Università degli Studi di Pisa

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea in Matematica

Alcune conseguenze non banali del Teorema di Hahn-Banach

13 giugno 2014

Tesi di Laurea Magistrale

Candidato

Andrea Rossi

rossi@mail.dm.unipi.it

Relatore

Prof. Paolo Acquistapace

Università di Pisa

Controrelatore

Prof. Alessandro Berarducci

Università di Pisa

Anno Accademico 2013/2014

Indice

Introduzione

Il teorema di Hahn-Banach è uno dei principali risultati, se non il principale, dell'analisi funzionale lineare. Esso garantisce l'estendibilità a tutto lo spazio di funzionali lineari limitati, definiti in un sottospazio di uno spazio normato, mostrando in questo modo che vi è una tale varietà e quantità di funzionali lineari e continui di un generico spazio normato da rendere interessante e fruttuoso lo studio degli spazi duali. Inoltre la versione geometrica del teorema di Hahn-Banach è di grande importanza nell'analisi convessa. L'impatto del teorema di Hahn-Banach è di fondamentale rilevanza in molti settori dell'analisi matematica, dall'analisi complessa alla teoria delle equazioni alle derivate parziali, dalla teoria ergodica all'analisi convessa e all'ottimizzazione, e in altri campi ancora.

I primi due capitoli di questa tesi sono dedicati all'analisi dei legami del teorema di Hahn-Banach con altri significativi enunciati dell'analisi funzionale; vedremo in particolare il collegamento del teorema con il paradosso di Banach-Tarski, da cui segue l'importante corollario di esistenza di un insieme non misurabile secondo Lebesgue, e con l'assioma della scelta. La dimostrazione del teorema di Hahn-Banach si basa classicamente sull'uso del lemma di Zorn, il quale è equivalente all'assioma della scelta nell'ambito della teoria degli insiemi secondo gli assiomi di Zermelo-Fraenkel; al contrario, il teorema di Hahn-Banach non implica l'assioma della scelta, anche se non daremo la dimostrazione completa di questo fatto.

Nel primo capitolo proveremo varie forme equivalenti del teorema di Hahn-Banach, con particolare interesse alla versione del teorema nell'ambito delle algebre booleane. Grazie a questa specifica versione del teorema, mostreremo nel secondo capitolo che il paradosso di Banach-Tarski è conseguenza del teorema di Hahn-Banach, e quindi a sua volta non implica l'assioma della scelta. Dal paradosso di Banach-Tarski dedurremo, come corollario, l'esistenza di un sottoinsieme di $\mathbb{R}^2$ e di un sottoinsieme di $\mathbb{R}^3$ entrambi non misurabili secondo Lebesgue.

Il terzo e ultimo capitolo è rivolto alla ricerca e alla studio di presupposti sullo spazio normato, su cui è definita l'estensione del funzionale data dal teorema di Hahn-Banach, che garantiscano l'unicità di tale estensione, e di condizioni sufficienti e necessarie sullo spazio normato di arrivo affinché valga un teorema di estensione lineare continua che generalizzi quello di Hahn-Banach. In altre parole, per quanto riguarda quest'ultimo problema, analizzeremo la possibilità di generalizzare il teorema al caso in cui l'operatore da estendere non sia un funzionale a valori reali, ma sia invece a valori in uno spazio di Banach generico. Individueremo una classe particolare di spazi di Banach per cui l'enunciato di Hahn-Banach è ancora valido, e mostreremo alcuni controesempi.

1 I Teoremi di Hahn-Banach

1.1 Introduzione alle algebre booleane

In seguito risulterà necessaria una forma del teorema di Hahn-Banach in un particolare ambito: quello delle algebre booleane, che presentiamo ora.

Definizione 1.1. Un insieme $(X, \preceq)$ parzialmente ordinato si dice un reticolo se, per ogni coppia di suoi elementi x, y, il sottoinsieme $\{x, y\} \subseteq X$ ammette estremo superiore e inferiore in X, intesi rispettivamente come il minimo dei maggioranti e il massimo dei minoranti rispetto all'ordine parziale $\preceq$. Con la notazione $(X, \preceq, \wedge, \vee)$ indichiamo un reticolo in cui sono definite due operazioni binarie $\vee, \wedge : X \times X \longrightarrow X$ tali che

$$(*) \qquad \inf\{x, y\} = x \wedge y, \qquad \sup\{x, y\} = x \vee y,$$

e per le quali assumiamo i seguenti assiomi:

$$\begin{aligned} x \wedge y = y \wedge x, &\qquad x \vee y = y \vee x, \\ x \wedge (y \wedge z) = (x \wedge y) \wedge z, &\qquad x \vee (y \vee z) = (x \vee y) \vee z, \\ (x \wedge y) \vee y = y, &\qquad (x \vee y) \wedge y = y. \end{aligned}$$

In più diciamo che il reticolo è distributivo se per ogni $x, y, z \in X$ valgono

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \qquad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

Definizione 1.2. Sia $(X, \preceq, \wedge, \vee)$ un reticolo avente un elemento minimo e uno massimo, indicati rispettivamente con i simboli 0 e 1. Essi risultano (grazie a $(*)$) elementi neutri per $\vee$ e $\wedge$, ovvero per ogni $x \in X$ si ha

$$0 \vee x = x \quad \text{e} \quad 1 \wedge x = x;$$

di conseguenza (grazie alla terza riga di assiomi) $0 \wedge x = 0$ e $1 \vee x = 1$, da cui

$$x \wedge x = x \quad \text{e} \quad x \vee x = x.$$

Indichiamo un tale reticolo con la sestupla $(X, \preceq, \wedge, \vee, 0, 1)$.
Dato $x \in X$, chiamiamo complementare di x un elemento $y \in X$ che soddisfa

$$x \wedge y = 0 \quad \text{e} \quad x \vee y = 1.$$

Ad esempio 0 e 1 sono tra loro complementari, visto che $0 \wedge 1 = 0$ e $0 \vee 1 = 1$.
Un reticolo $(X, \preceq, \wedge, \vee, 0, 1)$ si dice complementato se ogni suo elemento ha almeno un complementare.

Osservazione 1.3. Se $(X, \preceq, \wedge, \vee, 0, 1)$ è distributivo, allora ogni suo elemento x ha al massimo un complementare. Infatti siano $y, y^{'}$ due complementari di x, allora $x \wedge y = x \wedge y^{'} = 0,\ x \vee y = x \vee y^{'} = 1$. Quindi, per distributività,

$$y^{'} = y^{'} \vee 0 = y^{'} \vee (x \wedge y) = \left(y^{'} \vee x\right) \wedge \left(y^{'} \vee y\right) = 1 \wedge \left(y^{'} \vee y\right) = y^{'} \vee y,$$

$$y = y \wedge 1 = y \wedge \left(x \vee y^{'}\right) = (y \wedge x) \vee \left(y \wedge y^{'}\right) = 0 \vee \left(y \wedge y^{'}\right) = y \wedge y^{'}.$$

Scambiando i ruoli di $y^{'}$ e y si ha $y^{'} = y^{'} \wedge y,\ y = y^{'} \vee y$ dunque

$$y = y \wedge y^{'} = y^{'} \wedge y = y^{'}.$$

Se $(X, \preceq, \wedge, \vee, 0, 1)$ è distributivo, per ogni $x \in X$ indichiamo il complementare di x, quando esiste (se esiste è unico per l'osservazione 1.3), con il simbolo x^C.
In tal caso, per definizione valgono

$$x \wedge x^C = 0, \quad x \vee x^C = 1, \quad \left(x^C\right)^C = x$$

(quindi elementi diversi hanno complementari diversi) e grazie agli assiomi sopra elencati (quando i seguenti complementari esistono) valgono le regole di De Morgan

$$(x \wedge y)^C = x^C \vee y^C \quad \text{e} \quad (x \vee y)^C = x^C \wedge y^C.$$

Definizione 1.4. Un' *algebra booleana* $(X, \preceq, \wedge, \vee, 0, 1)$ è un reticolo complementato e distributivo.

Osservazione 1.5. Per definizione, le relazioni $(*)$ sono ovviamente equivalenti a

$$x \preceq y \quad \Longleftrightarrow \quad [x \wedge y = x \quad \text{e} \quad x \vee y = y].$$

Di conseguenza

$$(x \wedge y) \preceq x \preceq (x \vee y) \qquad \text{per ogni } x, y \in X.$$

In particolare,

$$x \vee y = 0 \quad \Longleftrightarrow \quad [x = 0 \quad \text{e} \quad y = 0],$$

dato che $x \preceq (x \vee y) = 0,\ y \preceq (x \vee y) = 0$, e 0 è elemento minimo rispetto a $\preceq$.

Per quanto detto, in un'algebra booleana ogni elemento x ha esattamente un complementare. In altre parole un'algebra booleana è una sestupla

$$(X, \preceq, \wedge, \vee, 0, 1)$$

dove X è un insieme parzialmente ordinato con un elemento massimo 1 e uno minimo 0, rispettivamente elementi neutri per le due operazioni binarie $\wedge$ e $\vee$ che verificano gli assiomi descritti, e un'operazione biunivoca $X \longrightarrow X$ di passaggio al complementare. Per semplicità indichiamo un'algebra booleana $(X, \preceq, \wedge, \vee, 0, 1)$ solo con il simbolo X, sottointendendo la struttura e le operazioni che la costituiscono.

Notiamo che la più piccola algebra booleana (non banale) è l'insieme con solo due elementi $X = \{0, 1\}$, dove $0 \preceq 1$.
L'esempio più importante e intuitivo di algebre booleane è quello delle algebre di sottoinsiemi $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ di un insieme Ω. Infatti vale la corrispondenza

$$\left(X, \preceq, \wedge, \vee, 0, 1, {}^C\right) \longleftrightarrow \left(\mathcal{A}, \subseteq, \cap, \cup, \emptyset, \Omega, {}^C\right)$$

dato che per definizione $\mathcal{A}$ contiene Ω e $\emptyset$, ed è chiusa per intersezione finita, unione finita e passaggio al complementare. È chiaro che $\subseteq$ è un ordine parziale sull'algebra $\mathcal{A}$ e per ogni $A, B \in \mathcal{A}$ valgono ovviamente

$$\inf \{A, B\} = A \cap B, \qquad \sup \{A, B\} = A \cup B.$$

Non solo: vedremo che in generale un'algebra booleana non differisce molto da un'algebra di sottoinsiemi. In effetti mostreremo che ogni algebra booleana è immagine omomorfa di un'algebra di sottoinsiemi di un opportuno insieme.
In ogni caso la differenza concettuale tra un'algebra di sottoinsiemi e un'algebra booleana consiste nel fatto che nel primo caso possiamo considerare i singoli punti che costituiscono ogni sottoinsieme, mentre in generale ciò non è possibile per un elemento di un'algebra booleana.

Premettiamo la seguente formalizzazione.

Definizione 1.6. Un *omomorfismo booleano* $f : X \longrightarrow Y$ tra due algebre booleane X e Y è un'applicazione che preserva le operazioni fondamentali, sarebbe a dire

$$f(0_X) = 0_Y, \quad f(1_X) = 1_Y,$$

$$f\left(x^C\right) = (f(x))^C \quad \forall x \in X,$$

$$f(x_1 \wedge_X x_2) = f(x_1) \wedge_Y f(x_2), \quad f(x_1 \vee_X x_2) = f(x_1) \vee_Y f(x_2) \quad \forall x_1, x_2 \in X.$$

Spesso ometteremo i pedici $_X$ e $_Y$ in quanto sarà chiaro dal contesto a cosa ci stiamo riferendo. Per quanto riguarda l'ordine parziale, automaticamente anche esso si conserva nel senso che $x_1 \preceq_X x_2$ implica $f(x_1) \preceq_Y f(x_2)$: infatti per definizione di algebra booleana e omomorfismo booleano vale

$$x_1 \preceq_X x_2 \quad \Leftrightarrow \quad [x_1 \wedge x_2 = x_1 \quad \text{e} \quad x_1 \vee x_2 = x_2] \Rightarrow$$

$$\Rightarrow \quad [f(x_1 \wedge x_2) = f(x_1) \quad \text{e} \quad f(x_1 \vee x_2) = f(x_2)] \quad \Leftrightarrow$$

$$\Leftrightarrow \quad [f(x_1) \wedge f(x_2) = f(x_1) \quad \text{e} \quad f(x_1) \vee f(x_2) = f(x_2)] \quad \Leftrightarrow \quad f(x_1) \preceq_Y f(x_2).$$

Un omomorfismo booleano bigettivo $f : X \longrightarrow Y$ si dice un isomorfismo booleano, nel qual caso usiamo la notazione $X \cong Y$.

Osservazione 1.7. Nella definizione di omomorfismo booleano $f : X \longrightarrow Y$ basta che f soddisfi $f\left(x^C\right) = (f(x))^C$ e $f(x_1 \vee x_2) = f(x_1) \vee f(x_2)$: le altre condizioni seguono di conseguenza.
Infatti siano $y_1 = x_1^C$, $y_2 = x_2^C$, allora si ha $f(y_1 \wedge y_2) = f\left(x_1^C \wedge x_2^C\right) = f\left((x_1 \vee x_2)^C\right) = (f(x_1 \vee x_2))^C = (f(x_1) \vee f(x_2))^C = (f(x_1))^C \wedge (f(x_2))^C = f\left(x_1^C\right) \wedge f\left(x_2^C\right) = f(y_1) \wedge f(y_2)$. A questo punto da $0 = 0 \wedge 0^C$ deriva $f(0) = f(0) \wedge (f(0))^C = 0$, per cui passando al complementare $f(1) = f\left(0^C\right) = (f(0))^C = 0^C = 1$.

Dopo questa prima presentazione delle algebre booleane, introduciamo altre definizioni e risultati che ci serviranno a breve.

Definizione 1.8. Data un'algebra booleana X, una sottoalgebra booleana S di X è un sottoinsieme $S \subseteq X$, contenente 0 e 1, chiuso rispetto alle operazioni $\wedge, \vee, ^C$. In particolare S stessa è ancora un'algebra booleana, considerando le restrizioni delle operazioni $\wedge, \vee, ^C$ definite su X.
Sia G un sottoinsieme di un'algebra booleana X: possiamo allora considerare la sottoalgebra booleana S generata da G, definita come la più piccola sottoalgebra booleana che contiene G.

Osservazione 1.9. Possiamo descrivere esplicitamente la sottoalgebra booleana S generata da un sottoinsieme G di un'algebra booleana X. Sia

$$G_C = \left\{x \in X : \; x \in G \quad \text{oppure} \quad x^C \in G\right\}.$$

Allora S coincide con l'insieme S_0 costituito da tutti e soli gli elementi della forma

$$s = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} g_{i,j}$$

al variare di $I, J_1, ..., J_I \in \mathbb{N}$ e dei $g_{i,j} \in G_C$. Una scrittura di questo tipo si dice forma normale per l'elemento s. In effetti tale insieme S_0 contiene G (essendo $G \subseteq G_C$), è chiuso per costruzione rispetto alle operazioni, ammette 0 e 1 come elementi (rispettivamente come casi degeneri di $\bigvee$ di zero elementi e di $\bigwedge$ di zero elementi). Perciò S_0 è una sottoalgebra di X contenente G. Non solo: è la più piccola contenente G. Infatti supponiamo per assurdo esista un'altra sottoalgebra S_1, contenente G, ma strettamente contenuta in S_0. Allora esiste un $s \notin S_1$ che si scrive in forma normale come $s = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} g_{i,j}$, dove i $g_{i,j}$ appartengono a G_C. Questo significa che almeno uno di questi $g_{i,j}$, che indichiamo semplicemente con g, non appartiene a S_1 (altrimenti, per chiusura rispetto a $\bigwedge$ e $\bigvee$, s apparterrebbe a S_1). Dato che $g \in G_C$ e $G \subseteq S_1$ si ha $g^C \in S_1$ e questo, per chiusura di S_1 rispetto al passaggio al complementare, implica $g \in S_1$, che è assurdo. Abbiamo in questo modo provato che S_0 è la sottoalgebra generata da G.

Siano X un'algebra booleana e x un suo elemento. Per ogni $n \in \mathbb{N}$ definiamo

$$x^{C^n} = \begin{cases} x & \text{se } n \text{ è pari,} \\ x^C & \text{se } n \text{ è dispari.} \end{cases}$$

Lemma 1.10. *(Criterio di estensione di Sikorski)*
Siano X un'algebra booleana, S una sua sottoalgebra booleana generata da $G \subseteq X$. Supponiamo esista una applicazione $f : G \longrightarrow Y$, dove Y è un'altra algebra booleana. Allora f può essere estesa ad un omomorfismo booleano $F : S \longrightarrow Y$ se e solo se soddisfa la condizione seguente:

$$\begin{array}{cc} \textit{se} & g_1^{C^{n_1}} \wedge g_2^{C^{n_2}} \wedge ... \wedge g_k^{C^{n_k}} = 0 \\ \textit{allora} & (f(g_1))^{C^{n_1}} \wedge (f(g_2))^{C^{n_2}} \wedge ... \wedge (f(g_k))^{C^{n_k}} = 0 \end{array}$$

per ogni $k \in \mathbb{N}_0$ e per ogni scelta di $g_1, ..., g_k \in G$ e $n_1, ..., n_k \in \mathbb{N}$.
In tal caso, l'estensione F è univocamente determinata.

Dimostrazione. Se l'omomorfismo booleano F è un'estensione di f, allora è ovvio che vale la condizione sopra scritta, visto che un omomorfismo booleano preserva $\wedge$ e il passaggio al complementare.
Vediamo il viceversa, più complicato. L'unicità dell'estensione $F : S \longrightarrow Y$ è immediata. Infatti, per ipotesi di omomorfismo booleano, F deve soddisfare l'equazione

$$(1) \qquad F(s) = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (f(g_{i,j}))^{C^{n(i,j)}} \quad \text{per} \quad s = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (g_{i,j})^{C^{n(i,j)}}.$$

Inoltre per l'osservazione 1.9, ogni $s \in S$ può essere scritto come combinazione opportuna di $g_{i,j}$ e $(g_{i,j})^C$, quindi ha una forma normale come quella scritta sopra; perciò l'omomorfismo F (se esiste) è univocamente determinato dall'uguaglianza (1). Per l'esistenza dobbiamo provare che (1) determina davvero una funzione, ovvero che F è ben definita. In generale $s \in S$ potrebbe avere più di una rappresentazione in forma normale, dunque è necessario verificare che il valore $F(s)$ non dipende da tale rappresentazione. Fatto ciò, concluderemo provando che F definita da (1) è a tutti gli effetti un omomorfismo booleano.
Supponiamo che

$$\bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (g_{i,j})^{C^{n(i,j)}} = s = \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} (h_{k,l})^{C^{m(k,l)}}$$

con i $g_{i,j}$ e gli $h_{k,l}$ elementi di G; siano

$$\varphi_1 = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (f(g_{i,j}))^{C^{n(i,j)}} \quad \text{e} \quad \varphi_2 = \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} (f(h_{k,l}))^{C^{m(k,l)}}.$$

Dobbiamo verificare che $\varphi_1 = \varphi_2$, sfruttando la condizione data dall'ipotesi. Sia

$$S_J = \left\{\sigma = (\sigma_1, ..., \sigma_I) \in \mathbb{N}^I : \quad 1 \leqslant \sigma_i \leqslant J_i \quad \text{per } i = 1, ..., I\right\}.$$

Per le regole di De Morgan e la proprietà distributiva, usando la prima forma normale su s, si ha

$$(2) \qquad s^C = \left[\bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (g_{i,j})^{C^{n(i,j)}}\right]^C = \bigwedge_{i=1}^{I} \bigvee_{j=1}^{J_i} (g_{i,j})^{C^{n(i,j)+1}} = \bigvee_{\sigma \in S_J} \bigwedge_{i=1}^{I} (g_{i,\sigma_i})^{C^{n(i,\sigma_i)+1}}.$$

Infatti nella prima espressione (delle ultime due) il generico elemento è della forma (omettendo gli esponenti per comodità)

$$(*) \qquad (g_{1,1} \vee ... \vee g_{1,J_1}) \wedge (g_{2,1} \vee ... \vee g_{2,J_2}) \wedge ... \wedge (g_{I,1} \vee ... \vee g_{I,J_I}),$$

mentre nella seconda espressione (delle ultime due) il generico elemento è $\bigvee$, al variare dei $\sigma \in S_J$, di elementi della forma

$$g_{1,\sigma_1} \wedge ... \wedge g_{I,\sigma_I}, \qquad \text{ove} \quad \sigma = (\sigma_1, ..., \sigma_I) \in S_J :$$

ciò corrisponde a prendere in $(*)$ g_{1,σ_1} nel primo $\bigvee$, g_{2,σ_2} nel secondo $\bigvee$, ..., g_{I,σ_I} nell'ultimo $\bigvee$. Al variare di $\sigma \in S_J$ abbiamo tutti i modi possibili di prendere un elemento da ogni $\bigvee$ che compone l'espressione $(*)$, dunque l'ultima uguaglianza scritta in (2) è effettivamente verificata.

Possiamo adesso sfruttare (2), la seconda forma normale di s e la proprietà distributiva del reticolo X, giungendo alla seguente espressione:

$$0 = s \wedge s^C = \left[\bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} (h_{k,l})^{C^{m(k,l)}} \right] \bigwedge \left[\bigvee_{\sigma \in S_J} \bigwedge_{i=1}^{I} (g_{i,\sigma_i})^{C^{n(i,\sigma_i)+1}} \right] =$$

$$= \bigvee_{\sigma \in S_J} \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[(g_{i,\sigma_i})^{C^{n(i,\sigma_i)+1}} \wedge (h_{k,l})^{C^{m(k,l)}} \right].$$

Come osservato in 1.5, questo significa che per ogni $\sigma \in S_J$ e $k = 1, ..., K$ vale

$$0 = \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[(g_{i,\sigma_i})^{C^{n(i,\sigma_i)+1}} \wedge (h_{k,l})^{C^{m(k,l)}} \right];$$

allora, per la condizione data dall'ipotesi, per ogni $\sigma \in S_J$ e $k = 1, ..., K$ si ha

$$0 = \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[(f(g_{i,\sigma_i}))^{C^{n(i,\sigma_i)+1}} \wedge (f(h_{k,l}))^{C^{m(k,l)}} \right].$$

Quindi

$$0 = \bigvee_{\sigma \in S_J} \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[(f(g_{i,\sigma_i}))^{C^{n(i,\sigma_i)+1}} \wedge (f(h_{k,l}))^{C^{m(k,l)}} \right] = (\varphi_1)^C \wedge \varphi_2,$$

dove nell'ultimo passaggio abbiamo utilizzato (all'indietro) la proprietà distributiva. Allo stesso modo scambiando le due rappresentazioni di s otteniamo $0 = \varphi_1 \wedge (\varphi_2)^C$ da cui, passando al complementare, $1 = (\varphi_1)^C \vee \varphi_2$. Le due condizioni

$$\begin{cases} 0 = (\varphi_1)^C \wedge \varphi_2 \\ 1 = (\varphi_1)^C \vee \varphi_2 \end{cases}$$

implicano $\varphi_2 = \left((\varphi_1)^C\right)^C = \varphi_1$, dunque F è ben definita.
Rimane da far vedere che F è un omomorfismo booleano, e a tal proposito mostriamo (vedi osservazione 1.7) che F preserva $\vee$ ed il passaggio al complementare. Il primo caso è immediato e segue dalla definizione (1). Inoltre F preserva anhe il passaggio al complementare, infatti

$$(F(s))^C = \bigvee_{\sigma \in S_J} \bigwedge_{i=1}^{I} (f(g_{i,\sigma_i}))^{C^{n(i,\sigma_i)+1}} = F\left(s^C\right),$$

dove per la prima uguaglianza abbiamo usato la (1), per la seconda la (2), e in entrambe le regole di De Morgan. □

Proposizione 1.11. *Sia X un'algebra booleana, allora esistono un'algebra di sottoinsiemi $\mathcal{S} \subseteq \mathcal{P}(\Omega)$ e un omomorfismo booleano surgettivo f tale che*

$$f : \mathcal{S} \longrightarrow X.$$

Dimostrazione. Per ogni $x \in X$, consideriamo la famiglia di sottoinsiemi di X

$$J_x = \{A \subseteq X : \ x \in A\} \subseteq \mathcal{P}(X).$$

Allora $J = \{J_x : x \in X\}$ è una famiglia di sottoinsiemi di $\mathcal{P}(X)$; sia $\mathcal{S}$ l'algebra di sottoinsiemi di $\mathcal{P}(X)$ generata da J.
Mostriamo che la mappa surgettiva $f_0 : J \longrightarrow X$ definita da $f_0(J_x) = x$ soddisfa la condizione di Sikorski. In questo modo, per il lemma 1.10, f_0 si estende univocamente ad un omomorfismo booleano surgettivo $f : \mathcal{S} \longrightarrow X$.
Siano $x_1, ..., x_m$ e $y_1, ..., y_n$ elementi di X, tali che

$$(*) \qquad (J_{x_1} \cap ... \cap J_{x_m}) \cap \left[(J_{y_1})^C \cap ... \cap (J_{y_n})^C\right] = \emptyset.$$

Dobbiamo provare che

$$(x_1 \wedge ... \wedge x_m) \wedge \left(y_1^C \wedge ... \wedge y_n^C\right) = 0.$$

Notiamo che, per definizione di J_x, $(*)$ si può riscrivere come

$$\{S \subseteq X : \quad \{x_1, ..., x_m\} \subseteq S \subseteq (X \setminus \{y_1, ..., y_n\})\} = \emptyset,$$

da cui $\{x_1, ..., x_m\}$ non è contenuto in $(X \setminus \{y_1, ..., y_n\})$. Questo significa che esiste almeno un $x_i \notin (X \setminus \{y_1, ..., y_n\})$, ovvero x_i coincide con uno dei y_j.
Dunque $x_i \wedge y_j^C = x_i \wedge x_i^C = 0$; a maggior ragione

$$(x_1 \wedge ... \wedge x_m) \wedge \left(y_1^C \wedge ... \wedge y_n^C\right) = 0$$

come volevasi dimostrare. □

Definizione 1.12. Siano X un'algebra booleana e I un suo sottoinsieme.
Diciamo che I è un *ideale* di X se I contiene 0 e soddisfa

$$i_1, i_2 \in I \Rightarrow i_1 \vee i_2 \in I, \qquad i \in I, x \in X \Rightarrow i \wedge x \in I.$$

Notiamo la somiglianza con la definizione di ideale di un anello (in tal caso $+$ e $\times$ giocano il ruolo di $\vee$ e $\wedge$). Un ideale I di X si dice proprio se è un sottoinsieme proprio di X: per definizione I è proprio se e solo se $1 \notin I$. Osserviamo che ogni

algebra booleana diversa da $\{0\}$ ammette sempre l'ideale proprio $\{0\}$, e che il nucleo di un omomorfismo booleano $f : X \longrightarrow Y$ è un ideale (per di più un ideale proprio se $Y \neq \{0\}$).
Nel caso particolare dell'algebra booleana $\mathcal{P}(\Omega)$ la definizione di ideale assume la seguente forma semplificata. Dato un insieme Ω, una famiglia non vuota $I \subseteq \mathcal{P}(\Omega)$ si dice un ideale di $\mathcal{P}(\Omega)$ se I è chiuso per unione finita e per ogni $S \in I$ si ha $\mathcal{P}(S) \subseteq I$. Un ideale I di $\mathcal{P}(\Omega)$ si dice proprio se $I \neq \mathcal{P}(\Omega)$: osserviamo che per definizione I è proprio se e solo se $\Omega \notin I$.

Osservazione 1.13. Analogamente al caso degli anelli, anche per le algebre booleane è possibile quozientare per un ideale. Vediamo come.
Dato un ideale I di un'algebra booleana X, introduciamo la relazione su X

$$x \sim y \quad \Longleftrightarrow \quad \left[\left(x \wedge y^C\right) \in I \quad \text{e} \quad \left(x^C \wedge y\right) \in I\right].$$

La relazione $\sim$ appena introdotta è una relazione di equivalenza in quanto è chiaramente simmetrica e riflessiva ($x \sim x$ perché $x \wedge x^C = 0 \in I$) e pure transitiva. Verifichiamo quest'ultima: siano $x \sim y$ e $y \sim z$, dobbiamo provare che $x \sim z$ o equivalentemente $\left(x \wedge z^C\right), \left(x^C \wedge z\right) \in I$. In effetti valgono queste due appartenenze (vediamo solo la prima), visto che

$$x \wedge z^C = \left(x \wedge z^C\right) \wedge \left(y^C \vee y\right) = \left(x \wedge z^C \wedge y^C\right) \vee \left(z^C \wedge x \wedge y\right) =$$
$$\left[\left(x \wedge z^C\right) \wedge \left(x \wedge y^C\right)\right] \vee \left[\left(x \wedge z^C\right) \wedge \left(z^C \wedge y\right)\right] = (x \wedge z^C) \wedge \left[(x \wedge y^C) \vee (y \wedge z^C)\right] \in I,$$

dove abbiamo usato le proprietà di chiusura dell'ideale I e le ipotesi $x \sim y,\ y \sim z$.
Quindi $\sim$ è una relazione di equivalenza che decompone X in classi di equivalenza disgiunte: denotiamo con il simbolo $[x]$ la classe di equivalenza (modulo $\sim$) di un elemento $x \in X$. Indichiamo con X/I l'insieme di queste classi di equivalenza.
Il quoziente X/I eredita la struttura booleana da X: infatti, definendo le operazioni (indotte dalle corrispondenti operazioni di X)

$$[x] \wedge [y] = [x \wedge y], \quad [x] \vee [y] = [x \vee y], \quad [x]^C = [x^C],$$

possiamo verificare che quelle appena introdotte sono operazioni ben definite (nel senso che il risultato non dipende dai rappresentanti scelti per la classe) e soddisfano tutti gli assiomi. Pertanto X/I è una algebra booleana.
Dalla definizione delle operazioni su X/I segue immediatamente che la proiezione canonica sul quoziente $\pi : X \longrightarrow X/I$, che manda ogni elemento nella rispettiva classe di equivalenza ($\pi(x) = [x]$), è un omomorfismo booleano.
Per di più, se $f : X \longrightarrow Y$ è un omomorfismo booleano surgettivo allora f induce un isomorfismo booleano

$$[f] : X/\ker(f) \longrightarrow Y, \quad \text{dato da} \quad [f]([x]) = f(x).$$

Definizione 1.14. Sia X un'algebra booleana diversa da $\{0\}$, si dice *carica di probabilità* su X una mappa $\mu : X \longrightarrow [0,1]$ tale che $\mu(1) = 1$ e per la quale valga $\mu\,(x \vee y) = \mu\,(x) + \mu\,(y)$ per ogni coppia di elementi $x, y \in X$ tali che $x \wedge y = 0$. In altre parole si tratta di una misura di probabilità finitamente additiva sui disgiunti (intesi come gli elementi che soddisfano $x \wedge y = 0$) rispetto all'operazione $\vee$.

Certamente se μ è una carica di probabilità su un'algebra booleana, allora $\mu(0) = 0$ (dato che 0 è disgiunto da se stesso); perciò l'immagine di una carica di probabilità μ contiene sempre $\{0,1\}$. Inoltre per ogni $x \in X$ vale $1 = \mu(1) = \mu(x) + \mu(x^C)$ (per definizione x e x^C sono disgiunti e $x \vee x^C = 1$). Vediamo ora altre proprietà di μ.

Lemma 1.15. *Sia $\mu : X \longrightarrow [0,1]$ una carica di probabilità; allora per ogni $x, y \in X$ si ha*

$$(i) \quad x \preceq y \Rightarrow \mu(x) \leqslant \mu(y), \qquad (ii) \quad \mu(x \vee y) \leqslant \mu(x) + \mu(y).$$

Dimostrazione. (i) Dato che $\mu \geqslant 0$, si ha $\mu(x) \leqslant \mu(x) + \mu(y \wedge x^C) = \mu(y)$ in quanto x e $y \wedge x^C$ sono disgiunti (per associatività di $\wedge$) e $x \vee (y \wedge x^C) = x \vee y$ (per la proprietà distributiva); a questo punto $x \vee y = y$ perché per ipotesi $x \preceq y$.
(ii) Allo stesso modo $\mu(x \vee y) = \mu(x) + \mu(y \wedge x^C) \leqslant \mu(x) + \mu(y)$, dove nell'ultimo passaggio abbiamo usato (i) e il fatto che $\left(y \wedge x^C\right) \preceq y$. $\square$

Osservazione 1.16. Se $\mu : X \longrightarrow [0,1]$ è una carica di probabilità su un'algebra booleana X, allora $\ker(\mu)$ è un ideale proprio di X. Intanto è proprio perché $1 \notin I$ ($\mu(1) = 1$). Ovviamente $0 \in \ker(\mu)$ (essendo $\mu(0) = 0$), e per il lemma precedente $\ker(\mu)$ è chiuso internamente per $\vee$ (usando (ii), se $x_1, x_2 \in \ker(\mu)$ si ha $\mu(x_1 \vee x_2) \leqslant \mu(x_1) + \mu(x_2) = 0 + 0 = 0$, cioè $(x_1 \vee x_2) \in \ker(\mu)$) ed esternamente rispetto a $\wedge$ (usando (i), se $x \in \ker(\mu)$, $y \in X$ si ha $\mu(x \wedge y) \leqslant \mu(x) = 0$, cioè $(x \wedge y) \in \ker(\mu)$).

Proposizione 1.17. *Data una famiglia di algebre booleane $\{X_i\}_{i \in I}$, esistono un'algebra booleana, detta somma booleana di $\{X_i\}_{i \in I}$ e indicata con il simbolo $\sum_{i \in I} X_i$, e una famiglia $\{f_i\}_{i \in I}$ di omomorfismi booleani $f_i : X_i \longrightarrow \sum_{i \in I} X_i$ che soddisfano la seguente proprietà universale: se B è un'algebra booleana per cui per ogni $i \in I$ esiste un omomorfismo booleano $g_i : X_i \longrightarrow B$, allora esiste un unico omomorfismo booleano $g : \sum_{i \in I} X_i \longrightarrow B$ tale che per ogni $i \in I$ valga $g \circ f_i = g_i$ (ovvero esiste un unico g che fa commutare il diagramma).*

Per questo risultato e per approfondimenti sulla teoria delle algebre booleane, rimandiamo a [4] e [7].

1.2 Il ruolo dell'assioma della scelta

L'assioma della scelta (AC) è lo strumento fondamentale nella dimostrazione del teorema di Hahn-Banach (HB).
Viceversa, HB non implica l'assioma della scelta, come spiegheremo nell'ultimo paragrafo di questo primo capitolo. Ciò significa che suppore valido HB è una condizione più debole di AC.
Vediamo ora come AC, anzi una delle sue forme equivalenti, il lemma di Zorn, entra in gioco nella dimostrazione di HB.

Lemma 1.18. *(Zorn) Sia Σ un insieme non vuoto parzialmente ordinato in cui ogni catena ammette maggiorante in Σ; allora Σ ha un elemento massimale.*

Definizione 1.19. Detto X un $\mathbb{R}$-spazio vettoriale, una funzione $p : X \longrightarrow \mathbb{R}$ si dice subadditiva e positivamente omogenea se

$$p(x+y) \leqslant p(x) + p(y) \qquad \forall x, y \in X,$$

$$p(tx) = tp(x) \qquad \forall x \in X, \quad \forall t \geqslant 0.$$

Teorema 1.20. *(HB classico) (AC)*
Siano X un $\mathbb{R}$-spazio vettoriale, Y un suo sottospazio, $p : X \longrightarrow \mathbb{R}$ una funzione subadditiva e positivamente omogenea.
Sia $\lambda_0 : Y \longrightarrow \mathbb{R}$ una mappa lineare tale che $\lambda_0 \leqslant p$ su Y. Allora λ_0 si estende ad una mappa lineare $\lambda : X \longrightarrow \mathbb{R}$ che soddisfa $\lambda \leqslant p$ su tutto X.

Dimostrazione. Sia Σ l'insieme delle estensioni lineari h di λ_0 definite su un sottospazio $D(h)$ di X, con $Y \subseteq D(h)$ e $h \leqslant p$ su $D(h)$. Σ è ovviamente non vuoto dato che per ipotesi $\lambda_0 \in \Sigma$. Su Σ definiamo l'ordine parziale

$$h_1 \preceq h_2 \quad \Longleftrightarrow \quad [D(h_1) \subseteq D(h_2) \quad \text{e} \quad h_2 \text{ estende } h_1] \,.$$

Mostriamo che ogni catena in Σ ammette maggiorante in Σ, così da poter applicare il lemma di Zorn. Sia $H = (h_i)_{i \in I} \subseteq \Sigma$ una catena, cioè un sottoinsieme totalmente ordinato; allora possiamo costruire l'ulteriore estensione

$$h : \bigcup_{i \in I} D(h_i) \longrightarrow \mathbb{R}, \qquad \text{con} \quad h(x) = h_i(x) \quad \text{se } x \in D(h_i).$$

In questo modo h è l'elemento maggiorante per la catena, e $h \in \Sigma$ per definizione di Σ, sfruttando l'ordinamento totale della catena e le proprietà degli $h_i \in \Sigma$.

Per il lemma di Zorn esiste in Σ un'estensione massimale λ di λ_0.
Se $D(\lambda) = X$ abbiamo finito perché λ è l'estensione di λ_0 su tutto X con le proprietà richieste. Mostriamo che questo è l'unico caso possibile.
Supponiamo $D(\lambda) \subsetneq X$, esiste allora $x_0 \in X \setminus D(\lambda)$. Posto $D(h) = \langle D(\lambda), x_0 \rangle$ il sottospazio generato da $D(\lambda)$ e x_0, definiamo l'estensione lineare $h : D(h) \longrightarrow \mathbb{R}$ di λ come

$$h(x + tx_0) = \lambda(x) + t\alpha \qquad \forall x \in D(\lambda), \quad \forall t \in \mathbb{R},$$

dove $\alpha \in \mathbb{R}$ è una costante che può essere scelta, come vedremo, in modo che $h \in \Sigma$.
Ma questo è assurdo, dato che λ per Zorn è l'elemento massimale in Σ.
Proviamo per concludere che esiste una costante α in modo che

$$\lambda(x) + t\alpha \leqslant p(x + tx_0) \quad \forall x \in D(\lambda), \quad \forall t \in \mathbb{R},$$

o in modo equivalente che per ogni $t > 0$ valgano

$$\lambda(x) + t\alpha \leqslant p(x + tx_0) \quad \text{ e } \quad \lambda(x) - t\alpha \leqslant p(x - tx_0).$$

Quindi basta che esista α tale che

$$\frac{\lambda(x_1) - p(x_1 - t_1x_0)}{t_1} \leqslant \alpha \leqslant \frac{p(x_2 + t_2x_0) - \lambda(x_2)}{t_2} \qquad \forall\ x_1, x_2 \in D(\lambda),\ \forall\ t_1, t_2 > 0$$

ovvero (ponendo $x^{'} = x_1/t_1, x^{''} = x_2/t_2$ e sfruttando l'ipotesi di omegeneità su p)

$$\sup_{x^{'} \in D(\lambda)} \left\{\lambda(x^{'}) - p(x^{'} - x_0)\right\} \leqslant \alpha \leqslant \inf_{x^{''} \in D(\lambda)} \left\{p(x^{''} + x_0) - \lambda(x^{''})\right\}.$$

Un tale α esiste visto che per ogni $x^{'}, x^{''} \in D(\lambda)$ si ha

$$\lambda(x^{'})+\lambda(x^{''}) = \lambda(x^{'}+x^{''}) \leqslant p(x^{'}+x^{''}) = p(x^{'}-x_0+x_0+x^{''}) \leqslant p(x^{'}-x_0)+p(x^{''}-x_0),$$

che implica

$$\lambda(x^{'}) - p(x^{'} - x_0) \leqslant p(x^{''} + x_0) - \lambda(x^{''}) \qquad \forall\ x^{'}, x^{''} \in D(\lambda).$$

$\square$

Mostreremo in seguito che due notevoli teoremi, tipicamente dimostrati tramite AC, sono conseguenze solo del teorema di Hahn-Banach, il quale rappresenta come spiegheremo una condizione più debole dell'assioma della scelta.
I teoremi di cui ci occuperemo sono il paradosso di Banach-Tarski e l'esistenza di un insieme non misurabile secondo Lebesgue.

1.3 Formulazioni equivalenti di HB

Il classico teorema di Hahn-Banach provato nel paragrafo precedente possiede molte forme equivalenti fra loro, alcune delle quale in ambiti non convenzionali.
In questa trattazione risulterà fondamentale la versione del teorema nell'ambito delle algebre booleane. Presentiamo ora le diverse versioni.

Ricordiamo velocemente le ben note definizioni per gli spazi normati e il loro duale. Con $(X, \|.\|)$ indichiamo un $\mathbb{R}$-spazio vettoriale X dotato di una norma $\|.\|$, cioè di una funzione $\|.\| : X \longrightarrow [0, \infty[$ avente le seguenti proprietà:
(i) $\quad \|x\| = 0 \quad \Longleftrightarrow \quad x = 0;$
(ii) $\quad \|\lambda x\| = |\lambda| \, \|x\| \quad \forall \, x \in X, \quad \forall \, \lambda \in \mathbb{R};$
(iii) $\quad \|x + y\| \leqslant \|x\| + \|y\| \quad \forall \, x, y \in X.$
La coppia $(X, \|.\|)$ (con queste proprietà) si dice spazio normato.
Uno spazio normato completo si dice spazio di Banach.

Dati due $\mathbb{R}$-spazi normati $(X, \|.\|_X)$ e $(Y, \|.\|_Y)$, un operatore lineare $\Lambda : X \longrightarrow Y$ si dice limitato se esiste una costante $M \geqslant 0$ per cui $\|\Lambda(x)\|_Y \leqslant M \, \|x\|_X$ per ogni $x \in X$. Un operatore lineare tra due spazi normati è limitato se e solo se è continuo. Indichiamo con $\mathcal{L}(X, Y)$ lo spazio dei funzionali lineari e limitati da X a Y, che è uno spazio normato mediante

$$\|\Lambda\|_{\mathcal{L}(X,Y)} = \sup_{0 \neq x \in X} \frac{\|\Lambda(x)\|_Y}{\|x\|_X}.$$

Inoltre, se Y è completo, allora $\Big(\mathcal{L}(X, Y), \|.\|_{\mathcal{L}(X,Y)}\Big)$ è uno spazio di Banach.
Denotiamo con $X^* = \mathcal{L}(X, \mathbb{R})$ lo spazio duale dell'$\mathbb{R}$-spazio normato $(X, \|.\|)$, definito come l'insieme degli operatori lineari e limitati da X in $\mathbb{R}$.
Per quanto detto, X^* è uno spazio di Banach con la norma

$$\|\Lambda\|_{X^*} = \sup_{0 \neq x \in X} \frac{|\Lambda(x)|}{\|x\|_X}.$$

Con questa premessa, possiamo subito considerare una diversa versione di HB, applicata per l'appunto a spazi normati.

Teorema 1.21. *(HB per spazi normati)*
Siano $(X, \|.\|)$ *un* $\mathbb{R}$*-spazio normato,* Y *un suo sottospazio,* $\lambda \in Y^*$. *Allora* λ *si estende a un funzionale* $\Lambda \in X^*$ *che soddisfa* $\|\Lambda\|_{X^*} = \|\lambda\|_{Y^*}$.

Dimostrazione. Basta considerare la funzione (definita su X) $p(x) = \|x\| \, \|\lambda\|_{Y^*}$, che è ovviamente subadditiva e positivamente omogenea. Per definizione di norma $\|.\|_{Y^*}$ si ha $\lambda(y) \leqslant \|y\| \, \|\lambda\|_{Y^*} = p(y)$ perciò per HB classico (teorema 1.20) λ si estende a $\Lambda \in X^*$ che verifica $\Lambda(x) \leqslant p(x) = \|x\| \, \|\lambda\|_{Y^*}$ per ogni $x \in X$. Di conseguenza $\|\Lambda\|_{X^*} \leqslant \|\lambda\|_{Y^*}$, anzi vale proprio l'uguaglianza perché Λ estende λ (quindi stiamo considerando l'estremo superiore su un insieme più ampio).

□

Proposizione 1.22. *(Funzionali unitari per vettori assegnati)*
Dato $(X, \|.\|)$ *un* $\mathbb{R}$*-spazio normato diverso da* $\{0\}$*, sia* $x_0 \in X$. *Allora esiste* $\Lambda \in X^*$ *tale che* $\|\Lambda\|_{X^*} = 1$ *e* $\Lambda(x_0) = \|x_0\|$.

Dimostrazione. Detto $Y = \langle x_0 \rangle$ il sottospazio generato da x_0, definiamo su Y la funzione $\lambda(rx_0) = r \, \|x_0\|$. È chiaro che λ è lineare e limitata con $\|\lambda\|_{Y^*} = 1$: allora per HB per spazi normati (teorema 1.21) esiste Λ, estensione lineare di λ su tutto X, che ha ancora norma unitaria. Infine, essendo un'estensione, $\Lambda(x_0) = \lambda(x_0) = \|x_0\|$.

□

Proposizione 1.23. *(Separazione di sottospazi chiusi)*
Siano B *un sottospazio chiuso di un* $\mathbb{R}-$*spazio normato* X *ed* η *un elemento di* $X \backslash B$. *Allora esiste una mappa in* X^* *che si annulla su* B *ma non su* η.

Dimostrazione. Sia $Q = X/B$ lo spazio quoziente, che è normato mediante

$$\|q\|_Q = \inf \{ \|x\|_X : \ \pi(x) = q \} ,$$

dove $\pi : X \longrightarrow Q$ è la proiezione canonica sul quoziente (in effetti si verifica facilmente che $\|\|_Q$ eredita le proprietà di norma da $\|\|_X$).
Dato che $\eta \notin B$, $\pi(\eta) \neq 0$ in Q. Allora per la proposizione 1.22 esiste $\Lambda \in Q^*$ tale che $\Lambda(\pi(\eta)) \neq 0$. Si conclude osservando che la composizione $\Lambda \circ \pi$ appartiene a X^*, non si annulla su η ma ovviamente si annulla su B visto che proiettiamo sul corrispondente quoziente. □

Teorema 1.24. *(Misura di Luxemburg)*
Per ogni insieme non vuoto Ω *e ogni ideale proprio* I *di* $\mathcal{P}(\Omega)$ *esiste una carica di probabilità* μ *su* $\mathcal{P}(\Omega)$ *che assume valore 0 su ogni elemento di* I*.*

Dimostrazione. Per ogni $E \in I$, sia $1_E : \Omega \longrightarrow [0,1]$ la funzione caratteristica corrispondente. In particolare $1_\Omega = 1$ è la funzione costante 1. Detto X lo spazio di Banach delle funzioni limitate da X a $\mathbb{R}$ con la norma del sup, sia B il sottospazio chiuso generato dalle 1_E, al variare di $E \in I$.
Mostriamo che $1 \notin B$. Sia g un qualsiasi elemento in B, ovvero una funzione del tipo $g = \sum_{i=1}^n c_i 1_{E_i}$ dove $n \in \mathbb{N}_0$, $c_i \in \mathbb{R}$, $E_i \in I$ per ogni $i = 1, ..., n$. L'insieme $E = \bigcup_{i=1}^n E_i$, per definizione di ideale proprio, è un elemento di I diverso da Ω, perciò esiste $\omega \in \Omega \setminus E$. Per costruzione $g(\omega) = 0$. Dunque $g \neq 1_\Omega$, ossia $1 \notin B$.
Per la proposizione 1.23 sulla separazione di sottospazi chiusi esiste quindi un funzionale lineare $\Lambda : X \longrightarrow \mathbb{R}$ che si annulla su B ma non sulla funzione costante 1.
Per ogni $S \subseteq \Omega$ poniamo

$$\nu(S) = \Lambda(1_S).$$

ν è una funzione finitamente additiva; infatti dati S_1 e S_2 sottoinsiemi disgiunti di Ω si ha (indicando con $\sqcup$ l'unione disgiunta)

$$\nu(S_1 \sqcup S_2) = \Lambda(1_{S_1 \sqcup S_2}) = \Lambda(1_{S_1}) + \Lambda(1_{S_2}) = \nu(S_1) + \nu(S_2)$$

grazie alla linearità di Λ. Perciò ν è una carica a valori reali su $\mathcal{P}(\Omega)$ che si annulla su ogni elemento $E \in I$ e che non si annulla su Ω. Di conseguenza la parte positiva di questa carica, data da

$$\nu^+(S) = \sup \{\nu(A) : \ A \subseteq S\}$$

è positiva ($\nu^+(S) \geqslant \nu(\emptyset) = \Lambda(1_\emptyset) = \Lambda(0) = 0$) su $\mathcal{P}(\Omega)$, si annulla su ogni elemento $E \in I$ ma non si annulla su Ω. Verifichiamo che ν^+ è finitamente additiva: dati S_1 e S_2 sottoinsiemi disgiunti di Ω, vale

$$\nu^+(S_1 \sqcup S_2) = \sup \{\nu(A) : \ A \subseteq (S_1 \sqcup S_2)\} =$$

$$= \sup \{\nu(A) : \ A = A_1 \sqcup A_2, \quad \text{dove } A_1 \subseteq S_1, \ A_2 \subseteq S_2\} = \nu^+(S_1) + \nu^+(S_2).$$

Poniamo infine, riscalando,

$$\mu(S) = \frac{\nu^+(S)}{\nu^+(\Omega)}.$$

In questo modo $\mu : \mathcal{P}(\Omega) \longrightarrow [0,1]$ è la carica di probabilità con le proprietà richieste.
$\square$

Teorema 1.25. *(HB per algebre booleane)*
Ogni algebra booleana diversa da $\{0\}$ *ammette una carica di probabilità.*

Dimostrazione. Sia X un'algebra booleana diversa da $\{0\}$. Grazie alla proposizione 1.11 esiste un omomorfismo surgettivo di algebre booleane $f : \mathcal{S} \longrightarrow X$, dove $\mathcal{S} \subseteq \mathcal{P}(\Omega)$ è un'algebra di sottoinsiemi di un certo insieme Ω. Per surgettività, come sottolineato nella parte finale dell'osservazione 1.13, $\{0\} \neq X \cong \mathcal{S}/\ker(f)$.
Perciò $\ker(f)$ è un sottoinsieme proprio in $\mathcal{S}$, anzi un ideale proprio perché (per definizione di $\ker(f)$) contiene 0 ed è chiuso internamente rispetto a $\vee$ ed esternamente rispetto a $\wedge$.
Allora per il teorema 1.24 esiste una carica di probabilità μ su $\mathcal{P}(\Omega)$ che assume valore 0 su ogni elemento di $\ker(f)$. Per restrizione $\mu : \mathcal{S} \longrightarrow [0,1]$ è ancora una carica di probabilità che assume valore 0 su ogni elemento di $\ker(f)$. Proprio per questo possiamo passare al quoziente e risulta ben definita la funzione (indicando con $[A] = A + \ker(f)$ la classe di resto modulo $\ker(f)$)

$$\nu : \mathcal{S}/\ker(f) \longrightarrow [0,1] \quad \text{dove} \quad \nu([A]) := \mu(A).$$

In questo modo ν è una carica di probabilità su $\mathcal{S}/\ker(f)$ perché eredita questa proprietà da μ: infatti $\nu([\Omega]) = \mu(\Omega) = 1$ e

$$\nu\left([A_1] \sqcup [A_2]\right) = \mu(A_1 \sqcup A_2) = \mu(A_1) + \mu(A_2) = \nu([A_1]) + \nu([A_2]).$$

Dunque $X \cong \mathcal{S}/\ker(f)$ ammette una carica di probabilità ν. □

Corollario 1.26. *Sia* X *un'algebra booleana diversa da* $\{0\}$*. Allora per ogni ideale proprio* I *di* X *esiste una carica di probabilità* μ *su* X *che si annulla su* I*.*

Dimostrazione. Dato che I è un ideale proprio (cioè $1 \notin I$), X/I è un'algebra booleana diversa da $\{0\}$. Sia $\pi : X \longrightarrow X/I$ la proiezione sul quoziente; chiamiamo ν la carica di probabilità su X/I data dal teorema 1.25.
La composizione $\mu = \nu \circ \pi : X \longrightarrow [0,1]$, che si annulla su ogni elemento di I per proiezione, è una carica di probabilità su X. Infatti $\mu(1) = \nu(\pi(1)) = \nu([1]) = 1$; d'altra parte se $x_1 \wedge x_2 = 0$ allora (applicando π)

$$\pi(x_1 \wedge x_2) = \pi(x_1) \wedge \pi(x_2) = \pi(0) = [0],$$

che implica, dato che μ è una carica di probabilità su X/I,

$$\mu(x_1 \vee x_2) = \nu(\pi(x_1 \vee x_2)) = \nu(\pi(x_1) \vee \pi(x_2)) = \nu(\pi(x_1)) + \nu(\pi(x_2)) = \mu(x_1) + \mu(x_2).$$

□

Corollario 1.27. *Supponiamo che ogni algebra booleana diversa da* $\{0\}$ *ammetta una carica di probabilità. Allora per ogni insieme non vuoto* Ω *e ogni ideale proprio* I *di* $\mathcal{P}(\Omega)$ *esiste una carica di probabilità* μ *su* $\mathcal{P}(\Omega)$ *che assume valore 0 su ogni elemento di* I. *In altre parole, i teoremi 1.24 e 1.25 sono equivalenti.*

Per completare l'equivalenza delle forme del teorema fin qui presentate, dobbiamo presentare un nuovo concetto: il limite di Banach.

Definizione 1.28. Un insieme parzialmente ordinato $(\Delta, \sqsubseteq)$ si dice diretto se per ogni $\delta_1, \delta_2 \in \Delta$ esiste $\delta \in \Delta$ tale che $\delta_1 \sqsubseteq \delta$ e $\delta_2 \sqsubseteq \delta$.

Dato un insieme diretto $(\Delta, \sqsubseteq)$, sia

$$B(\Delta, \mathbb{R}) = \{\text{funzioni limitate da } \Delta \text{ a } \mathbb{R}\},$$

ossia lo spazio dei net limitati definiti su Δ a valori in $\mathbb{R}$.
Per ogni $f \in B(\Delta, \mathbb{R})$ esistono (dato che f è a valori reali)

$$\liminf_{\delta \in \Delta} f(\delta) = \sup_{\alpha \in \Delta} \inf_{\beta \sqsupseteq \alpha} f(\beta), \qquad \limsup_{\delta \in \Delta} f(\delta) = \inf_{\alpha \in \Delta} \sup_{\beta \sqsupseteq \alpha} f(\beta).$$

In ogni caso $\liminf_{\delta \in \Delta}(f) \leqslant \limsup_{\delta \in \Delta}(f)$. Diciamo che f converge se e solo se $\liminf_{\delta \in \Delta}(f)$ e $\limsup_{\delta \in \Delta}(f)$ coincidono, nel qual caso chiamiamo limite di f il loro valore comune, indicato con $\lim_{\delta \in \Delta} f(\delta)$ o più semplicemente con $\lim(f)$. Osserviamo che possiamo interpretare il limite come un operatore lineare dal sottospazio dei net convergenti $C(\Delta, \mathbb{R})$ $(\subseteq B(\Delta, \mathbb{R}))$ a $\mathbb{R}$. Tale operatore $\lim : C(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$ è positivo nel senso che $\lim(f) \geqslant 0$ per ogni net convergente $f \geqslant 0$.

Definizione 1.29. Sia $LIM : B(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$ un funzionale lineare.
Diciamo che LIM è un *limite di Banach* se è un operatore positivo che estende il limite ordinario $\lim : C(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$. Ovvero $LIM(f) \geqslant 0$ per ogni $f \geqslant 0$ e $LIM(f) = \lim(f)$ per ogni $f \in C(\Delta, \mathbb{R})$.

Lemma 1.30. *Sia* $LIM : B(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$ *un funzionale lineare; allora le seguenti condizioni sono equivalenti:*
(i) $LIM(f) \leqslant \limsup_{\delta \in \Delta} f(\delta)$ *per ogni* $f \in B(\Delta, \mathbb{R})$*;*
(ii) $\liminf_{\delta \in \Delta} f(\delta) \leqslant LIM(f) \leqslant \limsup_{\delta \in \Delta} f(\delta)$ *per ogni* $f \in B(\Delta, \mathbb{R})$*;*
(iii) LIM *è un limite di Banach.*

Dimostrazione. È ovvio che (ii) implica (i), anzi sono equivalenti: infatti

$$\liminf_{\delta \in \Delta} f(\delta) = -\limsup_{\delta \in \Delta} -f(\delta) \leqslant -LIM(-f) = LIM(f),$$

sfruttando la linearità dell'operatore LIM e la condizione (i) per $-f$.
Chiaramente (ii) implica (iii), perché quando $\liminf_{\delta\in\Delta}(f)$ e $\limsup_{\delta\in\Delta}(f)$ coincidono, per la (ii) si ha $LIM(f) = \lim(f)$ e la positività è ovvia.
Per concludere proviamo che (iii) implica (i). Fissiamo $f \in B(\Delta, \mathbb{R})$. Per ogni $\varepsilon \in \Delta$, sia $u(\varepsilon) = \sup_{\delta \sqsupseteq \varepsilon} f(\delta)$: osserviamo che il net $\{u(\varepsilon) : \varepsilon \in \Delta\}$ converge (decrescendo) a $\limsup_{\delta\in\Delta} f(\delta)$. Inoltre essendo un estremo superiore vale $u(\varepsilon) \geqslant f(\varepsilon)$ per ogni ε, cioè $u - f \geqslant 0$. Per la condizione (iii) sappiamo che LIM è un limite di Banach, ossia un operatore lineare positivo che estende il limite ordinario sui net convergenti, perciò

$$LIM(f) \leqslant LIM(u) = \lim_{\varepsilon\in\Delta} u(\varepsilon) = \limsup_{\delta\in\Delta} f(\delta).$$

□

Se per ogni insieme diretto $(\Delta, \preceq)$ esiste un limite di Banach, sintetizziamo allora questo fatto dicendo che siamo nell'ipotesi di ″Esistenza dei limiti di Banach″.

Proposizione 1.31. *Supponiamo che per ogni insieme non vuoto Ω e ogni ideale proprio I di $\mathcal{P}(\Omega)$ esista una carica di probabilità μ su $\mathcal{P}(\Omega)$ che assume valore 0 su ogni elemento di I. Allora è garantita l'esistenza dei limiti di Banach.*

Dimostrazione. Fissato un arbitrario insieme diretto $(\Delta, \preceq)$ non vuoto, consideriamo la seguente famiglia di sottoinsiemi di Δ:

$$I = \left\{A^C \subseteq \Delta : \quad \exists\, \delta_0 \in \Delta \ \text{ tale che } \ A \supset \{\delta \in \Delta : \delta \succeq \delta_0\}\right\}.$$

In altre parole I è l'insieme dei complementari dei sottoinsiemi A di Δ contenenti insiemi della forma $\{\delta \in \Delta : \delta \succeq \delta_0\}$, al variare di $\delta_0 \in \Delta$.
Mostriamo che I è un ideale proprio di Δ. È chiaro che I è chiuso per intersezione finita: se $A_1^C, A_2^C \in I$ allora $A_1 \supset \{\delta \in \Delta : \delta \succeq \delta_1\}$ e $A_2 \supset \{\delta \in \Delta : \delta \succeq \delta_2\}$; scelto $\delta_0 \in \Delta$ tale che $\delta_0 \succeq \delta_1$ e $\delta_0 \succeq \delta_2$ (un tale δ_0 esiste perché Δ è diretto), si ha $(A_1 \cup A_2) \supset \{\delta \in \Delta : \delta \succeq \delta_0\}$, ovvero $(A_1 \cup A_2)^C = \left(A_1^C \cap A_2^C\right) \in I$. Inoltre se $B^C \subseteq A^C$ e $A^C \in I$ vale $B \supset A \supset \{\delta \in \Delta : \delta \succeq \delta_0\}$ quindi $B^C \in I$. Abbiamo provato che I è un ideale di Δ. Infine è ovvio che l'insieme vuoto non può contenere nessun insieme della forma $\{\delta \in \Delta : \delta \succeq \delta_0\}$ (questi insiemi sono non vuoti perché contengono almeno δ_0, per la riflessività di $\preceq$). Dunque $\emptyset^C = \Delta \notin I$ e I è proprio.
Per ipotesi, esiste una carica di probabilità μ su tutto $\mathcal{P}(\Delta)$ che assume valore 0 su ogni elemento dell'ideale I. Possiamo allora definire l'integrale di un net in $B(\Delta, \mathbb{R})$ rispetto alla carica di probabilità μ. Per un net $u \in B(\Delta, \mathbb{R})$ semplice, ovvero

della forma $u = \sum_{i=1}^{n} u_i 1_{U_i}$ (dove per ogni $i = 1, ..., n,\ u_i$ appartiene a $\mathbb{R}$, U_i è un sottoinsieme di Δ e 1_{U_i} è la relativa funzione caratteristica) poniamo

$$\int_\Delta u(\delta) d\mu(\delta) = \sum_{i=1}^{n} u_i \mu(U_i).$$

Grazie alla densità dei net semplici in $B(\Delta, \mathbb{R})$, estendiamo la definizione di integrale a un qualunque net limitato u, passando al limite: più precisamente, se $\{u_n\}_{n\in\mathbb{N}}$ è una successione di net semplici che converge a u allora definiamo

$$\int_\Delta u(\delta) d\mu(\delta) = \lim_{n\to+\infty} \int_\Delta u_n(\delta) d\mu(\delta)$$

(si può verificare che tale limite esiste e non dipende dalla successione approssimante scelta). Con questa premessa, è ben definito il funzionale $LIM : B(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$

$$LIM(u) = \int_\Delta u(\delta) d\mu(\delta).$$

Per costruzione LIM risulta lineare e positivo. Dato che la carica μ si annulla sugli elementi dell'ideale I, osserviamo che per ogni $A \in I$ si ha $LIM(1_A) = \mu(A) = 0$. Fissiamo un elemento A di I e un net $g \in B(\Delta, \mathbb{R})$. Se g si annulla sul complementare di A, allora $-\|g\|_\infty 1_A \leqslant g \leqslant \|g\|_\infty 1_A$, dunque

$$0 = LIM(-\|g\|_\infty 1_A) \leqslant LIM(g) \leqslant LIM(\|g\|_\infty 1_A) = 0,$$

da cui $LIM(g) = 0$ per ogni g limitata che si annulla sul complementare di un sottoinsieme A di I. In particolare dato $u \in B(\Delta, \mathbb{R})$, il net $u - u1_{A^C}$ si annulla proprio sul complementare di A, perciò $LIM(u - u1_{A^C}) = 0$ e per linearità

$$(*) \qquad LIM(u) = LIM(u1_{A^C}) \qquad \forall\ u \in B(\Delta, \mathbb{R}), \quad \forall\ A \in I.$$

Mostriamo che LIM è proprio un limite di Banach sull'insieme diretto Δ: grazie al lemma 1.30 basta provare che $LIM(u) \leqslant \limsup_{\delta\in\Delta} u(\delta)$ per ogni $u \in B(\Delta, \mathbb{R})$. Fissato un qualsiasi numero reale $r > \limsup_{\delta\in\Delta} u(\delta)$, è sufficiente provare che $r \geqslant LIM(u)$. Per come è stato scelto r, esiste un δ_0 tale che $r > u(\delta)$ per ogni $\delta \succeq \delta_0$. In altre parole, posto $A^C = \{\delta \in \Delta :\ \delta \succeq \delta_0\}$, ovviamente $A \in I$ e $r1_{A^C} \geqslant u1_{A^C}$. Allora, tramite $(*)$ e per le proprietà di LIM, possiamo concludere che

$$LIM(u) = LIM(u1_{A^C}) \leqslant LIM(r1_{A^C}) = r\mu(A^C) = r(\mu(\Delta) - \mu(A)) = r.$$

□

Lemma 1.32. *(Estensione finita)*
Siano X un $\mathbb{R}-$spazio vettoriale e Y un suo sottospazio. Siano $\lambda_0 : Y \longrightarrow \mathbb{R}$ lineare, $p : X \longrightarrow \mathbb{R}$ subadditiva e positivamente omogenea tali che $\lambda_0 \leqslant p$ su Y.
Supponiamo infine che esista un sottoinsieme finito $S \subseteq X$ tale che $X =$Span$(Y \cup S)$.
Allora λ_0 può essere estesa ad una mappa lineare $\lambda : X \longrightarrow \mathbb{R}$ che soddisfa $\lambda \leqslant p$ su tutto X.

Dimostrazione. Si prova per induzione sulla cardinalità dell'insieme finito S; è chiaro che possiamo assumere che S contenga un solo elemento s.
Mostriamo che, grazie alle ipotesi su p e λ_0, vale

$$\sup_{w\in Y,\ t<0} \frac{p(w+ts)-\lambda_0(w)}{t} \leqslant \inf_{v\in Y,\ r>0} \frac{p(v+rs)-\lambda_0(v)}{r}.$$

Infatti ciò è equivalente a provare che per ogni $w, v \in Y,\ r > 0,\ t < 0$ vale

$$r\lambda_0(w) - rp(w+ts) \leqslant -tp(v+rs) + t\lambda_0(v)$$

o ancora

$$\lambda_0(rw - tv) \leqslant rp(w+ts) - tp(v+rs).$$

Quest'ultima disuguaglianza è verificata dato che

$$\lambda_0(rw-tv) \leqslant p(rw-tv) = p(rw+rts-rts-tv) \leqslant$$

$$\leqslant p(rw+rts) + p(-tv-trs) = rp(w+ts) - tp(v+rs)$$

usando la linearità di λ_0, le proprietà di p, e il fatto che $\lambda_0 \leqslant p$.
Allora esiste un numero reale $\lambda(s)$ tale che

$$\sup_{w\in Y,\ t<0} \frac{p(w+ts)-\lambda_0(w)}{t} \leqslant \lambda(s) \leqslant \inf_{v\in Y,\ r>0} \frac{p(v+rs)-\lambda_0(v)}{r}.$$

Definiamo l'estensione λ di λ_0 così:

$$\lambda(y+rs) = \lambda_0(y) + r\lambda(s) \qquad \forall\ y \in Y, \quad \forall\ r \in \mathbb{R}.$$

È ovvio che $\lambda : X \longrightarrow \mathbb{R}$ è lineare. Infine per costruzione (cioè per la proprietà di $\lambda(s)$) risulta $\lambda \leqslant p$ su X. $\square$

Proposizione 1.33. *Esistenza dei limiti di Banach $\Longrightarrow$ HB classico.*

Dimostrazione. Siano fissati $p : X \longrightarrow \mathbb{R}$ subadditiva e positivamente omogenea e $\lambda_0 : Y \longrightarrow \mathbb{R}$ lineare definita sul sottospazio $Y \subsetneq X$ sul quale si ha $\lambda_0 \leqslant p$.
Il nostro obiettivo è estendere λ_0 ad una mappa lineare λ su X in modo che valga $\lambda \leqslant p$ su tutto X. Per ogni sottoinsieme finito $S \subseteq X$, indichiamo con χ_S l'insieme delle estensioni $g : X \longrightarrow \mathbb{R}$ di λ_0 tali che $-p(-x) \leqslant g(x) \leqslant p(x)$ per ogni $x \in X$, e la cui restrizione $g\,|_{Span(Y \cup S)}$ sia lineare. Grazie al lemma 1.32 sappiamo che χ_S è non vuoto per ogni $S \subseteq X$ finito. Inoltre la famiglia degli insiemi χ_S è chiusa per intersezione finita, dato che $\chi_S \cap \chi_T = \chi_{S \cup T}$. Definiamo ora

$$\Delta = \{(g, S) : \; S \text{ è un sottoinsieme finito di } X \;\text{ e }\; g \in \chi_S\}$$

e introduciamo su di esso l'ordine parziale $\preceq$: $(g_1, S_1) \preceq (g_2, S_2)$ se e solo se $S_1 \subseteq S_2$. Con tale ordine parziale Δ risulta un insieme diretto (segue da $\chi_S \cap \chi_T = \chi_{S \cup T}$). Possiamo allora sfruttare l'ipotesi di esistenza dei limiti di Banach e considerare quindi $LIM : B(\Delta, \mathbb{R}) \longrightarrow \mathbb{R}$. Per ogni $x \in X$ definiamo una funzione $\psi_x : \Delta \longrightarrow \mathbb{R}$ ponendo $\psi_x(g, S) = g(x)$. Notiamo che ogni ψ_x è limitata, poiché $-p(-x) \leqslant \psi_x(\delta) \leqslant p(x)$ per ogni $\delta \in \Delta$. Sia $\lambda : X \longrightarrow \mathbb{R}$ definita da

$$\lambda(x) = LIM(\psi_x).$$

La funzione λ appena introdotta è proprio l'estensione cercata. Innanzitutto vale $\lambda \leqslant p$ su tutto X, infatti (dato che $p(x)$ non dipende da δ)

$$\lambda(x) = LIM(\psi_x) \leqslant LIM(p(x)) = p(x).$$

In effetti λ è un'estensione di λ_0: se $x \in Y$ allora (per definizione di ψ_x) vale $\psi_x(\delta) = \lambda_0(x)$ per ogni $\delta \in \Delta$, quindi per ogni $x \in Y$ si ha

$$\lambda(x) = LIM(\psi_x) = LIM(\lambda_0(x)) = \lambda_0(x).$$

Resta da provare la linearità di λ. Fissiamo $x, y \in X,\ a, b \in \mathbb{R}$. Per tutti gli elementi $\delta = (g, S)$ sufficientemente grandi in $(\Delta, \preceq)$ possiamo supporre $x, y \in S$ e in questo modo g è lineare su $Span\,(x, y)$. In particolare vale $g(ax + by) = ag(x) + bg(y)$, dunque $\psi_{ax+by}(\delta) = a\psi_x(\delta) + b\psi_y(\delta)$ per ogni δ sufficientemente grande in $(\Delta, \preceq)$. Concludiamo sfruttando la linearità dell'operatore LIM:

$$\lambda(ax + by) = LIM(\psi_{ax+by}) = LIM(a\psi_x + b\psi_y) =$$
$$= aLIM(\psi_x) + bLIM(\psi_y) = a\lambda(x) + b\lambda(y).$$

□

Corollario 1.34. *Dato che il teorema 1.24 sulla misura di Luxemburg è conseguenza di HB classico, tutti i teoremi presentati in questo paragrafo sono formulazioni equivalenti del teorema di Hahn-Banach.*

1.4 Hahn-Banach non è equivalente all'assioma della scelta

Completiamo questa sezione mostrando alcuni risultati che suggeriscono che il teorema di Hahn-Banach, che è una conseguenza dell'assioma della scelta, non equivale ad esso. Non arriveremo però alla dimostrazione rigorosa di questo fatto, che è alquanto tecnica e richiede argomenti che esulano dai nostri scopi. Per una trattazione completa rimandiamo a [5].

Ricordando la definizione 1.12 di ideale di un'algebra booleana, vediamo una interessante caratterizzazione.

Definizione 1.35. Siano I_0 un ideale di un'algebra booleana X e x un elemento di X. Il più piccolo ideale I contenente I_0 e x, detto ideale generato da I_0 e x, è l'insieme degli elementi $y \in X$ tali che esiste $i \in I_0$ per cui $y \preceq (i \vee x)$.

Dalla definizione si verifica facilmente che si tratta effettivamente di un ideale che ovviamente contiene I_0 e x (dato che $x = 0 \vee x$ e per ogni $i \in I_0$ vale $i \preceq (i \vee x)$). Chiaramente se $x \in I_0$ l'ideale generato è ancora I_0.

Osservazione 1.36. Detto I l'ideale generato da I_0 e x, si ha

$$I = X \quad \text{se e solo se} \quad x^C \in I_0.$$

Infatti se $x^C \in I_0 \subseteq I$ allora $x \vee x^C = 1 \in I$, perciò $I = X$.
Viceversa se $I = X$, allora $1 \in I$ e per definizione di ideale generato esiste $i \in I_0$ tale che $1 \preceq (i \vee x)$, ovvero $1 = (i \vee x)$ perché 1 è l'elemento massimo in X. Mostrando che $x^C = (x^C \wedge i)$ avremmo $x^C \in I_0$ ossia la tesi visto che, essendo un ideale, I_0 è chiuso esternamente rispetto a $\wedge$. Questo è vero dato che

$$x^C = x^C \wedge 1 = x^C \wedge (i \vee x) = (x^C \wedge i) \vee (x^C \wedge x) = (x^C \wedge i) \vee 0 = x^C \wedge i.$$

Definizione 1.37. Un ideale proprio $\mathcal{M}$ di un'algebra booleana X si dice massimale se non è strettamente contenuto in un altro ideale proprio.

Grazie all'assioma della scelta, per ogni algebra booleana diversa da $\{0\}$ esiste almeno un ideale massimale, come proveremo nella proposizione 1.40. D'altra parte osserviamo che a priori, cioè senza supporre AC, non è detto che un'algebra booleana ammetta un ideale massimale.

Lemma 1.38. *Siano X un'algebra booleana e $\mathcal{M}$ un suo ideale proprio, allora le seguenti condizioni sono equivalenti:*
(i) $\mathcal{M}$ è massimale,
(ii) per ogni $x \in X$, $\mathcal{M}$ contiene esattamente un elemento tra x e x^C.

Dimostrazione. Sia $\mathcal{M}$ un ideale proprio che contiene esattamente un elemento tra x e x^C. Supponiamo esista un ideale I che contiene strettamente $\mathcal{M}$: allora necessariamente I contiene due elementi tra loro complementari; ma un ideale è chiuso rispetto a $\vee$, e dunque $x \vee x^C = 1 \in I$, cioè $I = X$. Questo significa che l'unico ideale che contiene strettamente $\mathcal{M}$ è tutto X, ovvero che $\mathcal{M}$ è massimale.
Viceversa supponiamo $\mathcal{M}$ massimale. Sicuramente $\mathcal{M}$ non può contenere due elementi tra loro complementari, altrimenti per chiusura rispetto a $\vee$, 1 apparterrebbe a $\mathcal{M}$, cioè $\mathcal{M} = X$, e questo non è possibile perché $\mathcal{M}$ è proprio. Sia x un elemento che non appartiene a $\mathcal{M}$, allora l'ideale I generato da $\mathcal{M}$ e da x contiene strettamente $\mathcal{M}$; quindi, per ipotesi di massimalità, $I = X$ e questo significa $x^C \in \mathcal{M}$ per l'osservazione 1.36. □

Proposizione 1.39. *Data un'algebra booleana X diversa da $\{0\}$, esiste un ideale massimale di X se e solo se X ammette una carica di probabilità μ che assume solo valore 0 e 1.*

Dimostrazione. Sia $\mu : X \longrightarrow \{0,1\}$: allora in questo caso $\ker(\mu)$, che sappiamo essere un ideale proprio per l'osservazione 1.16, è addirittura massimale.
Infatti dato che μ assume solo valore 0 e 1, allora esattamente uno tra x e x^C appartiene a $\ker(\mu)$ (dato che $\mu(x) + \mu(x^C) = 1$). Perciò per il lemma precedente $\ker(\mu)$ è un ideale massimale.
Per il viceversa, sia $\mathcal{M}$ l'ideale massimale. Per il lemma 1.38 vale

$$X = \mathcal{M} \sqcup \left\{x^C : \ x \in M\right\}.$$

Possiamo allora definire una carica di probabilità $\mu : X \longrightarrow [0,1]$, data da

$$\mu(x) = \begin{cases} 0 & \text{se } x \in \mathcal{M}, \\ 1 & \text{se } x^C \in \mathcal{M}. \end{cases}$$

Ovviamente $\mu(1) = 1$ perché $\mathcal{M}$ è un ideale proprio. Verifichiamo che, essendo $\mathcal{M}$ un ideale proprio, μ è finitamente additiva sui disgiunti, considerando caso per caso.
Se $x, y \in \mathcal{M}$ anche $(x \vee y) \in \mathcal{M}$, quindi $\mu(x \vee y) = 0 = \mu(x) + \mu(y)$.
Se $x \in \mathcal{M}$, $y \notin \mathcal{M}$ si ha $(x \vee y) \notin \mathcal{M}$ (altrimenti $(x \vee y) \wedge y = y \in \mathcal{M}$ che è assurdo), quindi $\mu(x \vee y) = 1 = \mu(x) + \mu(y)$. Nell'ultimo caso, due elementi $x, y \notin \mathcal{M}$ non possono essere disgiunti altrimenti, passando al complementare, risulterebbe $1 = 0^C = (x \wedge y)^C = x^C \vee y^C$, cioè $1 \in \mathcal{M}$ che è assurdo.
Dunque μ è una carica di probabilità su X che assume solo valore 0 e 1.

□

Proposizione 1.40. *(AC) Ogni algebra booleana diversa da* $\{0\}$ *ammette una carica di probabilità che assume solo valore* 0 *e* 1*.*

Dimostrazione. Fissata un'algebra booleana X diversa da $\{0\}$, sia Σ l'insieme dei suoi ideali propri. Dato che $\{0\}$ è un ideale proprio, Σ è un insieme non vuoto. Chiaramente Σ è parzialmente ordinato tramite $\subseteq$.
Vogliamo applicare il lemma di Zorn a Σ concludendo così che esiste un ideale $\mathcal{M}$ massimale di X. Per applicare Zorn dobbiamo prima verificare che ogni catena in Σ ammette elemento maggiorante in Σ. Sia quindi $I_1 \subseteq I_2 ... \subseteq I_n \subseteq ...$ una catena di ideali propri di X: mostriamo che $I = \bigcup_{n\in\mathbb{N}} I_n$ è ancora un ideale proprio di X, quindi è l'elemento maggiorante della catena che stavamo cercando. Per ogni $n \in \mathbb{N}$, I_n è un ideale proprio, quindi $0 \in I_n$ ma $1 \notin I_n$ per ogni n. Ciò significa che $0 \in I$ e che $1 \notin I$, quindi I è un sottoinsieme proprio. Proviamo che è un ideale: dati $x, y \in I$, per definizione di I (ed essendo $\{I_n\}_{n\in\mathbb{N}}$ una catena), esiste un indice abbastanza grande n_0 tale che $x, y \in I_{n_0}$ da cui $(x \vee y) \in I_{n_0} \subseteq I$. Allo stesso modo proviamo che I è chiuso esternamente rispetto a $\wedge$. Allora $I = \bigcup_{n\in N} I_n$ è un ideale proprio di X: il lemma di Zorn ci garantisce l'esistenza in X di un ideale massimale $\mathcal{M}$. Concludiamo grazie alla proposizione 1.39. $\square$

Osservazione 1.41. Chiaramente, per dimostrare la proposizione 1.40 è sufficiente supporre che esista un ideale massimale dell'algebra booleana (in tal caso la proposizione vale senza supporre AC, che nella dimostrazione sopra serve per l'appunto ad assicurare l'esistenza di un ideale massimale).

Ricapitoliamo i diversi risultati ottenuti sulle cariche di probabilità delle algebre booleane. Da una parte il teorema 1.25 di Hahn-Banach per algebre booleane, equivalente al classico HB, assicura che
″Ogni algebra booleana diversa da $\{0\}$ ammette una carica di probabilità″.
D'altra parte la proposizione 1.40, conseguenza di AC, afferma che
″Ogni algebra booleana diversa da $\{0\}$ ammette una carica di probabilità che assume solo valore 0 e 1″.

Queste formulazioni non solo ci confermano che il teorema di Hahn-Banach è conseguenza dell'assioma della scelta (come già sapevamo), ma potrebbero indurci a ritenere che HB non è equivalente a AC.
In effetti è così, cioè non sono equivalenti, ma il problema è più intricato di quello che potrebbe sembrare, e provare la non equivalenza di HB e AC è tutt'altro che semplice e richiede metodi particolari. In dettaglio, come dimostrato rigorosamente in [5] (capitolo 7), l'esistenza di un ideale massimale per ogni algebra boolena diversa da $\{0\}$, che è conseguenza di AC come visto nella proposizione 1.40, non è equivalente all'assioma della scelta. D'altra parte l'ipotesi di esistenza di un ideale massimale per ogni algebra boolena diversa da $\{0\}$ basta (vedi osservazione 1.41) per provare la proposizione 1.40, che a sua volta implica il teorema 1.25, forma equivalente del classico HB. Dunque HB non può essere equivalente all'assioma della scelta, altrimenti l'ipotesi di esistenza di un ideale massimale sarebbe equivalente ad AC, e questo è assurdo.
In conclusione il teorema di Hahn-Banach rappresenta logicamente un'ipotesi più debole rispetto all'assioma della scelta.

2 Il Paradosso di Banach-Tarski

2.1 Introduzione e formalizzazione del paradosso

Nel corso dei secoli la nozione di infinito ha spesso condotto a costruzioni apparentemente paradossali o contro-intuitive, in molte delle quali sembra sia possibile cambiare la grandezza di oggetti tramite mosse o operazioni che conservano la grandezza. In un famoso esempio, Galileo mostrò che l'insieme dei numeri naturali può essere messo in corrispondenza uno ad uno con l'insieme dei quadrati interi, benché l'insieme dei non quadrati, e a maggior ragione l'insieme di tutti i naturali, sembri molto più numeroso di quello dei quadrati. Egli dedusse pertanto che i classici attributi di uguale, maggiore e minore non erano applicabili a insiemi infiniti, e fu per questo precursore della teoria della cardinalità di Cantor. Nella sua celebre osservazione, Galileo provò come, partendo dall'insieme dei naturali, fosse possibile suddividerlo in due suoi sottoinsiemi, ciascuno avente la stessa dimensione dell'insieme dei numeri naturali.

L'idea della duplicazione presentata in questo esempio è alla base anche dell'importante, e molto più recente (1924), paradosso di Banach-Tarski. La scoperta di paradossali decomposizioni risale all'inizio dello scorso secolo, ed è andata di pari passo con la formalizzazione della teoria della misura.
Il paradosso di Banach-Tarski è il famoso paradosso di duplicazione della sfera, il quale stabilisce che è possibile prendere una sfera, suddividerla in un numero finito di pezzi disgiunti e, solo tramite isometrie, riassemblare i pezzi fino ad ottenere due sfere della stessa grandezza di quella originaria.
Tale duplicazione vale sia per la sfera o più precisamente per la superficie sferica $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \ |x| = 1\}$, che per la sfera "piena" $\mathbb{B} = \{x \in \mathbb{R}^3 : \ |x| \leqslant 1\}$.

Tale duplicazione è così controintuitiva e lontana dalla realtà fisica che ha assunto, fin dalla sua scoperta, l'appellativo di paradosso. Sembrerebbe infatti impossibile poter duplicare una sfera (raddoppiando così il volume totale) solo tramite isometrie, cioè traformazioni biunivoche che preservano le distanze.

Nonostante la sfera originaria e le due sfere di arrivo abbiano un volume ben definito, la spiegazione sta nel fatto che durante la duplicazione non vi è nessun volume da preservare, perché alcuni dei sottoinsiemi in cui la prima sfera risulta suddivisa sono non misurabili, ovvero sono insiemi a cui non può essere applicata la nozione di volume. Non si può quindi parlare di volume dei pezzi tagliati sulla sfera, ma solo di una nozione più generale, definita per tutti i sottoinsiemi di $\mathbb{R}^3$, che è la misura esterna (tridimensionale) di Lebesgue, la quale da una parte estende il concetto di volume per gli insiemi misurabili ma dall'altra, considerata su tutte le parti di $\mathbb{R}^3$, non è neppure finitamente additiva sui disgiunti.

Ciò è possibile, e così è stato dimostrato la prima volta, utilizzando l'assioma della scelta. È nostro obiettivo provare che tale contro-intuitiva duplicazione sia soltanto conseguenza del teorema di Hahn-Banach, di cui l'assioma della scelta rappresenta una condizione sufficiente ma non necessaria.

In questo ambito un ruolo centrale sarà interpretato dalle azioni di gruppi liberi. Introduciamo concetti e definizioni necessarie alla formalizzazione del problema.
Ricordiamo che un gruppo G agisce su un insieme X se ad ogni $g \in G$ corrisponde una funzione biunivoca da X in X, ancora denotata con la lettera g, tale che per ogni $g, h \in G$ e per ogni $x \in X$ valga $g(h(x)) = (gh)(x)$, e $1(x) = x$, dove 1 è l'identità del gruppo G. Fissati $X_0 \subseteq X$ e $g \in G$ scriviamo gX_0 per indicare l'insieme degli elementi $g(x_0)$ al variare di $x_0 \in X_0$; analogamente, fissati $x \in X$ e $G_0 \subseteq G$ scriviamo G_0x per indicare l'insieme dei $g_0(x)$ al variare di $g_0 \in G_0$.
Diciamo che un'azione di gruppo G su X è libera se per ogni $x \in X,\ g \in G$ si ha $g(x) = x$ se e solo se $g = 1$, cioè se l'azione non ha punti fissi (non banali).
Infine possiamo definire tramite l'azione di gruppo una relazione di equivalenza su X: $x, x' \in X$ sono equivalenti se esiste $g \in G$ tale che $x' = g(x)$; queste classi di equivalenza, che formano una partizione dell'insieme X, si dicono $G-$orbite o semplicemente orbite (dell'azione).

Definizione 2.1. Siano G un gruppo che agisce su un insieme X, E un sottoinsieme di X. E si dice *G-paradossale* se per alcuni $m, n \in \mathbb{N}_0$ esistono sottoinsiemi di E a due a due disgiunti $A_1, ..., A_n, B_1, ..., B_m$ e elementi $g_1, ..., g_n, h_1, ...h_m \in G$ tali che

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

In altre parole E è $G-$paradossale se può essere duplicato mediante l'azione di G.

Detto G_n il gruppo delle isometrie euclidee di $\mathbb{R}^n$, denotiamo con SO_n il sottogruppo di G_n delle traformazioni ortogonali aventi determinante 1, ovvero il gruppo delle rotazioni della superficie sferica $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$.

Con le definizioni introdotte, il paradosso di Banach-Tarski assume la forma:

Teorema 2.2. *(Paradosso di Banach-Tarski)*
La superficie sferica $\mathbb{S}^2$ è SO_3-paradossale. La sfera $\mathbb{B}$ è G_3-paradossale.

La dimostrazione del paradosso, nella forma di questo teorema, e come implicazione di HB, è l'obiettivo di questo capitolo. Abbiamo bisogno di alcuni risultati preliminari; in particolare dobbiamo sapere quali insiemi sono paradossali e di quali proprietà godono.

2.2 L'esempio di partenza: i gruppi liberi sono paradossali

Ogni gruppo agisce naturalmente su se stesso tramite traslazione sinistra.
Dunque il primo quesito da porci è quali gruppi siano paradossali (secondo la definizone data) rispetto a questa prima e naturalissima azione: il primo esempio è il gruppo libero di rango 2.
Ricordiamo che il gruppo libero F generato dall'insieme M è il gruppo di tutte le parole con un numero finito di lettere, ottenute usando solo lettere appartenenti a $\{\sigma, \sigma^{-1} : \ \sigma \in M\}$, dove due parole sono equivalenti se una di esse può essere trasformata nell'altra tramite aggiunta o rimozione di coppie di lettere adiacenti della forma $\sigma\sigma^{-1}$ o $\sigma^{-1}\sigma$. Una parola priva di queste coppie di lettere è detta ridotta. Per semplicità possiamo considerare F come il gruppo formato da tutte e sole le parole ridotte. L'operazione di gruppo è la concatenazione (la concatenazione di due parole è sempre equivalente a una parola ridotta), e l'identità in F è la parola vuota, indicata col simbolo 1. La cardinalità di M (l'insieme che genera F) è detta il rango del gruppo libero F.

Osservazione 2.3. 1) Ogni gruppo finito non banale non è libero, perché per costruzione F è un gruppo con infiniti elementi.
2) Se F è un gruppo libero di rango $\geqslant 2$, allora F non è abeliano: infatti se esistono almeno due elementi $\sigma \neq \tau$ in M, allora $\sigma\tau$ e $\tau\sigma$ sono parole ridotte e distinte.
3) Un gruppo libero F di rango $\geqslant 2$ non è risolubile (vedi [10], appendice A) cioè non ammette una serie normale di sottogruppi $H_0 \lhd H_1 \lhd ... \lhd H_n = F$ in cui tutti i gruppi quoziente H_i/H_{i-1} $(i = 1, ..., n)$ sono abeliani.

Proposizione 2.4. *Un gruppo libero F di rango 2 è F-paradossale, dove F agisce su se stesso per moltiplicazione sinistra.*

Dimostrazione. Siano σ, τ i due generatori liberi di F che risulta quindi il gruppo delle parole (finite) ridotte costituite dalle lettere $\sigma, \sigma^{-1}, \tau, \tau^{-1}$. Detta ρ una di queste quattro lettere, sia $W(\rho)$ il sottoinsieme di F delle parole che iniziano (a sinistra) con ρ. In questo modo $F = \{1\} \sqcup W(\sigma) \sqcup W(\sigma^{-1}) \sqcup W(\tau) \sqcup W(\tau^{-1})$.
Consideriamo una generica parola h di F: se $h \notin W(\sigma)$, allora $h = \sigma(\sigma^{-1}h) \in \sigma W(\sigma^{-1})$; allo stesso modo se $h \notin W(\tau)$ allora $h \in \tau W(\tau^{-1})$. Pertanto

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

che è proprio la tesi. In dettaglio, vale la definizione 2.1 con $W(\sigma) = A_1, 1 = g_1, \ W(\sigma^{-1}) = A_2, \sigma = g_2, \ W(\tau) = B_1, 1 = h_1, \ W(\tau^{-1}) = B_2, \tau = h_2$. □

Proposizione 2.5. *(HB) Se F è un gruppo libero di rango 2 che agisce liberamente su un insieme X, allora X è F-paradossale.*

Dimostrazione. Sia $B = \sum \mathcal{P}(\mathcal{O}_i)$ la somma booleana (data dalla proposizione 1.17) delle algebre booleane $\mathcal{P}(\mathcal{O}_i)$ al variare di tutte le F-orbite $\mathcal{O}_i$. Per il teorema 1.25 (HB per algebre booleane) B ammette una carica di probabilità μ, in particolare vale $\mu(1) = 1$. Fissiamo adesso quattro sottoinsiemi $\{A_i\}_{i=1,2,3,4} \subseteq F$ a due a due disgiunti, quattro elementi $a_1, a_2, a_3, a_4 \in F$, ed una permutazione σ di $\{1, 2, 3, 4\}$ in modo che

$$(*) \qquad (A_i a_i^{-1}) \cup A_{\sigma(i)} = F \qquad \text{per } i = 1, 2, 3, 4.$$

Ad esempio, detti a_1, a_2 i generatori liberi di F, possiamo porre

$$a_3 = a_1^{-1},\ a_4 = a_2^{-1}, \quad \sigma(1) = 3,\ \sigma(3) = 1,\ \sigma(2) = 4,\ \sigma(4) = 2,$$

e definire A_i come l'insieme delle parole ridotte che terminano per a_i: in questo modo vale la proprietà $(*)$.
Fissato $x \in X$, notiamo che ogni elemento di $A_i x$ (per $i = 1, 2, 3, 4$) è ovviamente equivalente a x, perciò $A_i x \in \mathcal{P}(\mathcal{O}(x))$, dove $\mathcal{O}(x)$ è l'orbita che ha x come rappresentante. Allora, detto $j : \mathcal{P}(\mathcal{O}(x)) \longrightarrow B$ l'omomorfismo canonico di algebre booleane (j esiste per definizione di somma booleana), ha senso parlare di $\mu\left(j(A_i x)\right)$. Definito, per $i = 1, 2, 3, 4$, $X_i = \{x \in X : \mu(j(A_i x)) > 1/2\}$, siano

$$Y_1 = X \setminus (a_1 X_1 \cup a_2 X_2)\,, \quad Y_2 = X \setminus (a_3 X_3 \cup a_4 X_4)\,.$$

Ovviamente

$$a_1 X_1 \cup a_2 X_2 \cup Y_1 = X = a_3 X_3 \cup a_4 X_4 \cup Y_2.$$

Mostriamo che Y_1, Y_2 e gli insiemi X_i sono a due a due disgiunti: così facendo grazie alla doppia decomposizione appena scritta concludiamo che X è F-paradossale.
Innanzitutto gli X_i sono disgiunti. Infatti supponiamo ad esempio esista $x \in X_1 \cap X_2$; allora $\mu(j(A_1 x)) > 1/2$ e $\mu(j(A_2 x)) > 1/2$. Ma $A_1 x \cap A_2 x = \emptyset$ perché se esistessero $\alpha_1 \in A_1, \alpha_2 \in A_2$ tali che $\alpha_1(x) = \alpha_2(x)$ allora $\alpha_2^{-1}(\alpha_1(x)) = x$; quindi, essendo l'azione libera, avremmo $\alpha_1 = \alpha_2$, che è assurdo perché A_1 e A_2 sono disgiunti. Segue dunque che

$$\mu\left(j(A_1 x \sqcup A_2 x)\right) = \mu(j(A_1 x)) + \mu(j(A_2 x)) > 1,$$

assurdo poiché μ al massimo vale 1. Dunque tale x non esiste, e $X_1 \cap X_2 = \emptyset$.
$Y_1 \cap Y_2 = \emptyset$ perché vale $X = \bigcup_{i=1}^{4} a_i X_i$. Infatti fissato $x \in X$, essendo l'azione libera, gli insiemi $A_{\sigma(1)}x, A_{\sigma(2)}x, A_{\sigma(3)}x, A_{\sigma(4)}x$ sono a due a due disgiunti, perciò esiste i

per cui $\mu(j(A_{\sigma(i)}x)) < 1/2$ (altrimenti, per additività di μ sui disgiunti, risulterebbe $\mu\left(j(A_{\sigma(1)}x \sqcup A_{\sigma(2)}x \sqcup A_{\sigma(3)}x \sqcup A_{\sigma(4)}x)\right) \geqslant 2$, che è assurdo). Per $(*)$ si ha

$$\mu(j(A_i a_i^{-1}x \cup A_{\sigma(i)}x)) = \mu(j(Fx)) = \mu(j(\mathcal{O}(x)) = \mu(j(1)) = \mu(1) = 1$$

da cui, per subadditività della carica μ (vedi lemma 1.15),

$$\mu(j(A_i a_i^{-1}x)) \geqslant 1 - \mu(j(A_{\sigma(i)}x)) > 1/2.$$

Quindi $a_i^{-1}x \in X_i$, cioè $x \in a_i X_i$.
Per terminare dobbiamo provare che Y_1 e Y_2 sono disgiunti dagli X_i. A tal proposito verifichiamo che $X_i \subseteq a_j X_j$ per $i \neq \sigma(j)$: se $x \in X_i$ e $i \neq \sigma(j)$, sfruttando il fatto che gli A_i sono disgiunti, per $(*)$ vale $A_i \subseteq A_j a_j^{-1}$, da cui si deduce che $\mu(j(A_j a_j^{-1}x)) \geqslant \mu(j(A_i x)) > 1/2$ (essendo $x \in X_i$); in altre parole si ha $a_j^{-1}x \in X_i$ e quindi $x \in a_j X_j$.
Per quanto detto (ricordando la definizione di σ) si ha

$$(X_1 \cup X_2 \cup X_4) \subseteq a_1 X_1, \qquad (X_1 \cup X_2 \cup X_3) \subseteq a_2 X_2,$$

$$(X_2 \cup X_3 \cup X_4) \subseteq a_3 X_3, \qquad (X_1 \cup X_3 \cup X_4) \subseteq a_4 X_4.$$

Dunque

$$\bigcup_{i=1}^{4} X_i \subseteq [(a_1 X_1 \cup a_2 X_2) \cap (a_3 X_3 \cup a_4 X_4)] = Y_1^C \cap Y_2^C = (Y_1 \cup Y_2)^C.$$

In conclusione $X_1,\ X_2,\ X_3,\ X_4,\ Y_1,\ Y_2$ sono sottoinsiemi di X a due a due disgiunti tali che

$$a_1 X_1 \cup a_2 X_2 \cup Y_1 = X = a_3 X_3 \cup a_4 X_4 \cup Y_2.$$

Questo significa che X è F-paradossale.

□

Dato che un sottogruppo di un gruppo agisce liberamente per traslazione su tutto il gruppo, un'immediata conseguenza della proposizione 2.5 è:

Corollario 2.6. *Ogni gruppo G che ha un sottogruppo libero F di rango 2 è F-paradossale rispetto all'azione libera di traslazione sinistra. In particolare ciò vale per ogni gruppo libero non abeliano.*

2.3 Dal paradosso di Hausdorff a quello di Banach-Tarski

La prima opportunità di trovare un sottogruppo libero non abeliano nei gruppi G_n delle isometrie in $\mathbb{R}^n$ si ha per $n = 3$. Questo perché G_1 e G_2 sono gruppi risolubili (per dettagli rimandiamo a [10], appendice A), e pertanto non ammettono nessun sottogruppo libero non abeliano (dato che ogni sottogruppo di un gruppo risolubile è risolubile, mentre i gruppi liberi non abeliani non sono risolubili, come osservato in 2.3).

Invece è possibile esibire esplicitamente un sottogruppo libero non abeliano per G_3, anzi per il suo sottogruppo SO_3. Per la precisione si tratta del più semplice sottogruppo libero non abeliano: quello di rango 2.

Teorema 2.7. *Ci sono due rotazioni indipendenti in $\mathbb{R}^3$, ρ e φ, con assi passanti per l'origine. Quindi SO_3 possiede un sottogruppo libero di rango 2.*

Dimostrazione. Siano φ e ρ le due rotazioni antiorarie di $\mathbb{R}^3$ rispettivamente intorno all'asse z e all'asse x, entrambe di angolo $\arccos(1/3)$.
Le rotazioni $\varphi^{\pm 1}$ e $\rho^{\pm 1}$ sono rappresentate come applicazioni dalle seguenti equazioni:

$$\varphi^{\pm 1}(x,y,z) = \left(\frac{1}{3}x \mp \frac{2\sqrt{2}}{3}y, \pm\frac{2\sqrt{2}}{3}x + \frac{1}{3}y, z \right),$$

$$\rho^{\pm 1}(x,y,z) = \left(x, \frac{1}{3}y \mp \frac{2\sqrt{2}}{3}z, \pm\frac{2\sqrt{2}}{3}y + \frac{1}{3}z \right).$$

Vogliamo provare che nessuna parola ridotta e non banale in $\varphi^{\pm 1}$ e $\rho^{\pm 1}$ coincide con l'identità, cioè che le due rotazioni sono indipendenti. Ciò equivale alla tesi: il sottogruppo libero di rango 2 in SO_3 è quello che ha per generatori le rotazioni φ e ρ. Senza perdita di generalità possiamo restringere la nostra attenzione alle parole ridotte che terminano (a destra) con φ o φ^{-1}. Detta w una parola ridotta che termina con φ o φ^{-1}, supponiamo che w coincida con l'identità e cerchiamo un assurdo.
Proviamo per induzione sulla lunghezza della parola w che $w\,(1,0,0)$ è della forma $\left(a, b\sqrt{2}, c\right)/3^k$ dove $a, b, c \in \mathbb{Z}$ e b non è divisibile per 3.
Questo implica che $w\,(1,0,0) \neq (1,0,0)$ perciò w non può essere l'identità e siamo arrivati alla contraddizione voluta.
Il caso base è immediato: se w ha lunghezza 1, $w = \varphi^{\pm 1}$ quindi $w\,(1,0,0) = \left(1, \pm 2\sqrt{2}, 0\right)/3$, e $b = \pm 2$ non è divisibile per 3.

Se w ha lunghezza maggiore di 1, w è del tipo $\varphi^{\pm 1}w^{'}$ o $\rho^{\pm 1}w^{'}$ dove per ipotesi induttiva $w^{'}(1,0,0) = \left(a^{'}, b^{'}\sqrt{2}, c^{'}\right)/3^{k-1}$ con $a^{'}, b^{'}, c^{'} \in \mathbb{Z}$. Nel primo caso

$$\varphi^{\pm 1}w^{'}(1,0,0) = \varphi^{\pm 1}\left(\left(a^{'}, b^{'}\sqrt{2}, c^{'}\right)/3^{k-1}\right) = \left(a^{'} \mp 4b^{'}, \sqrt{2}\left(2a^{'} \pm b^{'}\right), 3c^{'}\right)/3^{k},$$

cioè è della forma $\left(a, b\sqrt{2}, c\right)/3^k$ con $a = a^{'} \mp 4b^{'}, b = 2a^{'} \pm b^{'}, c = 3c^{'}$. Analogamente per l'altro vale $\rho^{\pm 1}w^{'}(1,0,0) = \left(a, b\sqrt{2}, c\right)/3^k$ con $a = 3a^{'}, b = b^{'} \mp 2c^{'}, c = c^{'} \pm 4b^{'}$. In entrambi i casi per le relazioni scritte osserviamo che i coefficienti a, b, c sono sempre interi dato che per ipotesi induttiva lo sono $a^{'}, b^{'}, c^{'}$.
Resta da mostrare che partendo da un $b^{'}$ non divisibile per 3 (caso base) arriviamo ad un b che non è mai divisibile per 3 . Dobbiamo analizzare quattro casi per w: $\varphi^{\pm 1}\rho^{\pm 1}v, \rho^{\pm 1}\varphi^{\pm 1}v, \varphi^{\pm 1}\varphi^{\pm 1}v, \rho^{\pm 1}\rho^{\pm 1}v$, dove v è una parola che può essere anche vuota (a causa di eventuali cancellazioni).
Nei primi due casi, usando le notazioni e le equazioni precedenti, abbiamo $b = b^{'} \mp 2c^{'}$ dove 3 divide $c^{'}$ oppure $b = b^{'} \pm 2a^{'}$ dove 3 divide $a^{'}$. Quindi b non è divisibile per 3 dato che non lo è $b^{'}$. Per gli altri 2 casi, detti $a^{''}, b^{''}, c^{''}$ i soliti coefficienti nella scrittura di $v(1,0,0)$, vale $b = 2b^{'} - 9b^{''}$; perciò concludiamo come prima che b non è divisibile per 3. Ad esempio per il terzo caso, applicando 2 volte $\varphi^{\pm 1}$, si ha

$$b = b^{'} \pm 2a^{'} = b^{'} \pm 2\left(a^{''} \mp 4b^{''}\right) = b^{'} + \left(b^{''} \pm 2a^{''}\right) - 9b^{''} = 2b^{'} - 9b^{''}.$$

□

Osservazione 2.8. Detto F il sottogruppo libero di rango 2 di SO_3 (generato da φ e ρ) descritto nel precedente teorema, notiamo che ogni elemento di F (identità esclusa), essendo una rotazione di $\mathbb{R}^3$, ha per punti fissi i punti del proprio asse; perciò la proposizione 2.5 non può essere direttamente applicata perché l'azione del gruppo libero di rotazioni F sulla sfera non è libera.
Vediamo ora come si può aggirare questa difficoltà concentrandosi su un opportuno sottoinsieme di $\mathbb{S}^2$ per ottenere un'azione finalmente libera.

Teorema 2.9. *(Paradosso di Hausdorff) (HB)*
Esiste un sottoinsieme numerabile D di $\mathbb{S}^2$ tale che $\mathbb{S}^2 \setminus D$ è SO_3-paradossale.

Dimostrazione. Ogni rotazione (diversa dall'identità) del sottogruppo libero F ha esattamente due punti fissi su $\mathbb{S}^2$, precisamente i punti d'intersezione tra la superficie sferica e l'asse di rotazione. Sia $D \subset \mathbb{S}^2$ la famiglia di tutti questi punti fissi al variare delle rotazioni (diverse dall'identità) di F. Dato che F, un gruppo libero di rango 2, è numerabile (lo si può vedere come unione di una famiglia numerabile di insiemi finiti), anche D è numerabile.
Adesso, se $P \in \mathbb{S}^2 \setminus D$ e $g \in F$, allora $g(P) \in \mathbb{S}^2 \setminus D$. È ovvio che $g(P) \in \mathbb{S}^2$ (perché g è una rotazione con asse passante per l'origine perciò conserva la distanza 1 del punto P dall'origine), inoltre non può appartenere a D. Infatti se $g(P)$ appartenesse a D ciò vorrebbe dire che esiste $h \in F$ (diversa dall'identità) tale che $h(g(P)) = g(P)$; quindi $(g^{-1}hg)(P) = P$ e pertanto $P \in D$, assurdo perché abbiamo supposto $P \notin D$. Dunque il gruppo libero di rango 2 agisce su $\mathbb{S}^2 \setminus D$ senza punti fissi non banali, ossia l'azione di F su $\mathbb{S}^2 \setminus D$ è libera.
Quindi possiamo applicare la proposizione 2.5, grazie alla quale possiamo affermare che $\mathbb{S}^2 \setminus D$ è $F-$paradossale. A maggior ragione $\mathbb{S}^2 \setminus D$ è SO_3-paradossale (dato che F è un sottogruppo di SO_3). □

Definizione 2.10. Sia G un gruppo che agisce su un insieme X. Due sottoinsiemi $A, B \subseteq X$ si dicono $G-$*equidecomponibili* se possono essere partizionati nello stesso numero finito di pezzi G-congruenti. Formalmente,

$$A = \bigsqcup_{i=1}^{n} A_i, \qquad B = \bigsqcup_{i=1}^{n} B_i,$$

ed esistono $g_1, ..., g_n \in G$ tali che per ogni $i = 1, ..., n$ valga $g_i A_i = B_i$.

Osservazione 2.11. Si verifica facilmente che la relazione di $G-$equidecomponibilità, indicata con $\backsim_G$, è una relazione di equivalenza su $\mathcal{P}(X)$.
Riprendendo la definizione 2.1, se E è $G-$paradossale possiamo senz'altro supporre (eventualmente restringendo i sottoinsiemi $A_1, ..., A_n, B_1, ..., B_m$) che

$$\bigsqcup_{i=1}^{n} g_i A_i = E = \bigsqcup_{j=1}^{m} h_j B_j.$$

Dunque $E \subseteq X$ è $G-$paradossale se e solo se contiene due insiemi disgiunti A, B tali che $A \backsim_G E$ e $B \backsim_G E$.

Proposizione 2.12. *Supponiamo che G agisca su X e siano $E \subseteq E^{'}$ sottoinsiemi $G-$equidecomponibili di X. Se E è $G-$paradossale, anche $E^{'}$ è $G-$paradossale.*

Dimostrazione. Come osservato in 2.11, l'ipotesi che E è $G-$paradossale vuol dire che esistono $A, B \subseteq E \left(\subseteq E^{'}\right)$ disgiunti tali che $A \backsim_G E$ e $B \backsim_G E$. Inoltre $E \backsim_G E^{'}$, allora per transitività di $\backsim_G$ si ha $A \backsim_G E^{'}$ e $B \backsim_G E^{'}$, cioè la tesi. □

Teorema 2.13. *Se D è un sottoinsieme numerabile di $\mathbb{S}^2$, allora $\mathbb{S}^2$ e $\mathbb{S}^2 \setminus D$ sono SO_3-equidecomponibili.*

Dimostrazione. Basta mostrare che esiste una rotazione $\rho \in SO_3$ tale che gli insiemi $D, \rho(D), ..., \rho^n(D), ...$ siano a due a due disgiunti. Questo è sufficiente, perché in tal caso possiamo decomporre $\mathbb{S}^2$ e $\mathbb{S}^2 \setminus D$ in due parti SO_3-congruenti in questo modo:

$$\mathbb{S}^2 = \left[D_0 \sqcup \left(\mathbb{S}^2 \setminus D_0\right)\right] \ \backsim_{SO_3} \ \left[\rho(D_0) \sqcup \left(\mathbb{S}^2 \setminus D_0\right)\right] = \mathbb{S}^2 \setminus D,$$

$$\text{dove} \qquad D_0 = \bigsqcup_{n=0}^{\infty} \rho^n(D).$$

Poiché D è un sottoinsieme numerabile di $\mathbb{S}^2$, possiamo considerare una retta r passante per l'origine che non interseca D. Siano θ un angolo e $n \in \mathbb{N}_0$; fissata r poniamo $\rho_{r,n\theta} = \rho^n_{r,\theta}$ la rotazione di angolo $n\theta$ radianti avente per asse la retta r.
Adesso possiamo definire A come l'insieme degli angoli θ per cui esiste un $n \in \mathbb{N}_0$ ed un punto $P \in D$ tali che $\rho_{r,n\theta}(P) \in D$.
Dato che D è numerabile, anche A è numerabile visto che

$$A = \bigcup_{n \in \mathbb{N}_0} \bigcup_{P \in D} \left\{\theta : \ \rho_{r,n\theta}(P) \in D\right\}.$$

Ciò significa che esiste un θ non appartenente ad A ovvero, detto $\rho = \rho_{r,\theta}$, risulta

$$\rho^n(D) \cap D = \emptyset \qquad \text{per ogni } n > 0.$$

In particolare, se $0 \leqslant m < n$, vale ovviamente $\rho^{n-m}(D) \cap D = \emptyset$. Infine, come ultima conseguenza, si ha $\rho^n(D) \cap \rho^m(D) = \emptyset$ per ogni coppia di $n > m$ (infatti se così non fosse esisterebbero $P_1, P_2 \in D$ tali che $\rho^n(P_1) = \rho^m(P_2)$ ovvero $\rho^{n-m}(P_1) = P_2 \in D$ quindi $\rho^{n-m}(D) \cap D \neq \emptyset$, che è assurdo). □

Teorema 2.14. *(Paradosso di Banach-Tarski) (HB)*
La superficie sferica $\mathbb{S}^2$ è SO_3-paradossale. La sfera $\mathbb{B}$ è G_3-paradossale.

Dimostrazione. Il paradosso di Hausdorff (teorema 2.9) garantisce l'esistenza di un sottoinsieme numerabile D di $\mathbb{S}^2$ tale che $\mathbb{S}^2 \setminus D$ è SO_3-paradossale. Per il teorema precedente (2.13) $\mathbb{S}^2$ e $\mathbb{S}^2 \setminus D$ sono SO_3-equidecomponibili; allora possiamo applicare la proposizione 2.12 e concludere che $\mathbb{S}^2$ è SO_3-paradossale.
Osserviamo che la duplicazione di $\mathbb{S}^2$ (appena provata) implica, grazie alla corrispondenza radiale

$$P \longrightarrow \{\alpha P : \ 0 < \alpha \leqslant 1\},$$

una duplicazione di $\mathbb{B} \setminus \{0\}$, dunque $\mathbb{B} \setminus \{0\}$ è SO_3-paradossale e a maggior ragione G_3-paradossale.
Infine basta mostrare che $\mathbb{B} \setminus \{0\}$ e $\mathbb{B}$ sono G_3-equidecomponibili. Di conseguenza, per la proposizione 2.12, $\mathbb{B}$ risulta G_3-paradossale, cioè la tesi.
Sia $P = (0, 0, 1/2)$, consideriamo una rotazione ρ avente ordine infinito e asse passante per P ma non per l'origine. Quindi $\rho \in G_3 \setminus SO_3$. Posto $D = \{\rho^n(0) : \ n \in \mathbb{N}\}$, per le proprietà di ρ si ha

$$\rho(D) = D \setminus \{0\}.$$

Allora, indicato con D^C il complementare di D in $\mathbb{B}$, otteniamo

$$\mathbb{B} \setminus \{0\} = \left[(D \setminus \{0\}) \sqcup D^C\right] = \left[\rho(D) \sqcup D^C\right],$$

dunque

$$\mathbb{B} \setminus \{0\} = \left[\rho(D) \sqcup D^C\right] \ \backsim_{G_3} \ \left[D \sqcup D^C\right] = \mathbb{B},$$

cioè $B \setminus \{0\}$ e B sono G_3-equidecomponibili e la tesi è provata. □

2.4 Un corollario importante: l'esistenza di un insieme non misurabile

È noto che la misura di Lebesgue su $\mathbb{R}^n$ non è definita su tutti i sottoinsiemi dello spazio $\mathbb{R}^n$, ovvero che esistono insiemi non misurabili. Tipicamente la dimostrazione di questo fatto si basa sull'assioma della scelta. Per esempio nel caso $1-$dimensionale della misura di Lebesgue su $\mathbb{R}$, è l'assioma della scelta che permette di costruire l'insieme di Vitali, e di provarne la propria non misurabilità.

Vogliamo ora mostrare come, in realtà, sia possibile dedurre l'esistenza di un insieme non misurabile secondo Lebesgue grazie al teorema di Hahn-Banach, sfruttando il paradosso della duplicazione della sfera appena provato senza l'ausilio di AC. In questo contesto sarà naturale considerare la misura 3-dimensionale di Lebesgue, che è definita sulla $\sigma-$algebra dei misurabili. Questa classe, a priori, cioè senza usare l'assioma della scelta, potrebbe essere anche tutto $\mathcal{P}(\mathbb{R}^3)$.
Mostreremo che ciò non è possibile o, detto in altro modo, che esiste un sottoinsieme di $\mathbb{R}^3$ non misurabile secondo Lebesgue.

Definizione 2.15. Sia G un gruppo che agisce su un insieme X.
Una carica di probabilità $\mu : \mathcal{P}(X) \longrightarrow [0,1]$ si dice $G-$*invariante* se

$$\mu(gA) = \mu(A) \qquad \forall g \in G,\ \forall A \in \mathcal{P}(X).$$

Proposizione 2.16. *(HB)*
Non esiste una carica di probabilità SO_3-invariante definita su $\mathcal{P}(\mathbb{S}^2)$.

Dimostrazione. Sia $\mu : \mathcal{P}(\mathbb{S}^2) \longrightarrow [0,1]$ una carica di probabilità sull'algebra booleana $\mathcal{P}(\mathbb{S}^2)$ (una tale μ esiste per il teorema 1.25, la versione di HB per algebre booleane).
Per definizione di carica di probabilità vale $\mu(\mathbb{S}^2) = 1$.
Mostriamo che μ non può essere SO_3-invariante, utilizzando la duplicazione di $\mathbb{S}^2$ provata nel paragrafo precedente. Il teorema di Hahn-Banach implica il paradosso di Banach-Tarski (teorema 2.14), quindi esistono dei sottoinsiemi di $\mathbb{S}^2$ a due a due disgiunti $A_1, ..., A_n, B_1, ..., B_m$ e delle rotazioni $g_1, ..., g_n, h_1, ..., h_m$ di SO_3 tali che

$$\bigcup_{i=1}^{n} g_i A_i = \mathbb{S}^2 = \bigcup_{j=1}^{m} h_j B_j.$$

Essendo una carica di probabilità, rispetto all'unione μ è finitamente subadditiva (per il lemma 1.15) e finitamente additiva sui disgiunti (per definizione). Allora

$$(*) \qquad \mu(\mathbb{S}^2) \geqslant \mu\left(\left(\bigsqcup_{i=1}^{n} A_i\right) \bigsqcup \left(\bigsqcup_{j=1}^{m} B_j\right)\right) = \sum_{i=1}^{n} \mu(A_i) + \sum_{j=1}^{m} \mu(B_j).$$

D'altra parte vale anche

$$\mu(\mathbb{S}^2) = \mu\left(\bigcup_{i=1}^{n} g_i A_i\right) \leqslant \sum_{i=1}^{n} \mu(g_i A_i) \quad \text{e} \quad \mu(\mathbb{S}^2) = \mu\left(\bigcup_{j=1}^{m} h_j B_j\right) \leqslant \sum_{j=1}^{m} \mu(h_j B_j).$$

Supponiamo che μ sia SO_3-invariante. Ciò significa che le due disuguaglianze appena scritte assumono la forma

$$\mu(\mathbb{S}^2) \leqslant \sum_{i=1}^{n} \mu(g_i A_i) = \sum_{i=1}^{n} \mu(A_i) \quad \text{e} \quad \mu(\mathbb{S}^2) \leqslant \sum_{j=1}^{m} \mu(h_j B_j) = \sum_{j=1}^{m} \mu(B_j).$$

Utilizzando la disuguaglianza $(*)$, si conclude che

$$\sum_{i=1}^{n} \mu(A_i) + \sum_{j=1}^{m} \mu(B_j) \leqslant \sum_{i=1}^{n} \mu(A_i) \quad \text{e} \quad \sum_{i=1}^{n} \mu(A_i) + \sum_{j=1}^{m} \mu(B_j) \leqslant \sum_{j=1}^{m} \mu(B_j).$$

Dato che la carica μ è a valori non negativi, questo è possibile se e solo se

$$\sum_{i=1}^{n} \mu(A_i) = 0 = \sum_{j=1}^{m} \mu(B_j).$$

Per quanto detto risulterebbe allora $\mu(\mathbb{S}^2) = 0$, che è assurdo. In definitiva una carica di probabilità $\mu : \mathcal{P}(\mathbb{S}^2) \longrightarrow [0,1]$ non può essere SO_3-invariante. $\square$

Il nostro intento è quello di esibire un'opportuna misura μ di $\mathbb{R}^3$, invariante per rotazioni e con la proprietà $\mu(\mathbb{S}^2) = 1$, che supponiamo definita su tutto $\mathcal{P}(\mathbb{R}^3)$.
A questo punto la restrizione di questa misura a $\mathcal{P}(\mathbb{S}^2)$ risulta una carica di probabilità SO_3-invariante. Ma questo è assurdo per la proposizione precedente, dunque la misura scelta non può essere definita su tutto $\mathcal{P}(\mathbb{R}^3)$, ossia esiste un insieme di $\mathbb{R}^3$ non misurabile secondo la misura μ. Possiamo inoltre affermare che tale insieme sia contenuto in $\mathbb{S}^2$, dato che, ragionando allo stesso modo, non è possibile che $\mathcal{P}(\mathbb{S}^2)$ sia contenuto nella $\sigma-$ algebra degli insiemi $\mu-$misurabili.
Non possiamo direttamente considerare la misura m_3 di Lebesgue su $\mathbb{R}^3$, perché rispetto a m_3 la superficie sferica $\mathbb{S}^2$ è trascurabile.
Sarà decisivo nel nostro caso considerare la misura H_2 di Hausdorff in $\mathbb{R}^3$: infatti essa (anzi una sua versione normalizzata) avrà la proprietà richiesta $H_2(\mathbb{S}^2) = 1$.
Inoltre, dopo aver provato l'esistenza di un sottoinsieme di $\mathbb{S}^2$ non misurabile per la misura H_2, proietteremo tale insieme sul piano $\mathbb{R}^2$ e la misura H_2, stavolta in $\mathbb{R}^2$, ci farà da tramite con la misura m_2 di Lebesgue in $\mathbb{R}^2$, visto che esse coincidono a meno di una costante moltiplicativa.

Presentiamo brevemente le misure di Hausdorff in $\mathbb{R}^N$.

Le misure di Hausdorff si definiscono a partire dalla nozione base di diametro di un insieme. Se $E \subseteq \mathbb{R}^N$, il *diametro* di E è il numero non negativo (eventualmente infinito)

$$diam\ E = \begin{cases} 0 & \text{se } E = \emptyset, \\ \sup\{|x-y| : \ x, y \in E\} & \text{altrimenti.} \end{cases}$$

Sia $E \subseteq \mathbb{R}^N$, e siano p, $\delta > 0$. Definiamo

$$H^*_{p,\delta}(E) = \inf \left\{ \sum_{n\in\mathbb{N}} (diam\ U_n)^p : \ U_n \text{ aperti}, \ diam\ U_n < \delta, \ E \subseteq \bigcup_{n\in\mathbb{N}} U_n \right\}.$$

Per misurare in modo accurato l'insieme E quelli che contano sono i ricoprimenti con aperti di diametro piccolo: quindi un'approssimazione ottimale della misura di E si ottiene al limite per δ che tende a 0^+. Per $p > 0$ fissato, più δ è piccolo, più la quantità $H^*_{p,\delta}$ è grande, essendo l'estremo inferiore di un insieme più piccolo; in particolare:

Definizione 2.17. Fissati $E \subseteq \mathbb{R}^N$ e $p > 0$, la *misura esterna p-dimensionale di Hausdorff* H^*_p di E è data da

$$H^*_p(E) = \lim_{\delta\to 0^+} H^*_{p,\delta}(E) = \sup_{\delta>0} H^*_{p,\delta}(E).$$

Segue immediatamente dalla defінzione che H^*_p è a valori non negativi, monotona e subadditiva. La classe degli insiemi H_p-misurabili è

$$M_{H_p} = \left\{ E \subseteq \mathbb{R}^N \ : \ H^*_p(A) = H^*_p(A \cap E) + H^*_p(A \cap E^C) \quad \forall A \subseteq \mathbb{R}^N \right\}.$$

Osserviamo che in questa definizione quella che conta è la disuguaglianza $\geqslant$, dato che l'altra è sempre verificata per subadditività di H^*_p. Inoltre, per definizione, se $H^*_p(E) = 0$ allora E è H_p-misurabile: cioè H_p è completa.

Definizione 2.18. La *misura di Hausdorff di indice* p è $H_p = H^*_p|_{M_{H_p}}$.

Si può provare che M_{H_p} è una $\sigma-$algebra (contenente la $\sigma-$ algebra $\mathcal{B}$ dei boreliani) e che H_p è numerabilmente additiva sugli elementi disgiunti di M_{H_p}, quindi H_p è effettivamente una misura su $\mathbb{R}^N$.
È chiaramente invariante per traslazioni e isometrie lineari (dato che il diametro di un insieme, su cui si basa la definizione, è invariante per queste trasformazioni).

Ricordiamo che anche la misura di Lebesgue m_N su $\mathbb{R}^N$ si definisce a partire dalla misura esterna di Lebesgue m_N^* per mezzo del teorema di Carathéodory.
La misura m_N di Lebesgue su $\mathbb{R}^N$ è $m_N = m_N^*|_{\mathcal{M}_N}$, dove

$$\mathcal{M}_N = \left\{ E \subseteq \mathbb{R}^N \ : \ m_n^*(A) = m_n^*(A \cap E) + m_n^*(A \cap E^C) \quad \forall A \subseteq \mathbb{R}^N \right\}$$

è la $\sigma-$algebra dei misurabili secondo la misura di Lebesgue $N-$dimensionale.

Richiamiamo alcuni risultati.

Lemma 2.19. *Sia $E \subseteq \mathbb{R}^N$, allora $E \in \mathcal{M}_N$ se e solo se per ogni $\varepsilon > 0$ esiste un aperto A_ε di $\mathbb{R}^N$, contenente E, tale che*

$$m_N(A_\varepsilon \setminus E) < \varepsilon.$$

Teorema 2.20. *Esiste una costante α_N tale che per ogni insieme $E \subseteq \mathbb{R}^N$ si ha*

$$H_N^*(E) = \alpha_N m_N^*(E)$$

$$\textit{con} \qquad \alpha_N = \left(\frac{2}{\sqrt{\pi}}\right)^N \Gamma\left(\frac{N}{2}+1\right),$$

dove Γ è la funzione di Eulero, $\quad \Gamma(p) = \int_0^\infty x^{p-1}e^{-x}\,dx, \quad p > 0.$

Osservazione 2.21. Si può dimostrare che $\alpha_N = (m_N(B_0))^{-1}$, dove B_0 è la palla di $\mathbb{R}^N$ di diametro unitario. Ovvero la differenza tra le misure N-dimensionali di Hausdorff e Lebesgue è che la prima assegna misura 1 alla palla di diametro unitario, mentre la seconda assegna misura 1 al cubo di lato unitario (per i parallelepipedi la misura di Lebesgue è il volume N-dimensionale).

Corollario 2.22. *Su $\mathbb{R}^N$ la σ-algebra dei misurabili secondo la misura m_N di Lebesgue coincide con la σ-algebra dei misurabili secondo la misura H_N di Hausdorff.*

Dimostrazione. Per il teorema precedente le misure esterne H_N^* e m_N^* coincidono a meno di costante, allora ovviamente

$$\mathcal{M}_N = \left\{ E \subseteq \mathbb{R}^N \ : \ m_n^*(A) = m_n^*(A \cap E) + m_n^*(A \cap E^C) \quad \forall A \subseteq \mathbb{R}^N \right\} =$$

$$= \left\{ E \subseteq \mathbb{R}^N \ : \ H_N^*(A) = H_N^*(A \cap E) + H_N^*(A \cap E^C) \quad \forall A \subseteq \mathbb{R}^N \right\} = M_{H_N}.$$

□

Proposizione 2.23. *(HB)*
Esiste un sottoinsieme di $\mathbb{S}^2$ *non misurabile secondo* H_2.

Dimostrazione. Per quanto detto, H_2 è una misura su $\mathbb{R}^3$ invariante per rotazioni e definita sulla $\sigma-$algebra $M_{H_2} \subseteq \mathcal{P}(\mathbb{R}^3)$. Inoltre, è noto che per le superfici Σ sufficientemente regolari vale

$$H_2(\Sigma) = a(\Sigma),$$

dove con $a(\Sigma)$ intendiamo l'area della superficie. In particolare per la superficie sferica si ha

$$H_2(\mathbb{S}^2) = a(\mathbb{S}^2) = 4\pi.$$

Pertanto, a meno di normalizzare la misura H_2 (e con la normalizzazione ovviamente la classe dei misurabili M_{H_2} non viene alterata), possiamo concludere che $H_2(\mathbb{S}^2) = 1$.
Detto ciò, supponiamo per assurdo che M_{H_2} contenga $\mathcal{P}(\mathbb{S}^2)$.
In tal caso, restringendo H_2 a $\mathcal{P}(\mathbb{S}^2)$, otteniamo (essendo H_2 una misura, dunque additiva sui disgiunti) una carica di probabilità

$$H_2|_{\mathcal{P}(\mathbb{S}^2)} : \mathcal{P}(\mathbb{S}^2) \longrightarrow [0,1]$$

che è SO_3-invariante (dato che H_2 è invariante per rotazioni e SO_3 è proprio il gruppo delle rotazioni di $\mathbb{S}^2$). Questo è assurdo per la proposizione 2.16 (HB).
Quindi esiste un sottoinsieme della superficie sferica che non appartiene a M_{H_2}. $\square$

Osservazione 2.24. Sia E il sottoinsieme di $\mathbb{S}^2$ non misurabile secondo H_2. Possiamo decomporre E in

$$E = \left[E \cap (\mathbb{S}^2)^+\right] \cup \left[E \cap (\mathbb{S}^2)^-\right],$$

dove $(\mathbb{S}^2)^+$ è la parte superiore della superficie sferica (cioè l'insieme degli $(x,y,z) \in \mathbb{S}^2$ tali che $z \geqslant 0$) e $(\mathbb{S}^2)^-$ è quella inferiore.
Dato che $E \notin M_{H_2}$ e M_{H_2} è ovviamente chiusa per unione (addirittura numerabile, essendo una $\sigma-$algebra), ne consegue che almeno uno dei due insiemi in cui è stato decomposto E non è misurabile secondo H_2.
Senza perdita di generalità possiamo allora suppore che l'insieme $E \notin M_{H_2}$, dato dalla proposizione precedente (2.23), sia contenuto in $(\mathbb{S}^2)^+$.

Indichiamo con $\pi : (\mathbb{S}^2)^+ \longrightarrow \mathbb{R}^2$ l'applicazione definita da $\pi(x, y, z) = (x, y)$. Detta $\varphi = \pi^{-1}$ la sua inversa, si verifica che φ è lipschitziana di costante L.

Proposizione 2.25. *Sia $E \subseteq (\mathbb{S}^2)^+$. Se $E \notin M_{H_2}$ allora $\pi(E) \notin \mathcal{M}_2$.*

Dimostrazione. Detto $D = \pi(E)$, supponiamo per assurdo che $D \in \mathcal{M}_2$. Allora per il lemma 2.19 per ogni $\varepsilon > 0$ esiste un aperto $A_\varepsilon \supseteq D$ tale che $m_2(A_\varepsilon \setminus D) < \varepsilon$. Fissato $\varepsilon > 0$, è ovvio che $\varphi(A_\varepsilon) \setminus E = \varphi(A_\varepsilon) \setminus \varphi(D) = \varphi(A_\varepsilon \setminus D)$. Di conseguenza

$$H_2^*\left(\varphi(A_\varepsilon) \setminus E\right) = H_2^*\left(\varphi(A_\varepsilon \setminus D)\right) = \sup_{\delta>0} H_{2,\delta}^*\left(\varphi(A_\varepsilon \setminus D)\right) =$$

$$= \sup_{\delta>0} \inf \left\{ \sum_{n\in\mathbb{N}} (diam\ U_n)^2 : \ U_n \text{ aperti di } \mathbb{R}^3,\ diam\ U_n < \delta,\ \varphi(A_\varepsilon \setminus D) \subseteq \bigcup_{n\in\mathbb{N}} U_n \right\}$$

$$\leqslant \sup_{\delta>0} \inf \left\{ \sum_{n\in\mathbb{N}} (diam\ R_n)^2 : \ R_n \text{ aperti di } (\mathbb{S}^2)^+,\ diam\ R_n < \delta,\ \varphi(A_\varepsilon \setminus D) \subseteq \bigcup_{n\in\mathbb{N}} R_n \right\}$$

$$\leqslant \sup_{\delta>0} \inf \left\{ \sum_{n\in\mathbb{N}} (diam\ \varphi(I_n))^2 : \ I_n \text{ aperti di } \mathbb{R}^2,\ diam\ I_n < \delta,\ (A_\varepsilon \setminus D) \subseteq \bigcup_{n\in\mathbb{N}} I_n \right\}$$

$$\leqslant L^2 \sup_{\delta>0} \inf \left\{ \sum_{n\in\mathbb{N}} (diam\ I_n)^2 : \ I_n \text{ aperti di } \mathbb{R}^2,\ diam\ I_n < \delta,\ (A_\varepsilon \setminus D) \subseteq \bigcup_{n\in\mathbb{N}} I_n \right\}$$

$$= L^2 H_2^*\left(A_\varepsilon \setminus D\right) = \alpha_2 L^2 m_2\left(A_\varepsilon \setminus D\right) < \alpha_2 L^2 \varepsilon,$$

dove nell'ultima uguaglianza abbiamo usato il teorema 2.20.
Adesso, per ogni $\varepsilon = \frac{1}{n}$ al variare di $n \in \mathbb{N}_0$, esiste un aperto $A_n \supseteq D$ tale che $m_2(A_n \setminus D) < \frac{1}{n}$. Consideriamo l'intersezione $B = \bigcap_{n\in\mathbb{N}_0} A_n$: ovviamente B è un boreliano di $\mathbb{R}^2$ contenente D, da cui $\varphi(B) \supseteq \varphi(D) = E$. Osserviamo che, per le proprietà di φ, essa manda boreliani in boreliani: dunque $\varphi(B) \in \mathcal{B} \subset M_{H_2}$. Ripetendo il ragionamento fatto sopra si ottiene per ogni $n \in \mathbb{N}_0$

$$H_2^*\left(\varphi(B) \setminus E\right) \leqslant \alpha_2 L^2 m_2\left(B \setminus D\right) < \frac{\alpha_2 L^2}{n},$$

quindi

$$H_2^*\left(\varphi(B) \setminus E\right) = 0,$$

cioè E differisce dall'insieme H_2-misurabile $\varphi(B)$ per un insieme di misura H_2 nulla. Poiché H_2 è una misura completa, $H_2^*\left(\varphi(B) \setminus E\right) = 0$ implica $(\varphi(B) \setminus E) \in M_{H_2}$. Infine M_{H_2} è chiusa per differenza (essendo una $\sigma-$algebra), dunque concludiamo che E appartiene a M_{H_2}. Ma questo è assurdo per ipotesi: allora $\pi(E) \notin \mathcal{M}_2$. $\square$

Corollario 2.26. *(HB)*
Esiste un insieme di $\mathbb{R}^3$ *non misurabile secondo Lebesgue.*

Dimostrazione. Per la proposizione 2.23 e l'osservazione 2.24 esiste un sottoinsieme E di $(\mathbb{S}^2)^+$ che non appartiene a M_{H_2}.
Utilizzando la proposizione 2.25, otteniamo che $\pi(E) \notin \mathcal{M}_2$. A partire da questo insieme di $\mathbb{R}^2$ costruiamo l'insieme di $\mathbb{R}^3$

$$F = \pi(E) \times [0, 1].$$

L'insieme F non può essere misurabile secondo la misura m_3 di Lebesgue, altrimenti per quasi ogni $z \in [0, 1]$ la sezione

$$F^z = \big\{(x, y) \in \mathbb{R}^2 : (x, y, z) \in F\big\} = \pi(E)$$

risulterebbe misurabile rispetto a m_2. Ma questo è assurdo perché $\pi(E) \notin \mathcal{M}_2$.
Quindi $\pi(E) \times [0, 1]$ è un sottoinsieme di $\mathbb{R}^3$ non misurabile secondo Lebesgue. □

Osservazione 2.27. Grazie al paradosso di Banach-Tarski (teorema 2.14) la sfera $\mathbb{B}$ è G_3-paradossale pertanto, ragionando esattamente come nella dimostrazione di 2.16, deduciamo che non esiste una carica di probabilità G_3-invariante su $\mathcal{P}(\mathbb{B})$. Allora la misura m_3 di Lebesgue non può essere definita su tutto $\mathcal{P}(\mathbb{R}^3)$, altrimenti la restrizione di m_3 a $\mathcal{P}(\mathbb{B})$ risulterebbe una carica di probabilità G_3-invariante su $\mathcal{P}(\mathbb{B})$ (dato che m_3 è invariante per isometrie e $m_3(\mathbb{B}) = 1$ a meno di normalizzazione).
Per lo stesso motivo m_3 non può essere definita su una $\sigma-$algebra che contiene $\mathcal{P}(\mathbb{B})$.
In altre parole abbiamo provato, ancora una volta grazie al solo teorema di Hahn-Banach (cioè senza ricorrere all'assioma della scelta), che esiste un insieme di $\mathbb{R}^3$ (in questo caso un sottoinsieme di $\mathbb{B}$) non misurabile secondo Lebesgue.
La conclusione è la stessa del corollario 2.26, e in questo caso è stata immediata.
D'altra parte considerando la duplicazione paradossale di $\mathbb{S}^2$ e passando per la misura H_2 di Hausdorff abbiamo provato anche l'esistenza di un insieme di $\mathbb{R}^2$ non misurabile secondo Lebesgue (stiamo parlando di $\pi(E)$ della dimostrazione di 2.26), fatto certamente non banale (non potendo usare AC) dato che il paradosso di Banach-Tarski non vale né in $\mathbb{R}^2$ né in $\mathbb{R}$ (vedi [10], capitolo 10).

3 Varianti e generalizzazioni di Hahn-Banach

3.1 Condizioni per l'unicità dell'estensione

In generale il teorema di Hahn-Banach, nella sua formulazione per spazi normati (teorema 1.21), non assicura l'unicità dell'estensione del funzionale.
In effetti vi sono spazi normati X e funzionali $\lambda_0 \in Y^*$, dove Y è un sottospazio di X, che ammettono due diverse estensioni $\lambda_1, \lambda_2 \in X^*$ del funzionale $\lambda_0 \in Y^*$, per di più con la proprietà

$$\|\lambda_0\|_{Y^*} = \|\lambda_1\|_{X^*} = \|\lambda_2\|_{X^*}.$$

Vediamo un esempio.

Esempio 3.1. Indichiamo con $P[0,1]$ lo spazio dei polinomi definiti su $[0,1]$, che è normato con $\|p\|_\infty = \sup_{t\in[0,1]} |p(t)|$.
Detto $P_3[0,1]$ il sottospazio dei polinomi di grado minore o uguale a 3, consideriamo il funzionale $\lambda_0 : P_3[0,1] \longrightarrow \mathbb{R}$ definito da

$$\lambda_0(p) = \frac{1}{6}\left[p(0) + 4p\left(\frac{1}{2}\right) + p(1)\right] \qquad \forall\, p \in P_3[0,1].$$

Chiaramente λ_0 è un funzionale lineare e limitato con

$$\|\lambda_0\|_{(P_3[0,1])^*} = \sup_{0\neq p\in P_3[0,1]} \frac{|\lambda_0(p)|}{\|p\|_\infty} = 1.$$

Infatti per definizione di $\|.\|_\infty$ si ha $\|\lambda_0\| \leqslant 1$, e tale sup viene effettivamente raggiunto dal polinomio costante $p = 1$.
Esibiamo due diversi funzionali appartenenti a $P[0,1]^*$ che estendono λ_0, aventi entrambi norma unitaria, ovvero la stessa di λ_0. I funzionali che fanno al caso nostro sono i seguenti $\lambda_1, \lambda_2 : P[0,1] \longrightarrow \mathbb{R}$, dati da

$$\lambda_1(f) = \frac{1}{6}\left[f(0) + 4f\left(\frac{1}{2}\right) + f(1)\right], \qquad \lambda_2(f) = \int_0^1 f(t)\, dt, \qquad \forall f \in P[0,1].$$

Da una parte λ_1 è l'ovvia estensione di λ_0 a tutto lo spazio, e mantiene per quanto detto la proprietà $\|\lambda_1\|_{P[0,1]^*} = 1$.
Dall'altra verifichiamo che il funzionale lineare λ_2 (dato dell'integrale scritto sopra) sia effettivamente una estensione di λ_0. Indicato con $p(t) = at^3 + bt^2 + ct + d$ un qualunque polinomio di grado al più 3, vale

$$\int_0^1 \left(at^3 + bt^2 + ct + d\right)\, dt = \frac{a}{4} + \frac{b}{3} + \frac{c}{2} + d = \frac{1}{6}\left[6d + 3c + 2b + \frac{3}{2}a\right] =$$

$$= \frac{1}{6}\left[d + 4\left(\frac{a}{8} + \frac{b}{4} + \frac{c}{2} + d\right) + (a+b+c+d)\right] = \frac{1}{6}\left[p(0) + 4p\left(\frac{1}{2}\right) + p(1)\right].$$

Per la norma si ha ovviamente $\|\lambda_2\|_{P[0,1]^*} = 1$, dato che per ogni $f \in P[0,1]$ possiamo maggiorare $\left|\int_0^1 f(t)\ dt\right|$ con $\|f\|_\infty$ e ancora una volta il sup è raggiunto dal polinomio costante $p = 1$.
Infine verifichiamo che sono diversi: basta valutarli nel polinomio t^4 e osservare che

$$\lambda_1(t^4) = \frac{1}{6}\left[0 + 4 \cdot \frac{1}{16} + 1\right] = \frac{5}{24} \neq \frac{1}{5} = \int_0^1 t^4\ dt = \lambda_2(t^4).$$

In ogni caso, esistono spazi normati per cui vale l'unicità dell'estensione di HB. A tal proposito mostriamo prima una classe molto speciale di spazi normati, in cui tale estensione esiste esplicitamente (cioè senza richiedere AC) ed è garantita l'unicità. In seguito caratterizzeremo gli spazi normati dotati di questa proprietà di unicità per l'estensione HB.

Una prima classe di spazi normati per cui vale l'unicità dell'estensione di HB è quella degli spazi di Hilbert.
Ricordiamo che uno spazio di Hilbert $(\mathcal{H}, \langle , \rangle)$ è uno spazio normato completo in cui la norma è indotta dal prodotto scalare $\langle , \rangle$ secondo la formula

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Dato un $\mathbb{R}-$spazio normato X, la norma di X è indotta da un prodotto scalare se e solo se vale l'identità del parallelogramma

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \qquad \forall\ x, y \in X.$$

Dato un sottoinsieme M di uno spazio di Hilbert $(\mathcal{H}, \langle,\rangle)$, indichiamo con $M^\perp$ il sottospazio ortogonale a M, rispetto al prodotto scalare, definito da

$$M^\perp = \{x \in \mathcal{H} : \ \langle x, y\rangle \ = 0 \quad \forall\ y \in M\}\,.$$

Notiamo che, a causa della linearità e della continuità del prodotto scalare, $M^\perp$ è un sottospazio chiuso di $\mathcal{H}$, comunque sia fatto l'insieme M.

Per provare esistenza e unicità dell'estensione, abbiamo bisogno, rispettivamente, del teorema delle proiezioni e del teorema di rappresentazione di Riesz, due ben noti risultati validi per spazi di Hilbert, che richiamiamo qui sotto.

Teorema delle proiezioni: siano $\mathcal{H}$ uno spazio di Hilbert e M un suo sottospazio chiuso, allora sono univocamente determinate due proiezioni lineari e limitate

$$P : \mathcal{H} \longrightarrow M, \qquad Q : \mathcal{H} \longrightarrow M^\perp \qquad \text{tali che}$$

i) per ogni $x \in \mathcal{H} \quad x = P(x) + Q(x)$;
ii) $x \in M \Leftrightarrow [P(x) = x,\ Q(x) = 0]\,, \qquad x \in M^\perp \Leftrightarrow [P(x) = 0,\ Q(x) = x]\,;$
iii) per ogni $x \in \mathcal{H} \quad \|x\|^2 = \|P(x)\|^2 + \|Q(x)\|^2\,.$
P e Q si chiamano le proiezioni ortogonali su M e $M^\perp$, rispettivamente.

Teorema di rappresentazione di Riesz: dato uno spazio di Hilbert $\mathcal{H}$, sia $L \in \mathcal{H}^*$, allora esiste uno ed un solo $z \in \mathcal{H}$ tale che

$$L(x) = \ \langle x, z\rangle \quad \forall\ x \in \mathcal{H};$$

ed inoltre si ha

$$\|L\|_{\mathcal{H}^*} = \|z\|\,.$$

È interessante notare che, per l'esistenza dell'estensione di funzionali definiti su sottospazi di uno spazio di Hilbert, non è necessario ricorrere all'assioma della scelta dato che, come ci apprestiamo a mostrare, tale estensione può essere descritta esplicitamente grazie al teorema delle proiezioni. Per questo motivo enunciamo di nuovo il teorema di HB in questo caso molto particolare, e lo dimostriamo, potendo garantire anche l'unicità dell'estensione.

Teorema 3.2. *(Esistenza e unicità dell'estensione di HB per spazi di Hilbert) Siano $(\mathcal{H}, \langle , \rangle)$ un $\mathbb{R}$-spazio di Hilbert, Y un suo sottospazio e $\lambda_0 \in Y^*$. Allora λ_0 si estende a un funzionale $\lambda \in \mathcal{H}^*$ che soddisfa $\|\lambda\|_{\mathcal{H}^*} = \|\lambda_0\|_{Y^*}$.*
Inoltre, tale estensione è unica.

Dimostrazione. Sia $\overline{Y}$ la chiusura di Y in X. Essendo $\mathbb{R}$ completo, esiste un'unica estensione lineare e limitata $\overline{\lambda} : \overline{Y} \longrightarrow \mathbb{R}$ di λ_0, che per di più soddisfa

$$\left\|\overline{\lambda}\right\|_{\overline{Y}^*} = \|\lambda_0\|_{Y^*}.$$

Infatti, per ogni $y \in \overline{Y}$, esiste una successione $\{y_n\}_{n\in\mathbb{N}}$ di elementi di Y che converge a y: dato che

$$|\lambda_0(y_n) - \lambda_0(y_m)| \leqslant \|\lambda_0\|_{Y^*} \|y_n - y_m\| \qquad \forall\, n, m \in \mathbb{N},$$

per completezza di $\mathbb{R}$ esiste un numero reale a cui la successione $\{\lambda_0(y_n)\}_{n\in\mathbb{N}}$ converge. Poniamo allora

$$\overline{\lambda}(y) = \lim_{n\to+\infty} \lambda_0(y_n),$$

che è ben definita (tale limite non dipende dalla successione $\{y_n\}$).
Osserviamo che $\overline{\lambda}$ è un'estensione di λ_0 perché se $y \in Y$ possiamo direttamente considerare la successione costante $y_n = y$ per ogni n, da cui la successione costante $\{\lambda_0(y_n)\}$ converge banalmente a $\lambda_0(y)$.
Si verifica poi facilmente che tale estensione è lineare, limitata con norma uguale a quella di λ_0. È inoltre l'unica estensione lineare e limitata di λ_0 definita su $\overline{Y}$. Supponiamo infatti esista un'altra estensione lineare e limitata $\lambda' : \overline{Y} \longrightarrow \mathbb{R}$ di λ_0. Per ogni $y \in \overline{Y} \setminus Y$, sia $\{y_n\} \subset Y$ una successione convergente a y. Allora per continuità di λ' e per definizione di $\overline{\lambda}$ si ha

$$\lambda'(y) = \lim_{n\to+\infty} \lambda'(y_n) = \lim_{n\to+\infty} \lambda_0(y_n) = \overline{\lambda}(y),$$

visto che $\lambda'|_Y = \lambda_0$.
Dato che $\mathcal{H}$ è uno spazio di Hilbert possiamo, sfruttando il teorema delle proiezioni, decomporre ogni $x \in \mathcal{H}$ nella forma

$$x = P(x) + Q(x),$$

dove P e Q sono rispettivamente le proiezioni sui sottospazi chiusi $\overline{Y}$ e $\overline{Y}^\perp$.
Definiamo allora il funzionale $\lambda : \mathcal{H} \longrightarrow \mathbb{R}$ in questo modo:

$$\lambda(x) = \overline{\lambda}(P(x)) \qquad \forall\, x \in \mathcal{H}.$$

Mostriamo che λ è l'estensione di λ_0 con le proprietà richieste. La linearità è ovvia (composizione di mappe lineari). Osserviamo poi che per ogni $y \in Y$ vale

$$\lambda(y) = \overline{\lambda}(P(y)) = \overline{\lambda}(y) = \lambda_0(y),$$

quindi λ è effettivamente una estensione di λ_0. Infine dobbiamo verificare che $\|\lambda\|_{\mathcal{H}^*} = \|\lambda_0\|_{Y^*} < \infty$, da cui segue anche la limitatezza di λ.
Per questo basta notare che

$$\|\lambda_0\|_{Y^*} = \sup_{0 \neq y \in Y} \frac{|\lambda_0(y)|}{\|y\|} \leqslant \sup_{0 \neq x \in \mathcal{H}} \frac{|\lambda(x)|}{\|x\|} = \|\lambda\|_{\mathcal{H}^*} = \sup_{0 \neq x \in \mathcal{H}} \frac{\left|\overline{\lambda}(P(x))\right|}{\|x\|} \leqslant$$

$$\leqslant \sup_{0 \neq x \in \mathcal{H}} \frac{\left|\overline{\lambda}(P(x))\right|}{\|P(x)\|} = \left\|\overline{\lambda}\right\|_{\overline{Y}^*} = \|\lambda_0\|_{Y^*},$$

avendo utilizzato la definizione di λ, il fatto che estende λ_0, e la ovvia disuguaglianza $\|P(x)\| \leqslant \|x\|$.
Per l'unicità dell'estensione, possiamo assumere, senza perdita di generalità, che il sottospazio Y sia chiuso (dato che l'estensione $\overline{\lambda}$ del funzionale a $\overline{Y}$ è unica).
Supponiamo quindi esista un altro funzionale $\lambda_1 \in \mathcal{H}^*$ che estende λ_0 e con la proprietà $\|\lambda_1\|_{\mathcal{H}^*} = \|\lambda_0\|_{Y^*}$: vogliamo provare che $\lambda_1 = \lambda$.
Per il teorema di rappresetazione di Riesz esiste uno ed un solo $z \in \mathcal{H}$ tale che

$$\lambda_1(x) = \langle x, z \rangle \quad \forall\, x \in \mathcal{H}, \quad \text{di conseguenza} \quad \|\lambda_1\|_{\mathcal{H}^*} = \|z\|.$$

Siccome λ_1 è un'estensione di λ_0, vale

$$\lambda_1(y) = \lambda_0(y) = \langle y, z \rangle = \langle y, Pz \rangle \qquad \forall\, y \in Y,$$

dove nell'ultima uguaglianza abbiamo sfruttato il fatto che $z - P(z) = Q(z) \in Y^\perp$ e la linearità del prodotto scalare. Di conseguenza (direttamente dalla disuguaglianza di Cauchy-Schwarz) vale $\|\lambda_0\|_{Y^*} = \|P(z)\|$. Abbiamo ottenuto

$$\|P(z)\| = \|\lambda_0\|_{Y^*} = \|\lambda_1\|_{\mathcal{H}^*} = \|z\|,$$

ovvero (vedi teorema delle proiezioni) $P(z) = z$. Dunque per ogni $x \in \mathcal{H}$ (ricordando che le proiezione ortogonali sono operatori autoaggiunti)

$$\lambda_1(x) = \langle x, z \rangle = \langle x, P(z) \rangle = \langle P(x), z \rangle = \lambda_1(P(x)) = \lambda_0(P(x)) = \lambda(x),$$

cioè $\lambda = \lambda_1$ come volevasi dimostrare. $\square$

A completamento di questa sezione, vediamo un teorema generale che fornisce una condizione sufficiente e anche necessaria affinché uno spazio normato X ammetta un'unica estensione di HB, ovvero che per ogni sottospazio Y di X e per ogni $\lambda_0 \in Y^*$ esista un'unica estensione $\lambda \in X^*$ che soddisfa $\|\lambda\|_{X^*} = \|\lambda_0\|_{Y^*}$.
Nell'enunciato del prossimo teorema sintetizziamo questa proprietà con la frase $''X$ ammette un'unica estensione di Hahn-Banach$''$.
Ricordiamo che l'esistenza è garantita proprio dal teorema di HB, che nella sua forma generale richiede AC. Per questo ci concentriamo solo sull'unicità, sottointendendo HB.

Premettiamo la seguente definizione.

Definizione 3.3. Un $\mathbb{R}-$spazio normato è *strettamente convesso* se per ogni coppia di suoi elementi distinti $x \neq y$, entrambi di norma unitaria, si ha

$$\left\| \frac{x+y}{2} \right\| < 1$$

(in generale vale solo $\left\| \frac{x+y}{2} \right\| \leqslant \frac{\|x\|+\|y\|}{2} = 1$, per la subadditività della norma).

Esempio 3.4. 1) Ogni spazio di Hilbert è strettamente convesso.
Ciò segue direttamente dalla identità del parallelogramma. Infatti se $x \neq y$ ed entrambi hanno norma unitaria, si ha $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 = 4$, ovvero

$$\left\| \frac{x+y}{2} \right\|^2 = 1 - \left\| \frac{x-y}{2} \right\|^2 < 1.$$

2) Il duale di $(P[0,1], \|.\|_\infty)$ non è strettamente convesso.
Ripercorrendo l'esempio 3.1, esistono due diversi funzionali $\lambda_1 \neq \lambda_2 \in P[0,1]^*$, entrambi di norma unitaria. Nel dettaglio

$$\lambda_1(f) = \frac{1}{6}\left[f(0) + 4f\left(\frac{1}{2}\right) + f(1) \right], \qquad \lambda_2(f) = \int_0^1 f(t)\,dt, \qquad \forall f \in P[0,1].$$

Dato che per il polinomio costante $p(t) = 1$ vale ovviamente $\frac{\lambda_1+\lambda_2}{2} = 1$ si ha

$$\left\| \frac{\lambda_1+\lambda_2}{2} \right\|_{P[0,1]^*} = 1,$$

dunque $P[0,1]^*$ non è strettamente convesso.

Teorema 3.5. *(Unicità dell'estensione di Taylor-Foguel)*
Sia X un $\mathbb{R}-$spazio normato. Allora X ammette un'unica estensione di HB se e solo se lo spazio duale X^* è strettamente convesso.

Dimostrazione. Dato un sottospazio Y di X, e un funzionale $\lambda_0 \in Y^*$, che possiamo assumere di norma unitaria, supponiamo esistano due funzionali $\lambda_1 \neq \lambda_2 \in X^*$ che estendono λ_0 e con la proprietà

$$\|\lambda_1\|_{X^*} = \|\lambda_2\|_{X^*} = \|\lambda_0\|_{Y^*} = 1.$$

Proviamo che allora X^* non è strettamente convesso.
Trattandosi di estensioni, per ogni $y \in Y$ vale ovviamente

$$\frac{\lambda_1(y) + \lambda_2(y)}{2} = \lambda_0(y),$$

da cui segue

$$1 = \|\lambda_0\|_{Y^*} = \left\|\frac{\lambda_1 + \lambda_2}{2}\right\|_{Y^*} \leqslant \left\|\frac{\lambda_1 + \lambda_2}{2}\right\|_{X^*}.$$

Pertanto X^* non è strettamente convesso.
Viceversa, assumendo l'unicità dell'estensione di HB, siano $\lambda_1 \neq \lambda_2 \in X^*$ due diversi funzionali di norma unitaria. La tesi da provare è

$$\left\|\frac{\lambda_1 + \lambda_2}{2}\right\|_{X^*} < 1.$$

Detto Y il sottospazio proprio di X in cui i due funzionali coincidono, ovvero

$$Y = \{y \in X : \quad \lambda_1(y) = \lambda_2(y)\},$$

consideriamo il funzionale lineare e limitato $\lambda : Y \longrightarrow \mathbb{R}$ definito da

$$\lambda(y) = \lambda_1(y) \qquad \forall\, y \in Y.$$

Affermiamo che

$$\|\lambda\|_{Y^*} < 1;$$

infatti da una parte è ovvio che $\|\lambda\|_{Y^*} \leqslant \|\lambda_1\|_{X^*} = 1$, dall'altra non è possibile che valga $\|\lambda\|_{Y^*} = 1$, altrimenti λ_1 e λ_2 sarebbero due diverse estensioni di HB dello stesso funzionale λ.
Per ipotesi esiste un unico $\widehat{\lambda} \in X^*$ tale che

$$\widehat{\lambda}|_Y = \lambda = \lambda_1|_Y = \lambda_2|_Y \quad \text{e} \quad \left\|\widehat{\lambda}\right\|_{X^*} = \|\lambda\|_{Y^*} < 1.$$

Sia $x_0 \in X \setminus Y$. Per definizione di Y ciò significa che $\lambda_1(x_0) \neq \lambda_2(x_0)$, comunque esiste un numero reale $\mu = \mu(x_0)$ tale che

$$\mu\lambda_1(x_0) + (1-\mu)\lambda_2(x_0) = \widehat{\lambda}(x_0).$$

Dato che ogni $x \in X$ può essere scritto nella forma $x = y + \alpha x_0$, per opportuni $y \in Y$ e $\alpha \in \mathbb{R}$, per linearità dei funzionali in gioco vale

$$\widehat{\lambda}(x) = \widehat{\lambda}(y + \alpha x_0) = \widehat{\lambda}(y) + \alpha\widehat{\lambda}(x_0) = \lambda_1(y) + \alpha\left[\mu\lambda_1(x_0) + (1-\mu)\lambda_2(x_0)\right] =$$

$$= \left[\mu\lambda_1(y) + (1-\mu)\lambda_1(y)\right] + \left[\mu\lambda_1(\alpha x_0) + (1-\mu)\lambda_2(\alpha x_0)\right] =$$

$$= \mu\left[\lambda_1(y + \alpha x_0)\right] + (1-\mu)\left[\lambda_2(y + \alpha x_0)\right] = \mu\lambda_1(x) + (1-\mu)\lambda_2(x),$$

cioè

$$\widehat{\lambda} = \mu\lambda_1 + (1-\mu)\lambda_2 \qquad \text{per un opportuno } \mu \in \mathbb{R}.$$

Motiviamo che $\mu \in]0,1[$. Infatti, se fosse $\mu \geqslant 1$, avremmo

$$\lambda_1 = \frac{1}{\mu}\widehat{\lambda} + \frac{\mu-1}{\mu}\lambda_2,$$

che implicherebbe

$$\|\lambda_1\|_{X^*} \leqslant \frac{1}{\mu}\left\|\widehat{\lambda}\right\|_{X^*} + \frac{\mu-1}{\mu}\|\lambda_2\|_{X^*} < \frac{1}{\mu} + \frac{\mu-1}{\mu} = 1, \quad \text{che è assurdo.}$$

Allo stesso modo $\mu \leqslant 0$ implicherebbe l'altra contraddizione $\|\lambda_2\|_{X^*} < 1$.
Infine eseguiamo la verifica finale della stretta convessità.
Se $0 < \mu < \frac{1}{2}$ oppure $\frac{1}{2} < \mu < 1$ allora, rispettivamente,

$$\frac{\lambda_1 + \lambda_2}{2} = \frac{1}{2(1-\mu)}\widehat{\lambda} + \frac{1-2\mu}{2(1-\mu)}\lambda_2, \quad \text{oppure} \quad \frac{\lambda_1 + \lambda_2}{2} = \frac{1}{2\mu}\widehat{\lambda} + \frac{2\mu-1}{2\mu}\lambda_2,$$

che in ambedue i casi implica, analogamente a prima,

$$\left\|\frac{\lambda_1 + \lambda_2}{2}\right\|_{X^*} < 1.$$

Rimane il caso $\mu = \frac{1}{2}$, che è quello più immediato: in tal caso

$$\frac{\lambda_1 + \lambda_2}{2} = \widehat{\lambda} \qquad \text{da cui} \qquad \left\|\frac{\lambda_1 + \lambda_2}{2}\right\|_{X^*} = \left\|\widehat{\lambda}\right\|_{X^*} < 1.$$

Quindi X^* è strettamente convesso. $\square$

Osservazione 3.6. Il teorema generale appena provato è coerente con quanto visto per gli spazi di Hilbert. Non solo, il teorema di unicità dell'estensione valido per spazi di Hilbert è a questo punto un immediato corollario di quest'ultimo. Infatti, se $\mathcal{H}$ è uno spazio di Hilbert, allora (per il teorema di rappresentazione di Riesz) $\mathcal{H}^*$ è isomorfo a $\mathcal{H}$ che è strettamente convesso; quindi possiamo applicare il teorema. Inoltre, come avevamo notato nell'esempio 3.1, lo spazio normato dei polinomi $(P[0,1], \|.\|_\infty)$ non ammette estensione unica di HB per un funzionale lineare definito sul sottospazio $P_3[0,1]$: anche questo adesso può essere dedotto dal teorema, dato che, come osservato in 3.4, $P[0,1]^*$ non è strettamente convesso.

3.2 Il problema dell'estensione lineare e continua: esempi e controesempi

In analisi funzionale il teorema di Hahn-Banach rappresenta un importante e classico risultato che garantisce la possibilità di estendere, in modo lineare e limitato, un funzionale a valori nella retta reale, addirittura lasciando inalterata la norma del funzionale. In questo ambito, è naturale chiederci quanto sia importante il ruolo giocato da $\mathbb{R}$, e se sia possibile estendere in modo lineare e continuo, alla maniera di HB, un funzionale a valori in un qualsiasi spazio normato, o per lo meno a valori in qualche particolare spazio normato.

Molto spesso il tentativo di estendere in modo lineare e continuo un funzionale definito su un sottospazio Y di X a valori in un generico spazio Z, in maniera che l'immagine dell'estensione sia ancora contenuta in Z, fallisce. Mostriamo, a tal proposito, un primo esempio davvero immediato di operatore che non ammette una tale estensione lineare e continua.

Esempio 3.7. Sia Y un sottospazio non chiuso di uno spazio normato X.
Ovviamente l'operatore identità $Id : Y \longrightarrow Y$ è lineare e continuo. L'identità non può però ammettere in questo caso un'estensione lineare e continua su tutto lo spazio X la cui immagine sia ancora Y, altrimenti l'insieme dei punti fissi

$$\{x \in X : \ Id(x) = x\} = Y$$

risulterebbe, per continuità, un chiuso di X: assurdo per l'ipotesi su Y.

Visto che questo controesempio vale nella sua massima generalità, d'ora in avanti ci concentriamo su funzionali $f \in \mathcal{L}(Y, Z)$, dove Y è un sottospazio chiuso di uno spazio $(X, \|.\|)$, che da qui in poi supponiamo essere uno spazio di Banach. Così facendo il sottospazio Y, essendo chiuso, risulterà automaticamente a sua volta uno spazio di Banach rispetto alla norma $\|.\|_Y$ (la restrizione della norma di X al sottospazio Y).

Se da una parte una tale estensione continua può non esistere per funzionali a valori in sottospazi non chiusi, la risposta è ancora negativa, in generale, anche per funzionali a valori in sottospazi chiusi.
I controesempi con un sottospazio chiuso sono più complicati, come vedremo nel corso del paragrafo, esibendone uno importante.

Ci sono anche degli spazi di Banach Z in cui, come per $\mathbb{R}$, non solo esiste l'estensione lineare e continua del funzionale a valori in Z e tale estensione è ancora a valori in Z, ma viene anche preservata la norma. Caratterizzeremo questi spazi molto particolari per cui vale l'analoga proprietà di estensione lineare continua descritta dal teorema di Hahn-Banach.

Per cominciare vediamo come questa proprietà valga ancora considerando operatori a valori in $\left(\mathbb{R}^N, \|.\|_\infty\right)$, dove per un $x = (x_1, ..., x_N) \in \mathbb{R}^N$ si considera la norma $\|x\|_\infty = \max_{1\leqslant i\leqslant N} |x_i|$. Osserviamo che, nonostante l'estensione in X^* esista indipendentemente dalla norma scelta, la norma $\|.\|_\infty$ è fondamentale affinché sia preservata la norma dell'operatore.

Proposizione 3.8. *Fissato $N \in \mathbb{N}_0$, siano X un $\mathbb{R}$-spazio normato, Y un suo sottospazio, $g \in \mathcal{L}(Y, \mathbb{R}^N)$. Allora g si estende a un funzionale $f \in \mathcal{L}(X, \mathbb{R}^N)$.*
Per di più, considerando su $\mathbb{R}^N$ la norma $\|.\|_\infty$, l'estensione f ha la proprietà aggiuntiva $\|f\|_{\mathcal{L}(X,\mathbb{R}^N)} = \|g\|_{\mathcal{L}(Y,\mathbb{R}^N)}$.

Dimostrazione. Siano N, X e Y fissati come nelle ipotesi. Poiché $g \in \mathcal{L}(Y, \mathbb{R}^N)$, esistono N opportuni funzionali $y_1^*, ..., y_N^* \in Y^*$ tali che

$$g(y) = (y_1^*(y), ..., y_N^*(y)) \qquad \forall\, y \in Y.$$

Applicando il teorema di HB ad ognuno dei funzionali y_i^*, otteniamo $x_1^*, ..., x_N^* \in X^*$, rispettivamente estensioni di $y_1^*, ..., y_N^*$ che soddisfano $\|x_i^*\|_{X^*} = \|y_i^*\|_{Y^*}$ per ogni $i = 1, ..., N$. Per costruzione il funzionale $f \in \mathcal{L}(X, \mathbb{R}^N)$ definito da

$$f(x) = (x_1^*(x), ..., x_N^*(x)) \qquad \forall\, x \in X$$

è ovviamente una estensione di g.
Inoltre, considerando su $\mathbb{R}^N$ la norma $\|.\|_\infty$, per tale f vale

$$\|f\|_{\mathcal{L}(X,\mathbb{R}^N)} = \sup_{0\neq x\in X} \frac{\|f(x)\|_\infty}{\|x\|} = \sup_{0\neq x\in X} \frac{\max_{1\leqslant i\leqslant N} |x_i^*(x)|}{\|x\|} =$$

$$= \max_{1\leqslant i\leqslant N} \|x_i^*\|_{X^*} = \max_{1\leqslant i\leqslant N} \|y_i^*\|_{Y^*} = \|g\|_{\mathcal{L}(Y,\mathbb{R}^N)}\,.$$

□

Come esempio di spazio di dimensione infinita che soddisfa la proprietà di estensione di HB, vediamo lo spazio di Banach delle successioni reali limitate

$$\ell^\infty = \big\{x = \{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R} : \quad \{x_n\} \quad \text{è limitata}\big\}$$

munito della norma $\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$.
La dimostrazione di questo fatto è un'altra applicazione del teorema di HB, analoga a quanto visto nel caso di $\mathbb{R}^N$.
Successivamente proveremo che il suo sottospazio chiuso

$$c_0 = \big\{x = \{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R} : \quad \{x_n\} \quad \text{è infinitesima}\big\}$$

non soddisfa questa proprietà, ossia che esistono uno spazio di Banach X, un suo sottospazio chiuso Y ed un funzionale $f \in \mathcal{L}(Y, c_0)$ che non può essere esteso in modo lineare e continuo in maniera che l'immagine dell'estensione sia ancora contenuta in c_0. Anche in questa occasione basterà considerare l'identità $Id : c_0 \longrightarrow c_0$: vedremo perché non può essere estesa ad un'applicazione lineare e continua $I : \ell^\infty \longrightarrow c_0$.

Proposizione 3.9. *Siano X un $\mathbb{R}$-spazio di Banach, Y un suo sottospazio chiuso, $g \in \mathcal{L}(Y, \ell^\infty)$. Allora g si estende a un funzionale $f \in \mathcal{L}(X, \ell^\infty)$ che soddisfa $\|f\|_{\mathcal{L}(X,\ell^\infty)} = \|g\|_{\mathcal{L}(Y,\ell^\infty)}$.*

Dimostrazione. Fissati X, Y come nell'ipotesi, sia $g \in \mathcal{L}(Y, \ell^\infty)$. Allora per una qualche successione di funzionali $\{y_n^*\}_{n\in\mathbb{N}} \subseteq Y^*$ vale

$$g(y) = \{y_n^*(y)\}_{n\in\mathbb{N}} \qquad \forall\, y \in Y,$$

di conseguenza per la norma abbiamo

$$\|g\|_{\mathcal{L}(Y,\ell^\infty)} = \sup_{0\neq y\in Y} \frac{\|g(y)\|_\infty}{\|y\|} = \sup_{0\neq y\in Y} \frac{\sup_{n\in\mathbb{N}} |y_n^*(y)|}{\|y\|} = \sup_{n\in\mathbb{N}} \|y_n^*\|_{Y^*}\,.$$

Applicando il teorema di Hahn-Banach ad ogni y_n^*, otteniamo per ogni n una estensione $x_n^* \in X^*$ di y_n^* con $\|x_n^*\|_{X^*} = \|y_n^*\|_{Y^*}$. Basta quindi definire $f : X \longrightarrow \ell^\infty$ ponendo

$$f(x) = \{x_n^*(x)\}_{n\in\mathbb{N}}\,.$$

Per costruzione f è un'estensione lineare e limitata di g che soddisfa

$$\|f\|_{\mathcal{L}(X,\ell^\infty)} = \sup_{n\in\mathbb{N}} \|x_n^*\|_{X^*} = \sup_{n\in\mathbb{N}} \|y_n^*\|_{Y^*} = \|g\|_{\mathcal{L}(Y,\ell^\infty)}\,.$$

$\square$

Per quanto riguarda lo spazio c_0 delle successioni reali infinitesime, il fatto che l'identità $Id : c_0 \longrightarrow c_0$ non possa essere estesa ad un'applicazione $I \in \mathcal{L}(\ell^\infty, c_0)$ seguirà facilmente dai prossimi due lemmi.

Lemma 3.10. *Ogni insieme A infinito numerabile possiede una famiglia più che numerabile di sottoinsiemi infiniti $\{A_i\}_{i\in I}$ tale che ogni coppia di elementi in $\{A_i\}_{i\in I}$ ammette intersezione finita.*

Dimostrazione. Senza perdita di generalità possiamo identificare A con l'insieme numerabile e denso dei numeri razionali $\mathbb{Q}$. Per ogni numero irrazionale θ, esiste una successione di razionali $\{q_n\}_{n\in\mathbb{N}}$ che converge a θ. Allora gli insiemi della forma

$$A_\theta = \Big\{\{q_n\}_{n\in\mathbb{N}} : \quad \lim_{n\to\infty} q_n = \theta\Big\} \subset \mathbb{Q}$$

verificano il lemma essendo infiniti, in numero più che numerabile, e presi due di essi $A_{\theta 1}$ e $A_{\theta 2}$ (con $\theta_1 \neq \theta_2$) la loro intersezione non può essere infinita, altrimenti le successioni convergerebbero allo stesso irrazionale. □

Dato un sottoinsieme A di $\mathbb{N}$, denotiamo con $\ell^\infty(A)$ il sottospazio di ℓ^∞ dato da

$$\ell^\infty(A) = \big\{x = \{x_n\}_{n\in\mathbb{N}} \in \ell^\infty : \quad x_n = 0 \quad \forall\, n \notin A\big\}$$

ovvero il sottospazio delle successioni limitate con supporto contenuto in A.

Lemma 3.11. *Sia $T \in \mathcal{L}(\ell^\infty, \ell^\infty)$ tale che $T|_{c_0} = 0$. Allora esiste un sottoinsieme infinito $A \subseteq \mathbb{N}$ tale che $T|_{\ell^\infty(A)} = 0$.*

Dimostrazione. Usiamo la famiglia più che numerabile $\{A_i\}_{i\in I}$ di sottoinsiemi infiniti di $\mathbb{N}$, data dal lemma 3.10. Supponiamo per assurdo che per ognuno di questi insiemi esista un elemento non nullo $x_i \in \ell^\infty(A)$ tale che $T(x_i) \neq 0$. Possiamo supporre (eventualmente normalizzando) che $\|x_i\|_\infty = 1$ per ogni $i \in I$.
Dunque esiste $n \in \mathbb{N}$ per cui l'insieme

$$I_n = \{i \in I : \ (T(x_i))_n \neq 0\}$$

(con $(T(x_i))_n$ chiaramente intendiamo l'$n-$esima componente di $T(x_i) \in \ell^\infty$) è più che numerabile. Infatti se per ogni $n \in \mathbb{N}$ l'insieme I_n fosse al più numerabile, allora anche l'unione

$$\bigcup_{n\in\mathbb{N}} I_n = \{i \in I : \quad \text{esiste } n \text{ per cui } \ (T(x_i))_n \neq 0\} = I$$

risulterebbe al più numerabile, che è assurdo per l'ipotesi su I.
Fissato questo n, allo stesso modo esiste $k \in \mathbb{N}$ tale che

$$I_{n,k} = \left\{ i \in I : \ \left|\left(T(x_i)\right)_n\right| \geqslant \frac{1}{k} \right\}$$

è più che numerabile. Fissiamo questo $k \in \mathbb{N}$ (anche n è fissato). Per ogni $i \in I_{n,k}$ sia α_i con $|\alpha_i| = 1$ tale che $\alpha_i \left(T(x_i)\right)_n = \left|\left(T(x_i)\right)_n\right|$.
Preso un sottoinsieme finito $F \subset I_{n,k}$, consideriamo la successione y data da

$$y = \sum_{i \in F} \alpha_i x_i.$$

Poiché, per il lemma 3.10, l'intersezione dei supporti di due diversi x_i è finita, possiamo decomporre y nella forma

$$y = u + v,$$

dove v ha supporto finito e quindi è una successione infinitesima, mentre il rimanente u ha la proprietà $\|u\|_\infty \leqslant 1$, visto che per ogni n la sua componente $n-$esima è del tipo $\alpha_j (x_j)_n$ (ovvero non presenta somme, perché tali somme, presenti nelle componenti di y in corrispondenza dell'intersezione dei supporti, vengono opportunamente scaricate su v) e certamente $|\alpha_j (x_j)_n| = |(x_j)_n| \leqslant \|x_j\|_\infty = 1$. Dato che per ipotesi T è lineare e nulla sulle successioni infinitesime vale

$$\|T(y)\|_\infty = \|T(u+v)\|_\infty = \|T(u) + T(v)\|_\infty \leqslant \|T(u)\|_\infty + \|T(v)\|_\infty =$$

$$= \|T(u)\|_\infty = \frac{\|T(u)\|_\infty}{\|u\|_\infty} \|u\|_\infty \leqslant \|T\|_{\mathcal{L}(\ell^\infty, \ell^\infty)}.$$

Di conseguenza, ricordando che $F \subset I_{n,k} = \left\{ i \in I : \ \left|\left(T(x_i)\right)_n\right| \geqslant \frac{1}{k} \right\}$, abbiamo

$$\sum_{i \in F} k^{-1} \leqslant \sum_{i \in F} \left|\left(T(x_i)\right)_n\right| = \sum_{i \in F} \alpha_i \left(T(x_i)\right)_n = \left(T(y)\right)_n \leqslant \|T(y)\|_\infty \leqslant \|T\|_{\mathcal{L}(\ell^\infty, \ell^\infty)},$$

cioè $m \leqslant k \|T\|_{\mathcal{L}(\ell^\infty, \ell^\infty)}$, avendo indicato con m la cardinalità dell'insieme finito F. Siccome questa relazione è valida per un sottoinsieme finito di $I_{n,k}$ di cardinalità arbitraria (la cardinalità è arbitraria perché $I_{n,k}$ è un insieme di indici infinito, addirittura più che numerabile) otteniamo $\|T\|_{\mathcal{L}(\ell^\infty, \ell^\infty)} = \infty$, che è assurdo per l'ipotesi di limitatezza di T: da cui la tesi. $\square$

Proposizione 3.12. *L'identità $Id : c_0 \longrightarrow c_0$ non può essere estesa a un operatore lineare e continuo $P : \ell^\infty \longrightarrow c_0$.*

Dimostrazione. Supponiamo per assurdo esista l'estensione $P \in \mathcal{L}(\ell^\infty, c_0)$ dell'identità su c_0. Indichiamo con I l'identità su ℓ^∞.
Possiamo applicare allora il lemma 3.11 all'operatore $(I - P) \in \mathcal{L}(\ell^\infty, c_0)$, dato che per costruzione $(I - P)\,|_{c_0} = 0$. Di conseguenza per tale lemma esiste un sottoinsieme infinito $A \subseteq \mathbb{N}$ tale che $(I - P)\,|_{\ell^\infty(A)} = 0$, ossia

$$P(x) = x \qquad \forall\ x \in \ell^\infty(A).$$

Siccome P estende l'identità su c_0, ciò significa che

$$\ell^\infty(A) \subset c_0.$$

Questo è assurdo: infatti la successione $\{x_n\}_{n\in\mathbb{N}}$ definita da

$$x_n = \begin{cases} 0 & \text{se } n \notin A \\ 1 & \text{se } n \in A \end{cases}$$

è una successione di $\ell^\infty(A)$ ma non può essere infinitesima, visto che A è infinito, perché per ogni $m \in \mathbb{N}$ esiste un indice $n \geqslant m$ tale che $n \in A$ cioè $x_n = 1$.
Pertanto $Id : c_0 \longrightarrow c_0$ non può essere estesa a un funzionale lineare e continuo definito su tutto ℓ^∞ e a valori in c_0. □

3.3 Spazi isometricamente iniettivi e generalizzazione del teorema

Nell'ambito della auspicabile generalizzazione del teorema di HB risulta naturale introdurre la seguente definizione.

Definizione 3.13. Uno spazio di Banach Z si dice *iniettivo* se per ogni spazio di Banach X, per ogni sottospazio chiuso Y di X e per ogni $g \in \mathcal{L}(Y, Z)$, esiste una estensione $f \in \mathcal{L}(X, Z)$ di g.
Inoltre uno spazio iniettivo Z si dice *isometricamente iniettivo* se per di più esiste un'estensione $f \in \mathcal{L}(X, Z)$ di g che soddisfa $\|f\|_{\mathcal{L}(X,Z)} = \|g\|_{\mathcal{L}(Y,Z)}$.

Osservazione 3.14. Per definizione uno spazio è isometricamente iniettivo se per esso vale la generalizzazione del teorema di Hahn-Banach.
Nel paragrafo precedente abbiamo provato (proposizioni 3.8 e 3.9) che gli spazi $\mathbb{R}^N$ e ℓ^∞, entrambi considerati con norma infinito, sono isometricamente iniettivi, e che invece c_0 non è neppure iniettivo (vedi proposizione 3.12).

Il nostro intento è quello di esibire altri spazi isometricamente iniettivi.
Riusciremo infine ad ottenere una caratterizzazione di questi spazi.

Detto K uno spazio compatto, indichiamo con $\mathcal{C}(K)$ lo spazio delle funzioni continue a valori reali definite sul compatto K. È uno spazio di Banach con la norma

$$\|f\|_{\mathcal{C}(K)} = \sup_{x \in K} |f(x)| = \max_{x \in K} |f(x)|$$

(tale massimo esiste perché l'immagine del compatto K tramite f, per continuità, è un compatto di $\mathbb{R}$, dunque chiuso e limitato).

Gli spazi su cui si articola questa teoria costituiscono una particolare sottofamiglia dei $\mathcal{C}(K)$. Vediamo qual è la proprietà di cui abbiamo bisogno.

Definizione 3.15. Diciamo che uno spazio $\mathcal{C}(K)$ possiede *ordine completo* se per ogni coppia di suoi sottoinsiemi non vuoti $A, B \subseteq \mathcal{C}(K)$ tali che $f \leqslant g$ per ogni $f \in A$ e per ogni $g \in B$, esiste una funzione $h \in \mathcal{C}(K)$ che soddisfa

$$f \leqslant h \leqslant g \qquad \forall\, f \in A,\ \forall\, g \in B.$$

Osservazione 3.16. (1) Notiamo che uno spazio $\mathcal{C}(K)$ possiede ordine completo se e solo se per esso vale l'analogo dell'assioma di completezza per $\mathbb{R}$. Questo assicura l'esistenza dell'estremo superiore, ossia il minimo dei maggioranti, per ogni sottoinsieme non vuoto A di $\mathcal{C}(K)$ che ammetta maggiorante. Infatti sia B l'insieme non vuoto dei maggioranti di A; allora in uno spazio $\mathcal{C}(K)$ avente ordine completo è univocamente determinata una funzione $h \in \mathcal{C}(K)$ con la proprietà $f \leqslant h \leqslant g \quad \forall\, f \in A,\ \forall\, g \in B$. Di conseguenza h è il minimo dei maggioranti, che viene indicato con $\sup A$.
Osserviamo che per definizione $h = \sup A$ è una funzione continua, quindi può non coincidere con l'estremo superiore puntuale $\widehat{h}(s) = \sup_{f\in A} f(s)$, che può non essere una funzione continua.
Allo stesso modo esiste anche l'estremo inferiore $\inf A$ per ogni sottoinsieme non vuoto A di $\mathcal{C}(K)$ che ammetta minorante.
Nella ricerca di spazi isometricamenti iniettivi l'esistenza di estremo superiore e inferiore è un ingrediente essenziale: per questo è importante la definizione data.
(2) La definizione data può essere estesa in modo naturale ad ogni spazio con adeguata struttura di ordine come per esempio ℓ^∞, che possiede ordine completo.
In realtà ℓ^∞ è uno spazio del tipo $\mathcal{C}(K)$, dato che può essere identificato con $\mathcal{C}(\beta\mathbb{N})$, che quindi possiede automaticamente ordine completo; ricordiamo che $\beta\mathbb{N}$ è la compattificazione di Stone-Čech di $\mathbb{N}$ dotato della topologia discreta, ossia $\beta\mathbb{N}$ è l'unico spazio compatto e di Hausdorff, contenente $\mathbb{N}$ come sottospazio denso, tale che ogni funzione limitata e continua su $\mathbb{N}$ si estende a una funzione continua su $\beta\mathbb{N}$.
(3) Gli spazi compatti K tali che il corrispondente spazio $\mathcal{C}(K)$ possieda ordine completo sono caratterizzati dalla seguente proprietà topologica: la chiusura di ogni insieme aperto è aperta; spazi K di questo tipo sono detti estremamente disconnessi.

Sia Y un sottospazio di uno spazio di Banach X.
Riprendendo la definizione 1.19, una mappa $V : Y \longrightarrow \mathcal{C}(K)$ è subadditiva e positivamente omogenea se soddisfa le due relazioni

$$V(x+y) \leqslant V(x) + V(y) \qquad \forall x, y \in Y,$$

$$V(ty) = tV(y) \qquad \forall y \in Y, \quad \forall t \geqslant 0.$$

Una mappa subadditiva e positiva omogenea $V : X \longrightarrow \mathcal{C}(K)$ è minimale se non esiste una mappa subadditiva e positivamente omogenea $U : X \longrightarrow \mathcal{C}(K)$, diversa da V, per cui $U(x) \leqslant V(x)$ per ogni $x \in X$.

Lemma 3.17. *Siano X uno spazio di Banach e Y un suo sottospazio. Supponiamo che $V : X \longrightarrow \mathcal{C}(K)$ e $W : Y \longrightarrow \mathcal{C}(K)$ siano mappe subadditive e positivamente omogenee, tali che*

$$W(y) + V(-y) \geqslant 0 \qquad \forall\, y \in Y.$$

Se $\mathcal{C}(K)$ possiede ordine completo allora la mappa $V \wedge W : X \longrightarrow \mathcal{C}(K)$ data da

$$(V \wedge W)(x) = \inf\{V(x-y) + W(y) : \ y \in Y\}$$

è subadditiva e positivamente omogenea.

Dimostrazione. Per ogni fissato $x \in X$ per ipotesi abbiamo

$$V(x-y) + W(y) \geqslant V(-y) - V(-x) + W(y) \geqslant -V(-x) \qquad \forall\, y \in Y,$$

quindi $-V(-x)$ è un minorante per l'insieme $\{V(x-y) + W(y) : \ y \in Y\} \subset \mathcal{C}(K)$, che ammette estremo inferiore, per l'osservazione 3.16, dato che $\mathcal{C}(K)$ possiede ordine completo. È allora ben definita la mappa

$$(V \wedge W)(x) = \inf\{V(x-y) + W(y) : \ y \in Y\}.$$

Si verica facilmente che $V \wedge W$ è subadditiva e positivamente omogenea. □

Osserviamo che per definizione, valutando $V(x-y) + W(y)$ in $y = 0 \in Y$, vale

$$(V \wedge W)(x) \leqslant V(x) \quad \text{per ogni } x \in X.$$

Lemma 3.18. *Siano X uno spazio di Banach e $V : X \longrightarrow \mathcal{C}(K)$ subadditiva e positivamente omogenea. Se $\mathcal{C}(K)$ possiede ordine completo allora esiste una mappa subadditiva e positivamente omogenea minimale $W : X \longrightarrow \mathcal{C}(K)$.*
In particolare vale $W(x) \leqslant V(x)$ per ogni $x \in X$.

Dimostrazione. Vogliamo utilizzare il lemma di Zorn sull'insieme

$$\Sigma = \{U : X \longrightarrow \mathcal{C}(K) : \ U \text{ è subadditiva, posit. omogenea e } U \leqslant V\},$$

che è ovviamente non vuoto ($V \in \Sigma$) e parzialmente ordinato.
Se ogni catena $\Phi = (U_i)_{i \in I}$ in Σ ammette minorante in Σ, possiamo concludere per il lemma di Zorn che esiste un elemento minimale in Σ, ovvero la tesi.
Dobbiamo quindi fare questa verifica. Sia $\Phi = (U_i)_{i \in I}$ una catena in Σ. Osserviamo che per ogni $i \in I$ vale per subadditività $0 \leqslant U_i(x + (-x)) \leqslant U_i(x) + U_i(-x)$ per tutti gli $x \in X$, quindi, essendo $(U_i)_{i \in I} \subset \Sigma$,

$$U_i(x) \geqslant -U_i(-x) \geqslant -V(-x).$$

Questo significa che per ogni $x \in X$ l'insieme $\{U_i(x) : \ i \in I\} \subset \mathcal{C}(K)$ ha un minorante e di conseguenza, siccome $\mathcal{C}(K)$ possiede ordine completo, è ben definita la mappa

$$U_\Phi(x) = \inf_{i \in I} U_i(x).$$

Proviamo che U_Φ, il minorante della catena, appartiene a Σ. Dato che $U_i \in \Sigma$ per ogni $i \in I$, è ovvio che U_Φ è positivamente omogenea e $U_\Phi \leqslant V$.
Resta da verificare la subadditività di U_Φ. Possiamo assumere, senza perdita di generalità, che dati $i \neq j \in I$ valga $U_i \leqslant U_j$. Allora per ogni $x, y \in X$ abbiamo

$$U_\Phi(x+y) \leqslant U_i(x+y) \leqslant U_i(x) + U_i(y) \leqslant U_j(x) + U_i(y),$$

da cui $U_\Phi(x+y) \leqslant U_j(x) + U_\Phi(y)$. Passando ancora all'estremo inferiore otteniamo

$$U_\Phi(x+y) \leqslant U_\Phi(x) + U_\Phi(y).$$

□

Lemma 3.19. *Dati uno spazio di Banach X e un spazio $\mathcal{C}(K)$ avente ordine completo, sia $V : X \longrightarrow \mathcal{C}(K)$ subadditiva e positivamente omogenea. Se V è minimale, allora V è lineare.*

Dimostrazione. Fissato $x \in X$, sia $Y = \langle x \rangle$ il sottospazio generato da x.
La mappa $W : Y \longrightarrow \mathcal{C}(K)$ definita da

$$W(\lambda x) = -\lambda V(-x)$$

è chiaramente lineare. Inoltre per ogni $\lambda \in \mathbb{R}$ vale $W(\lambda x) \geqslant -\lambda V(-x)$; infatti
se $\lambda \geqslant 0 : \ W(\lambda x) = -\lambda V(-x) = -V(-\lambda x)$;
se $\lambda < 0 : \ W(\lambda x) = -\lambda V(-x) = V(\lambda x) \geqslant -V(-\lambda x)$
(perché per subadditività di V si ha $V(\lambda x) + V(-\lambda x) \geqslant V(0) = 0$).
Pertanto possiamo applicare il lemma 3.17, che definisce su X la mappa subadditiva e positivamente omogenea

$$\left(V \wedge W\right)(x) = \inf_{\lambda \in \mathbb{R}} \{V(x - \lambda x) + W(\lambda x)\} \leqslant V(x).$$

Per minimalità di V, $(V \wedge W) = V$: in particolare $V|_Y = (V \wedge W)\,|_Y \leqslant W$
(basta sostituire $\lambda = 1$ nella formula che definisce $(V \wedge W)$).
Di conseguenza vale $V(x) \leqslant W(x) = -V(-x)$, e per subadditività anche $V(x) \geqslant -V(-x)$. Allora $V(x) = -V(-x)$. Dato che x fissato inizialmente è arbitrario, l'ultima relazione vale per ogni $x \in X$. Ne segue che V è lineare; infatti è omogenea e, sfruttando la subadditività, abbiamo per ogni $x, y \in X$

$$V(x) + V(y) = -V(-x) - V(-y) \leqslant -V(-x-y) = V(x+y) \leqslant V(x) + V(y).$$

□

Teorema 3.20. *(Goodner, Nachbin)*
Sia K uno spazio compatto e di Hausdorff. Allora $\mathcal{C}(K)$ è isometricamente iniettivo se e solo se $\mathcal{C}(K)$ possiede ordine completo.

Dimostrazione. Supponiamo che $\mathcal{C}(K)$ possieda ordine completo.
Sia Y un sottospazio chiuso di uno spazio di Banach X. Fissiamo un operatore $S \in \mathcal{L}(Y, \mathcal{C}(K))$, che possiamo senz'altro supporre di norma unitaria.
Vogliamo estendere in modo lineare e continuo l'operatore S a tutto X, preservandone la norma unitaria.
Poiché $\|S\|_{\mathcal{L}(Y,\mathcal{C}(K))} = 1$, è immediato che per ogni $x \in Y$ valga

$$-\|x\| \leqslant (S(x))\,(k) \leqslant \|x\| \qquad \forall\, k \in K,$$

o equivalentemente, indicata con 1 la funzione costante 1 definita su K,

$$-\|x\| \cdot 1 \leqslant S(x) \leqslant \|x\| \cdot 1.$$

Questo significa che $S(x) \geqslant -V_0(-x)$ per ogni $x \in Y$, dove $V_0 : X \longrightarrow \mathcal{C}(K)$ è la mappa subadditiva e positivamente omogenea data da $V_0(x) = \|x\| \cdot 1$.
Siamo nelle ipotesi del lemma 3.17, che ci permette di definire la mappa subadditiva e positivamente omogenea $V = (V_0 \wedge S) : X \longrightarrow \mathcal{C}(K)$: data da

$$V(x) = \inf \{V_0(x-y) + S(y) : \; y \in Y\}.$$

Il lemma 3.18 ci garantisce l'esistenza di una mappa $T : X \longrightarrow \mathcal{C}(K)$ subadditiva e positivamente omogena minimale, che soddisfa per definizione

$$T \leqslant V \quad \text{e} \quad T|_Y \leqslant S,$$

ed è addirittura lineare per il lemma 3.19.
Per ogni $x \in Y$ abbiamo allora $T(x) \leqslant S(x)$ e $-T(x) = T(-x) \leqslant S(-x) = -S(-x)$, quindi $T|_Y = S$, cioè T è l'estensione lineare di S a tutto lo spazio X.
Verifichiamo che la norma di T è ancora unitaria. Da una parte, essendo T una estensione di S, è ovvio che $\|T\|_{\mathcal{L}(X,\mathcal{C}(K))} \geqslant \|S\|_{\mathcal{L}(Y,\mathcal{C}(K))} = 1$.
Dall'altra $T \leqslant V \leqslant V_0$, che implica per ogni $x \in X$

$$T(x) \leqslant V_0(x) = \|x\| \cdot 1 \quad \text{e} \quad T(-x) \leqslant V_0(-x) = \|-x\| \cdot 1 = \|x\| \cdot 1.$$

Dunque $\|T\|_{\mathcal{L}(X,\mathcal{C}(K))} \leqslant 1$, anzi per quanto detto vale l'uguaglianza.

Viceversa, supponiamo che lo spazio $\mathcal{C}(K)$ sia isometricamente iniettivo.
Sia $\ell^\infty(K)$ lo spazio delle funzioni limitate a valori reali definite sul compatto K, che è di Banach con la stessa norma $\|.\|_{\mathcal{C}(K)}$: chiaramente $\mathcal{C}(K)$ è sottospazio chiuso di $\ell^\infty(K)$. Per l'ipotesi su $\mathcal{C}(K)$ sappiamo che l'identità $Id : \mathcal{C}(K) \longrightarrow \mathcal{C}(K)$ si estende ad un operatore lineare e continuo

$$P : \ell^\infty(K) \longrightarrow \mathcal{C}(K) \quad \text{che soddisfa} \quad \|P\|_{\mathcal{L}(\ell^\infty(K),\mathcal{C}(K))} = \|Id\|_{\mathcal{L}(\mathcal{C}(K),\mathcal{C}(K))} = 1.$$

Dati due sottoinsiemi non vuoti A e B di $\mathcal{C}(K)$ con la proprietà $f \leqslant g$ per ogni $f \in A$ e per ogni $g \in B$, vogliamo provare che esiste una opportuna funzione $h \in \mathcal{C}(K)$ che separa i due sottoinsiemi, sarebbe a dire

$$f \leqslant h \leqslant g \qquad \forall\ f \in A,\ \forall\ g \in B.$$

Indichiamo con a l'estremo superiore in senso puntuale delle funzioni f di A, ossia

$$a(s) = \sup_{f \in A} f(s) \qquad \forall\ s \in K.$$

La funzione a così definita è limitata, cioè $a \in \ell^\infty(K)$, dato che per definizione $f \leqslant a \leqslant g \quad \forall\ f \in A,\ \forall\ g \in B$, e tutte le f, g sono ovviamente limitate.
Se a fosse continua, avremmo finito. In generale però a non è continua.
Mostriamo che in ogni caso esiste un funzione continua che separa A da B.
La funzione separatrice è

$$h = P(a) \in \mathcal{C}(K),$$

ovviamente continua per definizione di P.
Per concludere dobbiamo verificare l'effettiva proprietà di separazione di h. A questo scopo, fissiamo $\lambda \in \mathbb{R}$ e una funzione positiva b appartenente a $\ell^\infty(K)$ (con funzione positiva b intendiamo che b è a valori non negativi e non è identicamente nulla).
Indicata con 1 la funzione costante 1 definita su K, vale

$$\|1 - \lambda b\|_{\mathcal{C}(K)} \leqslant 1 \quad \text{se e solo se} \quad 0 \leqslant \lambda \leqslant \frac{2}{\|b\|_{\mathcal{C}(K)}} :$$

infatti $\|1 - \lambda b\| \leqslant 1 \Leftrightarrow -1 \leqslant 1 - \lambda b(s) \leqslant 1$ per ogni $s \in K$, ossia $0 \leqslant \lambda b(s) \leqslant 2$ per ogni $s \in K$, e dividendo per il numero positivo $b(s)$ otteniamo $0 \leqslant \lambda \leqslant \frac{2}{b(s)}$ per ogni $s \in K$ tale che $b(s) > 0$, che è equivalente a $0 \leqslant \lambda \leqslant \frac{2}{\|b\|}$.
Dato che P ha norma unitaria e $P(1) = 1$ (ovvio, essendo $P|_{\mathcal{C}(K)} = Id$), allora per ogni b positiva appartenente a $\ell^\infty(K)$ abbiamo

$$\|P(1 - \lambda b)\|_{\mathcal{C}(K)} \leqslant 1 \qquad \text{per ogni} \quad 0 \leqslant \lambda \leqslant \frac{2}{\|b\|_{\mathcal{C}(K)}}.$$

Quest'ultima relazione deriva dalla precedente perché

$$\|P(1-\lambda b)\| \leqslant \|1-\lambda b\| \, \|P\| = \|1-\lambda b\| \leqslant 1 \qquad \text{per ogni} \quad 0 \leqslant \lambda \leqslant \frac{2}{\|b\|_{\mathcal{C}(K)}}.$$

Proviamo che la mappa P deve essere positiva, nel senso che $b \geqslant 0$ implica $P(b) \geqslant 0$.
Dalla positività di P discende poi la sua monotonia, perché se $b_1 \geqslant b_2$ applicando la mappa P lineare e positiva otteniamo banalmente $P(b_1) \geqslant P(b_2)$.
Ovviamente, per linearità di P, se b è la funzione nulla $P(b) = P(0) = 0$.
Sia allora b limitata positiva. Supponiamo per assurdo che esista $s \in K$ per cui $P(b)(s) < 0$: per tale s e un qualunque $\lambda > 0$ risulta

$$P(1-\lambda b)(s) = 1 - \lambda P(b)(s) > 1 \quad \text{quindi} \quad \|P(1-\lambda b)\| > 1,$$

che è assurdo scegliendo $0 < \lambda \leqslant \frac{2}{\|b\|_{\mathcal{C}(K)}}$.
Ne consegue positività e monotonia per l'operatore P.
Proprio dalla monotonia, e dal fatto che P estende l'identità su $\mathcal{C}(K)$, possiamo infine dedurre la proprietà di separazione di $h = P(a)$.
Infatti $f \leqslant a$ per ogni $f \in A$, da cui $f = P(f) \leqslant P(a) = h$. Allo stesso modo $a \leqslant g$ per ogni $g \in B$, da cui $h = P(a) \leqslant P(g) = g$. Dunque abbiamo provato la tesi. □

Abbiamo provato che gli spazi $\mathcal{C}(K)$ aventi ordine completo sono isometricamente iniettivi. È naturale chiederci se esistano spazi isometricamente iniettivi che non siano del tipo $\mathcal{C}(K)$.

Ricordiamo che un'applicazione lineare $T : X \longrightarrow Z$ tra due spazi normati è un isomorfismo se è continua, biunivoca e la sua inversa è continua.
Un'applicazione lineare T che preserva la norma, ovvero $\|T(x)\|_Z = \|x\|_X$ per ogni $x \in X$, si chiama isometria. Chiaramente, un'isometria $T \in \mathcal{L}(X,Y)$ è iniettiva (se $T(x_0) = T(x_1)$, per linearità si ha $T(x_0 - x_1) = 0$; pertanto, essendo T un'isometria $\|x_0 - x_1\| = \|0\| = 0$, che implica $x_0 = x_1$) e ha norma unitaria in $\mathcal{L}(X,Y)$.
Se un isomorfismo $T : X \longrightarrow Z$ è pure un'isometria, T si dice un isomorfismo isometrico. In tal caso X e Z si dicono isometricamente isomorfi.

Il seguente e definitivo risultato sancisce che ogni spazio isometricamente iniettivo è della forma $\mathcal{C}(K)$, dove quest'ultimo deve avere ordine completo per il teorema precedente.

Teorema 3.21. *(Kelley)*
Uno spazio di Banach è isometricamente iniettivo se e solo se è isometricamente isomorfo ad uno spazio $\mathcal{C}(K)$ avente ordine completo.

Per la dimostrazione di questo teorema conclusivo, abbiamo bisogno di richiamare vari concetti. Stiamo parlando della topologia debole*, della convessità, della nozione di operatore aggiunto. Enunciamo alcuni importanti risultati collegati a questi ambiti ed un teorema significativo sul duale degli spazi $\mathcal{C}(K)$: questi saranno necessari nel corso della dimostrazione, piuttosto elaborata, del teorema di Kelley.

Siano X uno spazio normato e X^* il suo duale. Fissato un elemento $x \in X$ definiamo il funzionale lineare $x^{**} : X^* \longrightarrow \mathbb{R}$ (quindi x^{**} appartiene al biduale di X) dato da

$$x^{**}(f) = f(x) \qquad \forall f \in X^*.$$

La topologia debole* su X^* si definisce come la topologia meno fine su X^* che rende continui tutti gli elementi del biduale del tipo x^{**}, al variare di $x \in X$. Osserviamo che la topologia debole* è localmente convessa e di Hausdorff. Il principale risultato riguardante la topologia debole* è rappresentato dal teorema di Banach-Alaoglu: se X è uno spazio normato, allora $B_{X^*} = \{x^* \in X^* : \ \|x^*\|_{X^*} \leqslant 1\}$ è un compatto nella topologia debole*.

Un sottoinsieme S di uno spazio vettoriale X è detto convesso se per ogni $x, y \in S$ e per ogni $t \in [0,1]$ vale $tx + (1-t)y \ \in S$.
Dato $S \subseteq X$, si definisce inviluppo convesso di S, e si indica con $co(S)$, il più piccolo insieme convesso che contiene S (ovviamente un tale insieme esiste sempre, in quanto X è convesso e l'intersezione arbitraria di convessi è convessa). L'inviluppo convesso di S può essere descritto analiticamente da

$$co(S) = \left\{ \sum_{i=1}^{n} t_i x_i : \quad \{x_i\}_{i=1}^{n} \subseteq S, \ t_i \geqslant 0, \ \sum_{i=1}^{n} t_i = 1, \ n \in \mathbb{N}_0 \right\}.$$

Se X è dotato di un topologia τ, indichiamo con $\overline{co}^{\tau}(S)$ l'inviluppo convesso chiuso di S, ovvero il più piccolo $\tau-$chiuso convesso contenente S. Negli spazi vettoriali topologici vale $\overline{co}^{\tau}(S) = \overline{co(S)}^{\tau}$ (in generale vale soltanto la disuguaglianza $\subseteq$).
Sia S un insieme convesso. Un punto $x \in S$ è detto punto estremo di S se l'insieme $S \setminus \{x\}$ è convesso; in altre parole se x è un punto estremo di S e si ha $x = tx_1 + (1-t)x_2$, con $t \in (0,1)$, allora $x = x_1 = x_2$. Denotiamo con $\partial_e S$ l'insieme (eventualmente vuoto) dei punti estremi di S. Vediamo ora due importanti teoremi che mettono in relazione punti estremi e inviluppo convesso chiuso.
Teorema di Krein-Milman: sia X uno spazio vettoriale topologico localmente convesso; se $K \subseteq X$ è un compatto convesso (non vuoto) allora K è l'inviluppo convesso chiuso dei suoi punti estremi, in particolare K ammette punti estremi.
Teorema di Milman: siano X uno spazio vettoriale topologico localmente convesso e K un suo sottoinsieme compatto e chiuso, allora ogni punto estremo di $\overline{co}(K)$ appartiene a K, in simboli $\partial_e\overline{co}(K) \subseteq K$.

Dati due spazi normati X e Y, sia $T : X \longrightarrow Y$ un operatore lineare e limitato. L'operatore lineare $T^* : Y^* \longrightarrow X^*$ dato da

$$T^*(\varphi) = \varphi \circ T \qquad \forall \varphi \in Y^*,$$

si chiama operatore aggiunto di T; esso è caratterizzato dalla relazione

$$\left(T^*(\varphi)\right)(x) = \varphi(T(x)) \qquad \forall x \in X,\ \forall \varphi \in Y^*.$$

Inoltre per l'operatore aggiunto vale

$$\|T^*\|_{\mathcal{L}(Y^*,X^*)} = \|T\|_{\mathcal{L}(X,Y)}\,.$$

Prima di passare alla dimostrazione del teorema 3.21, dobbiamo citare un importante teorema che fornisce un isomorfismo isometrico per il duale degli spazi $\mathcal{C}(K)$.
Teorema di Riesz-Markov-Kakutani: sia K uno spazio compatto e di Hausdorff, allora $\mathcal{C}(K)^*$ è isometricamente isomorfo allo spazio $\mathcal{M}(K)$ di tutte le misure regolari con segno, finite e boreliane su K, con la norma $\|\mu\|_{\mathcal{M}(K)} = |\mu|(K)$; l'isomorfismo è dato da

$$\mu(f) = \int_K f \ d\mu \qquad \forall f \in \mathcal{C}(K).$$

Dimostrazione. (Kelley)
Per il teorema 3.20, se X è uno spazio di Banach isometricamente isomorfo ad uno spazio $\mathcal{C}(K)$ avente ordine completo, allora X è isometricamente iniettivo.
Dobbiamo provare l'implicazione contraria. Fissato uno spazio di Banach X isometricamente iniettivo, il nostro obiettivo è quello di identificare X (tramite un isomorfismo isometrico) con un opportuno spazio $\mathcal{C}(K)$; così facendo, $\mathcal{C}(K)$ risulterà isometricamente iniettivo, perciò grazie al teorema 3.20 avrà ordine completo e avremmo la tesi.
Il compatto K, come vedremo, sarà un particolare sottoinsieme della palla unitaria chiusa $B_{X^*} = \{x^* \in X^* : \ \|x^*\|_{X^*} \leqslant 1\}$ dello spazio duale X^*.
Per il teorema di Banach-Alaoglu, B_{X^*} è compatta nella topologia debole*.
Consideriamo l'insieme $\partial_e B_{X^*}$ dei punti estremi di B_{X^*} rispetto alla topologia debole* (osserviamo che $\partial_e B_{X^*}$ è non vuoto, come conseguenza del teorema di Krein-Milman, dato che la topologia debole* è localmente convessa e B_{X^*} è un compatto in tale topologia). Esiste un sottoinsieme aperto U di $\partial_e B_{X^*}$, massimale tra tutti gli aperti di $\partial_e B_{X^*}$ aventi la proprietà $U \cap (-U) = \emptyset$. Si tratta di un'applicazione del lemma di Zorn; infatti, posto $\Sigma = \{U \text{ aperto } \subseteq \partial_e B_{X^*} : \ U \cap (-U) = \emptyset\}$, ovviamente Σ non è vuoto perché contiene l'insieme vuoto, e ogni catena di aperti siffatti ammette come maggiorante in Σ la loro unione. Quindi esiste U elemento massimale in Σ.
Sia K la chiusura debole* di U in B_{X^*}: K, rispetto a questa topologia, è certamente compatto (essendo un chiuso contenuto nel compatto B_{X^*}) e di Hausdorff (dato che la topologia debole* è di Hausdorff). Osserviamo che, per costruzione,

$$(K \cap \partial_e B_{X^*}) \cap (-U) = \emptyset.$$

Infatti, supponiamo esista $x \in [K \cap \partial_e B_{X^*} \cap (-U)] = \left[\overline{U}^* \cap \partial_e B_{X^*} \cap (-U)\right]$. Poiché $\partial_e B_{X^*} \cap (-U)$ è aperto in $\partial_e B_{X^*}$, esiste un intorno V di x, aperto in $\partial_e B_{X^*}$, tale che $x \in V \subseteq [\partial_e B_{X^*} \cap (-U)]$. Allora risulta $V \subseteq (-U)$ e, essendo $x \in \overline{U}^*$, $V \cap U \neq \emptyset$. Ciò è assurdo perché $\emptyset \neq (V \cap U) \subseteq [(-U) \cap U] = \emptyset$.
Di conseguenza $(K \cap \partial_e B_{X^*}) \cap (-U) = \emptyset$, da cui segue $K \cap (-U) = \emptyset$.
Proviamo che

$$\partial_e B_{X^*} \subseteq (K \cup (-K)) .$$

Infatti, supponiamo esista un elemento $x^* \in \partial_e B_{X^*} \setminus (K \cup (-K))$. Per chiusura di $(K \cup (-K))$, esiste un intorno aperto e convesso di 0 (per la topologia debole*), che indichiamo con V e che possiamo supporre bilanciato (ossia $\alpha V \subseteq V$ per ogni $|\alpha| \leqslant 1$) e tale che

$$x^* \notin V \quad \text{ e } \quad (x^* + V) \cap (K \cup (-K)) = \emptyset.$$

Poniamo ora

$$U_1 = U \cup [(x^* + V) \cap \partial_e B_{X^*}]$$

osservando che U_1 è un aperto che contiene strettamente U, dato che $x^* \notin K = \overline{U}^*$. Mostriamo che $U_1 \cap (-U_1) = \emptyset$, che è assurdo per massimalità di U. Supponiamo $y^* \in U_1 \cap (-U_1)$, allora $y^* \notin U$ oppure $-y^* \notin U$ (altrimenti U e $-U$ non sarebbero disgiunti); senza perdita di generalità diciamo che $y^* \notin U$ (se necessario basta scambiare y^* con $-y^*$). Pertanto $y^* \in U_1 \setminus U$ e ciò significa $y^* \in (x^* + V)$. Sappiamo che quest'ultimo insieme non interseca $(K \cup (-K))$, quindi $y^* \notin (K \cup (-K))$, che implica $y^* \notin (-U)$ (essendo $-K = \overline{(-U)}^*$). Perciò vale anche $y^* \in -U_1 \setminus -U$, da cui $y^* \in (-x^* - V)$. Riepilogando,

$$y^* \in [(x^* + V) \cap (-x^* - V)] \, ,$$

ovvero $0 \in (2x^* + 2V)$, che implica $x^* \in V$: assurdo per le ipotesi su V.
Allora $U_1 \cap (-U_1) = \emptyset$, in contraddizione con la massimalità di U. Abbiamo dunque provato che

$$\partial_e B_{X^*} \subseteq (K \cup (-K)) \, .$$

Di conseguenza abbiamo $\overline{co}^* (\partial_e B_{X^*}) \subseteq \overline{co}^* (K \cup (-K))$; d'altra parte, per il teorema di Krein-Milman vale $\overline{co}^* (\partial_e B_{X^*}) = B_{X^*}$. Allora $B_{X^*} \subseteq \overline{co}^* (K \cup (-K))$. Anzi

$$B_{X^*} = \overline{co}^* (K \cup (-K)) \, ,$$

dato che B_{X^*} è un convesso chiuso (debole*) contenente $(K \cup (-K))$.
Fissiamo $x \in X$. Ovviamente, essendo K un compatto contenuto in B_{X^*}, vale la disuguaglianza $\sup_{x^* \in B_{X^*}} |x^*(x)| \geqslant \max_{x^* \in K} |x^*(x)|$. Grazie al fatto che B_{X^*} è l'inviluppo convesso chiuso di $(K \cup (-K))$, vale anche la disuguaglianza opposta e possiamo allora affermare che

$$\sup_{x^* \in B_{X^*}} |x^*(x)| = \max_{x^* \in K} |x^*(x)| \, .$$

Mostriamo infatti che $|x^*(x)| \leqslant \max_{x^* \in K} |x^*(x)|$ per ogni $x^* \in B_{X^*}$. Fissato un qualunque $x^* \in B_{X^*}$, dato che $B_{X^*} = \overline{co}^* (K \cup (-K)) = \overline{co\,(K \cup (-K))}^*$, esiste una successione $\{x_n^*\}_{n \in \mathbb{N}}$ di elementi di $co\,(K \cup (-K))$ che converge a x^*, dove ogni x_n^* è della forma

$$x_n^* = \sum_{j=1}^{k_n} \lambda_{j,n} \, y_{j,n}^*, \quad \text{con} \quad \sum_{j=1}^{k_n} \lambda_{j,n} = 1, \quad \lambda_{j,n} \geqslant 0, \quad y_{j,n}^* \in (K \cup (-K)) \, .$$

Allora

$$|x^*(x)| = \left| \lim_{n \to \infty} x_n^*(x) \right| \leqslant \limsup_{n \to \infty} \left[\sum_{j=1}^{k_n} \lambda_{j,n} \, \left| y_{j,n}^*(x) \right| \right] \leqslant$$

$$\leqslant \limsup_{n\to\infty} \left(\max_{x^* \in K} |x^*(x)| \right) = \max_{x^* \in K} |x^*(x)| \, .$$

Inoltre

$$\sup_{x^* \in B_{X^*}} |x^*(x)| = \|x\|$$

perché se $x = 0$ allora $\|0\| = 0 = x^*(0)$ per ogni x^*, mentre per $x \neq 0$ si ha

$$\sup_{x^* \in B_{X^*}} |x^*(x)| = \|x\| \sup_{x^* \in B_{X^*}} \frac{|x^*(x)|}{\|x\|} \leqslant \|x\| \sup_{x^* \in B_{X^*}} \|x^*\|_{X^*} \leqslant \|x\|$$

(dove l'ultimo passaggio segue per definizione di B_{X^*}) e tale maggiorante per l'estremo superiore è effettivamente raggiunto dalla mappa $\varphi \in X^*$, data dalla proposizione 1.22, avente norma unitaria (quindi $\varphi \in B_{X^*}$) e tale che $\varphi(x) = \|x\|$.
Allora per ogni $x \in X$ si ha

$$\|x\| = \max_{x^* \in K} |x^*(x)| \, .$$

Sia $J : X \longrightarrow \mathcal{C}(K)$ la mappa lineare così definita:

$$J(x) = \widehat{x}, \qquad \text{dove} \qquad \widehat{x}(x^*) = x^*(x) \qquad \forall \; x^* \in K.$$

Osserviamo che, per definizione di $\widehat{x}$, risulta

$$\|x\| = \max_{x^* \in K} |x^*(x)| = \max_{x^* \in K} |\widehat{x}(x^*)| = \|\widehat{x}\|_{\mathcal{C}(K)} = \|J(x)\|_{\mathcal{C}(K)} \, ,$$

quindi J è un'isometria; in particolare è iniettiva e $\|J\|_{\mathcal{L}(X, \mathcal{C}(K))} = 1$.
Dunque $J : X \longrightarrow J(X)$ è un'isometria bigettiva; consideriamo l'inversa $J^{-1} : J(X) \longrightarrow X$, che è anche essa una mappa lineare ed un'isometria, in particolare ha norma unitaria.
Vogliamo provare che in realtà $J(X) = \mathcal{C}(K)$, ossia che $J : X \longrightarrow \mathcal{C}(K)$ è un isomorfismo isometrico: così avremo la tesi.
Per ipotesi sappiamo che X è uno spazio isometricamente iniettivo, allora esiste una estensione lineare e continua $T : \mathcal{C}(K) \longrightarrow X$ di J^{-1} che ha la stessa norma di J^{-1}; più precisamente

$$T(\widehat{x}) = x \qquad \forall \; x \in X, \qquad \text{con} \qquad \|T\|_{\mathcal{L}(\mathcal{C}(K), X)} = 1.$$

Sia $T^* : X^* \longrightarrow \mathcal{C}(K)^*$ l'operatore aggiunto di T, che possiamo interpretare come una mappa $T^* : X^* \longrightarrow \mathcal{M}(K)$, dato che, per il teorema di Riesz-Markov-Kakutani, $\mathcal{M}(K)$ è isometricamente isomorfo a $\mathcal{C}(K)^*$ tramite ($\lambda \in \mathcal{M}(K)$)

$$\lambda(f) = \int_K f(x) \; d\lambda \qquad \forall \; f \in \mathcal{C}(K).$$

Una delle proprietà significative dell'operatore aggiunto di una mappa lineare continua è quella di conservare la norma, ossia $\|T^*\|_{\mathcal{L}(X^*,\mathcal{M}(K))} = \|T\|_{\mathcal{L}(\mathcal{C}(K),X)} = 1$. Fissiamo $u^* \in U \subseteq \partial_e B_{X^*}$: essendo un punto estremo, u^* appartiene alla frontiera di B_{X^*}, ovvero $\|u^*\|_{X^*} = 1$. Posto $T^*u^* = \mu \in \mathcal{M}(K)$, vale

$$\|\mu\|_{\mathcal{M}(K)} = \|T^*u^*\|_{\mathcal{M}(K)} \leqslant \|u^*\|_{X^*} = 1.$$

Fissiamo un intorno aperto debole* V di u^*, relativo a K; poniamo $K_0 = K \setminus V$. Possiamo definire due mappe $v^*, w^* \in X^*$, date da

$$v^*(x) = \int\limits_V x^*(x)\ d\mu(x^*), \qquad w^*(x) = \int\limits_{K_0} x^*(x)\ d\mu(x^*).$$

Per queste mappe vale la seguente maggiorazione della norma:

$$\|v^*\|_{X^*} \leqslant |\mu|(V), \qquad \|w^*\|_{X^*} \leqslant |\mu|(K_0).$$

Infatti per entrambe (lo proviamo solo v^*, allo stesso modo per w^*) vale

$$\|v^*\|_{X^*} = \sup_{x\neq 0} \frac{|v^*(x)|}{\|x\|} = \sup_{x\neq 0} \frac{|\int_V x^*(x)\ d\mu(x^*)|}{\|x\|} \leqslant \sup_{x\neq 0} \frac{\int_V |x^*(x)|\ d|\mu|(x^*)}{\|x\|} \leqslant$$

$$\leqslant \sup_{x\neq 0} \frac{[\max_{x^*\in K} |x^*(x)|]\,|\mu|(V)}{\|x\|} = \sup_{x\neq 0} \frac{\|x\|\,|\mu|(V)}{\|x\|} = |\mu|(V).$$

Ne consegue che v^* e w^* appartengono a B_{X^*}, essendo

$$\|v^*\|_{X^*} \leqslant |\mu|(V) \leqslant |\mu|(K) = \|\mu\|_{\mathcal{M}(K)} \leqslant 1,$$

dove nell'ultima uguaglianza abbiamo usato la definizione di norma in $\mathcal{M}(K)$.
Per definizione di operatore aggiunto T^*, e utilizzando le definizioni di T e $\widehat{x}$, cioè $T(\widehat{x}) = x$ e $\widehat{x}(x^*) = x^*(x)$, otteniamo

$$v^*(x) + w^*(x) = \int\limits_V x^*(x)\ d\mu(x^*) + \int\limits_{K_0} x^*(x)\ d\mu(x^*) = \int\limits_K x^*(x)\ d\mu(x^*) =$$

$$= \mu(\widehat{x}) = (T^*(u^*))(\widehat{x}) = u^*(T^*(\widehat{x})) = u^*(x),$$

quindi

$$v^* + w^* = u^*.$$

Di conseguenza

$$\|v^*\|_{X^*} + \|w^*\|_{X^*} = 1;$$

infatti,

$$1 = \|u^*\|_{X^*} = \|v^* + w^*\|_{X^*} \leqslant \|v^*\|_{X^*} + \|w^*\|_{X^*} \leqslant$$

$$\leqslant |\mu|(V) + |\mu|(K_0) = |\mu|(K) = \|\mu\|_{\mathcal{M}(K)} \leqslant 1$$

(da cui seguono anche $\|v^*\|_{X^*} = |\mu|(V)$ e $\|w^*\|_{X^*} = |\mu|(K_0)$).
Dato che $\|v^*\|_{X^*} + \|w^*\|_{X^*} = 1$, almeno uno tra v^* e w^* non è l'operatore nullo.
Se uno dei due è nullo, per esempio $w^* = 0$, allora $u^* = v^*$ e $|\mu|(V) = 1$.
Altrimenti, cioè supponendo entrambi non nulli, da $\|v^*\|_{X^*} + \|w^*\|_{X^*} = 1$ deduciamo che

$$u^* = v^* + w^* = \|v^*\|_{X^*} \frac{v^*}{\|v^*\|_{X^*}} + \|w^*\|_{X^*} \frac{w^*}{\|w^*\|_{X^*}}$$

è combinazione convessa di due punti $\frac{v^*}{\|v^*\|_{X^*}}, \frac{w^*}{\|w^*\|_{X^*}} \in B_{X^*}$.
Ma $u^* \in U \subset \partial_e B_{X^*}$, allora per definizione di punto estremo deve essere

$$u^* = \frac{v^*}{\|v^*\|_{X^*}} = \frac{w^*}{\|w^*\|_{X^*}}.$$

Avendo supposto v^* e w^* entrambi non nulli, sia $\alpha = |\mu|(K_0) = \|w^*\|_{X^*}$, quindi $\alpha > 0$. Allora

$$u^*(x) = \frac{w^*}{\|w^*\|_{X^*}} = \alpha^{-1} \int_{K_0} x^*(x)\ d\mu(x^*), \qquad x \in X;$$

in particolare,

$$|u^*(x)| \leqslant \max_{x^* \in K_0} |x^*(x)|, \qquad x \in X,$$

(visto che $|u^*(x)| = \alpha^{-1} \left| \int_{K_0} x^*(x)\ d\mu(x^*) \right| \leqslant \alpha^{-1} \left[\max_{x^* \in K_0} |x^*(x)| \right] \alpha)$.
Sfruttando l'ultima espressione ottenuta per u^*, proviamo che

$$u^* \in \overline{co}^* \left(K_0 \cup (-K_0) \right).$$

Dato che K_0 è un chiuso nel compatto K, anche K_0 è un compatto debole*. Fissato $\delta > 0$, grazie alle compattezza di K_0, il suo ricoprimento aperto $\{B(x^*, \delta),\ x^* \in K_0\}$ costituito dalle palle aperte di raggio δ centrate nei punti di K_0, ammette un sottoricoprimento finito di K_0, che indichiamo con $\mathcal{U}_\delta = \{B(x_i^*, \delta),\ 1 \leqslant i \leqslant N\}$. A partire da $\mathcal{U}_\delta$ possiamo ottenere un sottoricoprimento finito $\mathcal{U}_\delta'$ di K_0 costituito da sottoinsiemi disgiunti; basta infatti porre $\mathcal{U}_\delta' = \left\{ \widehat{B_i},\ 1 \leqslant i \leqslant N \right\}$, dove

$$\widehat{B_1} = B(x_1^*, \delta), \qquad \widehat{B_i} = B(x_i^*, \delta) \setminus \left(\bigcup_{j=1}^{i-1} \widehat{B_j} \right) \quad \text{per } 2 \leqslant i \leqslant N.$$

Mediante il ricoprimento $\mathcal{U}_\delta'$ possiamo decomporre u^* nella forma

$$u^* = \alpha^{-1} \int_{K_0} x^*\ d\mu(x^*) = \sum_{i=1}^{N} \alpha^{-1} \int_{\widehat{B_i}} x^*\ d\mu(x^*) =$$

$$= \sum_{i=1}^{N} \frac{\mu(\widehat{B_i})}{\alpha} x_i^* \quad + \quad \sum_{i=1}^{N} \frac{1}{\alpha} \int_{\widehat{B_i}} (x^* - x_i^*)\ d\mu(x^*), \qquad \text{da cui}$$

$$\left\| u^* - \left(\sum_{i=1}^{N} \frac{\mu(\widehat{B_i})}{\alpha} x_i^* \right) \right\|_{X^*} \leqslant \quad \left\| \sum_{i=1}^{N} \frac{1}{\alpha} \int_{\widehat{B_i}} (x^* - x_i^*)\ d\mu(x^*) \right\|_{X^*} \leqslant$$

$$\leqslant \quad \sum_{i=1}^{N} \|x^* - x_i^*\|_{X^*} \frac{|\mu| (\widehat{B_i})}{\alpha} \leqslant \quad \delta \sum_{i=1}^{N} \frac{|\mu| (\widehat{B_i})}{\alpha} = \delta \frac{|\mu| (K_0)}{\alpha} = \delta.$$

Scegliendo $\delta = 1/n$, al variare di $n \in \mathbb{N}_0$, e indicando con x_i^{*n} i centri delle palle di raggio $1/n$ del corrispondente ricoprimento $\mathcal{U}_{1/n}$ e con $\widehat{B_i^n}$ gli insiemi di $\mathcal{U}'_{1/n}$ otteniamo

$$\left\| u^* - \left(\sum_{i=1}^{N_n} \frac{\mu(\widehat{B_i^n})}{\alpha} x_i^{*n} \right) \right\|_{X^*} \leqslant \frac{1}{n}, \qquad \text{ovvero}$$

$$u^* = \lim_{n \to +\infty} \left(\sum_{i=1}^{N_n} \frac{\mu(\widehat{B_i^n})}{\alpha} x_i^{*n} \right).$$

Per ogni n, se μ è positiva, $\sum_{i=1}^{N_n} \frac{\mu(\widehat{B_i^n})}{\alpha} x_i^{*n}$ è una combinazione convessa degli x_i^{*n}, quindi $u^* \in \overline{co(K_0)}^* = \overline{co}^*(K_0) \subseteq \overline{co}^* (K_0 \cup (-K_0))$.
In generale, per una misura μ con segno, poniamo per $i = 1, ..., N_n$

$$y_i^{*n} = \begin{cases} x_i^{*n} & \text{se } \mu(\widehat{B_i^n}) \geqslant 0, \\ -x_i^{*n} & \text{se } \mu(\widehat{B_i^n}) < 0. \end{cases}$$

Così facendo gli y_i^{*n} appartengono a $(K_0 \cup (-K_0))$ e

$$u^* = \lim_{n \to +\infty} \left(\sum_{i=1}^{N_n} \frac{\mu(\widehat{B_i^n})}{\alpha} x_i^{*n} \right) = \lim_{n \to +\infty} \left(\sum_{i=1}^{N_n} \frac{|\mu| (\widehat{B_i^n})}{\alpha} y_i^{*n} \right);$$

certamente $\sum_{i=1}^{N_n} \frac{|\mu|(\widehat{B_i^n})}{\alpha} y_i^{*n}$ è una combinazione convessa degli $y_i^{*n} \in (K_0 \cup (-K_0))$, visto che

$$|\mu| (\widehat{B_i^n}) \geqslant 0 \text{ per ogni } i, \qquad \sum_{i=1}^{N_n} \frac{|\mu| (\widehat{B_i^n})}{\alpha} = \frac{|\mu| (K_0)}{\alpha} = 1.$$

Questo significa che, per ogni n, $\sum_{i=1}^{N_n} \frac{|\mu|(\widehat{B_i^n})}{\alpha} y_i^{*n} \in co\,(K_0 \cup (-K_0))$; di conseguenza, essendo limite di elementi appartenenti all'inviluppo convesso di $(K_0 \cup (-K_0))$, vale

$$u^* \in \overline{co\,(K_0 \cup (-K_0))}^* = \overline{co}^* (K_0 \cup (-K_0)).$$

D'altra parte u^* è un punto estremo di B_{X^*} e $\overline{co}^* (K_0 \cup (-K_0)) \subseteq \overline{co}^* (K \cup (-K)) = B_{X^*}$, perciò u^* deve essere un punto estremo anche per $\overline{co}^* (K_0 \cup (-K_0))$.

Grazie al teorema di Milman deduciamo che $u^* \in (K_0 \cup (-K_0))$. Siccome $u^* \in V$ e V è disgiunto da K_0, si ha $u^* \in -K_0$ cioè $-u^* \in K_0$. Quindi

$$-u^* \in [K_0 \cap (-U)] \subseteq [K \cap (-U)]\,,$$

che è assurdo perché abbiamo provato inizialmente che K non interseca $-U$.
Allora $w^* = 0$, che implica $u^* = v^*$ e $\|v^*\|_{X^*} = |\mu|\,(V) = 1$. Quindi $|\mu(V)| = 1$ per ogni intorno aperto debole* V di u^*. Essendo μ una misura regolare si ha

$$\mu(\{u^*\}) = \inf\{\mu(V):\ K \supseteq V \supseteq \{u^*\}\,,\ V \text{ aperto misurabile secondo } \mu\} = \pm 1.$$

Pertanto μ coincide con δ_{u^*} oppure con $-\delta_{u^*}$, dove con δ_{u^*} indichiamo la misura centrata nel punto u^*, ovvero quella che su ogni boreliano B di K vale

$$\delta_{u^*}(B) = \begin{cases} 1 & \text{se } u^* \in B, \\ 0 & \text{altrimenti.} \end{cases}$$

In realtà μ non può essere uguale a $-\delta_{u^*}$, altrimenti per ogni $x \in X$ risulterebbe

$$u^*(x) = v^*(x) = \int\limits_V x^*(x)\ d\mu(x^*) = \int\limits_V x^*(x)\ d(-\delta_{u^*})(x^*) = -u^*(x),$$

ovvero $u^* = 0$, che è assurdo visto che $\|u^*\|_{X^*} = 1$.
Dunque $T^*(u^*) = \mu = \delta_{u^*}$ per $u^* \in U$. Siccome T^* è continua rispetto alla topologia debole*, vale

$$T^*(x^*) = \delta_{x^*} \qquad \forall\ x^* \in K.$$

Allora per ogni $f \in \mathcal{C}(K)$ si ha

$$x^*\left(T(f)\right) = \left(T^*(x^*)\right)(f) = \delta_{x^*}(f) = \int\limits_K f\ d\delta_{x^*} = f(x^*), \qquad \forall\ x^* \in K.$$

Finalmente quest'ultima relazione ci permette di provare che

$$J \circ T = Id|_{\mathcal{C}(K)}.$$

Infatti, per ogni $f \in \mathcal{C}(K)$ vale $J\left(T(f)\right) = \widehat{T(f)}$, dove

$$\widehat{T(f)}(x^*) = x^*\left(T(f)\right) = f(x^*) \quad \forall\ x^* \in K \quad \Longrightarrow \quad J\left(T(f)\right) = f.$$

Allora $J(X) = \mathcal{C}(K)$, dunque possiamo concludere che

$$J : X \longrightarrow \mathcal{C}(K) \qquad \text{è un isomorfismo isometrico.}$$

Come volevasi dimostrare, X è isometricamente isomorfo ad uno spazio $\mathcal{C}(K)$. $\square$

Con la prossima e conclusiva proposizione, che enunciamo soltanto, elenchiamo alcuni esempi non banali di spazi isometricamente iniettivi.
Per una dimostrazione rimandiamo a [2], proposizione 4.3.8.

Proposizione 3.22. (*i*) *Se* $\mathcal{C}(K)$ *è isometricamente isomorfo a uno spazio duale, allora* $\mathcal{C}(K)$ *è isometricamente iniettivo.*
(*ii*) *Se* $(\Omega, \mathcal{F}, \mu)$ *è spazio misurato con* μ σ*-finita, allora* $L^\infty(\Omega, \mathcal{F}, \mu)$ *è isometricamente iniettivo.*
(*iii*) *Per ogni spazio compatto e di Hausdorff* K*, lo spazio* $\mathcal{C}(K)^{**}$ *è isometricamente iniettivo.*

Riferimenti bibliografici

[1] Acquistapace, P., Appunti di Analisi Funzionale, http://www.dm.unipi.it/∼acquistp/anafun.pdf.

[2] Albiac, F., Kalton, N.J., Topics in Banach Spaces Theory, Springer Verlag (2006).

[3] Ciarlet, P.G., Linear and Nonlinear Functional Analysis with Applications, Society For Industrial and Applied Mathematics, U.S. (2013).

[4] Halmos, P., Givant, S., Introduction to Boolean algebras, Springer (2009).

[5] Jech, T.J., The Axiom of Choice, Academic Press (1973).

[6] Schechter, E., Handbook of Analysis and Its Foundations, Academic Press (1996).

[7] Sikorski, R., Boolean algebras, Springer (1969) .

[8] Sofi, M.A., Some problems in functional analysis inspired by Hahn Banach type theorems, Ann. Funct. Anal. 5 (2014) 1-29.

[9] Pawlikowski, J., The Hahn-Banach theorem implies the Banach-Tarski paradox, Fund. Math. 138 (1991) 21-22.

[10] Wagon, S., The Banach-Tarski Paradox, Cambridge University Press (1993).