Metodi Matematici per l'Ingegneria

Davide Guidetti
Dipartimento di Matematica
Piazza di Porta S. Donato, 5
40127 Bologna

0

Preliminari

In questo capitolo raccogliamo alcuni risultati preliminari al corso vero e proprio, che lo studente forse già conosce, magari in una forma un po' diversa. Può essere però utile disporre di un riferimento facilmente consultabile, e questo è lo scopo che qui ci prefiggiamo.

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

In questa sezione vogliamo fornire i rudimenti della teoria della misura e integrazione secondo Lebesgue. Poiché un sottoinsieme di $\mathbf{R}^n$ può avere misura infinita, è conveniente estendere le solite operazioni di somma e prodotto tra numeri reali non negativi al caso in cui un addendo o un fattore sia $+\infty$. Poniamo allora

$$[0, +\infty] := [0, +\infty[\cup \{+\infty\}, \tag{0.1.1}$$

ove $+\infty$ è un oggetto non appartenente a $[0, +\infty[$. Ammettiamo che

$$a < +\infty \qquad \forall a \in [0, +\infty[. \tag{0.1.2}$$

Poniamo poi, dato $a \in [0, +\infty]$,

$$a + (+\infty) = (+\infty) + a = +\infty \tag{0.1.3}$$

e

$$a \cdot (+\infty) = (+\infty) \cdot a = \begin{cases} +\infty & \text{se} \quad a \neq 0, \\ 0 & \text{se} \quad a = 0. \end{cases} \tag{0.1.4}$$

Se $(a_n)_{n\in\mathbf{N}}$ è una successione a valori in $[0,+\infty]$, definiamo la somma della serie

$$\sum_{n=1}^{\infty} a_n = \begin{cases} \text{solito se la serie converge in } \mathbf{R}, \\ +\infty \qquad \text{altrimenti.} \end{cases} \tag{0.1.5}$$

Saltuariamente utilizzeremo anche il simbolo $-\infty$, che penseremo soddisfacente

$$-\infty < a \qquad \forall a \in \mathbf{R}. \tag{0.1.6}$$

Porremo

$$[-\infty,+\infty] := \mathbf{R} \cup \{-\infty,+\infty\}. \tag{0.1.7}$$

Costruiamo ora la misura di Lebesgue partendo dagli intervalli limitati in $\mathbf{R}^n$ di cui definiamo il volume $n-$dimensionale vol_n.

Definizione 0.1.1 *Un intervallo limitato in* $\mathbf{R}^n$ è *il prodotto cartesiano di* n *intervalli limitati in* $\mathbf{R}$.

Se I è un intervallo limitato in $\mathbf{R}$, poniamo

$$\ell(I) := \begin{cases} \sup I - \inf I & \text{se} \quad I \neq \emptyset, \\ 0 & \text{se} \quad I = \emptyset. \end{cases} \tag{0.1.8}$$

Se I è un intervallo limitato in $\mathbf{R}^n$ e $I = I_1 \times ... \times I_n$, poniamo

$$vol_n(I) := \ell(I_1) \cdot .... \cdot \ell(I_n). \tag{0.1.9}$$

Definiamo ora la misura esterna di un generico sottoinsieme di $\mathbf{R}^n$.

Definizione 0.1.2 *Sia* $A \subseteq \mathbf{R}^n$. *Definiamo la misura esterna* $L_n^*(A)$ *di* A *come segue:*

$$\begin{gathered} L_n^*(A) := \inf\{\textstyle\sum_{k=1}^{\infty} vol_n(I^k) : I^k \text{intervallo } n- \text{ dimensionale } \forall k \in \mathbf{N}, \\ A \subseteq \cup_{k\in\mathbf{N}} I^k\}. \end{gathered} \tag{0.1.10}$$

Osserviamo che, se $L_n^*(A) = +\infty$ se, comunque si prenda una famiglia numerabile $(I^k)_{k\in\mathbf{N}}$ di intervalli limitati $n-$dimensionali che ricopre A, vale $\sum_{k=1}^{\infty} vol_n(I^k) = +\infty$.

La misura esterna L_n^* è per certi aspetti insoddisfacente. Infatti, sarebbe auspicabile che, dati due sottoinsiemi arbitrari e disgiunti A e B di $\mathbf{R}^n$, valesse

$$L_n^*(A \cup B) = L_n^*(A) + L_n^*(B).$$

e ciò non accade in generale. Si può solo dire che vale sempre

$$L_n^*(A \cup B) \leq L_n^*(A) + L_n^*(B),$$

ma anche nel caso in cui A e B sono disgiunti può accadere che valga la disuguaglianza stretta. Per ovviare a questo inconveniente, ci si limita a considerare una sottoclasse di sottoinsiemi di $\mathbf{R}^n$, i cosi detti *insiemi misurabili secondo Lebesgue*, in cui le cose vanno bene.

Definizione 0.1.3 *Sia $A \subseteq \mathbf{R}^n$. Diremo che A è* **misurabile** *secondo Lebesgue se $\forall E \subseteq \mathbf{R}^n$*

$$L_n^*(E) = L_n^*(E \cap A) + L_n^*(E \cap (\mathbf{R}^n \setminus A)).$$

Indichiamo con $\mathcal{M}_n$ la classe degli insiemi misurabili secondo Lebesgue in $\mathbf{R}^n$. Questa classe di insiemi è molto duttile rispetto alle solite operazioni insiemistiche:

Teorema 0.1.1 *(I) $\emptyset$ e $\mathbf{R}^n$ appartengono a $\mathcal{M}_n$;*

(II) se A e B appartengono a $\mathcal{M}_n$, allora $A \setminus B$ appartiene a $\mathcal{M}_n$;

(III) l'unione di una famiglia finita o numerabile di elementi di $\mathcal{M}_n$ appartiene a $\mathcal{M}_n$;

(IV) l'intersezione di una famiglia non vuota finita o numerabile di elementi di $\mathcal{M}_n$ appartiene a $\mathcal{M}_n$.

Definiamo ora la **misura di Lebesgue** L_n come la restrizione di L_n^* a $\mathcal{M}_n$.

Vale il seguente

Teorema 0.1.2 *Sia $\mathcal{I}$ un insieme del tipo $\{1, ...n\}$ per qualche $n \in \mathbf{N}$, oppure l'insieme dei numeri naturali $\mathbf{N}$. Allora:*

(I) $L_n(\emptyset) = 0$;

(II) se $A_i \in \mathcal{M}_n$ $\forall i \in \mathcal{I}$*, si ha*

$$L_n(\cup_{i\in\mathcal{I}} A_i) \leq \sum_{i\in\mathcal{I}} L_n(A_i);$$

(si osservi che $\cup_{i\in\mathcal{I}} A_i$ è *misurabile per il teorema 0.1.1(III));*
(III) se, di più, gli A_i *sono a due a due disgiunti, si ha*

$$L_n(\cup_{i\in\mathcal{I}} A_i) = \sum_{i\in\mathcal{I}} L_n(A_i);$$

(IV) se A *e* B *sono elementi di* $\mathcal{M}_n$ *e* $A \subseteq B$*, vale*

$$L_n(A) \leq L_n(B);$$

(V) se, di più , $L_n(B) < +\infty$*, vale*

$$L_n(B \setminus A) = L_n(B) - L_n(A).$$

Il risultato più interessante del teorema 0.1.2 è probabilmente (III): se gli A_i, a due a due disgiunti, sono in numero finito o al più un'infinità numerabile, la misura dell'unione coincide con la somma delle misure. Ci si può chiedere infine se, dato un sottoinsieme di $\mathbf{R}^n$, esistano dei criteri per stabilire se è misurabile. Il seguente risultato risponde a questa esigenza.

Teorema 0.1.3 *(I) Se* I è *un intervallo limitato* $n-$*dimensionale, allora* $I \in \mathcal{M}_n$ *e* $L_n(I) = vol_n(I)$*;*
(II) ogni sottoinsieme aperto di $\mathbf{R}^n$ è *misurabile;*
(III) ogni sottoinsieme chiuso di $\mathbf{R}^n$ è *misurabile.*

Esercizio 0.1.1 Sia A un sottoinsieme finito o numerabile di $\mathbf{R}^n$. Dimostrare che A è misurabile secondo Lebesgue e ha misura nulla.

Esercizio 0.1.2 Sia A un sottoinsieme di $\mathbf{R}^n$ tale che $L_n^*(A) = 0$. Dimostrare che A è misurabile secondo Lebesgue.

0.2 Teoria dell'integrazione secondo Lebesgue

Dopo la misura, passiamo all'integrazione in $\mathbf{R}^n$. Cominceremo con l'integrare le così dette *funzioni semplici* non negative.

Definizione 0.2.1 *Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$. Diremo che f* è *una funzione* **semplice** *se esistono $A_1, ..., A_m$ ($m \in \mathbf{N}$) sottoinsiemi misurabili di A a due a due disgiunti, la cui unione* è *A, e dei numeri reali reale $\alpha_1, ..., \alpha_m$ in modo che,*

$$f(x) = \alpha_i \forall x \in A_i, 1 \leq i \leq m.$$

Definizione 0.2.2 *Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$ semplice come nella definizione 0.2.1, con $\alpha_i \geq 0$ $\forall i \in \{1, ..., m\}$. Poniamo*

$$\int_A f(x)dx := \sum_{i=1}^{m} \alpha_i L_n(A_i).$$

Si può dimostrare che la definizione 0.2.2 è ben posta, nel senso che non dipende da come si decompone A, purché si conservi la proprietà che su ciascuna delle parti f è costante. Nel caso in cui $\alpha_i = 0$, in base alla definizione (0.1.4), il prodotto $\alpha_i L_n(A_i)$ vale comunque 0, anche nel caso in cui $L_n(A_i) = +\infty$. Ciò ha come conseguenza il fatto che su un insieme misurabile qualunque, anche di misura infinita, l'integrale della funzione identicamente nulla vale 0. In effetti, la scelta di (0.1.4) era proprio finalizzata a questo scopo.

Definiamo ora una classe molto vasta di funzioni, le così dette *funzioni misurabili.*

Definizione 0.2.3 *Siano $A \in \mathcal{M}_n$, $f : A \to [-\infty, +\infty]$ (vedi (0.1.7)). Diremo che f* è **misurabile** *se esiste una successione di funzioni semplici $(f_k)_{k \in \mathbf{N}}$ di dominio A tale che*

$$\lim_{k \to +\infty} f_k(x) = f(x) \qquad \forall x \in A.$$

È chiaro dalla definizione che ogni funzione semplice è misurabile. Ma, ad esempio, si può vedere che

Teorema 0.2.1 *Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$ continua. Allora f* è *misurabile.*

Anche la classe delle funzioni misurabili dispone di una notevole duttilità rispetto alle operazioni standard tra funzioni:

Teorema 0.2.2 *Siano* $A \in \mathcal{M}_n$, $f, g : A \to \mathbf{R}$ *misurabili,* $\phi : \mathbf{R} \to \mathbf{R}$ *continua. Allora:*

(I) $f + g$, fg, $\phi \circ f$ *sono misurabili;*
(II) se $g(x) \neq 0\ \forall x \in A$, *allora* $\frac{f}{g}$ è *misurabile;*
(III) $\phi \circ f$ è *misurabile.*

Da (I), prendendo $\phi(y) = cy$, segue in particolare che, se f è misurabile e a valori reali, lo è anche $cf\ \forall c \in \mathbf{R}$. Definiamo ora l'integrale per una funzione misurabile non negativa.

Definizione 0.2.4 *Siano* $A \in \mathcal{M}_n$, $f : A \to [0, +\infty]$. *Poniamo*

$$\int_A f(x)dx$$

$$:= \sup\{\int_A \phi(x)dx : \phi : A \to [0, +\infty[\text{ semplice }, 0 \leq \phi(x) \leq f(x) \forall x \in A\}.$$

Si può vedere che, nel caso in cui f sia semplice non negativa, la definizione 0.2.4 coincide con la definizione 0.2.2.

Enunciamo ora alcune proprietà fondamentali dell'integrale di una funzione misurabile non negativa.

Teorema 0.2.3 *Siano* $A \in \mathcal{M}_n$, $f, g : A \to [0, +\infty]$ *misurabili,* $\alpha \in [0, +\infty[$. *Allora:*

(I) $\int_A f(x)dx \in [0, +\infty]$;
(II) se $f(x) \leq g(x)\ \forall x \in A$, *si ha anche* $\int_A f(x)dx \leq \int_A g(x)dx$;
(III) $\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx$;
(IV) $\int_A \alpha f(x)dx = \alpha \int_A f(x)dx$;
(V) se $A = A_1 \cup ... \cup A_m$, *con* $A_1, ..., A_m$ *misurabili e a due a due disgiunti, allora*

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f(x)dx,$$

ove $\int_{A_j} f(x)dx := \int_{A_j} f_{|A_j}(x)dx$ *(si veda il seguente esercizio 0.2.1).*

Passiamo ora all'integrazione di funzioni che possono assumerare valori di segno qualunque. Cominciamo con la seguente

Definizione 0.2.5 *Siano* $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$. *Diremo che* f è **sommabile** *se* è *misurabile e*

$$\int_A |f(x)|dx < +\infty$$

(questo integrale ha senso in virtù del teorema 0.2.2 (III)*).*

Consideriamo ora le due funzioni:

$$\begin{cases} \phi_+ : \mathbf{R} \to \mathbf{R}, \\ \phi_+(x) = \begin{cases} x & \text{se} \quad x \geq 0, \\ 0 & \text{se} \quad x < 0, \end{cases} \end{cases} \tag{0.2.1}$$

$$\begin{cases} \phi_- : \mathbf{R} \to \mathbf{R}, \\ \phi_-(x) = \begin{cases} -x & \text{se} \quad x \leq 0, \\ 0 & \text{se} \quad x > 0. \end{cases} \end{cases} \tag{0.2.2}$$

È facile verificare che ϕ_+ e ϕ_- sono continue e non negative, $\phi_+(x)-\phi_-(x) = x \ \forall x \in \mathbf{R}$, $\phi_+(x) + \phi_-(x) = |x| \ \forall x \in \mathbf{R}$.

Sia ora $f : A \to \mathbf{R}$ misurabile. Poniamo

$$f_+ := \phi_+ \circ f, \qquad f_- := \phi_- \circ f. \tag{0.2.3}$$

Dal teorema 0.2.2 (III) si ha che f_+ e f_- sono misurabili non negative. Inoltre,

$$f = f_+ - f_-, \qquad |f| = f_+ + f_-. \tag{0.2.4}$$

Dal teorema 0.2.3(III) si ha che

$$\int_A |f(x)|dx = \int_A f_+(x)dx + \int_A f_-(x)dx. \tag{0.2.5}$$

Dunque, se f è sommabile,gli integrali a secondo membro di (0.2.5) sono entrambi reali. Tenuto conto allora della prima formula in (0.2.4), diviene naturale la seguente

Definizione 0.2.6 *Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$ sommabile. Poniamo*

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

Il teorema 0.2.3 ammette la seguente estensione alle funzioni sommabili:

Teorema 0.2.4 *Siano $A \in \mathcal{M}_n$, $f, g : A \to \mathbf{R}$ sommabili, $\alpha \in \mathbf{R}$. Allora:*
(I) se $f(x) \leq g(x)$ $\forall x \in A$, si ha anche $\int_A f(x)dx \leq \int_A g(x)dx$;
(II) $f + g$ è sommabile e

$$\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx;$$

(III) αf è *sommabile e*

$$\int_A \alpha f(x)dx = \alpha \int_A f(x)dx;$$

(IV) se $A = A_1 \cup ... \cup A_m$, *con* $A_1,...,A_m$ *misurabili e a due a due disgiunti, allora* $f_{|A_j}$ è *sommabile su* A_j *per* $j = 1, ..., m$ *e vale*

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f_{|A_j}(x)dx; \tag{0.2.6}$$

viceversa, se per $j = 1, ..., m$ $f_{|A_j}$ è *sommabile su* A_j, *allora* f è *sommabile su* A *e vale (0.2.6).*

In molte circostanze è importante disporre di una definizione di integrale per funzioni a valori complessi.

Definizione 0.2.7 *Siano* $A \in \mathcal{M}_n$, $f : A \to \mathbf{C}$. *Diremo che* f è **sommabile** *se lo sono le due funzioni a valori reali* $Re(f)$ *e* $Im(f)$. *In tal caso, poniamo*

$$\int_A f(x)dx := \int_A Re(f(x))dx + i \int_A Im(f(x))dx.$$

Parte del teorema 0.2.4 può essere estesa a funzioni a valori complessi:

Teorema 0.2.5 *Siano* $A \in \mathcal{M}_n$, $f, g : A \to \mathbf{C}$ *sommabili,* $\alpha \in \mathbf{C}$. *Allora:*

(I) $f + g$ è *sommabile e*

$$\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx;$$

(II) αf è *sommabile e*

$$\int_A \alpha f(x)dx = \alpha \int_A f(x)dx;$$

(III) se $A = A_1 \cup ... \cup A_m$, *con* $A_1,...,A_m$ *misurabili e a due a due disgiunti, allora* $f_{|A_j}$ è *sommabile su* A_j *per* $j = 1, ..., m$ *e vale*

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f_{|A_j}(x)dx; \tag{0.2.7}$$

viceversa, se per $j = 1, ..., m$ $f_{|A_j}$ è *sommabile su* A_j, *allora* f è *sommabile su* A *e vale (0.2.7);*

(IV) $|f|$ è *misurabile,* $\int_A |f(x)|dx < +\infty$ *e vale la disuguaglianza (0.2.8).*

Esercizio 0.2.1 Siano $A, B \in \mathcal{M}_n$ con $A \subseteq B$, $f : B \to [-\infty, +\infty]$ misurabile. Dimostrare che$f_{|A}$ è misurabile in A.

Esercizio 0.2.2 Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{R}$ sommabile. Dimostrare che

$$|\int_A f(x)dx| \leq \int_A |f(x)|dx. \tag{0.2.8}$$

Esercizio 0.2.3 Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{C}$ sommabile. Dimostrare che $|f|$ è sommabile su A (sugg.: $|f| \leq |Re(f)| + |Im(f)|$).

Esercizio 0.2.4 Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{C}$. Dimostrare che f è sommabile se e solo se $Re(f)$ e $Im(f)$ sono misurabili e $|f|$ è sommabile.

Esercizio 0.2.5 Siano $A_1, ..., A_m \in \mathcal{M}_n$, a due a due disgiunti, $f : \bigcup\limits_{1 \leq i \leq m} A_i$ $\to [-\infty, +\infty]$ tali che $f_{|A_i}$ è misurabile su A_i per ogni $i \in \{1, ..., m\}$. Dimostrare che f è misurabile.

Esercizio 0.2.6 Siano $A \in \mathcal{M}_n$, $f : A \to \mathbf{C}$. Dimostrare che, se f è sommabile, allora $\overline{f}$ è sommabile (ove, dato $z \in \mathbf{C}$, indichiamo con $\overline{z}$ il complesso coniugato di z). Vale inoltre

$$\int_A \overline{f(x)}dx = \overline{\int_A f(x)dx}. \tag{0.2.9}$$

0.3 Tecniche operative per il calcolo di integrali

In questa sezione presentiamo alcuni risultati di base atti a rendere spesso possibile il calcolo effettivi di integrali nel senso di Lebesgue. Presupponiamo nel lettore la conoscenza dell'integrale di Riemann per funzioni di una variabile su un intervallo chiuso e limitato.

Il primo risultato è in seguente

Teorema 0.3.1 *Siano $A \in \mathcal{M}_n$, con $L_n(A) = 0$, $f : A \to [0, +\infty]$. Allora f è misurabile e*

$$\int_A f(x)dx = 0. \tag{0.3.1}$$

Se $f : A \to \mathbf{C}$, allora f è sommabile e vale (0.3.1).

Dimostrazione Vedi gli esercizi 0.1.2. □

Consideriamo adesso il caso di A intervallo chiuso e limitato.

Teorema 0.3.2 *Sia a e b numeri reali, con $a < b$, $f : [a, b] \to \mathbf{R}$ integrabile secondo Riemann. Allora f è sommabile su $[a, b]$ e l'integrale nel senso di Lebesgue (definizione 0.2.6) coincide con l'integrale di Riemann.*

Passiamo ora a un risultato su intervalli semiaperti.

Teorema 0.3.3 *Siano $-\infty < a < b \leq +\infty$, $f : [a, b[\to [0, +\infty[$ integrabile secondo Riemann su ogni intervallo $[a, c]$ con $c \in]a, b[$. Allora f è misurabile non negativa in $[a, b[$ e vale*

$$\int_{[a,b[} f(x)dx = \lim_{c \to b} \int_a^c f(x)dx. \tag{0.3.2}$$

Osserviamo che il limite in (0.3.2) esiste (finito o infinito perché la funzione integrale $c \to \int_a^c f(x)dx$ è non decrescente in $[a, b[$. Un risultato analogo vale per funzioni definite su intervalli del tipo $]a, b]$, con $-\infty \leq a < b < +\infty$ (vedi l'esercizio 0.3.3).

Passiamo ora a funzioni a valori reali qualunque.

Teorema 0.3.4 *Siano $-\infty < a < b \leq +\infty[$, $f : [a, b[\to \mathbf{R}$ integrabile su $[a, c[$ per ogni $c \in]a, b[$ e $\lim_{c \to b} \int_a^c |f(x)|dx < +\infty$. Allora:*

(I) esiste in $\mathbf{R}$ $\lim_{c \to b} \int_a^c f(x)dx$;

(II) f è sommabile su $[a, b[$;

(III) $\int_{[a,b[} f(x)dx = \lim_{c \to b} \int_a^c f(x)dx$.

Passiamo adesso al caso multidimensionale con i classici teoremi di Tonelli, Fubini e del cambiamento di variabile. Premettiamo alcune notazioni. Supponiamo di lavorare in $\mathbf{R}^{m+n}$, con m e n in $\mathbf{N}$. Indichiamo con (x, y) il generico elemento di $\mathbf{R}^{m+n}$, con $x \in \mathbf{R}^m$ e $y \in \mathbf{R}^n$. Dato $x \in R^m$, poniamo

$$A_x := \{y \in \mathbf{R}^n : (x, y) \in A\}. \tag{0.3.3}$$

Teorema 0.3.5 *(di Tonelli) Siano $A \in \mathcal{M}_{m+n}$, $f : A \to [0, +\infty]$ misurabile. Sia poi $B \in \mathcal{M}_m$ tale che*

$$\{x \in \mathbf{R}^m : A_x \in \mathcal{M}_n, L_n(A_x) > 0\} \subseteq B.$$

Definiamo

$$\begin{cases} g : B \to [0, +\infty], \\ g(x) = \begin{cases} \int_{A_x} f(x,y)dy \text{ se } A_x \neq \emptyset, A_x \in \mathcal{M}_n, f(x,.) \text{ misurabile su } A_x, \\ 0 \text{ altrimenti.} \end{cases} \end{cases}$$

Allora:

(I) g è *misurabile su* B*;*

(II) $\int_A f(x,y)dxdy = \int_B g(x)dx$.

Teorema 0.3.6 *(di Fubini) Siano* $A \in \mathcal{M}_{m+n}$*,* $f : A \to \mathbf{C}$ *sommabile. Sia poi* $B \in \mathcal{M}_m$ *tale che*

$$\{x \in \mathbf{R}^m : A_x \in \mathcal{M}_n, L_n(A_x) > 0\} \subseteq B.$$

Definiamo

$$\begin{cases} g : B \to \mathbf{C}, \\ g(x) = \begin{cases} \int_{A_x} f(x,y)dy \text{ se } A_x \neq \emptyset, A_x \in \mathcal{M}_n, f(x,.) \text{ sommabile su } A_x, \\ 0 \text{ altrimenti.} \end{cases} \end{cases}$$

Allora:

(I) g è *sommabile su* B*;*

(II) $\int_A f(x,y)dxdy = \int_B g(x)dx$.

Passiamo ora al teorema di cambiamento di variabile.

Definizione 0.3.1 *Siano* Ω *un aperto in* $\mathbf{R}^n$*,* $T : \Omega \to \mathbf{R}^n$*. Diremo che* T è *un* **cambiamento di variabile** *se*

(I) T è *iniettiva;*

(II) T è *di classe* C^1*;*

(III) per ogni $x \in \Omega$ *il determinante* $det J_T(x)$ *della matrice jacobiana di* T *in* x è *diverso da* 0.

Teorema 0.3.7 *(del cambiamento di variabile) Siano* $T : \Omega \to \mathbf{R}^n$ *un cambiamento di variabile, con* Ω *aperto in* $\mathbf{R}^n$*,* $A \in \mathcal{M}_n$*,* $f : A \to [0, +\infty]$*, (risp.* $f : A \to \mathbf{C}$*). Supponiamo che* $A \subseteq T(\Omega)$*. Allora:*

(I) $T^{-1}(A) \in \mathcal{M}_n$*;*

(II) f è *misurabile su* A *(risp.* f è *sommabile su* A*) se e solo se* $x \to f(T(x))|det J_T(x)|$ è *misurabile (risp. sommabile) su* $T^{-1}(A)$*;*

(III) $\int_A f(y)dy = \int_{T^{-1}(A)} f(T(x))|det J_T(x)|dx$.

Esercizio 0.3.1 Dimostrare il teorema 0.3.1, ammettendo che ogni funzione di dominio A sia misurabile (sugg.: utilizzare il risultato dell' esercizio 0.1.2; cominciare a considerare il caso di f semplice non negativa).

Esercizio 0.3.2 Sia $f : [0,1] \to \mathbf{R}$, $f(x) = 1$ se $x \in \mathbf{Q}$, $f(x) = 0$ altrimenti. Verificare che f è sommabile su $[0,1]$ e ha integrale nullo (sugg.: utilizzare il risultato dell'esercizio 0.1.1). f è la cosi detta "funzione di Dirichlet", che non è integrabile secondo Riemann. Questo esempio prova che l'enunciato del teorema 0.3.2 non è invertibile.

Esercizio 0.3.3 Enunciare l'analogo del teorema 0.3.3 per un intervallo del tipo $]a,b]$, con $-\infty \leq a < b < +\infty$.

Esercizio 0.3.4 Estendere il teorema 0.3.4 al caso di funzioni a valori complessi.

0.4 Identificazione di funzioni misurabili coincidenti quasi dappertutto

Cominciamo con la seguente definizione, piuttosto suggestiva:

Definizione 0.4.1 *Sia $A \in \mathcal{M}_n$. Diremo che una certa proprietà $P(x)$ vale* **quasi dappertutto** *su A o per quasi ogni $x \in A$ se $\{x \in A : P(x)$ non vale $\}$ è misurabile e ha misura 0.*

Date f e g da A a $\mathbf{C}$, diremo che f è *equivalente a g* e scriveremo

$$f \sim g$$

se $f(x) = g(x)$ q. d. in A. Si verifica facilmente che $\sim$ è una relazione di equivalenza, vale a dire, date tre funzioni qualunque f, g e h da A a $\mathbf{C}$, si ha

(I) $f \sim f$ (proprietà riflessiva);

(II) se $f \sim g$, allora $g \sim f$ (proprietà simmetrica);

(III) se $f \sim g$ e $g \sim h$, allora $f \sim h$ (proprietà transitiva).

Data $f : A \to \mathbf{C}$, si definisce la *classe di equivalenza di* $[f]$ di f come

$$[f] := \{g : A \to \mathbf{C} : g \sim f\}. \tag{0.4.1}$$

Il seguente risultato sarà importante nel seguito:

Teorema 0.4.1 *Siano* $A \in \mathcal{M}_n$, $f, g, f_1, f_2, g_1, g_2 : A \to \mathbf{C}$, *con* $f \sim g$, $f_1 \sim g_1$, $f_2 \sim g_2$, $c \in \mathbf{C}$. *Allora:*

(I) $f_1 + g_1 \sim f_2 + g_2$;

(II) $cf \sim cg$;

(III) $Re(f) \sim Re(g)$, $Im(f) \sim Im(g)$;

(IV) se f *e* g *sono a valori reali e* f è *misurabile, anche* g è *misurabile;*

(V) se f è *sommbile, anche* g è *sommabile e si ha*

$$\int_A f(x)dx = \int_A g(x)dx.$$

Dimostrazione Verifichiamo solo (IV) lasciando il resto come esercizio al lettore (vedi l' esercizio 0.4.2). Sia $N := \{x \in A : f(x) \neq g(x)\}$. N è misurabile e ha misura 0. Allora g coincide con f in $A \setminus N$ ed è misurabile su N per teorema 0.3.1. Otteniamo allora la conclusione dal risultato dell'esercizio 0.2.5. $\square$

Dal teorema 0.4.1 segue che, se f e g sono funzioni definite su A a valori complessi, sono ben definite

$$[f] + [g] := [f + g] \tag{0.4.2}$$

e

$$\lambda[f] := [\lambda f], \tag{0.4.3}$$

nel senso che il risultato è in ciascun caso indipendente dalla scelta dei rappresentanti in ciascuna classe.

Esercizio 0.4.1 Provare che la relazione $\sim$ gode delle proprietà $(I) - (III)$.

Esercizio 0.4.2 Completare la dimostrazione del teorema 0.4.1.

0.5 Passaggio al limite sotto il segno di integrale

Vediamo ora senza dimostrazione alcuni fondamentali risultati di passaggio al limite sotto il segno di integrale. Siano $A \in \mathcal{M}_n$ e, per $k \in \mathbf{N}$, $f_k : A \to [0, +\infty]$ (risp. $f_k : A \to \mathbf{C}$), $f : A \to [0, +\infty]$ (risp. $f : A \to \mathbf{C}$) tali che per quasi ogni $x \in A$ vale $\lim\limits_{k \to +\infty} f_k(x) = f(x)$. Nell'ipotesi che le f_k ammettano tutte integrale nel senso della definizione 0.2.4 o della definizione 0.2.6, ci chiediamo quando sia possibile concludere che la f ammette integrale e vale $\lim\limits_{k \to +\infty} \int_A f_k(x)dx = \int_A f(x)dx$. Il problema è non banale, come dimostra il seguente

Esempio 0.5.1 Sia, per $k \in \mathbf{N}$,

$$\begin{cases} f_k : [0,1] \to \mathbf{R}, \\ f_k(x) = \begin{cases} k^2 x & \text{se} \quad x \in [0, \frac{1}{k}], \\ k - k^2(x - \frac{1}{k}) & \text{se} \quad x \in [\frac{1}{k}, \frac{2}{k}], \\ 0 & \text{se} \quad x \in [\frac{2}{k}, 1]. \end{cases} \end{cases}$$

È facile vedere che si ha $\lim_{k \to +\infty} f_k(x) = 0 \; \forall x \in [0,1]$. Ciò è ovvio se $x = 0$. Se $x > 0$, basta osservare che, se k è abbastanza grande, si ha $x > \frac{2}{k}$, per cui $f_k(x) = 0$.

Tuttavia, si ha $\int_{[0,1]} f_k(x)dx = 1 \; \forall k \in \mathbf{N}$.

Vediamo ora (senza dimostrazione) due teoremi classici di passaggio al limite. Premettiamo la seguente definizione di convergenza di una successione a valori complessi, che riprenderemo più avanti (si veda la successiva sezione 1.3): sia $(a_n)_{n \in \mathbf{N}}$ una successione in $\mathbf{C}$ e sia $a \in \mathbf{C}$. Scriveremo che $\lim_{n \to +\infty} a_n = a$ se $\forall \epsilon > 0$ esiste $n(\epsilon) \in \mathbf{N}$ tale che $\forall n \in \mathbf{N}$ con $n > n(\epsilon)$ si ha $|a_n - a| < \epsilon$.

Teorema 0.5.1 *(di Beppo Levi) Siano $A \in \mathcal{M}_n$ e, per $k \in \mathbf{N}$, $f_k : A \to [0, +\infty]$ misurabile. Valga inoltre $f_k(x) \leq f_{k+1}(x) \; \forall k \in \mathbf{N}, \forall x \in A$. Sia poi $f : A \to [0, +\infty]$, $f(x) = \lim_{k \to +\infty} f_k(x)$ (il limite esiste per la monotonia di $(f_k(x))_{k \in \mathbf{N}}$). Allora f è misurabile e a valori in $[0, +\infty]$, e vale*

$$\int_A f(x)dx = \lim_{k \to +\infty} \int_A f_k(x)dx.$$

Esempio 0.5.2 Sia, per $k \in \mathbf{N}$, $f_k : [1, +\infty[\to \mathbf{R}$, $f_k(x) = \frac{x^k e^{-x}}{1+x^k}$. f_k è misurabile perché continua, e non negativa. Se $x \geq 1$, si ha $x^k \leq x^{k+1}$, e, poiché $y \to \frac{y}{1+y}$ è non decrescente su $[0, +\infty[$, $\frac{x^k}{1+x^k} \leq \frac{x^{k+1}}{1+x^{k+1}}$. Dunque, $f_k(x) \leq f_{k+1}(x)$, per ogni $k \in \mathbf{N}$ e $x \in [1, +\infty[$. Si ha

$$\lim_{k \to +\infty} f_k(x) = \begin{cases} \frac{1}{2e} & \text{se} \quad x = 1, \\ e^{-x} & \text{se} \quad x > 1. \end{cases}$$

Poniamo allora

$$\begin{cases} f : [1, +\infty[\to \mathbf{R}, \\ f(x) = \begin{cases} \frac{1}{2e} & \text{se} \quad x = 1, \\ e^{-x} & \text{se} \quad x > 1. \end{cases} \end{cases}$$

Allora, per il teorema di Beppo Levi, si ha

$$\lim_{k\to+\infty}\int_{[1,+\infty[} f_k(x)dx = \int_{[1,+\infty[} f(x)dx$$

$$= \int_{\{1\}} \frac{1}{2e}dx + \int_{]1,+\infty[} e^{-x}dx$$

$$= \int_{]1,+\infty[} e^{-x}dx = \int_{\{1\}} e^{-x}dx + \int_{]1,+\infty[} e^{-x}dx$$

$$= \int_{[1,+\infty[} e^{-x}dx,$$

ove abbiamo usato il teorema 0.2.3(V) e il teorema 0.3.1, perché $L_1(\{1\}) = 0$. A questo punto, applicando il teorema 0.3.3, otteniamo

$$\int_{[0,+\infty[} e^{-x}dx = \lim_{c\to+\infty}\int_0^c e^{-x}dx$$

$$= \lim_{c\to+\infty}(1 - e^{-c}) = 1.$$

Passiamo ora a funzioni a valori complessi.

Teorema 0.5.2 *(di Lebesgue, della convergenza dominata) Siano $A \in \mathcal{M}_n$ e, per $k \in \mathbf{N}$, $f_k : A \to \mathbf{C}$ misurabile e $g : A \to [0,+\infty[$ sommabili. Valga inoltre $|f_k(x)| \le g(x)$ $\forall k \in \mathbf{N}$ e q. d. in A. Sia poi $f : A \to \mathbf{C}$, tale che $f(x) = \lim_{k\to+\infty} f_k(x)$ q. d. in A. Allora le f_k e f sono sommabili in A e vale*

$$\int_A f(x)dx = \lim_{k\to+\infty}\int_A f_k(x)dx.$$

Esempio 0.5.3 Sia, per $k \in \mathbf{N}$, $f_k : [0,+\infty[\to \mathbf{R}$, $f_k(x) = \sin(\frac{x}{n})e^{-x}$. f_k è misurabile perché continua. Poniamo $g : [0,+\infty[\to \mathbf{R}$, $g(x) = e^{-x}$. g è sommabile, non negativa e $|f_k(x)| \le g(x)$ $\forall k \in \mathbf{N}$, $\forall x \ge 0$. Sia ha infine $\lim_{k\to+\infty} f_k(x) = 0$ $\forall x \ge 0$. Segue allora dal teorema della convergenza dominata che per ogni k f_k è sommabile su $[0,+\infty[$ e vale

$$\lim_{k\to+\infty}\int_{[0,+\infty[} f_k(x)dx = \int_{[0,+\infty[} 0dx = 0.$$

Esercizio 0.5.1 Siano $A \in \mathcal{M}_n$, per $k \in \mathbf{N}$ $f_k : A \to [0, +\infty]$ misurabile, $s : A \to [0, +\infty]$, $s(x) = \sum_{k=1}^{\infty} f_k(x)$. Dimostrare che f è misurabile e che

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

Esercizio 0.5.2 Siano $A \in \mathcal{M}_n$ e per $k \in \mathbf{N}$ $f_k : A \to \mathbf{C}$ misurabile. Supponiamo inoltre che

(a) $L_n(A) < +\infty$;
(b) esiste $M \geq 0$ tale che $|f_k(x)| \leq M$ $\forall x \in A$;
(c) esiste $f : A \to \mathbf{C}$ tale che $f(x) = \lim\limits_{k \to +\infty} f_k(x)$ q. d. in A.

Dimostrare che le f_k e f sono sommabile e che

$$\int_A f(x)dx = \lim_{k \to +\infty} \int_A f_k(x)dx.$$

Capitolo 1

Spazi normati

1.1 Norme

Presupponiamo che il lettore conosca gli elementi di base dell'algebra lineare. Per completezza e per fissare le notazioni, cominciamo col richiamare alcuni di questi elementi. Vediamo, innazi tutto, la nozione di spazio vettoriale sul campo $K = \mathbf{R}$ o $\mathbf{C}$.

Definizione 1.1.1 *Siano X un insieme non vuoto, $+$ un'operazione su X (vale a dire, una funzione da $X \times X$ a X), $(\lambda, x) \to \lambda x$ un'applicazione da $K \times X$ a X, che chiameremo moltiplicazione per uno scalare. Diremo che X, con la somma $+$ e questo prodotto scalare,* è *uno spazio vettoriale su K se valgono le seguenti propriet*à *:*

(ASV1) $\forall x, y \in X$

$$y + x = x + y$$

(proprietà commutativa della somma);

(ASV2) $\forall x, y, z \in X$

$$x + (y + z) = (x + y) + z$$

(proprietà associativa della somma);

(ASV3) esiste un elemento $O \in X$ tale che, $\forall x \in X$

$$x + O = O + x = x$$

(esistenza di un elemento neutro per la somma);

(ASV4) $\forall x \in X$ esiste $-x \in X$ tale che

$$x + (-x) = (-x) + x = 0;$$

(esistenza di un inverso additivo per ogni $x \in X$);

(ASV5) $\forall \lambda,\ \mu \in K$, $\forall x \in X$

$$(\lambda + \mu)x = \lambda x + \mu x;$$

(ASV6) $\forall \lambda \in K$, $\forall x, y \in X$

$$\lambda(x + y) = \lambda x + \lambda y;$$

(ASV7) $\forall \lambda,\ \mu \in K$, $\forall x \in X$

$$\lambda(\mu x) = (\lambda\mu)x;$$

(ASV8) $\forall x \in X$

$$1x = x.$$

Osservazione 1.1.1 Come si fa solitamente, abbiamo indicato con lo stesso simbolo + la somma in X e la somma in K. Per esempio, in $(ASV5)$ il primo + indica la somma in K, il secondo la somma in X.

Esempio 1.1.1 Dati $K = \mathbf{R}$ o $\mathbf{C}$ e $n \in \mathbf{N}$, indichiamo con K^n l'insieme delle $n-$ple ordinate di elementi di K. Se $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$ sono elementi di K^n, poniamo

$$x + y := (x_1 + y_1, ..., x_n + y_n). \tag{1.1.1}$$

Se $\lambda \in K$, poniamo poi

$$\lambda x := (\lambda x_1, ..., \lambda x_n). \tag{1.1.2}$$

È ben noto che con la somma (1.1.1) e il prodotto per uno scalare (1.1.2) K^n è uno spazio vettoriale su K.

Esempio 1.1.2 Sia A un insieme non vuoto arbitrario. Indichiamo con $\mathcal{F}(A, K)$ l'insieme delle funzioni di dominio A a valori in K. Dati f e g in $\mathcal{F}(A, K)$ e $\lambda \in K$, poniamo

$$\begin{cases} f+g : A \to K, \\ (f+g)(a) = f(a) + g(a), \forall a \in A, \end{cases} \tag{1.1.3}$$

$$\begin{cases} \lambda f : A \to K, \\ (\lambda f)(a) = \lambda \cdot f(a), \forall a \in A, \end{cases} \tag{1.1.4}$$

con $\cdot$ prodotto in K. È facile vedere che $\mathcal{F}(A, K)$ con la somma (1.1.3) e il prodotto per uno scalare (1.1.4) è uno spazio vettoriale su K.

Osservazione 1.1.2 Dato uno spazio vettoriale X sul campo K e dato $Y \subseteq X$, si dice che Y è un **sottospazio** di X se, dati y_1, y_2 e y elementi di Y e $\lambda \in K$, si ha sempre $y_1 + y_2 \in Y$ e $\lambda y \in Y$. Non è difficile verificare che, se Y è un sottospazio di X, è a sua volta uno spazio vettoriale su K con le restrizioni della somma a $Y \times Y$ e del prodotto per uno scalare a $K \times Y$.

Se $A \subseteq \mathbf{R}^n$, lo spazio vettoriale $\mathcal{F}(A, \mathbf{R})$ possiede parecchi sottospazi interessanti. Ne indichiamo qualcuno.

$B(A, \mathbf{R})$ (dall'inglese "bounded") è lo spazio delle funzioni da A a $\mathbf{R}$ limitate.

$C(A, \mathbf{R})$ è lo spazio delle funzioni da A a $\mathbf{R}$ continue.

$BC(A, \mathbf{R})$ è lo spazio delle funzioni da A a $\mathbf{R}$ continue e limitate.

Se A è misurabile secondo Lebesgue, $\mathcal{M}(A, \mathbf{R})$ è lo spazio delle funzioni da A a $\mathbf{R}$ misurabili secondo Lebesgue.

Se A è misurabile secondo Lebesgue, $\mathcal{L}^1(A, \mathbf{R})$ è lo spazio delle funzioni da A a $\mathbf{R}$ sommabili.

Indichiamo invece con $\mathcal{L}^1(A)$ è lo spazio delle funzioni su A, in generale a valori complessi, sommabili. Otteniamo in questo caso uno spazio vettoriale su $\mathbf{C}$.

Definizione 1.1.2 *Sia X uno spazio vettoriale su K (= $\mathbf{R}$ o $\mathbf{C}$). Una* **norma** *in X é una funzione $\|.\| : X \to [0, +\infty[$, $x \to \|x\|$ $\forall x \in X$, tale che*

(I) $\forall \lambda \in K$, $\forall x \in X$

$$\|\lambda x\| = |\lambda| \|x\|;$$

(II) $\forall x, y \in X$

$$\|x + y\| \leq \|x\| + \|y\|;$$

(III) se $x \in X$ e $\|x\| = 0$, allora $x = 0$.

Esempio 1.1.3 Sia, dato $x = (x_1, ..., x_n) \in \mathbf{R}^n$,

$$\|x\| := (\sum_{j=1}^{n} x_j^2)^{\frac{1}{2}}.$$

$\|.\|$ è la norma euclidea in $\mathbf{R}^n$ e sappiamo dal corso di Analisi B che soddisfa le condizioni $(I) - (III)$ della definizione 1.1.2.

Non si tratta dell'unica norma in $\mathbf{R}^n$. Due norme diverse dalla norma euclidea sono, per $x = (x_1, ..., x_n) \in \mathbf{R}^n$,

$$\|x\|_1 := \sum_{j=1}^{n} |x_j|, \tag{1.1.5}$$

e

$$\|x\|_\infty := \max_{1 \leq j \leq n} |x_j|. \tag{1.1.6}$$

Per la verifica vedi il seguente esercizio 1.1.1.

Esempio 1.1.4 In analogia con (1.1.6), dato $A \subseteq \mathbf{R}^n$ (anche se in realtà si può prendere come A un generico insieme non vuoto), e posto $X := \mathcal{B}(A, \mathbf{R})$, definiamo in X la norma

$$\|x\|_\infty := \sup_A |f| = \sup\{|f(a)| : a \in A\}. \tag{1.1.7}$$

Verichiamo che $\|.\|_\infty$ è una norma in X.

Innanzi tutto è chiaro che $\|.\|_\infty$ è ben definita su X e a valori in $[0, +\infty[$. Passiamo in rassegna le proprietà $(I) - (III)$ della definizione 1.1.2.

(I) Se $a \in A$, $|\lambda f(a)| = |\lambda||f(a)| \leq |\lambda| \sup_A\{|f(a)| : a \in A\} = |\lambda|\|f\|_\infty$. Poiché ciò vale qualunque sia a, ricaviamo

$$\|\lambda f\|_\infty = \sup\{|\lambda f(a)| : a \in A\} \leq |\lambda|\|f\|_\infty. \tag{1.1.8}$$

Se $\lambda \neq 0$, osservando che $f = \lambda^{-1}(\lambda f)$ e applicando (1.1.8), otteniamo

$$\|f\|_\infty \leq |\lambda^{-1}|\|\lambda f\|_\infty,$$

da cui ricaviamo immediatamente (I). Il caso $\lambda = 0$ è immediato.

(II) Siano f e g elementi di X e $a \in A$. Allora

$$|f(a) + g(a)| \leq |f(a)| + |g(a)| \leq \|f\|_\infty + \|g\|_\infty.$$

Segue che

$$\|f+g\|_\infty = \sup\{|f(a)+g(a)| : a \in A\} \le \|f\|_\infty + \|g\|_\infty.$$

(III) Sia $f \in X$ tale che $\|f\|_\infty = 0$. Allora $\sup\{|f(a)| : a \in A\} = 0$. Ciò implica, evidentemente, che $f(a) = 0 \ \forall a \in A$. La funzione identicamente nulla è proprio lo "zero" dello spazio verttoriale X. Segue che vale anche (III).

Osservazione 1.1.3 È facile verificare che su ogni sottospazio di $\mathcal{B}(A, \mathbf{R})$ (come $BC(A, \mathbf{R})$) (1.1.7) definisce una norma.

Esempio 1.1.5 Siano $A \in \mathcal{M}_n$ (la classe dei sottoinsiemi misurabili secondo Lebesgue in $\mathbf{R}^n$). Poniamo, data $f \in \mathcal{L}^1(A)$,

$$\|f\|_1 := \int_A |f(x)|dx. \tag{1.1.9}$$

Se $\|.\|_1$ fosse una norma, dovrebbe valere il seguente fatto:

se $f \in \mathcal{L}^1(A)$ e $\int_A |f(x)|dx = 0$ allora f è identicamente nulla.

Ma si vede facilmente che ciò è falso. Per esempio, se $A = \mathbf{R}^n$, $B \in \mathcal{M}_n$, $B \ne \emptyset$ e $L_n(B) = 0$, poniamo

$$\begin{cases} f : A \to \mathbf{R}, \\ f(x) = \begin{cases} 1 & \text{se} \quad x \in B, \\ 0 & \text{se} \quad x \notin B. \end{cases} \end{cases}$$

f è una funzione semplice non negativa e non identicamente nulla e vale

$$\|f\|_1 = \int_A f(x)dx = 1 \cdot L_n(B) + 0 \cdot L_n(\mathbf{R}^n \setminus B) = 0.$$

Di fatto, vale il seguente

Lemma 1.1.1 *Siano $A \in \mathcal{M}_n$, $f : A \to [0, +\infty]$ misurabile. Le due condizioni seguenti sono equivalenti:*

(I) $\int_A f(x)dx = 0$;

(II) $f(x) = 0$ q. d. (quasi dappertutto) in A, vale a dire $\{x \in A : f(x) > 0\}$ è misurabile e ha misura 0.

Perciò , se $\|f\|_1 = 0$, possiamo solo dire che $|f(x)| = 0$ q. d. in A, e quindi che $f(x) = 0$ q. d. in A. Per aggirare questo inconveniente, invece di $\mathcal{L}^1(A)$, conviene considerare lo spazio vettoriale

$$L^1(A) := \{[f] : f \in \mathcal{L}^1(A)\}, \tag{1.1.10}$$

ove $[f]$ è la classe di equivalenza di f definita in (0.4.1). $L^1(A)$ è uno spazio vettoriale su $\mathbf{C}$ con le operazioni di somma e prodotto per uno scalare definite in (0.4.2) e (0.4.3). Data $[f] \in L^1(A)$, poniamo

$$\|[f]\|_1 := \int_A |f(x)|dx. \tag{1.1.11}$$

La definizione è ben posta per il teorema 0.4.1(V).

Facciamo vedere che $\|.\|_1$ è una norma in $L^1(A, \mathbf{R})$.

Innanzi tutto, se $\lambda \in \mathbf{C}$ e $f \in \mathcal{L}^1(A)$,

$$\begin{gathered}\|\lambda[f]\|_1 = \|[\lambda f]\|_1 \\ = \int_A |\lambda f(x)|dx = |\lambda| \int_A |f(x)|dx \\ = |\lambda|\|[f]\|_1.\end{gathered}$$

Se f e g sono elementi di $\mathcal{L}^1(A)$,

$$\begin{gathered}\|[f] + [g]\|_1 = \|[f + g]\|_1 \\ = \int_A |f(x) + g(x)|dx \le \int_A |f(x)|dx + \int_A |g(x)|dx \\ = \|[f]\|_1 + \|[g]\|_1.\end{gathered}$$

Infine, sia $\|[f]\|_1 = 0$. Allora $\int_A |f(x)|dx = 0$ e quindi $f(x) = 0$ q. d. . Ciò significa che f è equivalente alla funzione identicamente nulla e quindi $[f] = [0]$, che è lo zero di $L^1(A)$.

Osservazione 1.1.4 Anche se gli elementi di $L^1(A)$ ($A \in \mathcal{M}_n$) non sono funzioni, ma classi di funzioni coincidenti quasi dappertutto, è comodo e consueto parlare di una certa funzione f come elemento di $L^1(A)$. Si intenderà sempre, in realtà , la classe $[f]$.

Esercizio 1.1.1 Verificare che $\|.\|_1$ e $\|.\|_\infty$ definite in (1.1.5) e (1.1.6) sono norme in $\mathbf{R}^n$.

Esercizio 1.1.2 Verificare che, se $\|.\|$ è una norma in X, $\forall x, y \in X$ si ha

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Esercizio 1.1.3 Sia $f : [0, 2\pi] \to \mathbf{R}$, $f(x) = x - \sin(x) - 4\pi$. Calcolare $\|f\|_\infty$. Calcolare anche $\|f\|_1$

Esercizio 1.1.4 Sia $f : [-2, 1] \to \mathbf{R}$, $f(x) = x^3$. Calcolare $\|f\|_\infty$ e $\|[f]\|_1$.

1.2 Nozioni di carattere topologico in uno spazio normato

Introduciamo ora in un generico spazio normato una serie di nozioni di carattere topologico già viste in $\mathbf{R}^n$. In questa sezione X sarà uno spazio normato qualunque su $K = \mathbf{R}$ o $\mathbf{C}$. Indicheremo con $\|.\|$ la norma in X.

Siano $x \in X$ e $r > 0$. Poniamo

$$B(x, r) := \{y \in X : \|y - x\| < r\}. \tag{1.2.1}$$

$B(x, r)$ si chiama **palla aperta** di centro x e raggio r.

Siano $A \subseteq X$ e $x_0 \in X$. Diremo che x_0 appartiene all'**interno** di A e scriveremo

$$x_0 \in \mathring{A} \tag{1.2.2}$$

se esiste $r > 0$ tale che $B(x_0, r) \subseteq A$.

Dato poi $A \subseteq X$, diremo che A è **aperto** se $\mathring{A} = A$.

Dati $A \subseteq X$ e $x_0 \in X$, diremo che x_0 appartiene alla **frontiera** di A e scriveremo

$$x_0 \in Fr(A) \tag{1.2.3}$$

se ogni palla aperta $B(x_0, r)$, con $r > 0$, contiene sia elementi appartenenti ad A, sia elementi non appartenenti ad A.

Diremo che $A \subseteq X$ è **chiuso** se $Fr(A) \subseteq A$.

Dato poi $A \subseteq X$, chiameremo **chiusura** di A e indicheremo con la scrittura $\overline{A}$ l'insieme

$$\overline{A} := A \cup Fr(A). \tag{1.2.4}$$

Se $A \subseteq X$ e $\overline{A} = X$, si dice che A è **denso** in X.

Introduciamo infine la nozione di continuità .

Definizione 1.2.1 *Siano X e Y spazi normati con norme rispettivamente $\|.\|_X$ e $\|.\|_Y$, $A \subseteq X$, $f : A \to Y$, $x_0 \in A$. Diremo che f* è **continua** *in x_0 se $\forall \epsilon \in \mathbf{R}^+$ esiste $r(\epsilon) \in \mathbf{R}^+$ tale che, $\forall x \in A$ verificante $\|x - x_0\|_X < r(\epsilon)$ vale*

$$\|f(x) - f(x_0)\|_Y < \epsilon.$$

Diremo che f è *continua se* è *continua in ogni punto di A.*

È opportuno introdurre anche le nozioni di punto di accumulazione e di limite.

Definizione 1.2.2 *Siano $A \subseteq X$, con X spazio normato, e $x^0 \in X$. Diremo che x^0* è *un punto di accumulazione per A (e scriveremo $x^0 \in D(A)$) se $\forall r > 0$ $B(x^0, r)$ contiene qualche elemento di A distinto da x^0.*

Definizione 1.2.3 *Siano X e Y spazi normati, con norme, rispettivamente, $\|.\|_X$ e $\|.\|_Y$. Siano poi $A \subseteq X$, $x^0 \in D(A)$ e $l \in Y$. Scriveremo che $\lim_{x \to x^0} f(x) = l$ se $\forall \epsilon \in \mathbf{R}^+$ esiste $r(\epsilon) \in \mathbf{R}^+$ tale che, $\forall x \in A$ verificante $\|x - x_0\|_X < r(\epsilon)$, con $x \neq x^0$, vale*

$$\|f(x) - f(x_0)\|_Y < \epsilon.$$

Esempio 1.2.1 Consideriamo lo spazio vettoriale $X := C([0,1], \mathbf{R})$ normato con la norma $\|.\|_\infty$ (vedi (1.1.7)). Osserviamo che per il teorema di Weierstrass $C([0,1], \mathbf{R})$ coincide con $BC([0,1], \mathbf{R})$, e quindi questa norma è ben definita in X. Poniamo

$$\begin{cases} f : X \to X, \\ f(x)(t) := \int_0^t x(s)ds, \qquad x \in X. \end{cases} \tag{1.2.5}$$

Il teorema fondamentale del calcolo integrale assicura che $f(x)$ è continua e appartiene, quindi, a X. Verifichiamo che f è continua.

Siano $x_0 \in X$ e $\epsilon > 0$; dato $x \in X$, si ha, $\forall t \in [0,1]$,

$$|F(x)(t) - F(x_0)(t)| = |\int_0^t (x(s) - x_0(s))ds|$$

$$\leq \int_0^t |x(s) - x_0(s)|ds \leq t\|x - x_0\|_\infty$$

$$\leq \|x - x_0\|_\infty.$$

Ne segue che

$$\|F(x) - F(x_0)\|_\infty \leq \|x - x_0\|_\infty.$$

Posto dunque $r(\epsilon) := \epsilon$, si ha che, se $\|x - x_0\|_\infty < r(\epsilon)$, vale $\|F(x) - F(x_0)\|_\infty < \epsilon$.

Esercizio 1.2.1 Dimostrare che ogni palla aperta è un insieme aperto.

Esercizio 1.2.2 Dimostrare che, se $A = B(x_0, r)$, con $x_0 \in A$ e $r > 0$,

$$Fr(A) = \{x \in X : \|x - x_0\| = r\}.$$

Esercizio 1.2.3 Sia $A \subseteq X$. Dimostrare che A è aperto se e solo se

$$A \cap Fr(A) = \emptyset.$$

Esercizio 1.2.4 Sia $A \subseteq X$. Dimostrare che A è chiuso se e solo se

$$A = \overline{A}.$$

Esercizio 1.2.5 Siano $A \subseteq X$, $f : A \to \mathbf{C}$. Dimostrare che f è continua se e solo se $Re(f)$ e $Im(f)$ sono continue a valori reali (sugg.: per ogni $\lambda \in \mathbf{C}$ $\max\{|Re(\lambda)|, |Im(\lambda)|\} \leq |\lambda| \leq |Re(\lambda)| + |Im(\lambda)|$).

Esercizio 1.2.6 Verificare che, se X è normato con norma $\|.\|$, la norma è continua da X a $\mathbf{R}$. (Sugg.: utilizzare il risultato dell'esercizio 1.1.2).

Esercizio 1.2.7 Enunciare e dimostrare un risultato di unicità del limite.

Esercizio 1.2.8 Sia $f : A \subseteq X \to Y$, con X e Y spazi normati. Sia poi $x^0 \in A \cap D(A)$. Dimostrare che f è continua in x^0 se e solo se esiste $\lim_{x \to x^0} f(x)$ e coincide con $f(x^0)$.

1.3 Successioni in uno spazio normato e completezza

Cominciamo con la definizione di limite di una successione in uno spazio normato X (di norma $\|.\|$).

Definizione 1.3.1 *Siano $(x_n)_{n\in\mathbf{N}}$ una successione nello spazio normato X ed $l \in X$. Diremo che $(x_n)_{n\in\mathbf{N}}$* **converge** *a l o ha per* **limite** *l se $\forall \epsilon \in \mathbf{R}^+$ esiste $n(\epsilon) \in \mathbf{N}$ tale che per ogni $n \in \mathbf{N}$ con $n > n(\epsilon)$ si ha*

$$\|x_n - l\| < \epsilon.$$

Osservazione 1.3.1 La definizione 1.3.1 equivale a richiedere che

$$\lim_{n\to+\infty} \|x_n - l\| = 0.$$

Osservazione 1.3.2 Si vede facilmente con lo stesso argomento delle successioni a valori in $\mathbf{R}$ che, se il limite esiste, è unico (vedi esercizio 1.3.1).

Osservazione 1.3.3 La convergenza rispetto alla norma $\|.\|_\infty$ nell'ambito di $\mathcal{B}(A, \mathbf{R})$ (con A insieme non vuoto) si chiama anche **convergenza uniforme**.

Esempio 1.3.1 Siano $\delta \in]0, 1[$, $X = C([0, \delta], \mathbf{R})$ con la norma $\|.\|_\infty$. Poniamo, dati $n \in \mathbf{N}$, $t \in [0, \delta]$,

$$x_n(t) := t^n, \tag{1.3.1}$$

$$l(t) = 0. \tag{1.3.2}$$

Verifichiamo che la successione $(x_n)_{n\in\mathbf{N}}$ tende a l. Infatti

$$\|x_n - l\|_\infty = \sup_{t\in[0,\delta]} |x_n(t) - l(t)|$$

$$= \delta^n \to 0(n \to +\infty).$$

È evidente che la convergenza uniforme implica la convergenza puntuale, nel senso che, se $(x_n)_{n\in\mathbf{N}}$ è una successione in $\mathcal{B}(A, \mathbf{R})$ convergente uniformemente a l ($\in \mathcal{B}(A, \mathbf{R})$), allora per ogni $t \in A$ si ha

$$\lim_{n\to\infty} x_n(t) = l(t).$$

Non vale il viceversa. Consideriamo infatti lo spazio normato $\mathcal{B}([0, 1], \mathbf{R})$ con la norma $\|.\|_\infty$. Sia, per $n \in \mathbf{N}$, $x_n : [0, 1] \to \mathbf{R}$, definita ancora come in (1.3.1) e sia

$$\begin{cases} m : [0, 1] \to \mathbf{R}, \\ m(t) = \begin{cases} 0 & \text{se} \quad t \in [0, 1[, \\ 1 & \text{se} \quad t = 1. \end{cases} \end{cases}$$

Si vede subito che vale $\lim_{n\to+\infty} x_n(t) = m(t)\ \forall t \in [0,1]$. Tuttavia non si ha $\lim_{n\to+\infty} \|x_n - m\|_\infty = 0$. Infatti, per ogni $n \in \mathbf{N}$,

$$|x_n(t) - m(t)| = \begin{cases} t^n & \text{se} \quad t \in [0,1[, \\ 0 & \text{se} \quad t = 1. \end{cases}$$

Dunque $\|x_n - m\|_\infty = 1\ \forall n \in \mathbf{N}$.

Una prima importante conseguenza della convergenza di una successione è la sua limitatezza:

Teorema 1.3.1 *Sia* $(a_n)_{n\in\mathbf{N}}$ *convergente nello spazio normato* X. *Allora* $(a_n)_{n\in\mathbf{N}}$ è *limitata, nel senso che esiste* $M \geq 0$ *tale che*

$$\|a_n\| \leq M \qquad \forall n \in \mathbf{N}.$$

Dimostrazione Sia $\lim_{n\to\infty} a_n = l \in X$. Esiste $n(1) \in \mathbf{N}$ tale che, se $n > n(1)$, si ha $\|a_n - l\| < 1$. Di conseguenza, se $n > n(1)$,

$$\|a_n\| = \|(a_n - l) + l\| \leq \|l\| + \|a_n - l\| < \|l\| + 1.$$

Ponendo allora

$$M := \max\{\|a_1\|, ..., \|a_n\|, \|l\| + 1\},$$

si ha $\|a_n\| \leq M\ \forall n \in \mathbf{N}$. □

Vediamo ora certi importanti legami tra la nozione di limite di una successione e alcune delle nozioni introdotte nella sezione 1.2.

Teorema 1.3.2 *Siano* $A \subseteq X$ *spazio normato e* $x_0 \in X$. *Allora sono equivalenti:*

(I) $x_0 \in \overline{A}$*;*

(II) esiste una successione $(a_n)_{n\in\mathbf{N}}$ *a valori in* A *convergente a* x_0.

Dimostrazione Verifichiamo che da (I) segue (II). Sia $x_0 \in \overline{A}$. Costruiamo una successione a valori in A convergente a x_0. Se $x_0 \in \overline{A}$, allora $x_0 \in A$ oppure $x_0 \in Fr(A)$. Nel primo caso, possiamo prendere $a_n = x_0$ $\forall n \in \mathbf{N}$. Sia invece $x_0 \in Fr(A)$. Per definizione, per ogni $n \in \mathbf{N}$ esiste $a_n \in A \cap B(x_0, \frac{1}{n})$. Si ha

$$\|a_n - x_0\| < \frac{1}{n} \to 0 (n \to +\infty).$$

Dunque, da (I) segue (II).

Supponiamo invece che valga (II). Dobbiamo far vedere che $x_0 \in \overline{A}$. Consideriamo separatamente i due casi $x_0 \in A$ e $x_0 \notin A$. Nel primo caso, $x_0 \in \overline{A}$. Nel secondo caso, sia $r > 0$. Poiché $x_0 \notin A$, $B(x_0, r)$ contiene almeno un elemento non appartenente ad A: x_0 stesso. Scegliendo poi $n \in \mathbf{N}$ in modo che $\|a_n - x_0\| < r$, si ottiene $a_n \in A \cap B(x^0, r)$. Dunque $x_0 \in Fr(A)$. $\square$

Teorema 1.3.3 *Siano X e Y spazi normati, con norma rispettivamente $\|.\|_X$ e $\|.\|_Y$, $A \subseteq X$, $f : A \to Y$, $x_0 \in A$. Allora sono equivalenti:*

(I) f è continua in x_0;

(II) comunque si prenda una successione $(a_n)_{n\in\mathbf{N}}$ a valori in A e convergente in X a x_0, si ha che la successione $(f(a_n))_{n\in\mathbf{N}}$ converge in Y a $f(x_0)$.

Dimostrazione Proviamo che (I) implica (II). Sia $(a_n)_{n\in\mathbf{N}}$ a valori in A e convergente in X a x_0. Sia poi $\epsilon \in \mathbf{R}^+$. Poiché f è continua in x_0, esiste $r(\epsilon) > 0$ tale che, se $a \in A$ e $\|a - x_0\|_X < r(\epsilon)$, allora $\|f(a) - f(x_0)\|_Y < \epsilon$. Dalla definizione di convergenza di una successione, si ha che esiste $n(r(\epsilon)) \in \mathbf{N}$ tale che, se $n > n(r(\epsilon))$, si ha $\|a_n - x_0\|_X < r(\epsilon)$. Segue $\|f(a_n) - f(x_0)\|_Y < \epsilon$. Quindi vale (II).

Supponiamo, viceversa, che valga (II). Dobbiamo far vedere che f è continua in x_0. Ragioniamo per assurdo e supponiamo che non sia cosi̇ . Allora esiste $\epsilon > 0$ tale che, comunque si prenda $r > 0$ esiste $a \in A$ con $\|a - x_0\|_X < r$ e $\|f(a) - f(x_0)\|_Y \geq \epsilon$. Prendiamo $r = 1, \frac{1}{2}, ..., \frac{1}{n}, ...$ e scegliamo $a_n \in A \cap B(x_0, \frac{1}{n})$ con $\|f(a_n) - f(x_0)\|_Y \geq \epsilon$. Allora, evidentemente, $(a_n)_{n\in\mathbf{N}}$ converge a x_0, ma $(f(a_n))_{n\in\mathbf{N}}$ non converge a $f(x_0)$. Dunque non vale (II) e questa è una contraddizione. $\square$

Vediamo adesso la nozione di completezza di uno spazio normato. Cominciamo con le seguente

Definizione 1.3.2 *Siano X uno spazio normato e $(x_n)_{n\in\mathbf{N}}$ una successione in X. Diremo che $(x_n)_{n\in\mathbf{N}}$ è una* **successione di Cauchy** *se $\forall \epsilon \in \mathbf{R}^+$ esiste $n(\epsilon) \in \mathbf{N}$ tale che, per ogni possibile scelta di m e n naturali con $m > n(\epsilon)$ e $n > n(\epsilon)$, si ha*

$$\|a_n - a_m\| < \epsilon.$$

Osservazione 1.3.4 Ogni successione convergente è di Cauchy. Siano infatti $(x_n)_{n\in\mathbf{N}}$ una successione in X e $x_0 \in X$ tali che $\|x_n - x_0\| \to 0$

$(n \to +\infty)$. Siano poi $\epsilon > 0$ ed $n(\frac{\epsilon}{2}) \in \mathbf{N}$ tali che $\|x_n - x_0\| < \frac{\epsilon}{2}$ se $n > n(\frac{\epsilon}{2})$. Presi n e m naturali, entrambi maggiori di $n(\frac{\epsilon}{2})$, si ha allora

$$\begin{aligned}\|x_n - x_m\| &= \|(x_n - x_0) + (x_0 - x_m)\| \\ &\leq \|x_n - x_0\| + \|x_0 - x_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon.\end{aligned}$$

In generale, non è vero viceversa: in uno spazio normato possono esistere successioni di Cauchy che non sono convergenti.

Definizione 1.3.3 *Sia X uno spazio normato con norma $\|.\|$. Diremo che X con questa norma* è **completo** *se ogni successione di Cauchy* è *convergente. Gli spazi normati completi si chiamano anche* **spazi di Banach**.

Vediamo ora alcuni esempio di spazi di Banach. Il primo e fondamentale risultato è il seguente

Teorema 1.3.4 $\mathbf{R}$ *normato con il valore assoluto* è *completo.*

Il risultato si estende facilmente a $\mathbf{R}^n$:

Teorema 1.3.5 $\mathbf{R}^n$ *normato con la norma euclidea* è *completo.*

Dimostrazione Sia $(x^k)_{k\in\mathbf{N}}$ una successione di Cauchy in $\mathbf{R}^n$ rispetto alla norma euclidea, che indichiamo con $\|.\|$. Dobbiamo far vedere che la successione è convergente.

Sia, dato $k \in \mathbf{N}$, $x^k = (x_1^k, ..., x_n^k)$. Fissiamo $\epsilon > 0$. Allora esiste $k(\epsilon) \in \mathbf{N}$ tale che, se i e j sono numeri naturali maggiori di $k(\epsilon)$, si ha $\|x^i - x^j| < \epsilon$. Sia r un numero naturale minore o uguale a n. Consideriamo la successione in $\mathbf{R}$ $(x_r^k)_{k\in\mathbf{N}}$. Se i e j sono numeri naturali maggiori di $k(\epsilon)$, si ha

$$|x_r^i - x_r^j\| \leq \|x^i - x^j\| < \epsilon.$$

Dunque, per $r = 1, ..., n$, $(x_r^k)_{k\in\mathbf{N}}$ è una successione di Cauchy in $\mathbf{R}$. Poiché $\mathbf{R}$ è completo, ciascuna successione $(x_r^k)_{k\in\mathbf{N}}$ ammette un limite reale x_r^0. Poniamo

$$x^0 := (x_1^0, ..., x_n^0).$$

Allora la successione $(x^k)_{k\in\mathbf{N}}$ converge a x^0. Infatti

$$\|x^k - x^0\| = [\sum_{r=1}^n (x_r^k - x_r^0)^2]^{\frac{1}{2}} \to 0(k \to +\infty).$$

□

Consideriamo ora la norma $\|.\|_\infty$.

Teorema 1.3.6 *Sia $A \subseteq \mathbf{R}^n$ non vuoto. Lo spazio normato $B(A, \mathbf{R})$ con la norma $\|.\|_\infty$ è completo.*

Lo spazio normato $BC(A, \mathbf{R})$ con la norma $\|.\|_\infty$ è completo.

Dimostrazione Sia $(x^k)_{k \in \mathbf{N}}$ una successione di Cauchy in $B(A, \mathbf{R})$ rispetto alla norma $\|.\|_\infty$. Sia poi $\epsilon > 0$. Allora esiste $k(\epsilon) \in \mathbf{N}$ tale che, se i e j sono numeri naturali maggiori di $k(\epsilon)$, si ha $\|x^i - x^j\|_\infty < \epsilon$. Sia a un elemento qualunque di A. Consideriamo la successione in $\mathbf{R}$ $(x^k(a))_{k \in \mathbf{N}}$. Se i e j sono numeri naturali maggiori di $k(\epsilon)$, si ha

$$|x^i(a) - x^j(a)| \leq \|x^i - x^j\|_\infty < \epsilon.$$

Dunque, $\forall a \in A$, $(x^k(a))_{k \in \mathbf{N}}$ è una successione di Cauchy in $\mathbf{R}$. Poiché $\mathbf{R}$ è completo, ciascuna successione $(x^k(a))_{k \in \mathbf{N}}$ ammette un limite reale $x^0(a)$. Evidentemente, x^0 è una funzione da A a $\mathbf{R}$. Verifichiamo che $x^0 \in B(A, \mathbf{R})$ e che

$$\|x^k - x^0\|_\infty \to 0 (k \to +\infty).$$

Siano $\epsilon > 0$ e $k(\epsilon)$ come prima. Sia $j > k(\epsilon)$. Siano poi $i > k(\epsilon)$ e $a \in A$. Si ha

$$|x^i(a) - x^j(a)| \leq \|x^i - x^j\|_\infty < \epsilon. \tag{1.3.3}$$

Al limite per $i \to +\infty$ in (1.3.3), si ottiene $|x^0(a) - x^j(a)| \leq \epsilon \ \forall a \in A$. In particolare, dal caso $\epsilon = 1$ otteniamo $\forall a \in A$ e per un arbitrario $j > k(1)$,

$$|x^0(a)| = |(x^0(a) - x^j(a)) + x^j(a)|$$

$$\leq |x^0(a) - x^j(a)| + |x^j(a)| \leq 1 + \|x^j\|_\infty.$$

Perciò $x^0 \in B(A, \mathbf{R})$ e $\|x^0\|_\infty \leq 1 + \|x^j\|_\infty$. Inoltre, se $j > k(\frac{\epsilon}{2})$, abbiamo

$$\|x^0 - x^j\|_\infty \leq \frac{\epsilon}{2} < \epsilon.$$

Da ciò la prima conclusione.

Verifichiamo ora la completezza di $BC(A, \mathbf{R})$. Sia $(x^k)_{k \in \mathbf{N}}$ una successione di Cauchy in $BC(A, \mathbf{R})$. Evidentemente, è di Cauchy anche in $B(A, \mathbf{R})$ e quindi, per la prima parte del teorema, esiste $x^0 \in B(A, \mathbf{R})$ tale che $\|x^k - x^0\|_\infty \to 0$ $(k \to \infty)$. Per concludere, si tratta solo di verificare che $x^0 \in BC(A, \mathbf{R})$. Poiché sappiamo già che si tratta di una funzione limitata, resta solo da far vedere che è una funzione continua.

Sia $a^0 \in A$. Verifichiamo che x^0 è continua in A. Sia $\epsilon > 0$. Sia $k \in \mathbf{N}$ tale che $\|x^k - x^0\|_\infty < \frac{\epsilon}{3}$. Poiché x^k è continua, esiste $r > 0$ tale che, se $a \in A$ e $\|a - a^0\| < r$, allora $|x^k(a) - x^k(a^0)| < \frac{\epsilon}{3}$. Con questa scelta di a, si ha allora

$$\begin{aligned}
&|x^0(a) - x^0(a^0)| \\
&= |(x^0(a) - x^k(a)) + (x^k(a) - x^k(a^0)) + (x^k(a^0) - x^0(a^0))| \\
&\leq |x^0(a) - x^k(a)| + |x^k(a) - x^k(a^0)| + |x^k(a^0) - x^0(a^0)| \\
&\leq 2\|x^k - x^0\|_\infty + |x^k(a) - x^k(a^0)| < \epsilon.
\end{aligned}$$

Perciò x^0 è continua in a^0. □

Si potrebbe infine provare anche il seguente

Teorema 1.3.7 *Sia $A \in \mathcal{M}_n$. Allora $L^1(A)$, normato con $\|.\|_1$,* è *completo.*

Esempio 1.3.2 Dal teorema 1.3.6 segue che $C([0,1], \mathbf{R})$ è uno spazio di Banach con la norma $\|.\|_\infty$. Sia X lo spazio vettoriale delle funzioni polinomiali a coefficienti reali ristrette a $[0,1]$. Un celebre teorema dovuto a Weierstrass asserisce che X è denso in $C([0,1], \mathbf{R})$. In altre parole, in virtù del teorema 1.3.2, possiamo dire che, fissata arbitrariamente una funzione continua f di dominio $[0,1]$ e a valori reali, esiste una successione di funzioni polinomiali convergente uniformemente a f. Da ciò segue immediatamente che X con la norma $\|.\|_\infty$ non è completo. Per vederlo, basta prendere una successione $(x^k)_{k \in \mathbf{N}}$ di funzioni polinomiali convergente uniformemente a una funzione continua che non sia una funzione polinomiale (per esempio, a una funzione continua che non sia derivabile in qualche punto). Questa successione è di Cauchy in X in base all'osservazione 1.3.4, ma non è convergente in X.

Esercizio 1.3.1 Verificare che in ogni spazio normato l'eventuale limite di una successione è unico.

Esercizio 1.3.2 Verificare che in ogni spazio normato X se $(x_n)_{n \in \mathbf{N}}$ converge a x_0 e $(y_n)_{n \in \mathbf{N}}$ converge a y_0, allora $(x_n + y_n)_{n \in \mathbf{N}}$ converge a $x_0 + y_0$ e, se λ appartiene al campo K, $(\lambda x_n)_{n \in \mathbf{N}}$ converge a λx_0.

Esercizio 1.3.3 Con riferimento all'esempio 1.3.1, verificare che la successione $(x_n)_{n \in \mathbf{N}}$ converge a 0 in $L^1([0,1])$. Osservare che m è identificabile con la funzione identicamente nulla, nel senso che $[m] = [0]$.

Esercizio 1.3.4 Siano X, Y e Z tre spazi normati, $A \subseteq X$, $B \subseteq Y$, $f : A \to Y$, $g : B \to Z$, con $f(A) \subseteq B$. Sia poi $x_0 \in A$ tale che f è continua in x_0 e g è continua in $f(x_0)$. Provare che $g \circ f$ è continua in x_0. (Sugg.: applicare il teorema 1.3.3)

Esercizio 1.3.5 Verificare che $\mathbf{C}$, prendendo come norma il valore assoluto, è completo.

Esercizio 1.3.6 Sia $A \subseteq X$ spazio normato. Provare che le due condizioni seguenti sono equivalenti:

(I) A è chiuso;

(II) sia $l \in X$ tale che esiste una successione $(a_n)_{n\in\mathbf{N}}$ a valori in A convergente a l. Allora $l \in A$.

Esercizio 1.3.7 Studiare la convergenza delle seguenti successioni $(x_k)_{k\in\mathbf{N}}$ nello spazio $B(A, \mathbf{R})$ assegnato:

(a) $A = [0, +\infty[$; $x_k(t) = [2\sinh(2kt) - e^{2kt}]t$;

(b) $A =]0, 2]$; $x_k(t) = (\frac{t}{2})^k \ln(\frac{t}{2})$;

(c) $A = \mathbf{R}$; $x_k(t) = \sin(\frac{t}{k}) - \cos(\frac{t}{k})$;

(d) $A = [0, 2\pi]$; $x_k(t) = \sin(\frac{t}{k}) - \cos(\frac{t}{k})$;

(e) $A = [0, +\infty[$; $x_k(t) = \frac{\pi}{2} - \arctan(kt)$;

(f) $A = [1, +\infty[$; $x_k(t) = \frac{\pi}{2} - \arctan(kt)$;

(g) $A = \mathbf{R}^+$; $x_k(t) = \frac{t}{1+k^2t^2}$;

(h) $A = \mathbf{R}^+$; $x_k(t) = \frac{t}{t+k}$;

(i) $A = [0, 1]$; $x_k(t) = \frac{t}{t+k}$.

1.4 Spazi con prodotto interno, spazi di Hilbert

Cominciamo con la definizione di prodotto interno.

Definizione 1.4.1 *Siano X uno spazio vettoriale su K (al solito coincidente con $\mathbf{R}$ o C), $< .,. >$ un'applicazione da $X \times X$ a K (che associa dunque a ogni coppia ordinata (x, y) di $X \times X$ un elemento $< x, y >$ appartenente a K). Diremo che $< .,. >$ è un* **prodotto interno** *(o scalare) in X se valgono le seguenti condizioni:*

(PI1) $\forall y \in X$ $x \to < x, y >$ è lineare da X a K;

(P12) $< y, x > = \overline{< x, y >}$ $\forall x, y \in X$;

(P13) $< x, x >$ è un numero reale non negativo $\forall x \in X$;

(PI4) se $x \in X$ e $< x, x > = 0$, allora $x = 0$.

Osservazione 1.4.1 $(PI1)$ equivale a richiedere che $\forall \lambda_1, \lambda_2 \in K$, $\forall x^1$, x^2, $y \in X$ si abbia

$$< \lambda_1 x^1 + \lambda_2 x^2, y >= \lambda_1 < x^1, y > + \lambda_2 < x^2, y > .$$

Da $(PI1)$ e $(PI2)$ segue poi, $\forall \lambda_1, \lambda_2 \in K$, $\forall x, y^1, y^2 \in X$,

$$< x, \lambda_1 y^1 + \lambda_2 y^2 >=$$

$$\overline{< \lambda_1 y^1 + \lambda_2 y^2, x >} = \overline{\lambda_1 < y^1, x > + \lambda_2 < y^2, x >} =$$

$$= \overline{\lambda_1} < x, y^1 > + \overline{\lambda_2} < x, y^2 > .$$

Si usa dire che $\forall x \in X$ $y \to < x, y >$ è antilineare da X a K. Nel caso $K = \mathbf{R}$, ciò significa che un prodotto interno è lineare anche nel secondo argomento.

Esempio 1.4.1 Siano $X = \mathbf{R}^n$ e $K = \mathbf{R}$. Se $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$, poniamo

$$< x, y > := x_1 y_1 + ... + x_n y_n. \tag{1.4.1}$$

È facile verificare che l'applicazione definita in (1.4.1) è un prodotto scalare in $\mathbf{R}^n$ (che probabilmente sarà già noto al lettore!)

Esempio 1.4.2 Siano $X = \mathbf{C}^n$ e $K = \mathbf{C}$. Se $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$, poniamo

$$< x, y > := x_1 \overline{y_1} + ... + x_n \overline{y_n}. \tag{1.4.2}$$

L'applicazione definita in (1.4.2) è un prodotto scalare in $\mathbf{C}^n$. Lasciamo la facile verifica al lettore (vedi l'esercizio 1.4.1).

Esempio 1.4.3 Siano $A \in \mathcal{M}_n$ e $f : A \to \mathbf{C}$ misurabile. Diremo che f appartiene a $\mathcal{L}^2(A)$ se

$$\int_A |f(x)|^2 dx < +\infty.$$

Se a e b sono numeri reali non negativi, si ha

$$0 \leq (a - b)^2 = a^2 + b^2 - 2ab.$$

Ne segue la disuguaglianza

$$ab \leq \frac{1}{2}(a^2 + b^2). \tag{1.4.3}$$

Siano allora f e g elementi di $\mathcal{L}^2(A)$. Se $x \in A$, si ha

$$|f(x) + g(x)|^2 \leq (|f(x)| + |g(x)|)^2$$

$$= |f(x)|^2 + |g(x)|^2 + 2|f(x)||g(x)|$$

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

utilizzando (1.4.3). Perciò

$$\int_A |f(x) + g(x)|^2 dx$$

$$\leq \int_A 2(|f(x)|^2 + |g(x)|^2) dx$$

$$= 2(\int_A |f(x)|^2 dx + \int_A |g(x)|^2) < +\infty.$$

Dunque $f + g \in \mathcal{L}^2(A)$. È facile anche vedere che, se $\lambda \in \mathbf{C}$ e $f \in \mathcal{L}^2(A)$, anche $\lambda f \in \mathcal{L}^2(A)$. Ne deduciamo che $\mathcal{L}^2(A)$ è uno spazio vettoriale su $\mathbf{C}$ con le solite operazioni di somma e prodotto per uno scalare. È poi facile verificare che, se $f \in \mathcal{L}^2(A)$ e $f \sim g$ nel senso della definizione 0.4.1, allora anche $g \in \mathcal{L}^2(A)$ (vedi esercizio 1.4.2).

Definizione 1.4.2 *Sia $A \in \mathcal{M}_n$. Poniamo*

$$L^2(A) := \{[f] : f \in \mathcal{L}^2(A)\}.$$

Con somma e prodotto per uno scalare definiti in (0.4.2) e (0.4.3), $L^2(A)$ è uno spazio vettoriale su $\mathbf{C}$, come è facile verificare.

Siano ora f e g elementi di $\mathcal{L}^2(A)$. Applicando ancora (1.4.3), abbiamo, $\forall x \in A$,

$$|f(x)\overline{g(x)}| = |f(x)||g(x)| \leq \frac{1}{2}(|f(x)|^2 + |g(x)|^2).$$

Ne segue che $f\overline{g}$ è sommabile su A. Siano allora $[f]$ e $[g]$ elementi di $L^2(A)$. Poniamo

$$< [f], [g] > := \int_A f(x)\overline{g(x)} dx. \tag{1.4.4}$$

Si verifica facilmente (applicando il teorema 0.4.1) che (1.4.4) non dipende dalla scelta dei rappresentanti delle classe di equivalenza. Verifichiamo che $< .,. >$ è un prodotto scalare in $L^2(A)$.

Lasciamo al lettore la verifica di $(PI1)$.

Verifichiamo $(PI2)$. Si ha

$$< [g],[f] >=$$

$$= \int_A g(x)\overline{f(x)}dx = \int_A \overline{f(x)\overline{g(x)}}dx$$

$$= \overline{\int_A f(x)\overline{g(x)}dx}$$

(applicando la formula (0.2.9)

$$= \overline{< [f],[g] >}.$$

Verifichiamo $(PI3)$. Se $f \in \mathcal{L}^2(A)$,

$$< [f],[f] >= \int_A |f(x)|^2 dx \geq 0.$$

Verifichiamo infine $(PI4)$. Dire che $< [f],[f] >= 0$ equivale a dire che $\int_A |f(x)|^2 dx = 0$. Dal lemma 1.1.1 segue $|f(x)|^2 = 0$ q. d. in A, e dunque $f(x) = 0$ q. d. in A. Concludiamo che

$$[f] = [0].$$

Vediamo ora come su ogni spazio con prodotto interno si può introdurre in maniera naturale una norma.

Definizione 1.4.3 *Sia X uno spazio vettoriale su K su cui* è *fissato un prodotto interno $< .,. >$. Poniamo, $\forall x \in X$,*

$$\|x\| := \sqrt{< x,x >}.$$

Chiameremo l'applicazione $x \to \|x\|$ **norma associata** *al prodotto interno $< .,. >$.*

Osservazione 1.4.2 Osserviamo che la definizione 1.4.3 è ben posta in virtù della condizione $(PI3)$. Tra poco verificheremo che $\|.\|$ è effettivamente una norma secondo la definizione 1.1.2.

Per provare che $\|.\|$ è una norma utilizzeremo il seguente

Teorema 1.4.1 *Sia X uno spazio vettoriale su K col prodotto interno $<.,.>$. Sia poi $\|.\|$ come nella definizione 1.4.3. Allora $\forall x, y \in X$*

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

(Disuguaglianza di Cauchy-Schwarz)

Dimostrazione Siano $\alpha \in K$, che specificheremo nel seguito. Allora, $\forall t \in \mathbf{R}$, utilizzando $(PI1) - (PI3)$,

$$0 \leq < x + t\alpha y, x + t\alpha y >$$

$$= < x, x > + t(\alpha < y, x > + \overline{\alpha} < x, y >) + t^2 |\alpha|^2 < y, y >$$
$$= \|x\|^2 + 2tRe(\alpha < y, x >) + t^2 |\alpha|^2 \|y\|^2,$$

essendo

$$\alpha < y, x > + \overline{\alpha} < x, y > = \alpha < y, x > + \overline{\alpha < y, x >}$$
$$= 2Re(\alpha < y, x >).$$

Scegliamo

$$\alpha = \begin{cases} \frac{<x,y>}{|<x,y>|} & \text{se} \quad < x, y > \neq 0, \\ 1 & \text{se} \quad < x, y > = 0. \end{cases}$$

Osserviamo che in ogni caso $|\alpha| = 1$ e si ha

$$P(t) := < x + t\alpha y, x + t\alpha y > = \|x\|^2 + 2t| < x, y > | + t^2 \|y\|^2 \forall t \in \mathbf{R}.$$

P è un polinomio di grado non superiore a due a coefficienti reali e $P(t) \geq 0$ $\forall t \in \mathbf{R}$. Ciò implica che il suo discriminante è non positivo, vale a dire

$$| < x, y > |^2 - \|x\|^2 \|y\|^2 \leq 0,$$

da cui segue subito la conclusione. $\square$

Teorema 1.4.2 *Siano X uno spazio vettoriale su K e $<.,.>$ un prodotto interno in X. Allora la norma associata a $<.,.>$ (definizione 1.4.3) è una norma nel senso della definizione 1.1.2.*

Dimostrazione Indichiamo al solito con $\|.\|$ la norma associata a $< .,. >$. Siano $\lambda \in K$ e $x \in X$. Allora

$$\|\lambda x\| = \sqrt{< \lambda x, \lambda x >} = \sqrt{|\lambda|^2 < x, x >} = |\lambda| \|x\|.$$

Se x e y sono elementi di X, si ha poi

$$\|x + y\|^2 =< x + y, x + y >= \|x\|^2 + 2Re(< x, y >) + \|y\|^2$$

$$\leq \|x\|^2 + 2|Re(< x, y >)| + \|y\|^2 \leq \|x\|^2 + 2| < x, y > | + \|y\|^2$$

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$

(per Cauchy-Schwarz)

$$= (\|x\| + \|y\|)^2,$$

da cui, prendendo le radici quadrate,

$$\|x + y\| \leq \|x\| + \|y\|.$$

Infine, se $\|x\| = 0$, allora $< x, x >= 0$, da cui $x = 0$. $\square$

Esempio 1.4.4 Consideriamo l'esempio 1.4.1. Dato $x = (x_1, ..., x_n) \in \mathbf{R}^n$, otteniamo dal prodotto scalare la norma

$$\|x\| = \sqrt{< x, x >} = \sqrt{x_1^2 + ... + x_n^2},$$

che è la ben nota norma euclidea.

Esempio 1.4.5 Consideriamo l'esempio 1.4.2. Dato $x = (x_1, ..., x_n) \in \mathbf{C}^n$, otteniamo dal prodotto scalare la norma

$$\|x\| = \sqrt{< x, x >} = \sqrt{|x_1|^2 + ... + |x_n|^2}. \tag{1.4.5}$$

Esempio 1.4.6 Consideriamo l'esempio 1.4.3. Dal prodotto scalare definito in $L^2(A)$, otteniamo la norma

$$\|[f]\| = \sqrt{< [f], [f] >} = \sqrt{\int_A |f(x)|^2 dx}. \tag{1.4.6}$$

In questo caso la disuguaglianza di Cauchy-Schwarz si traduce nella seguente disuguaglianza integrale, valida per f e g in $\mathcal{L}^2(A)$:

$$|\int_A f(x)\overline{g(x)}dx| \leq \sqrt{\int_A |f(x)|^2 dx}\sqrt{\int_A |g(x)|^2 dx}. \tag{1.4.7}$$

Concludiamo con la definizione di spazio di Hilbert.

Definizione 1.4.4 *Sia X uno spazio normato con una norma $\|.\|$ associata a un certo prodotto interno $< .,. >$. Diremo che con tale norma X* è *uno* **spazio di Hilbert** *se* è *completo secondo la definizione 1.3.3.*

Abbiamo già visto (teorema 1.3.5) che $\mathbf{R}^n$ con la norma euclidea è completo ed è perciò uno spazio di Hilbert. Anche $\mathbf{C}^n$ con la norma 1.4.5 è completo (esercizio 1.4.3). Si potrebbe poi dimostrare il seguente

Teorema 1.4.3 *Sia $A \in \mathcal{M}_n$. Allora lo spazio $L^2(A)$ con la norma (1.4.6)* è *completo ed* è *quindi uno spazio di Hilbert.*

Esercizio 1.4.1 Verificare nei dettagli che l'applicazione (1.4.2) è un prodotto interno in $\mathbf{C}^n$. (Sugg.: si cominci col far vedere che, se z e v sono elementi di $\mathbf{C}$, allora $\overline{z\overline{v}} = \overline{z}v$. Si rammenti inoltre che $z\overline{z} = |z|^2$.)

Esercizio 1.4.2 Verificare che, se $f \in \mathcal{L}^2(A)$ e $f \sim g$, allora $g \in \mathcal{L}^2(A)$.

Esercizio 1.4.3 Verificare che $\mathbf{C}^n$ con la norma (1.4.5) è completo. (Sugg.: osservare che con tale norma $\mathbf{C}^n$ coincide con $\mathbf{R}^{2n}$ munito della corrispondente norma euclidea.)

1.5 Proiezioni ortogonali in uno spazio di Hilbert

Cominciamo con l'enunciare quello che è probabilmente il risultato più importante sugli spazi di Hilbert.

Ricordando l'interpretazione del prodotto scalare standard in $\mathbf{R}^n$, diremo che due elementi x e y in uno spazio con prodotto interno $< .,. >$ sono ortogonali se $< x, y >= 0$.

Teorema 1.5.1 *Siano H uno spazio di Hilbert, $A \subseteq H$ convesso e chiuso. Sia poi $x \in H$. Allora la funzione $y \to \|y - x\|$ ammette minimo in A. Tale minimo* è *assunto in un unico punto $P_A x \in A$.*

Osservazione 1.5.1 In sostanza il teorema 1.5.1 afferma che, se A è convesso e chiuso, per ogni punto $x \in H$ esiste un unico punto $P_A x$ di A che ha distanza minima da x. Nel caso in cui A sia un sottospazio chiuso di H, $P_A x$ può essere pensato come la proiezione ortogonale di x su A, in base al risultato seguente.

Teorema 1.5.2 *Siano H uno spazio di Hilbert, A un sottospazio chiuso di H. Sia poi $x \in H$. Allora, dato $y \in A$, le condizioni seguenti sono equivalenti:*
(I) $y = P_A x$;
(II) $< x - y, z >= 0$ $\forall z \in A$.

Dimostrazione Verifichiamo che, se $y = P_A x$, allora $< x - y, z >= 0$ $\forall z \in A$. Siano $z \in A$ e $t \in \mathbf{R}$. Poiché $y + tz \in A$ $\forall t \in \mathbf{R}$, la funzione

$$\begin{cases} P : \mathbf{R} \to \mathbf{R}, \\ P(t) := \|x - y - tz\|^2 \end{cases}$$

assume il minimo in $t = 0$. Si ha

$$P(t) = \|x - y\|^2 - 2tRe(< x - y, z >) + t^2\|z\|^2.$$

Dunque P è una funzione polinomiale. Deve essere $P'(0) = 0$, da cui

$$Re(< x - y, z >) = 0.$$

Se il campo K è $\mathbf{C}$, anche $iz \in A$ e dunque vale anche $Re(< x-y, iz >) = 0$; ma allora

$$Im(< x - y, z >) = Re(-i < x - y, z >) = Re(< x - y, iz >) = 0.$$

Con ciò abbiamo dimostrato che da (I) segue (II).

Sia invece $y \in A$ tale che $< x - y, z >= 0$ $\forall z \in A$. Dato allora $z \in A$, si ha

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2$$

$$= \|x - y\|^2 + 2Re(< x - y, y - z >) + \|y - z\|^2 = \|x - y\|^2 + \|y - z\|^2$$

(perché $y - z \in A$)

$$\geq \|x - y\|^2.$$

Perciò $y = P_A x$. $\square$

Vogliamo ora fornire una formula esplicita per P_A nel caso in cui A sia un sottospazio di dimensione finita di H.

Premettiamo la seguente

Definizione 1.5.1 *Sia $\{x^i : i \in \mathcal{I}\}$, con $\mathcal{I}$ arbitrario, insieme di indici un sottoinsieme dello spazio di Hilbert H. Diremo che $\{x^i : i \in \mathcal{I}\}$* è *un sistema* **ortonormale** *se, $\forall i, j \in \mathcal{I}$,*

$$< x^i, x^j >= \delta_{ij}$$

con δ_{ij} simbolo di Kronecker.

Osservazione 1.5.2 In sostanza, nella definizione 1.5.1 chiediamo che i vettori x^i siano a due a due ortogonali e abbiano norma unitaria.

Per esempio, la base canonica $\{e^i : i \in \{1, ..., n\}\}$ di $\mathbf{R}^n$ è un sistema ortonormale rispetto al prodotto scalare standard (vedi l'esempio 1.4.1). La stessa cosa vale in $\mathbf{C}^n$ con il prodotto interno descritto nell'esempio 1.4.2.

Come vedremo, sarà molto utile lavorare con una base ortonormale. A tale proposito, premettiamo il seguente

Teorema 1.5.3 *Sia A un sottospazio di dimensione finita dello spazio di Hilbert H. Allora A possiede una base ortonormale.*

Dimostrazione La dimostrazione che forniamo è costruttiva, nel senso che fornisce un metodo (il così detto **metodo di Gram-Schmidt**) per costruire una base ortonormale di A a partire da una base qualunque $\{f^i : 1 \leq i \leq n\}$, con n dimensione di A. Per semplicità , consideriamo il caso $n = 3$, anche se il metodo è del tutto generale e segue le linee del caso particolare che trattiamo.

Sia $\{f^1, f^2, f^3\}$ una base di A. Poniamo

$$e^1 := \frac{1}{\|f^1\|} f^1.$$

(si osservi che $f^1 \neq 0$). Si ha $\|e^1\| = 1$.

Cerchiamo un elemento $\epsilon^2 \in A$ della forma

$$\epsilon^2 = ce^1 + f^2$$

che sia ortogonale a e^1. Otteniamo

$$0 =< \epsilon^2, e^1 >= c+ < f^2, e^1 > .$$

Dobbiamo allora prendere

$$c = - < f^2, e^1 > .$$

Si osservi che, qualunque sia c, $\epsilon^2 \neq 0$, in quanto e^1 e f^2 sono linearmente indipendenti. Poniamo ora

$$e^2 := \frac{1}{\|\epsilon^2\|} \epsilon^2.$$

È chiaro che e^1 ed e^2 sono ortogonali e di norma unitaria. Sono inoltre entrambi combinazioni lineari di f^1 e f^2.

Cerchiamo infine un elemento $\epsilon^3 \in A$ della forma

$$\epsilon^2 = c_1 e^1 + c^2 e^2 + f^3$$

che sia ortogonale a e^1 ed e^2. Otteniamo

$$0 =< \epsilon^3, e^1 >= c_1 + < f^3, e^1 >,$$

$$0 =< \epsilon^3, e^2 >= c_2 + < f^3, e^2 > .$$

Dobbiamo allora prendere

$$c_1 = - < f^3, e^1 >, c_2 = - < f^3, e^2 > .$$

Si ha che $\epsilon^3 \neq 0$, in quanto f^3 non dipende linearmente da e^1 ed e^2, in quanto non dipende linearmente da f^1 ed f^2. Poniamo allora

$$e^3 := \frac{1}{\|\epsilon^3\|}\epsilon^3,$$

e $\{e^1, e^2, e^3\}$ è la base ortonormale cercata. $\square$

Esempio 1.5.1 Siano $H = L^2([-1,1])$, A il sottospazio delle classi di equivalenza contenenti la restrizione a $[-1,1]$ di qualche funzione polinomiale di grado non superiore a 2. È evidente che A è uno spazio di dimensione finita e una sua base è $\{[1], [t], [t^2]\}$. Nel seguito, data $f \in \mathcal{L}^2([0,1])$, scriveremo f al posto di $[f]$, per semplicità di notazione. Sia $f^1(t) = 1$, $(t \in [-1,1])$. Sia ha

$$\|f^1\| = (\int_{-1}^{1} 1^2 dt)^{\frac{1}{2}} = \sqrt{2}.$$

Poniamo allora $e^1(t) := \frac{1}{\sqrt{2}}$ $(t \in [-1,1])$. Seguendo il metodo generale della dimostrazione del teorema 1.5.3, poniamo

$$c := - < f^2, e^1 >= - \int_{-1}^{1} \frac{t}{\sqrt{2}} dt = 0.$$

Dunque,

$$\epsilon^2(t) = t, t \in [-1,1].$$

Si ha

$$\|\epsilon^2\| = (\int_{-1}^{1} t^2 dt)^{\frac{1}{2}} = \sqrt{\frac{2}{3}},$$

da cui

$$e^2(t) = \sqrt{\frac{3}{2}}t.$$

Poniamo infine

$$c_1 := - < f^3, e^1 >= - \int_{-1}^{1} \frac{t^2}{\sqrt{2}} dt = -\frac{\sqrt{2}}{3},$$

$$c_2 := - < f^3, e^2 >= - \int_{-1}^{1} t^2 (\sqrt{\frac{3}{2}} t) dt = 0.$$

Abbiamo dunque

$$\epsilon^3(t) = t^2 - \frac{1}{3}.$$

Si vede poi che $\|\epsilon^3\| = \frac{2}{3}\sqrt{\frac{2}{5}}$. Poniamo allora

$$e^3(t) = \frac{3}{2}\sqrt{\frac{5}{2}}(t^2 - \frac{1}{3}).$$

$\{e^1, e^2, e^3\}$ è una base ortonormale di A.

Vediamo ora una formula per P_A nel caso di A sottospazio di H di dimensione finita.

Teorema 1.5.4 *Siano H uno spazio di Hilbert, A un suo sottospazio di dimensione finita, $\{e^1, ..., e^n\}$ una base ortonormale di A. Allora A è chiuso. Inoltre, per ogni $x \in H$,*

$$P_A x = \sum_{i=1}^{n} < x, e^i > e^i. \tag{1.5.1}$$

Dimostrazione Il fatto che A sia chiuso è lasciato al lettore (esercizio 1.5.2). Per quanto riguarda (1.5.1), poiché $P_A x \in A$, è della forma

$$P_A x = c_1 e^1 + ... + c_n e^n,$$

con c_1, ..., c_n elementi del campo K. Sia $1 \leq j \leq n$. Dal teorema 1.5.2 segue

$$0 =< x - \sum_{i=1}^{n} c_i e^i, e^j >=< x, e^j > -c_j,$$

da cui la conclusione. □

Esempio 1.5.2 Siano $H = L^2([0,1])$ e A il sottospazio di H delle funzioni polinomiali di grado non superiore a 2. Sia $f(t) = \sin(t)$. f è un elemento di H. $P_A f$ può essere interpretato come l'elemento di A che approssima meglio f, nel senso che $\int_{[-1,1]} |f(t) - P(t)|^2 dt$, con P polinomio di grado non superiore a 2, è minimo proprio per $P = P_A f$. Nell'esempio 1.5.1 abbiamo determinato una base ortonormale $\{e^1, e^2, e^3\}$ di A. Utilizzando i calcoli già fatti, otteniamo allora che

$$P_A f(t) = \sum_{i=1}^{3} < f, e^i > e^i(t) =$$

$$= \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{\sin(s)}{\sqrt{2}} ds + \int_{-1}^{1} \sin(s) \sqrt{\frac{3}{2}} s ds \sqrt{\frac{3}{2}} t$$

$$+ \int_{-1}^{1} \sin(s) \frac{3}{2} \sqrt{\frac{5}{2}} (s^2 - \frac{1}{3}) ds \ \frac{3}{2} \sqrt{\frac{5}{2}} (t^2 - \frac{1}{3}) =$$

$$= 3(\sin(1) - \cos(1))t.$$

Concludiamo con il seguente risultato.

Teorema 1.5.5 *Sia $(V_n)_{n \in \mathbf{N}}$ una successione di sottospazi chiusi dello spazio di Hilbert H, tale che $V_n \subseteq V_{n+1}$ $\forall n \in \mathbf{N}$ e $V := \bigcup_{n \in \mathbf{N}} V_n$ è denso in H. Dato $n \in \mathbf{N}$, indichiamo con P_n la proiezione ortogonale su V_n. Allora, $\forall x \in H$ vale*

$$\lim_{n \to +\infty} P_n x = x.$$

Dimostrazione Sia $\epsilon > 0$. Applicando il teorema 1.3.2, possiamo dire che esiste $v_0 \in V$ tale che $\|x - v_0\| < \epsilon$. Sia $v_0 \in V_{n_0}$ con $n_0 \in \mathbf{N}$. Se $n > n_0$, si avrà allora, tenendo conto che $v_0 \in V_n$,

$$\|x - P_n x\| \leq \|x - v_0\| < \epsilon.$$

□

Esercizio 1.5.1 Dimostrare che, se A è un sottospazio chiuso di H, l'operatore $P_A : H \to A$ è lineare. (Sugg.: utilizzare il teorema 1.5.2)

Esercizio 1.5.2 Dimostrare che, se A è un sottospazio di dimensione finita di uno spazio di Hilbert H, allora A è chiuso. (Sugg.: sia $\{e^1, ..., e^n\}$ una base ortonormale di A. Se $x = x_1 e^1 + ... + x_n e^n$, allora

$$\|x\| = (|x_1|^2 + ... + |x_n|^2)^{\frac{1}{2}} = \|(x_1, ..., x_n)\|_{\mathbf{C}^n},$$

ove $\|.\|_{\mathbf{C}^n}$ è la norma in $\mathbf{C}^n$ introdotta nell'esempio 1.4.2. Si usi poi il risultato dell'esercizio 1.4.3.

Esercizio 1.5.3 Costruire una base del sottospazio delle funzioni polinomiali di grado non superiore a 3 che sia ortonormale rispetto al prodotto scalare in $L^2([-1, 1])$.

Esercizio 1.5.4 Siano x^1,...,x^n elementi di H a due a due ortogonali. Verificare che

$$\|x^1 + ... + x^n\|^2 = \|x^1\|^2 + ... + \|x^n\|^2.$$

Questa formula può essere considerata una generalizzazione del classico teorema di Pitagora.

1.6 Serie di Fourier

Cominciamo col ricordare che, dato $y \in \mathbf{R}$, vale per definizione

$$e^{iy} = \cos(y) + i\sin(y).$$

Si osservi che si ha

$$\overline{e^{ix}} = \cos(x) - i\sin(x) =$$

$$= \cos(-x) + i\sin(-x) = e^{-ix}.$$

Vale inoltre la fondamentale formula

$$e^{ix}e^{iy} = e^{i(x+y)} \qquad \forall x, y \in \mathbf{R}, \tag{1.6.1}$$

da cui segue

$$(e^{ix})^n = e^{inx} \qquad \forall n \in \mathbf{N}$$

Tenuto conto poi che $e^{i0} = 0$, da (1.6.1) si vede che

$$(e^{ix})^{-1} = e^{-ix} \qquad \forall x \in \mathbf{R}.$$

Osserviamo infine che

$$\cos(x) = Re(e^{ix}) = \frac{e^{ix} + \overline{e^{ix}}}{2} = \frac{e^{ix} + e^{-ix}}{2} \tag{1.6.2}$$

e

$$\sin(x) = Im(e^{ix}) = \frac{e^{ix} - \overline{e^{ix}}}{2i} = \frac{e^{ix} - e^{-ix}}{2i} \tag{1.6.3}$$

(1.6.2) e (1.6.3) sono le ben note **formule di Eulero**. Siano ora m e n numeri interi, con $m \neq n$. Si ha

$$\begin{gathered} \int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}dx = \\ = \int_{-\pi}^{\pi} e^{i(n-m)x}dx = \int_{-\pi}^{\pi} \cos((n-m)x)dx + i\int_{-\pi}^{\pi} \sin((n-m)x)dx = \\ = 0. \end{gathered} \tag{1.6.4}$$

Invece,

$$\int_{-\pi}^{\pi} e^{inx}e^{-inx}dx = 2\pi. \tag{1.6.5}$$

Da (1.6.4) e (1.6.5) si deduce che $\{\frac{e^{inx}}{\sqrt{2\pi}} : n \in \mathbf{Z}\}$ è un sistema ortonormale in $L^2([-\pi,\pi])$. Dato $n \in \mathbf{N}$, indichiamo con V_n il sottospazio di $L^2([-\pi,\pi])$ generato da $\{\frac{e^{ikx}}{\sqrt{2\pi}} : -n \leq k \leq n\}$. Se $f \in L^2([-\pi,\pi])$, la proiezione ortogonale P_nf di f su V_n è data da

$$P_nf(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\sum_{k=-n}^{n} e^{ik(t-s)}ds \tag{1.6.6}$$

Tenendo conto della formula di somma dei termini di una progressione geometrica, si ha, per $e^{i(t-s)} \neq 1$,

$$\sum_{k=-n}^{n} e^{ik(t-s)} = e^{-in(t-s)}\sum_{k=0}^{2n} e^{ik(t-s)} = e^{-in(t-s)}\frac{1-e^{i(2n+1)(t-s)}}{1-e^{i(t-s)}}$$

$$= e^{-in(t-s)}\frac{(1-e^{i(2n+1)(t-s)})(1-e^{-i(t-s)})}{|1-e^{i(t-s)}|^2}.$$

Si ha

$$|1-e^{i(t-s)}|^2 = [1-\cos(t-s)]^2 + \sin^2(t-s)$$

$$= 2 - 2\cos(t-s) = 4\sin^2(\frac{t-s}{2}).$$

Inoltre

$$e^{-in(t-s)}(1 - e^{i(2n+1)(t-s)})(1 - e^{-i(t-s)}) =$$

$$= e^{in(t-s)} + e^{-in(t-s)} - e^{i(n+1)(t-s)} - e^{-i(n+1)(t-s)}$$

$$= 2[\cos(n(t-s)) - \cos((n+1)(t-s))]$$

(per le formule di Eulero)

$$= 4\sin((n+\frac{1}{2})(t-s))\sin(\frac{t-s}{2})$$

(per le formule di prostaferesi). Concludiamo che, se $e^{i(t-s)} \neq 1$, si ha

$$\sum_{k=-n}^{n} e^{ik(t-s)} = \frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})}.$$

Prolungando arbitrariamente $\frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})}$ nei punti (in numero finito in $[-\pi,\pi]$) in cui $\sin(\frac{t-s}{2})$ si annulla, abbiamo ottenuto

$$P_n f(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})} f(s)ds. \tag{1.6.7}$$

Poniamo $V := \bigcup_{n\in\mathbf{N}} V_n$. Si potrebbe dimostrare che V è denso in $L^2([-\pi,\pi]$. Applicando allora il teorema 1.5.5, otteniamo il seguente

Teorema 1.6.1 *Sia $f \in L^2([-\pi,\pi]$. Allora la successione $(P_n f)_{n\in\mathbf{N}}$, con $P_n f$ dato da (1.6.7), converge a f, per $n \to +\infty$, in $L^2([-\pi,\pi])$.*

Si usa anche scrivere

$$f = \sum_{n=-\infty}^{+\infty} \frac{e^{int}}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-ins}ds, \tag{1.6.8}$$

nel senso che $f = \lim_{n\to+\infty} \sum_{k=-n}^{n} \frac{e^{ik\cdot}}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-ins}ds$ in $L^2([-\pi,\pi])$. Osserviamo che in generale non c'è nessuna convergenza puntuale. Si osservi, per altro, che, poiché gli elementi di $L^2([-\pi,\pi])$ sono classi di equivalenza di funzioni e non funzioni, il valore di $f(t)$, che dipende dal rappresentante, non ha senso. Ci si può però chiedere se, ad esempio, data $f \in C([-\pi, \pi])$, valga puntualmente $\lim_{n \to +\infty} P_n f(t) = f(t)$, con $P_n f(t)$ definita in (1.6.7). Si potrebbe dimostrare il seguente

Teorema 1.6.2 *Sia $f \in C([-\pi, \pi])$ tale che*
(I) $f(-\pi) = f(\pi)$;
(II) esistono C e α positivi tali che

$$|f(t) - f(s)| \leq C|t - s|^{\alpha} \tag{1.6.9}$$

$\forall t, s \in [-\pi, \pi]$.
Allora la successione $(P_n f)_{n \in \mathbf{N}}$ converge a f uniformemente in $[-\pi, \pi]$.

Osservazione 1.6.1 Applicando il teorema del valor medio, non è difficile verificare che, se $Re(f)$ e $Im(f)$ sono di classe C^1, allora l'ipotesi (II) del teorema 1.6.2 è soddisfatta, con $\alpha = 1$. In generale, si potrebbe verificare che, se $f \in C([-\pi, \pi])$ e vale (I), la conclusione del teorema 1.6.2 è falsa.

Osservazione 1.6.2 Data $f \in L^2([-\pi, \pi]$, chiameremo (1.6.8) **sviluppo di Fourier** classico di f.

In generale se $S := \{e^n : n \in \mathbf{N}\}$ è un sistema ortonormale nello spazio di Hilbert H tale che lo spazio vettoriale generato da S è denso in H, si ha che, $\forall f \in H$,

$$f = \lim_{n \to +\infty} \sum_{k=1}^{n} < f, e^k > e^k. \tag{1.6.10}$$

scriveremo

$$f = \sum_{n=1}^{\infty} < f, e^n > e^n \tag{1.6.11}$$

e parleremo anche in questo caso di sviluppo di Fourier di f. Vedremo un esempio significativo nella prossima sezione.

Esercizio 1.6.1 Verificare quanto asserito nella prima parte dell'osservazione 1.6.1.

Esercizio 1.6.2 Verificare che, se $f = \sum_{n=1}^{\infty} < f, e^n > e^n$ è uno sviluppo di Fourier di $f \in H$ nel senso di (1.6.10), si ha

$$\|f\|^2 = \sum_{n=1}^{\infty} | < f, e^n > |^2.$$

(identità di Bessel) (Sugg.: utilizzare i risultati degli esercizi 1.2.6 e 1.5.4.)

1.7 Un'applicazione degli sviluppi di Fourier all' equazione del calore

Consideriamo l'andamento nel tempo della temperatura u di una sbarra costituita di un materiale omogeneo. Conosciamo la temperatura all'istante iniziale $t = 0$ e un termostato mantiene la temperatura costante (per esempio a livello 0) agli estremi della sbarra. Schematizziamo la sbarra mediante con un intervallo $[0, l] \subseteq$, con $l > 0$. Se $t \geq 0$ e $x \in [0, l]$, indichiamo con $u(t, x)$ la temperatura in x all'istante $t \geq 0$. La fisica ci dice che u soddisfa la seguente equazione differenziale:

$$D_t u(t, x) = k^2 D_x^2 u(t, x), \qquad (t, x) \in]0, +\infty[\times[0, l], \tag{1.7.1}$$

con k costante positiva che dipende dal materiale. Valgano inoltre

$$u(0, x) = f(x), \qquad x \in [0, l] \tag{1.7.2}$$

e

$$u(t, 0) = u(t, l) = 0, \qquad t \geq 0. \tag{1.7.3}$$

Qui f è una funzione continua a valori reali di dominio $[0, l]$, tale che $f(0) = f(l) = 0$ (questa condizione assicura che (1.7.2) e (1.7.3) sono compatibili).

Per soluzione del sistema (1.7.1)-(1.7.2)-(1.7.3) intenderemo una funzione u continua in $[0, +\infty[\times[0, l]$, dotata in $]0, +\infty[\times[0, l]$ di derivate $D_t u$, $D_x u$, $D_x^2 u$ continue, soddisfacente le tre condizioni (1.7.1)-(1.7.2)-(1.7.3).

In questo caso semplice, applicheremo il così detto **metodo della separazione delle variabili**: cercheremo prima soluzioni di (1.7.1) e (1.7.3) che siano prodotto di una funzione di t con una funzione di x. Costruiremo poi una soluzione del problema mediante sviluppi in serie delle soluzioni particolari trovate.

Cominciamo allora a realizzare il nostro programma, cercando le funzioni della forma

$$v(t,x) = T(t)X(x), \qquad t \geq 0, x \in [0,l], \tag{1.7.4}$$

non identicamente nulle e soddisfacenti (1.7.1) e (1.7.3). Dovrà allora essere

$$\begin{cases} T'(t)X(x) = k^2 T(t)X''(x), & t > 0, x \in [0,l], \\ \quad X(0) = X(l) = 0. \end{cases} \tag{1.7.5}$$

Sia $t_0 > 0$ tale che $T(t_0) \neq 0$. Allora deve valere

$$\begin{cases} \lambda X(x) = X''(x), & x \in [0,l], \\ X(0) = X(l) = 0, \end{cases} \tag{1.7.6}$$

con $\lambda := \frac{T'(t_0)}{k^2 T(t_0)}$.

Nello stesso modo, se $x_0 \in]0,l[$ e $X(x_0) \neq 0$, si ha

$$T'(t) = \frac{k^2 X''(x_0)}{X(x_0)} T(t) = \lambda k^2 T(t). \tag{1.7.7}$$

Ci interessano le soluzioni non identicamente nulle. Consideriamo allora il problema (1.7.6). Non è difficile verificare che possiede soluzioni non nulle se e solo se λ è della forma

$$\lambda = -\frac{n^2\pi^2}{l^2}, n \in \mathbf{N}. \tag{1.7.8}$$

In tal caso le soluzioni sono le funzioni X della forma

$$X(x) = C\sin(\frac{n\pi x}{l}), \tag{1.7.9}$$

con $C \in \mathbf{R}$ arbitrario. Per $\lambda = -\frac{n^2\pi^2}{l^2}$, si ottengono le soluzioni di (1.7.7), che sono le funzioni della forma

$$T(t) = Ce^{-\frac{(n\pi k)^2 t}{l^2}}, \tag{1.7.10}$$

con $C \in \mathbf{R}$ arbitrario. Otteniamo dunque per ogni $n \in \mathbf{N}$ la funzione

$$v_n(t,x) = e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}), \tag{1.7.11}$$

che verifica (1.7.1) e (1.7.3). Cerchiamo allora una soluzione del sistema (1.7.1)-(1.7.2)-(1.7.3) nella forma

$$u(t,x) = \sum_{n=1}^{\infty} c_n v_n(t,x). \tag{1.7.12}$$

Operando formalmente, si ottiene

$$f(x) = u(0,x) = \sum_{n=1}^{\infty} c_n \sin(\frac{n\pi x}{l}). \tag{1.7.13}$$

Poniamo

$$\epsilon_n(x) := \sin(\frac{n\pi x}{l}). \tag{1.7.14}$$

Indichiamo con $< .,. >$ il prodotto scalare in $L^2([0,l])$. Se $m \neq n$, si ha

$$< \epsilon_n, \epsilon_m >= \int_0^l \sin(\frac{n\pi x}{l}) \sin(\frac{m\pi x}{l}) dx$$

$$= \frac{1}{2}[\int_0^l \cos(\frac{(n-m)\pi x}{l}) dx - \int_0^l \cos(\frac{(n+m)\pi x}{l}) dx] = 0,$$

mentre

$$< \epsilon_n, \epsilon_n >= \int_0^l \sin^2(\frac{n\pi x}{l}) dx = \frac{1}{2}\{\int_0^l [1 - \cos(\frac{2n\pi x}{l})] dx\} =$$

$$= \frac{l}{2}.$$

Se poniamo allora

$$e_n(x) = \sqrt{\frac{2}{l}} \sin(\frac{n\pi x}{l}), \tag{1.7.15}$$

otteniamo che $\{e_n : n \in \mathbf{N}\}$ costituisce un sistema ortonormale in $L^2([0,l])$. Ebbene, si potrebbe dimostrare che lo spazio vettoriale generato da $\{e_n : n \in \mathbf{N}\}$ è denso in $L^2([0,l])$. Ciò implica, in base all'osservazione 1.6.2, che ogni elemento $f \in L^2([0,l])$ può essere rapprentato nella forma (1.6.11), con la serie convergente in $L^2([0,l])$. Sarà allora naturale prendere in (1.7.12)

$$c_n = \sqrt{\frac{2}{l}} < f, e_n >= \frac{2}{l} \int_0^l \sin(\frac{n\pi y}{l}) f(y) dy, \tag{1.7.16}$$

da cui otteniamo

$$u(t,x) = \frac{2}{l}\sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}) \int_0^l \sin(\frac{n\pi y}{l}) f(y)dy \tag{1.7.17}$$

Osserviamo che, se $t > 0$, la serie nel secondo membro di (1.7.17) converge assolutamente, in quanto, $\forall n \in N$,

$$|e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}) \int_0^l \sin(\frac{n\pi y}{l}) f(y)dy|$$

$$\leq e^{-\frac{(n\pi k)^2 t}{l^2}} \int_0^l |f(y)|dy,$$

e la serie che ha quest'ultima espressione come termine $n-$esimo converge per il criterio della radice. Si potrebbe, di più , verificare che è lecito derivare rispetto a t e a x sotto il segno di serie in (1.7.17) e che la funzione così definita:

$$u(t,x) = \begin{cases} \frac{2}{l}\sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}) \int_0^l \sin(\frac{n\pi y}{l}) f(y)dy \\ \text{se } t > 0, x \in [0,l], \\ f(x) \\ \text{se } t = 0, x \in [0,l], \end{cases} \tag{1.7.18}$$

è una soluzione nel senso richiesto del sistema (1.7.1)-(1.7.2)-(1.7.3). Ci si può chiedere se ce ne sono altre.

Sia allora v una seconda soluzione dello stesso problema. Si vede allora subito che $z := u - v$ risolve (1.7.1)-(1.7.2)-(1.7.3) con $f = 0$. In conclusione, ci poniamo il seguente problema: sia $z \in C([0,+\infty[\times[0,l])$ dotata di derivate $D_t z$, $D_x z$, $D_x^2 z$ continue in $]0,+\infty[\times[0,l]$, tale che valgano

$$D_t z(t,x) = k^2 D_x^2 z(t,x), \quad \forall (t,x) \in]0,+\infty[\times[0,l], \tag{1.7.19}$$

$$z(0,x) = 0, \quad \forall x \in [0,l] \tag{1.7.20}$$

e

$$z(t,0) = z(t,l) = 0, \quad \forall t \geq 0. \tag{1.7.21}$$

Possiamo dire che z è identicamente nulla?. A tale scopo, introduciamo la funzione $E : [0,+\infty[\to \mathbf{R}$, così definita:

$$E(t) := \int_0^l z(t,x)^2 dx. \tag{1.7.22}$$

Dalle ipotesi di regolarità fatte su z, si potrebbe verificare che E è continua in $[0,+\infty[$ e derivabile in $]0,+\infty[$ e che, per $t > 0$, si ha

$$E'(t) = 2\int_0^l z(t,x) D_t z(t,x) dx.$$

Ne segue allora che

$$E'(t) = 2\int_0^l z(t,x) D_x^2 z(t,x) dx =$$

(integrando per parti e usando (1.7.21))

$$= -2\int_0^l (D_x z(t,x))^2 dx \leq 0.$$

Ne segue che E è non crescente in $[0,+\infty[$. Perciò , $\forall t \geq 0$,

$$0 \leq E(t) \leq E(0) = 0.$$

Concludiamo che $E(t) = 0$ $\forall t \geq 0$. Applicando il lemma 1.1.1, possiamo dire che $\forall t \geq 0$ $u(t,x) = 0$ q. d. in $[0,l]$. Usando allora il fatto che f è continua (vedi esercizio 1.7.1), possiamo dire che z è identicamente nulla.

Possiamo condensare quello che abbiamo visto nel seguente

Teorema 1.7.1 *Consideriamo il problema (1.7.1)-(1.7.2)-(1.7.3), con $f \in C([0,l],\mathbf{R})$, $f(0) = f(l) = 0$. Allora tale problema possiede un'unica soluzione u continua e a valori reali in $[0,+\infty[\times[0,l]$ e dotata inoltre delle derivate $D_t u$, $D_x u$, $D_x^2 u$ continue in $]0,+\infty[\times[0,l]$. Tale soluzione u ammette la rappresentazione (1.7.17).*

Esercizio 1.7.1 Sia $f \in C([0,l])$, tale che $f(x) = 0$ q. d. in $[0,l]$. Verificare che $f(x) = 0$ $\forall x \in [0,l]$.

Esercizio 1.7.2 Si consideri il problema (1.7.1)-(1.7.2), con la condizione

$$D_x u(t,0) = D_x u(t,l) = 0, \qquad t \geq 0 \tag{1.7.23}$$

al posto di (1.7.3). Con ragionamenti analoghi a quelli impiegati per il problema (1.7.1)-(1.7.2)-(1.7.3) si studi il problema (1.7.1)-(1.7.2)-(1.7.23).

Si utilizzi il fatto che lo spazio vettoriale generato da $\{\frac{1}{\sqrt{l}}\} \cup \{\sqrt{\frac{2}{l}}\cos(\frac{n\pi.}{l}) : n \in \mathbf{N}\}$ è denso in $L^2([0,l])$. In tal modo si dimostri almeno parzialmente il seguente

Teorema 1.7.2 *Consideriamo il problema (1.7.1)-(1.7.2)-(1.7.23), con* $f \in C([0,l],\mathbf{R})$. *Allora tale problema possiede un'unica soluzione* u *continua e a valori reali in* $[0,+\infty[\times[0,l]$ *e dotata inoltre delle derivate* $D_t u$, $D_x u$, $D_x^2 u$ *continue in* $]0,+\infty[\times[0,l]$. *Tale soluzione* u *ammette la rappresentazione*

$$u(t,x) = \begin{cases} \frac{2}{l}[2\int_0^l f(y)dy + \sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \cos(\frac{n\pi x}{l}) & \\ \int_0^l \cos(\frac{n\pi y}{l}) f(y)dy] & \text{se } t > 0, x \in [0,l], \\ f(x) & \text{se } t = 0, x \in [0,l]. \end{cases} \tag{1.7.24}$$

Capitolo 2

Funzioni di una variabile complessa

2.1 Funzioni olomorfe

Il valore assoluto è una norma su $\mathbf{C}$, pensato come spazio vettoriale su se stesso. Si ricordi che, come insieme, $\mathbf{C} = \mathbf{R}^2$ e si osservi che il valore assoluto complesso coincide con la norma euclidea in $\mathbf{R}^2$ stesso. Utilizzando allora ben noti risultati della teoria delle funzioni di più variabili reali, possiamo dire che:

Lemma 2.1.1 *Dati $A \subseteq \mathbf{C}$, $f : A \to \mathbf{C}$ e $z \in A$, f è continua in z se e solo $Re(f)$ e $Im(f)$ sono continue in z.*

Se $z \in D(A)$, $\lim_{v \to Z} f(v) = l$ equivale a

$$\lim_{v \to z} Re(f(v)) = Re(l), \lim_{v \to z} Im(f(v)) = Im(l).$$

Vediamo ora la definizione di derivata complessa.

Definizione 2.1.1 *Siano $A \subseteq \mathbf{C}$, A aperto, $f : A \to \mathbf{C}$, $z^0 \in A$. Diremo che f è* **derivabile in senso complesso** *in z^0 se esiste in $\mathbf{C}$*

$$f'(z^0) := \lim_{z \to z^0} \frac{f(z) - f(z^0)}{z - z_0} = \lim_{h \to 0} \frac{f(z^0 + h) - f(z^0)}{h}. \tag{2.1.1}$$

Chiameremo $f'(z^0)$ derivata (complessa) di f in z^0.

Definizione 2.1.2 *Siano $A \subseteq \mathbf{C}$, A aperto, $f : A \to \mathbf{C}$. Diremo che f è* **olomorfa** *in A se è derivabile in senso complesso in ogni punto di A e inoltre la funzione derivata $z \to f'(z)$ è continua su A.*

Esempio 2.1.1 Siano $n \in \mathbf{N}$, $f : \mathbf{C} \to \mathbf{C}$, $f(z) = z^n$. Allora f è olomorfa. Infatti, $\forall z \in \mathbf{C}$, si ha

$$\lim_{h\to 0} \frac{(z+h)^n - z^n}{h} = \lim_{h\to 0}[(z+h)^{n-1} + (z+h)^{n-2}z + ... + (z+h)z^{n-2} + z^{n-1}]$$
$$= nz^{n-1}.$$

Esempio 2.1.2 Sia $f : \mathbf{C} \to \mathbf{C}$, $f(z) = Re(z)$. Allora f non è derivabile in senso complesso in alcun punto di $\mathbf{C}$. Infatti, se $z \in \mathbf{C}$,

$$\lim_{h\to 0, h\in\mathbf{R}} \frac{Re(z+h) - Re(z)}{h} = 1,$$

mentre

$$\lim_{h\to 0, h\in\mathbf{R}} \frac{Re(z+ih) - Re(z)}{ih} = 0.$$

Valgono alcune proprietà già viste a suo tempo per la derivata reale:

Teorema 2.1.1 *Siano $A \subseteq \mathbf{C}$ aperto, $f, g : A \to \mathbf{C}$, $z^0 \in A$. Siano poi f e g derivabili in senso complesso in z^0. Allora:*
(I) f è continua in z^0;
(II) $f + g$ è derivabile in z^0 e

$$(f+g)'(z^0) = f'(z^0) + g'(z^0);$$

(III) fg è derivabile in z^0 e

$$(fg)'(z^0) = f'(z^0)g(z^0) + f(z^0)g'(z^0);$$

(IV) se $g(z) \neq 0$ $\forall z \in A$, $\frac{f}{g}$ è derivabile in z^0 e

$$(\frac{f}{g})'(z^0) = \frac{f'(z^0)g(z^0) - f(z^0)g'(z^0)}{g(z^0)^2}.$$

Teorema 2.1.2 *Siano A e B aperti in $\mathbf{C}$, $f : A \to \mathbf{C}$, $g : B \to \mathbf{C}$ tali che $f(A) \subseteq B$, $z^0 \in A$. Siano poi f derivabile in z^0 e g derivabile in $f(z^0)$. Allora $g \circ f$ è derivabile in z^0 e*

$$(g \circ f)'(z^0) = g'(f(z^0))f'(z^0).$$

Il prossimo, importantissimo, risultato stabilisce il legame fondamentale tra le parti reale e immaginaria di una funzione olomorfa.

Teorema 2.1.3 *Siano $A \subseteq \mathbf{C}$ aperto, $f : A \to \mathbf{C}$. Allora sono equivalenti:*
(I) f è olomorfa in A;
(II) siano $u := Re(f)$, $v := Im(f)$, pensate come funzioni da A aperto in $\mathbf{R}^2$ a $\mathbf{R}$. Allora u e v appartengono a $C^1(A)$ e valgono le seguenti condizioni (di **Cauchy-Riemann***):*

$$D_x u = D_y v, \tag{2.1.2}$$

$$D_y u = -D_x v. \tag{2.1.3}$$

Dimostrazione Sia $f : A \to \mathbf{C}$ olomorfa. Sia poi $z \in A$. Se $z = (x, y)$ ($x + iy$ in notazione algebrica), si ha

$$f'(z) = \lim_{h \to 0, h \in \mathbf{R}} \frac{f(z+h) - f(z)}{h} =$$
$$= \lim_{h \to 0, h \in \mathbf{R}} (\frac{u(x+h,y) - u(x,y)}{h} + i\frac{v(x+h,y) - v(x,y)}{h}).$$

Dal lemma 2.1.1 segue l'esistenza di $D_x u(x,y)$ e di $D_x v(x,y)$. Si ha inoltre

$$D_x u(x,y) = Re(f'(z)), D_x v(x,y) = Im(f'(z)). \tag{2.1.4}$$

Si ha poi

$$f'(z) = \lim_{h \to 0, h \in \mathbf{R}} \frac{f(z+ih) - f(z)}{ih} =$$
$$= \lim_{h \to 0, h \in \mathbf{R}} (\frac{v(x,y+h) - v(x,y)}{h} - i\frac{u(x,y+h) - u(x,y)}{h}).$$

Dal lemma 2.1.1 segue l'esistenza di $D_y u(x,y)$ e di $D_y v(x,y)$. Si ha inoltre

$$D_y u(x,y) = -Im(f'(z)), D_y v(x,y) = Re(f'(z)). \tag{2.1.5}$$

Confrontando (2.1.4) con (2.1.5), si ottiene che (I) implica (II).

Ammettiamo invece che valga (II). Poniamo

$$\alpha := D_x u(z) = D_y v(z),$$
$$\beta := D_x v(z) = -D_y u(z).$$

Indichiamo poi con $h = (h_1, h_2) = h_1 + ih_2$ il generico elemento di $\mathbf{C}$. Consideriamo gli sviluppi di Taylor al primo ordine di u e v con punto iniziale z: se $z + h \in A$, si ha

$$u(z+h) = u(x+h_1, y+h_2) = u(z) + \alpha h_1 - \beta h_2 + r(h),$$

$$v(z+h) = v(x+h_1, y+h_2) = v(z) + \beta h_1 + \alpha h_2 + s(h),$$

con

$$\lim_{h\to 0} \frac{r(h)}{|h|} = \lim_{h\to 0} \frac{s(h)}{|h|} = 0.$$

Segue allora

$$\frac{f(z+h)-f(z)}{h} = \frac{u(z+h)-u(z)+i(v(z+h)-v(z))}{h}$$

$$= \frac{\alpha h_1 - \beta h_2 + r(h) + i(\beta h_1 + \alpha h_2 + s(h))}{h} =$$

$$= \frac{(\alpha + i\beta)(h_1 + ih_2) + r(h) + is(h)}{h}$$

$$= \alpha + i\beta + \frac{r(h) + is(h)}{h}.$$

Da

$$|\frac{r(h)+is(h)}{h}| \leq \frac{|r(h)|}{|h|} + \frac{|s(h)|}{|h|} \to 0 (h \to 0)$$

segue allora l'esistenza di $f'(z)$ e l'identità

$$f'(z) = \alpha + i\beta.$$

□

Osservazione 2.1.1 Sia $f : A \to \mathbf{C}$ olomorfa, con A aperto connesso per archi in $\mathbf{C}$. Dalle condizioni di Cauchy Riemann segue immediatamente che i due campi vettoriali

$$U(x,y) := (u(x,y), -v(x,y)) \tag{2.1.6}$$

e

$$V(x,y) := (v(x,y), u(x,y)) \tag{2.1.7}$$

sono chiusi in A.

Viceversa, se U e V, definiti in (2.1.6) e (2.1.7), sono di classe C^1 e chiusi, $f := u + iv$ è olomorfa.

Concludiamo questa sezione con alcuni esempi importanti di funzioni olomorfe.

Esempio 2.1.3 (Funzione esponenziale complessa) Sia $z \in \mathbf{C}$, $z = x + iy$, con x e y reali. Poniamo

$$e^z := e^x e^{iy} = e^x(\cos(y) + i\sin(y)). \tag{2.1.8}$$

Le seguenti fondamentali proprietà sono di facile verifica. Ne lasciamo la dimostrazione al lettore (vedi l'esercizio 2.1.2).

$$\begin{array}{lll} (I) & \forall z, v \in \mathbf{C} & e^z e^v = e^{z+v} \\ (II) & \forall z \in \mathbf{C} & |e^z| = e^{Re(z)} > 0; \\ (III) & \forall z \in \mathbf{C} & (e^z)^{-1} = e^{-z}. \end{array} \tag{2.1.9}$$

Posto $f : \mathbf{C} \to \mathbf{C}$, $f(z) = e^z$, è facile vedere che f è olomorfa, in quanto soddisfa le condizioni di Cauchy-Riemann. Infatti, se $u := Re(f)$, $v := Im(f)$, si ha, dato $z = (x, y)$,

$$u(x, y) = e^x \cos(y),$$

$$v(x, y) = e^x \sin(y),$$

da cui

$$D_x u(x, y) = e^x \cos(y) = D_y v(x, y),$$

$$D_y u(x, y) = -e^x \sin(y) = -D_x v(x, y).$$

Inoltre, si ha

$$f'(z) = D_x u(z) + iD_x v(z) = e^x \cos(y) + ie^x \sin(y) = f(z).$$

Esempio 2.1.4 (Funzioni logaritmo) Sia A un aperto connesso per archi in $\mathbf{C} \setminus \{0\}$. La necessità di escludere 0 viene dalla proprietà (II) della funzione esponenziale. Una funzione logaritmo è una funzione $f : A \to \mathbf{C}$ continua, tale che $e^{f(z)} = z\ \forall z \in A$. Si potrebbe dimostrare che, se f è una funzione logaritmo, f è olomorfa e $f'(z) = \frac{1}{z}\ \forall z \in A$. Noi ci limitiamo a considerarne una particolare.

Sia $A := \mathbf{C} \setminus \{x \in R : x \leq 0\}$. Poniamo

$$f : A \to \mathbf{C},$$

$$f(z) = \ln(|z|) + iArg(z),$$

ove $Arg(z)$ è l'unico elemento di $arg(z)$ maggiore di $-\pi$ e minore di π. Dall'esercizio 2.1.3 si ha $e^{f(z)} = z\ \forall z \in A$. Se $\theta := Arg(z)$, deve essere allora $\cos(\theta) = \frac{Re(z)}{|z|}$, $\sin(\theta) = \frac{Im(z)}{|z|}$, $\theta \in]-\pi, \pi[$. Consideriamo il caso particolare $Re(z) > 0$. Allora $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ e vale

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} = \frac{Im(z)}{Re(z)}.$$

Se dunque $z = x + iy$, con $x > 0$, si ha

$$f(z) = \ln(\sqrt{x^2 + y^2}) + i \arctan(\frac{y}{x}).$$

Se dunque $u := Re(f)$ e $v := Im(f)$, si ha, per $x > 0$,

$$D_x u(x, y) = \frac{x}{x^2 + y^2} = D_y v(x, y),$$

$$D_y u(x, y) = \frac{y}{x^2 + y^2} = -D_x v(x, y),$$

da cui segue che f soddisfa le condizioni di Cauchy-Riemann, almeno in $\{z \in \mathbf{C} : Re(z) > 0\}$. Inoltre,

$$f'(z) = D_x u(z) + iD_x v(z) = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2} = \frac{1}{x + iy} = \frac{1}{z}.$$

Esempio 2.1.5 Siano $A \subseteq \mathbf{C} \setminus \{0\}$ aperto e connesso per archi, g una funzione logaritmo in A. Per $\alpha \in \mathbf{R}$, poniamo

$$f_\alpha(z) := e^{\alpha g(z)}.$$

Osserviamo che f_α è una possibile versione di z^α. Teniamo però conto del fatto che la sua definizione dipende dalla scelta della funzione logaritmo.

Per gli esempi 2.1.3 e 2.1.4 e il teorema 2.1.2, f_α è olomorfa e si ha, $\forall z \in A$,

$$f'_\alpha(z) =$$

$$= \alpha\frac{e^{\alpha g(z)}}{z} = \alpha\frac{e^{\alpha g(z)}}{e^{g(z)}} = \alpha e^{(\alpha-1)g(z)}$$

$$= \alpha f_{\alpha-1}(z).$$

Si diceva che in generale f_α dipende dalla scelta della funzione logaritmo g. Supponiamo però $\alpha \in \mathbf{Z}$. Se g_1 e g_2 sono funzioni logaritmo su A, si ha che $g_1(z) - g_2(z) = 2k\pi i \ \forall z \in A$, per un certo $k \in \mathbf{Z}$. Ne segue che

$$\frac{e^{\alpha g_1(z)}}{e^{\alpha g_1 2(z)}} = e^{\alpha(g_1(z)-g_2(z))} = e^{2k\alpha\pi i} = 1,$$

perché $k\alpha \in \mathbf{Z}$. Quindi, nel caso di $\alpha \in \mathbf{Z}$, f_α non dipende dalla funzione logaritmo scelta.

Esempio 2.1.6 Estendiamo le funzioni trigonometriche seno e coseno a $\mathbf{C}$. In base alle formule di Eulero, è naturale porre, dato $z \in \mathbf{C}$:

$$\sin(z) := \frac{e^{iz} - e^{-iz}}{2i}, \tag{2.1.10}$$

$$\sin(z) := \frac{e^{iz} + e^{-iz}}{2}. \tag{2.1.11}$$

È immediato verificare che le funzioni sin e cos sono olomorfe su tutto $\mathbf{C}$. Per alcune delle loro proprietà si veda anche l'esercizio 2.1.5.

Esercizio 2.1.1 Siano A un aperto connesso per archi in $\mathbf{C}$, $f, g : A \to \mathbf{C}$ olomorfe. Dimostrare che:

(I) se $Re(f)$ è costante, allora f è costante;

(II) se $Im(f)$ è costante, allora f è costante;

(III) se $f'(z) = g'(z) = 0 \ \forall z \in A$, allora $f - g$ è costante.

(Sugg.: utilizzare le condizioni di Cauchy-Riemann e noti risultati sui campi vettoriali)

Esercizio 2.1.2 Verificare la validità delle proprietà della funzione esponenziale complessa enunciate nell'esempio 2.1.3.

Esercizio 2.1.3 Sia $v \in \mathbf{C} \setminus \{0\}$. Dimostrare che l'equazione

$$e^z = v \tag{2.1.12}$$

ha infinite soluzioni in $\mathbf{C}$. Si tratta di tutti e soli i numeri complessi della forma

$$z = \ln(|v|) + i\theta, \tag{2.1.13}$$

con $\theta \in arg(v)$, ove $\arg(v) := \{\theta \in \mathbf{R} : e^{i\theta} = \frac{v}{|v|}\}$.

Osservare che $e^z = 1$ se e solo se $z = 2k\pi i$ per qualche $k \in \mathbf{Z}$.

Esercizio 2.1.4 Siano A un aperto connesso per archi contenuto in $\mathbf{C}\setminus\{0\}$, $f, g : A \to \mathbf{C}$ due funzioni logaritmo. Verificare che esiste $k \in \mathbf{Z}$ tale che

$$f(z) - g(z) = 2k\pi i$$

$\forall z \in A$.(Sugg.: usare i risultati degli esercizi 2.1.1 (III) e 2.1.3.)

Esercizio 2.1.5 Verificare che $\forall z, v \in \mathbf{C}$,

(I)

$$\sin'(z) = \cos(z), \cos'(z) = -\sin(z); \tag{2.1.14}$$

(II)

$$\sin^2(z) + \cos^2(z) = 1; \tag{2.1.15}$$

(III)

$$\sin(-z) = -\sin(z), \cos(-z) = \cos(z); \tag{2.1.16}$$

(IV)

$$\sin(v + z) = \sin(v)\cos(z) + \cos(v)\sin(z); \tag{2.1.17}$$

(V)

$$\cos(v + z) = \cos(v)\cos(z) - \sin(v)\sin(z); \tag{2.1.18}$$

(VI) $\sin(z) = 0$ se e solo se $z = k\pi$ per qualche $k \in \mathbf{Z}$;

(VII) $\cos(z) = 0$ se e solo se $z = k\frac{\pi}{2}$ per qualche $k \in \mathbf{Z}$.

Esercizio 2.1.6 Determinare

(I) $\{z \in \mathbf{C} : \sin(z) = 2\}$;

(II) $\{z \in \mathbf{C} : \cos(z) = 3\}$.

Il fatto che si trovino delle soluzioni implica che, al contrario di ciò che avviene in $\mathbf{R}$, in $\mathbf{C}$ sin e cos non sono a valori in $[-1, 1]$!

2.2 Integrali complessi

Molte proprietà importanti delle funzioni olomorfe sono espresse in termini di integrali complessi. Raccogliamo qui i principali fatti su questi integrali che utilizzeremo.

Definizione 2.2.1 *Sia $J = [a, b]$ un intervallo chiuso e limitato in* $\mathbf{R}$*. Una funzione $\alpha : [a, b] \to \mathbf{C}$ continua si chiama* **cammino continuo** *in* $\mathbf{C}$*. Se α è di classe C^1 (vedi la successiva osservazione 2.2.1, diremo che α è un* **cammino di classe** C^1*. Diremo poi che α è un* **cammino** C^1 **a tratti** *se è un cammino continuo ed esiste una scomposizione $\{a = a_0 < .... < a_k = b\}$ di $[a, b]$ tale che, per $j = 1, ..., k$, $\alpha_{[a_{j-1},a_j]}$ è di classe C^1.*

Se α è un cammino continuo la sua immagine si chiama **sostegno** *del cammino e sarà indicato con la scrittura $sost(\alpha)$.*

Se $\alpha : [a, b] \to \mathbf{C}$ è un cammino continuo tale che $\alpha(a) = \alpha(b)$, si dice che α è un cammino **chiuso**.

Osservazione 2.2.1 Con riferimento alla definizione 2.2.1, ricordiamo che $\mathbf{C} = \mathbf{R}^2$. Dunque, se $\alpha(t) := (\alpha_1(t), \alpha_2(t))$ $(t \in [a, b])$, dire che α è di classe C^1 equivale a dire che α_1 e α_2, che sono la sua parte reale e la sua parte immaginaria, sono di classe C^1. Inoltre, $\forall t \in [a, b]$, si ha

$$\alpha'(t) = (\alpha_1'(t), \alpha_2'(t)) = \alpha_1'(t) + i\alpha_2'(t). \tag{2.2.1}$$

Naturalmente, si definisce sempre $\alpha'(t)$ come $\lim\limits_{h\to 0} \frac{\alpha(t+h)-\alpha(t)}{h}$.

Esempio 2.2.1 Siano $z^0 \in C$, $r > 0$, $\alpha : [0, 2\pi] \to \mathbf{C}$, $\alpha(t) = z^0 + re^{it}$. α è un cammino che ha per sostegno la circonferenza di centro z^0 e raggio r. Si ha $\alpha(t) = Re(z^0) + r\cos(t) + i(Im(z^0) + r\sin(t))$. Perciò α è di classe C^1. Inoltre, $\forall t \in [0, 2\pi]$,

$$\alpha'(t) = -r\sin(t) + ir\cos(t) = ire^{it}.$$

Veniamo ora alla definizione degli integrali complessi.

Definizione 2.2.2 *Siano $\alpha : [a, b] \to \mathbf{C}$, un cammino di classe C^1, $f : \alpha([a, b]) \to \mathbf{C}$. Se $t \to f(\alpha(t))\alpha'(t)$ è sommabile in $[a, b]$, diremo che esiste l'integrale complesso*

$$\int_\alpha f(z)dz := \int_{[a,b]} f(\alpha(t))\alpha'(t)dt.$$

Esempio 2.2.2 Siano $z^0 \in \mathbf{C}$, $r > 0$, $\alpha : [0, 2\pi] \to \mathbf{C}$, $\alpha(t) := z^0 + re^{it}$. Sia $n \in \mathbf{Z}$. Calcoliamo $\int_\alpha (z - z^0)^n dz$. Per definizione,

$$\int_\alpha (z - z^0)^n dz$$

$$= \int_{[0,2\pi]} (z^0 + re^{it} - z^0)^n rie^{it} dt = r^{n+1} i \int_0^{2\pi} e^{i(n+1)t} dt$$

Se $n \neq -1$, si ha

$$\int_0^{2\pi} e^{i(n+1)t} dt = \int_0^{2\pi} \cos((n+1)t) dt + i \int_0^{2\pi} \sin((n+1)t) dt =$$

$$= [\frac{\sin((n+1)t)}{n+1}]_{t=0}^{t=2\pi} + i[-\frac{\cos((n+1)t)}{n+1}]_{t=0}^{t=2\pi} = 0.$$

Se invece $n = -1$, si ha

$$\int_0^{2\pi} e^{i(n+1)t} dt = \int_0^{2\pi} 1 dt = 2\pi.$$

In conclusione,

$$\int_\alpha (z - z^0)^n dz = \begin{cases} 0 & \text{se} \quad n \neq -1, \\ 2\pi i & \text{se} \quad n = -1. \end{cases}$$

È conveniente estendere la definizione di integrale complesso al caso di α C^1 a tratti:

Definizione 2.2.3 *Sia $\alpha : [a, b] \to \mathbf{C}$ un cammino C^1 a tratti, con $a = a_0 < a_1 < ... < a_k = b$ e $\alpha_{[a_0,a_1]}$, ..., $\alpha_{[a_{k-1},a_k]}$ di classe C^1. Indichiamo con α^j $(1 \leq j \leq k)$ la restrizione di α a $[t_{j-1}, t_j]$. Se $f : \alpha([a, b]) \to \mathbf{C}$ é tale che sono definiti tutti gli integrali $\int_{\alpha^1} f(z)dz$, ..., $\int_{\alpha^k} f(z)dz$, si pone*

$$\int_\alpha f(z)dz = \int_{\alpha^1} f(z)dz + ... + \int_{\alpha^k} f(z)dz.$$

Si potrebbe verificare che questa definizione è *indipendente dal modo in cui si decompone $[a, b]$.*

Il seguente risultato è di facile verifica (vedi l'esercizio 2.2.1).

Teorema 2.2.1 *Siano* $\alpha : [a,b] \to \mathbf{C}$ *un cammino* C^1 *a tratti,* $f, g : \alpha([a,b]) \to \mathbf{C}$ *tali che sono definiti* $\int_\alpha f(z)dz$ *e* $\int_\alpha g(z)dz$*,* r *e* s *numeri complessi. Allora:*

(I) è *definito* $\int_\alpha [rf(z) + sg(z)]dz$ *e coincide con* $r\int_\alpha f(z)dz + s\int_\alpha g(z)dz$*;*

(II) sia $c \in]a,b[$*. Indichiamo con* α^1 *e* α^2 *le restrizioni di* α *a (rispettivamente)* $[a,c]$ *e* $[c,b]$*. Allora sono definiti* $\int_{\alpha^1} f(z)dz$ *e* $\int_{\alpha^2} f(z)dz$ *e vale*

$$\int_\alpha f(z)dz = \int_{\alpha^1} f(z)dz + \int_{\alpha^2} f(z)dz.$$

Vediamo ora come gli integrali complessi sono sensibili ai cambiamenti di parametro. Cominciamo con una definizione:

Definizione 2.2.4 *Siano* $\alpha : [a,b] \to \mathbf{C}$ *e* $\beta : [c,d] \to \mathbf{C}$ *due cammini continui. Diremo che sono* **equivalenti** *se esiste* $u : [c,d] \to [a,b]$*, di classe* C^1*, con* $u'(t) \neq 0$ $\forall t \in [c,d]$ *e suriettiva su* $[a,b]$*, tale che*

$$\beta(t) = \alpha(u(t)) \qquad \forall t \in [c,d].$$

Diremo poi che α *e* β *sono* **positivamente equivalenti** *se* $u'(t) > 0$ $\forall t \in [c,d]$*.*

Osservazione 2.2.2 La definizione 2.2.4 è ben nota nell'ambito della teoria delle curve. Si può vedere che la relazione definita è effettivamente di equivalenza. L'ipotesi che $u'(t) \neq 0$ $\forall t \in [c,d]$ implica che $u'(t) > 0$ $\forall t \in [c,d]$, oppure $u'(t) < 0$ $\forall t \in [c,d]$ e quindi che u è strettamente monotona. Nel primo caso, il passaggio da α a β conserva il verso di percorrenza, nel secondo lo inverte.

Teorema 2.2.2 *Siano* α *e* β *cammini* C^1 *a tratti ed equivalenti. Sia* $f : \alpha([a,b]) = \beta([c,d]) \to \mathbf{C}$ *(vedi l'esercizio 2.2.2), tale che* è *definito* $\int_\alpha f(z)dz$*. Allora*

(I) è *definito anche* $\int_\beta f(z)dz$*.*

Sia poi u *l'applicazione da* $[c,d]$ *ad* $[a,b]$ *di classe* C^1 *descritta nella definizione 2.2.4. Allora:*

(II) se $u'(t) > 0$ $\forall t \in [c,d]$*, si ha*

$$\int_\beta f(z)dz = \int_\alpha f(z)dz;$$

(III) se $u'(t) < 0$ $\forall t \in [c,d]$*, si ha*

$$\int_\beta f(z)dz = -\int_\alpha f(z)dz.$$

Dimostrazione Supponiamo, per semplicità , che α e β siano di classe C^1.

Per il teorema di cambiamento di variabile per gli integrali (teorema 0.3.7) si ha

$$\int_\alpha f(z)dz =$$

$$= \int_{[a,b]} f(\alpha(t))\alpha'(t)dt = \int_{[c,d]} f(\alpha(u(s)))\alpha'(u(s))|u'(s)|ds$$

Nel caso (II), l'ultimo integrale scritto coincide con

$$\int_{[c,d]} f(\beta(s))\alpha'(u(s))u'(s)ds = \int_{[c,d]} f(\beta(s))\beta'(s)ds =$$

$$= \int_\beta f(z)dz.$$

Nel caso (III)

$$\int_{[c,d]} f(\alpha(u(s)))\alpha'(u(s))|u'(s)|ds =$$

$$-\int_{[c,d]} f(\beta(s))\alpha'(u(s))u'(s)ds = -\int_\beta f(z)dz.$$

□

Ricordiamo che, se α è un cammino continuo, $\alpha : [a,b] \to \mathbf{C}$, si definisce la sua lunghezza $l(\alpha)$ come

$$l(\alpha) := \sup\{\sum_{j=1}^{k} |\alpha(t_j) - \alpha(t_{j-1})| : a = t_0 < t_1 < ... < t_{k-1} < t_k = b\}. \tag{2.2.2}$$

È ben noto che, se α è di classe C^1, si ha

$$l(\alpha) = \int_a^b |\alpha'(t)|dt. \tag{2.2.3}$$

È assai utile la seguente stima:

Teorema 2.2.3 *Siano* $\alpha : [a,b] \to \mathbf{C}$ C^1 *a tratti,* $f : \alpha([a,b]) \to \mathbf{C}$ *limitata, tale che* è *definito* $\int_\alpha f(z)dz$. *Allora*

$$|\int_\alpha f(z)dz| \leq \sup|f| \cdot l(\alpha).$$

Dimostrazione Limitandosi al caso di α di classe C^1, si ha

$$|\int_\alpha f(z)dz|$$

$$= |\int_{[a,b]} f(\alpha(t))\alpha'(t)dt| \leq \int_{[a,b]} |f(\alpha(t))||\alpha'(t)|dt$$

(per il teorema 0.2.5 (IV))

$$\leq \sup|f| \cdot \int_{[a,b]} |\alpha'(t)|dt =$$

$$= \sup|f| \cdot l(\alpha).$$

□

Ricordiamo ora che, se A è un aperto connesso per archi in $\mathbf{R}^2$, $f : A \to \mathbf{R}^2$ un campo vettoriale e $\alpha : [a,b] \to \mathbf{R}^2$ è un cammino di classe C^1 con sostegno in A, è definito l'integrale curvilineo di seconda specie

$$\int f \cdot d\alpha := \int_{[a,b]} f(\alpha(t)) \cdot \alpha'(t)dt, \qquad (2.2.4)$$

ove $\cdot$ è il prodotto scalare standard in $\mathbf{R}^2$.

Siano $f(z) = (u(z), v(z)) = u(z) + iv(z)$ e $\alpha(t) = (\alpha_1(t), \alpha_2(t)) = \alpha_1(t) + i\alpha_2(t)$. Allora, per quanto riguarda l'integrale complesso $\int_\alpha f(z)dz$, si ha

$$\int_\alpha f(z)dz =$$

$$= \int_{[a,b]} f(\alpha(t))\alpha'(t)dt = \int_{[a,b]} [u(\alpha(t)) + iv(\alpha(t))][\alpha_1'(t) + i\alpha_2'(t)]dt =$$

$$= \int_{[a,b]} [u(\alpha(t))\alpha_1'(t) - v(\alpha(t))\alpha_2'(t)]dt + i\int_{[a,b]} [v(\alpha(t))\alpha_1'(t) + u(\alpha(t))\alpha_2'(t)]dt$$

$$= \int U \cdot d\alpha + i\int V \cdot d\alpha,$$

con $U(z) := (u(z), -v(z))$ e $V(z) := (v(z), u(z))$. Ricordiamo (osservazione 2.1.1) che, se f è olomorfa, allora i campi vettoriali U e V sono chiusi.

Introduciamo ora alcune importanti definizioni e risultati su cammini chiusi e campi vettoriali.

Definizione 2.2.5 *Siano A un aperto in $\mathbf{C}$, $\alpha, \beta : [a,b] \to \mathbf{C}$ due cammini continui e chiusi con sostegno in A. Diremo che α e β son $A-$***omotopi** *se esiste $F : [0,1] \times [a,b] \to A$ tale che:*

(I) F è continua;

(II) $F(0,t) = \alpha(t)$, $F(1,t) = \beta(t)$ $\forall t \in [a,b]$;

(III) $F(s,a) = F(s,b)$ $\forall s \in [0,1]$.

Un'applicazione F con le proprietà $(I) - (III)$ si chiama un' **omotopia** *tra α e β.*

Osservazione 2.2.3 Dire che α e β sono $A-$omotopi significa dire che è possibile deformare con continuità α sino a trasformarlo in β senza mai uscire da A. La deformazione deve essere tale che, istante per istante, il cammino resta sempre chiuso (quest'ultima richiesta è espressa rigorosamente dalla proprietà (III)).

Esempio 2.2.3 Siano $z^0 \in \mathbf{C}$, $A := \mathbf{C} \setminus \{z^0\}$, $0 < r_0 < r_1$, $\alpha^j : [0,2\pi] \to \mathbf{C}$, $\alpha^j(t) := z^0 + r_j e^{it}$ $(j \in \{0,1\})$. α^0 e α^1 sono $A-$omotopi. Basta porre:

$$\begin{cases} F : [0,1] \times [0,2\pi] \to A, \\ F(s,t) = z^0 + [r_0 + s(r_1 - r_0)]e^{it}, \quad (s,t) \in [0,1] \times [0,2\pi]. \end{cases}$$

Esempio 2.2.4 Siano $A := \mathbf{C}$, $\alpha : [a,b] \to \mathbf{C}$ un cammino chiuso, $\beta : [a,b] \to \mathbf{C}$, tale che $\beta(t) = z^0$ $\forall t \in [a,b]$, con $z^0 \in \mathbf{C}$. Diremo che β è un cammino **puntuale**. α e β sono A omotopi. Per vederlo, basta considerare

$$\begin{cases} F : [0,1] \times [a,b] \to A, \\ F(s,t) = (1-s)\alpha(t) + sz^0, \quad (s,t) \in [0,1] \times [a,b]. \end{cases}$$

Dunque, ogni cammino chiuso è $\mathbf{C}-$omotopo a un cammino puntuale. Se A è un aperto in $\mathbf{C}$, diremo che A è **semplicemente connesso** se ogni cammino chiuso con sostegno in A è $A-$omotopo a un cammino puntuale.

Si potrebbe dimostrare il seguente importante teorema:

Teorema 2.2.4 *Siano A un aperto connesso per archi in $\mathbf{R}^2$, $F : A \to \mathbf{R}^2$ un campo vettoriale chiuso, α e β cammini chiusi con sostegno in A e $A-$omotopi. Allora*

$$\int F \cdot d\alpha = \int F \cdot d\beta.$$

Il teorema 2.2.4 ha alcune importanti conseguenze per gli integrali di funzioni olomorfe:

Corollario 2.2.1 *Siano A un aperto connesso per archi in C, α e β cammini chiusi con sostegno in A e $A-$omotopi, $f : A \to \mathbf{C}$ olomorfa. Allora*

$$\int_\alpha f(z)dz = \int_\beta f(z)dz.$$

Dimostrazione Per quanto visto,

$$\int_\alpha f(z)dz = \int U \cdot d\alpha + i \int V \cdot d\alpha,$$

con U e V campi vettoriali chiusi. Il risultato segue allora immediatamente dal teorema 2.2.4. □

Corollario 2.2.2 *Siano A un aperto semplicemente connesso in $\mathbf{C}$, α un cammino chiuso con sostegno in A, $f : A \to \mathbf{C}$ olomorfa. Allora*

$$\int_\alpha f(z)dz = 0.$$

Dimostrazione Basta applicare il precedente corollario 2.2.1, prendendo come β un cammino puntuale.□

Concludiamo la sezione con una formula classica.

Nel seguito useremo la seguente notazione: per $z^0 \in \mathbf{C}$, $r > 0$, porremo

$$\begin{cases} C_r : [0, 2\pi] \to \mathbf{C}, \\ C_r(t) = z^0 + re^{it}, t \in [0, 2\pi]. \end{cases} \tag{2.2.5}$$

Teorema 2.2.5 *(Formula integrale di Cauchy) Siano A un aperto in $\mathbf{C}$, $z^0 \in A$, $r > 0$ tale che $\{z \in \mathbf{C} : |z - z^0| < r\} \subseteq A$, $f : A \to \mathbf{C}$ olomorfa. Sia poi α un cammino chiuso di classe C^1 a tratti $A \setminus \{z^0\}-$ omotopo a $C_r(z_0)$. Allora*

$$\int_\alpha \frac{f(z)}{z - z^0}dz = 2\pi i f(z^0).$$

Dimostrazione Poiché $z \to \frac{f(z)}{z-z^0}$ è olomorfa in $A \setminus \{z^0\}$, per il corollario 2.2.1,

$$\int_\alpha \frac{f(z)}{z - z^0}dz = \int_{C_r(z^0)} \frac{f(z)}{z - z^0}dz.$$

A sua volta, $C_r(z^0)$ è $A \setminus \{z^0\}-$omotopo a $C_\rho(z^0)$ per ogni $\rho \in]0, r]$. Ne segue che

$$\int_{C_r(z^0)} \frac{f(z)}{z - z^0}dz = \lim_{\rho \to 0} \int_{C_\rho(z^0)} \frac{f(z)}{z - z^0}dz$$

$$= \lim_{\rho \to 0} \int_0^{2\pi} f(z^0 + \rho e^{it}) i dt = 2\pi i f(z^0) + \lim_{\rho \to 0} \int_0^{2\pi} [f(z^0 + \rho e^{it}) - f(z^0)] i dt.$$

Dal teorema 2.2.3 otteniamo

$$| \int_0^{2\pi} [f(z^0 + \rho e^{it}) - f(z^0)] i dt| \leq \max_{t \in [0,2\pi]} |f(z^0 + \rho e^{it}) - f(z^0)| 2\pi \to 0 (\rho \to 0)$$

perché f è continua in z^0. Di qui la conclusione. □

Osservazione 2.2.4 Due cammini $A-$omotopi hanno per definizione lo stesso dominio. Utilizzando allora il teorema 2.2.2, può essere utile generalizzare leggermente il teorema 2.2.5, richiedendo soltanto che α sia positivamente equivalente a un cammino chiuso C^1 a tratti $A \setminus \{z^0\}-$ omotopo a $C_r(z^0)$. Questa versione un po' più generale non richiede più che α abbia come dominio $[0, 2\pi]$.

Per esempio, sia (per fissare le idee) $f : \mathbf{C} \to \mathbf{C}$ olomorfa. Consideriamo il cammino α C^1 a tratti seguente:

$$\alpha : [0, 4] \to \mathbf{C},$$

$$\alpha(t) = \begin{cases} 1 + t(i-1) & \text{se} \quad t \in [0, 1], \\ i + (t-1)(-1-i) & \text{se} \quad t \in [1, 2], \\ -1 + (t-2)(-i+i) & \text{se} \quad t \in [2, 3], \\ -i + (t-3)(1+i) & \text{se} \quad t \in [3, 4]. \end{cases}$$

α percorre una volta in senso antiorario la frontiera del quadrato di vertici 1, i, -1, $-i$. α è positivamente equivalente a $\beta : [0, 2\pi] \to \mathbf{C}$, $\beta(s) = \alpha(\frac{2s}{\pi})$. È intuitivamente chiaro che β è a sua volta $\mathbf{C} \setminus \{0\}-$omotopo a $C_1(0)$. Dunque si ha

$$f(0) = \frac{1}{2\pi i} \int_\alpha \frac{f(z)}{z} dz.$$

Utilizzando poi ancora i teoremi 2.2.2 e 2.2.1(II), possiamo anche scrivere:

$$\int_\alpha \frac{f(z)}{z} dz = \int_0^1 \frac{f(1 + t(i-1))}{1 + t(i-1)} (i-1) dt + \int_0^1 \frac{f(i + t(-1-i))}{i + t(-1-i)} (-1-i) dt$$

$$+ \int_0^1 \frac{f(-1 + t(-i+1))}{-1 + t(-i+1)} (-i+1) dt + \int_0^1 \frac{f(-i + t(1+i))}{-1 + t(1+i)} (1+i) dt.$$

Esercizio 2.2.1 Dimostrare il teorema 2.2.1.

Esercizio 2.2.2 Verificare che cammini equivalenti hanno lo stesso sostegno.

2.3 Funzioni analitiche

Premettiamo all'argomento principale di questa sezione alcuni risultati di base sulle serie a coefficienti complessi, che estendono note proprietà già viste per le serie a coefficienti reali.

Sia $(a_n)_{n\in\mathbf{N}}$ una successione di numeri complessi. Possiamo costruire a partire da essa una nuova successione, detta successione delle somme parziali o serie associata ad $(a_n)_{n\in\mathbf{N}}$, ponendo, per $n \in \mathbf{N}$,

$$s_n := a_1 + ... + a_n. \tag{2.3.1}$$

Indicheremo la serie con la notazione $\sum_{n=1}^{\infty} a_n$. Se la successione $(s_n)_{n\in\mathbf{N}}$ (vale a dire, la serie) ammette limite in $\mathbf{C}$ per $n \to +\infty$, diremo che la serie è convergente e chiameremo il limite somma della serie.

Se $a_n = x_n + iy_n$, con x_n e y_n reali per ogni $n \in \mathbf{N}$, si ha per ogni n

$$s_n = \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} x_k + i \sum_{k=1}^{n} y_k.$$

Poiché la convergenza di una successione a termini complessi equivale alla convergenza delle due successioni a termini reali delle parti reali e delle parti immaginarie, la serie $\sum_{n=1}^{\infty} a_n$ è convergente in $\mathbf{C}$ se e solo se le due serie $\sum_{n=1}^{\infty} x_n$ e $\sum_{n=1}^{\infty} y_n$ sono convergenti in $\mathbf{R}$.

Alle serie a termini in $\mathbf{C}$ è estendibile la nozione di assoluta convergenza.

Definizione 2.3.1 *Sia $a_n \in \mathbf{C}$ $\forall n \in \mathbf{N}$. Allora si dice che la serie $\sum_{n=1}^{\infty} a_n$ è assolutamente convergente se la serie $\sum_{n=1}^{\infty} |a_n|$ è convergente in $\mathbf{R}$.*

Nel prossimo teorema presentiamo l'estensione a $\mathbf{C}$ di alcune ben note proprietà delle serie a valori in $\mathbf{R}$.

Teorema 2.3.1 *Sia $(a_n)_{n\in\mathbf{N}}$ una serie a valori in $\mathbf{C}$, con $a_n = x_n + iy_n$ (x_n e y_n reali) $\forall n \in \mathbf{N}$. Allora*

(I) se la serie $\sum_{n=1}^{\infty} a_n$ è convergente, necessariamente si ha $\lim_{n\to+\infty} a_n = 0$;

(II) se la serie $\sum_{n=1}^{\infty} a_n$ è assolutamente convergente, è anche convergente.

Dimostrazione (I) Se la serie $\sum_{n=1}^{\infty} a_n$ è convergente, sono convergenti anche $\sum_{n=1}^{\infty} x_n$ e $\sum_{n=1}^{\infty} y_n$. Applicando allora il risultato corrispondente per serie a termini reali, si ha $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} y_n = 0$. Dunque

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty}(x_n + iy_n) = 0.$$

(II) Sia ha, $\forall n \in \mathbf{N}$,

$$|x_n| = |Re(a_n)| \leq |a_n|, |y_n| = |Im(a_n)| \leq |a_n|.$$

Ricordando allora ben note proprietà delle serie a termini reali non negativi, possiamo dire che, se la serie $\sum_{n=1}^{\infty} a_n$ è assolutamente convergente,lo sono anche $\sum_{n=1}^{\infty} x_n$ e $\sum_{n=1}^{\infty} y_n$. Dunque le serie $\sum_{n=1}^{\infty} x_n$ e $\sum_{n=1}^{\infty} y_n$ sono convergenti, da cui la conclusione. $\Box$

Veniamo ora alle definizioni di serie di potenze e di funzione analitica.

Definizione 2.3.2 *Una* **serie di potenze** è *una serie della forma*

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n,$$

dipendente dal parametro complesso z. Qui z_0 e a_n ($\forall n \in \mathbf{N}_0$) sono numeri complessi fissati.

Definizione 2.3.3 *Siano A un aperto in $\mathbf{C}$ e $f : A \to \mathbf{C}$. Diremo che f* è **analitica** *in A se $\forall z_0 \in A$ esistono $r > 0$ e una serie di potenze del tipo $\sum_{n=0}^{\infty} a_n > (z - z_0)^n$ tali che:*
a) $B(z_0, r) = \{z \in \mathbf{C} : |z - z_0| < r\} \subseteq A$;
b) la serie $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ converge per ogni $z \in B(z_0, r)$;
c) $\forall z \in B(z_0, r)$ si ha $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$.

Vediamo ora alcune proprietà delle serie di potenze. Introduciamo preliminarmente la nozione di raggio di convergenza.

Definizione 2.3.4 *Sia $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ una serie di potenze. Il suo* **raggio di convergenza** è *il*

$$\sup\{|z - z_0| : \sum_{n=0}^{\infty} a_n(z - z_0)^n \text{è convergente}\}.$$

Osservazione 2.3.1 Ovviamente, il raggio di convergenza di una serie di potenze è un elemento di $[0, +\infty]$. In particolare, dire che vale $+\infty$ equivale a dire che $\forall r > 0$ esiste $z \in \mathbf{C}$ tale che $|z-z_0| > r$ e la serie $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ è convergente.

La proprietà principale delle serie di potenze è espressa dal seguente

Teorema 2.3.2 *(Lemma di Abel) Sia* $z_1 \in \mathbf{C}$ *tale che la serie* $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ *è convergente. Allora* $\forall z \in \mathbf{C}$ *tale che* $|z - z_0| < |z_1 - z_0|$ *la serie* $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ *è assolutamente convergente.*

Dimostrazione Sia $z \in \mathbf{C}$ tale che $|z - z_0| < |z_1 - z_0|$ (si osservi che l'esistenza di z implica $|z_1 - z_0| > 0$). Per i teoremi 2.3.1(I) e 1.3.1 si ha che esiste $M \geq 0$ tale che

$$|a_n||z_1 - z_0|^n \leq M \qquad \forall n \in \mathbf{N}_0.$$

Ne segue che

$$|a_n(z - z_0)^n| = |a_n||z_1 - z_0|^n (\frac{|z - z_0|}{|z_1 - z_0|})^n \leq M(\frac{|z - z_0|}{|z_1 - z_0|})^n.$$

Poiché $\frac{|z-z_0|}{|z_1-z_0|} < 1$, la serie $\sum_{n=0}^{\infty} M(\frac{|z-z_0|}{|z_1-z_0|})^n$ è convergente. Di qui segue subito la conclusione. □

Dal lemma di Abel segue subito il seguente

Teorema 2.3.3 *Sia* $\rho \in [0, +\infty]$ *il raggio di convergenza della serie di potenze* $\sum_{n=0}^{\infty} a_n(z - z_0)^n$*. Allora:*

(I) se $\rho = 0$ *la serie converge solo per* $z = z_0$*;*

(II) se $0 < \rho < +\infty$*, la serie converge assolutamente se* $|z - z_0| < \rho$*, non converge se* $|z - z_0| > \rho$*;*

(III) se $\rho = +\infty$*, la serie converge assolutamente per ogni* $z \in \mathbf{C}$*.*

Dimostrazione Ci limitiamo al caso (II). Gli altri casi si possono trattare analogamente, con qualche semplificazione.

Sia $|z - z_0| < \rho$. Allora esiste $z_1 \in \mathbf{C}$ tale che $|z - z_0| < |z_1 - z_0|$ e la serie $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ è convergente (altrimenti avremmo $\rho \leq |z - z_0|$). Allora, per il lemma di Abel, la serie $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ è assolutamente convergente. Se invece $|z - z_0| > \rho$, la conclusione segue dalla definizione di raggio di convergenza. □

Per uso futuro enunciamo il seguente

Corollario 2.3.1 *Sia* $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ *una serie di potenze di raggio di convergenza* $\rho \in [0, +\infty]$ *e sia* $z_1 \in \mathbf{C}$*. Allora:*

(I) se $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ *è assolutamente convergente, si ha* $|z_1 - z_0| \leq \rho$*;*

(II) se $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ *non è assolutamente convergente, si ha* $|z_1 - z_0| \geq \rho$*.*

Dimostrazione Vedi l'esercizio 2.3.1. □

Esempio 2.3.1 Consideriamo la serie di potenze $\sum_{n=0}^{\infty} n!z^n$. Se $z \in \mathbf{R}^+$, segue subito dal criterio del rapporto che la serie non converge. Dal corollario 2.3.1 si ha che il raggio di convergenza è 0.

Esempio 2.3.2 Consideriamo la serie di potenze $\sum_{n=0}^{\infty} z^n$. Se $z \in [0, 1[$, la serie è convergente, mentre non è convergente se $z \in [1, +\infty]$. Di conseguenza, il raggio di convergenza può essere solo 1. Ragionando come nel caso di $z \in \mathbf{R}$, si vede che, se $|z| < 1$, la somma della serie vale $(1 - z)^{-1}$.

Esempio 2.3.3 Consideriamo la serie di potenze $\sum_{n=0}^{\infty} \frac{z^n}{n!}$. Dal criterio del rapporto segue subito che la serie è assolutamente convergente per ogni $z \in \mathbf{C}$. Dunque, in questo caso il raggio di convergenza vale $+\infty$.

Osservazione 2.3.2 Se il raggio di convergenza della serie di potenze $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ vale $\rho \in]0, +\infty[$ e $|z - z_0| = \rho$, il teorema 2.3.3 non ci dice se la serie converge. Vediamo allora alcuni esempi.

Cominciamo con la serie considerata nell'esempio 2.3.2, con raggio di convergenza 1. Se $|z| = 1$, si ha $|z^n| = 1$ $\forall n \in \mathbf{N}_0$. Di conseguenza, non è soddisfatta la condizione $\lim_{n \to +\infty} z^n = 0$, necessaria per la convergenza (vedi il teorema 2.3.1(I)).

Consideriamo invece la serie di potenze $\sum_{n=1}^{\infty} \frac{z^n}{n}$. Se $z = 1$, la serie non converge. Di conseguenza, il raggio di convergenza ρ è non superiore a 1. D'altra parte, se $z \in [0, 1[$, la serie converge per il criterio del rapporto. Concludiamo allora che il raggio di convergenza è ancora 1. Si osservi che, per il criterio di Leibniz, la serie converge per $z = -1$. Di fatto, si potrebbe dimostrare che si ha la convergenza per ogni $z \in \mathbf{C}$ con valore assoluto uguale a 1 e $z \neq 1$.

Consideriamo invece la serie di potenze $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$. Il raggio di convergenza è sempre 1 (lo si vede osservando che, se $r \in \mathbf{R}^+$, si ha la convergenza se $r \in [0, 1]$, non si ha la convergenza se $r > 1$). In generale, se $|z| = 1$, si ha $|\frac{z^n}{n^2}| = \frac{1}{n^2}$. Dunque, si ha la convergenza assoluta per ogni $z \in \mathbf{C}$ tale che $|z| = 1$.

Consideriamo la serie di potenze $\sum_{n=0}^{\infty} a_n(z - z_0)^n$. Derivando formalmente termine a termine si ottiene $\sum_{n=1}^{\infty} na_n(z - z_0)^{n-1}$, o anche, ponendo $n - 1 = m$, $\sum_{m=0}^{\infty}(m + 1)a_{m+1}(z - z_0)^m$. Chiameremo la serie di potenze $\sum_{n=0}^{\infty}(n + 1)a_{n+1}(z - z_0)^n$ **serie derivata**. Ci chiediamo cosa si può dire sulla convergenza della serie derivata. Vale il seguente

Teorema 2.3.4 *La serie derivata di una serie di potenze ha lo stesso raggio di convergenza della serie di partenza.*

Dimostrazione parziale Consideriamo la serie di potenze $\sum_{n=0}^{\infty} a_n(z - z_0)^n$. Indichiamo con ρ il suo raggio di convergenza, con ρ' il raggio di convergenza della serie derivata. Ci limitiamo a far vedere che $\rho \leq \rho'$. Ciò è ovvio se $\rho = 0$. Supponiamo dunque $\rho \in]0, +\infty]$. Sia allora $|z - z_0| < \rho$. Facciamo vedere che la serie $\sum_{n=0}^{\infty}(n + 1)a_{n+1}(z - z_0)^n$ è assolutamente convergente. Fissiamo $z_1 \in \mathbf{C}$ tale che $|z - z_0| < |z_1 - z_0| < \rho$. Allora la serie $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ è convergente. Ripetendo un ragionamento già svolto, otteniamo che esiste $M \geq 0$ tale che $|a_n(z_1 - z_0)^n| \leq M \ \forall n \in \mathbf{N}_0$. Segue allora che, $\forall n \in \mathbf{N}_0$,

$$|(n + 1)a_{n+1}(z - z_0)^n| = |a_{n+1}(z_1 - z_0)^{n+1}||\frac{n + 1}{z - z_0}||\frac{z - z_0}{z_1 - z_0}|^n \leq$$

$$\leq \frac{M}{|z_1 - z_0|}|(n + 1)|\frac{z - z_0}{z_1 - z_0}|^n.$$

La serie $\sum_{n=0}^{\infty} \frac{M}{|z_1-z_0|}|(n + 1)|\frac{z-z_0}{z_1-z_0}|^n$ è convergente, tenendo conto del fatto che $|\frac{z-z_0}{z_1-z_0}| < 1$, come conseguenza del criterio del rapporto.

Dunque, la serie derivata è assolutamente convergente ogni volta che $|z - z_0| < \rho$. Ne segue che $\rho \leq \rho'$. Infatti, se valesse $\rho' < \rho$, preso z tale che $\rho' < |z - z_0| < \rho$, si avrebbe che la serie derivata non converge in corrispondenza di z, in contraddizione con quanto ora visto. □

Il teorema 2.3.4 rende plausibile il seguente risultato, che ci limitiamo a enunciare.

Teorema 2.3.5 *Sia $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ una serie di potenze di raggio di convergenza $\rho > 0$. Poniamo $f : B(z_0, \rho) \to \mathbf{C}$ ($f : \mathbf{C} \to \mathbf{C}$ se $\rho = +\infty$), $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$. Allora f è olomorfa. Inoltre, per ogni $z \in B(z_0, \rho)$ ($z \in \mathbf{C}$ se $\rho = +\infty$) si ha*

$$f'(z) = \sum_{n=1}^{\infty} na_n(z - z_0)^{n-1} = \sum_{n=0}^{\infty}(n + 1)a_{n+1}(z - z_0)^n.$$

Corollario 2.3.2 *Siano $A \subseteq \mathbf{C}$ aperto, $f : A \to \mathbf{C}$ anilitica. Allora:*

(I) f è olomorfa, possiede derivate complesse di ogni ordine e le derivate sono tutte olomorfe;

(II) se, per un certo $r > 0$, $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ in $B(z_0, r)$, allora

$$a_n = \frac{f^{(n)}(z_0)}{n!} \qquad \forall n \in \mathbf{N}_0.$$

Dimostrazione La dimostrazione segue quasi immediatamente dal teorema 2.3.5. Infatti, se, per un certo $r > 0$, $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ in $B(z_0, r)$, il raggio di convergenza di $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ è almeno uguale a r. Segue allora dal teorema 2.3.5 che f è olomorfa in $B(z_0, r)$ e vale

$$f'(z) = \sum_{n=1}^{\infty} na_n(z - z_0)^{n-1} = \sum_{n=0}^{\infty}(n + 1)a_{n+1}(z - z_0)^n \forall z \in B(z_0, r). \quad (2.3.2)$$

Applicando adesso ancora i teoremi 2.3.4 e 2.3.5 alla serie derivata, ricaviamo che possiamo derivare f' in $B(z_0, r)$ e ottenere che $\forall z \in B(z_0, r)$ si ha

$$f''(z) = \sum_{n=1}^{\infty}(n+1)na_{n+1}(z-z_0)^{n-1} = \sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}(z-z_0)^n \quad (2.3.3)$$

e così via.

Quanto a (II), si ha, innanzi tutto, $a_0 = f(z_0)$ e, da (2.3.2) e (2.3.3),

$$f'(z_0) = a_1, f''(z_0) = 2a_2.$$

Derivando ulteriormente, si ottiene (II) in generale. □

Vale anche l'inverso del corollario 2.3.2:

Teorema 2.3.6 *Siano $A \subseteq \mathbf{C}$ aperto, $f : A \to \mathbf{C}$ olomorfa. Allora f è analitica in A.*

Dimostrazione parziale Siano $z_0 \in A$ e $r > 0$ tali che $\{z \in \mathbf{C} : |z - z_0| \leq r\} \subseteq A$. Sia poi $z \in \mathbf{C}$, tale che $|z - z_0| < r$. Allora è chiaro che $C_r(z_0)$ è $A \setminus \{z\}-$omotopo a $C_\rho(z)$ per $\rho > 0$ sufficientemente piccolo. Segue allora dalla formula integrale di Cauchy che

$$f(z) = \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(v)}{v - z} dv. \quad (2.3.4)$$

Se $|v - z_0| = r$, si ha

$$\frac{1}{v - z} = \frac{1}{(v - z_0) - (z - z_0)} = \frac{1}{v - z_0} \frac{1}{1 - \frac{z-z_0}{v-z_0}} =$$

$$= \frac{1}{v - z_0} \sum_{n=0}^{\infty} (\frac{z - z_0}{v - z_0})^n = \sum_{n=0}^{\infty} (v - z_0)^{-n-1}(z - z_0)^n,$$

ove abbiamo usato il fatto che $|\frac{z-z_0}{v-z_0}| < 1$. Sostituendo allora in (2.3.4), e ammettendo (andrebbe dimostrato) che si possa portare la serie fuori dall'integrale, otteniamo

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n, \tag{2.3.5}$$

con

$$a_n = \frac{1}{2\pi i} \int_{C_r(z_0)} f(v)(v - z_0)^{-n-1} dv. \tag{2.3.6}$$

□

Osservazione 2.3.3 Dal corollario 2.3.2 sappiamo già che a_n (definito in (2.3.6) coincide con $\frac{f^{(n)}(z_0)}{n!}$. Esaminando la dimostrazione del teorema 2.3.6 si ricava anche che la serie di potenze $\sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$ ha raggio di convergenza non inferiore a r per ogni $r > 0$ tale che $\{z \in \mathbf{C} : |z - z_0| \leq r\} \subseteq A$ e che, per r cosi̇̀ fatto, se $|z - z_0| < r$, vale

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n. \tag{2.3.7}$$

Osservazione 2.3.4 In base ai teoremi 2.3.2 e 2.3.6 le classi delle funzioni analitiche e olomorfe coincidono. Poiché una funzione analitica è dotata di derivate di ogni ordine e queste derivate sono tutte analitiche, ricaviamo il risultato (a prima vista sorprendente) che ogni funzione olomorfa è in realtà dotata di derivate complesse di ogni ordine e le derivate sono tutte olomorfe. Questo fenomeno non ha alcuna corrispondenza nel caso delle funzioni di una variabile reale. Sarebbe come dire che ogni funzione di classe C^1 è automaticamente di classe C^{∞}!

Esempio 2.3.4 Sia $f : \mathbf{C} \to \mathbf{C}$, $f(z) = e^z$. Sappiamo dall'esempio 2.1.3 che f è olomorfa. Dunque, in base al teorema 2.3.6, f è analitica in $\mathbf{C}$. Ricordiamo che $f'(z) = f(z)$, $\forall z \in \mathbf{C}$. Dunque $f^{(n)}(z) = e^z$ $\forall n \in \mathbf{N}_0$, $\forall z \in \mathbf{C}$. In particolare, $f^{(n)}(0) = 1$ $\forall n \in \mathbf{N}_0$. Otteniamo perciò , tenendo conto che il dominio di f è $\mathbf{C}$ e applicando l'osservazione 2.3.3, che

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \qquad \forall z \in \mathbf{C}. \tag{2.3.8}$$

Dalle formule (2.1.10) e (2.1.11) segue poi che, per ogni $z \in \mathbf{C}$,

$$\cos(z) =$$
$$= \frac{e^{iz} + e^{-iz}}{2} = \frac{1}{2} \sum_{n=0}^{\infty} \frac{(iz)^n + (-iz)^n}{n!}.$$

Se $n \in \mathbf{N}_0$,
$$(iz)^n + (-iz)^n = [1 + (-1)^n] i^n z^n.$$

Dunque, $(iz)^n + (-iz)^n = 0$ se n è dispari. Se invece $n = 2k$, con $k \in \mathbf{N}_0$, si ha
$$(iz)^n + (-iz)^n = 2i^{2k} z^{2k} = 2(-1)^k z^{2k}.$$

Segue che
$$\frac{1}{2} \sum_{n=0}^{\infty} \frac{(iz)^n + (-iz)^n}{n!} = \frac{1}{2} \sum_{k=0}^{\infty} 2(-1)^k \frac{z^{2k}}{(2k)!},$$

da cui la classica formula

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} \qquad \forall z \in \mathbf{C}. \tag{2.3.9}$$

Analogamente si prova che

$$\sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^{2k+1} z^{2k+1}}{(2k+1)!} \qquad \forall z \in \mathbf{C}. \tag{2.3.10}$$

Si veda in proposito l'esercizio 2.3.2.

Esercizio 2.3.1 Dimostrare il corollario 2.3.1.

Esercizio 2.3.2 Dimostrare la formula (2.3.10.

Esercizio 2.3.3 Dimostrare che

$$\cosh(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \qquad \forall x \in \mathbf{R}.$$

Esercizio 2.3.4 Dimostrare che

$$\sinh(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \qquad \forall x \in \mathbf{R}.$$

Esercizio 2.3.5 Sia log la funzione logaritmo considerata nell'esempio 2.1.4 di dominio $\mathbf{C}\setminus\{x \in \mathbf{R} : x \leq 0\}$. Dimostrare che, se $|z-1| < 1$ vale la formula

$$\log(z) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{(z-1)^n}{n}.$$

(Sugg: osservare che $f^{(n)}(z) = (-1)^{n-1}\frac{(n-1)!}{z^n}$ $(n \in \mathbf{N})$.

2.4 Singolarità isolate e sviluppi di Laurent

Le funzioni olomorfe ammettono uno sviluppo anche in qualche intorno di una loro singolarità isolata. Per chiarire bene il risultato che ci interessa, cominciamo col precisare che cosa intendiamo per singolarità isolata.

Definizione 2.4.1 *Siano A un aperto di $\mathbf{C}$, $f : A \to \mathbf{C}$ olomorfa, $z_0 \in \mathbf{C}$. Diremo che f ha una singolarità isolata in z_0 se $z_0 \notin A$, ma esiste $r > 0$ tale che $B(z_0, r) \setminus \{z_0\} \subseteq A$.*

Sia ora $a_n \in \mathbf{C}$ per ogni $n \in \mathbf{Z}$. Se le due serie $\sum_{n=0}^{\infty} a_n$ e $\sum_{n=1}^{\infty} a_{-n}$ sono convergenti poniamo

$$\sum_{n=-\infty}^{+\infty} a_n := \sum_{n=0}^{+\infty} a_n + \sum_{n=1}^{+\infty} a_{-n}. \tag{2.4.1}$$

Il risultato che ci interessa è il seguente:

Teorema 2.4.1 *Siano A un aperto di $\mathbf{C}$, $f : A \to \mathbf{C}$ olomorfa, z_0 una singolarità isolata di f. Allora esistono univocamente determinate due serie di potenze $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ e $\sum_{n=1}^{\infty} a_{-n}v^n$ tali che:*

(I) se $r > 0$ e $B(z_0, r) \setminus \{z_0\} \subseteq A$, la serie $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ ha raggio di convergenza almeno uguale a r;

(II) $\sum_{n=1}^{\infty} a_{-n}v^n$ ha raggio di convergenza $+\infty$;

(III) se $r > 0$ è tale che $B(z_0, r) \setminus \{z_0\} \subseteq A$, $\forall z \in B(z_0, r) \setminus \{z_0\}$ si ha

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n(z-z_0)^n. \tag{2.4.2}$$

Dimostrazione parziale Siano $r > 0$ tale che $B(z_0, r) \setminus \{z_0\} \subseteq A$ e $z \in B(z_0, r) \setminus \{z_0\}$. Fissiamo r_1 e r_2 positivi e tali che

$$r_1 < |z - z_0| < r_2 < r.$$

Per fissare le idee, supponiamo $Im(z) > Im(z_0)$. Sia $0 < \epsilon < Arg(z)$ (vedi l'esempio 2.1.4). Sia α un cammino C^1 a tratti che percorre una sola volta in successione:

$a)$ il segmento di estremi $z_0 + r_1 e^{i\epsilon}$ e $z_0 + r_2 e^{i\epsilon}$;

$b)$ l'arco di circonferenza di raggio r_2 e centro z_0 di estremi $z_0 + r_2 e^{i\epsilon}$ e $z_0 + r_2$, muovendosi in senso antiorario sulla circonferenza stessa;

$c)$ il segmento di estremi $z_0 + r_2$ e $z_0 + r_1$;

$d)$ l'arco di circonferenza di raggio r_1 e centro z_0 di estremi $z_0 + r_1$ e $z_0 + r_1 e^{i\epsilon}$, muovendosi in senso orario sulla circonferenza stessa;

Chiaramente, questo arco è $A-$omotopo a $C_r(z)$, per ogni $r > 0$ sufficientemente piccolo. Allora, per la formula integrale di Cauchy,

$$f(z) = \frac{1}{2\pi i} \int_\alpha \frac{f(v)}{v - z} dv.$$

Mandando ϵ a 0, i tratti $a)$ e $c)$ tendono ad annullarsi, il tratto $b)$ tende a coincidere con $C_{r_2}(z_0)$, il tratto $d)$ tende a coincidere con la circonferenza di centro z_0 e raggio r_1 percorsa in senso orario. Queste argomentazioni rendono plausibile, mandando ϵ a 0, la formula

$$f(z) = \frac{1}{2\pi i} \int_{C_{r_2}(z_0)} \frac{f(v)}{v - z} dv - \frac{1}{2\pi i} \int_{C_{r_1}(z_0)} \frac{f(v)}{v - z} dv. \tag{2.4.3}$$

Il primo integrale nel secondo membro di (2.4.3) può essere trattato come l'integrale (2.3.4): ponendo, per $n \in \mathbf{N}_0$,

$$a_n = \frac{1}{2\pi i} \int_{C_{r_2}(z_0)} f(v)(v - z_0)^{-n-1} dv, \tag{2.4.4}$$

si ha

$$\frac{1}{2\pi i} \int_{C_{r_2}(z_0)} \frac{f(v)}{v - z} dv = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Quanto al secondo integrale in (2.4.3), se $|v - z_0| = r_1$, si ha

$$\frac{1}{v - z} = \frac{1}{(v - z_0) - (z - z_0)} = -\frac{1}{z - z_0} \frac{1}{1 - \frac{v - z_0}{z - z_0}} =$$

$$= -\frac{1}{z - z_0} \sum_{n=0}^{\infty} (\frac{v - z_0}{z - z_0})^n = -\sum_{n=0}^{\infty} (v - z_0)^n (z - z_0)^{-n-1},$$

ove abbiamo usato il fatto che $|\frac{v-z_0}{z-z_0}| < 1$. Sostituendo allora in (2.4.3), e ammettendo (andrebbe dimostrato) che si possa portare la serie fuori dall'integrale, otteniamo

$$-\frac{1}{2\pi i}\int_{C_{r_1}(z_0)}\frac{f(v)}{v-z}dv = \sum_{n=1}^{\infty} a_{-n}(z-z_0)^{-n}, \tag{2.4.5}$$

con

$$a_{-n} = \frac{1}{2\pi i}\int_{C_{r_1}(z_0)} f(v)(v-z_0)^{n-1}dv, \tag{2.4.6}$$

per ogni $n \in \mathbf{N}$.

Non proseguiamo nei dettagli della dimostrazione. $\square$

Osservazione 2.4.1 Con riferimento alla dimostrazione (parziale) del teorema 2.4.1, possiamo dire (applicando il corollario 2.2.1) che, se $R > 0$ è tale che $B(z_0, R) \setminus \{z_0\} \subseteq A$, per ogni $z \in B(z_0, R) \setminus \{z_0\}$ vale la formula

$$\sum_{n=-\infty}^{+\infty} a_n(z-z_0)^n, \tag{2.4.7}$$

con

$$a_n = \frac{1}{2\pi i}\int_{C_r(z_0)} f(v)(v-z_0)^{-n-1}dv \qquad \forall n \in \mathbf{Z}, \tag{2.4.8}$$

ove r è un elemento arbitrario di $]0, R[$.

Lo sviluppo (2.4.7) prende il nome di **sviluppo di Laurent** di f intorno al punto z_0.

Passiamo ora alla classificazione dei punti singolari isolati delle funzioni olomorfe:

Definizione 2.4.2 *Siano A un aperto in $\mathbf{C}$, $f : A \to \mathbf{C}$ olomorfa, z_0 una singolarità isolata di f. Sia (2.4.7) lo sviluppo di Laurent di f intorno a z_0. Diremo che:*

*a) z_0 è una **singolarità eliminabile** per f se $a_n = 0$ $\forall n \in \mathbf{Z}$, $n < 0$;*

*b) z_0 è una **singolarità polare** per f se $\{n \in \mathbf{Z} : n < 0$ e $a_n \neq 0\}$ è non vuoto e finito;*

*c) z_0 è una **singolarità essenziale** per f se $\{n \in \mathbf{Z} : n < 0$ e $a_n \neq 0\}$ è infinito.*

Nel caso b) si dice anche che z_0 è un **polo** *per f. Diremo che il polo è di ordine n_0 se n_0 è il massimo numero naturale n tale che $a_{-n} \neq 0$. I poli di ordine 1 sono anche detti poli semplici.*

Osservazione 2.4.2 Nel caso a) esiste $R > 0$ tale che, se $|z - z_0| < R$, si ha $f(z) = \sum_{n=0}^{+\infty} a_n(z - z_0)^n$. È perciò chiaro che, se prolunghiamo f in z_0 ponendo $f(z_0) = a_0$, otteniamo una funzione olomorfa in $A \cup \{z_0\}$. Di qui la definizione di singolarità "eliminabile".

Osservazione 2.4.3 Se $\sum_{n=-\infty}^{+\infty} a_n(z-z_0)^n$ è lo sviluppo di Laurent di f in corrispondenza della songolarità isolata z_0, La serie di potenze $\sum_{n=1}^{\infty} a_{-n}v^n$ ha raggio di convergenza infinito. Di conseguenza, la serie $\sum_{n=1}^{\infty} a_{-n}(z - z_0)^{-n}$ converge per ogni $z \in \mathbf{C} \setminus \{z_0\}$ e la sua somma fornisce una funzione olomorfa su questo insieme.

Vediamo ora qualche esempio:

Esempio 2.4.1 Sia $f : \mathbf{C} \setminus \{0\} \to \mathbf{C}$, $f(z) = \frac{e^z-1}{z}$. Per ogni $z \in \mathbf{C} \setminus \{0\}$ si ha

$$f(z) = \frac{1}{z}(\sum_{n=0}^{\infty} \frac{z^n}{n!} - 1) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!}$$

Quindi 0 è una singolarità eliminabile per f. La funzione

$$g : \mathbf{C} \to \mathbf{C},$$

$$g(z) = \begin{cases} \frac{e^z-1}{z} & \text{se} \quad z \neq 0, \\ 1 & \text{se} \quad z = 0 \end{cases}$$

è olomorfa in tutto $\mathbf{C}$.

Esempio 2.4.2 Sia $f : \mathbf{C} \setminus \{0\} \to \mathbf{C}$, $f(z) = \frac{e^z}{z^2}$. Per ogni $z \in \mathbf{C} \setminus \{0\}$ si ha

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{n-2}}{n!} = \sum_{n=-2}^{\infty} \frac{z^n}{(n+2)!}.$$

Si ha perciò $a_n = 0$ se $n < -2$, ma $a_{-2} = a_{-1} = 1$. Perciò 0 è un polo di ordine 2 per f.

Esempio 2.4.3 Sia $f : \mathbf{C} \setminus \{0\} \to \mathbf{C}$, $f(z) = e^{\frac{1}{z}}$. Per ogni $z \in \mathbf{C} \setminus \{0\}$ si ha

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{-n}}{n!} = \sum_{n=-\infty}^{0} \frac{z^n}{(-n)!}.$$

Si ha perciò $a_n = \frac{1}{(-n)!}$ per ogni $n \in \mathbf{Z}$, $n \leq 0$. Dunque 0 è una singolarità essenziale per f.

2.5 Il teorema dei residui

In questa sezione presenteremo il così detto teorema dei residui, strumento fondamentale per il calcolo di integrali complessi, che ha molte applicazioni a svariati integrali reali.

Cominciamo con la definizione di residuo:

Definizione 2.5.1 *Siano A un aperto in $\mathbf{C}$, $f : A \to \mathbf{C}$ olomorfa, $z_0 \in \mathbf{C}$ una singolarità isolata per f. Chiameremo* **residuo** *di f in z_0 e indicheremo con la scrittura $Res(f, z_0)$ il coefficiente a_{-1} dello sviluppo di Laurent di f in z_0.*

Presentiamo un semplice risultato di calcolo del residuo nel caso di una singolarità polare.

Lemma 2.5.1 *Siano A un aperto in $\mathbf{C}$, $f : A \to \mathbf{C}$ olomorfa, $z_0 \in \mathbf{C}$ una singolarità isolata per f. Supponiamo che z_0 sia un polo di ordine non superiore a n ($n \in \mathbf{N}$). Allora*

$$Res(f, z_0) = \frac{1}{(n-1)!} \lim_{z \to z_0} (\frac{d}{dz})^{n-1}[(z-z_0)^n f(z)].$$

Dimostrazione parziale Ci limitiamo al caso $n = 3$. Allora esiste $r > 0$ tale che, se $0 < |z - z_0| < r$, si ha

$$f(z) = a_{-3}(z-z_0)^{-3} + a_{-2}(z-z_0)^{-2} + a_{-1}(z-z_0)^{-1} + a_0 + a_1(z-z_0) + ...,$$

da cui

$$(z-z_0)^3 f(z) = a_{-3} + a_{-2}(z-z_0) + a_{-1}(z-z_0)^2 + a_0(z-z_0)^3 + a_1(z-z_0)^4 + ...,$$

$$(\frac{d}{dz})^2[(z-z_0)^3 f](z) = 2a_{-1} + 6a_0(z-z_0) + 12a_1(z-z_0)^2 + ...,$$

e di qui la conclusione. □

All'esempio che segue premettiamo alcuni semplici risultati di utilità pratica. Cominciamo con una versione complessa del teorema dell' Hopital.

Lemma 2.5.2 *Siano $z_0 \in \mathbf{C}$, $r > 0$, $f, g : B(z_0, r) \setminus \{z_0\} \to \mathbf{C}$ olomorfe e tali che $\lim_{z \to z_0} f(z) = \lim_{z \to z_0} g(z) = 0$, $g(z) \neq 0$ $\forall z \in B(z_0, r) \setminus \{z_0\}$. Supponiamo inoltre che $g'(z) \neq 0$ $\forall z \in B(z_0, r) \setminus \{z_0\}$ ed esista in $\mathbf{C}$ $\lim_{z \to z_0} \frac{f'(z)}{g'(z)}$. Allora esiste $\lim_{z \to z_0} \frac{f(z)}{g(z)}$ e coincide con $\lim_{z \to z_0} \frac{f'(z)}{g'(z)}$.*

Dimostrazione parziale Si potrebbe verificare che i prolungamenti naturali di f e g a $B(z_0,r)$ sono funzioni olomorfe in $B(z_0,r)$. Supponiamo (chiamando ancora, per semplicità , f e g questi prolungamenti) , che si abbia $g'(z_0) \neq 0$. Allora, $\forall z \in B(z_0,r)$, $f(z) = \sum_{n=1}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n$ e $g(z) = \sum_{n=1}^{\infty} \frac{g^{(n)}(z_0)}{n!}(z-z_0)^n$. Poiché $g'(z_0) \neq 0$, si ha

$$\lim_{z \to z_0} \frac{f'(z)}{g'(z)} = \frac{f'(z_0)}{g'(z_0)}.$$

D'altra parte, per $z \neq z_0$,

$$\frac{f(z)}{g(z)} = \frac{(z-z_0)\sum_{n=0}^{\infty} \frac{f^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n}{(z-z_0)\sum_{n=0}^{\infty} \frac{g^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n}$$

$$= \frac{\sum_{n=0}^{\infty} \frac{f^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n}{\sum_{n=0}^{\infty} \frac{g^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n} \to \frac{f'(z_0)}{g'(z_0)}(z \to z_0).$$

$\square$

Il lemma 2.5.1 richiede poi una stima dell'ordine di un polo. Presentiamo un semplice risultato in questa direzione. Se $f : A \to \mathbf{C}$ è olomorfa (con A aperto in $\mathbf{C}$) e $z_0 \in A$, diremo che f ha in z_0 uno **zero di ordine** m ($m \in \mathbf{N}_0$) se $f^{(k)}(z_0) = 0$ per $k \in \mathbf{N}_0$, $k < m$, mentre $f^{(m)}(z_0) \neq 0$. Si osservi che, se $f(z_0) \neq 0$, allora f ha in z_0 uno zero di ordine 0.

Lemma 2.5.3 *Siano A un aperto di $\mathbf{C}$, $f, g : A \to \mathbf{C}$ olomorfe, $z_0 \in A$. Supponiamo che z_0 sia uno zero di ordine m per f, di ordine n per g con m e n elementi di $\mathbf{N}_0$. Sia poi $h : \{z \in A : g(z) \neq 0\} \to \mathbf{C}$ $h(z) = \frac{f(z)}{g(z)}$. Allora:*

(I) se $m \geq n$, h ha in z_0 una singolarità eliminabile;

(II) se $m < n$, h ha in z_0 un polo di ordine $n - m$.

Dimostrazione Se $r > 0$ è abbastanza piccolo, si ha, per $|z - z_0| < r$,

$$f(z) = \sum_{k=m}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k =$$

$$= (z-z_0)^m \sum_{r=0}^{\infty} \frac{f^{(r+m)}(z_0)}{(r+m)!}(z-z_0)^r := (z-z_0)^m k(z)$$

e
$$g(z) = \sum_{k=n}^{\infty} \frac{g^{(k)}(z_0)}{k!}(z-z_0)^k =$$
$$(z-z_0)^n \sum_{r=0}^{\infty} \frac{g^{(r+n)}(z_0)}{(r+n)!}(z-z_0)^r := (z-z_0)^n l(z).$$

Le funzioni k e l sono olomorfe in $\{z \in \mathbf{C} : |z-z_0| < r\}$, $k(z_0) = \frac{f^{(m)}(z_0)}{m!} \neq 0$ e $l(z_0) = \frac{g^{(n)}(z_0)}{n!} \neq 0$. Diminuendo (eventualmente) r, si può supporre $l(z) \neq 0$ se $|z-z_0| < r$. La funzione $\frac{k}{l}$ ammette un certo sviluppo $\sum_{k=0}^{\infty} c_k(z-z_0)^k$ in $B(z_0, r)$ con $c_0 \neq 0$. Dunque, se $0 < |z-z_0| < r$, si ha
$$h(z) = (z-z_0)^{m-n} \sum_{k=0}^{\infty} c_k(z-z_0)^k = \sum_{k=0}^{\infty} c_k(z-z_0)^{k+m-n} =$$
$$= \sum_{j=m-n}^{\infty} c_{j-m+n}(z-z_0)^j,$$
da cui la conclusione. □

Esempio 2.5.1 Siano $A := \{z \in \mathbf{C} : \sin(z) \neq 0\}$, $f : A \to \mathbf{C}$, $f(z) = \frac{e^z-1}{\sin^2(z)}$. Utilizzando il risultato dell'esercizio 2.1.5(VI), si ha subito che 0 è una singolarità isolata per f. Determiniamo $Res(f, 0)$. Il lemma 2.5.1 richiede una stima per eccesso dell'ordine del polo. La funzione $z \to \sin^2(z)$ ha in 0 uno zero di ordine 2, mentre la funzione $z \to e^z - 1$ ha in 0 uno zero semplice (vale a dire, di ordine 1).

Dunque, in virtù del lemma 2.5.3, f ha in 0 un polo di ordine 1 (o semplice). Allora, per il lemma 2.5.1, si ha
$$Res(f, 0) = \lim_{z\to 0} zf(z) = \lim_{z\to 0} \frac{z(e^z-1)}{\sin^2(z)}.$$

Per calcolare quest'ultimo limite, proviamo ad applicare il lemma 2.5.2. Derivando a numeratore e a denominatore, si ottiene $\frac{e^z-1+ze^z}{2\sin(z)\cos(z)}$, che è ancora della forma indeterminata $\frac{0}{0}$. Derivando ancora una volta, otteniamo $\frac{2e^z+ze^z}{2(\cos^2(z)-\sin^2(z))}$ che tende a 1 per $z \to 0$. Si ha allora $Res(f, 0) = 1$.

Al risultato principale premettiamo la nozione di **indice** di un cammino chiuso rispetto a un punto.

Definizione 2.5.2 *Siano α un cammino chiuso C^1 a tratti e $z_0 \in \mathbf{C}$, con z_0 non appartenente al sostegno di α. Chiamiamo indice di α rispetto a z_0 e indichiamo con la scrittura $ind(\alpha, z_0)$ il numero complesso*

$$ind(\alpha, z_0) := \frac{1}{2\pi i} \int_\alpha (z - z_0)^{-1} dz. \qquad (2.5.1)$$

Si potrebbero dimostrare le seguenti proprietà dell'indice:

Teorema 2.5.1 *Sia α un cammino chiuso C^1 a tratti. Allora*

(I) $\forall z \in \mathbf{C} \setminus sost(\alpha)$ $ind(\alpha, z_0) \in \mathbf{Z}$;

(II) se $A \subseteq \mathbf{C} \setminus sost(\alpha)$ e A è connesso per archi, allora $ind(\alpha, z)$ è il medesimo per ogni elemento z di A;

(III) se $A \subseteq \mathbf{C} \setminus sost(\alpha)$ e A è connesso per archi e non limitato, allora $ind(\alpha, z) = 0$ $\forall z \in A$.

Osservazione 2.5.1 Nella pratica, $ind(\alpha, z_0)$ indica quante volte α "gira intorno " a z_0, contando 1 ogni giro in senso antiorario, -1 ogni giro in senso orario.

Esempio 2.5.2 Siano $n \in \mathbf{N}$ e $\alpha : [0, 2\pi] \to \mathbf{C}$, $\alpha(t) = e^{int}$. Allora $sost(\alpha) = \{z \in \mathbf{C} : |z| = 1\}$. Sia $z_0 \in \mathbf{C}$, con $|z_0| \neq 1$. In base al teorema 2.5.1(III) applicato ad $A := \{z \in \mathbf{C} : |z| > 1\}$, si ha in tal caso $ind(\alpha, z_0) = 0$. Sia invece $|z_0| < 1$. Allora, applicando il teorema 2.5.1(II) ad $A := \{z \in \mathbf{C} : |z| < 1\}$, possiamo dire che

$$ind(\alpha, z_0) = ind(\alpha, 0) = \frac{1}{2\pi i} \int_\alpha z^{-1} dz =$$

$$= \frac{1}{2\pi i} \int_0^{2\pi} e^{-int} ine^{int} dt = n.$$

In effetti, α gira n volte in senso antiorario intorno a z_0.

Consideriamo invece $\beta : [0, 2\pi] \to \mathbf{C}$, $\beta(t) = e^{i(2\pi - t)} = e^{-it}$, che gira una volta in senso orario intorno a (ad esempio) 0. Allora si ha

$$ind(\beta, 0) = \frac{1}{2\pi i} \int_\beta z^{-1} dz =$$

$$= \frac{1}{2\pi i} \int_0^{2\pi} e^{it} (-i) e^{-it} dt = -1.$$

Ci servirà anche il seguente

Lemma 2.5.4 *Siano* A *un aperto in* $\mathbf{C}$, $f : A \to \mathbf{C}$ *olomorfa,* α *un cammino di classe* C^1 *a tratti con sostegno in* A. *Supponiamo che esista* $F : A \to \mathbf{C}$ *olomorfa tale che* $F'(z) = f(z)\ \forall z \in A$. *Allora*

$$\int_\alpha f(z)dz = 0.$$

Dimostrazione parziale Supponiamo che α sia di classe C^1. Allora, se $\alpha : [a, b] \to \mathbf{C}$,

$$\int_\alpha f(z)dz = \int_a^b f(\alpha(t))\alpha'(t)dt =$$

(applicando il teorema fondamentale del calcolo integrale)

$$= [F(\alpha(t)]_{t=a}^{t=b} = 0$$

perché $\alpha(a) = \alpha(b)$. $\square$

Veniamo ora al risultato principale di questa sezione, che fornisce (come si vedrà) un importante strumento di calcolo per gli integrali:

Teorema 2.5.2 *(dei residui) Siano* A *un aperto semplicemente connesso in* $\mathbf{C}$, z_1,...,z_n *elementi di* A, $f : A \setminus \{z_1, ..., z_n\} \to \mathbf{C}$ *olomorfa,* α *un cammino chiuso* C^1 *a tratti con sostegno in* $A \setminus \{z_1, ..., z_n\}$. *Allora*

$$\frac{1}{2\pi i}\int_\alpha f(z)dz = \sum_{i=1}^{n} Res(f, z_i)ind(\alpha, z_i). \tag{2.5.2}$$

Dimostrazione parziale Consideriamo, per $i = 1, ..., n$, lo sviluppo di Laurent di f in corrispondenza di z_i. Se questo sviluppo è $\sum_{n=-\infty}^{+\infty} a_{i,n}(z - z_i)^n$, poniamo

$$S_i(z) := \sum_{n=-\infty}^{-1} a_{i,n}(z - z_i)^n = \sum_{n=1}^{+\infty} a_{i,-n}(z - z_i)^{-n}.$$

Abbiamo già osservato (vedi l'osservazione 2.4.3) che S_i è olomorfa in $\mathbf{C} \setminus \{z_i\}$. Consideriamo ora la funzione $z \to f(z) - S_1(z) - ... - S_n(z)$, olomorfa in $A \setminus \{z_1, ..., z_n\}$. Osserviamo che ciascuno dei punti $z_1, ..., z_n$ è una singolarità eliminabile per questa funzione. Per vederlo, prendiamo (ad esempio) $i = 1$. Allora $f(z) - S_1(z) = \sum_{n=0}^{+\infty} a_{i,n}(z - z_i)^n$ per z abbastanza prossimo a z_1, mentre S_2, ..., S_n sono regolari in z_1. Allora, applicando il corollario 2.2.2, si ha

$$\frac{1}{2\pi i}\int_\alpha f(z)dz =$$

$$= \frac{1}{2\pi i}\int_\alpha (f(z) - S_1(z) - ... - S_n(z))dz + \frac{1}{2\pi i}\int_\alpha (S_1(z) + ... + S_n(z))dz$$

$$= \frac{1}{2\pi i}\int_\alpha (S_1(z) + ... + S_n(z))dz.$$

Per $i = 1, ..., n$, si ha

$$\frac{1}{2\pi i}\int_\alpha S_i(z)dz = \frac{1}{2\pi i}\int_\alpha \sum_{n=1}^{+\infty} a_{i,-n}(z - z_i)^{-n}dz$$

$$= \sum_{n=1}^{+\infty} a_{i,-n}\int_\alpha (z - z_i)^{-n}dz$$

(non giustifichiamo quest'ultimo passaggio). Consideriamo adesso, per $n \in \mathbf{N}$, l'integrale $\int_\alpha (z - z_i)^{-n}dz$. Se $n \neq 1$, si ha

$$(z - z_i)^{-n} = F'(z),$$

con $F(z) = (1 - n)^{-1}(z - z_i)^{1-n}$. Applicando allora il lemma 2.5.2, concludiamo che, se $n \neq 1$, $\int_\alpha (z - z_i)^{-n}dz = 0$. Si ha perciò che

$$\frac{1}{2\pi i}\int_\alpha S_i(z)dz = \frac{1}{2\pi i}a_{i,-1}\int_\alpha (z - z_i)^{-1}dz =$$

$$= Res(f, z_i)ind(\alpha, z_i).$$

□

Esempio 2.5.3 Sia $\alpha = C_2(2i)$. Vogliamo calcolare

$$\int_\alpha \frac{1}{z^4 - 8z^2 - 9}dz. \tag{2.5.3}$$

Gli zeri complessi del polinomio $P(z) := z^4 - 8z^2 - 9$ sono 3, -3, i e $-i$. Poniamo allora $f : \mathbf{C} \setminus \{3, -3, i, -i\} \to \mathbf{C}$, $f(z) = \frac{1}{z^4 - 8z^2 - 9}$. f è olomorfa e, applicando il teorema dei residui, possiamo scrivere che l'integrale in (2.5.3) coincide con

$$2\pi i[Res(f, 3)ind(\alpha, 3) + Res(f, -3)ind(\alpha, -3)+$$

$$+Res(f, i)ind(\alpha, i) + Res(f, -i)ind(\alpha, -i)].$$

Osserviamo ora che $|3 - 2i| = |-3 - 2i| = \sqrt{13} > 2$ e $|-i - 2i| = 3 > 2$. Dunque, per il teorema 2.5.1(III), si ha

$$ind(C_2(2i), 3) = ind(C_2(2i), -3) = ind(C_2(2i), -i) = 0.$$

Tenendo poi conto dell'osservazione 2.5.1, possiamo scrivere anche

$$ind(C_2(2i), i) = 1.$$

Dunque l'integrale in (2.5.3) coincide con $2\pi i Res(f, i)$. Per calcolare $Res(f, i)$, osserviamo che $P'(i) = -20i \neq 0$, per cui P ha in i uno zero semplice (o di molteplicità 1).

Segue dal lemma 2.5.3 che f ha un polo semplice in i. Dunque

$$Res(f, i) = \lim_{z \to i} (z - i) f(z) = \lim_{z \to i} \frac{1}{4z^3 - 16z} = \frac{i}{20}.$$

Perciò l'integrale in (2.5.3) vale $2\pi i \frac{i}{20} = -\frac{\pi}{10}$.

Esempio 2.5.4 Utilizziamo il teorema dei residui per calcolare

$$\int_{\mathbf{R}} \frac{1}{x^4 + 1} dx. \tag{2.5.4}$$

Osserviamo, innanzi tutto, che la funzione integranda è continua e quindi misurabile in $\mathbf{R}$. È inoltre sempre positiva. Dunque l'integrale esiste nel senso della definizione 0.2.4. Possiamo inoltre applicare il teorema di Beppo Levi per ottenere che (2.5.4) coincide con

$$\lim_{n \to +\infty} \int_{-n}^{n} \frac{1}{x^4 + 1} dx.$$

Infatti,

$$\int_{-n}^{n} \frac{1}{x^4 + 1} dx = \int_{\mathbf{R}} \frac{\chi_n(x)}{x^4 + 1} dx,$$

ove abbiamo indicato con χ_n la funzione caratteristica dell'intervallo $[-n, n]$. Infatti, per ogni $n \in \mathbf{N}$ e $\forall x \in \mathbf{R}$,

$$\frac{\chi_n(x)}{x^4 + 1} \leq \frac{\chi_{n+1}(x)}{x^4 + 1}$$

e

$$\lim_{n \to +\infty} \frac{\chi_n(x)}{x^4 + 1} = \frac{1}{x^4 + 1}.$$

Consideriamo adesso un cammino chiuso C^1 a tratti α_n ottenuto percorrendo in senso antiorario prima l'intervallo $[-n, n]$, poi la semicirconferenza $\{ne^{it} : t \in [0, \pi]\}$. La funzione $z \to \frac{1}{z^4+1}$ è olomorfa in $\mathbf{C} \setminus \{e^{i\frac{(2k+1)\pi}{4}} : 0 \leq k \leq 3, k \in$

$\mathbf{Z}\}$. Se $n > 1$, nessuno dei numeri complessi della forma $e^{i\frac{(2k+1)\pi}{4}}$ appartiene al sostegno di α_n. È chiaro anche che

$$ind(\alpha_n, e^{i\frac{\pi}{4}}) = ind(\alpha_n, e^{i\frac{3\pi}{4}}) = 1,$$

mentre

$$ind(\alpha_n, e^{i\frac{5\pi}{4}}) = ind(\alpha_n, e^{i\frac{7\pi}{4}}) = 0.$$

Segue allora dal teorema dei residui che

$$\int_{\alpha_n} \frac{1}{z^4+1} dz = 2\pi i [Res(f, e^{i\frac{\pi}{4}}) + Res(f, e^{i\frac{3\pi}{4}})],$$

con $f(z) = \frac{1}{z^4+1}$. Non è difficile verificare che f ha in $e^{i\frac{\pi}{4}}$ un polo di ordine 1. Applicando allora i lemmi 2.5.1 e 2.5.2 possiamo scrivere

$$Res(f, e^{i\frac{\pi}{4}}) = \lim_{z \to e^{i\frac{\pi}{4}}} \frac{z - e^{i\frac{\pi}{4}}}{z^4+1}$$

$$= \lim_{z \to e^{i\frac{\pi}{4}}} \frac{1}{4z^3} = \frac{e^{-i\frac{3\pi}{4}}}{4}.$$

Nello stesso modo

$$Res(f, e^{i\frac{3\pi}{4}}) = \lim_{z \to e^{i\frac{3\pi}{4}}} \frac{z - e^{i\frac{3\pi}{4}}}{z^4+1}$$

$$= \lim_{z \to e^{i\frac{3\pi}{4}}} \frac{1}{4z^3} = \frac{e^{-i\frac{9\pi}{4}}}{4}.$$

Dunque

$$\int_{\alpha_n} \frac{1}{z^4+1} dz = 2\pi i \left(\frac{e^{-i\frac{3\pi}{4}}}{4} + \frac{e^{-i\frac{9\pi}{4}}}{4}\right) = \frac{\pi}{\sqrt{2}}.$$

D'altra parte,

$$\int_{\alpha_n} \frac{1}{z^4+1} dz =$$

$$= \int_{-n}^{n} \frac{1}{x^4+1} dx + \int_{C_n^+(0)} \frac{1}{z^4+1} dz,$$

con

$$\begin{cases} C_n^+(0) : [0, \pi] \to \mathbf{C}, \\ C_n^+(0)(t) = ne^{it}. \end{cases} \qquad (2.5.5)$$

Se $|z| = n$ e $n > 1$, si ha, applicando il risultato dell'esercizio 1.1.2,

$$|z^4 + 1| \geq ||z^4| - 1|| = n^4 - 1 > 0,$$

per cui

$$|f(z)| \leq \frac{1}{n^4 - 1}$$

se $|z| = n > 1$. Applicando allora il teorema 2.2.3, abbiamo

$$|\int_{C_n^+(0)} \frac{1}{z^4+1} dz| \leq \frac{n\pi}{n^4 - 1} \to 0 (n \to +\infty).$$

Riepilogando, abbiamo, per $n \geq 2$,

$$\int_{-n}^{n} \frac{1}{x^4+1} dx + \int_{C_n^+(0)} \frac{1}{z^4+1} dz = \frac{\pi}{\sqrt{2}},$$

da cui, passando al limite per $n \to +\infty$,

$$\int_{\mathbf{R}} \frac{1}{x^4+1} dx = \frac{\pi}{\sqrt{2}}.$$

Esempio 2.5.5 Vogliamo calcolare

$$\int_{\mathbf{R}^+} \frac{x^\alpha}{1+x^2} dx, \tag{2.5.6}$$

con $\alpha \in \mathbf{R}$.

Qualunque sia α, se poniamo $f : \mathbf{R}^+ \to \mathbf{R}$, $f(x) = \frac{x^\alpha}{1+x^2}$, f è misurabile perché continua e non negativa. Dunque l'integrale è comunque definito. Si tratta di vedere, innanzi tutto, per quali valori di α è un numero reale. Si ha

$$\int_{\mathbf{R}^+} f(x)dx = \int_{]0,1]} f(x)dx + \int_{]1,+\infty[} f(x)dx. \tag{2.5.7}$$

La funzione $g(x) := x^\beta$ è sommabile su $]0, 1]$ se e solo se $\beta > -1$. Se $0 < x \leq 1$, si ha

$$f(x) \leq x^\alpha \leq 2f(x).$$

Dunque il primo integrale in (2.5.7) è finito se e solo se $\alpha > -1$. Inoltre, la funzione $g(x) := x^\beta$ è sommabile su $[1, +\infty[$ se e solo se $\beta < -1$. Se $x \geq 1$, si ha

$$f(x) \leq x^{\alpha-2}$$

e

$$x^{\alpha-2} = f(x)(1 + \frac{1}{x^2}) \leq 2f(x).$$

Dunque il secondo integrale in (2.5.7) è finito se e solo se $\alpha - 2 < -1$, cioè $\alpha < 1$. Concludiamo che l'integrale in (2.5.6) è finito se e solo se

$$-1 < \alpha < 1. \tag{2.5.8}$$

D'ora in poi supporremo che la condizione (2.5.8) sia soddisfatta.

Consideriamo adesso, per un cero $n \in \mathbf{N}$, un cammino chiuso C^1 a tratti α_n ottenuto percorrendo una volta in senso antiorario l'intervallo $[\frac{1}{n}, n]$, la semicirconferenza $\{z \in \mathbf{C} : |z| = n, Im(z) \geq 0\}$, l'intervallo $[-n, -\frac{1}{n}]$, la semicirconferenza $\{z \in \mathbf{C} : |z| = \frac{1}{n}, Im(z) \geq 0\}$. Estendiamo la funzione f a una funzione olomorfa in un opportuno aperto in $\mathbf{C}$. A tale scopo consideriamo la funzione logaritmo log definita su $A := \mathbf{C} \setminus \{iy : y \in \mathbf{R}, y \leq 0\}$ nel seguente modo:

$$log(z) = \ln(|z|) + i\theta, \theta \in arg(z) \cap] - \frac{\pi}{2}, \frac{3\pi}{2}[. \tag{2.5.9}$$

Poniamo poi, per ogni $\beta \in \mathbf{R}$ (ricordando l'esempio 2.1.5)

$$z^\beta := e^{\beta log(z)} \tag{2.5.10}$$

Consideriamo ora la funzione

$$\begin{cases} g : A \setminus \{i\} \to \mathbf{C}, \\ g(z) = \frac{z^\alpha}{1+z^2}. \end{cases}$$

g è olomorfa e la sua restrizione a $\mathbf{R}^+$ è f. Applicando allora il teorema dei residui, avremo, per $n \geq 2$,

$$\int_{\alpha_n} g(z)dz = 2\pi i Res(g, i).$$

Si vede facilmente che g ha in i un polo semplice. Dunque, applicando ancora il lemma 2.5.2 e ricordando l'esempio 2.1.5, si ha

$$Res(g, i) = \lim_{z \to i} (z - i)g(z)$$

$$= \lim_{z \to i} \frac{z^\alpha + (z - i)\alpha z^{\alpha-1}}{2z} = \frac{i^\alpha}{2i} = \frac{e^{\alpha log(i)}}{2i}$$

$$= \frac{e^{i\alpha\frac{\pi}{2}}}{2i},$$

da cui

$$\int_{\alpha_n} g(z)dz = \pi e^{i\alpha\frac{\pi}{2}}.$$

D'altra parte,

$$\int_{\alpha_n} g(z)dz =$$

$$= \int_{\frac{1}{n}}^{n} \frac{x^\alpha}{1+x^2}dx + \int_{C_n^+(0)} \frac{z^\alpha}{1+z^2}dz + \int_{-n}^{-\frac{1}{n}} \frac{x^\alpha}{1+x^2}dx - \int_{C_{\frac{1}{n}}^+(0)} \frac{z^\alpha}{1+z^2}dz.$$

Se $x < 0$, si ha

$$x^\alpha = e^{\alpha log(x)} = e^{\alpha(ln(-x)+i\pi)} = (-x)^\alpha e^{i\alpha\pi},$$

da cui

$$\int_{-n}^{-\frac{1}{n}} \frac{x^\alpha}{1+x^2}dx = e^{i\alpha\pi}\int_{-n}^{-\frac{1}{n}} \frac{(-x)^\alpha}{1+x^2}dx = e^{i\alpha\pi}\int_{\frac{1}{n}}^{n} \frac{x^\alpha}{1+x^2}dx.$$

Inoltre, se $|z| = R$, si ha

$$|z^\alpha| = |e^{\alpha log(z)}| = e^{Re(\alpha log(z))} = e^{\alpha ln(R)} = R^\alpha.$$

Se $R > 1$, si ha poi

$$|1+z^2| = |z^2 - (-1)| \geq ||z^2| - 1| = R^2 - 1,$$

mentre, se $R < 1$,

$$|1+z^2| = |1 - (-z^2)| \geq |1 - | - z^2|| = 1 - R^2.$$

Applicando allora il teorema 2.2.3, otteniamo

$$|\int_{C_n^+(0)} \frac{z^\alpha}{1+z^2}dz| \leq \frac{n^\alpha}{n^2-1}\pi n \to 0(n \to +\infty)$$

e

$$|\int_{C_{\frac{1}{n}}^+(0)} \frac{z^\alpha}{1+z^2}dz| \leq \frac{n^{-\alpha}}{1-n^{-2}}\pi n^{-1} \to 0(n \to +\infty),$$

tenendo conto di (2.5.8). Concludiamo, mandando n a $+\infty$, che

$$(1+e^{i\alpha\pi})\int_{\mathbf{R}^+}\frac{x^\alpha}{1+x^2}dx=\pi e^{i\alpha\frac{\pi}{2}},$$

da cui

$$\int_{\mathbf{R}^+}\frac{x^\alpha}{1+x^2}dx=\frac{\pi e^{i\alpha\frac{\pi}{2}}}{1+e^{i\alpha\pi}}=\frac{\pi}{2\cos(\frac{\alpha\pi}{2})}.$$

Vediamo infine un ultimo esempio che sarà utile nelle prossime sezioni.

Esempio 2.5.6 Dato $\xi\in\mathbf{R}$, calcoliamo

$$\int_{\mathbf{R}} e^{-x^2-ix\xi}dx. \tag{2.5.11}$$

Osserviamo, innanzi tutto, che

$$|e^{-x^2-ix\xi}|=e^{-x^2}$$

Si ha $e^{-x^2}=o(x^{-2})$ per $x\to\pm\infty$. Dunque, esiste $M\in\mathbf{R}^+$ tale che, se $|x|>M$, si ha $e^{-x^2}<x^{-2}$. Perciò ,

$$\int_{\mathbf{R}} e^{-x^2}dx\le\int_{-M}^{M} e^{-x^2}dx+\int_{\{|x|>M\}} x^{-2}dx<+\infty.$$

Dunque, in base al risultato dell'esercizio 0.2.4, la funzione integranda in (2.5.5) è sommabile per ogni $\xi\in\mathbf{R}$. Cominciamo allora col considerare il caso $\xi=0$. Poniamo, per comodità ,

$$I:=\int_{\mathbf{R}} e^{-x^2}dx.$$

Applicando il teorema 0.3.7 con $T:\mathbf{R}^+\times]0,2\pi[\to\mathbf{R}^2$, $T(\rho,\theta)=(\rho\cos(\theta),\rho\sin(\theta))$ e il teorema di Tonelli, si ha

$$\int_{\mathbf{R}^2} e^{-x^2-y^2}dxdy=\int_{\mathbf{R}^+\times]0,2\pi[} e^{-\rho^2}\rho d\rho d\theta$$

$$=2\pi\int_{\mathbf{R}^+} e^{-\rho^2}\rho d\rho=2\pi\lim_{c\to+\infty}\int_0^c e^{-\rho^2}\rho d\rho$$

$$=2\pi\lim_{c\to+\infty}\frac{1}{2}(1-e^{-c^2})=\pi.$$

Ma, ancora per il teorema di Tonelli,

$$\int_{\mathbf{R}^2} e^{-x^2-y^2} dxdy = I^2.$$

Poiché $I > 0$, deve essere $I = \sqrt{\pi}$. Consideriamo ora il caso $\xi \neq 0$. Osserviamo, innanzi tutto, che

$$\begin{aligned} x^2 + ix\xi &= (x^2 + 2x\frac{i\xi}{2} - \frac{\xi^2}{4}) + \frac{\xi^2}{4} \\ &= (x + \frac{i\xi}{2})^2 + \frac{\xi^2}{4}. \end{aligned}$$

Ne segue che

$$\int_{\mathbf{R}} e^{-x^2-ix\xi} dx = e^{-\frac{\xi^2}{4}} \int_{\mathbf{R}} e^{-(x+\frac{i\xi}{2})^2} dx.$$

Supponiamo adesso (per fissare le idee) $\xi > 0$. Consideriamo, dato $n \in \mathbf{N}$, un cammino α_n C^1 a tratti e orientato in senso orario che abbia come sostegno la frontiera del rettangolo di vertici $-n + \frac{i\xi}{2}$, $n + \frac{i\xi}{2}$, n, $-n$. In base al corollario 2.2.2, si ha

$$\int_{\alpha_n} e^{-z^2} dz = 0.$$

D'altra parte,

$$\int_{\alpha_n} e^{-z^2} dz =$$

$$= \int_{-n}^{n} e^{-(x+\frac{i\xi}{2})^2} dx - \int_{-n}^{n} e^{-x^2} dx - \int_0^{\xi} e^{-(n+iy)^2} idy + \int_0^{\xi} e^{-(-n+iy)^2} idy$$

Se $y \in [0, \xi]$, si ha

$$|e^{-(n+iy)^2}| = e^{y^2-n^2} \leq e^{\xi^2-n^2}.$$

Ne segue che

$$|\int_0^{\xi} e^{-(n+iy)^2} idy| \leq \xi e^{\xi^2-n^2} \to 0 (n \to +\infty).$$

Con un ragionamento simile, si vede che vale anche

$$\lim_{n \to +\infty} \int_0^{\xi} e^{-(-n+iy)^2} idy = 0.$$

Mandando allora n a $+\infty$, otteniamo

$$0 = \int_{\mathbf{R}} e^{-(x+\frac{i\xi}{2})^2} dx - \sqrt{\pi},$$

da cui
$$\int_{\mathbf{R}} e^{-(x+\frac{i\xi}{2})^2} dx = \sqrt{\pi}.$$
Con considerazioni analoghe, si vede che questa formula vale anche per $\xi < 0$. Possiamo allora concludere che per ogni $\xi \in \mathbf{R}$ l'integrale in (2.5.11) vale $\sqrt{\pi} e^{-\frac{\xi^2}{4}}$.

Esercizio 2.5.1 Calcolare i residui delle funzioni seguenti nel punto z_0 specificato:

(I) $\frac{\sin(z)}{z^3}$, $z_0 = 0$;

(II) $(2-z)^{-1} e^{\frac{1}{z}}$, $z_0 = 2$;

(III) $\frac{\sin(z)}{\cos(z)}$, $z_0 = \frac{\pi}{2}$;

(IV) $\frac{z^\alpha}{(z^\beta - 1)^2}$, $z_0 = 1$, con $z^\alpha = f_\alpha(z)$ e $z^\beta = f_\beta(z)$ (vedi l'esempio 2.1.5), prendendo come funzione logaritmo $g(z) = \ln(|z|) + iArg(z)$ per $Re(z) > 0$, ove $Arg(z)$ è l'argomento principale di z;

(V) $\frac{z^3}{sin^5(z)}$, $z_0 = 0$.

Esercizio 2.5.2 Calcolare

(I) $\int_0^{2\pi} \frac{1}{\cos(x)+2} dx$;

(II) $\int_0^{2\pi} \frac{1}{\cos(x)+\sin(x)+2} dx$;

(III) $\int_0^{2\pi} \frac{1}{\cos^2(x)+1} dx$;

(IV) $\int_0^{2\pi} \frac{\cos(x)}{\cos(x)+\sin(x)+2} dx$.

(Sugg.: usare le formule di Eulero e ricondurre gli integrali dati a integrali complessi su $C_1(0)$).

Esercizio 2.5.3 Calcolare

(I) $\int_{\mathbf{R}} \frac{1}{x^2+2x+2} dx$;

(II) $\int_{\mathbf{R}} \frac{1}{(x^2+1)^2} dx$;

(III) $\int_{\mathbf{R}} \frac{1}{x^6+1} dx$;

(IV) $\int_{\mathbf{R}} \frac{1}{x^8+1} dx$;

(V) $\int_{\mathbf{R}^+} \frac{x^2}{x^4+1} dx$.

Esercizio 2.5.4 Calcolare, dato $\alpha \in]-1,1[$,
$$\int_{\mathbf{R}^+} \frac{x^\alpha \ln(x)}{1+x^2} dx. \tag{2.5.12}$$

Cominciare col far vedere che la funzione integranda è sommabile su $\mathbf{R}^+$, usando il fatto che $\ln(x) = o(x^\epsilon)$ per $x \to +\infty$ e $\ln(x) = o(x^{-\epsilon})$ per $x \to 0$ per ogni $\epsilon > 0$.

Esercizio 2.5.5 Calcolare, verificandone preliminarmente l'esistenza,
(I) $\int_{\mathbf{R}^+} \frac{\ln(x)}{x^4+1} dx$;
(II) $\int_{\mathbf{R}^+} \frac{\ln^2(x)}{x^2+1} dx$;
(III) $\int_{\mathbf{R}^+} \frac{x^\alpha}{x^4+1} dx$ $(-1 < \alpha < 3)$.

2.6 Funzioni olomorfe e funzioni armoniche

Cominciamo con l'introdurre, in $\mathbf{R}^n$, $(n \in \mathbf{N})$ l'operatore di Laplace Δ:

$$\Delta := \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} = \frac{\partial^2}{\partial x_1^2} + ... \frac{\partial^2}{\partial x_n^2}. \tag{2.6.1}$$

L'operatore differenziale Δ riveste una notevole importanza nell'ambito della fisica-matematica.

Definizione 2.6.1 *Siano A un aperto in $\mathbf{R}^n$ $u \in C^2(A)$. Diremo che u è* **armonica** *in A se*

$$\Delta u(x) = 0 \qquad \forall x \in A.$$

Osservazione 2.6.1 Esempi di funzioni armoniche sono le funzioni polinomiali di grado non superiore a 1. Altri esempi sono le funzioni $u : \mathbf{R}^2 \to \mathbf{R}$, $u(x,y) = e^x \cos(y)$ e $v : \mathbf{R}^2 \to \mathbf{R}$, $v(x,y) = x^2 - y^2$, come si verifica facilmente.

Ne caso $n = 2$, le funzioni armoniche sono strettamente legate alle funzioni olomorfe:

Teorema 2.6.1 *Siano A un aperto in $\mathbf{C}$, f $A \to \mathbf{C}$ olomorfa, $u : A \to \mathbf{R}$, $u(z) = u(x,y) = Re(f(z))$. Allora u è armonica in A.*

Viceversa, siano A un aperto semplicemente connesso in $\mathbf{R}^2$ (che identifichiamo con $\mathbf{C}$), $u : A \to \mathbf{R}$ armonica. Allora esiste $f : A \to \mathbf{C}$ olomorfa tale che $u = Re(f)$.

Dimostrazione Siano $f : A \to \mathbf{C}$ olomorfa, $u = Re(f)$. Ricordiamo (vedi l'osservazione 2.3.4) che u è di classe C^∞, in quanto f ha derivate complesse di ogni ordine tutte olomorfe. Verifichiamo che $\Delta u(x,y) = 0 \ \forall (x,y) \in A$. Sia $v := Im(f)$. Allora, applicando le condizioni di Cauchy-Riemann e il teorema di Schwarz, si ha, $\forall (x,y) \in A$,

$$\Delta u(x,y) = \frac{\partial^2 u}{\partial x^2}(x,y) + \frac{\partial^2 u}{\partial y^2}(x,y)$$

$$= \frac{\partial^2 v}{\partial x \partial y}(x,y) - \frac{\partial^2 v}{\partial y \partial x}(x,y) = 0.$$

Siano ora A un aperto semplicemente connesso in $\mathbf{R}^2$ e $u : A \to \mathbf{R}$ armonica. Vogliamo costruire $f : A \to \mathbf{C}$ olomorfa tale che $u = Re(f)$. A tele scopo,in base al teorema 2.1.3, basta determinare $v \in C^1(A)$ tale che

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{2.6.2}$$

e porre

$$\begin{cases} f : A \to \mathbf{C}, \\ f(z) = u(z) + iv(z), \quad z \in A. \end{cases}$$

Consideriamo il campo vettoriale $F : A \to \mathbf{R}^2$,

$$F(x,y) := (-\frac{\partial u}{\partial y}(x,y), \frac{\partial u}{\partial x}(x,y)).$$

Poiché u è armonica, F è chiuso e quindi esatto, essendo A semplicemente connesso. Dunque, esiste $v \in C^1(A)$ tale soddisfacente (2.6.2). Con ciò , il risultato è completamente provato. $\square$

Osservazione 2.6.2 Abbiamo già osservato che la parte reale e la parte immaginaria di una funzione olomorfa sono di classe C^∞. Ora, tenuto conto del fatto che le palle aperte sono insiemi semplicemente connessi,in base al teorema 2.6.1, data $u : A \to \mathbf{R}$ armonica in A aperto di $\mathbf{R}^2$, possiamo costruire in ogni palla aperta contenuta in A una funzione olomorfa f in tale palla, la cui parte reale è u. Possiamo perciò concludere che, almeno nel caso $n = 2$, le funzioni armoniche sono di classe C^∞. Di fatto, si potrebbe dimostrare che ciò è vero in generale.

Osservazione 2.6.3 Se $f : A \to \mathbf{C}$ è olomorfa, (A aperto in $\mathbf{C}$), poiché $Im(f) = Re(-if)$, anche $Im(f)$ è armonica in A.

2.7 Principio del massimo e problema di Dirichlet per l'equazione di Laplace nella circonferenza unitaria di $\mathbf{R}^2$

In questa sezione presenteremo alcuni risultati relativi al problema di Dirichlet per l'equazione di Laplace. Questo problema può essere formulato come segue: siano A un aperto in R^n ($n \in \mathbf{N}$), $g : Fr(A) \to \mathbf{R}$ continua. Cerchiamo le eventuali funzioni $u : \overline{A} \to \mathbf{R}$, continue in $\overline{A}$ e di classe C^2 in A, tali che

$$\begin{cases} \Delta u(x) = 0, & \forall x \in A, \\ u(x') = g(x'), & \forall x' \in Fr(A). \end{cases} \tag{2.7.1}$$

Cominciamo con un primo importante risultato, il così detto **principio del massimo**.

Teorema 2.7.1 *(Principio del massimo) Siano A un aperto limitato in $\mathbf{R}^n$, $u \in C(\overline{A}) \cap C^2(A)$ a valori reali, tale che*

$$\Delta u(x) \geq 0 \qquad \forall x \in A. \tag{2.7.2}$$

Allora

$$\max_{\overline{A}} u = \max_{Fr(A)} u. \tag{2.7.3}$$

Dimostrazione Osserviamo, innanzi tutto, che, poiché A è limitato, $\overline{A}$ e $Fr(A)$ sono chiusi e limitati (vedi, per questo, gli esercizi 2.7.1 e 2.7.2). Dunque, per il teorema di Weierstrass, $\max_{\overline{A}} u$ e $\max_{Fr(A)} u$ esistono. Essendo $Fr(A) \subseteq \overline{A}$, possiamo immediatamente dire che

$$\max_{Fr(A)} u \leq \max_{\overline{A}} u. \tag{2.7.4}$$

Si tratta ora di rovesciare la disuguaglianza (2.7.4). A tale scopo, cominciamo col considerare il caso meno generale

$$\Delta u(x) > 0 \qquad \forall x \in A. \tag{2.7.5}$$

Faremo vedere che, in questo caso, u non possiede punti di massimo in A, da cui segue che i suoi punti di massimo sono necessariamente in $Fr(A)$ e perciò vale (2.7.3). Ragioniamo per assurdo, supponendo che valga (2.7.5) e che esista $x^0 \in A$, punto di massimo per u. Consideriamo la forma quadratica

$$Q(h) := \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_i x_j}(x^0) h_i h_j, \qquad h = (h_1, ..., h_n) \in \mathbf{R}^n. \tag{2.7.6}$$

Ricordiamo che Q è semidefinita negativa. Ne segue che, se $k \in \{1, ..., n\}$ e e^k è il $k-$esimo elemento della base canonica di $\mathbf{R}^n$,

$$0 \geq Q(e^k) = \frac{\partial^2 u}{\partial x_k^2}(x^0).$$

Dunque $\Delta u(x^0) \leq 0$, in contraddizione con (2.7.5).

Consideriamo ora il caso generale (2.7.2). Fissiamo $v \in C^2(\mathbf{R}^n)$, a valori reali, tale che $\Delta v(x) > 0 \ \forall x \in \mathbf{R}^n$. Una possibile scelta è (ad esempio), $v : \mathbf{R}^n \to \mathbf{R}$, $v(x_1, ..., x_n) = x_1^2$, che verifica $\Delta v(x) = 2 \ \forall x \in \mathbf{R}^n$. Dato $\epsilon > 0$, poniamo

$$u_\epsilon := u + \epsilon v. \tag{2.7.7}$$

Si ha

$$\Delta u_\epsilon(x) = \Delta u(x) + \epsilon \Delta v(x) > 0 \qquad \forall x \in A.$$

Dunque, utilizzando il caso particolare già trattato, abbiamo

$$\max_{\overline{A}} u_\epsilon = \max_{Fr(A)} u_\epsilon.$$

Questa identità implica che, qualunque sia $x \in \overline{A}$, si ha

$$u(x) + \epsilon v(x) \leq \max_{Fr(A)} (u + \epsilon v). \tag{2.7.8}$$

Se $x' \in Fr(A)$, vale

$$u(x') + \epsilon v(x') \leq \max_{Fr(A)} u + \epsilon \max_{Fr(A)} v,$$

che implica, qualunque sia $x \in \overline{A}$,

$$u(x) + \epsilon v(x) \leq \max_{Fr(A)} u + \epsilon \max_{Fr(A)} v. \tag{2.7.9}$$

Passando al limite per $\epsilon \to 0^+$ in (2.7.9), otteniamo allora che, qualunque sia $x \in \overline{A}$, si ha

$$u(x) \leq \max_{Fr(A)} u. \tag{2.7.10}$$

Da ciò segue la tesi. $\square$

Dal teorema 2.7.1 segue facilmente il seguente risultato di unicità :

Corollario 2.7.1 *Siano A un aperto limitato in $\mathbf{R}^n$, $u_0, u_1 \in C(\overline{A}) \cap C^2(A)$ a valori reali, tali che $\Delta u_0(x) = \Delta u_1(x)$ $\forall x \in A$, $u_0(x') = u_1(x')$ $\forall x' \in Fr(A)$. Allora $u_0(x) = u_1(x)$ $\forall x \in \overline{A}$.*

Dimostrazione Poniamo $u : \overline{A} \to \mathbf{R}$, $u(x) = u_0(x) - u_1(x)$. Allora $u \in C(\overline{A}) \cap C^2(A)$, è a valori reali, $\Delta u(x) = 0\,\forall x \in A$, $u(x') = 0\,\forall x' \in Fr(A)$. Per il principio del massimo, si ha

$$u(x) \leq 0 \qquad \forall x \in \overline{A}. \tag{2.7.11}$$

D'altra parte, vale anche $\Delta(-u)(x) = 0$ $\forall x \in A$, $-u(x') = 0$ $\forall x' \in Fr(A)$. Dunque, ancora per il principio del massimo, si ha $-u(x) \leq 0$ $\forall x \in \overline{A}$, cioè ,

$$u(x) \geq 0 \qquad \forall x \in \overline{A}. \tag{2.7.12}$$

Da (2.7.11) e (2.7.12) otteniamo la conclusione. □

Consideriamo adesso il problema (2.7.1) nel caso $A = \{z \in \mathbf{C} : |z| < 1\} = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < 1\}$.

Teorema 2.7.2 *Siano $A = \{z \in \mathbf{C} : |z| < 1\}$, $g \in C(Fr(A); \mathbf{R})$. Allora esiste un'unica $u \in C(\overline{A}; \mathbf{R}) \cap C^2(A)$ per cui vale (2.7.1).*

Dimostrazione parziale L'unicità segue dal corollario 2.7.1.

Per quanto riguarda l'esistenza, costruiremo euristicamente una certa soluzione del problema. Non verificheremo in tutti i dettagli che la funzione costruita è effettivamente soluzione.

Supponiamo allora che esista una soluzione u con le proprietà volute. In base al teorema 2.6.1, esiste una funzione olomorfa $f : A \to \mathbf{C}$, tale che $u = Re(f)$. Applicando il teorema 2.3.6 e l'osservazione 2.3.3, possiamo allora dire che, per ogni $z \in \mathbf{C}$, con $|z| < 1$, si ha

$$f(z) = \sum_{n=0}^{\infty} \alpha_n z^n, \tag{2.7.13}$$

con il raggio di convergenza della serie di potenze $\displaystyle\sum_{n=0}^{\infty} a_n z^n$ almeno uguale a 1. Dunque, se $r \in [0, 1[$ e $\theta \in] - \pi, \pi]$, si ha

$$\begin{array}{rcc} u(re^{i\theta}) & = & Re(f(re^{i\theta})) \\ & = & Re(\sum_{n=0}^{\infty} \alpha_n r^n e^{in\theta}) \\ & = & \sum_{n=0}^{\infty} a_n r^n \cos(n\theta) + \sum_{n=1}^{\infty} b_n r^n \sin(n\theta), \end{array} \tag{2.7.14}$$

ove avviamo posto

$$a_n := Re(\alpha_n), \qquad n \in \mathbf{N}_0, \tag{2.7.15}$$

$$b_n := -Im(\alpha_n), \qquad n \in \mathbf{N}. \tag{2.7.16}$$

Si tratta adesso di determinare i coefficienti $(a_n)_{n\in\mathbf{N}_0}$ e $(b_n)_{n\in\mathbf{N}}$. Ragionando formalmente, per $r = 1$, si ottiene:

$$\begin{array}{rcl} g(e^{i\theta}) & = & u(\cos(\theta), \sin(\theta)) \\ & = & \sum_{n=0}^{\infty} a_n \cos(n\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta) \\ & = & a_0 + \sum_{n=1}^{\infty} \frac{a_n - ib_n}{2} e^{in\theta} + \sum_{n=1}^{\infty} \frac{a_n + ib_n}{2} e^{-in\theta}. \end{array} \tag{2.7.17}$$

L'ultimo termine in (2.7.17) è lo sviluppo in serie di Fourier di $\theta \to g(e^{i\theta})$. Dunque

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{it})dt, \tag{2.7.18}$$

e, per ogni $n \in \mathbf{N}$,

$$a_n + ib_n = \frac{1}{\pi} \int_{-\pi}^{\pi} e^{int} g(e^{it})dt,$$

da cui

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(nt) g(e^{it})dt, \tag{2.7.19}$$

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin(nt) g(e^{it})dt, \tag{2.7.20}$$

Da (2.7.14), (2.7.18), (2.7.19), (2.7.20), otteniamo allora, per $r \in [0, 1[$ e $\theta \in]-\pi, \pi]$, portando la serie all'interno dell'integrale,

$$\begin{array}{rcl} u(re^{i\theta}) & = & \frac{1}{\pi} \int_{-\pi}^{\pi} \{\frac{1}{2} + \sum_{n=1}^{\infty} r^n [\cos(n\theta)\cos(nt) + \sin(n\theta)\sin(nt)] g(e^{it})dt \\ & = & \frac{1}{\pi} \int_{-\pi}^{\pi} [\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(n(\theta - t))] g(e^{it})dt \end{array} \tag{2.7.21}$$

Per $r \in [0, 1[$, $s \in \mathbf{R}$, si ha

$$\begin{array}{rcl} \frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(ns) & = & \frac{1}{2} + Re\{\sum_{n=1}^{\infty} r^n e^{ins}\} \\ & = & \frac{1}{2} + rRe\{\frac{e^{is}}{1 - re^{is}}\} \\ & = & \frac{1}{2} + r\frac{\cos(s) - r}{1 - 2r\cos(s) + r^2} \\ & = & \frac{1 - r^2}{2(1 - 2r\cos(s) + r^2)}. \end{array}$$

Abbiamo perciò ottenuto la seguente formula classica:

$$u(re^{i\theta}) = \frac{1 - r^2}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - 2r\cos(\theta - t) + r^2} g(e^{it})dt. \tag{2.7.22}$$

Ora, si potrebbe effettivamente dimostrare che la funzione $u : \overline{A} \to \mathbf{R}$, tale che

$$u(re^{i\theta}) = \begin{cases} \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(\theta-t)+r^2} g(e^{it})dt & \text{se} \quad 0 \leq r < 1, \\ g(e^{i\theta}) & \text{se} \quad r = 1 \end{cases}$$

appartiene a $C(\overline{A};\mathbf{R}) \cap C^2(A)$ ed è armonica in A. Qui noi ci limitiamo a verificare che si ha

$$\lim_{r\to 1} u(r,0) = g(1,0). \tag{2.7.23}$$

A tale scopo, osserviamo preliminarmente che, per ogni $r \in [0,1[$, si ha

$$\frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(s)+r^2} ds = 1. \tag{2.7.24}$$

Infatti, ammettendo di poter scambiare la serie con l'integrale, si ha

$$\begin{aligned} &\frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(s)+r^2} ds \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} (\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(ns))ds \\ &= \frac{1}{\pi}(\pi + \sum_{n=1}^{\infty} r^n \int_{-\pi}^{\pi} \cos(ns))ds) \\ &= 1. \end{aligned}$$

Inoltre,

$$\frac{1-r^2}{2\pi(1-2r\cos(s)+r^2)} > 0 \qquad \forall r \in [0,1[, s \in]-\pi,\pi],$$

essendo

$$1-2r\cos(s)+r^2 \geq 1-2r+r^2 = (1-r)^2.$$

Dunque,

$$\begin{aligned} |u(r,0)-g(1,0)| &= |\tfrac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \tfrac{1}{1-2r\cos(t)+r^2}(g(e^{it})-g(1,0))dt| \\ &\leq \tfrac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \tfrac{1}{1-2r\cos(t)+r^2}|g(e^{it})-g(1,0)|dt. \end{aligned} \tag{2.7.25}$$

Fissiamo $\epsilon \in \mathbf{R}^+$. Poiché g è continua, esiste $\delta \in]0,\frac{\pi}{2}[$, tale che, se $|t| < \delta$, si ha

$$|g(e^{it})-g(1,0)| < \frac{\epsilon}{2}.$$

Da (2.7.25) segue allora

$$\begin{array}{rcl}|u(r,0)-g(1,0)| & \leq & \frac{\epsilon}{2}\frac{1-r^2}{2\pi}\int_{-\delta}^{\delta}\frac{1}{1-2r\cos(t)+r^2}dt \\ & & +2\max\limits_{Fr(A)}|g|\frac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\frac{1}{1-2r\cos(t)+r^2}dt \\ & & +\int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt).\end{array} \tag{2.7.26}$$

Si ha

$$\frac{1-r^2}{2\pi}\int_{-\delta}^{\delta}\frac{1}{1-2r\cos(t)+r^2}dt \leq \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt = 1.$$

Inoltre, se $|t| \geq \delta$, si ha $\cos(t) \leq \cos(\delta) < 1$. Ne segue che

$$1-2r\cos(t)+r^2 \geq 1-2r\cos(\delta)+r^2$$

Con metodi elementari, si verifica che

$$\min_{[0,1]}(1-2r\cos(\delta)+r^2) = \sin^2(\delta) > 0.$$

Perciò

$$2\max_{Fr(A)}|g|\frac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\frac{1}{1-2r\cos(t)+r^2}dt + \int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt)$$

$$= 2\max_{Fr(A)}|g|\frac{1-r^2}{\pi}\int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt$$

$$\leq 2\max_{Fr(A)}|g|\frac{1-r^2}{\pi}\frac{\pi-\delta}{\sin^2(\delta)},$$

che tende a 0 per $r \to 1$.

Ne concludiamo che esiste $r(\epsilon) \in [0,1[$, tale che, se $r(\epsilon) < r < 1$, si ha

$$2\max_{Fr(A)}|g|\frac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\frac{1}{1-2r\cos(t)+r^2}dt + \int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt) < \frac{\epsilon}{2},$$

e quindi, se $r(\epsilon) < r < 1$,

$$|u(r,0)-g(1,0)| < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon.$$

□

Esercizio 2.7.1 Sia $A \subseteq \mathbf{R}^n$, A limitato. Provare che $\overline{A}$ e $Fr(A)$ sono limitati.

Esercizio 2.7.2 Sia $A \subseteq \mathbf{R}^n$. Provare che $Fr(A)$ è un chiuso.

2.8 Trasformazioni conformi e problema di Dirichlet per l'equazione di Laplace in $\mathbf{R}^2$

Vogliamo ora estendere il teorema 2.7.2 ad aperti più generali. Per farlo, utilizzeremo certi "cambiamenti di variabile", noti nella letteratura come "trasformazioni conformi".

Definizione 2.8.1 *Siano A e B aperti un $\mathbf{C}$. Una* **trasformazione conforme** *da B ad A* è *una funzione $\phi : B \to A$ tale che:*
(a) ϕ è *una biiezione tra B e A;*
(b) ϕ è *olomorfa;*
(c) $\phi'(z) \neq 0$ $\forall z \in B$.
Diremo che B è conformemente equivalente ad A se esiste una trasformazione conforme da B ad A.

Vale il seguente risultato, che non dimostriamo:

Teorema 2.8.1 *Siano A e B aperti un $\mathbf{C}$, ϕ una trasformazione conforme da B ad A. Allora la trasformazione inversa ϕ^{-1}* è *una trasformazione conforme da A a B.*

Osservazione 2.8.1 L'unico punto non ovvio nella dimostrazione del teorema 2.8.1 è che ϕ^{-1} è olomorfa. Supposto che valga questo fatto, la derivata di ϕ^{-1} si ricava facilmente: basta osservare che $\phi^{-1}(\phi(z)) = z$ $\forall z \in B$. Applicando allora la regola di derivazione di funzioni composte, si ottiene

$$(\phi^{-1})'(\phi(z))\phi'(z) = 1 \qquad \forall z \in B,$$

da cui

$$(\phi^{-1})'(\phi(z)) = \frac{1}{\phi'(z)} \qquad \forall z \in B, \tag{2.8.1}$$

che è un'estensione naturale della formula di derivazione di una funzione inversa vista in Analisi A.

Osservazione 2.8.2 Il termine "trasformazione conforme" vuole segnalare il fatto che queste applicazioni "conservano gli angoli tra le curve". Per attribuire un significato preciso a questa affermazione, consideriamo due cammini di classe C^1 α e β, con dominio $[-\delta, \delta]$ ($\delta > 0$), a sostegno in $\mathbf{R}^2$. Supponiamo che $\alpha(0) = \beta(0)$ e che i vettori $\alpha'(0)$ e $\beta'(0)$ siano entrambi non nulli. Allora possiamo definire come "angolo tra α e β in corrispondenza di

$t = 0$" l'angolo di misura $\theta \in [0, \pi]$ compreso tra i vettori tangenti $\alpha'(0)$ e $\beta'(0)$. Ricordando l'interpretazione geometrica del prodotto scalare in $\mathbf{R}^2$, si ha

$$\theta = \arccos(\frac{\alpha'(0) \cdot \beta'(0)}{\|\alpha'(0)\| \|\beta'(0)\|}),$$

ove abbiamo indicato con $\|.\|$ la norma euclidea in $\mathbf{R}^2$, coincidente con il valore assoluto nei complessi. Osserviamo che, se $z = (z_1, z_2)$ e $v = (v_1, v_2)$ sono elementi di $\mathbf{R}^2$, se indichiamo con zv il loro prodotto in $\mathbf{C}$, si ha

$$z \cdot v = z_1 v_1 + z_2 v_2 = Re(z\overline{v}).$$

Dunque,

$$\theta = \arccos(\frac{Re(\alpha'(0)\overline{\beta'(0)})}{|\alpha'(0)||\beta'(0)|}).$$

Siano adesso B un aperto in $\mathbf{C}$ contente i sostegni di α e β e $\phi : B \to A$ una trasformazione conforme, con A aperto in $\mathbf{C}$. Consideriamo i cammini $\phi \circ \alpha$ e $\phi \circ \beta$ e l'angolo θ' conpreso tra essi in corrispondenza di $t = 0$. Poiché $(\phi \circ \alpha)'(0) = \phi'(\alpha(0))\alpha'(0)$ e

$$(\phi \circ \beta)'(0) = \phi'(\beta(0))\beta'(0) = \phi'(\alpha(0))\beta'(0),$$

si ha

$$\begin{array}{rcl} \theta' & = & \arccos(\frac{Re(\phi'(\alpha(0))\alpha'(0)\overline{\phi'(\alpha(0))\beta'(0)})}{|\phi'(\alpha(0)|^2|\alpha'(0)||\beta'(0)|}) \\ & = & \arccos(\frac{Re(\alpha'(0)\overline{\beta'(0)})}{|\alpha'(0)||\beta'(0)|}) \\ & = & \theta. \end{array}$$

Esempio 2.8.1 Siano $B := \{z \in \mathbf{C} : |z| < 1\}$, $\phi : B \to \mathbf{C}$, $\phi(z) = \frac{1+z}{1-z}$. ϕ è olomorfa e, per ogni $z \in B$,

$$\phi'(z) = \frac{2}{(1-z)^2} \neq 0.$$

Sia $v \in \mathbf{C}$. Consideriamo l'equazione

$$\phi(z) = v, \qquad z \in B. \tag{2.8.2}$$

Se $v \neq -1$, l'equazione (2.8.2) ha in $\mathbf{C}$ l'unica soluzione $z = \frac{v-1}{v+1}$. Tale soluzione appartiene a B se e solo se

$$|v - 1| < |v + 1|,$$

equivalente a $Re(v) > 0$. Dunque ϕ è una trasformazione conforme tra B e $\{v \in \mathbf{C} : Re(v) > 0\}$.

Osservazione 2.8.3 Non è difficile verificare che la relazione di conforme equivalenza è di equivalenza tra gli aperti di $\mathbf{C}$ (esercizio 2.8.1).

A questo punto ci possiamo chiedere se, dati due aperti in $\mathbf{C}$, essi sono conformemente equivalenti. Un primo famoso risultato è il seguente classico teorema, dovuto a Riemann:

Teorema 2.8.2 *Sia A un aperto non vuoto in $\mathbf{C}$. Allora A è conformemente equivalente a $\{z \in \mathbf{C} : |z| < 1\}$ se e solo se è semplicemente connesso e non coincide con tutto $\mathbf{C}$.*

Per una dimostrazione, vedi W. Rudin "Analisi reale e complessa", (Boringhieri), capitolo 14.

Il teorema 2.8.2 ci dice che, se A è un aperto semplicemente connesso e non coincide con tutto $\mathbf{C}$, esiste una trasformazione conforme ϕ da A a $\{z \in \mathbf{C} : |z| < 1\}$, ma non ci dice nulla su un'eventuale estensione continua di ϕ a $\overline{A}$. Introducimao allora due definizioni:

Definizione 2.8.2 *Siano $E \subseteq \mathbf{C}$ e $F \subseteq \mathbf{C}$. Un* **omeomorfismo** *da E a F è una funzione $\phi : E \to F$, iniettiva e suriettiva, continua, con l'inversa ϕ^{-1} continua.*

Se esiste un omeomorfismo da E a F diremo che E è omeomorfo a F.

Osservazione 2.8.4 È facile vedere che la relazione "essere omeomorfo a" è di equivalenza tra i sottoinsiemi di $\mathbf{C}$. (esercizio 2.8.1). È inoltre immediato verificare che le trasformazioni conformi sono omeomorfismi.

Definizione 2.8.3 *Siano A un aperto in $\mathbf{C}$, $\beta \in Fr(A)$. Diremo che β è un* **punto semplice** *di $Fr(A)$ se gode della proprietà seguente: comunque si prenda una successione (α_n) a valori in A e convergente a β, è sempre possibile costruire un cammino continuo $\gamma : [0,1] \to \mathbf{C}$ e una successione $(t_n)_{n \in \mathbf{N}}$, con $0 \leq t_1 < t_2 < ...$, $t_n \to 1$, tali che $\gamma(t_n) = \alpha_n$ $\forall n \in \mathbf{N}$, $\gamma([0,1[) \subseteq A$ e $\gamma(1) = \beta$.*

Osservazione 2.8.5 È facile verificare che, se A è un aperto convesso in $\mathbf{C}$, ogni punto di $Fr(A)$ è semplice per $Fr(A)$ (vedi l'esercizio 2.8.3)). Vediamo invece un esempio di un punto di frontiera non semplice.

Sia

$$A := \{z \in \mathbf{C} : |z| < 1\} \setminus \{z \in \mathbf{R} : z \in [0,1[\}.$$

Poniamo $\beta = 1/2$ e, per $n \in \mathbf{N}$,

$$\alpha_n := \begin{cases} 1/2 + i/n & \text{se } n \text{ è dispari,} \\ 1/2 - i/n & \text{se } n \text{ è pari.} \end{cases}$$

Siano $\gamma : [0,1[\to \mathbf{C}$ continua e $(t_n)_{n\in\mathbf{N}}$ una successione soddisfacente $0 \leq t_1 < t_2 < ..., t_n \to 1$, tali che $\gamma(t_n) = \alpha_n$ $\forall n \in \mathbf{N}$, $\gamma([0,1[) \subseteq A$. Allora, necessariamente, qualunque sia $n \in \mathbf{N}$, esiste τ_n compreso tra t_n e t_{n+1}, per cui si ha

$$\gamma_1(\tau_n) \leq 0. \tag{2.8.3}$$

Da ciò segue che non può valere $\lim_{t\to 1}\gamma(t) = \beta$. Infatti, se così fosse, esisterebbe $\delta > 0$, tale che per ogni $t \in [0,1[$ soddisfacente $t > 1-\delta$, si avrebbe $|\gamma(t) - 1/2| < 1/2$, che implica $\gamma_1(t) > 0$. Poiché la successione $(t_n)_{n\in\mathbf{N}}$ tende a 1, qualunque sia δ, per n abbastanza grande, dovrebbe valere $\tau_n > 1-\delta$, e quindi $\gamma_1(\tau_n) > 0$, in contraddizione con (2.8.3).

Siamo ora in grado di enunciare il seguente risultato (per una dimostrazione, vedi Rudin *"Analisi reale e complessa"*, teorema 14.19):

Teorema 2.8.3 *Siano A un aperto di $\mathbf{C}$ limitato, semplicemente connesso, i cui punti di frontiera sono tutti semplici, ϕ una trasformazione conforme tra A e $\{z \in \mathbf{C} : |z| < 1\}$. Allora ϕ è estendibile a un omeomorfismo tra $\overline{A}$ e $\{z \in \mathbf{C} : |z| \leq 1\}$, la cui restrizione a $Fr(A)$ è un omeomorfismo tra $Fr(A)$ e $\{z \in \mathbf{C} : |z| = 1\}$.*

Torniamo ora al problema di Dirichlet per l'equazione di Laplace. Il seguente lemma rende conto del nostro interesse nei confronti delle trasformazioni conformi:

Lemma 2.8.1 *Siano B e A aperti in $\mathbf{C}$, $u \in C^2(A;\mathbf{R})$, $\phi : B \to \mathbf{C}$ olomorfa, con $\phi(B) \subseteq A$. Allora, per ogni $z \in A$,*

$$\Delta(u \circ \phi)(z) = \Delta u(\phi(z))|\phi'(z)|^2.$$

In particolare, se u è armonica, anche $u \circ \phi$ è armonica.

Dimostrazione La dimostrazione è un esercizio di derivazione. Indichiamo con ϕ_1 e ϕ_2 le componenti di ϕ. Allora, per ogni $z \in B$, si ha

$$\frac{\partial(u \circ \phi)}{\partial x}(z) = D_1 u(\phi(z))\frac{\partial \phi_1}{\partial x}(z) + D_2 u(\phi(z))\frac{\partial \phi_2}{\partial x}(z),$$

$$\begin{aligned}\frac{\partial^2(u\circ\phi)}{\partial x^2}(z) &= D_1^2u(\phi(z))\frac{\partial\phi_1}{\partial x}(z)^2+2D_{12}u(\phi(z))\frac{\partial\phi_1}{\partial x}(z)\frac{\partial\phi_2}{\partial x}(z)\\ &\quad +D_2^2u(\phi(z))\frac{\partial\phi_2}{\partial x}(z)^2+D_1u(\phi(z))\frac{\partial^2\phi_1}{\partial x^2}(z)\\ &\quad +D_2u(\phi(z))\frac{\partial^2\phi_2}{\partial x^2}(z).\end{aligned} \tag{2.8.4}$$

Analogamente, si ha

$$\begin{aligned}\frac{\partial^2(u\circ\phi)}{\partial y^2}(z) &= D_1^2u(\phi(z))\frac{\partial\phi_1}{\partial y}(z)^2+2D_{12}u(\phi(z))\frac{\partial\phi_1}{\partial y}(z)\frac{\partial\phi_2}{\partial y}(z)\\ &\quad +D_2^2u(\phi(z))\frac{\partial\phi_2}{\partial y}(z)^2+D_1u(\phi(z))\frac{\partial^2\phi_1}{\partial y^2}(z)\\ &\quad +D_2u(\phi(z))\frac{\partial^2\phi_2}{\partial y^2}(z).\end{aligned} \tag{2.8.5}$$

Dalle condizioni di Cauchy Riemann, otteniamo

$$\frac{\partial\phi_1}{\partial x}(z)^2+\frac{\partial\phi_1}{\partial y}(z)^2=\frac{\partial\phi_1}{\partial x}(z)^2+\frac{\partial\phi_2}{\partial x}(z)^2=|\phi'(z)|^2,$$

$$\frac{\partial\phi_1}{\partial x}(z)\frac{\partial\phi_2}{\partial x}(z)+\frac{\partial\phi_1}{\partial y}(z)\frac{\partial\phi_2}{\partial y}(z)=0,$$

$$\frac{\partial\phi_2}{\partial x}(z)^2+\frac{\partial\phi_2}{\partial y}(z)^2=\frac{\partial\phi_1}{\partial y}(z)^2+\frac{\partial\phi_1}{\partial x}(z)^2=|\phi'(z)|^2.$$

Inoltre, ϕ_1 e ϕ_2 sono armoniche per il teorema 2.6.1 e l'osservazione 2.6.3. Allora, sommando (2.8.4) e (2.8.5), si ottiene la conclusione. □

Possiamo ora enunciare e dimostrare il seguente

Teorema 2.8.4 *Sia A un aperto limitato, semplicemente connesso in R^2, i cui punti di frontiera sono tutti semplici, nel senso della definizione 2.8.3. Sia poi $g \in C(Fr(A),\mathbf{R})$. Allora il problema di Dirichlet per l'equazione di Laplace (2.7.1) possiede un'unica soluzione $u \in C(\overline{A};\mathbf{R}) \cap C^2(A)$.*

Dimostrazione L'unicità segue dal corollario 2.7.1.

Proviamo l'esistenza. Indichiamo con B $\{z \in \mathbf{C} : |z| < 1\}$. Per il teorema 2.8.3, esiste una trasformazione conforme ϕ tra A e B, estendibile a un omeomorfismo tra $\overline{A}$ e $\overline{B}$, che continuiamo a indicare con ϕ. Sia $h : Fr(B) \to \mathbf{R}$, $h(z) = g(\phi^{-1}(z))$, $z \in Fr(B)$. $h \in C(Fr(B),\mathbf{R})$, in quanto composizione di funzioni continue. Consideriamo il problema

$$\begin{cases} \Delta v(z) = 0, & z \in B,\\ v(z') = h(z'), & z' \in Fr(B).\end{cases}$$

Per il teorema 2.7.2, questo problema ha un'unica soluzione $v \in C(\overline{B},\mathbf{R}) \cap C^2(B)$. Poniamo $u : \overline{A} \to \mathbf{R}$, $u(z) = v(\phi(z))$. Allora $u \in C(\overline{A},\mathbf{R}) \cap C^2(A)$.

Per il lemma 2.8.1, u è armonica in A. Inoltre, se $z \in Fr(A)$, poiché $\phi(z) \in Fr(B)$, si ha

$$u(z) = v(\phi(z)) = h(\phi(z)) = g(z).$$

La dimostrazione è completa. □

Osservazione 2.8.6 Tranne casi molto semplici (vedi, ad esempio, gli esercizi 2.8.4 e 2.8.5) non si è in grado di scrivere un'espressione esplicita di una trasformazione conforme tra $B := \{z \in \mathbf{C} : |z| < 1\}$ e A, aperto semplicemente connesso diverso da $\mathbf{C}$. Un caso particolarmente importante, dal punto di vista applicativo, è quello di A poligono limitato. Utilizzando il risultato dell'esercizio 2.8.4, si può sostituire B con il semispazio $\{z \in \mathbf{C} : Im(z) > 0\}$. In tale caso trasformazioni conformi tra $\{z \in \mathbf{C} : Im(z) > 0\}$ e A possono essere ottenute come soluzioni di certe equazioni differenziali ordinarie nei complessi dette formule di Schwarz-Christoffel. Ciò consente di ricavare numericamente delle espressioni approssimate di tali trasformazioni. Per una discussione su tale argomento, vedi (ad esempio) T. W. Gamelin *"Complex Analysis"* (Springer) cap. XI.

Esercizio 2.8.1 1) Verificare che la relazione "essere omeomorfo" è di equivalenza tra i sottoinsiemi di $\mathbf{C}$.

2) Verificare che la relazione di conforme equivalenza è di equivalenza tra gli aperti di $\mathbf{C}$.

Esercizio 2.8.2 Siano E e F sottoinsiemi omeomorfi di $\mathbf{C}$. Verificare che:

1) se E è chiuso e limitato, F è chiuso e limitato;

2) se E è connesso per archi, F è connesso per archi;

3) se E è semplicemente connesso, F è semplicemente connesso.

Verificare (costruendo un esempio) che, se E è limitato, F non è necessariamente limitato.

Esercizio 2.8.3 Verificare che, se A è un aperto convesso in $\mathbf{C}$, ogni punto di $Fr(A)$ è semplice per $Fr(A)$.

(Sugg.: siano $\beta \in Fr(A)$ e $(\alpha_n)_{n\in\mathbf{N}}$ una successione a valori in A, convergente a β. Porre, per $n \in \mathbf{N}$, $t_n = 1 - 2^{-n}$e, se $t \in [1 - 2^{-n}, 1 - 2^{-n-1}]$,

$$\gamma(t) = \alpha_n + \frac{\alpha_{n+1} - \alpha_n}{t_{n+1} - t_n}(t - t_n).$$

Osservare che nell'intervallo $[t_n, t_{n+1}]$, γ descrive il segmento di estremi α_n e α_{n+1}. Verificare, usando il fatto che $\lim_{n\to+\infty} \alpha_n = \beta$, che si ha $\lim_{t\to 1} \gamma(t) = \beta$.)

Esercizio 2.8.4 Costruire, a partire dall'esempio 2.8.1, una trasformazione conforme tra $\{z \in \mathbf{C} : |z| < 1\}$ e $\{z \in \mathbf{C} : Im(z) > 0\}$. Scrivere anche l'espressione della trasformazione inversa. (Sugg.: costruire preliminarmente una trasformazione conforme tra $\{z \in \mathbf{C} : Re(z) > 0\}$ e $\{z \in \mathbf{C} : Im(z) > 0\}$)

Esercizio 2.8.5 Siano $z_0 \in \mathbf{C}$ e $r > 0$. Allora la funzione $\phi(z) := \frac{z-z_0}{r}$ è una trasformazione conforme tra $\{z \in \mathbf{C} : |z - z_0| < r\}$ e $\{z \in \mathbf{C} : |z| < 1\}$. Utilizzare questo fatto per scrivere un'espressione esplicita per la soluzione u del problema di Dirichlet

$$\begin{cases} \Delta u(z) = 0, & |z - z_0| < r, \\ u(z) = g(z), & |z - z_0| = r, \end{cases}$$

con $g \in C(\{z \in \mathbf{C} : |z - z_0| = r\}, \mathbf{R})$.

2.9 Un metodo di discretizzazione alle differenze finite per il problema di Dirichlet

In questa sezione presenteremo un metodo alle differenze finite per approssimare la soluzione del problema di Dirichlet per l'equazione di Laplace.

Cominciamo con l'introdurre una discretizzazione della derivata seconda in dimensione uno. Sia $u : I \to \mathbf{R}$, di classe C^2, con I intervallo con interno non vuoto in $\mathbf{R}$. Siano poi $x \in I$ e $h \in \mathbf{R}^+$ tali che $x \pm h \in I$. Poniamo

$$\delta_h u(x) := \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}. \tag{2.9.1}$$

$\delta_h u(x)$ è un'approssimazione di $u''(x)$ nel senso seguente:

Lemma 2.9.1 *Sia $u : I \to \mathbf{R}$, di classe C^2, con I intervallo con interno non vuoto in $\mathbf{R}$. Siano poi $x \in I$ e, per $h \in \mathbf{R}^+$ tale che $x \pm h \in I$, sia $\delta_h(x)$ l'espressione definita in (2.9.1). Allora*

$$\lim_{h \to 0} \delta_h u(x) = u''(x).$$

Dimostrazione La dimostrazione è una semplice applicazione della formula di Taylor: si ha

$$u(x+k) = u(x) + u'(x)k + \frac{u''(x)}{2}k^2 + r(k),$$

con $r(k) = o(k^2)$ per $k \to 0$. Dunque,

$$\delta_h u(x) = u''(x) + \frac{r(h) + r(-h)}{h^2} \to u''(x) \quad (h \to 0).$$

$\square$

Il lemma 2.9.1 ci dice che l'errore commesso sostituendo $u''(x)$ con $\delta_h u(x)$ tende a 0 per $h \to 0$. Per il seguito, è conveniente disporre di una stima quantitativa di tale errore. A questo scopo, è conveniente introdurre alcune notevoli classi di funzioni.

Definizione 2.9.1 *Siano A un aperto in $\mathbf{R}^n$, $m \in \mathbf{N}$, $\alpha \in]0,1[$, $u : A \to \mathbf{R}$.*
Diremo che $u \in C^\alpha(A)$ se esiste $L > 0$ tale che, $\forall x, y \in A$,

$$|u(x) - u(y)| \leq L||x - y||^\alpha. \tag{2.9.2}$$

In tale caso, diremo che u è **hölderiana** *di esponente α.*
Se (2.9.2) è *soddisfatta con $\alpha = 1$, si dice anche che u* è **lipschitziana**. *Indicheremo con $Lip(A)$ la classe delle funzioni lipschitziane di dominio A.*
Diremo che $u \in C^{m+\alpha}(A)$ se $u \in C^m(A)$ e le derivate di ordine m di u appartengono a $C^\alpha(A)$.

Lemma 2.9.2 *Sia $u : I \to \mathbf{R}$, di classe $C^{2+\alpha}$, con I intervallo con interno non vuoto in $\mathbf{R}$ e $\alpha \in]0,1[$. Siano poi $x \in I$ e, per $h \in \mathbf{R}^+$ tale che $x \pm h \in I$, sia $\delta_h(x)$ l'espressione definita in (2.9.1). Supponiamo che valga*

$$|u''(z) - u''(y)| \leq L|x - y|^\alpha,$$

per un certo $L > 0$. Allora

$$|\delta_h u(x) - u''(x)| \leq Lh^\alpha.$$

Dimostrazione Applicando più volte il teorema fondamentale del calcolo integrale, si ha

$$u(x + h) - 2u(x) + u(x - h)$$
$$= [u(x + h) - u(x)] - [u(x) - u(x - h)]$$
$$= \int_x^{x+h} u'(y)dy - \int_{x-h}^{x} u'(y)dy = \int_x^{x+h} [u'(y) - u'(y - h)]dy$$
$$= \int_x^{x+h} (\int_{y-h}^{y} u''(z)dz)dy$$

$$= u''(x)h^2 + \int_x^{x+h} (\int_{y-h}^{y} (u''(z) - u''(x))dz)dy.$$

Dunque,

$$|\delta_h u(x) - u''(x)|$$
$$= h^{-2} | \int_x^{x+h} (\int_{y-h}^{y} (u''(z) - u''(x))dz)dy|$$
$$\leq h^{-2} \int_x^{x+h} (\int_{y-h}^{y} |u''(z) - u''(x)|dz)dy$$
$$\leq Lh^{-2} \int_x^{x+h} (\int_{y-h}^{y} |z - x|^{\alpha} dz)dy.$$

Se $x \leq y \leq x + h$ e $y - h \leq z \leq y$, si ha

$$x - h \leq y - h \leq z \leq y \leq x + h,$$

per cui $|z - x| \leq h$. Ne segue che

$$|\delta_h u(x) - u''(x)|$$
$$\leq Lh^{-2} \int_x^{x+h} (\int_{y-h}^{y} h^{\alpha} dz)dy$$
$$= Lh^{\alpha}.$$

□

Introduciamo ora una discretizzazione dell'operatore di Laplace, suggerita naturalmente da (2.9.1).

Siano $B \subseteq \mathbf{R}^2$, $h > 0$, $z = (x, y) \in B$ tale che i punti $z \pm he^1 = (x \pm h, y)$ e $z \pm he^2 = (x, y \pm h)$ appartengano tutti a B. Sia poi $U : B \to \mathbf{R}$. Poniamo

$$\Delta_h U(z) \quad := \quad \begin{array}{l} \frac{U(x+h,y) - 2U(x,y) + U(x-h,y)}{h^2} \\ + \frac{U(x,y+h) - 2U(x,y) + U(x,y-h)}{h^2}. \end{array} \tag{2.9.3}$$

Introduciamo anche altre notazioni: sia $A \subseteq \mathbf{R}^2$, aperto. Poniamo

$$A_h := \{hj : j = (j_1, j_2) \in \mathbf{Z}^2, hj \in A\}, \tag{2.9.4}$$

$$A_h^{\circ} := \{z \in A_h : z \pm he^1 \in A_h, z \pm he^2 \in A_h\}, \tag{2.9.5}$$

$$\partial A_h := A_h \setminus A_h^{\circ}. \tag{2.9.6}$$

4.00in by 2.46in (fi2.2 scaled 0600)

Figura 2.1

(Si veda, per un esempio, la Figura 2.1, con $A := \{(x, y \in \mathbf{R}^2 : x^2 + y^2 < 4\}$ e $h = 1/2$. I punti segnati in grassetto sono gli elementi di ∂A_h, gli altri sono gli elementi di A_h°). Si può verificare ed è intuitivamente chiaro che, se $z \in \partial A_h$, esiste $z' \in Fr(A)$, tale che $\|z - z'\| < h$. Possiamo dunque pensare a ∂A_h come a una discretizzazione di $Fr(A)$.

Introduciamo ora il seguente problema discreto:

$$\begin{cases} \Delta_h U(z) = F(z), & \forall z \in A_h^\circ, \\ U(z) = G(z), & \forall z \in \partial A_h. \end{cases} \tag{2.9.7}$$

Qui $F : A_h^\circ \to \mathbf{R}$ e $G : \partial_h A \to \mathbf{R}$ sono funzioni assegnate. Naturalmente, le eventuali soluzioni di (**??**) avranno come dominio A_h. Osserviamo che, se A è limitato, A_h è finito per ogni $h \in \mathbf{R}^+$. Nel seguito supporremo sempre che h sia abbastanza piccolo, in modo che $A_h \neq \emptyset$. Se A_h è costituito da $N_h \geq 1$ elementi, (**??**) è un sistema lineare di N_h equazioni in N_h incognite.

Nel seguito, useremo la notazione seguente: dato $A \subseteq \mathbf{R}^2$, indicheremo con d_A il suo **diametro**, definito come segue:

$$d_A := \sup\{\|z - v\| : z, v \in A\}. \tag{2.9.8}$$

Teorema 2.9.1 *Sia A un aperto limitato in $\mathbf{R}^2$ e sia $h > 0$, tale che $A_h \neq \emptyset$. Allora:*

*(I) (principio del massimo) se $U : A_h \to \mathbf{R}$ soddisfa (**??**) e $F(z) \geq 0$ $\forall z \in A_h^\circ$, allora*

$$\max_{A_h} U = \max_{\partial A_h} G; \tag{2.9.9}$$

*(II) se U risolve (**??**), si ha*

$$\max_{A_h} |U| \leq (d_A^2/4) \max_{A_h^\circ} |F| + \max_{\partial A_h} |G|;$$

*(III) per ogni $F : A_h^\circ \to \mathbf{R}$ e per ogni $G : \partial A_h \to \mathbf{R}$ il problema (**??**) ha un'unica soluzione U.*

Dimostrazione (I) La conclusione è ovvia se $A_h^\circ = \emptyset$. Supponiamo ora $A_h^\circ \neq \emptyset$ e $F(z) > 0$ $\forall z \in A_h^\circ$. Allora è chiaro che nessun punto di massimo di U può stare in A_h°. Infatti, si verifica immediatamente da (**??**) che, se z è un punto di massimo per U e $z \in A_h^\circ$, si ha $\Delta_h U(z) \leq 0$. Vale perciò ancora (**??**).

Consideriamo ora il caso generale. Fissiamo $(x_0, y_0) \in A$. Poniamo $V : \mathbf{R}^2 \to \mathbf{R}$, $V(x,y) = \frac{(x-x_0)^2+(y-y_0)^2}{4}$. È immediato verificare che si ha $\Delta_h V(z) = 1\ \forall z \in \mathbf{R}^2$. Allora, per $\epsilon > 0$ e $z \in A_h^\circ$, si ha

$$\Delta_h(U + \epsilon V)(z) = \Delta_h U(z) + \epsilon > 0.$$

Ne segue che, per ogni $z \in A_h$, si ha

$$U(z) + \epsilon V(z) \leq \max_{\partial A_h}(G + \epsilon V) \leq \max_{\partial A_h} G + \epsilon \max_{\overline{A}} V.$$

Al limite per $\epsilon \to 0^+$ otteniamo (I).

(II) Poniamo

$$M := \max_{A_h^\circ} |F|.$$

Sia V la funzione introdotta nella dimostrazione di (I). Osserviamo che, per $z \in A_h^\circ$, si ha

$$\Delta_h(\pm U + MV)(z) = \pm F(z) + M \geq 0.$$

Dal principio del massimo, segue, per ogni $z \in A_h$,

$$\begin{aligned} &\pm U(z) \\ &\leq \pm U(z) + MV(z) \leq \max_{\partial A_h}(\pm G + MV) \\ &\leq \max_{\partial A_h}(\pm G) + M \max_{\overline{A}} V \\ &\leq \max_{\partial A_h} |G| + M(d_A^2/4), \end{aligned}$$

da cui (II).

(III) Poiché si tratta di un sistema lineare di N_h equazioni in N_h incognite, l'esistenza e l'unicità di una soluzione è equivalente alla sola unicità , che è a sua volta equivalente al fatto che, se F e G sono identicamente nulle, il problema ha solo la soluzione identicamente nulla. Ciò segue immediatamente da (II). $\square$

Teorema 2.9.2 *Siano A un aperto limitato di $\mathbf{R}^2$, $u \in C(\overline{A}) \cap C^2(A)$ soluzione del problema (2.7.1), con $g \in C(Fr(A), \mathbf{R})$. Supponiamo che*

(I) esista $L_1 > 0$ tale che $|u(z) - u(v)| \leq L_1|z - v|$, per ogni z, v in $\overline{A}$;

(II) esistano $L_2 > 0$, $\alpha \in]0,1[$, tali che, per ogni $z, v \in A$,

$$|D_x^2 u(z) - D_x^2 u(v)| \leq L_2 |z - v|^\alpha,$$

$$|D_y^2 u(z) - D_y^2 u(v)| \leq L_2 |z - v|^\alpha.$$

Sia $h \in \mathbf{R}^+$, tale che $A_h \neq \emptyset$. Per ogni $z \in \partial A_h$, scegliamo $z' \in Fr(A)$ tale che $\|z - z'\| < h$ e poniamo

$$G(z) := g(z').$$

Sia u_h la soluzione del problema

$$\begin{cases} \Delta_h u_h(z) = 0, & \forall z \in A_h^\circ, \\ u_h(z) = G(z), & \forall z \in \partial A_h. \end{cases} \tag{2.9.10}$$

Allora,

$$\max_{A_h} |u_h - u| \leq (d_A^2/2) L_2 h^\alpha + L_1 h.$$

Dimostrazione Per ogni $z \in A_h^\circ$ si ha

$$\Delta_h (u_h - u)(z) = \Delta u(z) - \Delta_h u(z).$$

Dal teorema **??** (II) segue allora

$$\max_{A_h} |u_h - u|$$

$$\leq (d_A^2/4) \max_{A_h^\circ} |\Delta u - \Delta_h u| + \max_{\partial A_h} |G - u|.$$

Dal lemma 2.9.2 otteniamo, per ogni $z = (x, y) \in A_h^\circ$,

$$|\Delta u(z) - \Delta_h u(z)|$$

$$\leq |D_x^2 u(x,y) - \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2}|$$

$$+ |D_y^2 u(x,y) - \frac{u(x,y+h) - 2u(x,y) + u(x,y-h)}{h^2}|$$

$$\leq 2 L_2 h^\alpha.$$

Sia invece $z \in \partial A_h$. Si ha

$$|G(z) - u(z)|$$

$$= |g(z') - u(z)| = |u(z') - u(z)|$$

$$\leq L_1 |z' - z| \leq L_1 h.$$

Con ciò il risultato è provato. □

Osservazione 2.9.1 La stima di $\max_{A_h}|u_h - u|$ ottenuta col teorema **??** è subordinata alla validità delle condizioni (I) e (II). Si potrebbe dimostrare che queste condizioni sono soddisfatte (per certe costanti positive L_1 e L_2) sotto opportune ipotesi di regolarità dell'aperto A e del dato g su $Fr(A)$. Concludiamo con un risultato preciso in tal senso.

Definizione 2.9.2 *Sia A un aperto limitato in $\mathbf{R}^2$. Diremo che A* è *regolare di classe $C^{2+\alpha}$ ($\alpha \in]0,1[$) se esistono B aperto in $\mathbf{R}^2$ e $\Phi \in C^{2+\alpha}(B)$ tali che:*
(I) $\overline{A} \subseteq B$;
(II) $A = \{z \in B : \Phi(z) < 0\}$;
(III) $\nabla\Phi(z) \neq 0$ per ogni $z \in B$ tale che $g(z) = 0$.

Osservazione 2.9.2 Con riferimento alla definizione **??**,si può verificare che $Fr(A) = \{z \in B : \Phi(z) = 0\}$. La condizione (III) implica (applicando il teorema di Dini) che $Fr(A)$ è genuinamente un oggetto unidimensionale.

Siamo ora in grado di enunciare (senza dimostrazione) il seguente

Teorema 2.9.3 *Sia no A un aperto limitato e regolare di classe $C^{2+\alpha}$ ($0 < \alpha < 1$) in $\mathbf{R}^2$, g una funzione a valori reali, di classe $C^{2+\alpha}$ in qualche aperto Ω contenente $Fr(A)$. Allora il problema (2.7.1) ha una soluzione (unica per il corollario 2.7.1) u. Tale soluzione u appartiene a $C^{2+\alpha}(A)$ e verifica le condizioni (I) e (II) del teorema* **??***, con certe costanti L_1 e L_2 stimabili in termini di A e di g.*

Esercizio 2.9.1 Sia A un aperto in $\mathbf{R}^n$. Verificare quanto segue:
(I) $\forall \alpha \in]0,1]$ $C^\alpha(A) \subseteq C(A,\mathbf{R})$;
(II) se A è limitato e $0 < \alpha \leq \beta < 1$,

$$Lip(A) \subseteq C^\beta(A) \subseteq C^\alpha(A) \subseteq BC(A;\mathbf{R});$$

(III) se A è convesso, $u \in C^1(A)$ e le derivate prime di A sono limitate, allora $u \in Lip(A)$ (Sugg.: utilizzare il teorema del valor medio);
(IV)verificare, più in generale, che, se $u \in C^1(A)$, allora u è lipschitziana su ogni palla la cui chiusura è contenuta in A.

Esercizio 2.9.2 Siano $0 < \alpha < \beta < 1$, $f : \mathbf{R}^+ \to \mathbf{R}$, $f(x) = x^\alpha$. Dimostrare che $u \in C^\alpha(\mathbf{R}^+) \setminus C^\beta(\mathbf{R}^+)$. (Sugg.: verificare preliminarmente che, se $x \geq 1$, allora $x^\alpha - 1 \leq (x-1)^\alpha$, facendo vedere che $x \to x^\alpha - 1 - (x-1)^\alpha$ è decrescente su $[1, +\infty[$)

Esercizio 2.9.3 Siano A un aperto connesso per archi in $\mathbf{R}^n$, $L \in \mathbf{R}^+$, $\alpha > 1$, $f : A \to \mathbf{R}$ tale che

$$|f(z) - f(v)| \leq L\|z - v\|^{\alpha}, \forall z, v \in A.$$

Verificare che f è costante. (Sugg.: far vedere che $\forall z \in A$, $\forall j \in \{1, ..., n\}$, $D_j f(z) = 0$.)

Capitolo 3

La trasformata di Fourier

3.1 La trasformata di Fourier in $L^1(\mathbf{R}^n)$

Ricordiamo (vedi l'esempio 1.1.5) che $L^1(\mathbf{R}^n)$ è lo spazio vettoriale delle classi di equivalenza di funzioni sommabili. Tale spazio è di Banach se munito della norma (1.1.9). Nel seguito, per semplicità di notazione, data $f \in \mathcal{L}^1(\mathbf{R}^n)$, scriveremo f in luogo di $[f]$ e penseremo agli elementi di $L^1(\mathbf{R}^n)$ come funzioni sommabili, identificando tra loro due funzioni che coincidono quasi dappertutto. Naturalmente, caso per caso, i vari risultati presentati dovranno essere invarianti rispetto alla relazione di equivalenza.

Introduciamo ora la definizione di **trasformata di Fourier** di $f \in L^1(\mathbf{R}^n)$.

Definizione 3.1.1 *Sia $f \in L^1(\mathbf{R}^n)$. Chiameremo trasformata di Fourier di f e indicheremo con la notazione $\hat{f}$ o $\mathcal{F}f$ la funzione*

$$\begin{cases} \hat{f} : \mathbf{R}^n \to \mathbf{C}, \\ \hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-i<x,\xi>} f(x)dx, \end{cases} \tag{3.1.1}$$

con $< .,. >$ prodotto scalare "standard" in $\mathbf{R}^n$ (vedi l'esempio 1.4.1).

Osservazione 3.1.1 Per ogni $\xi \in \mathbf{R}^n$ si ha $|e^{-i<x,\xi>} f(x)| = |f(x)|$. Ne segue che l'integrale in (**??**) è ben definito.

Definizione 3.1.2 *Sia $A \subseteq \mathbf{R}^n$. Indichiamo con $BC(A)$ lo spazio vettoriale delle funzioni $f : A \to \mathbf{C}$ continue e limitate. Introduciamo in $BC(A)$ la norma $\|.\|_\infty$ definita come segue:*

$$\|f\|_\infty := \sup\{|f(a)| : a \in A\}. \tag{3.1.2}$$

Osservazione 3.1.2 Si può verificare in maniera simile a $BC(A, \mathbf{R})$ che con la norma $\|.\|_\infty$ $BC(A)$ è uno spazio di Banach (vedi l'esercizio **??**).

Teorema 3.1.1 *(di Riemann-Lebesgue) Per ogni* $f \in L^1(\mathbf{R}^n)$ *la trasformata di Fourier* $\hat{f}$ *appartiene a* $BC(\mathbf{R}^n)$. *Si ha inoltre*

$$\lim_{\|\xi\| \to +\infty} \hat{f}(\xi) = 0, \tag{3.1.3}$$

nel senso che $\forall \epsilon > 0$ *esiste* $\delta(\epsilon) > 0$ *tale che, se* $\xi \in \mathbf{R}^n$ *e* $\|\xi\| > \delta(\epsilon)$, *vale* $|\hat{f}(\xi)| < \epsilon$.

Dimostrazione parziale Dal teorema 0.2.5(IV) segue

$$|\hat{f}(\xi)| \leq \int_{\mathbf{R}^n} |f(x)| dx = \|f\|_1. \tag{3.1.4}$$

Quindi $\hat{f}$ è limitata. Per dimostrare che è continua, possiamo applicare il teorema 1.3.4.

Sia $(\xi^k)_{k\in\mathbf{N}}$ una successione in $\mathbf{R}^n$ tale che $\lim\limits_{k\to+\infty} \xi^k = \xi^0 \in \mathbf{R}^n$. Per ogni $k \in \mathbf{N}$ si ha

$$|e^{-i<x,\xi^k>} f(x)| = |f(x)| \qquad \forall x \in \mathbf{R}^n.$$

Possiamo allora applicare il teorema della convergenza dominata, per concludere che

$$\lim_{k\to+\infty} \hat{f}(\xi^k) = \lim_{k\to+\infty} \int_{\mathbf{R}^n} e^{-i<x,\xi^k>} f(x) dx = \int_{\mathbf{R}^n} e^{-i<x,\xi^0>} f(x) dx = \hat{f}(\xi^0).$$

Perciò $\hat{f}$ è continua.

Non dimostriamo che vale (**??**). □

Corollario 3.1.1 *Indichiamo con* $\mathcal{F}$ *l'applicazione*

$$\begin{cases} \mathcal{F} : L^1(\mathbf{R}^n) \to BC(\mathbf{R}^n), \\ \mathcal{F}f = \hat{f}. \end{cases} \tag{3.1.5}$$

Allora $\mathcal{F}$ è *lineare e continua tra* $L^1(\mathbf{R}^n)$ *e* $BC(\mathbf{R}^n)$.

Dimostrazione Lasciamo al lettore la banale verifica della linearità (esercizio **??**). Facciamo vedere che $\mathcal{F}$ è continua. Sia $(f_k)_{k\in\mathbf{N}}$ una successione in $L^1(\mathbf{R}^n)$ tale che $\lim_{k\to+\infty} f_k = f$ in $L^1(\mathbf{R}^n)$. Dobbiamo verificare che $\lim_{k\to+\infty} \mathcal{F}f_k = \mathcal{F}f$ in $BC(\mathbf{R}^n)$.

Da (**??**) segue che $\|\mathcal{F}g\|_\infty \leq \|g\|_1$ $\forall g \in L^1(\mathbf{R}^n)$. Ciò implica che

$$\|\mathcal{F}f - \mathcal{F}f_k\|_\infty = \|\mathcal{F}(f - f_k)\|_\infty \leq \|f - f_k\|_1 \to 0$$

per $k \to +\infty$. $\square$

Esempio 3.1.1 Una conseguenza dell'esempio 2.5.6 è che, se $f(x) = e^{-x^2}$, allora

$$\hat{f}(\xi) = \sqrt{\pi}e^{-\frac{\xi^2}{4}}.$$

Esempio 3.1.2 Sia $f : \mathbf{R} \to \mathbf{R}$, $f(x) = \frac{1}{1+x^2}$. $f \in L^1(\mathbf{R})$. Calcoliamone la trasformata di Fourier. Si tratta di calcolare

$$\int_R \frac{e^{-ix\xi}}{1+x^2}dx \tag{3.1.6}$$

al variare di $\xi \in \mathbf{R}$. Osserviamo preliminarmente che, se $z \in \mathbf{C}$,

$$|e^{-iz\xi}| = e^{\xi Im(z)}.$$

Dunque, se $\xi \geq 0$, si ha $|e^{-iz\xi}| \leq 1$ se $Im(z) \leq 0$, se $\xi \leq 0$, si ha $|e^{-iz\xi}| \leq 1$ se $Im(z) \geq 0$. Supponiamo $\xi \geq 0$. Dato $n \in \mathbf{N}$, $n \geq 2$, consideriamo un cammino chiuso α_n orientato in senso orario che percorra una volta, in successione, $[-n, n]$ e la semicinconferenza $\{z \in \mathbf{C} : |z| = n, Im(z) \leq 0\}$. Poiché la funzione $f(z) := \frac{e^{-iz\xi}}{1+z^2}$ è olomorfa in $\mathbf{C} \setminus \{i, -i\}$, si ha

$$\int_{\alpha_n} f(z)dz = 2\pi i Res(f, -i)Ind(\alpha_n, -i) = -2\pi i Res(f, -i).$$

È facile verificare che $-i$ è un polo semplice per f. Dunque

$$Res(f, -i) = \lim_{z\to -i}(z + i)f(z) = \lim_{z\to -i}\frac{e^{-iz\xi}}{z - i} = i\frac{e^{-\xi}}{2}.$$

Dunque

$$\int_{\alpha_n} f(z)dz = \pi e^{-\xi}.$$

D'altra parte,

$$\int_{\alpha_n} f(z)dz = \int_{-n}^{n} \frac{e^{-ix\xi}}{1+x^2}dx - \int_{C_r^-(0)} f(z)dz,$$

con

$$\begin{cases} C_r^-(0) : [\pi, 2\pi] \to \mathbf{C}, \\ C_r^-(0)(t) = re^{it}, \end{cases} \tag{3.1.7}$$

per $r > 0$. Dal solito teorema 2.2.3 e dalle osservazioni precedenti, otteniamo

$$|\int_{C_r^0(0)} f(z)dz| \leq \frac{1}{n^2-1}\pi n \to 0 \qquad (n \to +\infty).$$

Dunque, al limite per $n \to +\infty$, otteniamo

$$\int_{\mathbf{R}} \frac{e^{-ix\xi}}{1+x^2}dx = \pi e^{-\xi}$$

se $\xi \geq 0$.

Il caso $\xi < 0$ può essere trattato integrando su un cammino che percorra in successione e in senso antiorario $[-n, n]$ e la semicinconferenza $\{z \in \mathbf{C} : |z| = n, Im(z) \geq 0\}$, per sfruttare la limitatezza di $e^{-iz\xi}$ su $\{z \in \mathbf{C} : Im(z) \geq 0\}$. Lasciamo al lettore il completamento del calcolo. Ci limitiamo a dire che vale la seguente formula:

$$\int_{\mathbf{R}} \frac{e^{-ix\xi}}{x^2+1}dx = \pi e^{-|\xi|} \qquad \forall \xi \in \mathbf{R}. \tag{3.1.8}$$

Esempio 3.1.3 Sia $f : \mathbf{R} \to \mathbf{R}$, $f(x) = \chi_+(x)e^{-x}$, ove χ_+ è la funzione caratteristica di $\mathbf{R}^+$. Si tratta di calcolare

$$\int_{R^+} e^{-x(1+i\xi)}dx \tag{3.1.9}$$

al variare di $\xi \in \mathbf{R}$. Applicando il teorema 0.3.4, si ottiene

$$\int_{R^+} e^{-x(1+i\xi)}dx = \lim_{c \to +\infty} \int_0^c e^{-x(1+i\xi)}dx = \lim_{c \to +\infty} \frac{1 - e^{-c(1+i\xi)}}{1+i\xi}.$$

Da $|e^{-c(1+i\xi)}| = e^{-c} \to 0$ per $c \to +\infty$ otteniamo

$$\hat{f}(\xi) = \frac{1}{1+i\xi} \qquad \forall \xi \in \mathbf{R}.$$

Se $\xi \geq 1$, si ha

$$|\hat{f}(\xi)| = \frac{1}{\sqrt{1+\xi^2}} = \frac{1}{\xi}\frac{1}{\sqrt{1+\frac{1}{\xi^2}}} \geq \frac{1}{\sqrt{2}\xi}.$$

Poiché $\int_{[1,+\infty[} \frac{1}{\sqrt{2}\xi} d\xi = +\infty$, $\hat{f}$ non è sommabile.

Questo esempio prova che, data $f \in L^1(\mathbf{R}^n)$, non vale necessariamente $\hat{f} \in L^1(\mathbf{R}^n)$.

Vediamo ora alcuni risultati di interazione della trasformata di Fourier con la derivata. A questi premettiamo la seguente convenzione: data $f \in L^1(\mathbf{R}^n)$, diremo che f è continua se esiste g continua tale che $g \sim f$. Ciò significa che la classe di equivalenza $[f]$ contiene una funzione continua, necessariamente unica (si rammenti l'esercizio 1.7.1 per il caso unidimensionale). Convenzioni analoghe varranno quando scriveremo che f è di classe C^1, ecc..

Proposition 3.1.1 *Sia $f \in L^1(\mathbf{R}^n)$ e di classe C^1. Supponiamo che, per un certo $j \in \{1,...,n\}$, $D_j f \in L^1(\mathbf{R}^n)$. Allora $\forall \xi \in \mathbf{R}^n$*

$$\mathcal{F}(D_j f)(\xi) = i\xi_j \hat{f}(\xi).$$

Dimostrazione parziale Dimostriamo il risultato nel caso $n = 1$, supponendo inoltre che valga $\lim_{x\to\pm\infty} f(x) = 0$.

Integrando per parti, si ha $\forall \xi \in \mathbf{R}$,

$$\hat{f'}(\xi) = \int_{\mathbf{R}} e^{-ix\xi} f'(x)dx = \lim_{n\to+\infty} \int_{-n}^{n} e^{-ix\xi} f'(x)dx =$$

$$= \lim_{n\to+\infty} (e^{-in\xi} f(n) - e^{in\xi} f(-n) + i\xi \int_{-n}^{n} e^{-ix\xi} f(x)dx) = i\xi \hat{f}(\xi).$$

□

Proposition 3.1.2 *Sia $f \in L^1(\mathbf{R}^n)$, tale che, per un certo $j \in \{1,...,n\}$, $x_j f \in L^1(\mathbf{R}^n)$. Allora $\hat{f}$ ammette derivata parziale $D_j \hat{f}(\xi)$ in ogni punto $\xi \in \mathbf{R}^n$; inoltre, $\forall \xi \in \mathbf{R}^n$,*

$$D_j \hat{f}(\xi) = -i\mathcal{F}(x_j f)(\xi).$$

Idea della dimostrazione Derivando formalmente rispetto a ξ_j, si ha

$$D_j\hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-i<x,\xi>}(-ix_j)f(x)dx.$$

□

Introduciamo ora alcune notazioni estremamente utili.

Chiameremo **multiindice** un elemento di $\mathbf{N}_0^n$, con $\mathbf{N}_0 := \{0\} \cup \mathbf{N}$. Se $\alpha := (\alpha_1, ..., \alpha_n)$ e $\beta := (\beta_1, ..., \beta_n)$, poniamo

$$\alpha + \beta := (\alpha_1 + \beta_1, ..., \alpha_n + \beta_n). \tag{3.1.10}$$

Scriveremo

$$\alpha \leq \beta \tag{3.1.11}$$

per indicare che $\alpha_j \leq \beta_j \ \forall j \in \{1, ..., n\}$ e

$$|\alpha| := \alpha_1 + ... + \alpha_n. \tag{3.1.12}$$

Chiameremo $|\alpha|$ **peso** del multiindice α.

Se $x = (x_1, ..., x_n) \in \mathbf{C}^n$ e $\alpha = (\alpha_1, ..., \alpha_n)$ è un multiindice, poniamo

$$x^\alpha := x_1^{\alpha_1} ... x_n^{\alpha_n}, \tag{3.1.13}$$

con la convenzione che $0^0 := 1$. Ad esempio, se $x = (0, 2i)$ e $\alpha = (0, 3)$, avremo

$$x^\alpha = (2i)^3 = -8i.$$

Infine, se A è un aperto di $\mathbf{R}^n$, $f \in C^m(A)$, per un certo $m \in \mathbf{N}$ e $|\alpha| \leq m$, poniamo

$$D^\alpha f := D_1^{\alpha_1} ... D_n^{\alpha_n} f, \tag{3.1.14}$$

ove D_j indica la derivata parziale rispetto alla $j-$esima variabile. Ricordiamo che, come conseguenza del teorema di Schwarz, le derivazioni rispetto alle singole variabili in (**??**) possono essere scambiate di ordine senza modificare il risultato finale. Ad esempio, per $n = 3$, se $\alpha = (1, 2, 0)$, si ha

$$D^\alpha f = D_1 D_2^2 f.$$

Si osservi che se β è un secondo multiindice tale che $|\alpha + \beta| \leq m$, vale

$$D^\alpha(D^\beta f) = D^{\alpha+\beta} f. \tag{3.1.15}$$

Vediamo ora una generalizzazione delle proposizioni **??** e **??**.

Corollario 3.1.2 *(I) Sia $f \in L^1(\mathbf{R}^n)$ e di classe C^m ($m \in \mathbf{N}$). Supponiamo che, per ogni multiindice α, con $|\alpha| \le m$, $D^\alpha f \in L^1(\mathbf{R}^n)$. Allora $\forall \xi \in \mathbf{R}^n$*

$$\mathcal{F}(D^\alpha f)(\xi) = (i\xi)^\alpha \hat{f}(\xi).$$

(II) Sia $f \in L^1(\mathbf{R}^n)$, tale che, per un certo $m \in \mathbf{N}$, $\|x\|^m f \in L^1(\mathbf{R}^n)$, con $\|.\|$ norma euclidea. Allora $\hat{f} \in C^m(\mathbf{R}^n)$; inoltre, per ogni multiindice α tale che $|\alpha| \le m$ si ha $(-ix)^\alpha f \in L^1(\mathbf{R}^n)$ e $\forall \xi \in \mathbf{R}^n$,

$$D^\alpha \hat{f}(\xi) = \mathcal{F}((-ix)^\alpha f)(\xi).$$

Dimostrazione Il risultato segue applicando più volte le proposizioni **??** e **??**.

Per quanto riguarda il punto (II), si osservi che, se $|\alpha| \le m$, allora $(-ix)^\alpha f \in L^1(\mathbf{R}^n)$. Infatti, $\forall x \in \mathbf{R}^n$, $\forall j \in \{1, ..., n\}$, $|x_j| \le \|x\|$, da cui $|(-ix)^\alpha f(x)| \le \|x\|^{|\alpha|}|f(x)|$. Se $\|x\| \le 1$, si ha $\|x\|^{|\alpha|}|f(x)| \le |f(x)|$. Se $\|x\| \ge 1$, vale $\|x\|^{|\alpha|}|f(x)| \le \|x\|^m|f(x)|$. Dunque, in ogni caso

$$|(-ix)^\alpha f(x)| \le (1 + \|x\|^m)|f(x)|.$$

□

Osservazione 3.1.3 In base al teorema di Riemann-Lebesgue, sotto le ipotesi del corollario **??**(II), se $|\alpha| \le m$, si ha che $D^\alpha \hat{f} \in BC(\mathbf{R}^n)$ e soddisfa la condizione (**??**).

Concludiamo questa sezione con la fondamentale **formula di inversione**.

Teorema 3.1.2 *Sia $f \in L^1(\mathbf{R}^n)$ tale che $\hat{f} \in L^1(\mathbf{R}^n)$. Allora*

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i<x,\xi>} \hat{f}(\xi) d\xi = \frac{1}{(2\pi)^n} \hat{\hat{f}}(-x) \qquad q.d. \quad in \quad \mathbf{R}^n. \tag{3.1.16}$$

Dimostrazione parziale Dimostriamo il risultato nell'ipotesi ulteriore che $f \in BC(\mathbf{R}^n)$.

Dal teorema della convergenza dominata, si ha

$$\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i<x,\xi>} \hat{f}(\xi) d\xi =$$

$$\lim_{k\to+\infty}\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i<x,\xi>-\frac{\|\xi\|^2}{k}}\hat{f}(\xi)d\xi =$$

$$= \lim_{k\to+\infty}\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i<x,\xi>-\frac{\|\xi\|^2}{k}}(\int_{\mathbf{R}^n} e^{-i<y,\xi>}f(y)dy)d\xi.$$

Poiché

$$|e^{i<x,\xi>-\frac{\|\xi\|^2}{k}}e^{-i<y,\xi>}f(y)| = e^{-\frac{\|\xi\|^2}{k}}|f(y)|$$

e, applicando il teorema di Tonelli,

$$\int_{\mathbf{R}^n\times\mathbf{R}^n} e^{-\frac{\|\xi\|^2}{k}}|f(y)|d\xi dy = (\int_{\mathbf{R}} e^{-\frac{t^2}{2}}dt)^n\int_{\mathbf{R}^n}|f(y)|dy < +\infty,$$

dal teorema di Fubini segue

$$\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i<x,\xi>-\frac{\|\xi\|^2}{k}}(\int_{\mathbf{R}^n} e^{-i<y,\xi>}f(y)dy)d\xi =$$

$$= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}(\int_{\mathbf{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}}d\xi)f(y)dy$$

Per il teorema di Fubini, si ha

$$\int_{\mathbf{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}}d\xi = \prod_{j=1}^n\int_{\mathbf{R}} e^{i(x_j-y_j)\xi_j-\frac{\xi_j^2}{k}}d\xi_j =$$

$$= k^{\frac{n}{2}}\prod_{j=1}^n\int_{\mathbf{R}} e^{i(x_j-y_j)\sqrt{k}t-t^2}dt = (k\pi)^{\frac{n}{2}}e^{-\frac{k\|x-y\|^2}{4}},$$

applicando il risultato dell'esempio 2.5.6. Dunque

$$\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}(\int_{\mathbf{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}}d\xi)f(y)dy =$$

$$= (\frac{k}{4\pi})^{\frac{n}{2}}\int_{\mathbf{R}^n} e^{-\frac{k\|x-y\|^2}{4}}f(y)dy = \pi^{-\frac{n}{2}}\int_{\mathbf{R}^n} e^{-\|z\|^2}f(x-\frac{2z}{\sqrt{k}})dz.$$

Passiamo adesso al limite, mandando k a $+\infty$. Poiché f è continua, si ha che, $\forall z\in\mathbf{R}^n$,

$$\lim_{k\to+\infty} e^{-\|z\|^2}f(x-\frac{2z}{\sqrt{k}}) = e^{-\|z\|^2}f(x).$$

Inoltre, qualunque sia $k \in \mathbf{N}$ e qualunque sia la scelta di x e z in $\mathbf{R}^n$, si ha

$$|e^{-\|z\|^2} f(x - \frac{2z}{\sqrt{k}})| \leq g(z) := \sup_{\mathbf{R}^n} |f| e^{-\|z\|^2}.$$

Poiché g è sommabile, possiamo (in base al teorema della convergenza dominata), passare al limite sotto il segno di integrale e concludere che

$$\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i<x,\xi>} \hat{f}(\xi) d\xi = \pi^{-\frac{n}{2}} \int_{\mathbf{R}^n} e^{-\|z\|^2} dz f(x) =$$

$$= f(x).$$

Con ciò la dimostrazione è conclusa. $\square$

Corollario 3.1.3 *L'applicazione $\mathcal{F}$, definita in (**??**), è iniettiva.*

Dimostrazione Poiché $\mathcal{F}$ è lineare, basta far vedere che il suo nucleo è costituito dallo zero dello spazio vettoriale $L^1(\mathbf{R}^n)$, che è $[0]$. Ciò segue immediatamente dalla formula di inversione (**??**). $\square$

Il seguente risultato sarà utile nelle sezioni successive:

Teorema 3.1.3 *Siano f e g elementi di $L^1(\mathbf{R}^n)$. Allora*

$$\int_{\mathbf{R}^n} \hat{f}(\xi) g(\xi) d\xi = \int_{\mathbf{R}^n} f(x) \hat{g}(x) dx. \tag{3.1.17}$$

Dimostrazione Innanzi tutto, gli integrali in (**??**) sono definiti. Infatti, per esempio, $\hat{f}$ è limitata. Dunque, $\forall \xi \in \mathbf{R}^n$,

$$|\hat{f}(\xi) g(\xi)| \leq \|\hat{f}\|_\infty |g(\xi)|.$$

Inoltre, poiché $(x, \xi) \to e^{-i<x,\xi>} f(x) g(\xi)$ è sommabile su $\mathbf{R}^n \times \mathbf{R}^n$, per il teorema di Fubini si ha

$$\int_{\mathbf{R}^n} \hat{f}(\xi) g(\xi) d\xi =$$

$$\int_{\mathbf{R}^n} (\int_{\mathbf{R}^n} e^{-i<x,\xi>} f(x) dx) g(\xi) d\xi = \int_{\mathbf{R}^n \times \mathbf{R}^n} e^{-i<x,\xi>} f(x) g(\xi) dx d\xi =$$

$$= \int_{\mathbf{R}^n} f(x) (\int_{\mathbf{R}^n} e^{-i<x,\xi>} g(\xi) d\xi) dx$$

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

$\square$

Esercizio 3.1.1 Verificare che lo spazio normato $BC(A)$ della definizione **??** è uno spazio di Banach.

Esercizio 3.1.2 Verificare che $\mathcal{F}$ è lineare da $L^1(\mathbf{R}^n)$ a $BC(\mathbf{R}^n)$.

Esercizio 3.1.3 Siano X e Y due spazi normati sullo stesso campo,con norme (rispettivamente) $\|.\|_X$ e $\|.\|_Y$. Sia poi $T : X \to Y$ lineare e tale che esiste $L \geq 0$ per cui si ha

$$\|Tx\|_Y \leq L\|x\|_X \qquad \forall x \in X. \tag{3.1.18}$$

Dimostrare che T è continua.

Esercizio 3.1.4 Siano X e Y due spazi normati sullo stesso campo,con norme (rispettivamente) $\|.\|_X$ e $\|.\|_Y$. Sia poi $T : X \to Y$ lineare. Dimostrare che T è continua se e solo se è continua in 0.

Esercizio 3.1.5 Sia $f \in L^1(\mathbf{R}^n)$. Dimostrare quanto segue, applicando il teorema di cambiamento di variabile:

(I) sia $A : \mathbf{R}^n \to \mathbf{R}^n$ lineare e invertibile; allora $f \circ A \in L^1(\mathbf{R}^n)$ e, $\forall \xi \in \mathbf{R}^n$, vale

$$\mathcal{F}(f \circ A)(\xi) = \frac{\hat{f}((A^{-1})^T\xi)}{|det(A)|},$$

ove abbiamo indicato con B^T la trasposta della matrice B;

(II) se $f(-x) = -f(x)$ q. d. in $\mathbf{R}^n$, allora $\hat{f}(-\xi) = -\hat{f}(\xi)$ $\forall \xi \in \mathbf{R}^n$;

(III) se $f(-x) = f(x)$ q. d. in $\mathbf{R}^n$, allora $\hat{f}(-\xi) = \hat{f}(\xi)$ $\forall \xi \in \mathbf{R}^n$;

(IV) se $x^0 \in \mathbf{R}^n$ e $g(x) = f(x - x^0)$ q. d. in $\mathbf{R}^n$, allora $g \in L^1(\mathbf{R}^n)$ e $\hat{g}(\xi) = e^{-i<x^0,\xi>}\hat{f}(\xi)$ $\forall \xi \in \mathbf{R}^n$.

Esercizio 3.1.6 Calcolare le trasformate di Fourier delle seguenti funzioni:

(I) $f(x) = \frac{1}{x^2+2x+2}$;

(II) $f(x) = \frac{1}{(x^2+1)^2}$;

(III) $f(x) = \frac{1}{x^4+1}$;

(IV) $f(x) = \frac{x}{(x^2+1)^2}$;

(V) $f(x, y) = \frac{1}{(x^2+1)(y^2+1)}$.

Esercizio 3.1.7 Sia $f \in L^1(\mathbf{R}^n)$ e $g(x) = \overline{f(x)}$ q. d. . Verificare che, $\forall \xi \in \mathbf{R}^n$,

$$\hat{g}(\xi) = \overline{\hat{f}(-\xi)}.$$

3.2 La classe $\mathcal{D}(\Omega)$

Apriamo ora una parentesi, per introdurre una classe di funzioni che svolgerà un ruolo ausiliario assai rilevante.

Cominciamo allora con lo specificare che, se Ω è un aperto in $\mathbf{R}^n$, indicheremo con $C^\infty(\Omega)$ la classe delle funzioni da Ω a valori complessi e dotate di tutte le derivate di ogni ordine continue su Ω.

Definizione 3.2.1 *Sia Ω un aperto in $\mathbf{R}^n$. Sia $f \in C^\infty(\Omega)$. Scriveremo che $f \in \mathcal{D}(\Omega)$ se esiste $K \subseteq \Omega$ chiuso e limitato, tale che $f(x) = 0$ se $x \in \Omega \setminus K$.*

È facile verificare che, se $f \in \mathcal{D}(\Omega)$, ogni derivata $D^\alpha f \in \mathcal{D}(\Omega)$, per ogni multiindice α. Si vede, inoltre, che, se f e g sono elementi di $\mathcal{D}(\Omega)$, anche $f + g \in \mathcal{D}(\Omega)$. Ciò segue dal fatto che, se K e L sono sottoinsiemi chiusi e limitati di Ω, se f è nulla al di fuori di K e g è nulla al di fuori di L, allora $f + g$ è nulla al di fuori di $K \cup L$. Si potrebbe dimostrare che $K \cup L$ è ancora chiuso e limitato. È infine chiaro che, se $f \in \mathcal{D}(\Omega)$ e $\alpha \in \mathbf{C}$, allora $\alpha f \in \mathcal{D}(\Omega)$. Dunque $\mathcal{D}(\Omega)$ costituisce uno spazio vettoriale di funzioni, con le solite operazioni di somma e prodotto per uno scalare.

La costruzione di esempi espliciti di elementi di $\mathcal{D}(\Omega)$ non è del tutto ovvia. A tale fine, cominciamo con il seguente

Lemma 3.2.1 *Sia $g : \mathbf{R} \to \mathbf{R}$, definita come segue:*

$$g(x) = \begin{cases} e^{-\frac{1}{x}} & se \quad x > 0, \\ 0 & se \quad x \leq 0. \end{cases} \tag{3.2.1}$$

Allora $g \in C^\infty(\mathbf{R})$.

Dimostrazione parziale Si vede subito che g è continua, in quanto $\lim_{x \to 0^+} e^{-\frac{1}{x}} = \lim_{y \to +\infty} e^{-y} = 0$. Si ha poi $g'(x) = 0$ se $x < 0$ e $D_- g(0) = 0$. Se $x > 0$, vale $g'(x) = e^{-\frac{1}{x}} x^{-2}$, mentre

$$D_+ g(0) = \lim_{x \to 0^+} \frac{g(x)}{x} = \lim_{x \to 0^+} e^{-\frac{1}{x}} x^{-1} = \lim_{y \to +\infty} e^{-y} y = 0.$$

Dunque g è derivabile in ogni punto di $\mathbf{R}$. g è di classe C^1, in quanto

$$\lim_{x \to 0^+} g'(x) = \lim_{x \to 0^+} e^{-\frac{1}{x}} x^{-2} = 0.$$

Il ragionamento si può iterare, sfruttando il fatto che, per ogni $n \in \mathbf{N}$, per $x > 0$, $g^{(n)}(x)$ è della forma $e^{-\frac{1}{x}} P_n(\frac{1}{x})$, ove P_n è una funzione polinomiale. □

Esempio 3.2.1 Costruiamo un primo esempio di elemento non identicamente nullo di $\mathcal{D}(\mathbf{R})$. Sia $\epsilon > 0$. Poniamo

$$\begin{cases} f_\epsilon : \mathbf{R} \to \mathbf{R}, \\ f_\epsilon(x) = g(\epsilon^2 - x^2), \end{cases} \tag{3.2.2}$$

ove g è la funzione definita in (**??**). f_ϵ è di classe C^∞ come composizione di funzioni di classe C^∞ e $f_\epsilon(x) > 0$ se $x \in]-\epsilon, \epsilon[$, mentre $f_\epsilon(x) = 0$ se $|x| \geq \epsilon$. Con riferimento alla definizione **??**, possiamo prendere $K = [-\epsilon, \epsilon]$. Chiaramente, f_ϵ è sommabile e $\int_\mathbf{R} f_\epsilon(x)dx \in \mathbf{R}^+$. Ponendo

$$\phi_\epsilon(x) := \frac{f_\epsilon(x)}{\int_\mathbf{R} f_\epsilon(y)dy}, \tag{3.2.3}$$

otteniamo un esempio di elemento di $\mathcal{D}(\mathbf{R})$ nullo al di fuori di $[-\epsilon, \epsilon]$ e con integrale uguale a 1.

Esempio 3.2.2 Siano a, b, c, d numeri reali, con $a < c < d < b$. Sia poi $\epsilon > 0$, tale che $a < c - 2\epsilon$ e $d + 2\epsilon < b$. Consideriamo la funzione $f :]a, b[\to \mathbf{R}$,

$$f(x) := \int_{c-\epsilon}^{d+\epsilon} \phi_\epsilon(y - x)dy,$$

con ϕ_ϵ definita in (**??**). Si può verificare che è lecito derivare sotto il segno di integrale e che $f \in C^\infty(]a, b[)$. Osserviamo ora che, se $x \in [c, d]$, poichè ϕ_ϵ è nulla al di fuori di $[-\epsilon, \epsilon]$, si ha che $\phi(y - x)$ è nulla al di fuori di $[c - \epsilon, d + \epsilon]$. Dunque

$$f(x) = \int_{c-\epsilon}^{d+\epsilon} \phi_\epsilon(y - x)dy = \int_\mathbf{R} \phi_\epsilon(y - x)dy = \int_\mathbf{R} \phi_\epsilon(y)dy = 1.$$

Se invece $x < c - 2\epsilon$, $\phi_\epsilon(y - x) = 0 \ \forall y \in [c - \epsilon, d + \epsilon]$. Dunque, in tal caso, $f(x) = 0$. Analogamente si vede che $f(x) = 0$ se $x > d + 2\epsilon$.

Riepilogando, $f \in \mathcal{D}(]a, b[)$, in quanto è identicamente nulla al di fuori di $[c - 2\epsilon, d + 2\epsilon]$ e vale 1 identicamente in $[c, d]$.

È chiaro che $\mathcal{D}(\Omega) \subseteq L^1(\Omega) \cap L^2(\Omega)$, in quanto, se $f \in \mathcal{D}(\Omega)$ e f è nulla al di fuori di K chiuso e limitato, vale

$$\int_\Omega |f(x)|dx \leq \max_K |f| L_n(K) < +\infty$$

perché ogni insieme limitato e misurabile ha misura di Lebesgue finita (vedi l'esercizio **??**). Da ultimo è utile sapere che $\mathcal{D}(\Omega)$ è denso in $L^p(\Omega)$, per $p \in \{1, 2\}$:

Teorema 3.2.1 *Sia Ω un aperto in $\mathbf{R}^n$. Allora $\mathcal{D}(\Omega)$ è denso negli spazi $L^1(\Omega)$ e $L^2(\Omega)$.*

Esercizio 3.2.1 Verificare che ogni sottoinsieme limitato e misurabile A di $\mathbf{R}^n$ ha misura finita. (Sugg.: verificare preliminarmente che esiste $L > 0$ tale che $A \subseteq [-L, L]^n$)

Esercizio 3.2.2 Siano, per $n \in \mathbf{N}$ e $j \in \{1, ..., n\}$, $a_j < c_j < d_j < b_j$. Costruire $f \in \mathcal{D}(\prod_{j=1}^n]a_j, b_j[$ identicamente uguale a 1 su $\prod_{j=1}^n [c_j, d_j]$.

3.3 La trasformata di Fourier in $L^2(\mathbf{R}^n)$

Abbiamo visto nella sezione **??** che la trasformata di Fourier può essere definita immediatamente tramite un integrale nello spazio $L^1(\mathbf{R}^n)$. Tuttavia, in tale spazio essa non gode di proprietà ottimali. Abbiamo visto, così , che non è a valori in $L^1(\mathbf{R}^n)$ (vedi l'esempio **??**), mentre la formula di inversione (teorema **??**) vale sotto ipotesi un po' macchinose. Vogliamo allora definire la trasformata di Fourier in $L^2(\mathbf{R}^n)$. Se, da una parte, la definizione è un po' più complicata, dall'altra in questo nuovo ambito la trasformata gode di proprietà assai migliori.

Cominciamo col seguente

Lemma 3.3.1 *Siano $f, g \in \mathcal{D}(\mathbf{R}^n)$. Allora:*
(I) $\hat{f} \in L^1(\mathbf{R}^n)$;
(II)

$$\int_{\mathbf{R}^n} f(x)\overline{g(x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}dx; \tag{3.3.1}$$

(formula di Parseval);
(III) $\hat{f} \in L^2(\mathbf{R}^n)$ e

$$\|\hat{f}\|_2 = (2\pi)^{\frac{n}{2}} \|f\|_2, \tag{3.3.2}$$

ove indichiamo con $\|.\|_2$ *la norma in* $L^2(\mathbf{R}^n)$.

Dimostrazione Poniamo

$$\Delta f := D_1^2 f + ... + D_n^2 f.$$

(ricordiamo che Δ è l'operatore di Laplace, vedi la sezione 2.6). Per ogni $m \in \mathbf{N}$ $\Delta^m f \in \mathcal{D}(\mathbf{R}^n)$. In base al corollario **??** (I), si ha, per $\xi \in \mathbf{R}^n$,

$$\mathcal{F}(\Delta^m f)(\xi) = (-1)^m \|\xi\|^{2m} \hat{f}(\xi).$$

Ciò implica che $\forall m \in \mathbf{N}$ esiste $C(m) \geq 0$ tale che

$$|\hat{f}(\xi)| \leq C(m)\|\xi\|^{-m}$$

$\forall \xi \in \mathbf{R}^n \backslash \{0\}$. Da ciò segue che $\hat{f}$ è sommabile in $\{\xi \in \mathbf{R}^n : \|\xi\| \geq 1\}$. Poiché è chiaramente sommabile in $\{\xi \in \mathbf{R}^n : \|\xi\| \leq 1\}$, possiamo concludere che $\hat{f}$ è sommabile e vale quindi la formula di inversione (**??**). Abbiamo allora

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

$$= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{\hat{f}}(-x)\overline{g(x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{\hat{f}}(x)\overline{g(-x)}dx.$$

Poniamo $h(x) := \overline{g(-x)}$. Per $\xi \in \mathbf{R}^n$, si ha

$$\hat{h}(\xi) = \int_{\mathbf{R}^n} e^{-i<x,\xi>}\overline{g(-x)}dx =$$

$$= \overline{\int_{\mathbf{R}^n} e^{-i<x,\xi>} g(x)dx} = \overline{\hat{g}(\xi)}.$$

Segue allora dal teorema **??** che

$$\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{\hat{f}}(x)\overline{g(-x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi.$$

Abbiamo perciò provato la formula di Parseval (**??**).

Il punto (III) segue immediatamente da (II), prendendo $g = f$. $\square$

Siamo ora in grado di provare il seguente

Lemma 3.3.2 *Sia $f \in L^2(\mathbf{R}^n)$. Consideriamo una successione $(f_k)_{k\in\mathbf{N}}$ in $\mathcal{D}(\mathbf{R}^n)$ tale che $\lim_{k\to\infty} f_k = f$ in $L^2(\mathbf{R}^n)$. Allora la successione $(\hat{f}_k)_{k\in\mathbf{N}}$* è *convergente in $L^2(\mathbf{R}^n)$ e il limite non dipende dalla successione scelta (purché, ovviamente, convergente a f).*

Dimostrazione In base al teorema **??**, esiste una successione $(f_k)_{k\in\mathbf{N}}$ in $\mathcal{D}(\mathbf{R}^n)$ tale che $\lim_{k\to\infty} f_k = f$ in $L^2(\mathbf{R}^n)$. Consideriamo la successione $(\hat{f}_k)_{k\in\mathbf{N}}$. Poiché $L^2(\mathbf{R}^n)$ è completo, per verificare che è convergente, basterà verificare che è di Cauchy. Sia allora $\epsilon > 0$; poiché $(f_k)_{k\in\mathbf{N}}$ è di Cauchy, esiste $n(\epsilon) \in \mathbf{N}$ tale che, se j e k sono interi maggiori di $n(\epsilon)$, vale $\|f_j - f_k\|_2 < \frac{\epsilon}{(2\pi)^{\frac{n}{2}}}$. Segue allora dal lemma **??** (III) che

$$\|\hat{f}_k - \hat{f}_j\|_2 = \|\mathcal{F}(f_k - f_j)\|_2 < \epsilon.$$

Verifichiamo ora che il limite della successione $(\hat{f}_k)_{k\in\mathbf{N}}$ non dipende dalla scelta di $(f_k)_{k\in\mathbf{N}}$. Sia $(g_k)_{k\in\mathbf{N}}$ un'altra successione in $\mathcal{D}(\mathbf{R}^n)$ tale che $\lim_{k\to\infty} g_k = f$ in $L^2(\mathbf{R}^n)$. Indichiamo provvisoriamente con l_1 e l_2 i limiti rispettivamente delle successioni $(\hat{f}_k)_{k\in\mathbf{N}}$ e $(\hat{g}_k)_{k\in\mathbf{N}}$. Dalla continuità della norma (esercizio 1.2.6) segue che

$$\|l_1 - l_2\|_2 = \lim_{k\to\infty} \|\hat{f}_k - \hat{g}_k\|_2$$

$$= \lim_{k\to\infty} \|\mathcal{F}(f_k - g_k)\|_2 = (2\pi)^{\frac{n}{2}} \lim_{k\to\infty} \|f_k - g_k\|_2 = (2\pi)^{\frac{n}{2}} \|f - f\|_2 = 0.$$

Dunque $l_1 = l_2$. □

Possiamo ora definire la trasformata di Fourier di $f \in L^2(\mathbf{R}^n)$:

Definizione 3.3.1 *Sia $f \in L^2(\mathbf{R}^n)$. Poniamo*

$$\mathcal{F}f = \hat{f} := \lim_{k\to\infty} \mathcal{F}f_k,$$

ove $(f_k)_{k\in\mathbf{N}}$ è *un'arbitraria successione in $\mathcal{D}(\mathbf{R}^n)$ convergente a f in $L^2(\mathbf{R}^n)$.*

Nel prossimo teorema elenchiamo le principali proprietà della trasformata di Fourier in $L^2(\mathbf{R}^n)$.

Teorema 3.3.1 (I) *L'applicazione $\mathcal{F} : f \to \mathcal{F}f$ di cui alla definizione* **??** è *lineare da $L^2(\mathbf{R}^n)$ in sé;*

*(II) per f e g in $L^2(\mathbf{R}^n)$ vale la formula di Parseval (**??**);*

(III) $\forall f \in L^2(\mathbf{R}^n)$ *vale l'identità* (**??**)*;*

(IV) $\mathcal{F}$ è *continua da* $L^2(\mathbf{R}^n)$ *in sé;*

(V) $\mathcal{F}$ è *una biiezione di* $L^2(\mathbf{R}^n)$ *in sé; l'applicazione inversa* $\mathcal{F}^{-1}$ *verifica la formula*

$$\mathcal{F}^{-1}f(\xi) = \frac{1}{(2\pi)^n}\hat{f}(-\xi) \qquad q.d.. \tag{3.3.3}$$

□

Dimostrazione parziale Lasciamo al lettore la dimostrazione di (I) (si tenga conto del risultato dell'esercizio 1.3.3).

(II) segue facilmente dal risultato dell'esercizio **??**. Infatti, se f e g sono elementi di $L^2(\mathbf{R}^n)$, esistono due successioni $(f_k)_{k\in\mathbf{N}}$ e $(g_k)_{k\in\mathbf{N}}$ a valori in $\mathcal{D}(\mathbf{R}^n)$ tali che si ha $\lim_{k\to+\infty} f_k = f$ e $\lim_{k\to+\infty} g_k = g$ in $L^2(\mathbf{R}^n)$. Allora

$$\int_{\mathbf{R}^n} f(x)\overline{g(x)}dx =< f,g >=$$

$$= \lim_{k\to+\infty} < f_k, g_k >= \lim_{k\to+\infty}\int_{\mathbf{R}^n} f_k(x)\overline{g_k(x)}dx =$$

$$= \lim_{k\to+\infty}\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} \hat{f}_k(\xi)\overline{\hat{g_k}(\xi)}d\xi =$$

$$= \lim_{k\to+\infty}\frac{1}{(2\pi)^n} < \hat{f}_k, \hat{g_k} >= \frac{1}{(2\pi)^n} < \hat{f}, \hat{g} >=$$

$$= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi.$$

(III) segue immediatamente da (II).

(IV) segue da (III) e dal risultato dell'esercizio **??**.

Proviamo (V). Per far vedere che $\mathcal{F}$ è iniettiva, basta provare che il suo nucleo è costituito esclusivamente dallo zero di $L^2(\mathbf{R}^n)$. Ciò segue immediatamente da (III). Per quanto riguarda la suriettività , sia $f \in L^2(\mathbf{R}^n)$ e $(f_k)_{k\in\mathbf{N}}$ la solita successione in $\mathcal{D}(\mathbf{R}^n)$ tale che $\lim_{k\to+\infty} f_k = f$ in $L^2(\mathbf{R}^n)$. Dalla formula di inversione (**??**), si ha

$$f_k(x) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i<x,\xi>}\hat{f}_k(\xi)d\xi =$$

$$= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{-i<x,\xi>}\hat{f}_k(-\xi)d\xi,$$

da cui $f_k = \mathcal{F}(\frac{1}{(2\pi)^n}\hat{f}_k(-.))$. Dalla continuità di $\mathcal{F}$ segue subito che $f = \mathcal{F}(\frac{1}{(2\pi)^n}\hat{f}(-.))$. Perciò $\mathcal{F}$ è una biiezione. Inoltre, $\forall k \in \mathbf{N}$ si ha $\mathcal{F}^{-1}f_k = \frac{1}{(2\pi)^n}\hat{f}_k(-.)$, per cui si ottiene (**??**) passando al limite in $L^2(\mathbf{R}^n)$ per $k \to +\infty$. □

Osservazione 3.3.1 Se $f \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$, disponiamo di due definizioni della trasformata di Fourier per f: la **??** e la **??**. Non sarebbe difficile verificare che in questo caso portano allo stesso risultato.

Esempio 3.3.1 Dall'esempio **??**, se $f(x) = \chi_+(x)e^{-x}$, si ha $\hat{f}(\xi) = g(\xi) = \frac{1}{1+i\xi}$. È facile verificare che f e g appartengono a $L^2(\mathbf{R})$. Vale $\hat{g} = \hat{\hat{f}}$.

Dal teorema **??** sappiamo che

$$f(x) = \frac{1}{2\pi}\hat{\hat{f}}(-x) \qquad q.d.$$

Dunque $\hat{g}(x)$ coincide quasi dappertutto con

$$(2\pi)^n f(-x) = \begin{cases} (2\pi)^n e^x & \text{se} \quad x \leq 0, \\ 0 & \text{se} \quad x > 0. \end{cases}$$

Esercizio 3.3.1 Sia H uno spazio di Hilbert, con prodotto scalare $< .,. >$. Siano poi $(x_n)_{n\in\mathbf{N}}$ e $(y_n)_{n\in\mathbf{N}}$ successioni in H tali che $\lim_{n\to+\infty} x_n = x$, $\lim_{n\to+\infty} y_n = y$. Provare che vale $\lim_{n\to+\infty} < x_n, y_n > = < x, y >$.

Esercizio 3.3.2 Completare la dimostrazione del teorema **??**.

3.4 Derivate deboli, convoluzione e trasformata di Fourier

Cominciamo con qualche risultato e definizione preliminare.

Definizione 3.4.1 *Siano Ω un aperto in $\mathbf{R}^n$, $f : \Omega \to \mathbf{C}$ misurabile. Diremo che $[f] \in L^1_{loc}(\Omega)$ (o che $[f]$ è* **localmente sommabile** *in Ω), se per ogni $H \subseteq \Omega$, con H chiuso e limitato, $f_{|H}$ è sommabile su H.*

Osservazione 3.4.1 Riguardo alla definizione **??**, come al solito parleremo di "f localmente sommabile", invece che di "$[f]$ localmente sommabile". Non è difficile verificare che questa definizione è invariante per modificazioni di f su insiemi di misura nulla.

Osservazione 3.4.2 È quasi ovvio che $L^1(\Omega) \subseteq L^1_{loc}(\Omega)$. Non è poi difficile verificare che, se $f \in L^2(\Omega)$, allora $f \in L^1_{loc}(\Omega)$. Infatti, dalla solita disuguaglianza (1.4.3) si ha

$$|f(x)| = |f(x)| \cdot 1 \leq \frac{1}{2}(|f(x)|^2 + 1) \qquad \forall x \in \Omega.$$

Perciò , se $H \subseteq \Omega$, con H chiuso e limitato, si ha

$$\int_H |f(x)|dx \leq \frac{1}{2}\int_H (|f(x)|^2 + 1)dx$$
$$\leq \frac{1}{2}(\int_\Omega |f(x)|^2 dx + L_n(H)) < +\infty.$$

È facile anche vedere che, se $f \in C(\Omega)$, allora $[f] \in L^1_{loc}(\Omega)$ (qui è il caso di distinguere f da $[f]$). Infatti, se $H \subseteq \Omega$, con H chiuso e limitato, si ha

$$\int_H |f(x)|dx \leq \max_H |f| L_n(H) < +\infty.$$

Vediamo ora (senza dimostrazione) un risultato estremamente utile.

Teorema 3.4.1 *Siano Ω un aperto di $\mathbf{R}^n$, f e g localmente sommabili in Ω. Allora sono equivalenti:*
(a) $f = g$;
(b) $\forall v \in \mathcal{D}(\Omega)$ si ha

$$\int_\Omega f(x)v(x)dx = \int_\Omega g(x)v(x)dx.$$

Dopo questi preparativi, siamo pronti per dare la definizione di derivata debole.

Definizione 3.4.2 *Siano Ω un aperto in $\mathbf{R}^n$, $f \in L^1_{loc}(\Omega)$ α un multiindice. Chiameremo $\alpha-$***derivata debole** *di f una funzione localmente sommabile g tale che $\forall v \in \mathcal{D}(\Omega)$ vale*

$$\int_\Omega g(x)v(x)dx = (-1)^{|\alpha|}\int_\Omega f(x)D^\alpha v(x)dx. \tag{3.4.1}$$

Osservazione 3.4.3 Una conseguenza immediata del teorema **??** è che, se f ammette derivata debole $\alpha-$derivata debole, questa è univocamente determinata q. d. . La indicheremo con $\partial^\alpha f$.

La definizione **??** è motivata dal seguente

Teorema 3.4.2 *Siano* Ω *aperto in* $\mathbf{R}^n$ *e* $f \in C^m(\Omega)$. *Allora, se* $\alpha \in \mathbf{N}_0^n$ *e* $|\alpha| \leq m$, f *ammette derivata debole* $\partial^\alpha f$ *e* $\partial^\alpha f(x) = D^\alpha f(x)$ *q. d.* .

Dimostrazione parziale Si tratta di far vedere che vale

$$\int_\Omega D^\alpha f(x)v(x)dx = (-1)^{|\alpha|} \int_\Omega f(x)D^\alpha v(x)dx$$

$\forall v \in \mathcal{D}(\Omega)$. Verifichiamo il risultato nel solo caso $n = 1$, $\Omega =]a,b[$, con $-\infty \leq a < b \leq +\infty$, e $\alpha = 1$.

Sia $v \in \mathcal{D}(\Omega)$. Allora esistono c e d reali, tali che $a < c < d < b$ e $v(x) = 0$ se $x \notin [c,d]$. Si osservi che valgono necessariamente $v(c) = v(d) = 0$ e $v'(x) = 0$ se $x \notin [c,d]$. Allora

$$\int_\Omega f'(x)v(x)dx = \int_c^d f'(x)v(x)dx =$$

$$= [f(x)v(x)]_{x=c}^{x=d} - \int_c^d f(x)v'(x)dx$$

$$= -\int_\Omega f(x)v'(x)dx.$$

□

Vediamo ora un esempio di esistenza della derivata debole senza che esista in tutti i punti la corrispondente derivata classica.

Esempio 3.4.1 Sia $f : \mathbf{R} \to \mathbf{R}$, $f(x) = |x|$. f non è derivabile in senso classico in 0. Verifichiamo che tuttavia ammette derivata (prima) debole ∂f. Tale derivata coincide quasi dappertutto con la funzione "segno" (sgn). Infatti, sia $v \in \mathcal{D}(\mathbf{R})$. Sia $v(x) \neq 0$ per $x \notin [c,d]$, per certi c e d reali. Possiamo supporre $c < 0 < d$. Allora

$$-\int_\mathbf{R} |x|v'(x)dx = -\int_c^d |x|v'(x)dx =$$

$$= \int_c^0 xv'(x)dx - \int_0^d xv'(x)dx$$

$$= [xv(x)]_{x=c}^{x=0} - \int_c^0 v(x)dx - [xv(x)]_{x=0}^{x=d} + \int_0^d v(x)dx =$$

$$= -\int_{[c,0[} v(x)dx + \int_{]0,d]} v(x)dx =$$

$$= \int_{\mathbf{R}} sgn(x)v(x)dx.$$

Veniamo ora all'interazione tra derivate deboli e trasformata di Fourier. Tenuto conto del corollario **??**(I), il seguente risultato diventa piuttosto naturale:

Teorema 3.4.3 *Siano $f \in L^2(\mathbf{R}^n)$ e $\alpha \in \mathbf{N}_0^n$. Allora, se f possiede derivata debole $\partial^\alpha f$ e tale derivata appartiene a $L^2(\mathbf{R}^n)$, si ha $\mathcal{F}(\partial^\alpha u) = (i\xi)^\alpha \mathcal{F} f$.*

Viceversa, supponiamo che $(i\xi)^\alpha \hat{f} \in L^2(\mathbf{R}^n)$. Allora f ammette derivata debole $\partial^\alpha f$. Inoltre, $\partial^\alpha f \in L^2(\mathbf{R}^n)$ e $\partial^\alpha f = \mathcal{F}^{-1}((i\xi)^\alpha \hat{f})$.

Dimostrazione Proviamo solo la seconda parte.

Poniamo $g := \mathcal{F}^{-1}((i\xi)^\alpha \hat{f})$. Sia poi $v \in \mathcal{D}(\mathbf{R}^n)$. Utilizzando la formula di Parseval e il risultato dell'esercizio **??**, otteniamo

$$\int_{\mathbf{R}^n} g(x)v(x)dx = \int_{\mathbf{R}^n} g(x)\overline{\overline{v(x)}}dx =$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} (i\xi)^\alpha \hat{f}(\xi)\hat{v}(-\xi)d\xi =$$

$$= (-1)^{|\alpha|} \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{f}(\xi)(-i\xi)^\alpha \hat{v}(-\xi)d\xi =$$

$$= (-1)^{|\alpha|} \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \hat{f}(\xi)\mathcal{F}(D^\alpha v)(-\xi)d\xi,$$

utilizzando il corollario **??**(I). Ancora per il risultato dell'esercizio **??** e la formula di Parseval, quest'ulimo integrale coincide con

$$(-1)^{|\alpha|} \int_{\mathbf{R}^n} f(x)D^\alpha v(x)dx,$$

da cui la conclusione. $\square$

Passiamo ora alla nozione di convoluzione.

Definizione 3.4.3 *Siano f e g elementi di $L^1_{loc}(\mathbf{R}^n)$. Diremo che f e g sono* **convolubili** *se per quasi ogni $x \in \mathbf{R}^n$ la funzione $y \to f(x-y)g(y)$ è sommabile in $\mathbf{R}^n$.*

In tal caso, chiamiamo **convoluzione** *di f e g e indichiamo col simbolo $f * g$, la funzione*

$$\begin{cases} f * g : \mathbf{R}^n \to \mathbf{C}, \\ (f*g)(x) = \begin{cases} \int_{\mathbf{R}^n} f(x-y)g(y)dy \\ \textit{se } y \to f(x-y)g(y) \text{ è } \textit{sommabile}, \\ 0 \textit{ altrimenti.} \end{cases} \end{cases} \tag{3.4.2}$$

Vale il seguente risultato:

Teorema 3.4.4 *(I) Siano f e g elementi di $L^1(\mathbf{R}^n)$. Allora sono convolubili e la convoluzione $f * g$ appartiene a $L^1(\mathbf{R}^n)$. Inoltre*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1. \tag{3.4.3}$$

(II) Siano $f \in L^1(\mathbf{R}^n)$ e $g \in L^2(\mathbf{R}^n)$. Allora sono convolubili e la convoluzione $f * g$ appartiene a $L^2(\mathbf{R}^n)$. Inoltre

$$\|f * g\|_2 \leq \|f\|_1 \|g\|_2. \tag{3.4.4}$$

(III) In ciascuno dei casi (I) e (II) vale la formula

$$\mathcal{F}(f * g) = \mathcal{F}f \mathcal{F}g. \tag{3.4.5}$$

Dimostrazione parziale Siano f e g in $L^1(\mathbf{R}^n)$. Dando per scontato che sono convolubili e che la convoluzione è misurabile, si ha, applicando il teorema di Tonelli,

$$\int_{\mathbf{R}^n} |(f*g)(x)|dx \leq \int_{\mathbf{R}^n} (\int_{\mathbf{R}^n} |f(x-y)||g(y)|dy)dx =$$

$$= \int_{\mathbf{R}^n} |g(y)|(\int_{\mathbf{R}^n} |f(x-y)|dx)dy = \int_{\mathbf{R}^n} |g(y)|(\int_{\mathbf{R}^n} |f(z)|dz)dy =$$
$$= \|f\|_1 \|g\|_1.$$

Inoltre, per ogni $\xi \in \mathbf{R}^n$,

$$\mathcal{F}(f*g)(\xi) = \int_{\mathbf{R}^n} e^{-i<x,\xi>} (\int_{\mathbf{R}^n} f(x-y)g(y)dy)dx =$$

$$= \int_{\mathbf{R}^n} (\int_{\mathbf{R}^n} e^{-i<x-y,\xi>} f(x-y) e^{-i<y,\xi>} g(y)dy)dx =$$

$$= \int_{\mathbf{R}^n} (\int_{\mathbf{R}^n} e^{-i<x-y,\xi>} f(x-y)dx) e^{-i<y,\xi>} g(y)dy =$$
$$= \int_{\mathbf{R}^n} (\int_{\mathbf{R}^n} e^{-i<z,\xi>} f(z)dz) e^{-i<y,\xi>} g(y)dy =$$
$$\mathcal{F}f(\xi)\mathcal{F}g(\xi).$$

Tralasciamo il caso $g \in L^2(\mathbf{R}^n)$. $\square$

Esercizio 3.4.1 Siano Ω un aperto in $\mathbf{R}^n$ e α e β multiindici. Supponiamo che esistano in $L^1_{loc}(\Omega)$ le derivate deboli $\partial^\alpha f$ e $\partial^\beta(\partial^\alpha f)$. Verificare che f ammette la derivata debole $\partial^{\alpha+\beta} f$ e vale

$$\partial^{\alpha+\beta} f = \partial^\beta(\partial^\alpha f).$$

Esercizio 3.4.2 Dimostrare che l'operazione di convoluzione in $L^1(\mathbf{R}^n)$ gode delle proprietà commutativa e associativa. Queste proprietà consentono, date f_1,...,f_m in $L^1(\mathbf{R}^n)$, di considerare la convoluzione $f_1 * ... * f_m$, ove l'ordine con cui le operazioni sono svolte non ha rilevanza per il risultato finale.

Sugg: usare il fatto che la trasformata di Fourier della convoluzione è il prodotto delle trasformate di Fourier e l'iniettività di $\mathcal{F}$.

3.5 Alcune applicazioni della trasformata di Fourier a problemi di equazioni differenziali

La trasformata di Fourier è utile per lo studio di molti problemi nei campi (ad esempio) delle derivate parziali e della probabilità . Qui presentiamo alcune semplici applicazioni a equazioni a derivate parziali.

Esempio 3.5.1 (**Equazione di Helmoltz in $\mathbf{R}^n$**) Sia $\lambda \in \mathbf{C}$. Consideriamo il problema

$$(\lambda - \Delta)u(x) = f(x), \qquad x \in \mathbf{R}^n, \tag{3.5.1}$$

ove indichiamo con Δ l'operatore di Laplace

$$\Delta := \sum_{j=1}^{n} D_j^2 = D_1^2 + ... + D_n^2. \tag{3.5.2}$$

Considereremo il caso $\lambda \in \mathbf{R}^+$. Introduciamo alcune notevoli classi di (classi di equivalenza di) funzioni:

Definizione 3.5.1 *Sia $m \in \mathbf{N}$. Indichiamo con $H^m(\mathbf{R}^n)$ l'insieme degli elementi f in $L^2(\mathbf{R}^n)$ dotati di derivata debole $\partial^\alpha f$ appartenente a $L^2(\mathbf{R}^n)$ per ogni multiindice α tale che $|\alpha| \leq m$.*

Queste classi di funzioni sono un caso particolare dei così detti "spazi di Sobolev", che tanta importanza hanno assunto nell'analisi moderna.

Vale il seguente

Teorema 3.5.1 *Sia $\lambda \in \mathbf{R}^+$. Allora per ogni $f \in L^2(\mathbf{R}^n)$ il problema (**??**) possiede un'unica soluzione u nella classe $H^2(\mathbf{R}^n)$. La u risolve il problema nel senso che*

$$\lambda u - \sum_{j=1} \partial_j^2 u = f,$$

ove abbiamo indicato con $\partial_j^2 u$ la corrispondente derivata debole di u.

Dimostrazione Per il teorema **??**, se esiste una soluzione u in $H^2(\mathbf{R}^n)$, deve valere

$$(\lambda + \|\xi\|^2)\mathcal{F}(u) = \mathcal{F}(f),$$

da cui

$$u = \mathcal{F}^{-1}((\lambda + \|\xi\|^2)^{-1}\mathcal{F}(f)). \tag{3.5.3}$$

Osserviamo che la u definita in (**??**) è un elemento di $L^2(\mathbf{R}^n)$. Infatti,

$$\int_{\mathbf{R}^n} |((\lambda + \|\xi\|^2)^{-1}\mathcal{F}(f)(\xi)|^2 d\xi$$

$$\leq \lambda^{-2} \int_{\mathbf{R}^n} |\mathcal{F}(f)(\xi)|^2 d\xi = \frac{(2\pi)^n}{\lambda^2} \int_{\mathbf{R}^n} |f(x)|^2 dx < +\infty$$

per il teorema **??**(III).

Inoltre $u \in H^2(\mathbf{R}^n)$. Per vederlo, basta applicare il teorema **??**. Infatti, sia $\alpha \in \mathbf{N}_0^n$, con $|\alpha| \leq 2$. Allora

$$(i\xi)^\alpha \mathcal{F}u(\xi) = \frac{(i\xi)^\alpha}{\lambda + \|\xi\|^2}\mathcal{F}f(\xi).$$

La funzione $\xi \to \frac{(i\xi)^\alpha}{\lambda + \|\xi\|^2}\mathcal{F}f(\xi)$ appartiene a $L^2(\mathbf{R}^n)$, in quanto

$$|\frac{(i\xi)^\alpha}{\lambda + \|\xi\|^2}\mathcal{F}f(\xi)| \leq \frac{\|\xi\|^{|\alpha|}}{\lambda + \|\xi\|^2}|\mathcal{F}f(\xi)|. \tag{3.5.4}$$

Il primo fattore nel secondo membro di (**??**) è limitato come funzione di ξ, in quanto $|\alpha| \leq 2$. Perciò

$$\frac{\|\xi\|^{|\alpha|}}{\lambda + \|\xi\|^2}|\mathcal{F}f(\xi)| \leq C|\mathcal{F}f(\xi)|,$$

per un certo C positivo.

Dunque, se $|\alpha| \leq 2$, u ha derivata debole in $L^2(\mathbf{R}^n)$ e appartiene quindi a $H^2(\mathbf{R}^n)$.

Esempio 3.5.2 (**Problema di Cauchy per l'equazione del calore o equazione di diffusione**) Cerchiamo una funzione u delle variabili (t, x), con $t \geq 0$, $x \in \mathbf{R}^n$, soddisfacente le seguenti condizioni:

$$\begin{cases} D_t u(t,x) = \Delta_x u(t,x), & t > 0, x \in \mathbf{R}^n, \\ u(0,x) = f(x), & x \in \mathbf{R}^n. \end{cases} \tag{3.5.5}$$

ove abbiamo indicato con Δ_x l'operatore di Laplace rispetto alle variabili $x = (x_1, ..., x_n)$.

Eseguiamo i calcoli formalmente, ponendo

$$U(t,\xi) := \int_{\mathbf{R}^n} e^{-i<x,\xi>} u(t,x)dx,$$

per $t > 0$ e $\xi \in \mathbf{R}^n$. Applichiamo, sostanzialmente, in maniera formale la trasformata rispetto alle varibili x a $u(t,.)$ per ogni $t \geq 0$. Scambiando questa trasformata con la derivata temporale D_t, otteniamo

$$\begin{cases} D_t U(t,\xi) = -\|\xi\|^2 U(t,\xi), & t > 0, \xi \in \mathbf{R}^n, \\ U(0,\xi) = \hat{f}(\xi), & \xi \in \mathbf{R}^n. \end{cases} \tag{3.5.6}$$

da cui

$$U(t,\xi) = e^{-t\|\xi\|^2}\hat{f}(\xi), \qquad t \geq 0, \xi \in \mathbf{R}^n. \tag{3.5.7}$$

Continuando a procedere formalmente, ricordiamo che, in base al teorema **??**(III), il prodotto è la trasformata di Fourier della convoluzione. Calcoliamo allora l'antitrasformata di $\xi \to e^{-t\|\xi\|^2}$, per $t > 0$. Si ha, per $x \in \mathbf{R}^n$,

$$\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i<x,\xi>} e^{-t\|\xi\|^2} d\xi =$$

$$= \frac{1}{(2\pi)^n}\prod_{j=1}^n \int_{\mathbf{R}} e^{ix_j\xi_j} e^{-t\xi_j^2} d\xi_j = \frac{1}{(4\pi^2 t)^{\frac{n}{2}}}\prod_{j=1}^n \int_{\mathbf{R}} e^{i\frac{x_j\eta_j}{\sqrt{t}}} e^{-\eta_j^2} d\eta_j =$$

$$= \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{\|x\|^2}{4t}},$$

applicando il risultato dell'esempio 2.5.6.

Abbiamo allora ottenuto la seguente soluzione "formale" del problema (**??**):

$$u(t,x) = \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbf{R}^n} e^{-\frac{\|x-y\|^2}{4t}} f(y)dy & \text{se} \quad t > 0, \\ f(x) & \text{se} \quad t = 0. \end{cases} \tag{3.5.8}$$

Enunciamo adesso un risultato preciso. Premettiamo la definizione di soluzione classica:

Definizione 3.5.2 *Sia $f \in C(\mathbf{R}^n)$. Per* **soluzione classica** *di (**??**) intendiamo una funzione u di dominio $[0,+\infty[\times\mathbf{R}^n$ continua, dotata in $]0,+\infty[\times\mathbf{R}^n$ delle derivate $D_t u$, e $D_x^\alpha u$ per $|\alpha| \le 2$, continue in $]0,+\infty[\times\mathbf{R}^n$.*

Vale il seguente

Teorema 3.5.2 *Sia $f \in BC(\mathbf{R}^n)$. Allora il problema (**??**) possiede un'unica soluzione u classica tale che $u(t,.)$ è limitata in $\mathbf{R}^n$ $\forall t \ge 0$.*

*Tale soluzione u ammette la rappresentazione (**??**).*

Ci limitiamo a osservare che, se $f \in BC(\mathbf{R}^n)$, la funzione in (**??**) è ben definita, in quanto per ogni $t > 0$ e per ogni $x \in \mathbf{R}^n$ la funzione $y \to e^{-\frac{\|x-y\|^2}{4t}} f(y)$ è sommabile in $\mathbf{R}^n$.

Esempio 3.5.3 (**Problema di Cauchy per l'equazione delle onde in dimensione spaziale** 1) Cerchiamo una funzione u delle variabili (t,x), con $t \ge 0$, $x \in \mathbf{R}$, soddisfacente le seguenti condizioni:

$$\begin{cases} D_t^2 u(t,x) = D_x^2 u(t,x), & t > 0, x \in \mathbf{R}, \\ u(0,x) = f(x), & x \in \mathbf{R}, \\ D_t u(0,x) = g(x), & x \in \mathbf{R}. \end{cases} \tag{3.5.9}$$

Eseguiamo i calcoli formalmente, ponendo, al solito,

$$U(t,\xi) := \int_{\mathbf{R}} e^{-ix\xi} u(t,x)dx,$$

per $t > 0$ e $\xi \in \mathbf{R}$. Applichiamo formalmente la trasformata rispetto alla varibile x a $u(t,.)$ per ogni $t \ge 0$. Scambiando questa trasformata con la derivata temporale D_t^2, otteniamo

$$\begin{cases} D_t^2 U(t,\xi) = -\xi^2 U(t,\xi), & t > 0, \xi \in \mathbf{R}, \\ U(0,\xi) = \hat{f}(\xi), & \xi \in \mathbf{R}, \\ D_t U(0,\xi) = \hat{g}(\xi), & \xi \in \mathbf{R}, \end{cases} \tag{3.5.10}$$

da cui, almeno per $\xi \neq 0$,

$$U(t,\xi) = \cos(t\xi)\hat{f}(\xi) + \frac{\sin(t\xi)}{\xi}\hat{g}(\xi). \tag{3.5.11}$$

Vogliamo adesso antitrasformare. Procedendo sempre formalmente, si ha

$$\frac{1}{2\pi}\int_{\mathbf{R}} e^{ix\xi}\cos(t\xi)\hat{f}(\xi)d\xi$$

$$= \frac{1}{2}(\frac{1}{2\pi}\int_{\mathbf{R}} e^{i(x+t)\xi}\hat{f}(\xi)d\xi + \frac{1}{2\pi}\int_{\mathbf{R}} e^{i(x-t)\xi}\hat{f}(\xi)d\xi) =$$

$$= \frac{f(x+t)+f(x-t)}{2}.$$

Quanto al secondo addendo nel secondo membro di (**??**), è facile verificare che, se $a > 0$ e χ_a è la funzione caratteristica di $[-a,a]$, si ha, per $\xi \neq 0$,

$$\hat{\chi_a}(\xi) = 2\frac{\sin(a\xi)}{\xi}.$$

Dunque, almeno formalmente, l'antitrasformata di Fourier di $\frac{\sin(t\xi)}{\xi}\hat{g}(\xi)$ è

$$\frac{1}{2}(\chi_t * g)(x) = \frac{1}{2}\int_{x-t}^{x+t} g(y)dy.$$

In conclusione, abbiamo ottenuto

$$u(t,x) = \frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy. \tag{3.5.12}$$

La formula (**??**) è nota come **formula di D'Alembert**.

Procediamo ora in maniera simile a quanto fatto nel caso dell'equazione del calore.

Definizione 3.5.3 *Siano* $f, g \in C(\mathbf{R}^n)$. *Per* **soluzione classica** *di (**??**) intendiamo una funzione* u *in* $C([0,+\infty[\times\mathbf{R}^n)$ *, dotata in* $]0,+\infty[\times\mathbf{R}^n$ *delle derivate continue* $D_t u$,$D_t^2 u$ *e* $D_x^\alpha u$ *per* $|\alpha| \leq 2$*, con* $D_t u$ *prolungabile con continuità a* $[0,+\infty[\times\mathbf{R}^n$.

Vale il seguente

Teorema 3.5.3 *Siano $f \in C^2(\mathbf{R}^n)$, $g \in C^1(\mathbf{R}^n)$. Allora il problema (**??**) possiede un'unica soluzione u classica.*

*Tale soluzione u ammette la rappresentazione (**??**).*

Osservazione 3.5.1 Al contrario di ciò che accade per l'equazione del calore (vedi il teorema **??**), per l'equazione delle onde vale un risultato di esistenza e unicità senza alcuna ipotesi di limitatezza o di crescenza per i dati del problema.

Esercizio 3.5.1 Scrivere esplicitamente sotto forma di integrale la soluzione del problema (**??**) nel caso unidimensionale. Osservare che la soluzione non è unica se si ammettono soluzioni non appartenenti a $L^2(\mathbf{R})$.

Esercizio 3.5.2 Studiare l'equazione del calore non omogenea

$$\begin{cases} D_t u(t,x) = \Delta_x u(t,x) + f(t,x), & t > 0, x \in \mathbf{R}^n, \\ u(0,x) = u_0(x), & x \in \mathbf{R}^n. \end{cases} \tag{3.5.13}$$

Ricavare, applicando formalmente la trasformata di Fourier rispetto alle variabili spaziali x, la soluzione

$$u(t,x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbf{R}^n} e^{-\frac{\|x-y\|^2}{4t}} u_0(y)dy + \int_0^t \frac{1}{(4\pi(t-s))^{\frac{n}{2}}} (\int_{\mathbf{R}^n} e^{-\frac{\|x-y\|^2}{4(t-s)}} f(s,y)dy)ds. \tag{3.5.14}$$

Si potrebbe dimostrare che, se $u_0 \in BC(\mathbf{R}^n)$, $f \in BC([0,+\infty[\times\mathbf{R}^n)$ e ammette derivata $D_t f \in BC([0,+\infty[\times\mathbf{R}^n)$, allora (**??**) è l'unica soluzione classica di (**??**) (nel senso della definizione **??**) tale che per ogni $t \geq 0$ $u(t,.)$ è limitata in $\mathbf{R}^n$.

Esercizio 3.5.3 Verificare che, se $f \in C^2(\mathbf{R}^n)$ e $g \in C^1(\mathbf{R}^n)$, allora (**??**) è effettivamente soluzione classica di (**??**).

Capitolo 4

Elementi di calcolo delle probabilità

4.1 Spazi di probabilità

Nello studio di vari fenomeni in ambito naturale, scientifico, fisico, sociale, ecc. si presentano spesso delle situazioni in cui non si è in grado di predire esattamente (in maniera deterministica) come si svilupperà una certa situazione. In molti casi di tale genere si parla allora di fenomeni di tipo aleatorio o probabilistico.

Nella creazione di modelli matematici in questo ambito, si è dimostrato conveniente il metodo seguente: si considera l'insieme Ω di tutti i possibili risultati di un certo esperimento (qui la parola "esperimento" va intesa in un senso molto lato). Si associa poi a ogni elemento di una certa classe di situazioni concrete un opportuno sottoinsieme di Ω. Si individua così una famiglia $\mathcal{A}$ di sottoinsiemi di Ω. A ciascun elemento A di questa famiglia si attribuisce poi un numero reale $P(A)$ compreso tra 0 e 1, chiamato "probabiltà di A", che indica in che misura ci si può aspettare il verificarsi della situazione concreta corrispondente ad A. Naturalmente, le singole situazioni che si possono presentare sono correlate tra loro. Ciò si ripercuote nella richiesta che la misura di probabilità P soddisfi alcune proprietà di carattere generale.

Inoltre, le operazioni insiemistiche standard nella classe $\mathcal{A}$ ammettono delle interpretazioni naturali a livello di situazioni concrete. Ad esempio, l'intersezione in $\mathcal{A}$ rappresenterà il contemporaneo verificarsi dei casi concreti corrispondenti, mentre l'unione costituirà un modello del verificarsi di almeno uno dei casi in questione.

Vediamo ora qualche esempio.

Esempio 4.1.1 Supponiamo di voler costruire un modello dell'esperimento costituito dal lancio di un dado con le facce numerate da 1 a 6 e perfettamente equilibrato. L'insieme di tutti i possibili risultati è

$$\Omega := \{1, 2, 3, 4, 5, 6\}.$$

Allora si può , ad esempio, identificare la situazione concreta "il risultato è un numero pari" con l'insieme $A := \{2, 4, 6\}$ e il caso "il risultato è non superiore a 4" con $B = \{1, 2, 3, 4\}$. $A \cup B = \{1, 2, 3, 4, 6\}$ corrisponderà a "il risultato è un numero pari o non superiore a 4", mentre $A \cap B = \{2, 4\}$ rappresenterà "il risultato è un numero pari non superiore a 4".

Esempio 4.1.2 Supponiamo di avere a disposizione un certo dispositivo (per esempio elettronico) e di voler considerare il primo istante in cui il dispositivo si guasta a partire dalla sua accensione. Se indichiamo con 0 l'istante di accensione, potremo rappresentare il risultato con un numero reale non negativo, che indicherà l'istante in cui il dispositivo si guasta espresso in un'opportuna unità di misura temporale. Prendiamo allora come insieme Ω di tutti i possibili risultati sperimentali l'intervallo $[0, +\infty[$. L'insieme $[1, 2]$ rappresenterà allora il fatto che l'apparecchio si guasta in qualche istante tra 1 e 2.

Si accennava in precedenza al fatto che le principali operazioni insiemistiche ammettono un'interpretazione naturale a livello di situazioni concrete. Allora, se indichiamo con $\mathcal{A}$ la classe dei sottoinsiemi di Ω di cui vogliamo definire una probabilità , richiederemo che $\mathcal{A}$ sia chiusa rispetto a tali operazioni. L'ipotesi che si fa in generale è che $\mathcal{A}$ costituisca una $\sigma-$algebra. Sarà comodo nel seguito porre, dato $A \subseteq \Omega$, con Ω "insieme ambiente" fissato dal contesto, porre

$$A^c := \Omega \setminus A. \tag{4.1.1}$$

Chiameremo A^c **complementare** di A in Ω.

Definizione 4.1.1 *Sia Ω un insieme. Indichiamo con $\mathcal{P}(\Omega)$ l'***insieme potenza** *di Ω, vale a dire, l'insieme dei sottoinsiemi di Ω. Sia $\mathcal{A} \subseteq \mathcal{P}(\Omega)$. Diremo che $\mathcal{A}$ è una $\sigma-$algebra se soddisfa le seguenti condizioni:*

(I) $\Omega \in \mathcal{A}$*;*

(II) se $A_n \in \mathcal{A}\ \forall n \in \mathbf{N}$*, allora* $\bigcup_{n\in\mathbf{N}} A_n \in \mathcal{A}$*;*

(III) se $A \in \mathcal{A}$*, allora* $A^c \in \mathcal{A}$*.*

Dalla definizione segue il fatto che una $\sigma-$algebra è chiusa rispetto alle principali operazioni insiemistiche:

Teorema 4.1.1 *Sia* $\mathcal{A}$ *una* $\sigma-$*algebra nell'insieme* Ω*. Allora:*

(I) $\emptyset \in \mathcal{A}$*;*

(II) se $n \in \mathbf{N}$ *e* $A_i \in \mathcal{A}$ *per* $i = 1, ..., n$*, allora* $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$*;*

(III) se $\mathcal{I} = \mathbf{N}$ *oppure* $\mathcal{I} = \{1, ..., n\}$ *per qualche* $n \in \mathbf{N}$ *e* $A_i \in \mathcal{A}\ \forall i \in \mathcal{I}$*, allora* $\bigcap_{i\in\mathcal{I}} A_i \in \mathcal{A}$*;*

(IV) se A *e* B *sono elementi di* $\mathcal{A}$*, lo è anche* $A \setminus B$*.*

Dimostrazione parziale (I) Basta osservare che $\emptyset = \Omega^c$ e utilizzare (I) e (III) nella definizione **??**.

(II) segue da (II) della definizione **??** e da (I) osservando che, se poniamo $A_i := \emptyset$ per ogni $i \in \mathbf{N}$ con $i > n$, si ha

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i\in\mathbf{N}} A_i.$$

Omettiamo la dimostrazione di (III) (vedi l'esercizio **??**).

(IV) segue dal fatto che

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

e dal punto (III). $\square$

Consideriamo adesso il numero $P(A)$ che dovrebbe esprimere la "probabilità " del verificarsi della situazione concreta corrispondente ad $A \in \mathcal{A}$. P è chiaramente una funzione da $\mathcal{A}$ a $[0, 1]$. Chiederemo che P sia una misura di probabilità , nel senso seguente:

Definizione 4.1.2 *Siano* Ω *un insieme non vuoto e* $\mathcal{A}$ *una* $\sigma-$*algebra in* Ω*. Una* **misura di probabilità** *su* $\mathcal{A}$ *è una funzione* $P : \mathcal{A} \to [0, 1]$ *tale che:*

(I) $P(\Omega) = 1$*;*

(II) se $A_n \in \mathcal{A}\ \forall n \in \mathbf{N}$ *e gli* A_n *sono a due a due disgiunti per valori distinti di* n*, allora*

$$P(\bigcup_{n\in\mathbf{N}} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Anche qui dalla definizione si possono ricavare ulteriori proprietà delle misure di probabilità :

Teorema 4.1.2 *Siano Ω un insieme non vuoto, $\mathcal{A}$ una $\sigma-$algebra in Ω, P una misura di probabilità in $\mathcal{A}$. Allora:*

(I) $P(\emptyset) = 0$;

(II) se $n \in \mathbf{N}$, $A_i \in \mathcal{A}$ per $i = 1, ..., n$ e gli A_i sono a due a due disgiunti, allora

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i);$$

(III) se A e B sono elementi di $\mathcal{A}$ e $A \subseteq B$, si ha

$$P(B \setminus A) = P(B) - P(A).$$

Dimostrazione (I) Applicando la proprietà (II) della definizione **??** con $A_n = \emptyset \ \forall n \in \mathbf{N}$, si ottiene subito che da $P(\emptyset) > 0$ segue $P(\emptyset) = +\infty$. Ciò è incompatibile con la definizione di misura di probabilità .

(II) segue subito da (I) e da (II) della definizione **??**, osservando (come già fatto) che, se poniamo $A_i = \emptyset$ per $i > n$, si ha $\bigcup_{i \in \mathbf{N}} A_i = \bigcup_{i=1}^{n} A_i$.

(III) Si ha $B = A \cup (B \setminus A)$ e $A \cap (B \setminus A) = \emptyset$. Segue allora da (II) che

$$P(B) = P(A) + P(B \setminus A),$$

da cui la conclusione. $\square$

Osservazione 4.1.1 Dal (III) del teorema **??** segue che, se $A \subseteq B$, allora $P(A) \leq P(B)$.

Possiamo ora dare la definizione di spazio di probabilità :

Definizione 4.1.3 *Uno* **spazio di probabilità** *è una terna $(\Omega, \mathcal{A}, P)$, ove Ω è un insieme non vuoto, $\mathcal{A}$ è una $\sigma-$algebra in Ω, P è una misura di probabilità di dominio $\mathcal{A}$.*

Passiamo adesso ad alcuni esempi concreti.

Esempio 4.1.3 (**Spazi di probabilità finiti** Sia Ω un insieme con un numero finito ($n \in \mathbf{N}$) di elementi. Siano questi $\{\omega_1, ..., \omega_n\}$. Prendiamo come σ-algebra la totalità $P(\Omega)$ dei sottoinsiemi di Ω. Vogliamo definire una misura di probabilità su $P(\Omega)$. Cominciamo allora a considerare gli insiemi $\{\omega_1\}$, ...,$\{\omega_n\}$ con un solo elemento. Se attribuiamo a $\{\omega_j\}$ ($1 \leq j \leq n$) la probabilità $p_j \in [0,1]$, la proprietà (II) del teorema **??** impone che

$$1 = P(\Omega) = p_1 + ... + p_n. \tag{4.1.2}$$

Viceversa, se la condizione (**??**) è soddisfatta, non è difficile provare che esiste una e una sola misura di probabilità su $P(\Omega)$ tale che $P(\{\omega_j\}) = p_j$ per $j = 1, ..., n$. Dovrà valere, evidentemente, se $A \subseteq \Omega$,

$$P(A) = \sum_{\omega_j \in A} p_j. \tag{4.1.3}$$

Riconsideriamo, così , l'esempio **??**. Si tratterà di attribuire una probabilità p_j al fatto che il risultato del lancio sia j, con $1 \leq j \leq 6$. Le condizioni da soddisfare sono $p_j \in [0,1]$ per $1 \leq j \leq n$, e $\sum_{j=1}^{6} p_j = 1$. È chiaro allora che, se riteniamo che il dado sia ben equilibrato e garantisca quindi l'equiprobabilità dei singoli risultati, l'unica possibilità è porre $p_j = \frac{1}{6}$ per $j = 1, ..., 6$. Se invece riteniamo che questa condizione non sia soddisfatta, attribuiremo valori p_j eventualmente diversi tra loro, conservando però la condizione (**??**).

Nel caso di un dado bilanciato, per ogni $A \subseteq \{1, ..., 6\}$ la condizione (**??**) impone

$$P(A) = \frac{\sharp(A)}{6},$$

ove $\sharp(A)$ indica la cardinalità (numero di elementi) dell'insieme A.

Un altro esempio nello stesso ordine di idee è il seguente: consideriamo il risultato del lancio di due dadi equilibrati. L'insieme Ω dei risultati può essere schematizzato mediante l'insieme di coppie ordinate $\{(i,j) : 1 \leq i, j \leq 6\}$ ed è quindi costituito da 36 elementi. Se $A \subseteq \Omega$, otteniamo uno spazio di probabilità ponendo

$$P(A) := \frac{\sharp(A)}{36}.$$

Ad esempio, consideriamo l'evento "il punteggio totale è 2". Allora $A = \{(1,1)\}$, da cui $P(A) = \frac{1}{36}$. Se indichiamo invece con B l'evento "il punteggio totale è 7", avremo $B = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$, da cui $P(B) = \frac{6}{36} = \frac{1}{6}$.

Esempio 4.1.4 Costruiamo ora un modello probabilistico relativo all'esempio **??**. Poniamo $\Omega := \mathbf{R}^+$ e indichiamo con $\mathcal{A}$ la classe dei sottoinsiemi di Ω misurabili secondo Lebesgue. Dal teorema 0.1.1 segue subito che $\mathcal{A}$ è una $\sigma-$algebra. Quanto alla misura di probabilità , fissiamo $g : \mathbf{R}^+ \to [0, +\infty[$ misurabile, tale che

$$\int_{\mathbf{R}^+} g(t)dt = 1 \tag{4.1.4}$$

e poniamo, dato $A \in \mathcal{A}$,

$$P(A) := \int_A g(t)dt. \tag{4.1.5}$$

Non è difficile verificare che P è una misura di probabilità . Qui ci limitiamo a dimostrare la proprietà (II) della definizione **??**. Sia $(A_n)_{n\in\mathbf{N}}$ una successione di elementi di $\mathcal{A}$ a due a due disgiunti. Indichiamo con χ_n la funzione caratteristica di A_n. Evidentemente, se $A := \bigcup_{n\in\mathbf{N}} A_n$, si ha che per ogni $x \in \Omega$ la funzione caratteristica χ_A di A soddisfa

$$\chi_A(x) = \sum_{n=1}^{\infty} \chi_n(x).$$

Dunque

$$P(A) = \int_A g(t)dt = \int_{\mathbf{R}^+} g(t)\chi_A(t)dt =$$
$$= \int_{\mathbf{R}^+} \sum_{n=1}^{\infty} g(t)\chi_n(t)dt =$$

(applicando il risultato dell'esercizio 0.5.1)

$$= \sum_{n=1}^{\infty} \int_{\mathbf{R}^+} g(t)\chi_n(t)dt = \sum_{n=1}^{\infty} \int_{A_n} g(t)dt =$$
$$= \sum_{n=1}^{\infty} P(A_n).$$

Per esempio, sia $\lambda > 0$. Poniamo $g(t) := \lambda e^{-\lambda t}$. Con questa scelta, se A è l'evento "la apparecchio si guasta in qualche istante tra 1 e 2 (nell'unità di misura temporale prefissata), identificheremo A con l'intervallo $]1, 2[$. Dunque

$$P(A) = \int_{]1,2[} \lambda e^{-\lambda t}dt = e^{-\lambda} - e^{-2\lambda}.$$

Esercizio 4.1.1 Dimostrare il punto (III) del teorema **??**. Il punto cruciale è costituito dalla formula

$$\bigcap_{i\in\mathcal{I}} A_i = (\bigcup_{i\in\mathcal{I}} A_i^c)^c. \tag{4.1.6}$$

Esercizio 4.1.2 Costruire un modello probabilistico del doppio lancio di un dado non truccato. Calcolare la probabilità di ottenere almeno un 6.

Esercizio 4.1.3 Siano A e B due eventi nello spazio di probabilità (Ω, $\mathcal{A}$, P). Verificare che

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

4.2 Elementi di calcolo combinatorio

In questa sezione introduciamo alcuni semplici elementi di calcolo combinatorio, che permettono di determinare la cardinalità di certi insiemi finiti, allo scopo di calcolare la probabilità di alcuni eventi nel caso di Ω finito.

Il risultato principale è il seguente

Teorema 4.2.1 *Siano N e K insiemi finiti con (rispettivamente) n e k elementi. Qui n e k sono numeri naturali e $k \leq n$. Allora:*
(I) il numero delle applicazioni iniettive da K a N è $\frac{n!}{(n-k)!}$;
(II) il numero di sottoinsiemi di N con k elementi è $\binom{n}{k}$, con

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Dimostrazione Per quanto riguarda (I), sia $K = \{a_1, ..., a_k\}$. Volendo elencare tutte le applicazioni $f : K \to N$ iniettive, possiamo scegliere $f(a_1)$ in n modi diversi, $f(a_2)$ in $n-1$ modi diversi ($f(a_2)$ può essere un elemento qualunque diverso da $f(a_1)$), ..., $f(a_k)$ in $n-k+1$ modi diversi. Perciò il numero di applicazioni iniettive da K a N è $n(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$.

Quanto a (II), sia K' un generico sottoinsieme di N con k elementi. Le funzioni iniettive da K a N la cui immagine è K' sono evidentemente le biiezioni da K a K'. Applicando (I), possiamo dire che sono $k!$. Dunque, moltiplicando $k!$ per il numero di sottoinsiemi di N con k elementi, otterremo la cardinalità dell'insieme delle funzioni iniettive da K a N, vale a dire $\frac{n!}{(n-k)!}$. Di qui la conclusione. $\square$

Esempio 4.2.1 Qual è la probabilità di azzeccare una terzina vincente al lotto con una singola giocata ?

Su ogni singola ruota vengono estratti settimanalmente cinque numeri su un totale di 90. Il numero complessivo delle terzine coincide con il numero di sottoinsiemi di tre elementi di un insieme di 90 elementi ed è pari (in base al teorema **??**) a $\binom{90}{3}$. Ciascuna terzina ha ovviamente la stessa probabilità di essere estratta. Il numero delle terzine vincenti è chiaramente $\binom{5}{3} = 10$. Dunque la probabilità di azzeccare una terzina vincente è

$$\frac{10}{\binom{90}{3}} = \frac{10 \times 3! \times 87!}{90!} \cong 8,5 \times 10^{-5}.$$

Esempio 4.2.2 Un'urna contiene 5 palline rosse e 10 palline bianche. Ne vengono estratte a caso 5 senza reimbussolamento. Qual è la probabilità di estrarre esattamente 3 palline rosse?

Sia Ω la famiglia dei sottoinsiemi di 5 elementi dell'insieme delle $10+5 = 15$ palline. Si ha $\sharp(\Omega) = \binom{15}{5}$. Il numero degli elementi di Ω con 3 palline rosse e 2 bianche è $\binom{5}{3}\binom{10}{2}$. Dunque la probabilità cercata è

$$\frac{\binom{5}{3}\binom{10}{2}}{\binom{15}{5}} \cong 0,15.$$

Concludiamo questa sezione con una formula ben nota:

Corollario 4.2.1 *(Formula del binomio di Newton) Siano a e b numeri complessi ed $n \in \mathbf{N}$. Allora*

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^{n-j} b^j.$$

Dimostrazione Sviluppando il prodotto $(a+b) \cdot \ldots \cdot (a+b)$ (n fattori) mediante la scelta dell'uno o dell'altro tra a o b, si ottiene una somma contenente termini del tipo $c_j a^{n-j} b^j$, con $0 \leq j \leq n$. c_j indica quante sono le possibili scelte ottenute prendendo $n-j$ volte a e j volte b. Ora, a ciascuna scelta si può far corrispondere la famiglia dei fattori in cui si è preso b, e dunque un sottoinsieme di j elementi di un insieme di n elementi. Concludiamo perciò, applicando il teorema **??**(II) che $c_j = \binom{n}{j}$, da cui la conclusione.

Esercizio 4.2.1 Calcolare la probabilità che, prese n persone a caso ($2 \leq n \leq 365$) almeno due di esse festeggino il compleanno nello stesso giorno.

(Sugg.: trascurare eventuali nascite il 29 febbraio di un anno bisestile. Cercare di calcolare la probabilità dell'evento complementare. Abbastanza sorprendentemente, si può vedere che già con sole 23 persone la probabilità cercata è superiore a $\frac{1}{2}$!)

4.3 Probabilità condizionata e indipendenza

Definizione 4.3.1 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità,* A *e* B *elementi di* $\mathcal{A}$*, con* $P(A) > 0$*. Definiamo la* **probabilità condizionata** $P(B|A)$ *di* B *dato* A *come segue:*

$$P(B|A) := \frac{P(B \cap A)}{P(A)}.$$

Osservazione 4.3.1 Intuitivamente, $P(B|A)$ misura la probabilità del verificarsi di B, nell'ipotesi che si verifichi A. Si osservi che, se A e B sono incompatibili tra loro (vale a dire, $A \cap B = \emptyset$), si ha $P(B|A) = 0$. Inoltre, $P(A|A) = 1$.

Esempio 4.3.1 Una popolazione è composta per il 40 per cento da fumatori (F), per il 60 per cento da non fumatori (N). Si sa che il 25 per cento dei fumatori è affetto da una certa malattia cronica delle vie respiratorie, mentre la stessa malattia colpisce solo il 7 per cento dei non fumatori. Qual è la probabilità che un individuo affetto dalla malattia sia un fumatore?

Dobbiamo calcolare $P(F|M) = \frac{P(F\cap M)}{P(M)}$. Sappiamo che $P(M|F) = \frac{1}{4}$. Si ha

$$P(M) = P(F \cap M) + P(N \cap M) =$$

$$= P(M|F)P(F) + P(M|N)P(N) = \frac{1}{4} \cdot \frac{2}{5} + \frac{7}{100} \cdot \frac{3}{5} = \frac{71}{500}.$$

Inoltre

$$P(F \cap M) = P(M|F)P(F) = \frac{1}{4} \cdot \frac{2}{5} = \frac{1}{10}.$$

Ne segue che

$$P(F|M) = \frac{\frac{1}{10}}{\frac{71}{500}} = \frac{50}{71} \cong 0,70.$$

Osservazione 4.3.2 Siano A_1,..., A_n e B eventi con probabilità positiva in $(\Omega, \mathcal{A}, P)$. Supponiamo che A_1,...,A_n costituiscano una partizione di Ω, nel senso che sono a due a due disgiunti e la loro unione è Ω. Supponiamo inoltre che $P(B) > 0$. Sia $j \in \{1, ..., n\}$. Osserviamo che B è unione disgiunta degli eventi $B \cap A_k$ $(1 \leq k \leq n)$. Allora

$$\begin{gathered} P(A_j|B) = \\ = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j \cap B)}{P(A_j)}\frac{P(A_j)}{P(B)} = P(B|A_j)\frac{P(A_j)}{\sum_{k=1}^{n} P(B \cap A_k)} = \\ = \frac{P(B|A_j)P(A_j)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}. \end{gathered} \tag{4.3.1}$$

L'identità ricavata in (**??**) costituisce la così detta **formula di Bayes**. Utilizzando questa formula, potevamo ricavare più rapidamente il risultato dell'esempio **??**. Si ha infatti

$$\begin{gathered} P(F|M) = \\ = \frac{P(M|F)P(F)}{P(M|F)P(F) + P(M|N)P(N)} = \\ = \frac{\frac{1}{4} \cdot \frac{2}{5}}{\frac{1}{4} \cdot \frac{2}{5} + \frac{7}{100} \cdot \frac{3}{5}} = \frac{50}{71}. \end{gathered}$$

La formula di Bayes è nota anche come formula di probabilità delle cause, perchè nelle applicazioni gli eventi A_j costituiscono potenziali cause del verificarsi di B e si vuole valutare qual è la causa più probabile.

Definizione 4.3.2 *Siano A e B eventi nello spazio di probabilità* $(\Omega, \mathcal{A}, P)$*. Diremo che A e B sono* **indipendenti** *se $P(A \cap B) = P(A)P(B)$.*

Osservazione 4.3.3 Se $P(A) > 0$, l'indipendenza tra A e B equivale a

$$P(B|A) = P(B).$$

Il senso della definizione **??** è il seguente: il verificarsi di A non muta la probabilità del verificarsi di B. È vero anche viceversa se $P(B) > 0$.

Esempio 4.3.2 Consideriamo un mazzo di 40 carte, con i soliti quattro semi. Ne estraiamo una a caso. Siano A l'evento "la carta estratta è un asso", B l'evento "la carta estratta è un denaro". Verifichiamo che A e B sono indipendenti.

Si ha $P(A) = \frac{4}{40} = \frac{1}{10}$, $P(B) = \frac{10}{40} = \frac{1}{4}$. $A \cap B$ è l'evento "viene estratto l'asso di denari". Quindi

$$P(A \cap B) = \frac{1}{40} = \frac{1}{10} \cdot \frac{1}{4} = P(A)P(B).$$

Dunque A e B sono indipendenti.

Esempio 4.3.3 Si consideri il lancio di tre dadi equilibrati. Dunque

$$\Omega = \{(i, j, k) : i, j, k \in \{1, 2, 3, 4, 5, 6\}\}.$$

Ogni singola tripla ordinata ha probabilità $\frac{1}{6^3} = \frac{1}{216}$.

Sia A l'evento "la somma $i + j + k$ è uguale a 6". A è identificabile con l'insieme di triple

$$\{(1, 1, 4), (1, 2, 3), (1, 3, 2), (1, 4, 1), (2, 1, 3), (2, 2, 2), \\ (2, 3, 1), (3, 1, 2), (3, 2, 1), (4, 1, 1)\}.$$

Quindi

$$P(A) = \frac{10}{216} = \frac{5}{108} \cong 0,046.$$

Sia B l'evento "i,j,k sono a due a due distinti". La cardinalità di B coincide con il numero di funzioni iniettive da un insieme di 3 elementi a un insieme di 6 elementi e vale, tenuto del teorema **??**,

$$\frac{6!}{(6-3)!} = \frac{6!}{3!} = 120.$$

Quindi

$$P(B) = \frac{120}{216} = \frac{5}{9} \cong 0,56.$$

Si ha

$$A \cap B = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}.$$

Quindi

$$P(A \cap B) = \frac{6}{6^3} = \frac{1}{36} \cong 0,028.$$

Invece,

$$P(A)P(B) = \frac{5}{108} \cdot \frac{5}{9} = \frac{25}{972} \cong 0,026.$$

Dunque, A e B non sono indipendenti. Da $P(A)P(B) < P(A \cap B)$, ricaviamo $P(B) < P(B|A)$. Ciò significa che il verificarsi di A favorisce il verificarsi di B.

Ha interesse definire l'indipendenza per famiglie costituite da più di due eventi:

Definizione 4.3.3 *Sia $\{A_i : i \in \mathcal{I}\}$ una famiglia di eventi in un certo spazio di probabilità $(\Omega, \mathcal{A}, P)$, dipendente dal parametro i in $\mathcal{I}$. Diremo che gli A_i sono indipendenti se, comunque si prendono $i_1, ..., i_n$ $(n \in \mathbf{N})$ in $\mathcal{I}$, a due a due distinti, si ha*

$$P(A_{i_1} \cap ... \cap A_{i_n}) = P(A_{i_1}) \cdot ... \cdot P(A_{i_n}).$$

Esempio 4.3.4 *(Processi di Bernoulli)* Vogliamo qui descrivere un semplice modello matematico applicabile a molte situazioni concrete.

Supponiamo di ripetere n volte un certo esperimento, sempre sotto le medesime condizioni. Introduciamo l'ipotesi che i risultati delle singole prove non si influenzino tra di loro. Siamo interessati solo a due aspetti complementari dell'esperimento, che possiamo definire "successo" (S) e "insuccesso" (I). Sia p $(\in [0,1])$ la probabilità che attribuiamo a S in ciascuna singola prova. Di conseguenza, attribuiremo a I la probabilità $q := 1 - p$. Poniamo allora

$$\Omega := \{\omega = \{\omega_1, ..., \omega_n) : \omega_1, ..., \omega_n \in \{S, I\}\}. \tag{4.3.2}$$

Osserviamo che Ω è finito ed è costituito da 2^n elementi. Seguendo lo schema generale dell'esempio **??** e ponendo quindi $\mathcal{A} = \mathcal{P}(\Omega)$, si tratta di attribuire una probabilità p_ω a ciascuna $n-$pla $\omega^0 = (\omega_1^0, ..., \omega_n^0)$, in modo tale che $\sum\limits_{\omega \in \Omega} p_\omega = 1$. Per $j = 1, ..., n$, poniamo

$$A_j := \{\omega \in \Omega : \omega_j = \omega_j^0\}. \tag{4.3.3}$$

È naturale porre

$$P(A_j) = \begin{cases} p & \text{se} \quad \omega_j^0 = S, \\ q & \text{se} \quad \omega_j^0 = I. \end{cases} \tag{4.3.4}$$

Osservando che

$$\{\omega^0\} = A_1 \cap ... \cap A_n$$

e tenendo conto che, poiché le singole prove non si influenzano tra loro, A_1,...,A_n andranno supposti indipendenti, poniamo

$$p_{\omega^0} := P(A_1) \cdot ... \cdot P(A_n) = p^m q^{n-m}, \tag{4.3.5}$$

ove m rappresenta il numero dei successi nella sequenza $\omega_1^0, ..., \omega_n^0$.

Verifichiamo adesso che, con la posizione **??**, $\sum\limits_{\omega \in \Omega} p_\omega = 1$. A tale scopo, osserviamo, innanzi tutto, che le $n-$ple contenenti esattamente m successi

$(0 \leq m \leq n)$ possono essere messe in corrispondenza biunivoca con i sottoinsiemi di m elementi di un insieme di n elementi (basta far corrispondere a ciascuna sequenza ω l'insieme $\{i_1, ..., i_m\}$, con $1 \leq i_1 < ... < i_m \leq n$ tale che $\omega_j = S$ se $j \in \{i_1, ..., i_m\}$). Perciò , in base al teorema **??** (II), le $n-$ple contenenti esattamente m successi sono $\binom{n}{m}$. Ne segue, applicando la formula del binomio di Newton, che

$$\sum_{\omega \in \Omega} p_\omega = \sum_{m=0}^{n} \binom{n}{m} p^m q^{n-m} = (q+p)^n = 1^n = 1.$$

Concludendo, avremo, per $A \subseteq \Omega$,

$$P(A) = \sum_{\omega \in A} p_\omega. \tag{4.3.6}$$

Con la posizione (**??**) si potrebbe allora verificare che vale (**??**) e che gli eventi A_j $(1 \leq j \leq n)$ definiti in (**??**) sono indipendenti (si veda per questo gli esercizi **??** e **??**).

Osservazione 4.3.4 Consideriamo ancora lo schema descritto nell'esempio **??**. Per $0 \leq m \leq n$, indichiamo con B_m l'evento "nella serie di prove si ottengono esattamente m successi". Per quanto visto, si ha

$$P(B_m) = \binom{n}{m} p^m q^{n-m}. \tag{4.3.7}$$

Ci chiediamo: qual è il numero di successi più probabile, o, in altri termini, per quali m $P(B_m)$ è massimo? Per rispondere a questa domanda, consideriamo la disequazione

$$P(B_m) \leq P(B_{m+1}), \quad 0 \leq m \leq n-1. \tag{4.3.8}$$

Applicando la formula (**??**), è facile verificare che la condizione (**??**) è equivalente a

$$m \leq np - q, \tag{4.3.9}$$

e che vale la disuguaglianza stretta se $m < np - q$. Osserviamo preliminarmente che, poiché $0 \leq p \leq 1$ e $q = 1 - p$, $np - q \leq n$ e $np - q = n$ se e solo se $p = 1$ e $q = 0$. Inoltre:

(I) se $np - q < 0$, si ha $P(B_0) > P(B_1) > ... > P(B_n)$; il risultato più probabile è $m = 0$;

(II) supponiamo $0 \leq np-q < n$ e $np-q \notin \mathbf{Z}$; indichiamo con m_0 la parte intera di $np-q$ $(0 \leq m_0 < n)$; allora $P(B_0) < ... < P(B_{m_0}) < P(B_{m_0+1}) > ... > P(B_n)$; in questo caso, il risultato più probabile è $m_0 + 1$;

(III) supponiamo $0 \leq np-q < n$ e $np-q = m_0 \in \mathbf{Z}$; allora $P(B_0) < ... < P(B_{m_0}) = P(B_{m_0+1}) > ... > P(B_n)$; in questo caso, i risultati più probabili sono (alla pari) m_0 e $m_0 + 1$;

(IV) supponiamo $np - q = n$; come già osservato, ciò equivale a $p = 1$ e $q = 0$; in tale caso $P(B_0) = ... = P(B_{n-1}) = 0$, mentre $P(B_n) = 1$; evidentemente, il risultato più probabile è $m = n$.

Esempio 4.3.5 Presentiamo un primo esempio di processo di Bernoulli.

Supponiamo di lanciare 50 volte una moneta perfettamente equilibrata. Consideriamo "successo" il risultato "testa" (T), insuccesso il risultato "croce" (C). Allora avremo $n = 50$, $p = q = \frac{1}{2}$. Se $0 \leq m \leq 50$, la probabilità di ottenere esattamente m teste (in 50 lanci) vale $\binom{50}{m}2^{-50}$. In questo caso,

$$np - q = 50 \cdot \frac{1}{2} - \frac{1}{2} = \frac{49}{2} = 24,5.$$

In base alle considerazioni dell'osservazione **??**, il numero di teste più probabile è 25. La probabilità di ottenere esattamente 25 teste vale

$$\binom{50}{25}2^{-50} \cong 0,11.$$

Esercizio 4.3.1 Dimostrare che, con la posizione (**??**), vale (**??**).

Sugg.: supponiamo che, ad esempio, $\omega_j^0 = S$. Per $i = 1, ..., n$, indichiamo con C_i il sottoinsieme degli elementi ω tali che $\omega_j = S$ e $\omega_k = S$ per i elementi k della sequenza $1, ..., n$. Se $\omega \in C_i$, si ha $p_\omega = p^i q^{n-i}$. Si può verificare che C_i ha $\binom{n-1}{i-1}$ elementi. Di conseguenza

$$\begin{aligned} P(A_j) &= \textstyle\sum_{i=1}^{n} \binom{n-1}{i-1} p^i q^{n-i} &&= \textstyle\sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} q^{n-1-k} \\ &= p(q+p)^{n-1} &&= p. \end{aligned}$$

Esercizio 4.3.2 Dimostrare che gli eventi $A_1, ..., A_n$ definiti in (**??**) sono indipendenti.

Sugg.: si tratta di far vedere che, se $1 \leq j_1 < ... < j_r \leq n$ $(2 \leq r \leq n)$, si ha

$$P(A_{j_1} \cap ... \cap A_{j_r}) = P(A_{j_1}) \cdot ... \cdot P(A_{j_r}) = p^s q^{r-s},$$

ove s indica la cardinalità di $\{j \in \{j_1, ..., j_r\} : \omega_j^0 = S\}$. Per $i = s, ..., n - (r - s) = n - r + s$, indichiamo con C_i il sottoinsieme degli elementi $\omega \in \Omega$ tali che $\omega_j = \omega_j^0$ per $j \in \{j_1, ..., j_r\}$ e $\omega_j = S$ per i elementi j della sequenza $1, ..., n$. Se $\omega \in C_i$, si ha $p_\omega = p^i q^{n-i}$. Si può verificare che C_i ha $\binom{n-r}{i-s}$ elementi. Di conseguenza

$$\begin{aligned} P(A_{j_1} \cap ... \cap A_{j_r}) &= \textstyle\sum_{i=s}^{n-r+s} \binom{n-r}{i-s} p^i q^{n-i} &&= \textstyle\sum_{k=0}^{n-r} \binom{n-r}{k} p^{k+s} q^{n-s-k} \\ &= p^s q^{r-s} (q+p)^{n-r} &&= p^s q^{r-s}. \end{aligned}$$

Esercizio 4.3.3 Si giocano alla roulette i numeri 3, 13, 22. Supponiamo di sapere che il gioco è truccato e che il numero vincente è sempre dispari. Qual è la probabilità che esca uno dei numeri giocati? Si tenga presente che i risultati possibili (a parte l'imbroglio) sono i numeri interi da 0 a 36.

Esercizio 4.3.4 Cinque monete false sono mescolate con 9 monete autentiche. Se ne estrae una a caso.

(I) Calcolare la probabilità che venga estratta una moneta falsa.

Se si estraggono due monete (senza reimbussolamento), calcolare la probabilità che

(II) una sia falsa e una sia autentica;
(III) siano entrambe false;
(IV) siano entrambe autentiche.

Esercizio 4.3.5 Un'urna contiene A pietre bianche e B pietre nere. Una seconda urna contiene C pietre bianche e D nere ($A, B, C, D \in \mathbf{N}$). Si estrae una pietra dalla prima urna e la si trasferisce nella seconda. Poi si estrae una pietra dalla seconda. Calcolare la probabilità dei seguenti eventi:

(I) la prima pietra estratta è bianca;
(II) la prima pietra estratta è nera;
(III) la seconda estratta è bianca, se la prima estratta è bianca;
(IV) la seconda estratta sia bianca, se la prima estratta è nera.

Esercizio 4.3.6 Si estraggono due pietre con reimbussolamento da un'urna che ne contiene 4 rosse e 2 bianche. Calcolare la probabilità dei seguenti eventi:

(I) entrambe le pietre estratte sono bianche;
(II) entrambe le pietre estratte sono rosse;
(III) le pietre estratte hanno lo stesso colore;
(IV) almeno una delle pietre estratte è rossa.

Esercizio 4.3.7 Il senato americano comprende due senatori per ciascuno dei 50 stati. Si estrae a sorte un comitato di 50 senatori. Qual è la probabilità che comprenda entrambi i senatori dell'Alaska?

Trovare la probabilità che comprenda entrambi i senatori dell'Alaska nell'ipotesi che ne comprenda almeno uno.

Esercizio 4.3.8 Lanciamo due volte un dado bilanciato. Siano A, B, C gli eventi:

A: "il risultato del primo lancio è dispari";

B: "il risultato del secondo lancio è dispari":

C: "la somma dei risultati è dispari".

Verificare che A, B, C, sono a due a due indipendenti, ma non complessivamente indipendenti.

Esercizio 4.3.9 Un correttore di bozze esamina un testo contenente 20 errori. Ogni volta che incappa in un errore, la probabilità che lo riconosca è uguale a $\frac{3}{4}$.

Calcolare il numero più probabile di errori non corretti dopo una lettura.

Calcolare il medesimo valore dopo due letture.

Esercizio 4.3.10 I bulloni prodotti da una ditta sono difettosi con probabilità 1/10 e vengono messi in commercio in confezioni da quattro pezzi ciascuna. Qual è la probabilità che in una confezione ci siano più di due bulloni difettosi?

Esercizio 4.3.11 Sapendo che il 20 per cento dei passeggeri che hanno prenotato non si presenta alla partenza, una compagnia aerea accetta fino a 55 prenotazioni su un volo con la capienza di 50 posti. Qual è la probabilità che (almeno) un passeggero che ha prenotato e si presenta alla partenza debba restare a terra?

Esercizio 4.3.12 Un sacchetto contiene 9 monete con testa e croce e una moneta con due teste. Ne estraiamo una a caso, la lanciamo 6 volte e otteniamo 6 volte testa. Qual è la probabilità di aver estratto la moneta con due teste?

Esercizio 4.3.13 Un test per individuare una malattia risulta positivo nel 99 % dei casi se applicato a una persona affetta, nel 2% dei casi se applicato a una persona non affetta. Tenuto conto che la percentuale di persone affette è valutata intorno allo 0,1% della popolazione, calcolare la probabilità che una persona sia affetta se il test risulta positivo.

4.4 Variabili aleatorie

Definizione 4.4.1 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità,* $X : \Omega \to \mathbf{R}$. *Diremo che* X è *una* **variabile aleatoria semplice** *se esistono* A_1,...,A_m *(*$m \in \mathbf{N}$*) elementi di* $\mathcal{A}$ *a due a due incompatibili, la cui unione è* Ω*, e dei numeri reali* $\alpha_1, ..., \alpha_m$ *in modo che,*

$$X(\omega) = \alpha_i \quad \forall \omega \in A_i, \quad 1 \leq i \leq m.$$

Diremo che X è *una* **variabile aleatoria reale** *(v.a.r.) se esiste una successione di variabili aleatorie semplici* $(X_k)_{k \in \mathbf{N}}$ *di dominio* Ω*, tale che*

$$\lim_{k \to +\infty} X_k(\omega) = X(\omega) \qquad\qquad \forall \omega \in \Omega.$$

Osservazione 4.4.1 Le definizioni di variabile aleatoria semplice e di variabile aleatoria reale presentano una forte analogia con le definizioni di funzione semplice (definizione 0.2.1) e di funzione misurabile (definizione 0.2.3). Di fatto, se $\Omega \in \mathcal{M}_n$ per qualche $n \in \mathbf{N}$ e $\mathcal{A}$ è la classe dei sottoinsiemi misurabili secondo Lebesgue in Ω, le variabili aleatorie semplici sono proprio le funzioni semplici, le variabili aleatorie reali sono le funzioni misurabili.

Definizione 4.4.2 *Sia* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità. Una* **variabile aleatoria** $n-$**dimensionale** *su* Ω è *una funzione* $X : \Omega \to \mathbf{R}^n$*, con*

$$X(\omega) = (X_1(\omega), ..., X_n(\Omega)),$$

tale che ciascuna componente $X_j(\omega)$ *(*$1 \leq j \leq n$*) è una v.a.r..*

Esempio 4.4.1 Consideriamo il lancio di un dado a sei facce ben equilibrato. Siano $\Omega := \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $P(A) := \sharp(A)/6$, $\forall A \in \mathcal{A}$. Poniamo

$$X : \Omega \to \mathbf{R},$$
$$X(\omega) = \begin{cases} 0 & \text{se } \omega \text{ è pari,} \\ 1 & \text{se } \omega \text{ è dispari.} \end{cases}$$

X è una v.a. semplice, in quanto, se poniamo $A_1 := \{2, 4, 6\}$ e $A_2 = \{1, 3, 5\}$, A_1 e A_2 sono eventi incompatibili con unione Ω, $X(\omega) = 1$ se $\omega \in A_1$, $X(\omega) = 0$ se $\omega \in A_2$.

Esempio 4.4.2 Si lancia un dardo su un bersaglio circolare di raggio r ($r \in \mathbf{R}^+$). Schematizziamo l'esperimento ponendo

$$\Omega := \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \le r^2\},$$

$$\mathcal{A} := \{A \subseteq \Omega : A \in \mathcal{M}_2\}, \tag{4.4.1}$$

$$P(A) := \tfrac{L_2(A)}{\pi r^2}, \qquad A \in \mathcal{A}.$$

Questa scelta di P significa, evidentemente, che si ritiene che la probabilità che il dardo finisca in A è direttamente proporzionale all'area $L_2(A)$ di A.

È facile verificare che $(\Omega, \mathcal{A}, P)$ è uno spazio di probabilità . Dato $\omega \in \Omega$, indichiamo con $X(\omega)$ la distanza di ω dall'origine $(0, 0)$. Ovviamente, $X(\omega) = \|\omega\|$, ove abbiamo indicato con $\|.\|$ la norma euclidea. In base all'osservazione **??** e al teorema 0.2.1, X è una v.a.r..

Un altro esempio di v.a.r. collegata allo stesso esperimento è

$$\begin{gathered} Y : \Omega \to \mathbf{R}, \\ Y(\omega_1, \omega_2) = |\omega_1|, \end{gathered} \tag{4.4.2}$$

che misura la distanza del punto ω centrato dal dardo dall'asse ω_1.

Infine, se poniamo

$$\begin{gathered} Z : \Omega \to \mathbf{R}^2, \\ Z(\omega) = (X(\omega), Y(\omega)), \end{gathered} \tag{4.4.3}$$

otteniamo una variabile aleatoria bidimensionale.

Introduciamo ora alcune notazione: siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e $X : \Omega \to \mathbf{R}^n$; se $B \subseteq \mathbf{R}^n$, poniamo

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}. \tag{4.4.4}$$

Nel caso di X v.a.r., poniamo, dato $a \in \mathbf{R}$,

$$\{X < a\} := \{\omega \in \Omega : X(\omega) < a\}. \tag{4.4.5}$$

Analoghe notazioni saranno utilizzate per altri tipi di disuguaglianza.

Il prossimo teorema, di notevole inportanza, ha un contenuto puramente insiemistico:

Teorema 4.4.1 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità ,* $X : \Omega \to \mathbf{R}^n$ *una variabile aleatoria,* $\{B_i : i \in \mathcal{I}\}$ *una famiglia di sottoinsiemi di* $\mathbf{R}^n$ *dipendente dal parametro* $i \in \mathcal{I}$, $B \subseteq \mathbf{R}^n$. *Allora:*
(I) $\{X \in \cup_{i\in\mathcal{I}} B_i\} = \cup_{i\in\mathcal{I}}\{X \in B_i\}$;
(II) $\{X \in \cap_{i\in\mathcal{I}} B_i\} = \cap_{i\in\mathcal{I}}\{X \in B_i\}$;
(III) $\{X \in B^c\} = \{X \in B\}^c$.

Dimostrazione (I) Dire che $\omega \in \{X \in \cup_{i\in\mathcal{I}} B_i\}$ significa dire che $X(\omega) \in \cup_{i\in\mathcal{I}} B_i$ e quindi che esiste i_0 in $\mathcal{I}$ tale che $X(\omega) \in B_{i_0}$. Dunque $\omega \in \{X \in B_{i_0}\} \subseteq \cup_{i\in\mathcal{I}}\{X \in B_i\}$. Viceversa, se $\omega \in \cup_{i\in\mathcal{I}}\{X \in B_i\}$, esiste $i_0 \in \mathcal{I}$ tale che $\omega \in \{X \in B_{i_0}\}$. Dunque $X(\omega) \in B_{i_0}$ e quindi $X(\omega) \in \cup_{i\in\mathcal{I}} B_i$, da cui la conclusione.

(II) Dire che $\omega \in \{X \in \cap_{i\in\mathcal{I}} B_i\}$ significa dire che $X(\omega) \in \cap_{i\in\mathcal{I}} B_i$. Ciò equivale a $X(\omega) \in B_i$ per ogni $i \in \mathcal{I}$, a $\omega \in \{X \in B_i\}$ per ogni i e quindi a $\omega \in \cap_{i\in\mathcal{I}}\{X \in B_i\}$.

(III) $\omega \in \{X \in B^c\}$ se e solo se $X(\omega) \in B^c$. Ciò equivale a $\omega \in \Omega$ e $X(\omega) \notin B$, a sua volta equivalente a $\{X \in B\}^c$.
□

Dati una variabile aleatoria $n-$dimensionale X e $B \subseteq \mathbf{R}^n$, ci chiediamo ora se abbia senso parlare della probabilità di $\{X \in B\}$. Qui si presenta il seguente problema: sotto quali condizioni, dato $B \subseteq \mathbf{R}^n$, $\{X \in B\}$ è un evento, vale a dire, $\{X \in B\}$ appartiene alla $\sigma-$algebra $\mathcal{A}$? Vedremo che esiste una vasta classe di sottoinsiemi di $\mathbf{R}^n$ per cui ciò vale: i così detti "boreliani" in $\mathbf{R}^n$.

Cominciamo con un risultato preliminare:

Teorema 4.4.2 *Siano* Ω *e* $\mathcal{I}$ *insiemi,* $\mathcal{I} \neq \emptyset$ *e* $\forall i \in \mathcal{I}$ *sia* $\mathcal{A}_i$ *una* $\sigma-$*algebra in* Ω. *Allora* $\cap_{i\in\mathcal{I}}\mathcal{A}_i$ *è una* $\sigma-$*algebra in* Ω.

Dimostrazione La dimostrazione è veramente banale. Qui ci limitiamo a provare che, se $A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$ per ogni $n \in \mathbf{N}$, allora $\cup_{n\in\mathbf{N}} A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$. Sia $i \in \mathcal{I}$. Allora, per ogni $n \in \mathbf{N}$, $A_n \in \mathcal{A}_i$. Poiché $\mathcal{A}_i$ è una $\sigma-$algebra, $\cup_{n\in\mathbf{N}} A_n \in \mathcal{A}_i$. Ciò vale per ogni i. Quindi $\cup_{n\in\mathbf{N}} A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$. □

Definizione 4.4.3 *Sia* $n \in \mathbf{N}$*; Indichiamo con* $\mathcal{B}(\mathbf{R}^n)$ *l'intersezione di tutte le* $\sigma-$*algebre in* $\mathbf{R}^n$ *che contengono la classe degli insiemi aperti.*

Osservazione 4.4.2 In virtù del teorema **??**, $\mathcal{B}(\mathbf{R}^n)$ è una $\sigma-$algebra in $\mathbf{R}^n$ e può essere pensata come la più piccola $\sigma-$algebra in $\mathbf{R}^n$ che contiene tutti gli insiemi aperti. Gli elementi di $\mathcal{B}(\mathbf{R}^n)$ sono detti **boreliani** in $\mathbf{R}^n$.

Si potrebbe poi dimostrare che, in uno spazio normato arbitrario, i chiusi sono esattamente i complementari degli aperti. Quindi anche tutti gli insiemi chiusi sono boreliani, in virtù delle proprietà delle $\sigma-$algebre (vedi il teorema **??**).

Si osservi che tra le $\sigma-$ algebre che contengono la classe degli insiemi aperti c'è $\mathcal{M}_n$. Dunque, i boreliani sono tutti misurabili secondo Lebesgue. Di fatto, $\mathcal{B}(\mathbf{R}^n)$ è una classe più ristretta di $\mathcal{M}_n$, che contiene però ancora tutti i sottoinsiemi di $\mathbf{R}^n$ che si incontrano nelle applicazioni.

Vale il seguente

Teorema 4.4.3 *Siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità , $X : \Omega \to \mathbf{R}^n$ una variabile aleatoria $n-$dimensionale. Allora, $\forall B \in \mathcal{B}(\mathbf{R}^n)$, $\{X \in B\} \in \mathcal{A}$.*

Definizione 4.4.4 *Siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità , $X : \Omega \to \mathbf{R}$ una v. a. r.. La* **funzione di ripartizione** *(o funzione distribuzione) F_X di X è la seguente*

$$\begin{cases} F_X : \mathbf{R} \to \mathbf{R}, \\ F_X(t) = P(X \le t), \quad t \in \mathbf{R}. \end{cases} \tag{4.4.6}$$

Osservazione 4.4.3 Nella definizione **??** abbiamo scritto $P(X \le t)$ invece di $P(\{X \le t\})$, oppure $P(\{\omega \in \Omega : X(\omega) \le t\})$. In generale, useremo le notazioni $P(X \in B)$, $P(X \ge t)$, ecc. invece di $P(\{X \in B\})$, $P(\{X \ge t\})$, ecc. ogni volta che questa semplificazione notazionale non darà luogo a equivoci. Osserviamo che $\{x \in \mathbf{R} : x \le t\}$ è un boreliano perché è un insieme chiuso. Perciò , in base al teorema **??**, la definizione **??** è ben posta.

Esempio 4.4.3 Consideriamo l'esempio **??**. Si verifica facilmente che si ha

$$\{X \le t\} = \begin{cases} \emptyset & \text{se} \quad t < 0, \\ \{2, 4, 6\} & \text{se} \quad 0 \le t < 1, \\ \Omega & \text{se} \quad t \ge 1. \end{cases}$$

Quindi

$$F_X(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ 1/2 & \text{se} \quad 0 \le t < 1, \\ 1 & \text{se} \quad t \ge 1. \end{cases} \tag{4.4.7}$$

Esempio 4.4.4 Determiniamo la funzione di ripartizione della variabile aleatoria dell'esempio **??**. Sia $t \in \mathbf{R}$. Si ha

$$\{X \leq t\} = \begin{cases} \emptyset & \text{se} \quad t < 0, \\ \{\omega \in \Omega : \|\omega\| \leq t\} & \text{se} \quad 0 \leq t \leq r, \\ \Omega & \text{se} \quad t > r. \end{cases}$$

Quindi

$$F_X(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ (t/r)^2 & \text{se} \quad 0 \leq t < r, \\ 1 & \text{se} \quad t \geq r. \end{cases} \tag{4.4.8}$$

Prima di enunciare e, almeno parzialmente, dimostrare le proprietà principali della funzione di ripartizione di una v. a. r., presentiamo senza dimostrazione due importanti proprietà delle misure di probabilità .

Teorema 4.4.4 *Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità . Sia poi $(A_n)_{n\in\mathbf{N}}$ una successione di eventi. Allora:*
(I) se $A_n \subseteq A_{n+1}$ $\forall n \in \mathbf{N}$ e $A = \cup_{n\in\mathbf{N}} A_n$, allora

$$P(A) = \lim_{n\to+\infty} P(A_n);$$

(II) se $A_{n+1} \subseteq A_n$ $\forall n \in \mathbf{N}$ e $A = \cap_{n\in\mathbf{N}} A_n$, allora

$$P(A) = \lim_{n\to+\infty} P(A_n).$$

Osservazione 4.4.4 In ciascuno dei due casi del teorema **??** $\lim\limits_{n\to+\infty} P(A_n)$ esiste perché la successione $(P(A_n))_{n\in\mathbf{N}}$ è monotona (non decrescente nel primo caso, non crescente nel secondo).

Teorema 4.4.5 *Siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità , $X : \Omega \to \mathbf{R}$ una v. a. r.. Allora:*
(I) $\forall t \in \mathbf{R}$, $0 \leq F_X(t) \leq 1$;
(II) $\forall a, b \in \mathbf{R}$, con $a < b$,

$$P(a < X \leq b) = F_X(b) - F_X(a);$$

(III) F_X è monotona non decrescente;
(IV) F_X è continua a destra, vale a dire, $\forall t \in \mathbf{R}$, $\lim\limits_{s\to t^+} F_X(s) = F_X(t)$;
(V) $\lim\limits_{t\to-\infty} F_X(t) = 0$ e $\lim\limits_{t\to+\infty} F_X(t) = 1$.

Dimostrazione (I) è ovvia.

(II) Si ha $\{X \le a\} \subseteq \{X \le b\}$ e $\{X \le b\} \setminus \{X \le a\} = \{a < X \le b\}$. Quindi (II) segue dal teorema **??**(III).

(III) Se $a < b$, da (II) si ha $F(b) - F(a) \ge 0$.

(IV) Sia $t \in \mathbf{R}$. Allora $\{X \le t\} = \cap_{n\in\mathbf{N}}\{X \le t + 1/n\}$. Infatti, se $\omega \in \Omega$ e $X(\omega) \le t$, allora $X(\omega) \le t + 1/n$ $\forall n \in \mathbf{N}$. Viceversa, da $X(\omega) \le t + 1/n$ $\forall n \in \mathbf{N}$, segue $X(\omega) \le t$. Infatti, se $X(\omega) > t$, prendendo n sufficientemente grande si avrebbe $t + 1/n < F_X(\omega)$. Poiché $\forall n \in \mathbf{N}$ $\{X \le t + 1/(n+1)\} \subseteq \{X \le t + 1/n\}$, dal teorema **??**(II) otteniamo

$$F_X(t) = P(X \le t) = \lim_{n\to+\infty} P(X \le t + 1/n) = \lim_{n\to+\infty} F_X(t + 1/n). \quad (4.4.9)$$

Sia ora $\epsilon \in \mathbf{R}^+$. Per (**??**), esiste $n_0 \in \mathbf{N}$, tale che $F_X(t + 1/n_0) < F_X(t) + \epsilon$. Dunque, se $0 < s < 1/n_0$, da (III), si ha

$$F_X(t) - \epsilon < F_X(t) \le F_X(s) \le F_X(t + 1/n_0) < F_x(t) + \epsilon.$$

(V) Per ogni $n \in \mathbf{N}$, $\{X \le -n-1\} \subseteq \{X \le -n\}$. Poiché $\cap_{n\in\mathbf{N}}\{X \le -n\} = \emptyset$, si ha, applicando ancora il teorema **??**(II),

$$\lim_{n\to+\infty} F_X(-n) = \lim_{n\to+\infty} P(X \le -n) = 0. \quad (4.4.10)$$

Il primo limite in (V) segue allora da (**??**) e da (III).

Inoltre, per ogni $n \in \mathbf{N}$, $\{X \le n\} \subseteq \{X \le n+1\}$. Poiché $\cup_{n\in\mathbf{N}}\{X \le n\} = \Omega$, si ha,applicando il teorema **??**(I),

$$\lim_{n\to+\infty} F_X(n) = \lim_{n\to+\infty} P(X \le n) = 1. \quad (4.4.11)$$

Il secondo limite in (V) segue allora da (**??**) e da (III). □

Osservazione 4.4.5 L'esempio **??** mostra che la funzione di ripartizione, sempre continua a destra, può non essere continua. Nel caso in questione, si ha, ad esempio,

$$\lim_{s\to0^-} F_X(s) = 0 \ne F_X(0) = 1/2.$$

Si veda per questo anche l'esercizio **??**.

La funzione di ripartizione è definita solo per v. a. r. Per le v.a. $n-$dimensionali ($n \in \mathbf{N}$), possiamo introdurre la nozione di legge:

Definizione 4.4.5 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità ,* $X : \Omega \to \mathbf{R}^n$ *una variabile aleatoria* $n-$*dimensionale. Definiamo* **legge** *di* X *la funzione*

$$\begin{cases} Q_X : \mathcal{B}(\mathbf{R}^n) \to \mathbf{R}, \\ Q_X(B) = P(X \in B), \quad B \in \mathcal{B}(\mathbf{R}^n). \end{cases} \tag{4.4.12}$$

Teorema 4.4.6 *La legge di una variabile aleatoria* $n-$*dimensionale* X *è una misura di probabilità sulla* $\sigma-$*algebra* $\mathcal{B}(\mathbf{R}^n)$.

Dimostrazione Innanzi tutto, è chiaro da (**??**) che $Q_X(B) \in [0,1]$ $\forall B \in \mathcal{B}(\mathbf{R}^n)$. Inoltre,

$$Q_X(\mathbf{R}^n) = P(X \in \mathbf{R}^n) = P(\Omega) = 1.$$

Sia infine $\{B_k : k \in \mathbf{N}\}$ una successione di boreliani in $\mathbf{R}^n$ a due a due disgiunti. È chiaro allora che anche gli eventi $\{X \in B_k\}$ $(k \in \mathbf{N})$ sono a due a due disgiunti. Perciò , in virtù del teorema **??**(I), si ha

$$\begin{array}{lcl} Q_X(\cup_{k\in\mathbf{N}} B_k) & = & P(X \in \cup_{k\in\mathbf{N}} B_k) \\ = P(\cup_{k\in\mathbf{N}}\{X \in B_k\}) & = & \sum_{k=1}^{\infty} P(X \in B_k) \\ & = & \sum_{k=1}^{\infty} Q_X(B_k). \end{array}$$

Esempio 4.4.5 Se $X : \Omega \to \mathbf{R}^n$ è una variabile aleatoria che assume solo un numero finito di valori $\{\alpha^1, ..., \alpha^r\}$, si avrà , per ogni $B \in \mathcal{B}(\mathbf{R}^n)$,

$$Q_X(B) = P(X \in B) = \sum_{\alpha^j \in B} P(X = \alpha^j). \tag{4.4.13}$$

Il risultato dell'esempio **??** si può generalizzare introducendo la nozione di variabile aleatoria discreta:

Definizione 4.4.6 *Sia* X *una variabile aleatoria* $n-$*dimensionale. Diremo che* X *è* **discreta** *se assume solo un numero finito o un'infinità numerabile di valori.*

Se X è una variabile aleatoria discreta, e assume l'insieme di valori $\{\alpha^j : j \in \mathbf{N}\}$, si verifica facilmente che, per ogni $B \in \mathcal{B}(\mathbf{R}^n)$, vale la formula (**??**). Al secondo membro di (**??**) ci sarà una somma finita o una serie, a seconda che l'insieme $\{j \in \mathbf{N} : \alpha^j \in B\}$ sia finito o numerabile.

Esempio 4.4.6 Supponiamo di lanciare una moneta equilibrata, fino a ottenere "testa" (T). Supponiamo anche che i risultati dei singoli lanci siano indipendenti tra loro. Poniamo

$$\Omega := \{T, CT, CCT, CCCT, ...\}. \tag{4.4.14}$$

Ω è l'insieme costituito da tutte le possibili successioni finite $\underbrace{CC...C}_{n}T$, con $n \in \mathbf{N}$, a cui aggiungiamo la sequenza costituita dal solo T e la successione infinita costante $CCCC....$ Sia $\omega \in \Omega$. Poniamo:

$$p_\omega := \begin{cases} 2^{-n-1} & \text{se} \quad \omega = \underbrace{CC...C}_{n}T, \\ 1/2 & \text{se} \quad \omega = T, \\ 0 & \text{se} \quad \omega \text{ è la successione identicamente uguale a C.} \end{cases} \tag{4.4.15}$$

Poniamo $\mathcal{A} = \mathcal{P}(\Omega)$ e, per $A \in \mathcal{A}$,

$$P(A) := \sum_{\omega \in A} p_\omega. \tag{4.4.16}$$

Osserviamo che

$$P(\Omega) = \sum_{n=0}^{\infty} 2^{-n-1} = \sum_{n=1}^{\infty} 2^{-n} = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

e non è difficile verificare che P è una misura di probabilità . Indichiamo con X la v. a. r. che assegna a ogni sequenza il numero dei lanci da cui è costituita. Dunque $X(T) = 1$, $X(CT) = 2$, ecc.. Se ω è la successione identicamente uguale a C, poniamo $X(\omega) = +\infty$. Con piccolo abuso di linguaggio (X assume il valore $+\infty$, anche se solo in corrispondenza di una sequenza con probabilità nulla), avremo che X è una v. a. r.. Sia allora $B =]10, +\infty[$. Evidentemente,

$$\begin{aligned} Q_X(B) &= P(X \in]10, +\infty[) = P(X > 10) = \textstyle\sum_{j=11}^{\infty} 2^{-j} \\ &= \textstyle\sum_{i=1}^{\infty} 2^{-(10+i)} = 2^{-10}. \end{aligned}$$

Calcoliamo la probabilità che X sia dispari. In questo caso, $B = \{1, 3, 5, ...\}$.

Allora

$$\begin{aligned} Q_X(B) &= P(X \in \{1,3,5,...\}) = \sum_{j\in\{1,3,5,...\}} 2^{-j} \\ &= \textstyle\sum_{i=0}^{\infty} 2^{-(2i+1)} = \frac{1}{2}\sum_{i=0}^{\infty} 4^{-i} \\ &= \frac{1}{2}\frac{1}{1-\frac{1}{4}} = \frac{2}{3}. \end{aligned}$$

Definizione 4.4.7 *Siano X una v. a. $n-$dimensionale, $f \in \mathcal{L}^1(\mathbf{R}^n)$, non negativa q. d.. Diremo che f è una* **densità** *di X se $\forall B \in \mathcal{B}(\mathbf{R}^n)$*

$$Q_X(B) = P(X \in B) = \int_B f(t)dt.$$

Osservazione 4.4.6 Si potrebbe far vedere che due densità della stessa variabile aleatoria coincidono q. d.. Osserviamo anche che, se f è densità di X,

$$\int_{\mathbf{R}^n} f(t)dt = Q_X(\mathbf{R}^n) = 1.$$

Inoltre, se $L_n(B) = 0$,

$$Q_X(B) = \int_B f(t)dt = 0.$$

In particolare, se X ammette densità , per ogni $\alpha \in \mathbf{R}^n$, $P(X = \alpha) = 0$.

Osservazione 4.4.7 Sia X una v. a. r. , che ammette la densità f. Allora, $\forall t \in \mathbf{R}$,

$$F_X(t) = P(X \le t) = \int_{]-\infty,t]} f(s)ds. \tag{4.4.17}$$

È vero anche viceversa; vale a dire, se esiste $f \in \mathcal{L}^1(\mathbf{R})$, quasi dappertutto non negativa, che verifica (**??**), è possibile dimostrare che f è densità di X. Dunque, se $B \in \mathcal{B}(\mathbf{R})$, si ha

$$Q_X(B) = P(X \in B) = \int_B f(s)ds.$$

Consideriamo cosi l'esempio **??**. Poniamo

$$f(t) = \begin{cases} 0 & \text{se} \quad t \in]-\infty, 0[\cup]r, +\infty[, \\ \frac{2t}{r^2} & \text{se} \quad t \in [0, r]. \end{cases}$$

Allora è chiaro che, per ogni $t \in \mathbf{R}$,

$$F_X(t) = \int_{]-\infty,t]} f(s)ds.$$

Quindi X ammette la densità f.

Esempio 4.4.7 Consideriamo l'intervallo $\Omega = [0, a[$ $(a > 0)$. Scegliamo "a caso" un punto ω in Ω. Con l'espressione "a caso" intendiamo affermare che la probabilità che il punto scelto appartenga a un certo sottoinsieme B di Ω misurabile secondo Lebesgue è direttamente proporzionale a $L_1(B)$. Possiamo allora porre:

$$\begin{array}{ll} \mathcal{A} := \{A \subseteq \Omega : A \in \mathcal{M}_1\}, & \\ P(A) = cL_1(A), & A \in \mathcal{A}, \end{array} \tag{4.4.18}$$

con $c \in \mathbf{R}^+$. c deve essere tale che $P(\Omega) = 1$. Quindi

$$1 = P(\Omega) = P([0, a[) = ca,$$

da cui $c = a^{-1}$. Dunque

$$P(A) = \frac{L_1(A)}{a}$$

per ogni $A \in \mathcal{A}$.

Consideriamo la v. a. r. X che fa corrispondere a $\omega \in \Omega$ il rapporto delle distanze di ω da 0 e da a. Perciò

$$\begin{cases} X : \Omega \to \mathbf{R}, \\ X(\omega) = \frac{\omega}{a-\omega}, \quad \omega \in \Omega. \end{cases}$$

Determiniamo la funzione di ripartizione F_X. Per $t \in \mathbf{R}$, si ha:

$$\{X \leq t\} = \begin{cases} \emptyset & \text{se} \quad t < 0, \\ [0, \frac{at}{1+t}] & \text{se} \quad t \geq 0, \end{cases}$$

da cui

$$F_X(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ \frac{t}{1+t} & \text{se} \quad t \geq 0. \end{cases}$$

Poniamo

$$f(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ \frac{1}{(1+t)^2} & \text{se} \quad t \geq 0. \end{cases}$$

Si ha

$$F_X(t) = \int_{]-\infty,t]} f(s)ds$$

per ogni $t \in \mathbf{R}$, per cui si può prendere f come densità di X.

Esempio 4.4.8 Facciamo riferimento agli esempi **??** e **??**. Poniamo

$$X(\omega) = \omega \quad (\omega \in \mathbf{R}^+).$$

X è la v. a. r. che specifica la durata temporale del dispositivo (o l'istante in cui cessa di funzionare). Evidentemente, per ogni $t \in \mathbf{R}$,

$$\{X \le t\} = \begin{cases} \emptyset & \text{se} \quad t \le 0, \\ ]0, t] & \text{se} \quad t > 0, \end{cases}$$

per cui

$$F_X(t) = \begin{cases} 0 & \text{se} \quad t \le 0, \\ \int_{]0,t]} g(s) ds & \text{se} \quad t > 0. \end{cases}$$

Quindi, se poniamo

$$\begin{cases} \tilde{g} : \mathbf{R} \to \mathbf{R}, \\ \tilde{g}(t) = \begin{cases} g(t) & \text{se } t > 0, \\ 0 & \text{se } t \le 0, \end{cases} \end{cases}$$

si ha che $\tilde{g}$ è una densità di X. Nel caso di $g(t) = \lambda e^{-\lambda t}$ ($\lambda \in \mathbf{R}^+$), si dice che X ammette legge esponenziale. Giustifichiamo questa scelta di g. Essa segue dalla seguente ipotesi: che la probabilità che l'apparecchio cessi di funzionare nell'intervallo temporale $]t, t+h]$ ($t, h \in \mathbf{R}^+$) se non si è ancora guastato all'istante t dipenda solo da h. Precisamente,

$$P(\{t < X \le t + h\} | \{X > t\}) = \phi(h), \tag{4.4.19}$$

con $\phi : \mathbf{R}^+ \to \mathbf{R}$. Si ha

$$P(\{t < X \le t + h\} | \{X > t\}) = \frac{P(t < X \le X + h)}{P(X > t)} = \frac{F_X(t + h) - F_X(t)}{1 - F_X(t)},$$

da cui

$$\frac{F_X(t + h) - F_X(t)}{h} = \frac{\phi(h)}{h}[1 - F_X(t)]. \tag{4.4.20}$$

Supponiamo ora che F_X sia derivabile in $\mathbf{R}^+$ e che $\lim_{h \to 0} \frac{\phi(h)}{h} = \lambda \in \mathbf{R}^+$. Passando allora al limite per $h \to 0$ in (**??**), otteniamo che F_X soddisfa in $\mathbf{R}^+$ l'equazione differenziale ordinaria

$$F'_X(t) = \lambda[1 - F_X(t)]. \tag{4.4.21}$$

Da ciò segue facilmente che F_X è della forma

$$F_X(t) = Ce^{-\lambda t} + 1,$$

per qualche $C \in \mathbf{R}$, $C < 0$. Dal fatto che $F_X(0) = 0$ e che supponiamo F_X continua in $\mathbf{R}$, otteniamo $C = -1$. Perciò ,

$$F_X(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ 1 - e^{-\lambda t} & \text{se } t \geq 0. \end{cases} \tag{4.4.22}$$

Da (**??**) si verifica facilmente che X ammette la densità f seguente:

$$f(t) = \begin{cases} 0 & \text{se} \quad t < 0, \\ \lambda e^{-\lambda t} & \text{se } t \geq 0. \end{cases} \tag{4.4.23}$$

Per concludere, ci si può chiedere quale sia il significato della condizione (**??**). Intuitivamente, il significato è il seguente: che, in un certo senso, l'usura del dispositivo in dipendenza della sua età è trascurabile. Infatti, questa condizione implica, in particolare, che la probabilità che un dispositivo con età t raggiunga senza guastarsi l'età $2t$ coincide con la probabilità che un dispositivo con età $2t$ raggiunga senza guastarsi l'età $3t$.

Veniamo ora alla fondamentale nozione di media di una v. a. r.. Cominciamo dal caso di una v. a. r. semplice.

Definizione 4.4.8 *Sia X una v. a. r. semplice sullo spazio di probabilità $(\Omega, \mathcal{A}, P)$. Supponiamo che $\Omega = A_1 \cup ... \cup A_m$, con A_1,...,A_m eventi a due a due disgiunti, e che $X(\omega) = \alpha_j$ se $\omega \in A_j$ $(1 \leq j \leq m)$. Chiamiamo allora* **media** *di X e indichiamo con la scrittura $E(X)$ il numero reale* $\sum_{j=1}^{m} \alpha_j P(A_j)$.

Osservazione 4.4.8 Si può verificare che la definizione di $E(X)$ non dipende dalla decomposizione $\{A_j : 1 \leq j \leq m\}$ di Ω scelta.

Osserviamo anche che, se $X(\omega) = \alpha \ \forall \alpha \in \omega$, $E(X) = \alpha$.

Esempio 4.4.9 Sia X la v. a. r. introdotta nell'esempio **??**. Si ha

$$E(X) = P(A_2) = \frac{1}{2}.$$

Esempio 4.4.10 Consideriamo un processo di Bernoulli con probabilità di successo e di insuccesso in ciascuna singola prova uguali a p e q rispettivamente (vedi l'esempio **??**). Sia X la v. a. r. che assegna alla sequenza $\omega = (\omega_1, ..., \omega_n)$ il numero dei successi che contiene. Dunque, se, per $m = 0, ..., n$, B_m è l'evento introdotto nell'osservazione **??**, abbiamo, in base a (**??**),

$$E(X) = \sum_{m=1}^{n} m \binom{n}{m} p^m q^{n-m}.$$

Passiamo ora a dare la definizione generale di media di una v. a. r.:

Definizione 4.4.9 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità ,* $X : \Omega \to \mathbf{R}$ *una v. a. r. tale che* $X(\omega) \geq 0 \ \forall \omega \in \Omega$. *Poniamo*

$$\begin{aligned} E(X) \quad &:= \sup\{E(Y) : Y : \Omega \to [0, +\infty[\text{ v. a. r. semplice },\\ &\quad 0 \leq Y(\omega) \leq X(\omega) \quad \forall \omega \in \Omega\}. \end{aligned} \tag{4.4.24}$$

Sia, infine, $X : \Omega \to \mathbf{R}$ *una v. a. r. generica. Poniamo*

$$X_+ := \phi_+ \circ X, \qquad\qquad X_- := \phi_- \circ X, \tag{4.4.25}$$

con $\phi_\pm$ *definite in (0.2.1)-(0.2.2). Si può verificare che* $X_\pm$ *sono v. a. r. non negative. Se* $E(X_+) < +\infty$ *e* $E(X_-) < +\infty$, *chiamiamo media di* X *il numero reale*

$$E(X) := E(X_+) - E(X_-). \tag{4.4.26}$$

Osservazione 4.4.9 La definizione **??** richiama le definizioni di integrale 0.2.4 e 0.2.6. Di fatto, sia le misure di probabilità che la misura di Lebesgue possono essere inquadrate in una teoria astratta dell'integrazione che le comprende entrambe. Non sarà dunque sorprendente constatare che, come apparirà nel prossimo teorema, la media ha proprietà analoghe all'integrale.

Teorema 4.4.7 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità ,* $X, Y : \Omega \to \mathbf{R}$ *v. a. r. che ammettono media. Allora:*

(I) $X + Y$ *è una v. a. r. che ammette media; inoltre* $E(X + Y) = E(X) + E(Y)$;

(II) se $\alpha \in \mathbf{R}$, αX *è una v. a. r. che ammette media e* $E(\alpha X) = \alpha E(X)$;

(III) se $X(\omega) \leq Y(\omega) \ \forall \omega \in \Omega$, $E(X) \leq E(Y)$;

(IV) in particolare, se $X(\omega) \geq 0 \ \forall \omega \in \Omega$, $E(X) \geq 0$.

Nel caso particolare delle funzioni semplici proponiamo la dimostrazione del teorema **??** come esercizio (vedi l'esercizio **??**).

Veniamo ora ad alcuni risultati utili per il calcolo delle medie delle v. a. r..

Teorema 4.4.8 *Sia X una v. a. r. discreta nello spazio di probabilità $(\Omega, \mathcal{A}, P)$. Supponiamo che $\Omega = \cup_{j \in \mathbf{N}} A_j$, con gli A_j eventi a due a due disgiunti e $X(\omega) = \alpha_j \in \mathbf{R}$ $\forall \omega \in A_j$. Allora X ammette media se e solo se la serie $\sum_{j=1}^{\infty} \alpha_j P(A_j)$ è assolutamente convergente. In tal caso,*

$$E(X) = \sum_{j=1}^{\infty} \alpha_j P(A_j).$$

Esempio 4.4.11 Consideriamo lo spazio di probabilità dell'esempio **??**. Sia ancora X la v. a. r. che conta il numero del lanci necessario per ottenere "testa". Abbiamo già visto che X assume tutti i possibili valori naturali, più il valore $+\infty$ con probabilità 0. Si ha, inoltre, che, $\forall j \in \mathbf{N}$, $P(X = j) = 2^{-j}$. Poiché la serie $\sum_{j=1}^{\infty} j 2^{-j}$ è convergente (lo si vede applicando, ad esempio, il criterio del rapporto), X ammette media e

$$E(X) = \sum_{j=1}^{\infty} j 2^{-j}. \tag{4.4.27}$$

Calcoliamo la somma della serie in (**??**). Consideriamo la serie di potenze $\sum_{j=0}^{\infty} z^j$, che ha raggio di convergenza 1 ed è tale che, $\forall z \in \mathbf{C}$, con $|z| < 1$,

$$\sum_{j=0}^{\infty} z^j = (1 - z)^{-1}. \tag{4.4.28}$$

Applicando il teorema 2.3.5, possiamo derivare in ciascuno dei due membri di (**??**) e ottenere

$$\sum_{j=1}^{\infty} j z^{j-1} = (1 - z)^{-2}, \tag{4.4.29}$$

che vale se $|z| < 1$. Moltiplicando per z, otteniamo subito, se $|z| < 1$,

$$\sum_{j=1}^{\infty} j z^j = z(1 - z)^{-2}. \tag{4.4.30}$$

Prendendo allora $z = 1/2$, otteniamo

$$E(X) = \sum_{j=1}^{\infty} j 2^{-j} = 2.$$

Per quanto rigurda, invece, le variabili aleatorie dotate di densità , vale il seguente

Teorema 4.4.9 *Sia X una v. a. r., che ammette la densità f. Allora X ammette media se e solo se $t \to tf(t)$ è sommabile su $\mathbf{R}$. In tal caso,*

$$E(X) = \int_{\mathbf{R}} tf(t)dt.$$

Esempio 4.4.12 Consideriamo una v. a. r. X che ammette legge esponenziale (vedi l'esempio **??**). Supponiamo che ammetta densità $f(t) = \lambda e^{-\lambda t}\chi_+(t)$, ove abbiamo indicato con χ_+ la funzione caratteristica di $[0, +\infty[$. Applicando il teorema **??**, otteniamo

$$E(X) \quad = \int_{[0,+\infty[} \lambda t e^{-\lambda t} dt = \lambda^{-1}.$$

Dunque il parametro λ rappresenta l'inverso della media di X.

Il teorema **??** ammette la seguente importante generalizzazione:

Teorema 4.4.10 *Siano X una variabile aleatoria $n-$dimensionale, che ammette la densità f, $g : \mathbf{R}^n \to \mathbf{R}$ continua. Allora $g(X) := g \circ X$ è una v. a. r. e ammette media se e solo se $x \to g(x)f(x)$ è sommabile su $\mathbf{R}^n$. In tal caso,*

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

Dimostrazione parziale Ci limitiamo a verificare il teorema nel caso di $g : \mathbf{R}^n \to \mathbf{R}$ semplice, con $g(x) = \sum_{j=1}^{m} \beta_j \chi_{B_j}(x)$, ove B_1,...,B_m sono boreliani in $\mathbf{R}^n$ a due a due disgiunti e β_1,...,β_m sono numeri reali. Osserviamo che una funzione g siffatta non è (in generale) continua, ma che ogni funzione continua può essere opportunamente approssimata da funzioni di questo tipo.

Poniamo, per $j = 1, ..., m$,

$$A_j := \{X \in B_j\}.$$

Osserviamo che gli A_j sono a due a due disgiunti. Allora, chiaramente, $g(X)$ coincide con la v. a. r. semplice $\sum_{j=1}^m \beta_j \chi_{A_j}$. Dunque

$$\begin{aligned} E(g(X)) &= \textstyle\sum_{j=1}^m \beta_j P(A_j) \\ &= \textstyle\sum_{j=1}^m \beta_j \int_{B_j} f(x)dx \\ &= \textstyle\int_{\mathbf{R}^n} \sum_{j=1}^m \beta_j \chi_{B_j}(x) f(x)dx \\ &= \textstyle\int_{\mathbf{R}^n} g(x)f(x)dx. \end{aligned}$$

Esempio 4.4.13 Consideriamo l'esempio **??**. Sia $X : \Omega \to \mathbf{R}^2$, $X(\omega) = \omega$ (Ricordiamo che $\Omega \subseteq \mathbf{R}^2$). X è una variabile aleatoria bidimensionale. Se $B \in \mathcal{B}(\mathbf{R}^2)$, si ha

$$\{X \in B\} = \Omega \cap B,$$

da cui

$$P(X \in B) = P(\Omega \cap B) = \frac{L_2(\Omega \cap B)}{\pi r^2} = \int_B f(x)dx,$$

con

$$f(x) = \begin{cases} (\pi r^2)^{-1} & \text{se} \quad x \in \Omega, \\ 0 & \text{se} \quad x \notin \Omega. \end{cases}$$

Sia X_1 la prima componente di X. Osserviamo che $X_1(\omega)^2 = g(X(\omega))$, con $g : \mathbf{R}^2 \to \mathbf{R}$, $g(x_1, x_2) = x_1^2$. Per il teorema **??**, si ha allora

$$\begin{aligned} E(X_1^2) &= \textstyle\int_{\mathbf{R}^2} x_1^2 f(x_1,x_2)dx_1dx_2 \\ &= (\pi r^2)^{-1} \textstyle\int_\Omega x_1^2 dx_1 dx_2 \\ &= (\pi r^2)^{-1} \textstyle\int_0^r (\int_0^{2\pi} (\rho\cos(\theta))^2 \rho d\theta)d\rho \\ &= (\pi r^2)^{-1} \textstyle\int_0^r \rho^3 d\rho \int_0^{2\pi} \cos^2(\theta)d\theta \\ &= \tfrac{r^2}{4}. \end{aligned}$$

Definizione 4.4.10 *Sia X una v. a. r., tale che $E(X^2) < +\infty$. Si chiama* **varianza** *di X il numero $\sigma^2(X)$ definito come*

$$\sigma^2(X) := E((X - E(X))^2).$$

Osservazione 4.4.10 Se $E(X^2) < +\infty$, allora X ammette media. Infatti, dalla solita elementare disuguaglianza (1.4.3), si ha

$$|X| = |X| \cdot 1 \leq \frac{1}{2}(X^2 + 1),$$

da cui

$$E(X_+) + E(X_-) = E(|X|) \leq \frac{1}{2}(E(X^2) + 1).$$

Osserviamo anche che

$$(X - E(X))^2 = X^2 - 2E(X)X + E(X)^2,$$

da cui ricaviamo l'utile formula

$$\sigma^2(X) = E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2. \tag{4.4.31}$$

Osservazione 4.4.11 Intuitivamente la varianza misura quanto una variabile aleatoria è sparsa rispetto alla sua media. Per esempio, è facile verificare che una v. a. costante ha varianza nulla (vedi l'esercizio **??**). Se invece X è una variabile aleatoria che assume solo i valori c e $-c$ ($c \in \mathbf{R}^+$) con probabilità $\frac{1}{2}$, dalla formula (**??**) si ha

$$\sigma^2(X) = c^2.$$

Esempio 4.4.14 Consideriamo una v. a. r. discreta X in un certo spazio di probabilità $(\Omega, \mathcal{A}, P)$. Sia $X(\omega) = \alpha_i$ se $\omega \in A_i$, con gli $(A_i)_{i\in\mathcal{I}}$ eventi a due a due disgiunti e $\mathcal{I} = \{1, ..., n\}$, oppure $\mathcal{I} = \mathbf{N}$. Supponiamo che X ammetta media. In base al teorema **??**, ciò significa che la serie $\sum_{i\in\mathcal{I}} |\alpha_i| P(A_i)$ è convergente e $E(X) = \sum_{i\in\mathcal{I}} \alpha_i P(A_i)$. Sempre in base al teorema **??**, si ha

$$E(X^2) = \sum_{i\in\mathcal{I}} \alpha_i^2 P(A_i).$$

Se tale somma è finita, si ha, applicando anche (**??**),

$$\sigma^2(X) = \sum_{i\in\mathcal{I}} (\alpha_i - E(X))^2 P(A_i) = \sum_{i\in\mathcal{I}} \alpha_i^2 P(A_i) - (\sum_{i\in\mathcal{I}} \alpha_i P(A_i))^2.$$

Esempio 4.4.15 Sia X una v. a. r. che ammette la densità f. Allora, applicando il teorema **??**, possiamo dire che

$$E(X^2) = \int_{\mathbf{R}} t^2 f(t) dt.$$

Qualora questa media sia finita, X ammetterà varianza e

$$\sigma^2(X) = \int_{\mathbf{R}} t^2 f(t) dt - (\int_{\mathbf{R}} t f(t) dt)^2.$$

Definizione 4.4.11 *Siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, $\mathcal{I}$ una famiglia di indici e, per ogni $i \in \mathcal{I}$, X_i una v. a. $n-$dimensionale di dominio*

Ω. Diremo che la famiglia $\{X_i : i \in \mathcal{I}\}$ è **indipendente** *(o che le variabili $\{X_i : i \in \mathcal{I}\}$ sono indipendenti) se, comunque si prendano $m \in \mathbf{N}$, $i_1, ..., i_m$ in $\mathcal{I}$ a due a due distinti, B_1,..., B_m boreliani in $\mathbf{R}^n$, si ha che gli eventi $\{X_{i_1} \in B_1\}$,...,$\{X_{i_m} \in B_m\}$ sono indipendenti. In altre parole*

$$P(X_{i_1} \in B_1, ..., X_{i_m} \in B_m) = P(X_{i_1} \in B_1) \cdot ... \cdot P(X_{i_m} \in B_m).$$

Esempio 4.4.16 Consideriamo un processo di Bernoulli (vedi l'esempio **??**). Indichiamo con Ω l'insieme delle $n-$le $\omega = (\omega_1, ..., \omega_i, ..., \omega_n)$, con $\omega_1, ..., \omega_n \in \{S, I\}$. Sia, per $i = 1, ..., n$, X_i una v. a. r. che dipende soltanto dal risultato della $i-$esima prova. In altre parole, supponiamo che

$$X_i(\omega_1, ..., \omega_i, ..., \omega_n) = f_i(\omega_i),$$

con $f_i(S) = \alpha_i$, $f_i(I) = \beta_i$. Per fissare le idee, supponiamo che, per $i = 1, ..., n$, $\alpha_i = 1$, $\beta_i = 0$. Verifichiamo che le variabili aleatorie $\{X_i : 1 \leq i \leq n\}$ sono indipendenti. Siano $1 \leq j_1 < ... < j_r \leq n$, con $1 \leq r \leq n$, per $i = 1, ..., r$ sia $B_i \in \mathcal{B}(\mathbf{R})$. Verifichiamo che

$$P(X_{j_1} \in B_1, ..., X_{j_r} \in B_r) = P(X_{j_1} \in B_1) \cdot ... \cdot P(X_{j_r} \in B_r). \qquad (4.4.32)$$

Osserviamo, innanzi tutto, che si ha, per $i \in \{1, ..., r\}$:

$$\{X_{j_i} \in B_i\} = \begin{cases} \emptyset & \text{se } B_i \cap \{0,1\} = \emptyset, \\ \{\omega \in \Omega : \omega_{j_i} = S\} & \text{se } B_i \cap \{0,1\} = \{1\}, \\ \{\omega \in \Omega : \omega_{j_i} = I\} & \text{se } B_i \cap \{0,1\} = \{0\}, \\ \Omega & \text{se } \{0,1\} \subseteq B_{j_i}. \end{cases} \qquad (4.4.33)$$

Dunque, se, al solito, indichiamo con p la probabilità di successo in ciascuna singola prova e con q la probabilità di insuccesso, si ha, per $i = 1, ..., r$:

$$P(X_{j_i} \in B_i) = \begin{cases} 0 & \text{se } B_i \cap \{0,1\} = \emptyset, \\ p & \text{se } B_i \cap \{0,1\} = \{1\}, \\ q & \text{se } B_i \cap \{0,1\} = \{0\}, \\ 1 & \text{se } \{0,1\} \subseteq B_i. \end{cases} \qquad (4.4.34)$$

Osserviamo allora che, se, per qualche i, $B_i \cap \{0,1\} = \emptyset$, il secondo membro di (**??**) è nullo. In questo caso si ha anche

$$\{X_{j_1} \in B_1\} \cap ... \cap \{X_{j_r} \in B_r\} = \emptyset,$$

per cui vale (**??**). Supponiamo invece che, per ciascun i, $B_i \cap \{0,1\} \neq \emptyset$. Poniamo:

$$\mathcal{I}_1 := \{i \in \{1,...,r\} : B_i \cap \{0,1\} = \{1\}\},$$

$$\mathcal{I}_2 := \{i \in \{1,...,r\} : B_i \cap \{0,1\} = \{0\}\},$$

$$\mathcal{I}_3 := \{i \in \{1,...,r\} : \{0,1\} \subseteq B_i\}.$$

Il prodotto nel secondo membro di (**??**) vale $p^{\sharp(\mathcal{I}_1)}q^{\sharp(\mathcal{I}_2)}$. Si ha invece

$$\{X_{j_1} \in B_1\} \cap ... \cap \{X_{j_r} \in B_r\} = \cap_{i \in \mathcal{I}_1 \cup \mathcal{I}_2} A_{j_i},$$

ove A_j è definito in (**??**) e ω^0 è un elemento qualunque di Ω tale che $\omega^0_{j_i} = S$ se $i \in \mathcal{I}_1$, $\omega^0_{j_i} = I$ se $i \in \mathcal{I}_2$. Poiché A_1,...,A_n sono indipendenti (vedi per questo l'esercizio **??**), si ha

$$P(\cap_{i \in \mathcal{I}_1 \cup \mathcal{I}_2} A_{j_i}) = \prod_{i \in \mathcal{I}_1 \cup \mathcal{I}_2} P(A_{j_i}) = p^{\sharp(\mathcal{I}_1)}q^{\sharp(\mathcal{I}_2)}.$$

Concludiamo questa sezione con alcune importanti proprietà di famiglie di v. a. indipendenti.

Teorema 4.4.11 *Siano X_1,..., X_n v. a. r. indipendenti nello spazio di probabilità $(\Omega, \mathcal{A}, P)$, dotate di densità f_1,...,f_n. Allora la v. a. $n-$dimensionale $X = (X_1, ..., X_n)$ ammette la densità*

$$(f_1 \otimes ... \otimes f_n)(x_1, ..., x_n) := f_1(x_1) \cdot ... \cdot f_n(x_n). \tag{4.4.35}$$

Dimostrazione parziale Ci limitiamo al caso $n = 2$. Si potrebbe dimostrare che $f_1 \otimes f_2$ è misurabile in $\mathbf{R}^2$. Dal teorema di Tonelli, segue poi subito che ha integrale uguale a 1. Sia $B \in \mathcal{B}(\mathbf{R}^2)$. Consideriamo solo il caso

$$B = B_1 \times B_2,$$

con B_1 e B_2 boreliani in $\mathbf{R}$. Si ha allora, sfruttando il fatto che X_1 e X_2 sono indipendenti e il teorema di Tonelli,

$$\begin{aligned} P(X \in B) = \ & P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1) \cdot P(X_2 \in B_2) \\ & = \int_{B_1} f_1(x_1)dx_1 \cdot \int_{B_2} f_2(x_2)dx_2 \\ & = \int_B (f_1 \otimes f_2)(x_1, x_2)dx_1dx_2. \end{aligned}$$

□

Teorema 4.4.12 *Siano X_1,..., X_n v. a. r. indipendenti nello spazio di probabilità $(\Omega, \mathcal{A}, P)$, dotate di densità f_1,...,f_n. Allora la v. a. r. $X_1 + ... + X_n$ ammette la densità $f_1 * ... * f_n$ (vedi l'esercizio* **??** *per la convoluzione di più di due funzioni).*

Dimostrazione parziale Ci limitiamo al caso $n = 2$. In virtù dell'osservazione **??**, basta provare che, $\forall t \in \mathbf{R}$,

$$P(X_1 + X_2 \leq t) = \int_{]-\infty,t]} (f_1 * f_2)(x)dx. \tag{4.4.36}$$

Poniamo $B_t := \{(x_1, x_2) \in \mathbf{R}^2 : x_1 + x_2 \leq t\}$. $B_t \in \mathcal{B}(\mathbf{R}^2)$ perché è chiuso. Allora, applicando il teorema **??** e il teorema di Tonelli, si ha

$$\begin{aligned} P(X_1 + X_2 \leq t) &= P((X_1, X_2) \in B_t) = \textstyle\int_{B_t} f_1(x_1) f_2(x_2) dx_1 dx_2 \\ &= \textstyle\int_{\mathbf{R}} (\int_{]-\infty,t-x_2]} f_1(x_1)dx_1) f_2(x_2)dx_2 \\ &= \textstyle\int_{\mathbf{R}} (\int_{]-\infty,t]} f_1(x - x_2)dx) f_2(x_2)dx_2 \\ &= \textstyle\int_{]-\infty,t]} (\int_{\mathbf{R}} f_1(x - x_2) f_2(x_2)dx_2)dx \\ &= \textstyle\int_{]-\infty,t]} (f_1 * f_2)(x)dx. \end{aligned}$$

□

Esempio 4.4.17 Consideriamo ancora l'esempio **??** (che riprendeva gli esempi **??** e **??**). Supponiamo che, nell'istante in cui il dispositivo cessa di funzionare, entri in funzione un secondo dispositivo con le stesse caratteristiche. Indichiamo con Y la durata del secondo dispositivo, regolata dalla stessa legge. Y avrà densità $f(t) = \lambda e^{-\lambda t} \chi_+(t)$, con χ_+ funzione caratteristica di $[0, +\infty[$. Supponiamo che i tempi di durata dei due dispositivo siano indipendenti. Sarà quindi naturale supporre X e Y v. a. r. indipendenti. Consideriamo la v. a. r. $X + Y$, che indica il tempo totale di funzionamento dei due dispositivo. Allora, in base al teorema **??**, $X + Y$ ammetterà densità $g := f * f$. Se $t \in \mathbf{R}$,

$$\begin{aligned} (f * f)(t) &= \lambda^2 \textstyle\int_{\mathbf{R}} e^{-\lambda(t-s)} \chi_+(t-s) e^{-\lambda s} \chi_+(s) ds \\ &= \lambda^2 t e^{-\lambda t} \chi_+(t). \end{aligned} \tag{4.4.37}$$

Teorema 4.4.13 *Siano X e Y v. a. r. indipendenti sullo stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$. Supponiamo che ammettano entrambe media. Allora anche XY ammette media e $E(XY) = E(X)E(Y)$.*

Dismostrazione parziale Supponiamo che X e Y ammettano densità , rispettivamente, f e g. Il caso di X e Y v. a. r. semplici è trattato nell'esercizio **??**. Poniamo $Z : \Omega \to \mathbf{R}^2$, $Z(\omega) = (X(\omega), Y(\omega))$. Per il teorema **??**, Z ammette la densità $f \otimes g$. Allora, per il teorema di Fubini, poiché $XY = F \circ Z$, con $F : \mathbf{R}^2 \to \mathbf{R}$, $F(x_1, x_2) = x_1 x_2$, si ha, applicando il teorema **??**,

$$\begin{array}{ll} E(XY) & = \int_{\mathbf{R}^2} x_1 x_2 f(x_1) g(x_2) dx_1 dx_2 \\ & = \int_{\mathbf{R}} x_1 f(x_1) dx_1 \cdot \int_{\mathbf{R}} x_2 g(x_2) dx_2 = E(X)E(Y). \end{array}$$

□

Teorema 4.4.14 *Siano $X_1, ..., X_n$ v. a. r. indipendenti sullo stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$. Supponiamo che X_1,...,X_n ammettano tutte varianza. Allora anche $X_1 + ... + X_n$ ammette varianza e*

$$\sigma^2(X_1 + ... + X_n) = \sigma^2(X_1) + ... + \sigma^2(X_n). \tag{4.4.38}$$

Dismostrazione parziale Ci limitiamo a verificare (**??**) nel caso $n = 2$. Si ha, utilizzando il teorema **??**,

$$\begin{array}{ll} \sigma^2(X_1 + X_2) & = E((X_1 + X_2)^2) - E(X_1 + X_2)^2 \\ & = E(X_1^2 + 2X_1X_2 + X_2^2) - [E(X_1) + E(X_2)]^2 \\ & = E(X_1^2) + 2E(X_1)E(X_2) + E(X_2^2) - E(X_1)^2 \\ & -2E(X_1)E(X_2) - E(X_2)^2 \\ & = \sigma^2(X_1) + \sigma^2(X_2). \end{array}$$

□

Esercizio 4.4.1 Provare le seguenti ulteriori proprietà della funzione di ripartizione F_X di una v. a. r. X: siano a e b reali, con $a < b$; allora:
(I) $P(a \le X \le b) = F_X(b) - F_X(a) + P(X = a)$;
(II) $P(a < X < b) = F_X(b) - F_X(a) - P(X = b)$;
(III) $P(a \le X < b) = F_X(b) - F_X(a) - P(X = b) + P(X = a)$;
(IV) esiste $\lim_{t \to a^-} F_X(t)$ e coincide con $P(X < a)$;
(V) $F_X(a) - \lim_{t \to a^-} F_X(t) = P(X = a)$;
(VI) F_X è continua in a se e solo se $P(X = a) = 0$.

Esercizio 4.4.2 Supponiamo di lanciare un dado equilibrato fino a ottenere 6, con i risultati dei singoli lanci indipendenti tra loro. Cotruire un modello probabilistico dell'esperimento. Sia X la variabile aleatoria che conta il numero dei lanci. Calcolare $P(X > 10)$.

Esercizio 4.4.3 Sia X una v. a. r.. Verificare che ammette media nel senso di (**??**) se e solo se $E(|X|) < +\infty$.

Esercizio 4.4.4 Dimostrare il teorema **??** nel caso di X e Y v. a. r. semplici. Sfruttare il fatto che è sempre possibile decomporre Ω in un numero finito di eventi $\{A_j : 1 \leq j \leq n\}$ a due a due disgiunti in modo tale che, su ciascuno degli A_j, sia X che Y sono costanti.

Esercizio 4.4.5 Verificare che una v. a. r. costante ha varianza nulla.

Esercizio 4.4.6 Calcolare la varianza della v. a. r. dell'esempio **??**.

Esercizio 4.4.7 Calcolare la varianza della v. a. r. X dell'esempio **??**. Ricavare preliminarmente, utilizzando ancora il teorema 2.3.5, la formula

$$\sum_{j=1}^{\infty} j^2 z^j = \frac{z+z^2}{(1-z)^3},$$

valida per ogni $z \in \mathbf{C}$, con $|z| < 1$.

Esercizio 4.4.8 Siano X e Y v. a. r. semplici e indipendenti sullo spazio di probabilità $(\Omega, \mathcal{A}, P)$. Verificare che $E(XY) = E(X)E(Y)$.

Sugg.: supponiamo che $\Omega = \cup_{j=1}^m A_j$,con gli A_j eventi a due a due disgiunti, e che $X(\omega) = \alpha_j$ se $\omega \in A_j$. Possiamo supporre che, per $j = 1, ..., m$, $A_j = \{X = \alpha_j\}$. Osserviamo che $X = \sum_{j=1}^{m} \alpha_j \chi_{A_j}$. Nello stesso modo, sia $\Omega = \cup_{k=1}^n B_k$,con i B_k eventi a due a due disgiunti, e che $X(\omega) = \beta_k$ se $\omega \in B_k$. Possiamo supporre che, per $k = 1, ..., n$, $B_k = \{Y = \beta_k\}$. Osserviamo che $Y = \sum_{k=1}^{n} \beta_k \chi_{B_k}$. Si ha allora

$$XY = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k \chi_{A_j} \chi_{B_k} = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k \chi_{A_j \cap B_k}.$$

Tenuto conto che

$$P(X = \alpha_j, Y = \beta_k) = P(X = \alpha_j) P(Y = \beta_k),$$

perché X e Y sono indipendenti, si ha allora

$$\begin{aligned} E(XY) &= \textstyle\sum_{j=1}^m \sum_{k=1}^n \alpha_j \beta_k E(\chi_{A_j \cap B_k}) \\ &= \textstyle\sum_{j=1}^m \sum_{k=1}^n \alpha_j \beta_k P(X = \alpha_j, Y = \beta_k) \\ &= \textstyle\sum_{j=1}^m \sum_{k=1}^n \alpha_j \beta_k P(X = \alpha_j) P(Y = \beta_k) \\ &= \textstyle\sum_{j=1}^m \alpha_j P(A_j) \cdot \sum_{k=1}^n \beta_k P(B_k) \\ &= E(X) \cdot E(Y). \end{aligned}$$

Esercizio 4.4.9 Siano X e Y v. a. r. semplici e indipendenti sullo spazio di probabilità $(\Omega, \mathcal{A}, P)$. Verificare che

$$\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y).$$

Esercizio 4.4.10 Il passo finale di un lungo calcolo richiede di calcolare la somma di tre interi X_1, X_2, X_3. Supponiamo che

(a) i calcoli di X_1, X_2, X_3 siano indipendenti;

(b) nel calcolo di ciascun X_i $(1 \le i \le 3)$ la probabilità di non aver commesso errori sia $p(\in]0,1[)$;

(c) l'errore possa essere solo in eccesso o in difetto di 1;

(d) la probabilità di un errore per eccesso coincida con quella di un errore per difetto.

Calcolare la probabilità che $X_1 + X_2 + X_3$ sia esatto (tenendo conto di eventuali compensazioni).

Esercizio 4.4.11 Una v. a. r. X ammette distribuzione di Poisson se può assumere solo valori interi non negativi e, $\forall k \in \mathbf{N}_0$,

$$P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!},$$

per un certo $\lambda \in \mathbf{R}^+$.

Calcolare $E(X)$ e $\sigma^2(X)$.

Esercizio 4.4.12 Una v. a. r. X ammette densità uniforme in $[a,b]$ $(a, b \in \mathbf{R}$, $a < b)$ se la sua densità è costante in $[a,b]$ e nulla al di fuori di $[a,b]$ stesso.

Calcolare $E(X)$ e $\sigma^2(X)$.

Esercizio 4.4.13 Calcolare la varianza di una v. a. r. con densità esponenziale $f(t) = \lambda e^{-\lambda t}\chi_+(t)$, con $\lambda \in \mathbf{R}^+$.

Esercizio 4.4.14 Siano X una v. a. r. tale che $E(X^2) < +\infty$, a,$b \in \mathbf{R}$, $Y := aX + b$. Verificare che

$$E(Y) = aE(X) + b, \tag{4.4.39}$$

$$\sigma^2(Y) = a^2\sigma^2(X). \tag{4.4.40}$$

Esercizio 4.4.15 Siano $m \in \mathbf{R}$ e $\sigma \in \mathbf{R}^+$. Poniamo

$$\begin{cases} f : \mathbf{R} \to \mathbf{R}, \\ f(t) = \frac{e^{-(t-m)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma)}. \end{cases} \tag{4.4.41}$$

Verificare che:

(I) $\forall m \in \mathbf{R}$, $\forall \sigma \in \mathbf{R}^+$, $\int_{\mathbf{R}} f(t)dt = 1$;

(II) se X è una v. a. r. che ammette f come densità , allora $E(X) = m$ e $\sigma^2(X) = \sigma^2$.

In tale caso, diremo che X è una v. a. r. con legge **normale**.

Esercizio 4.4.16 Sia X una v. a. r. con densità uniforme in $[0, 1]$. Verificare che le seguenti v. a. r. ammettono densità e calcolare tali densità :

(I) $3X + 1$;

(II) X^2;

(III) e^X.

Esercizio 4.4.17 Siano X e Y v. a. r. indipendenti nello spazio di probabilità $(\Omega, \mathcal{A}, P)$, con densità $t \to 2e^{-2t}\chi_{\mathbf{R}^+}$. Determinare la funzione di ripartizione di $2X - Y$.

(Sugg: per ogni $t \in \mathbf{R}$ pensare a $\{\omega \in \Omega : 2X(\omega) - Y(\omega) \leq t\}$ come a $\{\omega \in \Omega : (X(\omega), Y(\omega)) \in \{(x_1, x_2) \in \mathbf{R}^2 : 2x_1 - x_2 \leq t\}\}$).

4.5 Legge dei grandi numeri e teorema limite centrale

In questa sezione presenteremo due importanti risultati di carattere asintotico,vale a dire risultati che descrivono comportamenti al limite, per un certo parametro naturale n che tende a $+\infty$. Il loro interesse e la loro utilità sta nel fatto che consentono, in presenza di un numero molto grande di prove, di predire, con un certo margine di errore, un certo andamento medio complessivo delle prove stesse.

Cominciamo con una disuguaglianza molto semplice, ma spesso assai utile.

Teorema 4.5.1 *(Disuguaglianza di Chebyscev) Sia X una v. a. r. non negativa nello spazio di probabilità $(\Omega, \mathcal{A}, P)$, con $E(X) < +\infty$. Allora, $\forall \epsilon \in \mathbf{R}^+$,*

$$P(X \geq \epsilon) \leq E(X)/\epsilon.$$

Dimostrazione Consideriamo la v. a. r. semplice Y seguente:

$$\begin{cases} Y : \Omega \to \mathbf{R}, \\ Y(\omega) = \begin{cases} \epsilon & \text{se } X(\omega) \geq \epsilon, \\ 0 & \text{altrimenti.} \end{cases} & \omega \in \Omega. \end{cases}$$

Allora $Y(\omega) \leq X(\omega)$ $\forall \omega \in \Omega$. Segue che

$$\epsilon P(X \geq \epsilon) = E(Y) \leq E(X),$$

da cui la conclusione. □

Passiamo adesso al seguente fatto fondamentale:

Lemma 4.5.1 *Siano X e Y v. a. r. con la stessa legge (non necessariamente definite sullo stesso spazio di probabilità).*

Allora, se X ammette media, anche Y ammette media e $E(X) = E(Y)$.
Se X ammette varianza, anche Y ammette varianza e $\sigma^2(X) = \sigma^2(Y)$.

Dimostrazione parziale Ci limitiamo a osservare che, nel caso in cui X (e quindi Y) ammetta densità, il risultato segue dal teorema **??** e dall'esempio **??**. □

Veniamo ora al primo importante risultato di carattere asintotico.

Teorema 4.5.2 *(legge debole dei grandi numeri) Siano X_1,...,X_n v. a. r. indipendenti nello spazio di probabilità $(\Omega, \mathcal{A}, P)$. Supponiamo che ammettano la stessa legge μ e la stessa varianza σ^2. Poniamo*

$$\overline{X}_n := (X_1 + ... + X_n)/n. \tag{4.5.1}$$

Allora, $\forall \epsilon > 0$,

$$P(|\overline{X}_n - \mu| \geq \epsilon) \leq \sigma^2/(n\epsilon^2). \tag{4.5.2}$$

Dimostrazione Osserviamo, innanzi tutto, che

$$E(\overline{X}_n) = \mu.$$

Dunque, per la disuguaglianza di Chebyscev (teorema **??**), si ha

$$P(|\overline{X}_n - \mu| \geq \epsilon) = P((\overline{X}_n - \mu)^2 \geq \epsilon^2) \leq \sigma^2(\overline{X}_n)/\epsilon^2. \tag{4.5.3}$$

Dal risultato dell'esercizio **??** e dal teorema **??**, abbiamo

$$\sigma^2(\overline{X}_n) = \sigma^2(X_1 + ... + X_n)/n^2 = \sigma^2/n. \tag{4.5.4}$$

La conclusione segue allora da (**??**) e (**??**). □

Osservazione 4.5.1 Il significato intuitivo della legge dei grandi numeri è il seguente: per n che tende a $+\infty$, mantenendo costanti μ e σ^2, la v. a. $\overline{X}_n$ tende a coincidere con la media μ (che è un numero reale) delle X_1,...,X_n.

Esempio 4.5.1 Consideriamo il lancio ripetuto di un dado bilanciato. Sia, per $n \in \mathbf{N}$, S_n il numero dei 6 usciti in n lanci. Vogliamo stimare $P(252 < S_{1764} < 336)$.

Sia $\Omega := \{(\omega_1, ..., \omega_{1764} : \omega_j \in \{1,2,3,4,5,6\}\}$. Indichiamo, per $j \in \{1, ..., 1764\}$, con X_j la v. a. r. che vale 1 se $\omega_j = 6$, 0 altrimenti. Evidentemente, $S_{1764} = \sum_{j=1}^{1764} X_j$. Per $j = 1, ..., 1764$, le X_j sono v. a. indipendenti con la stessa legge. Si tratta infatti di v. a. semplici, con $P(X_j = 1) = 1/6$, $P(X_j = 0) = 5/6$, da cui segue $E(X_j) = 1/6$. Inoltre,

$$\sigma^2(X_j) = E(X_j^2) - (1/6)^2 = E(X_j) - 1/36 = 1/6 - 1/36 = 5/36.$$

Utilizzando le notazioni del teorema **??**, abbiamo $\overline{X}_{1764} = S_{1764}/1764$. Allora, per ogni $\epsilon \in \mathbf{R}^+$, in base alla legge debole dei grandi numeri, si ha

$$P(|\overline{X}_{1764} - 1/6| \geq \epsilon) \leq \frac{5/36}{1764\epsilon^2} \tag{4.5.5}$$

Ora,

$$\begin{aligned}\{252 < S_{1764} < 336\} &= \{\tfrac{1}{7} < \overline{X}_{1764} < \tfrac{4}{21}\} \\ &= \{-\tfrac{1}{42} < \overline{X}_{1764} - \tfrac{1}{6} < \tfrac{1}{42}\}\end{aligned}$$

Quindi, utilizzando (**??**) con $\epsilon = 1/42$, otteniamo

$$\begin{aligned}P(252 < S_{1764} < 336) &= P(|\overline{X}_{1764} - \tfrac{1}{6}| < \tfrac{1}{42}) \\ &= 1 - P(|\overline{X}_{1764} - \tfrac{1}{6}| \geq \tfrac{1}{42}) \\ &\geq 1 - \tfrac{42^2 \times 5/36}{1764} = \tfrac{31}{36} \cong 0,86.\end{aligned}$$

Passiamo ora a illustrare il secondo importante risultato di questa sezione, il così detto **teorema limite centrale**. Cominciamo con alcune definizioni e osservazioni.

Definizione 4.5.1 *Siano* $(X_n)_{n\in\mathbf{N}}$ *una successione di v. a. r. e* X *una v. a. r., non necessariamente definite sullo stesso spazio di probabilità. Indichiamo con* F_n *e* F *rispettivamente le funzioni di ripartizione di* X_n *e* X*. Diremo allora che la successione* $(X_n)_{n\in\mathbf{N}}$ *converge in legge a* X *e sciveremo*

$$X_n \overset{L}{\to} X \tag{4.5.6}$$

se $\forall t \in \mathbf{R}$ *che sia punto di continuità per* F *si ha*

$$\lim_{n\to\infty} F_n(t) = F(t).$$

Osservazione 4.5.2 Poiché F è non decrescente (teorema **??** (III)), si potrebbe dimostrare che l'insieme dei suoi punti di discontinuità è al più numerabile. Ricordiamo (esercizio 0.1.1) che un insieme numerabile ha misura di Lebesgue nulla. Dunque, la convergenza in legge implica la convergenza puntuale quasi dappertutto della successione delle funzioni di ripartizione $(F_n)_{n\in\mathbf{N}}$.

Osserviamo anche che, se X e Y sono v.a.r. con la stessa legge e vale

$$X_n \overset{L}{\to} X\ ,$$

vale anche

$$X_n \overset{L}{\to} Y\ .$$

Definizione 4.5.2 *Siano* $(\Omega, \mathcal{A}, P)$ *uno spazio di probabilità e* $X : \Omega \to \mathbf{C}$*. Diremo che* X è *una* **variabile aleatoria complessa** *(v.a.c.) se* $Re(X)$ *e* $Im(X)$ *sono v.a.r..*

Se $Re(X)$ *e* $Im(X)$ *ammettono media, definiamo* **media** *di* X *il numero complesso*

$$E(X) := E(Re(X)) + iE(Im(X)). \tag{4.5.7}$$

Definizione 4.5.3 *Sia* X *una v.a.r. sullo spazio di probabilità* $(\Omega, \mathcal{A}, P)$*. Definiamo* **funzione caratteristica** *di* X *la funzione*

$$\begin{cases} \phi_X : \mathbf{R} \to \mathbf{C}, \\ \phi_X(\xi) = E(e^{i\xi X}), \quad \xi \in \mathbf{R}. \end{cases} \tag{4.5.8}$$

Osservazione 4.5.3 In virtù del teorema **??**, per ogni $\xi \in \mathbf{R}$, la funzione di dominio Ω $\omega \to e^{i\xi X(\omega)}$ è una v.a.c., essendo

$$e^{i\xi X(\omega)} = \cos(\xi X(\omega)) + i\sin(\xi X(\omega)).$$

Ovviamente, poiché sin e cos sono funzioni limitate, le medie $E(\cos(\xi X))$ e $E(\sin(\xi X))$ sono definite. Ne segue che qualunque v.a.r. ammette funzione caratteristica.

Per inciso, la denominazione "funzione caratteristica" non sembra particolarmente felice, in quanto lo stesso termine è stato usato per indicare la funzione che vale 1 su un certo sottoinsieme del dominio, 0 al di fuori di esso. Tuttavia questa è la denominazione di uso comune. Nell'ambito del calcolo delle probabilità , per indicare le funzioni caratteristiche nel vecchio senso, si parla piuttosto di "funzioni indicatrici".

Osservazione 4.5.4 Se X è una v.a.r. che ammette la densità f, si ha, $\forall \xi \in \mathbf{R}$, applicando il teorema **??**,

$$\begin{aligned}\phi_X(\xi) &= E(e^{i\xi X}) = E(\cos(\xi X)) + iE(\sin(\xi X)) \\ &= \int_{\mathbf{R}} \cos(t\xi) f(t)dt + i \int_{\mathbf{R}} \sin(t\xi) f(t)dt = \int_{\mathbf{R}} e^{it\xi} f(t)dt \\ &= \hat{f}(-\xi),\end{aligned}$$

ove abbiamo indicato con $\hat{f}$ la trasformata di Fourier di f.

Enunciamo ora, senza dimostrazione, un risultato molto utile, che lega la convergenza in legge con la convergenza puntuale delle funzioni di ripartizione.

Teorema 4.5.3 *(di P. Levy) Siano $(X_n)_{n \in \mathbf{N}}$ una successione di v. a. r. e X una v. a. r., non necessariamente definite sullo stesso spazio di probabilità. Allora sono equivalenti:*

(I)

$$X_n \overset{L}{\to} X;$$

(II) $\forall \xi \in \mathbf{R}$, si ha

$$\lim_{n \to +\infty} \phi_{X_n}(\xi) = \phi_X(\xi).$$

Definizione 4.5.4 *Una v.a.r. X ammette* **distribuzione normale standard** *se ammette la densità $f(t) = e^{-t^2/2}/\sqrt{2\pi}$ ($t \in \mathbf{R}$).*

Osservazione 4.5.5 In base al risultato dell'esercizio **??**, se X ammette distribuzione normale standard, ha media nulla e varianza uguale a 1.

Siamo ora in grado di enunciare e parzialmente dimostrare il seguente classico risultato:

Teorema 4.5.4 *(Teorema limite centrale) Siano, per $n \in \mathbf{N}$, $(\Omega_n, \mathcal{A}_n, P_n)$ una successione di spazi di probabilità, X_{n1},...,X_{nn} v.a.r..Supponiamo che,*

*(I) per ogni $n \in \mathbf{N}$, $k \in \{1, ..., n\}$, le X_{nk} ammettano tutte la stessa legge, media $\mu \in \mathbf{R}$ e varianza $\sigma^2 \in \mathbf{R}^+$ (ancora indipendenti da n e k, in base al lemma **??**);*

(II) per ogni $n \in \mathbf{N}$, le v.a.r. X_{n1},...,X_{nn} siano indipendenti.

Poniamo, ancora per $n \in \mathbf{N}$,

$$S_n^* := (X_{n1} + ... + X_{nn} - n\mu)/(\sigma\sqrt{n}), \tag{4.5.9}$$

con $\sigma := \sqrt{\sigma^2}$.

Allora, la successione $(S_n^*)_{n \in \mathbf{N}}$ converge in legge a una v.a.r. normale standard.

Dimostrazione parziale Dimostriamo il teorema sotto l'ulteriore ipotesi che le X_{nk} ($n \in \mathbf{N}$, $k \in \{1, ..., n\}$) ammettano tutte la densità f.

Poniamo, per $t \in \mathbf{R}$, $g(t) := e^{-t^2/2}/\sqrt{2\pi}$. Lasciamo al lettore verificare che, per ogni $\xi \in \mathbf{R}$, si ha

$$\hat{g}(\xi) = e^{-\xi^2/2}.$$

Allora, in base al teorema di P. Levy e all'osservazione **??**, possiamo cercare di far vedere che, $\forall \xi \in \mathbf{R}$, vale

$$\lim_{n \to +\infty} \phi_{S_n^*}(\xi) = e^{-\xi^2/2}. \tag{4.5.10}$$

Poniamo, per $n \in \mathbf{N}$,

$$S_n := X_{n1} + ... + X_{nn}. \tag{4.5.11}$$

Per il teorema **??**, S_n ammette la densità

$$f_n := f * ... * f \quad (n \text{ fattori }). \tag{4.5.12}$$

Verifichiamo che S_n^* ammette la densità

$$f_n^*(t) := \sqrt{n}\sigma f_n(\sqrt{n}\sigma t + n\mu), t \in \mathbf{R}. \tag{4.5.13}$$

Infatti, per ogni $t \in \mathbf{R}$,

$$\begin{aligned} P_n(S_n^* \leq t) &= P_n(S_n \leq n\mu + \sqrt{n}\sigma t) = \textstyle\int_{]-\infty, n\mu + \sqrt{n}\sigma t]} f_n(s)ds \\ &= \textstyle\int_{]-\infty, t]} \sqrt{n}\sigma f_n(\sqrt{n}\sigma s + n\mu)ds. \end{aligned}$$

Perciò , applicando ancora l'osservazione **??**, si ha, $\forall \xi \in \mathbf{R}$,

$$\begin{aligned}\phi_{S_n^*}(\xi) &= \sqrt{n}\sigma \int_{\mathbf{R}} e^{it\xi} f_n(\sqrt{n}\sigma t + n\mu)dt = \int_{\mathbf{R}} e^{i(s-n\mu)\xi/(\sqrt{n}\sigma)} f_n(s)ds \\ &= e^{-i\sqrt{n}\mu\xi/\sigma} \hat{f}(-\xi/(\sqrt{n}\sigma))^n,\end{aligned} \tag{4.5.14}$$

ove nel passaggio finale abbiamo utilizzato il teorema **??** (III), in base al quale, per ogni $\eta \in \mathbf{R}$, $\hat{f}_n(\eta) = \hat{f}(\eta)^n$. Si tratta dunque di far vedere che, per ogni $\xi \in \mathbf{R}$,

$$\lim_{n\to+\infty} e^{-i\sqrt{n}\mu\xi/\sigma} \hat{f}(-\xi/(\sqrt{n}\sigma))^n = e^{-\xi^2/2}. \tag{4.5.15}$$

A tale scopo, cominciamo con l'esaminare la trasformata di Fourier $\hat{f}$. Poiché le v. a. r. X_{nk} ammettono media e varianza, in base al teorema **??**, le funzioni $t^j f$ $(j \in \{0, 1, 2\})$ sono sommabili in $\mathbf{R}$. Segue allora dal corollario **??** (II) che $\hat{f}$ è di classe C^2. Consideriamo allora lo sviluppo di Taylor di $\hat{f}$ con punto iniziale 0 (in generale, $\hat{f}$ è a valori complessi, ma un risultato analogo al teorema 3.7.1 di "Analisi matematica A" vale anche in questo caso). Si ha perciò

$$\hat{f}(\eta) = \hat{f}(0) + \hat{f}'(0)\eta + \hat{f}''(0)\eta^2/2 + r(\eta),$$

con $r(\eta) = o(\eta^2)$ per $\eta \to 0$. Poiché f è una densità , si ha

$$\hat{f}(0) = \int_{\mathbf{R}} f(t)dt = 1. \tag{4.5.16}$$

Dal corollario **??** e dal teorema **??** otteniamo inoltre

$$\hat{f}'(0) = -i \int_{\mathbf{R}} tf(t)dt = -i\mu \tag{4.5.17}$$

e

$$\begin{aligned}\hat{f}''(0) &= -\int_{\mathbf{R}} t^2 f(t)dt = -E(X_{nk}^2) = -[E(X_{nk}^2) - \mu^2] - \mu^2 \\ &-\sigma^2 - \mu^2.\end{aligned} \tag{4.5.18}$$

Da (**??**)-(**??**) segue allora

$$\hat{f}(\eta) = 1 - i\mu\eta - (\sigma^2 + \mu^2)\eta^2/2 + r(\eta), \tag{4.5.19}$$

con $r(\eta) = o(\eta^2)$ per $\eta \to 0$ e, per $\xi \in \mathbf{R}$ fissato,

$$\hat{f}(-\xi/(\sqrt{n}\sigma)) = 1 + i\mu\xi/(\sqrt{n}\sigma) - (\sigma^2 + \mu^2)\xi^2/(2n\sigma^2) + o(n^{-1})(n \to +\infty). \tag{4.5.20}$$

Indichiamo adesso con log la funzione logaritmo di dominio $\mathbf{C}\setminus]-\infty,0]$ tale che

$$\log(z)=\ln(|z|)+iArg(z),$$

con $Arg(z)\in]-\pi,\pi[\cap arg(z)$. log ammette lo sviluppo di Taylor

$$\log(1+z)=z-z^2/2+o(z^2)(z\to 0). \tag{4.5.21}$$

Dunque, poiché $\lim_{n\to+\infty}\hat{f}(-\xi/(\sqrt{n}\sigma))=1$, da (**??**)-(**??**) otteniamo, per $n\to+\infty$,

$$\begin{aligned}&\log(\hat{f}(-\xi/(\sqrt{n}\sigma)))\\&=\log(1+i\mu\xi/(\sqrt{n}\sigma)-(\sigma^2+\mu^2)\xi^2/(2n\sigma^2)+o(n^{-1}))\\&=i\mu\xi/(\sqrt{n}\sigma)-(\sigma^2+\mu^2)\xi^2/(2n\sigma^2)\\&-[i\mu\xi/(\sqrt{n}\sigma)+o(n^{-1/2})]^2/2+o(n^{-1})\\&=i\mu\xi/(\sqrt{n}\sigma)-\xi^2/(2n)+o(n^{-1}).\end{aligned} \tag{4.5.22}$$

Perciò ,

$$\begin{aligned}e^{-i\sqrt{n}\mu\xi/\sigma}\hat{f}(-\xi/(\sqrt{n}\sigma))^n&=e^{-i\sqrt{n}\mu\xi/\sigma}e^{n\log(\hat{f}(-\xi/(\sqrt{n}\sigma))}\\&=e^{-\xi^2/2+o(1)}\end{aligned} \tag{4.5.23}$$

che tende a $e^{-\xi^2/2}$ per $n\to+\infty$.

Perciò (**??**) è provata. □

Osservazione 4.5.6 Il teorema limite centrale è , almeno a prima vista, piuttosto sorprendente, in quanto il fatto che S_n^* converga in legge a una v. a. con distribuzione normale standard è largamente indipendente dalla legge delle v.a.r. X_{nk}. Come vedremo negli esempi, l'interesse del teorema sta essenzialmente nel fatto che costituisce uno strumento di calcolo piuttosto efficiente, grazie alle tavole della distribuzione normale standard, che si trovano in molti libri. Tali tavole riportano i valori di

$$\Phi(t):=(2\pi)^{-1/2}\int_{-\infty}^{t}e^{-s^2/2}ds \tag{4.5.24}$$

per un insieme discreto di t in $\mathbf{R}^+$.

Osservazione 4.5.7 Sotto le ipotesi del teorema **??**, si potrebbe in realtà dimostrare che

$$\lim_{n\to+\infty}P(S_n^*\le t)=\Phi(t)$$

uniformemente per $t\in\mathbf{R}$.

Esempio 4.5.2 Una compagnia di assicurazione deve stabilire il prezzo X di una certa polizza per infortuni. Supponiamo che la compagnia si aspetti 10.000 clienti e che il risarcimento (nel caso in cui capiti effettivamente l'infortunio) sia di 1.000 euro. Supponiamo anche di sapere che la probabilità che al singolo cliente capiti un infortunio sia stimata in $6\cdot 10^{-3}$. La compagnia vuole fissare il prezzo X in modo tale che la probabilità di un guadagno netto di almeno 10.000 euro sia non inferiore a 9/10.

Sia k il numero dei clienti che incorrono in un incidente. Se X è il prezzo della polizza, il guadagno netto è evidentemente,

$$10.000 \cdot X - 1.000 \cdot k. \tag{4.5.25}$$

Affinché il guadagno sia quanto voluto, deve perciò valere

$$10.000 \cdot X - 1.000 \cdot k \geq 10.000, \tag{4.5.26}$$

vale a dire,

$$k \leq 10(X-1). \tag{4.5.27}$$

Se k_0 ($\in \{0, ..., 10.000\}$), la probabilità che il numero di incidenti non superi k_0 (ammettendo che i clienti li subiscano indipendentemente l'uno dall'altro) vale, in base a (**??**),

$$\sum_{k=0}^{k_0} \binom{10.000}{k} (6 \cdot 10^{-3})^k (1 - 6 \cdot 10^{-3})^{10.000-k}, \tag{4.5.28}$$

numero di difficile valutazione. Seguiamo allora un'altra strada. Sia, per $j \in \{1, ..., 10.000\}$, X_j la v.a.r. che vale 1 se il $j-$esimo cliente subisce un incidente, 0 altrimenti. Supponiamo che le X_j siano indipendenti e con la stessa legge. Allora, se k indica il numero totale di incidenti, avremo

$$k = \sum_{j=1}^{10.000} X_j. \tag{4.5.29}$$

Si avrà , per ogni j:

$$\mu := E(X_j) = P(X_j = 1) = 6 \cdot 10^{-3}, \tag{4.5.30}$$

$$\sigma^2 := \sigma^2(X_j) = E(X_j^2) - E(X_j)^2 = 6 \cdot 10^{-3} - 3,6 \cdot 10^{-5} \cong 6 \cdot 10^{-3}. \tag{4.5.31}$$

Poniamo

$$\sigma := \sqrt{\sigma^2} \cong 7,7 \cdot 10^{-2}. \tag{4.5.32}$$

Poniamo infine

$$n := 10.000 \tag{4.5.33}$$

e

$$S_n^* := (k - n\mu)/(\sqrt{n}\sigma). \tag{4.5.34}$$

Allora, se $k_0 \in \{0, ..., 10.000\}$,

$$\begin{aligned} P(k \le k_0) &= P(S_n^* \le (k_0 - n\mu)/(\sqrt{n}\sigma)) \\ &\cong P(S_n^* \le (k_0 - 60)/(100 \cdot 7,7 \cdot 10^{-2})) \\ &= P(S_n^* \le (k_0 - 60)/7,7) \\ &:= P(S_n^* \le k_1). \end{aligned} \tag{4.5.35}$$

Il teorema limite centrale ci dice che

$$P(S_n^* \le k_1) \cong \Phi(k_1), \tag{4.5.36}$$

con Φ definita in (**??**). k_1 deve essere tale che

$$\Phi(k_1) \ge 9/10. \tag{4.5.37}$$

Ciò accade se

$$k_1 \ge 1,29, \tag{4.5.38}$$

che implica, in base a (**??**),

$$k_0 \ge 1,29 \cdot 7,7 + 60 \cong 69,9. \tag{4.5.39}$$

Dovrà perciò essere, in base a (**??**),

$$10(X - 1) \ge 69,9, \tag{4.5.40}$$

vale a dire,

$$X \ge 7,99. \tag{4.5.41}$$

Esempio 4.5.3 Lanciamo una moneta bilanciata 10.000 volte. Stimiamo la probabilità che il numero delle teste sia compreso tra 4.950 e 5.050.

Sia, per $j = 1, ..., 10.000$, X_j una v.a.r. che vale 1 se il $j-$esimo lancio dà testa, 0 se il $j-$esimo lancio dà croce. È ragionevole supporre che le X_j siano indipendenti. Hanno inoltre la stessa legge, con media $\mu = 1/2$, varianza $\sigma^2 = 1/4$. Se $n = 10.000$, $S_n := \sum_{j=1}^n X_j$ è il numero totale di teste. Dobbiamo valutare $P(4.950 < S_n \le 5.050)$. Poniamo,al solito

$$S_n^* := (S_n - n\mu)/(\sqrt{n}\sigma) = (S_n - 5.000)/50. \tag{4.5.42}$$

Allora, applicando il teorema limite centrale, si ha

$$\begin{aligned} P(4.950 < S_n \leq 5.050) &= P(-1 < S_n^* \leq 1) = P(S_n^* \leq 1) - P(S_n^* \leq -1) \\ &\cong \Phi(1) - \Phi(-1). \end{aligned}$$

Dalle tavole otteniamo

$$\Phi(1) \cong 0,8413.$$

Inoltre, se $t \geq 0$, si ha

$$\begin{aligned} \Phi(-t) &= (2\pi)^{-1/2} \int_{-\infty}^{-t} e^{-s^2/2} ds = (2\pi)^{-1/2} \int_t^{+\infty} e^{-s^2/2} ds \\ &= 1 - \Phi(t). \end{aligned}$$

Dunque,

$$\Phi(-1) = 1 - \Phi(1) \cong 0,1587.$$

Concludiamo che

$$P(4.950 < S_n \leq 5.050) \quad \cong 0,8413 - 0,1587 = 0,6826.$$

Esercizio 4.5.1 Una moneta bilanciata viene lanciata n volte. Indichiamo con X_n il numero delle teste uscito. Determinare n tale che $P(0,4 < X_n < 0,6)$ sia maggiore di $0,9$.

Esercizio 4.5.2 Sia, per $t \in \mathbf{R}$, $g(t) = e^{-t^2/2}/\sqrt{2\pi}$. Verificare che, per ogni $\xi \in \mathbf{R}$, si ha

$$\hat{g}(\xi) = e^{-\xi^2/2}.$$

4.6 Catene di Markov

In questa sezione accenneremo a un interessante tipo di processo stocastico discreto: le catene di Markov. Informalmente, per noi un processo stocastico discreto sarà una successione $(X_n)_{n \in \mathbf{N}_0}$ di v.a.r. definite sullo stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$. Intuitivamente, n rappresenta un parametro temporale discreto e $(X_n)_{n \in \mathbf{N}_0}$ descrive l'evoluzione nel tempo di un certo processo di carattere aleatorio.

Introduciamo la seguente

Definizione 4.6.1 *Siano $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, $(X_n)_{n \in \mathbf{N}_0}$ una successione di v.a.r. in $(\Omega, \mathcal{A}, P)$, $S := \{s_1, ..., s_m\}$ un sottoinsieme finito di $\mathbf{R}$. Diremo che $(X_n)_{n \in \mathbf{N}_0}$ è una* **catena di Markov stazionaria** *con insieme degli stati S se sono soddisfatte le seguenti condizioni:*

(I) per ogni $n \in \mathbf{N}_0$ X_n *è a valori in* S*;*

(II) siano $j_1,...,j_p$ *numeri interi, con* $0 \leq j_1 < ... < j_{p-1} < j_p$ $(p \in \mathbf{N}$*,* $p \geq 2)$*. Se* $s_{i_1},...,s_{i_p}$ *sono elementi di* S *(non necessariamente a due a due distinti) e* $P(X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) > 0$*, allora*

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = P(X_{j_p} = s_{i_p} | X_{j_{p-1}} = s_{i_{p-1}}); \tag{4.6.1}$$

(III) se s_i *e* s_j *sono elementi di* S *(non necessariamente distinti),* $n \in \mathbf{N}$ *e* $P(X_n = s_j) > 0$*,*

$$P(X_{n+1} = s_i | X_n = s_j) = a_{ij},$$

con a_{ij} *indipendente da* n*.*

Osservazione 4.6.1 Relativamente alla definizione **??**, l'insieme S rappresenta l'insieme degli stati che il sistema può assumere nel tempo.

Il punto (II) della definizione **??** è la così detta **proprietà di Markov** che caratterizza questi processi. Il suo significato intuitivo è il seguente: che la conoscenza del processo in istanti precedenti a j_{p-1} $(j_1,...,j_{p-2})$ non fornisce ulteriore informazione sulla sua evoluzione dopo j_{p-1} (all'istante j_p), rispetto alla sola conoscenza dello stato del processo all'istante j_{p-1}.

Infine, il punto (III) rappresenta la stazionarietà del processo, nel senso che la probabilità che $X_{n+1} = s_i$ se $X_n = s_j$ non dipende da n e quindi non varia nel tempo.

Osservazione 4.6.2 Sotto le condizioni (I)-(III) della definizione **??**, possiamo associare al processo la matrice $m \times m$ $A = (a_{ij})_{1 \leq i,j \leq m}$, tale che, se per un certo n, $P(X_n = s_j) > 0$,

$$a_{ij} = P(X_{n+1} = s_i | X_n = s_j). \tag{4.6.2}$$

A è una **matrice stocastica**, vale a dire, una matrice a termini reali non negativi, con la somma dei termini in ciascuna colonna uguale a 1. Infatti, se $P(X_n = s_j) > 0$, poiché

$$\cup_{1 \leq i \leq m} (\{X_{n+1} = s_i\} \cap \{X_n = s_j\}) = \{X_n = s_j\},$$

si ha

$$\begin{aligned} \textstyle\sum_{i=1}^m a_{ij} &= \textstyle\sum_{i=1}^m P(X_{n+1} = s_i | X_n = s_j) = \sum_{i=1}^m \frac{P(X_{n+1}=s_i, X_n=s_j)}{P(X_n=s_j)} \\ &= 1. \end{aligned}$$

Esempio 4.6.1 (Rovina del giocatore) Due giocatori disputano una serie di partite con risultati indipendenti tra loro e dallo stato iniziale. Supponiamo che

(I) la probabilità che il primo giocatore si aggiudichi una singola partita sia p ($\in [0,1]$), la probabilità che il secondo giocatore si aggiudichi una singola partita sia $q = 1 - p$;

(II) in ogni partita ci sia in palio un euro;

(III) i capitali dei due giocatori alla partenza siano, rispettivamente, a e b ($a, b \in \mathbf{N}_0$);

(IV) la sequenza si interrompa se uno dei due giocatori perde tutto il suo denaro.

Indichiamo con X_n il capitale del primo giocatore dopo n partite ($n \in \mathbf{N}_0$). X_n può assumere valori compresi tra 0 e $a + b$. Poniamo dunque

$$S := \{0, ..., a + b\}. \tag{4.6.3}$$

È intuitivamente chiaro che $(X_n)_{n \in \mathbf{N}_0}$ è una catena di Markov. Ci limitiamo a verificare che vale (II) della definizione **??** nel caso particolare $j_p = j_{p-1} + 1$. Consideriamo a questo proposito vari casi.

Vediamo, innanzi tutto, il caso

$$1 \leq s_{i_{p-1}} \leq a + b - 1. \tag{4.6.4}$$

In questo caso, entrambi i giocatori dispongono di denaro.
I valori di X_{j_1},...,$X_{j_{p-1}}$ dipendono solo da a, b e dall'andamento delle partite di indice non superiore a j_{p-1}, mentre il risultato della j_p-esima partita è indipendente da questi. Ne segue che

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = \begin{cases} p & \text{se } s_{i_p} = s_{i_{p-1}} + 1, \\ q & \text{se } s_{i_p} = s_{i_{p-1}} - 1, \\ 0 & \text{altrimenti }. \end{cases}$$

Allo stesso risultato si perviene considerando $P(X_{j_p} = s_{i_p} | X_{j_{p-1}} = s_{i_{p-1}})$.

Consideriamo il caso

$$s_{i_{p-1}} = 0. \tag{4.6.5}$$

In questo caso, il primo giocatore non può giocare. Si ha dunque

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = \begin{cases} 1 & \text{se } s_{i_p} = 0, \\ 0 & \text{altrimenti }. \end{cases}$$

Allo stesso risultato si perviene considerando $P(X_{j_p} = s_{i_p} | X_{j_{p-1}} = 0)$.

Un ragionamento analogo si applica, infine, al caso

$$s_{i_{p-1}} = a + b. \tag{4.6.6}$$

Abbiamo dunque una catena di Markov. La matrice A descritta in generale nell'osservazione **??** sarà la matrice $(a_{ij})_{0\leq i,j\leq a+b}$, tale che

$$a_{ij} = \begin{cases} 1 & \text{se } i = j = 0, \\ p & \text{se } 1 \leq j \leq a+b-1, i = j+1, \\ q & \text{se } 1 \leq j \leq a+b-1, i = j-1, \\ 1 & \text{se } i = j = a+b, \\ 0 & \text{negli altri casi.} \end{cases} \tag{4.6.7}$$

Ci poniamo ora le seguenti domande:
(I) qual è la probabilità che il secondo giocatore esaurisca il suo denaro?
(II) Qual è la probabilità che il primo giocatore esaurisca il suo denaro?
(III) Qual è la probabilità che il gioco continui indefinitamente?

Il primo caso corrisponde all'evento $A_a := \bigcup_{n\in\mathbf{N}_0}\{X_n = a+b\}$, il secondo all'evento $B_a := \bigcup_{n\in\mathbf{N}_0}\{X_n = 0\}$,il terzo a $C_a := (A_a \cup B_a)^c$. Abbiamo scritto a al piede di A e B perché, per rispondere alle domande (I)-(III), si rivelerà conveniente considerare il caso in cui la cifra posseduta dal primo giocatore all'inizio sia un generico $s \in \{0, ..., a+b\}$. Gli analoghi di A_a, B_a e C_a saranno A_s, B_s, C_s. Indichiamo con p_s, q_s, r_s le probabilità di A_s, B_s, C_s nei rispettivi spazi. Osserviamo allora, innanzi tutto, che

$$\begin{matrix} p_0 = 0, & q_0 = 1, & r_0 = 0, \\ p_{a+b} = 1, & q_{a+b} = 0, & r_{a+b} = 0. \end{matrix} \tag{4.6.8}$$

Inoltre,

$$p_s + q_s + r_s = 1 \quad \forall s \in \{0, ..., a+b\}. \tag{4.6.9}$$

Supponiamo ora di essere nel caso $X_0 = s \in \{1, ..., a+b-1\}$. Indichiamo con β l'evento "il primo giocatore si aggiudica la prima partita", con γ l'evento "il secondo giocatore si aggiudica la prima partita". Allora

$$\begin{aligned} p_s = P(A_s) &= P(A_s \cap \beta) + P(A_s \cap \gamma) = \tfrac{P(A_s\cap\beta)}{P(\beta)}P(\beta) + \tfrac{P(A_s\cap\gamma)}{P(\beta)}P(\gamma) \\ &= pP(A_s|\beta) + qP(A_s|\gamma). \end{aligned} \tag{4.6.10}$$

Se il primo giocatore si aggiudica la prima partita, passa dal capitale s al capitale $s+1$. Da (**??**) è quindi chiaro che, per $s = 1, ..., a+b-1$, si ha

$$p_s = p \cdot p_{s+1} + q \cdot p_{s-1}, \tag{4.6.11}$$

da cui, poiché $p+q=1$,

$$p(p_{s+1}-p_s)=q(p_s-p_{s-1}), 1\leq s\leq a+b-1. \tag{4.6.12}$$

Supponiamo ora $0<p<1$. Allora, da (**??**) segue

$$p_{s+1}-p_s=\frac{q}{p}(p_s-p_{s-1}), 1\leq s\leq a+b-1 \tag{4.6.13}$$

e quindi, da (**??**),per $s=1,...,a+b-1$,

$$p_{s+1}-p_s=\frac{q}{p}(p_s-p_{s-1})=(\frac{q}{p})^2(p_{s-1}-p_{s-2})=...=(\frac{q}{p})^s p_1. \tag{4.6.14}$$

Dunque, da (**??**) e (**??**) otteniamo

$$\begin{aligned} 1=p_{a+b}-p_0 &= \textstyle\sum_{s=0}^{a+b-1}(p_{s+1}-p_s)=\sum_{s=0}^{a+b-1}(\frac{q}{p})^s p_1 \\ &= \begin{cases} (a+b)p_1 & \text{se } p=q=1/2, \\ \frac{1-(q/p)^{a+b}}{1-q/p}p_1 & \text{se } p\neq q. \end{cases} \end{aligned} \tag{4.6.15}$$

Ne segue che

$$p_1=\begin{cases} \frac{1}{a+b} & \text{se } p=q=1/2, \\ \frac{1-q/p}{1-(q/p)^{a+b}} & \text{se } p\neq q \end{cases} \tag{4.6.16}$$

e, per $s=1,...,a+b$,

$$\begin{aligned} p_s &= \textstyle\sum_{j=0}^{s-1}(p_{j+1}-p_j)=\sum_{j=0}^{s-1}(\frac{q}{p})^j p_1 \\ &= \begin{cases} \frac{s}{a+b} & \text{se } p=q=1/2, \\ \frac{1-(q/p)^s}{1-(q/p)^{a+b}} & \text{se } p\neq q. \end{cases} \end{aligned} \tag{4.6.17}$$

Invertendo i ruoli dei due giocatori, otteniamo anche

$$q_s = \begin{cases} \frac{a+b-s}{a+b} & \text{se } p=q=1/2, \\ \frac{1-(p/q)^{a+b-s}}{1-(p/q)^{a+b}} & \text{se } p\neq q. \end{cases} \tag{4.6.18}$$

Da (**??**)-(**??**), segue subito che, per ogni $s\in\{0,...,a+b\}$,

$$p_s+q_s=1, \tag{4.6.19}$$

da cui

$$r_s=0. \tag{4.6.20}$$

Torniamo al caso generale, vale a dire, supponiamo solo che siano soddisfatte le condizioni (I)-(III) della definizione **??**. Consideriamo la matrice stocastica A definita in (**??**). Poiché si tratta di una matrice quadrata ($m \times m$), ha senso considerare le sue potenze A^l ($l \in \mathbf{N}_0$), ottenute tramite la maltiplicazione standard (righe per colonne) delle matrici. Introduciamo allora la notazione seguente: se $l \in \mathbf{N}_0$, $1 \leq i, j \leq m$, indichiamo con $a_{ij}^{(l)}$ il termine di posto (i, j) di A^l. Allora vale il seguente

Teorema 4.6.1 *Supponiamo di essere nelle ipotesi della definizione* **??**. *Sia* A *la matrice definita in (**??**). Siano poi* j_1,...,j_p *interi non negativi, tali che* $0 \leq j_1 < ... < j_{p-1} < j_p$ ($p \geq 2$), s_{i_1},...,s_{i_p} *elementi di* S *(non necessariamente a due a due distinti) e* $P(X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) > 0$. *Sia infine* $l = j_p - j_{p-1}$. *Allora*

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = a_{i_p, i_{p-1}}^{(l)}. \tag{4.6.21}$$

Dimostrazione parziale Per la proprietà di Markov, è sufficiente considerare il case $p = 2$. Scriviamo allora j al posto di j_1 e $j + l$ al posto di j_2. Ci limitiamo a provare il risultato per $l = 2$. Si ha allora

$$\begin{aligned} P(X_{j+2} = s_{i_2} | X_j = s_{i_1}) &= \tfrac{P(X_j = s_{i_1}, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} \\ &= \sum_{k=1}^{m} \tfrac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})}. \end{aligned} \tag{4.6.22}$$

Sia k tale che $P(X_j = s_{i_1}, X_{j+1} = s_k) > 0$. Allora

$$\begin{aligned} \tfrac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} &= \tfrac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1}, X_{j+1} = s_k)} \tfrac{P(X_j = s_{i_1}, X_{j+1} = s_k)}{P(X_j = s_{i_1})} \\ &= P(X_{j+2} = s_{i_2} | X_j = s_{i_1}, X_{j+1} = s_k) \\ &\times P(X_{j+1} = s_k | X_j = s_{i_1}) \\ &= a_{i_2, k} a_{k, i_1}. \end{aligned} \tag{4.6.23}$$

L'identità dimostrata con la formula (**??**) vale anche nel caso $P(X_j = s_{i_1}, X_{j+1} = s_k) = 0$. Infatti, in questo caso

$$\frac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} = 0,$$

mentre

$$\begin{aligned} a_{k, i_1} &= P(X_{j+1} = s_k | X_j = s_{i_1}) = \tfrac{P(X_j = s_{i_1}, X_{j+1} = s_k)}{P(X_j = s_{i_1})} \\ &= 0. \end{aligned}$$

Da (**??**) otteniamo allora

$$P(X_{j+2} = s_{i_2}|X_j = s_{i_1}) \quad = \textstyle\sum_{k=1}^m a_{i_2,k}a_{k,i_1} = a^{(2)}_{i_2,i_1}. \tag{4.6.24}$$

Esercizio 4.6.1 Relativamente al problema della rovina del giocatore, verificare la validità della proprietà di Markov nel caso $j_p = j_{p-1} + 2$.

Sugg. : ci limitiamo al caso $1 \le s_{j_{p-1}} \le a + b - 1$. Allora

$$\begin{aligned}
&P(X_{j_p} = s_{i_p}|X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) \\
&= \tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+2}=s_{i_p})}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}})} \\
&= \tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}+1,X_{j_{p-1}+2}=s_{i_p})}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}})} \\
&+\tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}-1,X_{j_{p-1}+2}=s_{i_p})}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}})} \\
&= \tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}+1,X_{j_{p-1}+2}=s_{i_p})}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}+1)} \\
&\quad\times\tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}+1)}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}})} \\
&+\tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}-1,X_{j_{p-1}+2}=s_{i_p})}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}-1)} \\
&\quad\times\tfrac{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}},X_{j_{p-1}+1}=s_{i_{p-1}}-1)}{P(X_{j_1}=s_{i_1},...,X_{j_{p-1}}=s_{i_{p-1}})}
\end{aligned}$$

$$\begin{aligned}
&= P(X_{j_{p-1}+2} = s_{i_p}|X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}, X_{j_{p-1}+1} = s_{i_{p-1}} + 1) \\
&\times P(X_{j_{p-1}+1} = s_{i_{p-1}} + 1|X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) \\
&+P(X_{j_{p-1}+2} = s_{i_p}|X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}, X_{j_{p-1}+1} = s_{i_{p-1}} - 1) \\
&\times P(X_{j_{p-1}+1} = s_{i_{p-1}} - 1|X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) \\
&= P(X_{j_{p-1}+2} = s_{i_p}|X_{j_{p-1}+1} = s_{i_{p-1}} + 1) \\
&\times P(X_{j_{p-1}+1} = s_{i_{p-1}} + 1|X_{j_{p-1}} = s_{i_{p-1}}) \\
&+P(X_{j_{p-1}+2} = s_{i_p}|X_{j_{p-1}+1} = s_{i_{p-1}} - 1) \\
&\times P(X_{j_{p-1}+1} = s_{i_{p-1}} - 1|X_{j_{p-1}} = s_{i_{p-1}}),
\end{aligned}$$

ove nell'ultimo passaggio abbiamo usato il caso $j_p = j_{p-1} + 1$.

Indice