ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea in Matematica

Il Teorema di Riesz per l'integrale di Riemann-Stieltjes

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.mo Prof.
Andrea Bonfiglioli

Presentata da:
Fabio Valenti Pettino

Anno Accademico 2023/2024

Introduzione

In questa tesi si intende analizzare il Teorema di Riesz per l'integrale di Riemann-Stieltjes, un risultato fondamentale in Analisi Funzionale che stabilisce una corrispondenza tra i funzionali lineari continui (su un intervallo compatto $[a,b] \subseteq \mathbb{R}$) e le funzioni di variazione limitata, attraverso l'uso degli integrali di Riemann-Stieltjes. Per affrontare questo argomento, è necessario esaminare dapprima l'integrale di Riemann-Stieltjes considerando funzioni integrande e integratrici entrambe limitate su un intervallo compatto $[a,b] \subseteq \mathbb{R}$.

Nonostante la teoria dell'integrale di Riemann-Stieltjes sia ormai ben consolidata, la sua introduzione risale al 1894, quando Stieltjes ne propose per la prima volta una definizione. Successivamente, nel 1909, Riesz ampliò e arricchì il lavoro di Stieltjes con il noto risultato che oggi prende il nome di Teorema di Riesz. Questo teorema ha avuto un impatto significativo sulla teoria degli spazi di Banach e sulla rappresentazione dei funzionali. Tuttavia, è importante notare che esistono diverse presentazioni dell'integrale di Riemann-Stieltjes nella letteratura, tra cui quella proposta da Apostol [1]. La tesi esaminerà in dettaglio le differenze tra queste varie definizioni.

Dopo aver delineato il contesto storico e teorico dell'integrale di Riemann-Stieltjes e del Teorema di Riesz, è opportuno approfondire la definizione formale di questo integrale, essenziale per la comprensione dei risultati che seguiranno. Siano $f, g : [a,b] \to \mathbb{R}$ due funzioni limitate, $P = \{x_0, x_1, ..., x_n\}$ una partizione dell'intervallo $[a,b]$ e sia t_k un punto dell'intervallo $[x_{k-1}, x_k]$, per ogni $k = 1, ..., n$. L'**integrale di Riemann-Stieltjes** di f rispetto a g è definito (se esiste) come il valore reale $\int_a^b f \, \mathrm{d}g$, tale che per ogni $\varepsilon > 0$ esiste $\delta(\varepsilon) > 0$ tale che, per ogni partizione P di $[a,b]$ purché sia $\|P\| := \max_k (x_k - x_{k-1}) < \delta(\varepsilon)$ e qualunque sia la scelta dei punti t_k, si ha

$$\left| \int_a^b f \, \mathrm{d}g - \sum_{k=1}^{n} f(t_k)(g(x_k) - g(x_{k-1})) \right| < \varepsilon.$$

Possiamo, quindi, comprendere meglio la rilevanza di questo strumento nel contesto dell'Analisi Funzionale grazie al risultato principale di questa tesi. Il **Teorema di Riesz** afferma che, dato un qualsiasi funzionale lineare continuo ϕ definito sullo spazio $C([a,b])$

(munito della usuale norma del sup), esiste una funzione a variazione limitata g su $[a,b]$ tale che:

$$\phi(f) = \int_a^b f \, \mathrm{d}g, \quad \forall\, f \in C([a,b]).$$

In altre parole, per lo spazio delle funzioni continue su un intervallo $[a,b]$, ogni funzionale lineare continuo può essere rappresentato come un integrale rispetto a una funzione a variazione limitata. Inoltre, il teorema stabilisce una corrispondenza tra i funzionali lineari continui e le funzioni a variazione limitata.

Per dimostrare il Teorema di Riesz per l'integrale di Riemann-Stieltjes, come proposto in [2] e [3], sarà necessario introdurre ulteriori strumenti e concetti. In particolare, data una funzione f definita su $[0,1]$, si definisce **polinomio di Bernstein** di grado n associato a f come

$$B_n(x) := \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Questi polinomi sono importanti per l'approssimazione di funzioni su $[0,1]$ e, per la dimostrazione del teorema, risulta fondamentale il fatto che i coefficienti abbiano come fattori delle immagini di f. Inoltre, saranno discussi i due Teoremi di Helly, che offrono risultati significativi sulla convergenza di famiglie di funzioni a variazione limitata. Il **primo Teorema di Helly** enuncia che, data una successione di funzioni a valori reali $(f_n)_{n\in\mathbb{N}}$ definite sull'intervallo $[a,b]$, se esiste $M>0$ tale che

$$\sup\{|f_n(x)|\,; n \in \mathbb{N}, x \in [a,b]\} \le M, \qquad V_a^b(f_n) \le M, \quad \forall\, n \in \mathbb{N},$$

allora esiste una sottosuccessione di $(f_n(x))_{n\in\mathbb{N}}$ che converge puntalmente su $[a,b]$ a una funzione $f(x)$ a variazione limitata. In altre parole, data una successione di funzioni equilimitate e con variazioni uniformemente dominate da una costante, allora si può estrarre una sottosuccessione puntualmente convergente ad una funzione a variazione limitata. Si tratta quindi di un teorema di relativa compattezza.

Il **secondo Teorema di Helly** stabilisce che, data una funzione f continua su $[a,b]$ e una successione di funzioni $(g_n(x))_{n\in\mathbb{N}}$ che converge puntualmente a una funzione $g(x)$ su $[a,b]$, se esiste $M>0$ tale che $V_a^b(g_n) < M$ per ogni n, allora

$$\lim_{n\to\infty} \int_a^b f \, \mathrm{d}g_n = \int_a^b f \, \mathrm{d}g.$$

Questo risultato è rilevante perché garantisce, sotto le condizioni di uniformità sulle variazioni totali delle funzioni integratrici, il passaggio del limite sotto integrale per le funzioni integratrici.

Oltre alla dimostrazione principale del Teorema di Riesz, sarà fornita una dimostrazione alternativa che utilizza il Teorema di Hahn-Banach [6]. Questa dimostrazione offre un approccio diverso e complementare alla rappresentazione dei funzionali lineari attraverso gli integrali di Riemann-Stieltjes. Infine, sarà considerata una generalizzazione del Teorema di Riesz per funzionali lineari continui sullo spazio delle funzioni continue definite su un insieme compatto arbitrario [7]. Questa estensione si basa su alcuni prerequisiti di Teoria della Misura, i quali non saranno trattati dettagliatamente in questa sede.

Indice

Capitolo 1

Proprietà generali dell'integrale di Riemann-Stieltjes

1.1 Funzioni a variazione limitata

Da qui, per brevità di notazione, adotteremo le seguenti terminologie e notazioni. Una partizione P di $[a, b]$ è un insieme finito di punti, diciamo $P = \{x_0, x_1, ..., x_n\}$ dove $a := x_0 < x_1 < ... < x_{n-1} < x_n =: b$. Diremo che una partizione Q di $[a, b]$ è più fine di un'altra partizione P di $[a, b]$, se $Q \supseteq P$. Data una qualsiasi funzione g definita sull'intervallo $[a, b]$, useremo la notazione Δg_k per indicare la differenza $\Delta g_k := g(x_k) - g(x_{k-1})$, dove gli x_k sono i punti che costituiscono la partizione di $[a, b]$ data da $\{a := x_0, x_1, ..., x_{n-1}, x_n =: b\}$.

Definizione 1.1 (**Funzione a variazione limitata e di variazione totale**). *Sia f una funzione definita su $[a, b]$. Sia $P = \{x_0, x_1, ..., x_n\}$ una partizione di $[a, b]$ e poniamo $\sum(P) := \sum_{k=1}^{n} |\Delta f_k|$ relativa alla partizione P. Definiamo variazione totale di f su $[a, b]$ il numero reale definito da*

$$V_a^b(f) := \sup\left\{ \sum(P) \mid P \text{ è partizione di } [a, b] \right\}.$$

Diciamo che f è una funzione a variazione limitata su $[a, b]$ se $V_a^b(f) < +\infty$.

Lemma 1.2. *Supponiamo f e g funzioni a variazione limitata su $[a, b]$. Allora $f + g$, $f - g$ e fg sono funzioni a variazione limitata su $[a, b]$. Inoltre si ha*

$$V(f \pm g) \leq V(f) + V(g) \qquad e \qquad V(fg) \leq \|g\|_\infty V(f) + \|f\|_\infty V(g),$$

dove $\|g\|_\infty := \sup\{|g(x)| : x \in [a, b]\}$ e $\|f\|_\infty := \sup\{|f(x)| : x \in [a, b]\}$.

Dimostrazione. I casi relativi a $f+g$ e $f-g$ seguono direttamente dalle definizioni. Appena più complicato è mostrare la tesi per fg. Poniamo $h(x) := f(x)g(x)$. Per ogni partizione P di $[a,b]$ abbiamo

$$\begin{aligned}|\Delta h_k| =& |f(x_k)g(x_k) - f(x_{k-1})g(x_{k-1})| \\ =& |f(x_k)g(x_k) - f(x_{k-1})g(x_k) + f(x_{k-1})g(x_k) - f(x_{k-1})g(x_{k-1})| \\ \leq& |f(x_k)g(x_k) - f(x_{k-1})g(x_k)| + |f(x_{k-1})g(x_k) - f(x_{k-1})g(x_{k-1})| \\ =& |\Delta f_k||g(x_k)| + |\Delta g_k||f(x_{k-1})| \leq \|g\|_\infty |\Delta f_k| + \|f\|_\infty |\Delta g_k|.\end{aligned}$$

Questo dimostra sia che h è a variazione limitata, sia $V(h) \leq \|g\|_\infty V(f) + \|f\|_\infty V(g)$. □

Lemma 1.3. *Se f è monotona su $[a,b]$, allora f è a variazione limitata su $[a,b]$.*

Dimostrazione. Se f crescente, per ogni partizione P di $[a,b]$ si ha $\Delta f_k \geq 0$, quindi

$$\sum_{k=1}^{n} |\Delta f_k| = \sum_{k=1}^{n} \Delta f_k = \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) = f(b) - f(a),$$

e similmente se f decrescente. □

Teorema 1.4. *Sia $f : [a,b] \to \mathbb{R}$ una funzione monotona. Allora l'insieme dei punti di discontinuità di f è al più numerabile.*

Dimostrazione. Si supponga f crescente e sia E l'insieme dei punti di discontinuità di f in (a,b). Per $x \in E$ si ha $f(x^-) < f(x^+)$; allora a ogni punto in x di E si può associare un numero razionale $r(x)$ tale che

$$f(x^-) < r(x) < f(x^+).$$

Per la monotonia di f, si ha che $x_1 < x_2$ implica $f(x_1^+) \leq f(x_2^-)$. Si deduce allora che $r(x_1) \neq r(x_2)$. Si stabilisce così una corrispondenza biunivoca tra l'insieme E e un sottoinsieme dei razionali. □

Lemma 1.5. *Una funzione f a variazione limitata su $[a,b]$ si può scrivere come differenza di due funzioni crescenti.*

1.2 Integrale di Riemann-Stieltjes

In questa sezione esamineremo due definizioni dell'integrale di Riemann-Stieltjes che non sono del tutto equivalenti. Dimostreremo i risultati seguendo la seconda definizione, meno generale, ma nelle trattazioni successive metteremo in luce le differenze tra le due definizioni.

Definizione 1.6. *Siano date $f : [a,b] \to \mathbb{R}$ una funzione limitata, $P = \{x_0, x_1, ..., x_n\}$ una partizione dell'intervallo $[a,b]$ e sia t_k un punto dell'intervallo $[x_{k-1}, x_k]$, per ogni $k = 1, ..., n$. Sia inoltre α una funzione limitata su $[a,b]$. Definiamo somma di Riemann-Stieltjes di f rispetto ad α una somma della forma*

$$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k.$$

Osservazione 1.7. Avvertiamo il lettore che la notazione $S(P, f, \alpha)$ per la somma di Riemann-Stieltjes è una notazione imprecisa in quanto omette la dipendenza dalla n-upla di valori $(t_1, ..., t_n)$. Per non appesantire ulteriormente la notazione tralasceremo questa dipendenza, ma è bene ricordarla per una più completa comprensione della definizione appena data.

Definizione 1.8 (Integrale di Riemann-Stieltjes). *Diciamo che f è Riemann-Stieltjes integrabile rispetto ad α su $[a,b]$, e scriviamo $f \in RS(\alpha)$ su $[a,b]$, se esiste un numero reale A tale per cui, per ogni $\varepsilon > 0$, esiste una partizione P_ε di $[a,b]$ tale che per ogni partizione P più fine di P_ε e per ogni scelta dei punti t_k in $[x_{k-1}, x_k]$, si ha*

$$|S(P, f, \alpha) - A| < \varepsilon.$$

Quando un tale numero A esiste, esso è unicamente determinato e lo si denota con $\int_a^b f \, \mathrm{d}\alpha$ o con $\int_a^b f(x) \, \mathrm{d}\alpha(x)$. Diciamo anche che l'integrale di Riemann-Stieltjes $\int_a^b f \, \mathrm{d}\alpha$ esiste.

Definizione 1.9 (Integrale di Stieltjes). *Sia $[a,b] \subseteq \mathbb{R}$ un intervallo compatto e $f : [a,b] \to \mathbb{R}$, $g : [a,b] \to \mathbb{R}$. Sia $P = \{a = x_0, x_1, ..., x_n = b\}$ una partizione dell'intervallo $[a,b]$ e t_j un punto arbitrario di $[x_{j-1}, x_j]$, $j = 1, 2, \ldots, n$.*

Si dice che f è integrabile secondo Stieltjes, o S-integrabile, relativamente a g su $[a,b]$, e scriviamo $f \in S(g)$ su $[a,b]$, se esiste un numero reale I tale che per ogni $\varepsilon > 0$ esiste $\delta(\varepsilon) > 0$ tale che per ogni partizione P di $[a,b]$ purché sia $\|P\| := \max_k (x_k - x_{k-1}) < \delta$ e qualunque sia la scelta dei punti t_j risulta

$$|I - S(P, f, g)| < \varepsilon.$$

Questo I, se esiste, si chiama integrale di Stieltjes della f rispetto a g su $[a,b]$ e si denota col simbolo $\int_a^b f \, \mathrm{d}g$.

Osservazione 1.10. Se $f \equiv 1$ allora, presa una qualsiasi partizione P,

$$S(P, 1, \alpha) = \sum_{k=1}^{n} \Delta\alpha_k = \alpha(b) - \alpha(a).$$

Per l'arbitrarietà di P, segue

$$\int_a^b f \, \mathrm{d}\alpha = \alpha(b) - \alpha(a).$$

Teorema 1.11. *Se $f_1, f_2 \in S(g)$ su $[a,b]$, allora $\alpha f_1 + \beta f_2 \in S(g)$ su $[a,b]$ per tutte le costanti α, β, e si ha*

$$\int_a^b (\alpha f_1 + \beta f_2) \, \mathrm{d}g = \alpha \int_a^b f_1 \, \mathrm{d}g + \beta \int_a^b f_2 \, \mathrm{d}g.$$

Dimostrazione. Poniamo $h := \alpha f_1 + \beta f_2$. Data una partizione Q di $[a,b]$, possiamo scrivere

$$\begin{aligned} S(Q,h,g) =& \sum_{k=1}^{n} h(t_k)\Delta g_k = \alpha \sum_{k=1}^{n} f_1(t_k)\Delta g_k + \beta \sum_{k=1}^{n} f_2(t_k)\Delta g_k \\ =& \alpha S(Q,f_1,g) + \beta S(Q,f_2,g). \end{aligned}$$

Si fissi ora un $\varepsilon > 0$, allora, poiché $f_1 \in S(g)$, esiste $\delta_1 > 0$ tale che per ogni partizione P con $\|P\| < \delta_1$ renda vera la disuguaglianza $\left| S(P,f_1,g) - \int_a^b f_1 \, \mathrm{d}g \right| < \varepsilon$. Analogamente esiste $\delta_2 > 0$ tale che per ogni partizione P con $\|P\| < \delta_2$ si ha $\left| S(P,f_2,g) - \int_a^b f_2 \, \mathrm{d}g \right| < \varepsilon$. Prendiamo $\delta = \min\{\delta_1, \delta_2\}$ e consideriamo una partizione P tale che $\|P\| < \delta$, quindi valgono le seguenti disuguaglianze

$$\begin{aligned} &\left| S(P,h,g) - \left(\alpha \int_a^b f_1 \, \mathrm{d}g + \beta \int_a^b f_2 \, \mathrm{d}g \right) \right| \\ &\quad = \left| \alpha S(P,f_1,g) + \beta S(P,f_2,g) - \left(\alpha \int_a^b f_1 \, \mathrm{d}g + \beta \int_a^b f_2 \, \mathrm{d}g \right) \right| \\ &\quad \le |\alpha| \left| S(P,f_1,g) - \int_a^b f_1 \, \mathrm{d}g \right| + |\beta| \left| S(P,f_2,g) - \int_a^b f_2 \, \mathrm{d}g \right| \\ &\quad < |\alpha|\varepsilon + |\beta|\varepsilon, \end{aligned}$$

e questo prova il teorema grazie all'arbitrarietà di ε. □

Teorema 1.12. *Se $f \in S(g_1)$ e $f \in S(g_2)$ su $[a,b]$, allora $f \in S(\alpha g_1 + \beta g_2)$ su $[a,b]$ per tutte le costanti α, β, e si ha*

$$\int_a^b f \, \mathrm{d}(\alpha g_1 + \beta g_2) = \alpha \int_a^b f \, \mathrm{d}g_1 + \beta \int_a^b f \, \mathrm{d}g_2.$$

Dimostrazione. La dimostrazione è analoga a quella del Teorema 1.11. □

Teorema 1.13. *Siano f e g due funzioni su $[a,b]$ e sia $c \in (a,b)$. Se $f \in S(g)$ su $[a,b]$, su $[a,c]$ e su $[c,b]$ allora*

$$\int_a^c f \, \mathrm{d}g + \int_c^b f \, \mathrm{d}g = \int_a^b f \, \mathrm{d}g. \tag{1.1}$$

Dimostrazione. Fissato $\varepsilon > 0$ esiste $\delta > 0$ tale che se $P = \{a = x_0, x_1, ..., x_m = c\}$ e $Q = \{c = x_m, x_{m+1}, ..., x_{m+n} = b\}$ sono due scomposizioni finite di $[a, c]$ e $[c, b]$ con $\|P\|, \|Q\| < \delta$ risulta

$$\left| S(P \cup Q, f, g) - \int_a^b f\, \mathrm{d}g \right| < \frac{\varepsilon}{3}$$
$$\left| S(P, f, g) - \int_a^c f\, \mathrm{d}g \right| < \frac{\varepsilon}{3}$$
$$\left| S(Q, f, g) - \int_c^b f\, \mathrm{d}g \right| < \frac{\varepsilon}{3}.$$

Perciò

$$\left| \int_a^b f\, \mathrm{d}g - \left[\int_a^c f\, \mathrm{d}g + \int_c^b f\, \mathrm{d}g \right] \right| < \varepsilon,$$

e quindi, per l'arbitrarietà di ε, segue l'affermazione del teorema. □

Teorema 1.14. *Sia $c \in (a, b)$. Se due dei tre integrali in (1.2) esistono secondo la Definizione 1.8, allora esiste anche il terzo e vale*

$$\int_a^c f\, \mathrm{d}\alpha + \int_c^b f\, \mathrm{d}\alpha = \int_a^b f\, \mathrm{d}\alpha. \tag{1.2}$$

Si noti che le ipotesi del Teorema 1.13 sono più forti di quelle del Teorema 1.14. È però notevole osservare che i primi due S-integrali in (1.1) possono esistere senza che esista il terzo. Dunque, per l'integrale secondo Stieltjes non vale l'analogo del Teorema 1.14 con le sue ipotesi; vediamolo nel seguente esempio.

Esempio 1.15. Sia

$$f(x) = \begin{cases} 0 & x \in [-1, 0] \\ 1 & x \in (0, 1], \end{cases} \quad \text{e} \quad g(x) = \begin{cases} 0 & x \in [-1, 0) \\ 1 & x \in [0, 1]. \end{cases}$$

Allora $\int\limits_{-1}^{0} f\, \mathrm{d}g$ e $\int\limits_{0}^{1} f\, \mathrm{d}g$ esistono e sono entrambi nulli. Però $f \notin S(g)$ su $[-1, 1]$; se infatti consideriamo una qualsiasi partizione P di $[-1, 1]$ allora

$$\sum_{k=1}^{n} f(t_k)\Delta g_k = \sum_{k=1}^{n} f(t_k)(g(x_k) - g(x_{k-1}))$$
$$= \begin{cases} f(t_k)(\underbrace{g(0)}_{1} - \underbrace{g(x_{k-1})}_{0}) + f(t_{k+1})(\underbrace{g(x_{k+1})}_{1} - \underbrace{g(0)}_{1}) = f(t_k) = 0 & \text{se } 0 = x_k \in P \\ & \text{se } 0 \notin P \text{ e} \\ f(t_k)(\underbrace{g(x_k)}_{1} - \underbrace{g(x_{k-1})}_{0}) = f(t_k) & \text{se } x_{k-1} < 0 < x_k, \end{cases}$$

dove nel primo caso si avrebbe $x_{k-1} \leq t_k \leq x_k := 0 \leq t_{k+1} \leq x_{k+1}$, mentre nel secondo caso $f(t_k)$ vale 0 o 1 rispettivamente se $t_k \leq 0$ o se $t_k > 0$, poiché qui si ha $x_{k-1} \leq t_k, 0 \leq x_k$. Questo prova che f non è S-integrabile perché per qualunque δ si possono costruire della partizioni P di $[a,b]$ con $\|P\| < \delta$ sia contenenti 0, sia non contenenti 0 e con $f(t_k) = 1$.

Teorema 1.16. *Sia f funzione continua su $[a,b]$ e g funzione a variazione limitata su $[a,b]$; allora $f \in S(g)$.*

Dimostrazione. Possiamo supporre g crescente poichè ogni funzione a variazione limitata è differenza di due funzioni crescenti. Considero una partizione $P = \{x_0, x_1, ..., x_n\}$ di $[a,b]$ e definisco $m_k := \min\limits_{[x_{k-1},x_k]} f(x)$ e $M_k := \max\limits_{[x_{k-1},x_k]} f(x)$. Siano

$$z = \sum_{k=1}^{n} m_k \Delta g_k, \qquad Z = \sum_{k=1}^{n} M_k \Delta g_k.$$

É ovvio che $z \leq S(P,f,g) \leq Z$ per ogni scelta dei punti t_k. Si verifica facilmente che la somma z non decresce e Z non cresce aggiungendo nuovi punti alla partizione. Da questo segue che nessuna delle somme z possa essere maggiore delle somme Z. Denotiamo con I il sup dell'insieme z di tutte le somme inferiori: $I := \sup z$. Per ogni partizione è vero che $z \leq I \leq Z$ e di conseguenza $|S(P,f,g) - I| < Z - z$. Poiché f è continua su un compatto e quindi uniformemente continua, scegliendo $\varepsilon > 0$ si trova $\delta > 0$ tale che $|f(x'') - f(x')| < \varepsilon$ con $|x'' - x'| < \delta$, allora si ha

$$M_k - m_k < \varepsilon \qquad k = 1, 2, \cdots, n,$$

per $\|P\| < \delta$. Perciò

$$Z - z < \varepsilon \left[g(b) - g(a)\right],$$

e dunque

$$|S(P,f,g) - I| < \varepsilon \left[g(b) - g(a)\right]$$

per $\|P\| < \delta$. In altre parole $I = \int_a^b f \, \mathrm{d}g$. □

Osservazione 1.17. Sia g a variazione limitata su $[a,b]$. Condizione necessaria e sufficiente affinchè sia $\int_a^b f \, \mathrm{d}g = 0$ per ogni $f \in C([a,b])$ è che sia $g(b) = g(a)$ e $g(x) = g(a)$ in ogni punto x di (a,b) in cui g è continua.

Supponiamo $\int_a^b f \, \mathrm{d}g = 0$ per ogni $f \in C([a,b])$. Anzitutto, essendo $\int_a^b \mathrm{d}g = 0$ risulta $g(b) - g(a) = 0$. Sia poi $c \in (a,b)$ e g continua in c; posto

$$f_n(x) = \begin{cases} 1 & x \in [a,c] \\ 1 - n(x-c) & x \in (c, c+\frac{1}{n}) \\ 0 & x \in [c+\frac{1}{n}, b], \end{cases}$$

per tutti gli $n \in \mathbb{N}$ per cui $c + \frac{1}{n} < b$. Risultando $f_n \in C([a,b])$ si ha

$$0 = \int_a^b f_n \, \mathrm{d}g = \int_a^c \mathrm{d}g + \int_c^{c+\frac{1}{n}} [1 - n(x-c)] \, \mathrm{d}g;$$

ma

$$\int_a^c \mathrm{d}g = g(c) - g(a),$$

$$\left| \int_c^{c+\frac{1}{n}} [1 - n(x-c)] \, \mathrm{d}g \right| \leq V_c^{c+\frac{1}{n}}(g) \overset{n\to\infty}{\longrightarrow} 0;$$

perciò $g(c) - g(a) = 0$.

Viceversa se g è variazione limitata su $[a,b]$, $g(a) = g(b)$ e $g(x) = g(a)$ in ogni punto x di (a,b) in cui g è continua, allora $\int_a^b f \, \mathrm{d}g = 0$ per ogni $f \in C([a,b])$. Infatti, per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni scomposizione finita $\{a := x_0, x_1, ..., x_p =: b\}$ di $[a,b]$ per cui $\max\limits_{k=1,2,...,p} (x_k - x_{k-1}) < \delta$ risulta

$$\left| \sum_{k=1}^{p} f(t_k)[g(x_k) - g(x_{k-1})] - \int_a^b f \, \mathrm{d}g \right| < \varepsilon \quad \forall \, t_k \in [x_{k-1}, x_k], k = 1, 2, \ldots, p.$$

Poiché l'insieme dei punti di continuità di g è denso in $[a,b]$ per il Teorema 1.4, scegliendo $x_1, \ldots, x_{p-1}$ in tale insieme, si ha $\left| \int_a^b f \, \mathrm{d}g \right| < \varepsilon$, e quindi, per l'arbitrarietà di ε, $\int_a^b f \, \mathrm{d}g = 0$.

Pertanto se g, h sono a variazione limitata su $[a,b]$, $\int_a^b f \, \mathrm{d}g = \int_a^b f \, \mathrm{d}h$ per ogni $f \in C([a,b])$ se e solo se $g(b) - h(b) = g(a) - h(a)$ e $g(x) - h(x) = g(a) - h(a)$ in tutti i punti x di (a,b) in cui g e h sono continue (più generale in cui $g - h$ è continua).

Lemma 1.18. *Se la funzione f è continua su $[a,b]$ e α è a variazione limitata su $[a,b]$, allora*

$$\left| \int_a^b f \, \mathrm{d}\alpha \right| \leq \|f\|_\infty V_a^b(\alpha)$$

Dimostrazione. Presa un'arbitraria partizione P di $[a,b]$, vale

$$|S(P, f, \alpha)| = \left| \sum_{k=1}^{n} f(t_k) \Delta\alpha_k \right| \leq \|f\|_\infty \sum_{k=1}^{n} \Delta\alpha_k \leq \|f\|_\infty V_a^b(\alpha).$$

La tesi è così provata. □

Teorema 1.19. *Sia $f(x)$ una funzione continua in $[a,b]$ e sia $\alpha(x)$ una funzione costante a tratti, ovvero esistono $c_1, \ldots, c_m \in [a,b]$ tali che $a < c_1 < c_2 < \ldots < c_m < b$ e α è costante in $(a, c_1), (c_k, c_{k-1})$ per $k = 1, 2, \ldots, m-1$ e in (c_m, b). Allora vale*

$$\int_a^b f \, \mathrm{d}\alpha = f(a) \left[\alpha(a^+) - \alpha(a) \right] + \sum_{k=1}^{m} f(c_k) \left[\alpha(c_k^+) - \alpha(c_k^-) \right] + f(b) \left[\alpha(b) - \alpha(b^-) \right].$$

Dimostrazione. Si deduce facilmente che

$$\begin{aligned} V_a^b(\alpha) &= \big[\alpha(a^+) - \alpha(a)\big] + \sum_{k=1}^{m} \big\{ \big[\alpha(c_k) - \alpha(c_k^-)\big] + \big[\alpha(c_k^+) - \alpha(c_k)\big] \big\} + \big[\alpha(b) - \alpha(b^-)\big] \\ &= \big[\alpha(a^+) - \alpha(a)\big] + \sum_{k=1}^{m} \big[\alpha(c_k^+) - \alpha(c_k^-)\big] + \big[\alpha(b) - \alpha(b^-)\big]\,, \end{aligned}$$

quindi α è a variazione limitata su $[a, b]$. Ponendo $c_0 = a, c_{m+1} = b$, si ha

$$\int_a^b f \,\mathrm{d}\alpha = \sum_{k=0}^{m} \int_{c_k}^{c_{k+1}} f \,\mathrm{d}\alpha = \sum_{k=0}^{m} \big\{ f(c_k)\,\big[\alpha(c_k^+) - \alpha(c_k)\big] + f(c_{k+1})\,\big[\alpha(c_{k+1}) - \alpha(c_{k+1}^-)\big] \big\}\,.$$

Questo conclude la dimostrazione. □

Osservazione 1.20. L'ipotesi di continuità di f garantisce l'esistenza dell'integrale. Infatti si può dimostrare che se la funzione integratrice è discontinua in un punto interno di $[a, b]$ la continuità della funzione integranda è condizione sufficiente all'esistenza dell'integrale. Nel seguente esempio mostriamo che l'integrale può non esistere se f è discontinua.

Esempio 1.21. Siano

$$\alpha(x) := \begin{cases} 1 & x = 0 \\ 0 & x \neq 0, \end{cases} \quad \text{e} \quad f(x) := \begin{cases} 2 & x = 0 \\ 1 & x \neq 0. \end{cases}$$

Prendiamo l'intervallo $[-1, 1]$ come intervallo di integrazione. Considerando una partizione P contenente lo 0, si ha

$$S(P, f, \alpha) = f(t_k)\left(\underbrace{\alpha(0)}_{1} - \underbrace{\alpha(x_{k-1})}_{0}\right) + f(t_{k+1})\left(\underbrace{\alpha(x_{k+1})}_{0} - \underbrace{\alpha(0)}_{1}\right) = f(t_k) - f(t_{k+1}),$$

dove $x_{k-1} \leq t_k \leq x_k := 0 \leq t_{k+1} \leq x_{k+1}$. Notiamo che in base alla scelta di t_k e t_{k+1}, il valore dell'integrale varia tra 0 (se entrambi sono diversi da zero), 1 (se solo $t_{k+1} \neq 0$) e -1 (se solo $t_k \neq 0$). Quindi l'integrale non può esistere.

1.3 Confronto tra le definizioni di integrale di Riemann-Stieltjes

In questa sezione, esamineremo la relazione tra due definizioni dell'integrale di Riemann-Stieltjes, dove la Definizione 1.9 è caratterizzata dall'uso del parametro di finezza $\delta(\varepsilon)$ anziché la finezza della partizione come invece nella Definizione 1.8. Questa differenza mostra che la Definizione 1.8 è più generale, in quanto la Definizione 1.9 può essere considerata un caso particolare di 1.8. Tuttavia, dimostreremo la non equivalenza delle definizioni fornendo un controesempio più avanti.

Teorema 1.22. *Se una funzione è integrabile secondo Stieltjes nel senso della Definizione 1.9 allora è integrabile anche secondo Riemann-Stieltjes nel senso della Definizione 1.8 e gli integrali coincidono.*

Dimostrazione. Data $f : [a, b] \to \mathbb{R}$ e $\alpha : [a, b] \to \mathbb{R}$ (limitate), supponiamo che esista I tale che, per ogni dato $\varepsilon > 0$, esista un $\delta_\varepsilon > 0$ tale che per ogni partizione P di $[a, b]$ con $\|P\| < \delta_\varepsilon$ e per ogni scelta di $t_k \in [x_{k-1}, x_k]$ per ogni $k = 1, ..., n$, si ha

$$\left|\lambda - \sum_{k=1}^{n} f(t_k)\Delta\alpha_k\right|.$$

Se δ_ε è come sopra, esiste certamente un $n \in \mathbb{N}$ sufficientemente grande affinché $\frac{b-a}{n} < \delta_\varepsilon$; consideriamo la partizione $P_\varepsilon := \{a := x_0, x_1, ..., x_n =: b\}$ ottenuta suddividendo $[a, b]$ in n parti uguali a $\frac{b-a}{n}$. Ne segue che $\|P_\varepsilon\| = \frac{b-a}{n} < \delta_\varepsilon$. Consideriamo una generica partizione P di $[a, b]$ tale che $P \supseteq P_\varepsilon$, come vuole la Definizione 1.8. Banalmente $\|P\| \leq \|P_\varepsilon\| < \delta_\varepsilon$, quindi segue immediatamente la tesi del teorema poiché l'ipotesi della Definizione 1.9 è quindi verificata in particolare per ogni partizione P di $[a, b]$ tale che $P \supseteq P_\varepsilon$. □

Osservazione 1.23. Il fatto cruciale su cui si basa la dimostrazione precedente è che per tutte le partizioni Q, P di $[a, b]$ tali che $Q \supseteq P$, vale $\|Q\| \leq \|P\|$, proprio grazie alla definizione di parametro di finezza. Non è vero però il viceversa. Banalmente, considerate le partizioni di $[0, 1]$, date da $Q := \{0, \frac{1}{3}, \frac{2}{3}, 1\}$ e $P := \{0, \frac{1}{2}, 1\}$, si ha che

$$\|Q\| = \frac{1}{3} < \frac{1}{2} = \|P\| \qquad \text{ma} \qquad Q \not\supseteq P \text{ e } Q \not\subseteq P.$$

La Osservazione 1.23 fa sospettare che le Definizioni 1.9 e 1.8 non siano equivalenti. Per provarlo in modo rigoroso, dobbiamo provare che ci sono delle funzioni f che sono RS-integrabili rispetto a g su un determinato intervallo $[a, b]$ secondo la Definizione 1.8, ma che non sono S-integrabili secondo la Definizione 1.9.

Teorema 1.24. *Dato $c \in (a, b)$, fissiamo tre numeri $\alpha(a), \alpha(c), \alpha(b)$ e definiamo*

$$\alpha(x) := \begin{cases} \alpha(a), & \text{se } x \in [a, c) \\ \alpha(c), & \text{se } x = c \\ \alpha(b), & \text{se } x \in (c, b]. \end{cases}$$

Sia inoltre f definita su $[a, b]$ in modo tale che almeno una delle due funzioni f o α sia continua a destra di c e almeno una delle due funzioni f o α sia continua a sinistra di c. Allora si ha che $f \in RS(\alpha)$ su $[a, b]$ e vale

$$\int_a^b f \, \mathrm{d}\alpha = f(c) \left(\alpha(c+) - \alpha(c-)\right).$$

Osservazione 1.25. Il teorema resta valido anche se $c = a$ o se $c = b$ a patto di scrivere $\alpha(c)$ rispettivamente per $\alpha(c-)$ nel primo caso e $\alpha(c+)$ nel secondo.

Dimostrazione. Per mostrare l'esistenza dell'integrale $\int_a^b f \, d\alpha$ andiamo preliminarmente a studiare la somma di Riemann-Stieltjes $S(P, f, \alpha)$ per una opportuna partizione P di $[a, b]$. Consideriamo quindi, per ogni $\varepsilon > 0$, una partizione P_ε di $[a, b]$ tale che $c \in P_\varepsilon$. A questo punto tutte le partizioni $P \supseteq P_\varepsilon$ sono forzate a contenere c. Dal momento che stiamo mostrando l'esistenza dell'integrale, richiedere che c appartenga alla partizione P_ε non è affatto restrittivo.

Supponiamo quindi che $c \in P$ e consideriamo la rispettiva somma $S(P, f, \alpha)$. Diciamo che c è il nodo i-esimo, cioè $x_i := c$. Allora

$$S(P, f, \alpha) = f(t_i)\Delta\alpha_i + f(t_{i+1})\Delta\alpha_{i+1},$$

poiché tutti gli altri termini sono nulli. Inoltre il valore di α nei nodi a destra e a sinistra di c sono rispettivamente $\alpha(c+)$ e $\alpha(c-)$. Riscrivendo quindi la somma, otteniamo

$$S(P, f, \alpha) = f(t_i)\Big(\alpha(c) - \alpha(c-)\Big) + f(t_{i+1})\Big(\alpha(c+) - \alpha(c)\Big),$$

dove $t_i \leq c \leq t_{i+1}$. Poniamo ora

$$\Delta := S(P, f, \alpha) - f(c)\Big(\alpha(c+) - \alpha(c-)\Big),$$

e notiamo che possiamo scriverla come

$$\begin{aligned}
\Delta =& S(P, f, \alpha) - f(c)[\alpha(c+) - \alpha(c-)] \\
=& f(t_i)[\alpha(c) - \alpha(c-)] + f(t_{i+1})[\alpha(c+) - \alpha(c)] - f(c)[\alpha(c+) - \alpha(c-)] \\
=& \alpha(c+)[f(t_{i+1}) - f(c)] + \alpha(c)[f(t_i) - f(t_{i+1})] + \alpha(c-)[f(c) - f(t_i)] \\
=& \alpha(c+)[f(t_{i+1}) - f(c)] + \alpha(c)[f(t_i) - f(c) + f(c) - f(t_{i+1})] + \alpha(c-)[f(c) - f(t_i)] \\
=& \alpha(c+)[f(t_{i+1}) - f(c)] - \alpha(c)[f(t_{i+1}) - f(c)] + \alpha(c)[f(t_i) - f(c)] \\
& - \alpha(c-)[f(t_i) - f(c)] \\
=& [\alpha(c+) - \alpha(c)][f(t_{i+1}) - f(c)] + [\alpha(c) - \alpha(c-)][f(t_i) - f(c)].
\end{aligned} \tag{1.3}$$

Applicando la disuguaglianza triangolare a (1.3) si ha

$$|\Delta| \leq |\alpha(c+) - \alpha(c)||f(t_{i+1}) - f(c)| + |\alpha(c) - \alpha(c-)||f(t_i) - f(c)|.$$

Da qui si ottiene facilmente la tesi analizzando singolarmente i vari casi: infatti se f è continua (bilateralmente) in c, per ogni $\varepsilon > 0$ esiste un $\delta_\varepsilon > 0$ tale che per ogni $\|P\| < \delta_\varepsilon$ si ha

$$|f(t_{i+1}) - f(c)| < \varepsilon \qquad \text{e} \qquad |f(t_i) - f(c)| < \varepsilon,$$

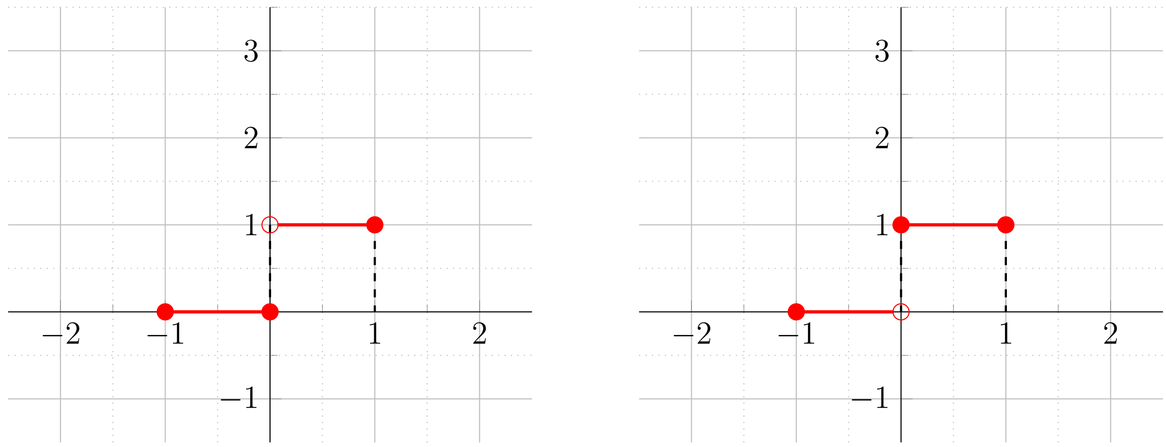
da cui $|S(P, f, \alpha) - f(c)(\alpha(c+) - \alpha(c-))| = |\Delta| < \varepsilon|\alpha(c+) - \alpha(c)| + \varepsilon|\alpha(c) - \alpha(c-)|$, che equivale a dire $\int_a^b f \, \mathrm{d}\alpha$ esiste e vale $f(c)\,(\alpha(c+) - \alpha(c-))$. Se invece f è discontinua in c sia a destra che a sinistra, allora per ipotesi deve essere $\alpha(c+) = \alpha(c) = \alpha(c-)$ e si ha $\Delta = 0$ per ogni P, in particolare $\int_a^b f \, \mathrm{d}\alpha = f(c)\,(\alpha(c+) - \alpha(c-))$.

Nel caso in cui f è continua a destra di c ma discontinua a sinistra di c, si ha che solo $\alpha(c-) = \alpha(c)$ ma, per ogni $\varepsilon > 0$, esiste $\delta_\varepsilon > 0$ tale che per $\|P\| < \delta_\varepsilon$ si ha $|f(t_{i+1}) - f(c)| < \varepsilon$, quindi $|\Delta| < \varepsilon|\alpha(c+) - \alpha(c)|$.

Similmente, per f continua a sinistra di c ma discontinua a destra di c, per ipotesi $\alpha(c+) = \alpha(c)$ da cui $|\Delta| < \varepsilon|\alpha(c-) - \alpha(c)|$. Questo conclude la prova del teorema. □

Esempio 1.26. Riprendiamo le funzioni f e g come nell'Esempio 1.15,

$$f(x) = \begin{cases} 0 & x \in [-1, 0] \\ 1 & x \in (0, 1], \end{cases} \quad \text{e} \quad g(x) = \begin{cases} 0 & x \in [-1, 0) \\ 1 & x \in [0, 1]. \end{cases}$$



(a) Grafico di f **(b)** Grafico di α

Abbiamo visto che $f \notin S(g)$ su $[-1, 1]$. Osserviamo però che le funzioni f e α rispettano le ipotesi del Teorema 1.24: infatti α è una funzione a gradino continua a destra di 0 ma discontinua a sinistra di 0, mentre f è continua a sinistra di 0 e discontinua a destra di 0. Questo permette di utilizzare il Teorema 1.24 ottenendo come valore dell'integrale di Riemann-Stieltjes

$$\int_{-1}^{1} f \, \mathrm{d}\alpha = f(0)(\underbrace{\alpha(0+)}_{1} - \underbrace{\alpha(0-)}_{0}) = 0.$$

Quindi, in simboli,

$$f \in RS(\alpha) \text{ in } [-1, 1] \qquad \text{ma} \qquad f \notin S(\alpha) \text{ in } [-1, 1].$$

Capitolo 2

Teoremi di Helly e polinomio di Bernstein

2.1 Polinomio di Bernstein

Definizione 2.1. *Sia f una funzione a valori reali definita su $[0,1]$. Il polinomio*

$$B_n(x) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}$$

è detto polinomio di Bernstein di grado n associato a f.

Osservazione 2.2. Poiché vale

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k},$$

scegliendo $p = x$ e $q = 1-x$ allora per ogni $x \in \mathbb{R}$

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1.$$

Lemma 2.3. *Per ogni x reale si ha*

$$\sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^k (1-x)^{n-k} \leq \frac{n}{4}.$$

Dimostrazione. Derivando rispetto a z l'identità

$$\sum_{k=0}^{n} \binom{n}{k} z^k = (1+z)^n \tag{2.1}$$

e moltiplicando per z si ha

$$\sum_{k=0}^{n} k \binom{n}{k} z^k = nz(1+z)^{n-1}. \tag{2.2}$$

Derivando e moltiplicando per z, si ottiene

$$\sum_{k=0}^{n} k^2 \binom{n}{k} z^k = nz(1+nz)(1+z)^{n-2}. \tag{2.3}$$

Ponendo $z = \frac{x}{1-x}$ nelle equazioni (2.1), (2.2) e (2.3), poi moltiplicando per $(1-x)^n$ si trova rispettivamente

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1, \tag{2.4}$$

$$\sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} = nx, \tag{2.5}$$

$$\sum_{k=0}^{n} k^2 \binom{n}{k} x^k (1-x)^{n-k} = nx(1-x+nx). \tag{2.6}$$

Moltiplicando (2.4) per n^2x^2, (2.5) per $-2nx$ e sommando le risultanti equazioni si ottiene

$$\sum_{k=0}^{n} (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k} = nx(1-x).$$

Per provare il lemma, basta osservare che per ogni x reale

$$x(1-x) \leq \frac{1}{4}.$$

Infatti

$$4x^2 - 4x + 1 = (2x-1)^2 \geq 0 \qquad \text{ossia} \qquad 1 \geq 4x(1-x).$$

Questo conclude la prova. $\square$

Teorema 2.4 (Teorema di Bernstein). *Se la funzione f è continua su $[0,1]$, allora $(B_n)_{n\in\mathbb{N}}$ converge uniformemente a f (per n che tende a $+\infty$) su $[0,1]$.*

Dimostrazione. Sia $M = \max|f(x)|$ e sia $\varepsilon > 0$. Poiché f è continua su un compatto, essa è uniformemente continua, quindi

$$\exists\, \delta > 0 \,:\, |f(x'') - f(x')| < \varepsilon \text{ se } |x'' - x'| < \delta.$$

Preso $x \in [0,1]$, segue dall'osservazione precedente che

$$f(x) = \sum_{k=0}^{n} f(x) \binom{n}{k} x^k (1-x)^{n-k},$$

dunque, per ogni $x \in [0,1]$,

$$|B_n(x) - f(x)| \leq \sum_{k=0}^{n} \left| f\left(\tfrac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k}. \tag{2.7}$$

Fissato $x \in [0,1]$, separiamo tutti i numeri $k = 0,1,2,\ldots,n$ in due insiemi A, B (dipendenti da x) in questo modo:

$$\begin{aligned} k \in A \quad &\text{se} \left|\frac{k}{n} - x\right| < \delta, \\ k \in B \quad &\text{se} \left|\frac{k}{n} - x\right| \geq \delta. \end{aligned}$$

Se $k \in A$, allora $\left|f\left(\frac{k}{n}\right) - f(x)\right| < \varepsilon$, perciò

$$\begin{aligned} \sum_A \left|f\left(\tfrac{k}{n}\right) - f(x)\right| \binom{n}{k} x^k (1-x)^{n-k} &\leq \varepsilon \sum_A \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \varepsilon \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \varepsilon. \end{aligned} \tag{2.8}$$

Se $k \in B$, allora $\frac{(k-nx)^2}{n^2\delta^2} \geq 1$, e quindi

$$\begin{aligned} &\sum_B \left|f\left(\tfrac{k}{n}\right) - f(x)\right| \binom{n}{k} x^k (1-x)^{n-k} \\ &\qquad \leq \sum_B 2M \binom{n}{k} \cdot 1 \cdot x^k (1-x)^{n-k} \\ &\qquad \leq \frac{2M}{n^2\delta^2} \sum_B \binom{n}{k} (k-nx)^2 x^k (1-x)^{n-k} \\ &\qquad \leq \frac{2M}{n^2\delta^2} \sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^k (1-x)^{n-k} \leq \frac{M}{2n\delta^2}, \end{aligned} \tag{2.9}$$

dove nell'ultimo passaggio, abbiamo usato il Lemma 2.3. Dalle disequazioni (2.7), (2.8) e (2.9), si ottiene che per ogni $x \in [0,1]$ vale

$$|B_n(x) - f(x)| < \varepsilon + \frac{M}{2n\delta^2}.$$

Di conseguenza, se $n > \frac{M}{2\varepsilon\delta^2}$ si ha

$$|B_n(x) - f(x)| < 2\varepsilon, \quad \forall\, x \in [0,1],$$

e questo termina la dimostrazione. □

Corollario 2.5 (**Teorema di approssimazione di Weierstrass**). *Data f funzione continua a valori reali su $[a,b]$, esiste una successione di polinomi $(p_n)_{n\in N}$ convergente uniformemente a f su $[a,b]$. In altre parole, i polinomi sono densi nello spazio delle funzioni continue $C([a,b])$.*

2.2 Primo Teorema di Helly

Teorema 2.6 (Primo Teorema di Helly). *Sia $(f_n)_{n\in\mathbb{N}}$ una successione di funzioni a valori reali definite sull'intervallo $[a,b]$ ed esista $M > 0$ tale che*

$$\sup\{|f_n(x)|\,;n\in\mathbb{N}, x\in[a,b]\}\le M, \qquad V_a^b(f_n)\le M, \quad \forall\, n\in\mathbb{N}.$$

Allora esiste una sottosuccessione di $(f_n(x))_{n\in\mathbb{N}}$ convergente in ogni punto di $[a,b]$ a una funzione $f(x)$ a variazione limitata. In altre parole, data una successione di funzioni equilimitate e con variazioni uniformemente dominate da una costante, allora si può estrarre una sottosuccessione puntualmente convergente ad una funzione a variazione limitata.

Dimostrazione. Per ogni funzione $f_n(x)$ della successione, poniamo

$$\varphi_n(x) = V_a^x(f_n), \quad \psi_n(x) = \varphi_n(x) - f_n(x), \qquad \forall\, n\in\mathbb{N}, \quad \forall\, x\in[a,b].$$

Le funzioni φ_n e ψ_n sono entrambe crescenti e inoltre

$$|\varphi_n(x)|\le M, \quad |\psi_n(x)|\le 2M, \quad \forall\, x\in[a,b].$$

Sia $X = \{x_1, x_2, \ldots\}$ un sottoinsieme numerabile denso in $[a,b]$ e tale che $a, b\in X$. Poiché $|\varphi_n(x_1)|\le M$, per Bolzano-Weierstrass la successione $(\varphi_n(x_1))_{n\in\mathbb{N}}$ ammette una sottosuccessione $(\varphi_{1,n}(x_1))_{n\in\mathbb{N}}$ convergente. Analogamente $|\varphi_{1,n}(x_2)|\le M$ quindi $(\varphi_{1,n}(x_2))_{n\in\mathbb{N}}$ ammette una sottosuccessione $(\varphi_{2,n}(x_2))_{n\in\mathbb{N}}$ convergente. Continuando così si ha

$\varphi_{1,1}(x), \ldots, \varphi_{1,n}(x), \ldots$ convergente per $x = x_1$,

$\varphi_{2,1}(x), \ldots, \varphi_{2,n}(x), \ldots$ convergente per $x = x_1$ e per $x = x_2$,

$$\vdots$$

$\varphi_{m,1}(x), \ldots, \varphi_{m,n}(x), \ldots$ convergente per $x = x_1, \ldots, x = x_m$, etc.

La successione (di tipo diagonale di Cantor) $(\varphi_{n,n}(x))_{n\in\mathbb{N}}$ converge per $x = x_k$ per ogni $k\in\mathbb{N}$. Poniamo

$$\varphi(x) = \begin{cases} \lim\limits_{n\to\infty}\varphi_{n,n}(x) & \forall\, x\in X \\ \sup\{\varphi(x_k); x_k < x\} & \forall\, x\in[a,b]\setminus X. \end{cases}$$

La funzione φ è crescente, infatti:

- se $x_h, x_k\in X$ con $x_h < x_k$, poiché $\varphi_{n,n}$ è una funzione crescente per ogni $n\in\mathbb{N}$ fissato, risulta che $\varphi(x_h)\le\varphi(x_k)$;

- se $x_h < x$ con $x_h \in X$ e $x \in [a,b] \setminus X$, si ha, per come è definita, $\varphi(x_h) \leq \varphi(x)$ e analogamente se $x \in [a,b] \setminus X, x_k \in X$ e $x < x_k$, si ha $\varphi(x) \leq \varphi(x_k)$;

- se $x', x'' \in [a,b] \setminus X$ e $x' < x''$, per densità di X in $[a,b]$ esiste $x_j \in X$ tale che $x' < x_j < x''$ e quindi $\varphi(x') \leq \varphi(x_j) \leq \varphi(x'')$.

Risulta $\lim\limits_{n\to\infty} \varphi_{n,n}(x) = \varphi(x)$ per ogni $x \in X$; consideriamo allora $x \in [a,b] \setminus X$ e distinguiamo due casi.

Supponiamo che φ sia continua in x. Se $\varepsilon > 0$, allora, per continuità, esiste $\delta > 0$ tale che, per ogni $y \in [x-\delta, x+\delta]$, $|\varphi(x) - \varphi(y)| < \varepsilon$. Poiché X è denso in $[a,b]$ esistono $x_h, x_k \in X \cap [x-\delta, x+\delta]$ tali che $x_h < x < x_k$. Dunque si ha

$$\varphi(x) - \varphi(x_h) < \varepsilon, \quad \varphi(x_k) - \varphi(x) < \varepsilon,$$

e quindi

$$\varphi(x) - \varepsilon < \varphi(x_h) \leq \varphi(x) \leq \varphi(x_k) < \varphi(x) - \varepsilon.$$

Ora $\lim\limits_{n\to\infty} \varphi_{n,n}(x_h) = \varphi(x_h)$ e $\lim\limits_{n\to\infty} \varphi_{n,n}(x_k) = \varphi(x_k)$, quindi esiste $\bar{n} \in \mathbb{N}$ tale che

$$\varphi(x) - \varepsilon < \varphi_{n,n}(x_h) \leq \varphi_{n,n}(x) \leq \varphi_{n,n}(x_k) < \varphi(x) + \varepsilon \quad \forall\, n \in \mathbb{N}, n > \bar{n},$$

ovvero

$$|\varphi_{n,n}(x) - \varphi(x)| < \varepsilon \quad \forall\, n \in \mathbb{N}, n > \bar{n}.$$

Dunque $\lim\limits_{n\to\infty} \varphi_{n,n}(x) = \varphi(x)$.

Supponiamo ora φ discontinua in x. Poiché φ è crescente, per il Teorema 1.4, i suoi punti di discontinuità costituiscono un insieme E al più numerabile. Allora, con lo stesso procedimento precedente, da $(\varphi_{n,n})_{n\in\mathbb{N}}$ si può estrarre una sottosuccessione (φ_{μ_n}) convergente in ogni punto di E. D'altra parte per ogni $x \in [a,b] \setminus E$ si ha $\lim\limits_{n\to\infty} \varphi_{\mu_n}(x) = \varphi(x)$. Poniamo allora

$$\widetilde{\varphi}(x) = \begin{cases} \varphi(x) & \forall\, x \in [a,b] \setminus E \\ \lim\limits_{n\to\infty} \varphi_{\mu_n}(x) & \forall\, x \in E. \end{cases}$$

La funzione $\widetilde{\varphi}(x)$ è anch'essa crescente su $[a,b]$. Infatti per ogni $x', x'' \in [a,b]$, con $x' < x''$, si ha $\varphi_{\mu_n}(x') \leq \varphi_{\mu_n}(x'')$ e passando al limite per n che tende a $+\infty$ si ha $\widetilde{\varphi}(x') \leq \widetilde{\varphi}(x'')$.

Adesso, dalla successione $(\psi_{\mu_n})_{n\in\mathbb{N}}$, con lo stesso procedimento già seguito, si può estrarre una sottosuccessione $(\psi_{\eta_n})_{n\in\mathbb{N}}$ convergente in ogni punto di $[a,b]$ a una funzione $\widetilde{\psi}$ crescente su $[a,b]$. Visto che $f_{\eta_n}(x) = \varphi_{\eta_n}(x) - \psi_{\eta_n}$, la successione $(f_{\eta_n})_{n\in\mathbb{N}}$ converge in ogni punto di $[a,b]$ a una funzione a variazione limitata. $\square$

2.3 Secondo Teorema di Helly

Teorema 2.7 (Secondo Teorema di Helly). *Sia $f : [a,b] \to \mathbb{R}$ continua e $(g_n(x))_{n\in\mathbb{N}}$ una successione di funzioni che converge a una funzione $g(x)$ in ogni punto di $[a,b]$. Se esiste $M > 0$ tale che $V_a^b(g_n) < M$ per ogni n, allora*

$$\lim_{n\to\infty} \int_a^b f \, \mathrm{d}g_n = \int_a^b f \, \mathrm{d}g.$$

Dimostrazione. Per prima cosa proviamo che $V_a^b(g) \leq M$. Presa $\{x_0, x_1, ..., x_m\}$ una partizione di $[a,b]$, si ha

$$\sum_{k=0}^{m-1} |g_n(x_{k+1}) - g_n(x_k)| < M,$$

e, per n che tende a $+\infty$, si deduce

$$\sum_{k=0}^{m-1} |g(x_{k+1}) - g(x_k)| \leq M.$$

Poiché la partizione era arbitraria, dall'ultima disequazione segue che $V_a^b(g) \leq M$. Poiché f è continua su un compatto, essa è uniformemente continua, cioè

$$\forall\, \varepsilon > 0 \quad \exists\, \delta(\varepsilon) > 0 \ : \qquad |f(x') - f(x'')| < \frac{\varepsilon}{3M} \quad \forall\, x', x'' \in [a,b], \ |x' - x''| < \delta.$$

Fissato $\varepsilon > 0$, consideriamo allora una partizione $\{x_0, x_1, ..., x_m\}$ di $[a,b]$ tale che $x_{k+1} - x_k < \delta$ per $k = 0, 1, \ldots, m-1$. Allora

$$\begin{aligned}\int_a^b f \, \mathrm{d}g &= \sum_{k=0}^{m-1} \int_{x_k}^{x_{k+1}} f(x) \, \mathrm{d}g \\ &= \sum_{k=0}^{m-1} \int_{x_k}^{x_{k+1}} (f(x) - f(x_k)) \, \mathrm{d}g + \sum_{k=0}^{m-1} f(x_k) \int_{x_k}^{x_{k+1}} \mathrm{d}g.\end{aligned}$$

Si ha

$$\begin{aligned}\left| \sum_{k=0}^{m-1} \int_{x_k}^{x_{k+1}} (f(x) - f(x_k)) \, \mathrm{d}g \right| &\overset{1.18}{\leq} \sum_{k=0}^{m-1} \|f(x) - f(x_k)\|_\infty V_{x_k}^{x_{k+1}}(g) \\ &< \frac{\varepsilon}{3M} \sum_{k=0}^{m-1} V_{x_k}^{x_{k+1}}(g) = \frac{\varepsilon}{3M} V_a^b(g) \leq \frac{\varepsilon}{3}.\end{aligned}$$

Inoltre $\int\limits_{x_k}^{x_{k+1}} \mathrm{d}g \overset{1.10}{=} g(x_{k+1}) - g(x_k)$. Perciò esiste $\eta \in \mathbb{R}$, con $|\eta| < 1$, tale che

$$\int_a^b f \, \mathrm{d}g = \sum_{k=0}^{m-1} f(x_k) \left[g(x_{k+1}) - g(x_k) \right] + \frac{\varepsilon}{3}\eta. \tag{2.10}$$

Analogamente, per ogni $n \in \mathbb{N}$ esiste $\eta_n \in \mathbb{R}$, con $|\eta_n| < 1$, tale che

$$\int_a^b f \,\mathrm{d}g_n = \sum_{k=0}^{m-1} f(x_k)\,[g_n(x_{k+1}) - g_n(x_k)] + \frac{\varepsilon}{3}\eta_n. \tag{2.11}$$

Poiché $\lim\limits_{n\to\infty} g_n(x) = g$ per ogni $x \in [a, b]$, esiste n_0 tale che

$$\left|\sum_{k=0}^{m-1} f(x_k)\,[g_n(x_{k+1}) - g_n(x_k)] - \sum_{k=0}^{m-1} f(x_k)\,[g(x_{k+1}) - g(x_k)]\right| < \frac{\varepsilon}{3}, \tag{2.12}$$

per ogni $n > n_0$. Allora segue da (2.10), (2.11) e (2.12) che, per ogni $n > n_0$,

$$\left|\int_a^b f \,\mathrm{d}g_n - \int_a^b f \,\mathrm{d}g\right| < \varepsilon.$$

Perciò

$$\int_a^b f \,\mathrm{d}g = \lim_{n\to\infty} \int_a^b f \,\mathrm{d}g_n.$$

Questo conclude la prova. $\square$

Capitolo 3

Teorema di Riesz

3.1 Funzionali lineari

Definizione 3.1. *Sia X uno spazio normato su $\mathbb{R}$. Si definisce funzionale lineare ogni funzione $\phi : X \to \mathbb{R}$ tale che*

1) $\phi(x_1 + x_2) = \phi(x_1) + \phi(x_2)$;

2) $\phi(ax) = a\phi(x)$*;*

per ogni $x, x_1, x_2 \in X$ e per ogni $a \in \mathbb{R}$.

Teorema 3.2. *Sia X uno spazio normato su $\mathbb{R}$ e ϕ un funzionale lineare. Allora ϕ è continuo in ogni punto oppure in nessun punto di X. Inoltre ϕ è continuo su X se e solo se è limitato, cioè se esiste $M > 0$ tale che*

$$|\phi(x)| \leq M \|x\|, \quad \forall\, x \in X.$$

Dimostrazione. Supponiamo ϕ continua in $x_0 \in X$; allora

$$\forall\, \varepsilon > 0 \quad \exists\, \delta_\varepsilon > 0 : \qquad |\phi(x) - \phi(x_0)| < \varepsilon \quad \forall\, x \in X : \|x - x_0\| < \delta_\varepsilon.$$

Fissiamo ora ad arbitrio un punto $x_1 \in X$ e sia $\|x - x_1\| < \delta_\varepsilon$. Allora

$$\|(x + x_0 - x_1) - x_0\| < \delta_\varepsilon \Rightarrow |\phi(x) - \phi(x_1)| = |\phi(x + x_0 - x_1) - \phi(x_0)| < \varepsilon.$$

Ciò prova che ϕ è continuo in x_1 e quindi è provata la prima affermazione del teorema. Supponiamo ora $|\phi(x)| \leq M \|x\|$ per ogni $x \in X$; Poiché $\phi(0) = 0$ si ha

$$|\phi(x) - \phi(0)| \leq M \|x - 0\|,$$

e quindi ϕ è continuo in 0 e dunque lo è in ogni punto per la prima parte. Viceversa, sia ϕ continuo in 0; fissato un ε esiste allora un $\delta > 0$ tale che $|\phi(x)| < \varepsilon$ per $\|x\| < \delta$. Sia ora $x \neq 0$, consideriamo $y = \frac{\delta}{2\|x\|}x$; risulta

$$\|y\| = \frac{\delta}{2} < \delta \Rightarrow |\phi(y)| < \varepsilon \Leftrightarrow \frac{\delta}{2\,\|x\|}\,|\phi(x)| < \varepsilon \Leftrightarrow |\phi(x)| < \frac{2\varepsilon}{\delta}\,\|x\|\,.$$

Con ciò è provata la seconda affermazione. □

Definizione 3.3. *Se X è uno spazio normato su $\mathbb{R}$ e ϕ è un operatore lineare continuo da X a $\mathbb{R}$, si definisce norma di ϕ il numero*

$$\|\phi\| = \sup_{\|x\|=1} |\phi(x)| = \sup_{x\neq 0} \frac{|\phi(x)|}{\|x\|}$$

Osservazione 3.4. Se ϕ è un funzionale lineare limitato è evidente che $\|\phi\| \leq M$, infatti

$$\|\phi\| = \sup_{x\neq 0} \frac{|\phi(x)|}{\|x\|} \leq \sup_{x\neq 0} \frac{M\,\|x\|}{\|x\|} = M.$$

3.2 Teorema di Riesz

Osservazione 3.5. Lo spazio delle funzioni continue su intervallo compatto $C([a,b])$ dotato della norma $\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$ è uno spazio normato.

Teorema 3.6 (**Teorema di Riesz**)**.** *Sia ϕ un funzionale lineare limitato definito nello spazio $C([a,b])$. Allora esiste una funzione a variazione limitata g su $[a,b]$ tale che*

$$\phi(f) = \int_a^b f\,\mathrm{d}g, \quad \forall\, f \in C([a,b]).$$

Dimostrazione. La funzione $f \mapsto \int_a^b f\,\mathrm{d}g$ è un funzionale lineare per il Teorema 1.11 ed è limitato, dunque continuo, poiché per ogni f funzione continua nell'intervallo $[a,b]$ e per ogni funzione g a variazione limitata vale

$$\left|\int_a^b f\,\mathrm{d}g\right| \overset{1.18}{\leq} V_a^b(g)\,\|f\|_\infty\,.$$

Viceversa, supponiamo che ϕ sia un funzionale lineare continuo. A meno di un cambiamento affine della variabile x possiamo supporre $a = 0$ e $b = 1$. Ricordiamo l'Osservazione 2.2:

$$\forall\, x \in \mathbb{R} \quad \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1.$$

Per $x \in [0,1]$, ogni addendo della somma è non negativo. Quindi qualsiasi siano $\varepsilon_{k,n} = \pm 1$ per $k = 0, 1, \dots, n$, si ha

$$\left| \sum_{k=0}^{n} \varepsilon_{k,n} \binom{n}{k} x^k (1-x)^{n-k} \right| \leq 1.$$

Definendo $f_{k,n}(x) := \binom{n}{k} x^k (1-x)^{n-k}$ e $h_n := \sum\limits_{k=0}^{n} \varepsilon_{k,n} f_{k,n}$ per $x \in [0,1]$ e $k = 0, 1, \dots, n$. Quindi l'ultima disuguaglianza implica che $\|h_n\|_\infty \leq 1$. Perciò si ha

$$\left| \sum_{k=0}^{n} \varepsilon_{k,n} \phi\left(f_{k,n}\right) \right| = \left|\phi(h_n)\right|,$$

e poiché il funzionale lineare è limitato esiste M tale che

$$|\phi(h_n)| \leq M \|h_n\|_\infty \leq M.$$

Scegliendo gli $\varepsilon_{k,n} = \operatorname{sgn}(\phi(f_{k,n}))$, si ottiene che tutti gli addendi di $\sum\limits_{k=0}^{n} \varepsilon_{k,n} \phi(f_{k,n})$ sono non negativi, pertanto si ha

$$\sum_{k=0}^{n} |\phi(f_{k,n})| \leq M. \tag{3.1}$$

Per ogni $n \in \mathbb{N}$ consideriamo la funzione costante a tratti g_n così definita

$$g_n(x) = \begin{cases} 0 & x = 0 \\ \phi(f_{0,n}) & 0 < x < \frac{1}{n} \\ \phi(f_{0,n}) + \phi(f_{1,n}) & \frac{1}{n} \leq x < \frac{2}{n} \\ \vdots & \\ \sum\limits_{k=0}^{n-1} \phi(f_{k,n}) & \frac{n-1}{n} \leq x < 1 \\ \sum\limits_{k=0}^{n} \phi(f_{k,n}) & x = 1. \end{cases}$$

La funzione g_n è tale che $|g_n(x)| \leq M$ per ogni $n \in \mathbb{N}$, per ogni $x \in [0,1]$. Inoltre $V_0^1(g_n) \leq M$ per ogni $n \in \mathbb{N}$. Infatti, presa una partizione $P = \{x_0, x_1, ..., x_m\}$ di $[0,1]$, è ovvio che, se x_i, x_{i+1} (con $1 < i < m$) appartengono allo stesso intervallo su cui g_n è costante,

$$g_n(x_i) - g_n(x_{i-1}) = 0.$$

Posso quindi considerare solo partizioni che contengono al massimo un punto per ogni intervallo su cui g_n è costante. Allora definisco una funzione $l : \{0, 1, \dots, m\} \to \{0, 1, \dots, n\}$ dove $l(i)$ è il massimo indice della sommatoria associata a $g_n(x_i)$, quindi

$l(0) = 0$ e $l(m) = n$. Inoltre, per come consideriamo le partizioni e per com'è definita g_n, l è strettamente crescente. Pertanto

$$\begin{aligned}\sum_{i=1}^{m} |g_n(x_i) - g_n(x_{i-1})| &= \sum_{i=1}^{m} \left| \sum_{k=0}^{l(i)} \phi(f_{k,n}) - \sum_{k=0}^{l(i-1)} \phi(f_{k,n}) \right| \\ &= \sum_{i=1}^{m} \left| \sum_{k=l(i-1)+1}^{l(i)} \phi(f_{k,n}) \right| \\ &\leq \sum_{i=1}^{m} \sum_{k=l(i-1)+1}^{l(i)} |\phi(f_{k,n})|,\end{aligned}$$

e poiché l è strettamente crescente

$$\sum_{i=1}^{m} \sum_{k=l(i-1)+1}^{l(i)} |\phi(f_{k,n})| \leq \sum_{k=0}^{n} |\phi(f_{k,n})| \overset{(3.1)}{\leq} M.$$

Passando al sup si ha dunque $V_0^1(g_n) \leq M$. Per il Teorema 2.6 da $(g_n)_{n\in\mathbb{N}}$ si può estrarre una sottosuccessione $(g_{n_k})_{k\in\mathbb{N}}$ convergente in ogni punto di $[0,1]$ a una funzione g a variazione limitata.

Sia $f : [0,1] \to \mathbb{R}$, $f \in C([0,1])$; allora per il Teorema 1.19,

$$\begin{aligned}\int_0^1 f \,\mathrm{d}g_n &= f(0)\left[g_n(0^+) - g_n(0)\right] + \sum_{k=1}^{n-1} f\left(\tfrac{k}{n}\right)\left[g_n(\tfrac{k}{n}^+) - g_n(\tfrac{k}{n}^-)\right] + f(1)\left[g_n(1) - g_n(1^-)\right] \\ &= f(0)\phi(f_{0,n}) + \sum_{k=1}^{n-1} f(\tfrac{k}{n})\phi(f_{k,n}) + f(1)\phi(f_{n,n}) \\ &= \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right)\phi(f_{k,n}) = \phi\left[B_n(x)\right],\end{aligned}$$

dove B_n è il polinomio di Bernstein di grado n associato a f. Per il Teorema di Bernstein 2.4, $\lim_{n\to\infty} \|f - B_n\|_\infty = 0$. Perciò

$$\lim_{n\to\infty} |\phi(B_n) - \phi(f)| \leq \lim_{n\to\infty} M \|B_n - f\|_\infty = 0.$$

Dunque

$$\lim_{n\to\infty} \int_0^1 f \,\mathrm{d}g_n = \phi(f).$$

Per il Teorema 2.7 si ha quindi

$$\phi(f) = \lim_{n\to\infty} \int_0^1 f \,\mathrm{d}g_n = \int_0^1 f \,\mathrm{d}g.$$

Questo conclude la prova. □

Osservazione 3.7. Da $\left|\int_0^1 f\,\mathrm{d}g\right| \le V_0^1(g)\,\|f\|_\infty$, si ha $|\phi(f)| \le V_0^1(g)\,\|f\|_\infty$, e quindi $\|\phi\| \le V_0^1(g)$. D'altra parte si ha la sottosuccessione $(g_{n_k})_{k\in\mathbb{N}}$ che, come provato nella dimostrazione, è tale che $V_0^1(g_{n_k}) \le M$ e dunque, per l'Osservazione 3.4, $V_0^1(g_{n_k}) \le \|\phi\|$. Presa una partizione $P = \{x_0, x_1, ..., x_m\}$ di $[0,1]$ arbitraria, si ha

$$\sum_{i=1}^{m} |g_{n_k}(x_i) - g_{n_k}(x_{i-1})| \le \|\phi\|,$$

e quindi, per k che tende a $+\infty$,

$$\sum_{i=1}^{m} |g(x_i) - g(x_{i-1})| \le \|\phi\|,$$

da cui segue $V_0^1(g) \le \|\phi\|$. Perciò $\|\phi\| = V_0^1(g)$.

Osservazione 3.8. La funzione g non è univocamente determinata. Infatti dall'Osservazione 1.17 segue che se $\phi(f) = \int_0^1 f\,\mathrm{d}g$ per ogni funzione $f \in C([0,1])$, allora anche $\phi(f) = \int_0^1 f\,\mathrm{d}h$ per ogni funzione $f \in C([0,1])$ per ogni h a variazione limitata su $[0,1]$ tale che $g(1) - h(1) = g(0) - h(0)$ e $g(x) - h(x) = g(0) - h(0)$ in ogni punto $x \in (0,1)$ ove g e h sono continue.

Osservazione 3.9. Nella dimostrazione del teorema di Riesz, si fa notare l'importanza dell'uso specifico del polinomio di Bernstein come polinomio approssimante. Anche se il Teorema di approsimazione di Weierstrass garantisce l'esistenza di un polinomio approssimante uniformemente per ogni funzione continua su un intervallo compatto, il polinomio di Bernstein è particolarmente significativo in questa dimostrazione. Questo perché i coefficienti del polinomio di Bernstein hanno come termini le immagini della funzione f in alcuni punti e ciò ci consente di ricondurci direttamente alla definizione di integrale di Riemann-Stieltjes. Pertanto, l'impiego del polinomio di Bernstein risulta essenziale nel connettere l'integrale di Riemann-Stieltjes e la funzione ϕ.

Capitolo 4

Dimostrazione Alternativa e Generalizzazione del Teorema di Riesz

4.1 Dimostrazione alternativa con Hahn-Banach

Definizione 4.1. *Dato un insieme parzialmente ordinato $(A, \preceq)$, si dice catena ogni sottoinsieme $B \subseteq A$ tale che la relazione d'ordine ridotta a B costituisce un insieme totalmente ordinato.*

Definizione 4.2. *Sia $(A, \preceq)$ un insieme parzialmente ordinato, si dice elemento massimale l'elemento $a \in A$ tale che, per ogni $x \in A$, $a \leq x$ implica $x = a$.*

Lemma 4.3 (**Lemma di Zorn**). *Se A è un insieme parzialmente ordinato in cui ogni catena ha un maggiorante, allora A ha un elemento massimale.*

Teorema 4.4 (**Teorema di Hahn-Banach**). *Sia Z un sottospazio di uno spazio vettoriale X su $\mathbb{R}$ e sia p una funzione a valori reali su X tale che per ogni $x, y \in X$*

$$p(x + y) \leq p(x) + p(y) \qquad e \qquad p(\alpha x) = \alpha p(x) \quad per\ \alpha \geq 0.$$

Se f è un funzionale lineare su Z e $f(x) \leq p(x)$ per ogni $x \in Z$, allora esiste un funzionale lineare F su X tale che $F(x) = f(x)$ per $x \in Z$ e $F(x) \leq p(x)$ per ogni $x \in X$.

Dimostrazione. Sia $\mathcal{F}$ l'insieme delle estensioni lineari g di f tali che $g \leq p$ su D_g (dominio di g); $\mathcal{F}$ è parzialmente ordinato definendo la relazione

$$g \preceq g' \Longleftrightarrow D_g \subseteq D_{g'},\ g(x) = g'(x) \quad \forall\, x \in D_g.$$

Si noti che $\mathcal{F} \neq \varnothing$ poiché $f \in \mathcal{F}$. Se $\{g_s : s \in S\}$ è una catena in $\mathcal{F}$, sia g un funzionale lineare definito su $\bigcup\limits_{s \in S} D_{g_s}$ ponendo $g(x) = g_s(x)$ per $x \in D_{g_s}$. La buon positura di g segua dal fatto che $\{g_s : s \in S\}$ è una catena; infatti questo implica che per ogni $s, t \in S$, o $g_s \leq g_t$ o $g_t \leq g_s$. Di conseguenza, se $x \in D_{g_s} \cap D_{g_t}$, allora $g_s(x) = g_t(x)$. Pertanto, la definizione $g(x) = g_s(x)$ è indipendente dalla scelta di g_s. Quindi g è funzionale lineare tale che $g \leq p$ su D_g. Perciò, ogni catena in $\mathcal{F}$ ha un maggiorante e, di conseguenza, per il Lemma di Zorn, $\mathcal{F}$ ha un elemento massimale che chiamiamo F. Per completare la dimostrazione rimane da mostrare che il dominio di F è X.

Supponiamo per assurdo che $D_F \neq X$ e prendiamo $x_0 \in X \setminus D_F$. Allora, per $y, z \in D_F$

$$F(z) - F(y) = F(z - y) \leq p(z - y) = p(z + x_0 - y - x_0) \leq p(z + x_0) + p(-y - x_0).$$

Quindi, per ogni $y, z \in D_F$

$$-p(-y - x_0) - F(y) \leq p(z + x_0) - F(z).$$

Prendiamo β_0 tale che

$$\sup_{y \in D_F} \left[-p(-y - x_0) - F(y)\right] \leq \beta_0 \leq \inf_{z \in D_F} \left[p(z + x_0) - F(z)\right].$$

Definiamo un funzionale lineare sul sottospazio generato da $D_F \cup \{x_0\}$ come segue:

$$g(x + \alpha x_0) := F(x) + \alpha \beta_0, \quad x \in D_F,\ \alpha \in \mathbb{R}.$$

Allora, per $\alpha > 0$,

$$\begin{aligned} g(x + \alpha x_0) &= F(x) + \alpha \beta_0 \leq F(x) + \alpha \left[p(\tfrac{x}{\alpha} + x_0) - F(\tfrac{x}{\alpha})\right] \\ &= F(x) + \alpha p(\tfrac{x}{\alpha} + x_0) - \alpha F(\tfrac{x}{\alpha}) = F(x) + p(x + \alpha x_0) - F(x) \\ &= p(x + \alpha x_0) \end{aligned}$$

e, per $\alpha < 0$,

$$\begin{aligned} g(x + \alpha x_0) &= F(x) + \alpha \beta_0 \leq F(x) + \alpha \left[-p(-\tfrac{x}{\alpha} - x_0) - F(\tfrac{x}{\alpha})\right] \\ &= F(x) - \alpha p(-\tfrac{x}{\alpha} - x_0) - \alpha F(\tfrac{x}{\alpha}) = F(x) + p(x + \alpha x_0) - F(x) \\ &= p(x + \alpha x_0) \end{aligned}$$

Perciò g è un'estensione di F che appartiene a $\mathcal{F}$ ma è diversa da F poiché $D_F \subsetneq D_g$. Poiché F è un elemento massimale di $\mathcal{F}$ si ha un assurdo, dunque $D_F = X$. $\square$

Corollario 4.5. *Se f è un funzionale lineare limitato sul sottospazio Z dello spazio normato reale X, allora esiste un funzionale lineare limitato F su X tale che $F = f$ su Z e $\|F\| = \|f\|$.*

Dimostrazione. Poniamo $p(x) = \|f\| \, \|x\|$; allora per ogni $x, y \in X$

$$p(x+y) \leq p(x) + p(y) \quad \text{e} \quad p(\alpha x) = \alpha p(x) \quad \text{per}\, \alpha \geq 0.$$

Vale anche che, per ogni $x \in Z$,

$$f(x) \leq |f(x)| \leq \|f\| \, \|x\| = p(x).$$

Perciò, il Teorema 4.4 implica che esiste un funzionale lineare F su X tale che $F = f$ su Z e $F(x) \leq p(x)$ per ogni $x \in X$. Da questo si ha

$$-F(x) = F(-x) \leq p(-x) = p(x) \Rightarrow |F(x)| \leq \|f\| \, \|x\| \quad \forall\, x \in X.$$

Ne segue allora $\|F\| \leq \|f\|$.

Inoltre, $F = f$ su Z implica

$$|f(x)| = |F(x)| \leq \|F\| \, \|x\| \quad \forall\, x \in Z,$$

e dunque $\|f\| \leq \|F\|$. Si è quindi provato che $\|F\| = \|f\|$. □

Vediamo ora una dimostrazione alternativa del teorema di Riesz di stampo più vicino all'Analisi Funzionale.

Teorema 4.6 (**Teorema di Riesz**). *Sia ϕ un funzionale lineare limitato definito nello spazio $C([a,b])$. Allora esiste una funzione a variazione limitata g su $[a,b]$ tale che*

$$\phi(f) = \int_a^b f \, \mathrm{d}g, \quad \forall\, f \in C([a,b]).$$

Dimostrazione. Ricordiamo che $C([a,b])$ è un sottospazio dello spazio normato $B([a,b])$ delle funzioni limitate su $[a,b]$ con norma $\|.\|_\infty$. Il Corollario 4.5 di Hahn-Banach implica che ϕ può essere esteso a $B[a,b]$. Il funzionale lineare esteso, che denotiamo con $\widetilde{\phi}$, avrà la stessa norma di ϕ.

Definiamo una funzione g su $[a,b]$ come:

$$g(x) = \begin{cases} 0 & x = a \\ \widetilde{\phi}(\chi_{[a,x)}) & a < x < b \\ \widetilde{\phi}(\chi_{[a,b]}) & x = b. \end{cases}$$

Per come è definita g vale, per ogni $x, y \in [a,b]$ con $x < y$,

$$g(y) - g(x) = \widetilde{\phi}(\chi_{[x,y)}) \quad \text{se } y < b \quad \text{e} \quad g(y) - g(x) = \widetilde{\phi}(\chi_{[x,b]}) \quad \text{se } y = b. \tag{4.1}$$

Presa un partizione $P = \{x_0, x_1, ..., x_m\}$ di $[a, b]$, allora, ponendo $\sigma_i = \operatorname{sgn}(g(x_i) - g(x_{i-1}))$, si ha

$$\begin{aligned}\sum_{i=1}^{m} |g(x_i) - g(x_{i-1})| &= \widetilde{\phi}\left[\sigma_m \chi_{[x_{m-1},b]} + \sum_{i=1}^{m-1} \sigma_i \chi_{[x_{i-1},x_i)}\right] \\ &\leq \left\|\widetilde{\phi}\right\|_{B([a,b])} \left\|\sigma_m \chi_{[x_{m-1},b]} + \sum_{i=1}^{m-1} \sigma_i \chi_{[x_{i-1},x_i)}\right\|_{\infty}.\end{aligned} \tag{4.2}$$

Dato che P è una partizione, per ogni $x \in [a, b]$, x appartiene a un solo sottointervallo $[x_{i-1}, x_i)$ o $[x_{m-1}, b]$. Quindi, senza perdere di generalità, considerando $x \in [x_{i-1}, x_i)$, si ha

$$\left|\sigma_m \chi_{[x_{m-1},b]}(x) + \sum_{i=1}^{m-1} \sigma_i \chi_{[x_{i-1},x_i)}(x)\right| = |\sigma_i|.$$

Supponendo che i σ_i non siano tutti nulli (altrimenti la disuguaglianza sarebbe banale), si ha $|\sigma_i| = 1$ e quindi $\left\|\sigma_m \chi_{[x_{m-1},b]} + \sum_{i=1}^{m-1} \sigma_i \chi_{[x_{i-1},x_i)}\right\|_{\infty} = 1$. Pertanto

$$\begin{aligned}\sum_{i=1}^{m} |g(x_i) - g(x_{i-1})| &\overset{(4.2)}{\leq} \left\|\widetilde{\phi}\right\|_{B([a,b])} \left\|\sigma_m \chi_{[x_{m-1},b]} + \sum_{i=1}^{m-1} \sigma_i \chi_{[x_{i-1},x_i)}\right\|_{\infty} \\ &= \left\|\widetilde{\phi}\right\|_{B([a,b])} = \|\phi\|_{C([a,b])}.\end{aligned}$$

Questo prova che g è a variazione limitata e

$$V_a^b(g) \leq \|\phi\|. \tag{4.3}$$

Per ogni $f \in C([a, b])$ e $\varepsilon > 0$, scegliamo $\delta > 0$ tale che $|x - y| < \delta$ implica $|f(x) - f(y)| < \frac{\varepsilon}{2}$ per ogni $x, y \in [a, b]$. Questo è possibile poiché f è uniformemente continua. Usando il fatto che $f \in S(g)$ per il Teorema 1.16, scegliamo una qualsiasi partizione P di $[a, b]$ con $\|P\| < \delta$ tale che

$$\left|S(P, f, g) - \int_a^b f \,\mathrm{d}g\right| \leq \frac{\varepsilon \|\phi\|}{2}.$$

Definiamo una funzione $s \in B([a, b])$ così:

$$s = f(t_{m-1})\chi_{[x_{m-1},b]} + \sum_{i=1}^{m-1} f(t_{i-1})\chi_{[x_{i-1},x_i)}.$$

Per (4.1), si ha

$$\begin{aligned}\widetilde{\phi}(s) &= f(t_{m-1})\widetilde{\phi}(\chi_{[x_{m-1},b]}) + \sum_{i=1}^{m-1} f(t_{i-1})\widetilde{\phi}(\chi_{[x_{i-1},x_i)}) \\ &= f(t_{m-1})(g(b) - g(x_{m-1})) + \sum_{i=1}^{m-1} f(t_{i-1})(g(x_i) - g(x_{i-1})) \\ &= \sum_{i=1}^{m} f(t_{i-1})(g(x_i) - g(x_{i-1}))\end{aligned}$$

È facile comprendere dall'ipotesi di continuità uniforme di f che $\|f - s\|_\infty \leq \frac{\varepsilon}{2}$ o equivalentemente $|f(y) - s(y)| < \frac{\varepsilon}{2}$ per ogni $y \in [a, b]$. Pertanto,

$$\begin{aligned}\left|\phi(f) - \int_a^b f \, \mathrm{d}g\right| &\leq \left|\phi(f) - \widetilde{\phi}(s)\right| + \left|\sum_{i=1}^{m} f(t_{i-1})(g(x_i) - g(x_{i-1})) - \int_a^b f \, \mathrm{d}g\right| \\ &\leq \left|\phi(f) - \widetilde{\phi}(s)\right| + \frac{\varepsilon \|\phi\|}{2} = \left|\widetilde{\phi}(f - s)\right| + \frac{\varepsilon \|\phi\|}{2} \\ &\leq \|\phi\| \, \|f - s\|_\infty + \frac{\varepsilon \|\phi\|}{2} \leq \varepsilon \|\phi\| \, .\end{aligned}$$

Poiché f e ε sono arbitrari, segue che $\phi(f) = \int_a^b f \, \mathrm{d}g$ per ogni $f \in C([a, b])$. Inoltre dal Lemma 1.18, $\|\phi\| \leq V_a^b(g)$ e ricordando (4.3) ne deduciamo

$$\|\phi\| = V_a^b(g).$$

La tesi è così verificata. □

4.2 Una generalizzazione del Teorema di Riesz

In questa sezione, verrà presentata una versione alternativa del Teorema di Riesz che estende le ipotesi alle funzioni continue definite su un insieme compatto arbitrario. L'integrale di Riemann-Stieltjes è definito su un intervallo compatto. Pertanto, nel teorema considereremo l'integrale di Lebesgue, che è definito per funzioni continue su un qualsiasi insieme compatto.

Osservazione 4.7. Si può provare che ogni misura di Borel in uno spazio metrico è regolare. Nella trattazione verrà usata solo la regolarità interna, ovvero, per ogni insieme di Borel A e ogni $\varepsilon > 0$ esiste un insieme compatto F_ε tale che $F_\varepsilon \subseteq A$ e $\mu(A \setminus F_\varepsilon) < \varepsilon$. [8] Teorema 7.1.7.

Definizione 4.8. *Una funzione* $\alpha : [a, b] \to \mathbb{R}$ *si dice normalizzata, se* $\alpha(a) = 0$ *ed* α *è continua da destra nell'intervallo* (a, b).

Lemma 4.9. *Data una funzione* α *monotona e normalizzata nell'intervallo* $[a, b]$*, allora esiste unica misura* μ *di Borel associata ad essa in modo tale che valga*

$$\int_a^b f \, \mathrm{d}\alpha = \int_{[a,b]} f \, \mathrm{d}\mu, \quad \forall\, f \in C([a, b]).$$

Dimostrazione. Per $a \leq s \leq t \leq b$ sia

$$\mathcal{F}_0 = \left\{ \bigcup_{\text{finita}} (s_k, t_k] : (s_k, t_k] \subset [a, b] \text{ due a due disgiunti} \right\}.$$

$\mathcal{F}_0$ è un'algebra di sottoinsiemi di $[a,b]$ e perciò possiamo definire una funzione sui suoi insiemi come

$$\mu_0\left(\bigcup_{\text{finita}}(s_k,t_k]\right)=\sum_{\text{finita}}\alpha(t_k)-\alpha(s_k).$$

Inoltre, μ_0 si estende in modo unico a una misura nella più piccola σ-algebra che contiene $\mathcal{F}_0$ per il Teorema di estensione Carathéodory. Si può provare che, per ogni funzione continua f, vale

$$\int_a^b f\,\mathrm{d}\alpha=\int\limits_{[a,b]} f\,\mathrm{d}\mu,$$

dove l'integrale di sinistra è un integrale di Riemann-Stieltjes, mentre l'integrale sulla destra è un integrale nel senso di Lebesgue. □

Definizione 4.10. *Un funzionale lineare $L : C(K)\to\mathbb{R}$ si dice positivo se $L(f)\geq 0$ per ogni $f\geq 0$.*

Osservazione 4.11. La positività di un funzionale lineare L implica la continuità di L. Infatti, presa la funzione caratteristica χ_K su K, essa è continua su K. Inoltre si ha $|f(x)|\leq\|f\|_\infty\chi_K$, perciò

$$\|f\|_\infty\chi_K\pm f(x)\geq 0\Rightarrow\|f\|_\infty L(\chi_K)\pm L(f)\geq 0,$$

dunque $|L(f)|\leq L(\chi_K)\|f\|_\infty$.

Teorema 4.12 (Teorema di Riesz generalizzato). *Sia $K\subseteq\mathbb{R}$ compatto e sia $l : C(K)\to\mathbb{R}$ un funzionale lineare positivo. Allora esiste una unica misura di Borel finita μ su K tale che $\mu(K)=\|l\|$ e*

$$l(f)=\int_K f\,\mathrm{d}\mu,\quad\forall f\in C(K).$$

Dimostrazione. La dimostrazione si svolge in tre fasi.

(i) *Rappresentazione integrale.* Sia $[a,b]\supseteq K$. Sia $r : C([a,b])\to C(K)$ l'operatore di restrizione, cioè, per ogni $f\in C([a,b])$, $r(f)(x)=f(x)$ per $x\in K$. Si osservi che r è un operatore lineare limitato e, quindi, ha senso considerare il suo aggiunto, definito come segue $r^t : C(K)^*\to C([a,b])^*$, dove $r^t(l)(f)=l(r(f))$ per $f\in C([a,b])$. Preso l un funzionale lineare positivo in $C(K)$ e definiamo

$$L(f)=r^t(l)(f)=l(rf).$$

Poichè l e r sono funzionali lineari positivi, anche L lo è e esso è quindi continuo per l'Osservazione 4.11. Allora, per il Teorema di Riesz, esiste una funzione α monotona e normalizzata, come si deduce dalla definizione della funzione integratrice nella dimostrazione di questo capitolo. Mentre, per il Lemma 4.9, ad α è associata una misura μ di Borel tale che

$$L(f) = r^t(l)(f) = \int_a^b f \,\mathrm{d}\alpha = \int_{[a,b]} f \,\mathrm{d}\mu,$$

per ogni $f \in C([a,b])$. Denotiamo $K^c := [a,b] \setminus K$.

Sia $\varepsilon > 0$ e, per l'Osservazione 4.7, esiste $F_\varepsilon \subseteq K^c$ tale che $\mu(K^c \setminus F_\varepsilon) < \varepsilon$. Consideriamo la funzione

$$\widetilde{f}(x) = \frac{d(x, F_\varepsilon)}{d(x, F_\varepsilon) + d(x, K)},$$

dove $d(x, A) = \inf_{y \in A} |x - y|$. $\widetilde{f} \in C([a,b])$ ed è tale che $\widetilde{f}(x) = 1$ se $x \in K$, $\widetilde{f}(x) = 0$ se $x \in F_\varepsilon$ e $\left\|\widetilde{f}\right\|_\infty \leq 1$. Allora

$$L(\widetilde{f}) = \int_{[a,b]} \widetilde{f} \,\mathrm{d}\mu = \int_K \mathrm{d}\mu + \int_{K^c \setminus F_\varepsilon} \widetilde{f} \,\mathrm{d}\mu + \underbrace{\int_{F_\varepsilon} \widetilde{f} \,\mathrm{d}\mu}_{0}. \tag{4.4}$$

Segue che

$$0 \leq \int_{K^c \setminus F_\varepsilon} \widetilde{f} \,\mathrm{d}\mu \leq \int_{K^c \setminus F_\varepsilon} \mathrm{d}\mu = \mu(K^c \setminus F_\varepsilon) < \varepsilon. \tag{4.5}$$

Dunque da (4.4) e (4.5) si ottiene

$$L(\widetilde{f}) < \int_K \mathrm{d}\mu + \varepsilon = \mu(K) + \varepsilon.$$

Pertanto si ha

$$\mu(K) + \mu(K^c) = \int_{[a,b]} \mathrm{d}\mu = L(\chi_{[a,b]}) = L(\widetilde{f}) < \mu(K) + \varepsilon,$$

dove l'ultima uguaglianza segue da $r(\widetilde{f}) = r(\chi_{[a,b]})$. Poiché $\mu(K) < \infty$, vale $0 \leq \mu(K^c) < \varepsilon$ e di conseguenza, per l'arbitrarietà di ε, $\mu(K^c) = 0$.

Per concludere, sia $f \in C(K)$ e f^* un'estensione continua di f all'intervallo $[a,b]$. È possibile prendere tale estensione, per esempio, in questo modo: poiché K^c è un sottoinsieme aperto di $[a,b]$, è al più un'unione numerabile di intervalli aperti (α_i, β_i) intersecati con l'intervallo $[a,b]$ a due a due disgiunti. Per $x \in (\alpha_i, \beta_i)$ definiamo

$$f^*(x) = (1-t)f(\alpha_i) + tf(\beta_i),$$

se $x = \alpha_i(1-t) + t\beta_i$ per $t \in (0,1)$. f^* è continua nell'intervallo $[a,b]$ perché su K coincide con f e su K^c consiste in una retta a tratti. Allora si ha

$$l(f) = l(r(f^*)) = L(f^*) = \int_a^b f^* \, \mathrm{d}\alpha = \int_{[a,b]} f^* \, \mathrm{d}\mu = \int_K f^* \, \mathrm{d}\mu = \int_K f \, \mathrm{d}\mu, \tag{4.6}$$

come volevasi dimostrare.

(ii) *Norma.* Consideriamo $f \in C(K)$ tale che $\|f\|_\infty = 1$. Da quanto provato prima, si ha

$$|l(f)| = \left| \int_K f \, \mathrm{d}\mu \right| \leq \|f\|_\infty \, \mu(K) = \mu(K),$$

e passando al sup si trova $\|l\| \leq \mu(K)$. D'altra parte

$$\|l\| \geq |l(\chi_K)| = \left| \int_K \chi_K \, \mathrm{d}\mu \right| = \mu(K).$$

Possiamo quindi concludere che $\mu(K) = \|l\|$.

(iii) *Unicità.* Supponiamo esistano due misure finite, μ e ν, che soddisfano la tesi. Poiché μ e ν sono misure regolari, basterà mostrare che $\mu(C) = \nu(C)$ per ogni sottoinsieme chiuso C di K. Quindi sia $C \neq \emptyset$ e $C \subseteq K$; poniamo

$$f_n(x) := \max\{0, 1 - n \cdot d(x, C)\} \quad \forall\, n \in \mathbb{N} \,\forall\, x \in K.$$

Queste funzioni sono continue e limitate da 0 e 1. Perciò $f_n \in C(K)$ per ogni n. Si noti che (f_n) è una successione decrescente che tende a χ_C. Dunque, per ogni n, si ha $\int_K f_n \, \mathrm{d}\mu = \int_K f_n \, \mathrm{d}\nu$, e usando la convergenza dominata si conclude che

$$\mu(C) = \lim_{k\to\infty} \int_K f_n \, \mathrm{d}\mu = \lim_{k\to\infty} \int_K f_n \, \mathrm{d}\nu = \nu(C),$$

quindi μ e ν sono la stessa misura.

La prova è così conclusa. □

Osservazione 4.13. È possibile rappresentare i funzionali positivi che agiscono su $C(K)$ come integrali di Riemann-Stieltjes. Ciò segue immediatamente dalla catena di equazioni (4.6). Si faccia attenzione al fatto che non è possibile usare direttamente f per definire l'integrale di Riemann-Stieljes, ma una qualsiasi estensione di f funziona. Questo integrale è indipendente dall'estensione di f.

Bibliografia

[1] T. M. Apostol, *Mathematical Analysis*, 2nd ed., Addison-Wesley, Massachusetts, USA, 1975.

[2] B. Pini: *Secondo corso di analisi matematica*, Bologna : Cooperativa libraria universitaria, 1972.

[3] I. P. Natanson: *Theory of functions of a real variable*, Tradotto dal russo da Leo F. Boron ; Con la collaborazione editoriale di, e con note di by Edwin Hewitt, New York : F. Ungar, 1955.

[4] P. Cannarsa, T. D'aprile, *Introduzione alla teoria della misura e all'analisi funzionale*, Springer-Verlag Italia, Milano, 2008.

[5] N. B. Haaser, J. A. Sullivan, *Real analysis*, New York : Van Nostrand Reinhold, 1971.

[6] A. Narkawicz: *A Formal Proof Of The Riesz Representation Theorem* NASA Langley Research Center, Journal of Formalized Reasoning Vol. 4, No. 1, 2011.

[7] R. Del Rio, A. L. Franco, J. A. Lara: *A short proof of F. Riesz representation Theorem*, Universidad Nacional Autónoma de México, 2017.

[8] V. I. Bogachev, *Measure theory. Vol. II*, Springer-Verlag, Berlino, 2007.