\mu = \int_X f^+\, d\mu + \int_X g^+\, d\mu + \int_X h^-\, d\mu.$$

Poiché tutti i termini di questa uguaglianza sono finiti, ne deduciamo

$$\int_X h^+ \, d\mu - \int_X h^- \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu + \int_X g^+ \, d\mu - \int_X g^- \, d\mu,$$

ovvero

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

Passiamo ora a dimostrare l'omogeneità dell'integrale. Sia $\alpha \geq 0$ e $\alpha f = \alpha f^+ - \alpha f^-$. Ne otteniamo $\alpha f + \alpha f^- = \alpha f^+$. Usando l'additività dell'integrale appena dimostrata, abbiamo che $\int_X \alpha f \, d\mu + \int_X \alpha f^- \, d\mu = \int_X \alpha f^+ \, d\mu$. Osserviamo ora che le funzioni $f^\pm$ sono non negative, e che $\alpha \geq 0$. Possiamo allora applicare la proposizione 5.1.6**(c)** e dedurne che $\int_X \alpha f \, d\mu = \alpha \int_X f^+ \, d\mu - \alpha \int_X f^- \, d\mu = \alpha \int_X f \, d\mu$. Prendiamo ora $\alpha = -1$. Abbiamo che $(\alpha f)^\pm = (-f)^\pm = f^\mp$. Ne otteniamo che $\int_X -f \, d\mu = \int_X f^- \, d\mu - \int_X f^+ \, d\mu = -\int_X f \, d\mu$. Sia ora $\alpha < 0$. Allora $\int_X \alpha f \, d\mu = \int_X -|\alpha| f \, d\mu = -\int_X |\alpha| f \, d\mu = -|\alpha| \int_X f \, d\mu = \alpha \int_X f \, d\mu$. Abbiamo quindi dimostrato il teorema nel caso in cui f e g sono funzioni a valori reali e α, β numeri reali.

Sia ora $f = u + iv$ una funzione a valori complessi. Abbiamo, utilizzando la definizione di integrale e l'omogeneità dell'integrale di funzioni a valori reali, che $\int_X if \, d\mu = \int_X (-v + iu) \, d\mu = \int_X (-v) \, d\mu + i \int_X u \, d\mu = -\int_X v \, d\mu + i \int_X u \, d\mu = i\Big(\int_X u \, d\mu + i \int_X v \, d\mu\Big) = i \int_X f \, d\mu$. Se $g = w + it$ è un'altra funzione a valori complessi, abbiamo, utilizzando la definizione di integrale e l'additività dell'integrale di funzioni a valori reali che $\int_X (f+g) \, d\mu = \int_X (u+w) \, d\mu + i \int_X (v+t) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$. Sia ora $\alpha = \alpha_1 + i\alpha_2$ e u a valori reali. Otteniamo che $\int_X \alpha u \, d\mu = \int_X (\alpha_1 u + i\alpha_2 u) \, d\mu = \alpha_1 \int_X u \, d\mu + i\alpha_2 \int_X u \, d\mu = \alpha \int_X u \, d\mu$. Ne segue che $\int_X \alpha f \, d\mu = \int_X (\alpha u + i\alpha v) \, d\mu = \int_X \alpha u \, d\mu + \int_X i\alpha v \, d\mu = \alpha \int_X u \, d\mu + i\alpha \int_X v \, d\mu = \alpha \int_X f \, d\mu$, e il teorema è dimostrato. □

Teorema 5.4.3. *Sia $f \in L^1(\mu)$. Allora*

$$|\int_X f \, d\mu| \leq \int_X |f| \, d\mu.$$

Dimostrazione. Sia $z = \int_X f \, d\mu \in \mathbb{C}$. Esiste $\alpha \in \mathbb{C}$ tale che $|\alpha| = 1$ e $\alpha z = |z|$. Sia $u = \Re(\alpha f)$ la parte reale della funzione αf. Ovviamente $u \leq |\alpha f| = |f|$. Ne deduciamo

$$|\int_X f \, d\mu| = \alpha \int_X f \, d\mu = \int_X \alpha f \, d\mu = \int_X u \, d\mu \leq \int_X |f| \, d\mu.$$

(si osservi che essendo $\int_X \alpha f \, d\mu$ un numero reale deve necessariamente essere $\int_X v \, d\mu = 0$). □

Teorema 5.4.4. *Sia f una funzione misurabile, ed $A \subset X$ un sottoinsieme misurabile.*

Allora

(a) *(**Diseguaglianza di Tchebyshev**) $f\colon X \to [0,+\infty]$ implica $\mu(\{ x \in X \mid f(x) \geq \alpha \}) \leq \frac{1}{\alpha}\int_X f\,d\mu$;*

(b) *Sia $f \in L^1(\mu)$. Allora $|f(x)| < +\infty$ per quasi ogni $x \in X$;*

(c) *$f\colon X \to [0,+\infty]$ e $\int_A f\,d\mu = 0$ implica $f(x) = 0$ per quasi ogni $x \in A$;*

(d) *$f \in L^1(\mu)$ e $\int_A f\,d\mu = 0$ per ogni A misurabile implica $f(x) = 0$ per quasi ogni $x \in X$;*

(e) *$f \in L^1(\mu)$ e $|\int_X f\,d\mu| = \int_X |f|\,d\mu$ implica che esiste $\beta \in \mathbb{C}$ tale che $|\beta| = 1$ e $\beta f(x) = |f(x)|$ per quasi ogni $x \in X$.*

Dimostrazione. **(a)**. Se $A = \{ x \in X \mid f(x) \geq \alpha \}$, abbiamo che A è misurabile e

$$\int_X f\,d\mu \geq \int_A f\,d\mu \geq \alpha \int_A d\mu = \alpha\,\mu(A),$$

da cui la tesi.

(b). Se $A_n = \{ x \in X \mid |f(x)| < +\infty \}$, allora, dalla Diseguaglianza di Tchebyshev **(a)** segue che $\mu(A_n) \leq \frac{1}{n}\int_X |f|\,d\mu \to 0$ per $n \to +\infty$. D'altra parte $A_{n+1} \subset A_n$ e quindi, per la Proposizione 2.2.4**(d)** $0 = \lim_{n\to+\infty}\mu(A_n) = \mu(\bigcap_n A_n)$. Poichè $\{ x \in X \mid |f(x)| = +\infty \} = \bigcap_n A_n$, la tesi segue.

(c). Se $A_n = \{ x \in A \mid f(x) > \frac{1}{n} \}$, abbiamo che A_n è misurabile e, per la diseguaglianza di Tchebyshev, $\mu(A_n) \leq n\int_X f\,d\mu = 0$. Quindi $\mu(A_n) = 0$. Poiché $\{ x \in A \mid f(x) \neq 0 \} = \bigcup A_n$ e la successione $A_1, A_2, \ldots$ è crescente, si ha che $\mu(\{ x \in A \mid f(x) \neq 0 \}) = \lim_n \mu(A_n) = 0$.

(d). Sia $f \in L^1(\mu)$, $f = u + iv$. Sia $A = \{ x \in X \mid u > 0 \}$. Poiché $u = u^+$ in A, e $0 = \int_A f\,d\mu = \int_A u^+\,d\mu + i\int_X v\,d\mu$, ne deduciamo subito che $\int_A u^+\,d\mu = 0$ e quindi $u^+(x) = 0$ per quasi ogni $x \in A$. Poiché $u^+(x) = 0$ per ogni $x \in X \setminus A$, ne deduciamo che $u^+(x) = 0$ per quasi ogni $x \in X$. Analogamente si mostra che $u^-(x) = 0$ e $v^\pm(x) = 0$ per quasi ogni $x \in X$.

(e). Poniamo $z = \int_X f\,d\mu$, e sia $\beta \in \mathbb{C}$ tale che $|\beta| = 1$ e $\beta z = |z|$. Allora

$$|\int_X f\,d\mu| = \beta\int_X f\,d\mu = \int_X \beta f\,d\mu = \int_X \Re(\beta f)\,d\mu.$$

Ne deduciamo che $\int_X \Re(\beta f)\,d\mu = \int_X |f|\,d\mu$, da cui $\int_X (|f| - \Re(\beta f))\,d\mu = 0$. Poiché la funzione $|f|(x) - \Re(\beta f)(x)$ è non negativa, segue da **(c)** che $|f|(x) = \Re(\beta f)(x)$ per quasi ogni $x \in X$. Ne segue che $\Im(\beta f)(x) = 0$ e $|f|(x) = \beta f(x)$ per quasi ogni $x \in X$. □

5.5. Teorema della Convergenza dominata

Teorema 5.5.1 (Teorema della convergenza dominata di Lebesgue)**.** *Consideriamo una successione $f_1, f_2, \dots$ di funzioni in $L^1(\mu)$, e supponiamo che*

1. $f_n(x) \to f(x)$ *per quasi ogni* $x \in X$*;*

2. *Esiste* $g \in L^1(\mu)$ *tale che* $|f_n(x)| \leq g(x)$ *per quasi ogni* $x \in X$ *e per ogni* $n \in \mathbb{N}$*.*

Allora

(a) $f \in L^1(\mu)$*;*

(b) $\int_X |f_n - f| \, d\mu \to 0$*;*

(c) $\int_X f_n \, d\mu \to \int_X f \, d\mu$*.*

Dimostrazione. Poiché $|f(x)| \leq g(x)$ q.o., abbiamo subito che $f \in L^1(\mu)$. Poiché $|f_n(x) - f(x)| \leq 2g(x)$, abbiamo che anche $|f_n(x) - f(x)| \in L^1(\mu)$. Sia, per ogni n, $g_n(x) = 2g(x) - |f_n(x) - f(x)|$. La successione di funzioni $g_1, g_2, \dots$ è allora una successione di funzioni misurabili non negative. Ovviamente vale $\liminf_n g_n(x) = 2g(x)$. Possiamo allora applicare a tale successione il lemma di Fatou. Ne segue che

$$\begin{aligned}\int_X 2g \, d\mu &\leq \liminf_n \int_X (2g - |f_n - f|) \, d\mu \\ &= \int_X 2g \, d\mu - \limsup_n \int_X |f_n - f| \, d\mu,\end{aligned}$$

da cui si ottiene

$$\limsup_n \int_X |f_n - f| \, d\mu \leq 0.$$

Poiché $\int_X |f_n - f| \, d\mu \geq 0$, ne segue subito che $\int_X |f_n - f| \, d\mu \to 0$. Infine

$$\left|\int_X f_n \, d\mu - \int_X f \, d\mu\right| = \left|\int_X (f_n - f) \, d\mu\right| \leq \int_X |f_n - f| \, d\mu \to 0.$$

□

Osservazione 5.5.2. Il teorema precedente ammette una piccola generalizzazione, utile però nelle applicazioni, in cui la funzione $g(x)$ è rimpiazzata da una successione $g_n(x)$ con la proprietà che esiste una funzione $g \in L^1$ tale che

$$\int_X |g - g_n| \, d\mu \to 0 \qquad \text{per} \quad n \to \infty$$

e tale che $|f_n(x)| \leq g_n(x)$, $|f(x)| \leq g(x)$. Di nuovo, se $f_n(x)$ converge puntualmente q.o. a $f(x)$ si può passare con il limite sotto il segno di integrale, ottenendo

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

Per dimostrarlo, poniamo $h_n(x) = 2g(x) - |f(x) - f_n(x)|$. Abbiamo che $h_n(x) \to 2g(x)$. Inoltre $h_n^+(x) \leq 2g(x)$ e $h_n^+(x) \to 2g(x)^+ = 2g(x)$. Per il teorema della convergenza dominata $\int_X h_n^+\, d\mu \to \int_X 2g\, d\mu$. Abbiamo poi che

$$h_n^-(x) = (2g(x) - g_n(x) + g_n(x) - |f(x) - f_n(x)|)^-.$$

Poiché $g_n(x)+g(x)-|f(x)-f_n(x)| \geq 0$ per ipotesi, $h_n^-(x) \leq (g(x)-g_n(x))^-$. Quindi

$$\int_X h_n^-\, d\mu \leq \int_X (g - g_n)^-\, d\mu \to 0,$$

e ne deduciamo che $\int_X h_n\, d\mu \to \int_X 2g\, d\mu$, ovvero

$$\int_X |f_n - f|\, d\mu \to 0,$$

e l'osservazione è dimostrata.

Dimostriamo ora qualche conseguenza del teorema della convergenza dominata.

Teorema 5.5.3. *Sia $f_1, f_2, \ldots : X \to \mathbb{C}$ una successione di funzioni misurabili tali che $\sum_{n=1}^{\infty}\int_X |f_n|\, d\mu < +\infty$.*

Allora la serie $f(x) = \sum_{n=1}^{\infty} f_n(x)$ converge per quasi ogni $x \in X$, $f \in L^1(\mu)$ e

$$\int_X f\, d\mu = \sum_{n=1}^{\infty}\int_X f_n\, d\mu.$$

Dimostrazione. Sia S_n l'insieme in cui è definita f_n, e $S = \bigcap S_n$. Allora S è misurabile, $\mu(S^c) = 0$, e per tutti gli $x \in S$, è ben definita la funzione

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

(che potrebbe avere valore $+\infty$ in alcuni punti.) Vale

$$\int_S \phi\, d\mu = \sum_{n=1}^{\infty}\int_S |f_n|\, d\mu < +\infty.$$

Quindi $\phi(x) < +\infty$ per quasi ogni $x \in S$, e quindi per quasi ogni $x \in X$. Ne segue che la serie $\sum_{n=1}^{\infty} f(x)$ è assolutamente convergente per quasi ogni $x \in X$. Posto $g_N(x) = \sum_{n=1}^{N} f_n(x)$, vale $|g_N|(x) \leq \phi(x) \in L^1(\mu)$ e $g_N(x) \to f(x)$ per quasi ogni $x \in X$. Il teorema della convergenza dominata implica allora che $\int_X g_N\, d\mu = \sum_{n=1}^{N}\int_X f_n\, d\mu \to \int_X f\, d\mu$. □

Il seguente teorema ci dice che, se le medie di f prendono valori in un chiuso S, allora f prende quasi sempre valori in quello stesso insieme. Ad esempio, se tutte le medie di f sono positive, allora anche f è (q.o.) positiva.

Teorema 5.5.4. *Supponiamo $\mu(X) < \infty$ e $f \in L^1(\mu)$. Sia S un sottoinsieme chiuso di $\mathbb{C}$ e supponiamo che, per ogni misurabile A di misura $\mu(A) > 0$, le medie integrali*

$$M_A(f) = \frac{1}{\mu(A)} \int_A f \, d\mu$$

appartengono ad S.

Allora $f(x) \in S$ per quasi ogni $x \in X$.

Dimostrazione. Essendo S^c un aperto, esiste una successione $B_1, B_2, \ldots$ di dischi aperti, $B_n \subset S^c$ tali che $\bigcup_n B_n = S^c$. Mostriamo che $f^{-1}(B_n)$ ha misura nulla per ogni n, da cui segue che $\mu(f^{-1}(S^c)) \leq \sum_{n=1}^{\infty} \mu(f^{-1}(B_n)) = 0$. Consideriamo allora $B = B(z, r) \subset S^c$. Sia $A = f^{-1}(B)$ e supponiamo, per assurdo, che $\mu(A) > 0$. Allora

$$|M_A(f) - z| = |\frac{1}{\mu(A)} \int_A (f - z) \, d\mu| \leq \frac{1}{\mu(A)} \int_A |f - z| \, d\mu \leq r,$$

in contraddizione con l'ipotesi che $M_A(f) \in S$ per ogni A di misura positiva. $\square$

Teorema 5.5.5 (Teorema di Vitali-Carathéodory)**.** *Sia $f\colon X \to \mathbb{R}$ una funzione $L^1(\mu)$, dove μ è una misura di Borel regolare tale che $\mu(K) < +\infty$ per ogni compatto $K \subset X$.*

Allora per ogni $\epsilon > 0$ esistono due funzioni u e $v\colon X \to \mathbb{R}$ tali che

(a) *$u(x) \leq f(x) \leq v(x)$ in tutto X;*

(b) *u è semicontinua superiormente e superiormente limitata;*

(c) *v è semicontinua inferiormente e inferiormente limitata;*

(d) *$\int_X (v - u) \, d\mu < \epsilon$.*

Dimostrazione. Possiamo sempre supporre f non identicamente nulla.

Consideriamo prima il caso $f \geq 0$. Sappiamo che in tal caso esiste una successione crescente di funzioni semplici e misurabili $s_n(x) \to f(x)$ in tutti i punti di X. Ponendo $s_0 = 0$ e $t_n(x) = s_n(x) - s_{n-1}(x)$, abbiamo che $f(x) = \sum_{n=1}^{\infty} t_n(x)$. Poiché t_n è una funzione semplice, possiamo anche scrivere $f(x) = \sum_{n=1}^{\infty} c_n \chi_{E_n}(x)$. Abbiamo immediatamente che

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} c_n \, \mu(E_n) < +\infty.$$

Essendo $\mu(E_n) < +\infty$ per ogni n, segue dalla regolarità della misura μ che per ogni n esistono un compatto K_n e un aperto V_n tali che $K_n \subset E_n \subset V_n$ con $\mu(V_n \setminus K_n) < \frac{\epsilon}{c_n 2^{n+1}}$. Con tale scelta è immediato verificare che la serie $\sum_{n=1}^{\infty} c_n \, \mu(V_n)$ è convergente, e quindi esiste $N \in \mathbb{N}$ tale che $\sum_{n=N+1}^{\infty} c_n \, \mu(V_n) < \epsilon/2$. Poniamo

$$v = \sum_{n=1}^{\infty} c_n \chi_{V_n} \qquad\qquad u = \sum_{n=1}^{N} c_n \chi_{K_n}.$$

Poiché v è l'estremo superiore di una famiglia di funzioni semicontinue inferiormente (le combinazioni lineari finite di funzioni caratteristiche di aperti) è semicontinua inferiormente (si veda a tale proposito la proposizione 1.5.4). Invece u è semicontinua superiormente in quanto combinazione lineare finita di funzioni semicontinue superiormente (le funzioni caratteristiche di chiusi). Poiché vale $\chi_{K_n}(x) \leq \chi_{E_n}(x) \leq \chi_{V_n}(x)$, si verifica subito che $u(x) \leq f(x) \leq v(x)$. Abbiamo poi che

$$\begin{aligned} v(x) - u(x) &= \sum_{n=1}^{N} c_n (\chi_{V_n}(x) - \chi_{K_n}(x)) + \sum_{n=N+1}^{\infty} c_n \chi_{V_n} \\ &\leq \sum_{n=1}^{\infty} c_n (\chi_{V_n}(x) - \chi_{K_n}(x)) + \sum_{n=N+1}^{\infty} c_n \chi_{V_n}, \end{aligned}$$

e quindi

$$\int_X (v-u) \, d\mu \leq \sum_{n=1}^{\infty} c_n \, \mu(V_n \setminus K_n) + \sum_{n=N+1}^{\infty} c_n \, \mu(V_n) < \epsilon.$$

Che u sia limitata superiormente segue subito dal fatto che u è una funzione semplice, mentre v è limitata inferiormente da 0.

Nel caso generale abbiamo $f = f^+ - f^-$. Poiché $f^{\pm}(x) \geq 0$, troviamo u_1 e u_2 semicontinue superiormente, v_1 e v_2 semicontinue inferiormente tali che $u_1 \leq f^+ \leq v_1$ e $u_2 \leq f^- \leq v_2$. È allora immediato verificare che le funzioni $u = u_1 - v_2$ e $v = v_1 - u_2$ sono le funzioni cercate. $\square$

5.6. Integrazione secondo Riemann e secondo Lebesgue

Mostriamo ora che ogni funzione integrabile secondo Riemann è anche integrabile secondo Lebesgue, e che in tal caso gli integrali coincidono. Indicheremo con $\int_E f \, dx$ l'integrale di Lebesgue di f sull'insieme E, e con $\mathcal{R}\int_a^b f \, dx$ l'integrale di Riemann di f nell'intervallo $[a,b]$. Data una funzione $f\colon [a,b] \to \mathbb{R}$ limitata e una partizione $\mathcal{S} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ dell'intervallo $[a,b]$, poniamo $m_i = \inf\{\, f(x) \mid x \in [x_{i-1}, x_i] \,\}$, $M_i = \sup\{\, f(x) \mid x \in [x_{i-1}, x_i] \,\}$, $\mathcal{R}_l(f,\mathcal{S}) = \sum_{i=1}^{n} m_i (x_{i+1} - x_i)$ e $\mathcal{R}_u(f,\mathcal{S}) =$

$\sum_{i=1}^n M_i(x_{i+1} - x_i)$. Ricordiamo che f è Riemann integrabile se per ogni $\epsilon > 0$ esiste una partizione $\mathcal{S}$ tale che $\mathcal{R}_u(f, \mathcal{S}) - \mathcal{R}_l(f, \mathcal{S}) < \epsilon$.

Teorema 5.6.1. *Supponiamo che $f\colon [a, b] \to \mathbb{R}$ sia una funzione limitata. Allora*

(a) *f è integrabile secondo Riemann se e solo se f è quasi ovunque continua (cioè tale che per quasi ogni $x \in [a, b]$ la funzione f è continua in x);*

(b) *se f è Riemann integrabile, allora è misurabile, Lebesgue integrabile e coincidono i valori dell'integrale secondo Riemann e secondo Lebesgue, cioè $\int_{[a,b]} f\,dx = \mathcal{R}\int_a^b f\,dx$.*

Dimostrazione. Consideriamo una partizione $\mathcal{S} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ dell'intervallo $[a, b]$. Posto $m_i = \inf\{\, f(x) \mid x \in [x_{i-1}, x_i]\,\}$ e $M_i = \sup\{\, f(x) \mid x \in [x_{i-1}, x_i]\,\}$, abbiamo che la funzione semplice $l(x) = \sum_{i=1}^n m_i \chi_{[x_{i-1},x_i]}$ è misurabile e tale che $\int_{[a,b]} l\,dx = \sum_{i=1}^n m_i(x_{i+1} - x_i) = \mathcal{R}_l(f, \mathcal{S})$. Analogamente, posto $u(x) = \sum_{i=1}^n M_i \chi_{[x_{i-1},x_i]}$, u è una funzione semplice misurabile tale che $\int_{[a,b]} u\,dx = \sum_{i=1}^n M_i(x_{i+1} - x_i) = \mathcal{R}_u(f, \mathcal{S})$. Ovviamente $l(x) \le u(x)$ in tutti i punti x.

Supponiamo che f sia Riemann integrabile. Dimostriamo che f è misurabile, integrabile secondo Lebesgue, quasi ovunque continua e che gli integrali secondo Riemann e Lebesgue coincidono.

Fissato $n \in \mathbb{N}$ troviamo una partizione $\mathcal{S}_n$ di $[a, b]$ tale che $\mathcal{R}_u(f, \mathcal{S}_n) - \mathcal{R}_l(f, \mathcal{S}_n) < \frac{1}{n}$. Ciò significa che $0 \le \int_{[a,b]}(u_n - l_n)\,dx < \frac{1}{n}$. Se supponiamo, come possiamo sempre fare, che $\mathcal{S}_n \subset \mathcal{S}_{n+1}$, è facile mostrare che l_n è una successione crescente e che u_n è una successione decrescente. Sia u il limite di u_n e l il limite delle l_n. Ovviamente u e v sono misurabili, e per il lemma di Fatou

$$0 \le \int_{[a,b]} (u - l)\,dx \le \liminf_n \int_{[a,b]} (u_n - l_n)\,dx = 0.$$

Quindi $u(x) = l(x)$ per quasi ogni x. Infine, da $u(x) \ge f(x) \ge l(x)$, abbiamo che $f(x) = u(x)$ per quasi ogni x, da cui segue la misurabilità di f. L'uguaglianza degli integrali è allora immediata. Sia Z l'insieme di misura nulla tale che $f(x) = l(x) = u(x)$ per ogni $x \in [a, b] \setminus Z$. Notiamo che anche l'insieme $\mathcal{S} = \bigcup_k \mathcal{S}_k$, in quanto insieme numerabile, è di misura nulla. Mostriamo che f è continua in tutti i punti di $[a, b] \setminus (\mathcal{S} \cup Z)$. Infatti, se non fosse continua in un tale x, esisterebbe un $\epsilon_0 > 0$ tale che $\sup_U f - \inf_U f > \epsilon_0$ qualunque sia l'intorno U di x. Ne segue che $u_k(x) - l_k(x) \ge \epsilon_0$ per ogni k e quindi che $u(x) \ne l(x)$, contro l'ipotesi $x \notin Z$.

Consideriamo ora una funzione f limitata e quasi ovunque continua. Mostriamo che è Riemann integrabile. Sia $\mathcal{S}_n$ una successione di partizioni di $[a, b]$ il cui diametro tende a zero. Definite come prima u_n e l_n, abbiamo che $u_n(x) - l_n(x) \to 0$ se $x \notin \bigcup_n \mathcal{S}_n$ e se x è un punto di continuità per f. Quindi u_n e l_n convergono a f in quasi ogni punto di $[a,b]$. Dal teorema della convergenza dominata abbiamo subito che $\int_{[a,b]}(u_n - l_n)\,d\mu \to 0$, ovvero l'integrabilità secondo Riemann. □

Abbiamo visto che l'integrabilità secondo Riemann implica quella secondo Lebesgue. L'integrabilità in senso improprio implica l'integrabilità secondo Lebesgue solo nel caso in cui la funzione considerata è non negativa (si veda l'esercizio 2 per un controesempio con una funzione che cambia segno). Infatti vale

Proposizione 5.6.2. *Sia $f\colon (a,b] \to [0,+\infty]$ una funzione integrabile secondo Riemann in nell'intervallo $[a+\epsilon, b]$ per ogni $\epsilon > 0$ e tale che*

$$\lim_{\epsilon\to 0^+} \mathcal{R}\int_{a+\epsilon}^{b} f\,dx = I \in [0,+\infty]\,.$$

Allora $f\colon [a,b] \to [0,+\infty]$ è misurabile e $\int_{[a,b]} f\,dx = I$.

Dimostrazione. È chiaro che per ogni $\epsilon > 0$ la funzione f_ϵ definita da $f_\epsilon(x) = f(x)$ per $x \in [a+\epsilon, b]$, $f_\epsilon(x) = 0$ in $[a, a+\epsilon)$ è misurabile (segue dal teorema precedente). Possiamo assumere che $f(a) = 0$. Allora $f_\epsilon(x) \to f(x)$ per ogni $x \in [a,b]$ e quindi f è misurabile. Inoltre f_ϵ è una successione non decrescente, e quindi possiamo applicare il teorema di convergenza monotona e dedurne che

$$\int_{[a,b]} f\,dx = \lim_{\epsilon\to 0^+}\int_{[a,b]} f_\epsilon\,dx = \lim_{\epsilon\to 0^+}\int_{[a+\epsilon,b]} f\,dx = \lim_{\epsilon\to 0^+}\mathcal{R}\int_{a+\epsilon}^{b} f\,dx = I.$$

□

5.7. Il termine mancante nel lemma di Fatou

Il teorema 5.5.1 è stato dimostrato usando il lemma di Fatou. È interessante notare che il teorema 5.5.1 può essere a sua volta utilizzato per generalizzare il lemma di Fatou nel modo seguente. Sia f_n una successione di funzioni nonnegative che convergono puntualmente a una funzione f. Come visto nell'osservazione 5.3.2, i limiti e gli integrali non possono essere scambiati in quanto, intuitivamente, la successione può "scappare all'infinito". Il teorema successivo è preso da [1] e rende precisa l'intuizione, dando un termine correttivo che cambia la diseguaglianza in un'uguaglianza. Anche se non verrà usato successivamente, è interessante di per sé come teorema di teoria della misura ed è stato efficacemente utilizzato per risolvere alcuni problemi del calcolo delle variazioni. Il lettore può consultare l'articolo originale per un versione piú generale in cui, fra le altre cose, $f \mapsto |f|^p$ è rimpiazzata da una classe piú ampia di funzioni, $f \mapsto j(f)$.

Teorema 5.7.1. *Sia f_j una successione di funzioni a valori complessi su uno spazio di misura che converge puntualmente q.o. a una funzione f. Si assuma, inoltre, che le f_j siano uniformemente p-integrabili per un qualche $0 < p < +\infty$, cioè che*

$$\int_X |f_j(x)|^p \, d\mu < C \qquad \textit{per} \quad j = 1, 2, \dots$$

e per qualche costante C. Allora

$$\lim_{j\to\infty} \int_X \big||f_j(x)|^p - |f_j(x) - f(x)|^p - |f(x)|^p\big| \, d\mu = 0. \tag{5.7.1}$$

Osservazione 5.7.2. Premettiamo alla dimostrazione le seguenti osservazioni.

(a) Per il lemma di Fatou, $\int |f|^p \, d\mu \leq C$;

(b) Applicando la diseguaglianza triangolare a (5.7.1) otteniamo che

$$\int |f_j|^p \, d\mu = \int |f|^p \, d\mu + \int |f - f_j|^p \, d\mu + o(1), \tag{5.7.2}$$

ove $o(1)$ indica un termine che tende a zero per $j \to \infty$. Quindi il termine correttivo è dato da $\int |f - f_j|^p \, d\mu$, che misura di quanto la successione "scappa". Una conseguenza ovvia della (5.7.2), per ogni $0 < p < \infty$, è che se $\int |f - f_j|^p \, d\mu \to 0$ e se $f_j \to f$ q.o., allora

$$\int |f_j|^p \, d\mu \to \int |f|^p \, d\mu.$$

(Questo si può dimostrare anche direttamente sotto la sola ipotesi che $\int |f - f_j|^p \, d\mu \to 0$. Quando $1 \leq p < \infty$ questo è una conseguenza banale della diseguaglianza di Minkowski (7.2.3). Quando $0 < p < 1$ segue dalla diseguaglianza elementare $|a + b|^p \leq |a|^p + |b|^p$ valida per tutti i numeri complessi a e b.) Un'altra conseguenza della (5.7.2), valida per $0 < p < \infty$, è che se $\int |f^j|^p \, d\mu \to \int |f|^p \, d\mu$ e $f^j \to f$ q.o., allora

$$\int |f - f^j|^p \, d\mu \to 0.$$

Dimostrazione. Assumiamo, per il momento, valida la seguente famiglia di diseguaglianze (5.7.3): per ogni $\epsilon > 0$ esiste una costante C_ϵ tale che per tutti i numeri $a, b \in \mathbb{C}$

$$\big||a + b|^p - |b|^p\big| \leq \epsilon |b|^p + C_\epsilon |a|^p. \tag{5.7.3}$$

Scriviamo $f_j = f + g_j$, cosí che $g_j \to 0$ puntualmente q.o. per ipotesi. Affermiamo che la quantità

$$G_{j,\epsilon} = \big(\big||f + g_j|^p - |g_j|^p - |f|^p\big| - \epsilon |g_j|^p\big)^+ \tag{5.7.4}$$

soddisfa $\lim_{j\to\infty} \int_X G_{j,\epsilon} \, d\mu = 0$. Qui, come al solito, $(h)^+$ denota la parte positiva della funzione h. Per mostrare ciò, osserviamo innanzi tutto che

$$\begin{aligned} \big||f + g_j|^p - |g_j|^p - |f|^p\big| &\leq \big||f + g_j|^p - |g_j|^p\big| + |f|^p \\ &\leq \epsilon |g_j|^p + (1 + C_\epsilon)|f|^p \end{aligned}$$

e quindi $G_{j,\epsilon} \leq (1 + C_\epsilon)|f|^p$. Inoltre $G_{j,\epsilon} \to 0$ puntualmente q.o. e quindi l'affermazione segue direttamente dal teorema della convergenza dominata 5.5.1. Poi

$$\int_X \big||f + g_j|^p - |g_j|^p - |f|^p\big| \, d\mu \leq \epsilon \int_X |g_j|^p \, d\mu + \int_X G_{j,\epsilon} \, d\mu.$$

Dobbiamo mostrare che $\int_X |g_j| \, d\mu$ è uniformemente limitato. Infatti

$$\int_X |g_j|^p \, d\mu = \int_X |f - f_j|^p \, d\mu \leq 2^p \int_X (|f|^p + |f_j|^p) \, d\mu \leq 2^{p+1} C.$$

Quindi

$$\limsup_{j \to +\infty} \int_X \big||f + g_j|^p - |g_j|^p - |f|^p\big| \, d\mu \leq \epsilon D,$$

e il teorema segue per l'arbitrarietà di ϵ.

Resta da dimostrare la (5.7.3). La funzione $t \mapsto |t|^p$ è convessa se $p > 1$. Quindi $|a + b|^p \leq (|a| + |b|)^p = ((1 - \lambda)\frac{|a|}{1-\lambda} + \lambda\frac{|b|}{\lambda})^p \leq (1 - \lambda)^{1-p}|a|^p + \lambda^{1-p}|b|^p$ per ogni $0 < \lambda < 1$. La scelta $\lambda = (1 + \epsilon)^{-1/(p-1)}$ da la (5.7.3) nel caso $p > 1$. Se $0 < p \leq 1$ abbiamo la semplice diseguaglianza $|a + b|^p - |b|^p \leq |a|^p$ la cui dimostrazione è lasciata al lettore. □

Esercizi Capitolo 5

(1) Sia X un insieme e $\#$ la misura che ad ogni sottoinsieme $A \subset X$ associa la sua cardinalità $\#(A)$ (si veda l'esercizio 4). Data $f\colon X \to [0, +\infty]$, mostrare che $\int_X f \, d\# = \sum_{x \in X} f(x)$. In che senso va intesa tale somma se X non è numerabile? Mostrare che $f \in L^1(X)$ implica che $f(x) \neq 0$ solo per un insieme al piú numerabile di punti $x \in X$.

(2) Consideriamo la funzione $f(x) = \frac{1}{x} \sin \frac{\pi}{x}$. Mostrare che
   (a) $f \in L^1([\epsilon, 1])$ per tutti gli $0 < \epsilon < 1$;
   (b) $\lim_{\epsilon \to 0} \int_\epsilon^1 f(x) \, dx$ esiste finito;
   (c) $f \notin L^1([0, 1])$.

(3) Sia $f\colon \mathbb{R}^2 \to \mathbb{R}$ una funzione continua. Dimostrare che la funzione

$$g(x) = \int_0^1 f(x, y) \, dy$$

è continua.

(4) Sia $f\colon \mathbb{R}^2 \to \mathbb{R}$ una funzione tale che
   (a) $\forall x \in \mathbb{R}$ $f(x, \cdot)$ è misurabile e $|f(x, y)| \leq g(y)$ dove $g \in L^1(\mathbb{R})$;
   (b) per quasi ogni $y \in \mathbb{R}$ $f(x, y)$ è continua in x.
   Dimostrare che $g(x) = \int_{\mathbb{R}} f(x, y) \, dy$ è una funzione continua.

(5) Sia $g\colon \mathbb{R}^2 \to \mathbb{R}$ tale che
   (a) per ogni $x \in \mathbb{R}$ $g(x, y)$ è misurabile in y e $\int_0^1 |g(x, y)| \, dy < +\infty$.

(b) per quasi ogni $y \in \mathbb{R}$ $g(x,y)$ è derivabile in x e $|\frac{\partial g}{\partial x}(x,y)| \leq h(y)$, $h \in L^1([0,1])$.

Mostrare che

- per ogni $x \in \mathbb{R}$ $\frac{\partial g}{\partial x}(x,y)$ è misurabile in y;
- $f(x) = \int_0^1 g(x,y)\,dy$ è derivabile in x e

$$f'(x) = \int_0^1 \frac{\partial g}{\partial x}(x,y)\,dy.$$

(6) Sia $f_1, f_2, \dots$ una successione di funzioni nonnegative e misurabili definite in $E \subset \mathbb{R}$. Si assuma che $f_k(x) \to f(x)$ e $f_k(x) \leq f(x)$ q.o. in E. Si dimostri che $\int_E f_k \to \int_E f$.

(7) Sia $f_1, f_2, \dots : X \to [-\infty, +\infty]$ una successione di funzioni misurabili tali che

(a) $f_n(x) \geq f_{n+1}(x)$ per q.o. $x \in X$;

(b) $f_1 \in L^1(\mu)$.

Si dimostri che

$$\int_X f_n\,d\mu \to \int_X f\,d\mu.$$

Si mostri anche che la tesi è falsa se 7b non vale.

(8) Sia $f\colon [a,b] \to [0,+\infty)$ una funzione misurabile e limitata. Si dimostri che

$$\int_{[a,b]} f\,d\mu = \inf\Big\{ \int_{[a,b]} s\,d\mu \;\Big|\; s \text{ semplice, } f \leq s \Big\}.$$

(Suggerimento: si cerchi di costruire una successione di funzioni semplici che convergono a f in modo monotono decrescente).

(9) Sia f una funzione misurabile e non-negativa sullo spazio di misura (X,μ). Mostrare che se f è integrabile, allora

$$\lim_{\lambda \to +\infty} \lambda \mu(\{f \geq \lambda\}) = 0. \tag{5.7.5}$$

Si mostri poi un esempio di una funzione non negativa, misurabile ma non integrabile in $[0,1]$ per cui valga la (5.7.5).

(10) Si dimostri l'assoluta continuità dell'integrale: $f \in L^1(\mu)$ implica che per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che $|\int_A f\,d\mu| < \epsilon$ qualunque sia l'insieme A misurabile tale che $\mu(A) < \delta$.

(11) Per quali $p, q \in (0,+\infty)$ si ha che $x^{-p}\log^{-q}(x) \in L^1([0,1/2])$?

(12) Sia $g_1, g_2, \dots : [0,+\infty) \to [0,+\infty]$ una successione di funzioni non crescenti convergenti puntualmente alla funzione g. Si dimostri che

$$\int_0^{+\infty} g_n(x)\,dx \to \int_0^{+\infty} g(x)\,dx$$

anche se l'integrale è quello di Riemann.

Bibliografia

[1] H. Brezis and E. H. Lieb, *A relation between pointwise convergence of functions and convergence of functionals*, Proc. Amer. Math. Soc. **88** (1983), 486–490.

[2] G. Cantor, *Über unendliche, linear Punktmannichfaltigkeiten*, Math. Ann. **20** (1882), 113–121.

[3] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics, no. 14, American Mathematical Society, 1997.

[4] G. Prodi, *Istituzioni di matematica*, McGraw-Hill Libri Italia, Milano, 1994.

[5] W. Rudin, *Real and complex analysis*, iii ed., McGraw-Hill, New York, 1966.

[6] R. L. Wheeden and A. Zygmund, *Measure and integral*, Marcel Dekker, New York, 1977.

Indice analitico