SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
Corso di Laurea Triennale in Matematica

Alcuni sviluppi del Teorema di Weierstrass

Relatore:
Prof. Giovanni Maria TROIANIELLO

Candidato:
Cesare Giulio ARDITO
Matricola 1388090

Sessione Autunnale
Anno Accademico 2012-2013
Dipartimento di Matematica 'Guido Castelnuovo'

Introduzione

La teoria dell'approssimazione è un settore dell'analisi matematica che si occupa di stabilire se e in che modo sia possibile approssimare funzioni con altre funzioni più semplici da esaminare o da calcolare. L'interesse per questo tipo di problemi si sviluppò nel diciannovesimo secolo, quando si passò dal definire una funzione solo esplicitamente al definirla anche tramite le sue proprietà. Infatti in mancanza di una formulazione esplicita diventa essenziale riuscire ad approssimare la funzione per poterla utilizzare in qualsiasi problema pratico.
La domanda a cui si cerca di dare una risposta è dunque se, dato un insieme di funzioni, sia possibile con esse approssimarne uno più grande. Questo è naturalmente legato al concetto di densità, e dunque l'intera teoria dell'approssimazione si riduce a cercare sottoinsiemi di funzioni "semplici" densi in insiemi più grandi.
La questione è, inoltre, intimamente legata alla regolarità delle funzioni in questione. Quando nel 1861 Weierstrass esibì il famoso esempio di una funzione ovunque continua e derivabile in nessun punto[a], molti matematici iniziarono a dubitare che la sola continuità fosse sufficiente a garantire il "buon comportamento" di una funzione. Esempi come quello di Weierstrass rappresentavano una *deplorevole epidemia*[b] ai danni della continuità per gli analisti dell'epoca, ma fu Weierstrass stesso a fornire una sorta di compensazione a questa mancanza, dimostrando nel 1881 la densità dei polinomi algebrici nell'insieme delle funzioni continue reali su un intervallo compatto. Dunque la presenza nell'insieme delle funzioni continue di elementi mai derivabili è compensata dal fatto che ognuno di essi può essere approssimato con precisione arbitraria da polinomi, classe di funzioni tra

[a] $f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$, con $0 < a < 1, b$ intero positivo dispari, $ab > 1 + \frac{3\pi}{2}$

[b] *"I recoil with dismay and horror at this lamentable plague of functions which do not have derivatives"*, "Mi ritraggo con sgomento e orrore di fronte a questa deplorevole epidemia di funzioni che non possiedono derivate" (Charles Hermite)

le più regolari.
Al teorema di approssimazione di Weierstrass sono seguite moltissime generalizzazioni riguardanti altre famiglie di funzioni o altri spazi, che hanno portato a una vasta quantità di risultati. Nella presente tesi proporremo il teorema e la dimostrazione così come li pubblicò il matematico tedesco nell'articolo originale; esamineremo poi altre dimostrazioni del teorema, includendo anche l'analogo caso trigonometrico, citando in particolare il lavoro di Bernstein e Fejér al riguardo. Vedremo qui uno dei tanti esempi di come in matematica esistano moltissime dimostrazioni della stessa proposizione, e come ognuna di esse apra nuove interessanti prospettive di applicazione e sviluppo. Dopo aver esaminato brevemente qualche generalizzazione del teorema, analizzeremo una delle sue conseguenze più importanti in una diverso approccio all'approssimazione, quello tramite funzionali. In questo contesto dimostreremo il teorema di Bohman-Korovkin, che caratterizza gli operatori lineari che generano successioni (non necessariamente polinomiali) approssimanti.
L'ultimo capitolo riguarderà invece il problema di cercare ulteriori sottoinsiemi densi nei polinomi, raffinando in tal modo l'insieme delle funzioni approssimanti. Proporremo due diverse dimostrazioni del teorema di Müntz, che caratterizza tali sottoinsiemi con una condizione insieme semplice ed elegante. Una di esse, dovuta a Rudin, utilizzerà strumenti di analisi complessa, evidenziando un sorprendente collegamento tra questo problema e la distribuzione degli zeri di una funzione olomorfa. Concluderemo, infine, analizzando qualche possibile generalizzazione del teorema.

Indice

Capitolo 1

Il teorema di Weierstrass

1.1 Preliminari

Prima di procedere a illustrare l'enunciato e la prima dimostrazione del teorema di Weierstrass, è opportuno ricordare alcune definizioni essenziali per la piena comprensione dei risultati che seguiranno. Fondamentale è il concetto di spazio normato.

Definizione 1.1.1. Uno spazio normato è una coppia $(V, \|.\|)$ dove V è uno spazio vettoriale su un campo K, e $\|.\|$ è una funzione da V in $\mathbb{R}^+$ con le seguenti proprietà:

- $\|x\| = 0 \iff x = 0$
- $\|\alpha x\| = |\alpha| \, \|x\| \quad \forall \alpha \in K, x \in V$
- $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$

Definizione 1.1.2. Due norme $\|.\|_1$ e $\|.\|_2$ su uno stesso spazio V si dicono equivalenti se esistono due costanti c e C tali che

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$$

$\forall x \in V$.

Due norme equivalenti inducono la stessa topologia. Inoltre due norme su uno spazio vettoriale di dimensione finita sono sempre equivalenti, in particolare dunque ogni spazio vettoriale di dimensione finita possiede la topologia standard di $\mathbb{R}^n$.

Ciò è falso in dimensione infinita: vedremo a breve un esempio di due norme non equivalenti.

Sebbene molti tra i risultati che illustreremo si possano considerare come casi particolari di teoremi più generali, in questa tesi tratteremo solamente spazi (almeno) normati. Omettiamo dunque di indicare come molti teoremi si generalizzino a spazi metrici o addirittura solo topologici, rimandando a [Car00] per una trattazione più dettagliata.
Definiamo ora gli spazi normati che esamineremo, ricordando le loro proprietà e andando a evidenziare eventuali particolarità.

Definizione 1.1.3. Detto $C[a, b]$ lo spazio delle funzioni continue su un intervallo reale chiuso e limitato, è possibile porvi diverse strutture di spazio normato, tra cui:

- $\|f\|_\infty = \sup\limits_{x\in[a,b]} |f(x)|$
- $\|f\|_1 = \int_a^b f(x)\, dx$

Osserviamo anzitutto che entrambe le norme sono ben definite: la prima per un teorema noto anch'esso come *teorema di Weierstrass* [a] , e la seconda perché ogni funzione continua e limitata è integrabile. Le proprietà delle norme si verificano facilmente.

Proposizione 1.1.4. *$\|.\|_\infty$ e $\|.\|_1$ non sono equivalenti.*

Dimostrazione. Consideriamo le funzioni $f_n(x) = x^n$. Ovviamente tali funzioni sono in $C[a, b]$ per ogni $a, b \in \mathbb{R}$. Considerando in particolare a=0, b=1 si ha

$$\|f_n\|_\infty = \max_{x\in[0,1]} |x^n| = 1$$

e $\forall n \in \mathbb{N}$

$$\|f_n\|_1 = \int_0^1 x^n\, dx = \frac{1}{n+1}$$

Supponiamo per assurdo che la tesi sia falsa: allora in particolare esiste una costante C tale che

$$\|f_n\|_\infty \leq C\|f_n\|_1$$

[a] Una funzione continua in un intervallo chiuso e limitato ammette massimo e minimo nell'intervallo

cioè tale che

$$n + 1 \leq C \quad \forall n \in \mathbb{N}$$

Essendo giunti a un assurdo, dobbiamo concludere che le norme non sono equivalenti. $\square$

Nel corso dell'intera trattazione, ove altrimenti specificato, considereremo lo spazio $C[a,b]$ con la norma $\|.\|_\infty$, detta anche *norma uniforme*, *norma infinito* o *norma del sup.*
Una naturale generalizzazione delle funzioni continue sono le funzioni *p-integrabili secondo Lebesgue.* Gli spazi L^p consentono infatti di trattare molte più situazioni delle sole funzioni continue, e proprio per questo motivo esamineremo diverse volte se e come si estendono a $L^p[a,b]$ i risultati ottenuti su $C[a,b]$. Ne ricordiamo la definizione

Definizione 1.1.5. Sia $p \in [1, +\infty)$, e sia (X, μ) uno spazio di misura[b]. $L^p(X)$ è lo spazio delle funzioni misurabili da X in $\mathbb{R}$ con la proprietà

$$\int_X |f|^p \ d\mu < +\infty$$

$(L^p, \|.\|_p)$ è uno spazio normato definendo $\|f\| = \left(\int_X |f|^p \ d\mu\right)^{1/p}$, purchè si considerino equivalenti due funzioni se esse differiscono solo per un insieme di misura nulla.

Una trattazione rigorosa sulla costruzione di questi spazi e sulle loro proprietà, a cui faremo spesso riferimento, può essere trovata in [Car00, Rud66].
Ricordiamo infine il concetto topologico di separabilità, che è essenziale ogniqualvolta si voglia parlare di approssimazione.

Definizione 1.1.6. Dato uno spazio metrico (X, d) e un suo sottoinsieme Y, Y si dice denso in X se $\overline{Y} = X$, dove $\overline{Y}$ indica la chiusura di Y. Equivalentemente

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists\, y \in Y \text{ tale che } d(x, y) < \varepsilon$$

Uno spazio metrico che possiede un simile sottoinsieme di cardinalità numerabile si dice *separabile.*

Un esempio di spazio separabile è l'insieme dei numeri reali $\mathbb{R}$. Infatti, detto $\mathbb{Q}$ l'insieme dei numeri razionali, si ha che esso è numerabile e denso in $\mathbb{R}$. Allo stesso modo è possibile

[b]si veda il primo capitolo di [Rud66] per un'introduzione rigorosa alla teoria della misura

mostrare che l'insieme dei polinomi a coefficienti reali $\mathbb{R}[x]$ è separabile, prendendo $\mathbb{Q}[x]$ come sottoinsieme.
Un esempio di spazio non separabile è dato dall'insieme delle serie formali $\mathbb{R}[[x]]$. Per dimostrare la non separabilità di uno spazio esistono alcune tecniche standard: ne vedremo una nel paragrafo 1.6.

1.2 Il teorema di Weierstrass

Il primo risultato importante che presentiamo è il teorema di approssimazione di Weierstrass, che afferma che i polinomi a coefficienti reali sono un sottoinsieme denso delle funzioni continue su un intervallo chiuso e limitato. Quest'ultima ipotesi, come vedremo più avanti, è essenziale.

Teorema 1.2.1. *Per ogni coppia di numeri reali (a,b) con* $a < b$ *e per ogni funzione continua* $f : [a, b] \longrightarrow \mathbb{R}$ *esiste un polinomio algebrico* p *con la seguente proprietà:*

$$|f(x) - p(x)| < \varepsilon \quad \forall x \in [a; b] \tag{1.1}$$

Premettiamo un lemma alla dimostrazione

Lemma 1.2.2. *Esiste un'isometria lineare tra* $C[0, 1]$ *e* $C[a, b]$ *che manda polinomi algebrici in polinomi algebrici*

Dimostrazione. Definiamo:

$$\begin{aligned} T :& C[0, 1] \longrightarrow C[a, b] \\ f(x) &\longmapsto f\left(\frac{x-a}{b-a}\right) \end{aligned}$$

Questa è evidentemente l'isometria cercata: sia $p_n(x) = \sum_{k=0}^{n} a_k x^k$, allora

$$T(p_n(x)) = \sum_{k=0}^{n} a_k \left(\frac{x-a}{b-a}\right)^k$$

è ancora un polinomio in x. □

Il lemma appena dimostrato ci permette di generalizzare molti risultati ottenuti sull'intervallo $[0,1]$ a un più generico intervallo $[a,b]$. D'ora in poi, quindi, ci limiteremo a considerare l'intervallo [0,1].
Dimostriamo ora il teorema di Weierstrass in $C[0,1]$. La dimostrazione è tratta da [Sch], che presenta l'enunciato moderno del teorema applicandovi però una tecnica dimostrativa sostanzialmente equivalente a quella dell'articolo originale.

Dimostrazione.
Definiamo un'estensione uniformemente continua di f a $\mathbb{R}$ nel seguente modo:

$$f(x) = \begin{cases} 0 & x < -1 \\ f(0)(x+1) & -1 \leq x < 0 \\ f(x) & 0 \leq x \leq 1 \\ f(1)(2-x) & 1 < x < 2 \\ 0 & x \geq 2 \end{cases}$$

Con un piccolo abuso di notazione, continueremo a indicare con f la funzione estesa a tutto $\mathbb{R}$.
Definiamo inoltre, dato $h > 0$

$$S_h(f)(x) = \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{+\infty} f(u) e^{-\left(\frac{u-x}{h}\right)^2} \, du$$

Vogliamo anzitutto mostrare che $\lim_{h \to 0} S_h(f)(x) = f(x)$ uniformemente in x.
Sia $\varepsilon > 0$, e sia δ tale che $|f(x) - f(y)| < \frac{\varepsilon}{2}$ per ogni $x, y \in \mathbb{R}$ con $|x-y| < \delta$. L'esistenza di un δ con tale proprietà è garantita dall'uniforme continuità di f.
Sia inoltre $M > 0$ tale che $|f(x)| < M$ per ogni x. Tale costante M esiste, poichè per come è definita l'estensione si ha $\sup_{x \in \mathbb{R}} f(x) = \max_{x \in [0,1]} f(x)$.
Ricordando che $\int_{-\infty}^{+\infty} e^{-u^2} = \sqrt{\pi}$, con un semplice calcolo si ha

$$\frac{1}{h\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\left(\frac{u-x}{h}\right)^2} \, du = 1$$

da cui segue

$$f(x) = \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{+\infty} f(x) e^{-\left(\frac{u-x}{h}\right)^2} \, du$$

Allora

$$
\begin{aligned}
|S_h(f)(x) - f(x)| &\leq \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{+\infty} |f(u) - f(x)|\, e^{-\left(\frac{u-x}{h}\right)^2}\, du \\
&\leq \frac{\varepsilon}{2} + \frac{1}{h\sqrt{\pi}} \int_{|x-u|\geq\delta} |f(u) - f(x)|\, e^{-\left(\frac{u-x}{h}\right)^2}\, du \\
&\leq \frac{\varepsilon}{2} + \frac{2M}{h\sqrt{\pi}} \int_{|x-u|\geq\delta} e^{-\left(\frac{u-x}{h}\right)^2}\, du \\
&\leq \frac{\varepsilon}{2} + \frac{2M}{h\sqrt{\pi}} \int_{|v|\geq\frac{\delta}{h}} e^{-v^2}\, dv \\
&\leq \frac{\varepsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}} \int_{|v|\geq\frac{\delta}{h}} |v|e^{-v^2}\, dv \\
&\leq \frac{\varepsilon}{2} + \frac{4Mh}{\delta\sqrt{\pi}} \int_{0}^{\infty} ve^{-v^2}\, dv \\
&= \frac{\varepsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}} < \varepsilon
\end{aligned}
$$

non appena $0 < h \leq \dfrac{\varepsilon\delta\sqrt{\pi}}{4M}$, e essendo $\displaystyle\int_0^\infty ve^{-v^2} = \frac{1}{2}$.

Sia ora h_0 tale che $|S_{h_0}(f)(x) - f(x)| < \frac{\varepsilon}{2}$. Essendo f nulla al di fuori di $[-2, 2]$

$$
S_{h_0}(f)(x) = \frac{1}{h_0\sqrt{\pi}} \int_{-2}^{2} f(u) e^{-\left(\frac{u-x}{h_0}\right)^2}\, du
$$

Sull'intervallo $\left[-\frac{4}{h_0}, \frac{4}{h_0}\right]$ la serie di Taylor di e^{-v^2} converge uniformemente, dunque possiamo approssimare l'esponenziale con precisione arbitraria tramite il suo polinomio di Taylor di un grado sufficientemente elevato. Sia N tale che

$$
\frac{1}{h_0\sqrt{\pi}} \left| e^{-\left(\frac{u-x}{h_0}\right)^2} - \sum_{k=0}^{N} \frac{(-1)^k}{k!} \left(\frac{u-x}{h_0}\right)^{2k} \right| < \frac{\varepsilon}{8M}
$$

per tutti gli x, u con $|x| \leq 2, |u| \leq 2$, essendo in tal caso $|u - x| < 4$.
Possiamo allora scrivere

$$
\left| S_{h_0} f(x) - \frac{1}{h_0\sqrt{\pi}} \int_{-2}^{2} f(u) \sum_{k=0}^{N} \frac{(-1)^k}{k!} \left(\frac{u-x}{h_0}\right)^{2k} \right| < \frac{\varepsilon}{2}
$$

che vale per ogni x con $|x| \leq 2$.

Se ora definiamo

$$p(x) = \sum_{k=0}^{N} \left(\frac{1}{h_0\sqrt{\pi}} \int_{-2}^{2} f(u) \frac{(-1)^k}{k!} \right) \left(\frac{u-x}{h_0} \right)^{2k}$$

esso è un polinomio di grado al più $2N$ con la seguente proprietà

$$|S_{h_0}(f)(x) - p(x)| < \frac{\varepsilon}{2} \quad \forall x \in [-2, 2]$$

che implica, per la convergenza uniforme degli $S_h(f)$ ad f

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

□

È facile intuire come un simile teorema possa avere moltissime applicazioni. La più immediata riguarda forse la risoluzione di problemi pratici in cui occorre trovare una funzione sconosciuta: molte definizioni di problema *ben posto* richiedono che tale funzione sia continua, e grazie al teorema appena dimostrato è possibile approssimarla con un polinomio arbitrariamente vicino ad essa. Un problema con bassa regolarità, dunque, si trasforma in un problema polinomiale a meno di un errore piccolo a piacere.
Un'ulteriore conseguenza è la seguente

Proposizione 1.2.3. *$C[a, b]$ è separabile*

Dimostrazione. Poiché i polinomi a coefficienti razionali sono un sottoinsieme denso e numerabile di $\mathbb{R}[x]$, allora $\mathbb{R}[x]$ è separabile. Per il teorema di Weierstrass $\mathbb{R}[x]$ è denso in $C[a, b]$, dunque quest'ultimo è anch'esso separabile. □

1.3 Polinomi trigonometrici

Per quanto le funzioni continue su un intervallo chiuso e limitato si presentino in moltissime situazioni reali, non sempre è possibile limitarsi a tale condizione. Spesso infatti il problema richiede la conoscenza di una soluzione continua, ad esempio, su tutto $\mathbb{R}$. In generale in questi casi non si trova un teorema analogo a quello di Weierstrass, tuttavia se si ha a che fare con funzioni periodiche (ad esempio se si sta lavorando con le serie di Fourier) esiste una versione equivalente ad esso, che però considera una differente classe di polinomi.

Definizione 1.3.1. Sia $T_n = span\{1, \sin x, \cos x, \sin 2x, \cos 2x, ..., \sin nx, \cos nx\}$ l'insieme delle combinazioni lineari a coefficienti reali delle funzioni trigonometriche indicate. Gli elementi di T_n si dicono polinomi trigonometrici (di grado al più n).

Il seguente lemma giustifica il nome "polinomio trigonometrico" che viene attribuito al generico $t_n(x) = \sum_{k=0}^{n} (a_k \cos(kx) + b_k \sin(kx))$, con $a_i, b_j \in \mathbb{R}$

Lemma 1.3.2. *Ogni polinomio trigonometrico è un polinomio algebrico nelle variabili* $\cos x, \sin x$*.*

Dimostrazione. È sufficiente mostrare che $\cos kx$ e $\sin kx$ possono essere scritti come polinomi in $\cos x, \sin x$. Dalla formula di ricorrenza

$$\cos kx + \cos(k-2)x = 2\cos(k-1)x \cos x$$

si ricava facilmente $\cos 2x = 2\cos^2 x - 1$, $cos 3x = 4\cos^3 x - 3\cos x$, o più in generale che $\cos nx$ è un polinomio di grado n in $\cos x$ con coefficiente $a_n = 2^{n-1}$.
Dall'identità trigonometrica

$$\sin(k+1)x - \sin(k-1)x = 2\cos kx \sin x$$

segue la tesi. □

Lo spazio di funzioni in cui studieremo i polinomi trigonometrici è $\tilde{C}[2\pi]$, definito come l'insieme delle funzioni continue $f : \mathbb{R} \to \mathbb{R}$ 2π-periodiche.
Possiamo definire la norma infinito su $\tilde{C}[2\pi]$ nel seguente modo:

$$\|f\|_\infty = \max_{x \in [0;2\pi]} |f(x)|$$

Osserviamo inoltre che ogni funzione in $\tilde{C}[2\pi]$ è uniformemente continua. Infatti è sufficiente restringersi all'intervallo $[0, 2\pi]$ e applicare il teorema di Heine-Cantor.

Premettiamo un lemma che mostra un primo legame tra le funzioni di $\tilde{C}[2\pi]$ e le funzioni continue su un intervallo chiuso e limitato.

Lemma 1.3.3. *Sia $f \in \tilde{C}[2\pi]$ una funzione pari. Allora $\forall \varepsilon > 0$ esiste un polinomio trigonometrico pari t tale che $\|f - t\|_\infty < \varepsilon$*

Dimostrazione. Sia $g(y) = f(\arccos y)$. Per $|y| \leq 1$ g è una funzione continua, in quanto composizione di funzioni continue. Allora per il teorema di Weierstrass (1.2.1) esiste un polinomio algebrico p con la proprietà

$$\max_{|y|\leq 1} |g(y) - p(y)| < \varepsilon$$

Sia allora $t(x) = p(\cos x)$. Questo è un polinomio trigonometrico per il lemma 1.3.2, e verifica la relazione .

$$\max_{x\in[0;2\pi]} |f(x) - p(x)| < \varepsilon$$

Inoltre non contiene potenze di $\sin x$, perciò è pari. □

Grazie a questo lemma possiamo ora enunciare e dimostrare l'equivalente del teorema di Weierstrass su $\tilde{C}[2\pi]$ come conseguenza del teorema 1.2.1

Teorema 1.3.4 (Teorema di Weierstrass per polinomi trigonometrici)**.**
Per ogni funzione continua $f \in \tilde{C}[2\pi]$ esiste un polinomio trigonometrico t con la seguente proprietà:

$$|f(x) - t(x)| < \varepsilon \quad \forall x \in \mathbb{R} \tag{1.2}$$

Dimostrazione. Le funzioni

- $a(x) = f(x) + f(-x)$
- $b(x) = (f(x) - f(-x)) \sin x$

sono entrambe pari. Allora per il lemma esistono $t_1(x)$ e $t_2(x)$ che approssimano $a(x)$ e $b(x)$, cioè

$$\begin{cases} f(x) + f(-x) = t_1(x) + r_1(x) \\ (f(x) - f(-x)) \sin x = t_2(x) + r_2(x) \end{cases}$$

con $\|r_1\|_\infty < \varepsilon$ e $\|r_2\|_\infty < \varepsilon$
Moltiplicando la prima equazione per $\sin^2 x$ e la seconda per $\sin x$ e sommandole otteniamo

$$f(x) \sin^2 x = t_3(x) + r_3(x)$$

con t_3 polinomio trigonometrico e $\|r_3\|_\infty < 2\varepsilon$
Poichè questa relazione vale per ogni f in $\tilde{C}[2\pi]$, deve valere anche per $f(y-\pi/2)$. Sostituendo nella relazione e da cui valutando in $y = x + \pi/2$ otteniamo

$$f(y-\pi/2)\sin^2 y = t_4(y) + r_4(y)$$
$$f(x)\cos^2 x = t_5(x) + r_5(x)$$

dove t_5 è ancora un polinomio trigonometrico e $\|r_5\|_\infty < 2\varepsilon$
Sommando le ultime due relazioni otteniamo infine

$$f(x) = t(x) + r(x)$$

dove $t(x) = t_3(x) + t_5(x)$, e $\|r\|_\infty < 2\varepsilon$
Allora $\|f-t\|_\infty < 2\varepsilon$. Per arbitrarietà di ε, il teorema è dimostrato. □

1.4 Dimostrazione di Fejér

Nella dimostrazione del teorema di Weierstrass per polinomi trigonometrici abbiamo utilizzato il teorema 1.2.1, che dunque implica il primo. Anche l'implicazione inversa è vera, e si dimostra in maniera analoga: i due teoremi sono, di fatto, equivalenti. Devono dunque esistere alcune dimostrazioni del teorema in $\tilde{C}[2\pi]$ indipendenti da 1.2.1. Una di queste è quella proposta dal matematico Lipót Fejér.
Ricordiamo anzitutto alcuni risultati della teoria dello sviluppo in serie di Fourier, necessaria alla trattazione. Queste e altre proprietà delle serie di Fourier, assieme alle relative dimostrazioni, possono essere trovate in [Car00, §15]

Definizione 1.4.1. Sia f limitata, Riemann-integrabile su $[-\pi,\pi]$ e $2\pi-$periodica.

$$\frac{a_0}{2} + \sum_{n=1}^{+\infty}(a_k \cos kx + b_k \sin kx)$$

dove
$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\; dt \quad \text{ e } \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt\; dt$$
si dice serie di Fourier di f.
Ricordando la definizione di esponenziale complesso $\left(e^{x+iy} = e^x\left(\cos y + i\sin y\right)\right)$, la serie

di Fourier si può scrivere anche come

$$\sum_{n\in\mathbb{Z}} c_k e^{ikx}$$

con $c_0 = \dfrac{a_0}{2}, \quad c_n = \dfrac{a_n - ib_n}{2}, \quad c_{-n} = \overline{c_n}$

Proposizione 1.4.2. *Se $f \in C^1(\mathbb{R}), f$ 2π -periodica allora la serie di Fourier di f converge uniformemente a f su tutto l'intervallo $[-\pi;\pi]$. Inoltre la convergenza è totale, ovvero*

$$\frac{|a_0|}{2} + \sum_{n=1}^{+\infty} (|a_n| + |b_n|) < +\infty$$

Proposizione 1.4.3. *Se $f \in C(\mathbb{R}), f$ 2π -periodica, f lipschitziana* [c] *allora la serie di Fourier di f converge uniformemente a f su tutto l'intervallo $[-\pi;\pi]$.*

Definizione 1.4.4 (Nucleo di Dirichlet)**.**
Detta $s_n(x) = \displaystyle\sum_{k=-n}^{n} c_k e^{ikx}$ l'n-esima somma parziale della serie di Fourier di f, il nucleo di Dirichlet $D_n(t)$ è un insieme numerabile di funzioni con la seguente proprietà

$$s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)D_n(t)\, dt$$

È possibile darne un'espressione esplicita: $D_n(t) = \dfrac{1}{2} + \displaystyle\sum_{k=1}^{n} \cos kt = \dfrac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{1}{2}t}$

Teorema 1.4.5 (Convergenza puntuale della serie di Fourier)**.**
Se $\exists \lim\limits_{x\to x_0^-} f(x)$ e $\exists \lim\limits_{x\to x_0^+} f(x)$ e inoltre
$\exists k > 0 : \left|f(x_0^{\pm} - t) - f(x_0^{\pm})\right| \leq kt \quad \forall t \in]0;\delta]$ per un δ opportuno allora

$$\sum_{k\in\mathbb{Z}} |c_k| e^{ikx_0} = \frac{f(x_0^+) - f(x_0^-)}{2}$$

In generale dunque la serie di Fourier di una funzione f solo continua non converge uniformemente alla funzione stessa; per farla convergere anche solo puntualmente è necessaria

[c] i.e. $\exists\, K > 0 : |f(x) - f(y)| \leq K|x-y| \quad \forall x, y \in \mathbb{R}$

un'ipotesi aggiuntiva (che sia localmente lipschitziana in un intorno del punto). Affinché si abbia la convergenza uniforme è necessaria una regolarità maggiore (lipschitzianità globale).

Per questo motivo le somme parziali della serie di Fourier non costituiscono una buona successione di polinomi trigonometrici per approssimare le funzioni continue. Tuttavia, come dimostrò Fejér, una buona successione è data dalle somme di Cesaro della serie di Fourier.

Definizione 1.4.6 (Somme di Cesaro)**.**

$$\sigma_n(x) = \frac{s_0(x) + \ldots + s_n(x)}{n+1}$$

si dice n-esima somma di Cesaro della successione s_n.

Proposizione 1.4.7.
Se $s_n \to s$, allora $\sigma_n \to s$

Dimostrazione.
Se s_n è convergente, allora è anche limitata. Sia $M : |s_n| \leq M \quad \forall n$.
Sia ora $\varepsilon > 0$ fissato. Poiché s_n è convergente possiamo scegliere un n_ε in modo che $|s_k - s| < \varepsilon$ per ogni $k > n_\varepsilon$. Allora

$$\begin{aligned}\sigma_N =& \frac{s_0 + \ldots + s_{N-1}}{N} \\ =& \frac{s_0 + \ldots + s_{n_\varepsilon}}{N} + \frac{s_{n_\varepsilon+1} + \ldots + s_{N-1}}{N}\end{aligned}$$

Per ogni $N > n_\varepsilon$ allora

$$-\left(\frac{n_\varepsilon}{N}\right) M + \left(\frac{N - n_\varepsilon}{N}\right)(s - \varepsilon) \leq \sigma_N \leq \left(\frac{n_\varepsilon}{N}\right) M + \left(\frac{N - n_\varepsilon}{N}\right)(s + \varepsilon)$$

cioé

$$s - 2\varepsilon \leq \sigma_N \leq s + 2\varepsilon$$

che dimostra la proposizione.

□

Osservazione 1.4.8. Il viceversa è falso: un controesempio è $s_n = (-1)^n$. Chiaramente $\nexists \lim\limits_{n\to+\infty} s_n$, ma $\lim\limits_{n\to+\infty} \sigma_n = 0$, essendo $|\sigma_n| \leq \frac{2}{n} \quad \forall n$

Dunque, in generale, le somme di Cesaro di una successione hanno proprietà di convergenza migliori della successione di partenza.

Definendo il nucleo di Fejér $K_n(t)$ per σ_n in modo analogo al nucleo di Dirichlet per s_n, è possibile darne un'espressione esplicita tramite $D_n(t)$. Infatti

$$\begin{aligned}\sigma_n(x) =& \frac{1}{n}\left(s_0(x) + ... + s_{n-1}(x)\right)\\ =& \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\left(\frac{1}{n}\sum_{k=0}^{n-1} D_k(t)\right)\, dt\\ =& \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)K_n(t)\, dt\end{aligned}$$

perciò

$$K_n(t) = \frac{1}{n}\sum_{k=0}^{n-1} D_k(t) = \frac{1}{n}\sum_{k=0}^{n-1}\frac{\sin(2k+1)t/2}{2\sin t/2} = \frac{\sin^2(nt/2)}{2n\sin^2(t/2)}$$

Per mostrare che nel caso della successione delle somme parziali della serie di Fourier le somme di Cesaro convergono uniformemente a f continua, dimostriamo un teorema più generale che elenca le proprietà che deve possedere un nucleo per poter generare tramite convoluzione successioni approssimanti per f.

Teorema 1.4.9.
Se una sequenza $\{k_n\}$ in $\tilde{C}[2\pi]$ soddisfa:

(a) $k_n \geq 0$

(b) $\displaystyle \frac{1}{\pi}\int_{-\pi}^{\pi} k_n(t)\, dt = 1$

(c) $\displaystyle \int_{\delta\leq|t|\leq\pi} k_n(t)\, dt \xrightarrow{n\to+\infty} 0 \quad \forall\delta > 0$

Allora

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)k_n(t)\, dt \textit{ converge uniformemente a } f(x)$$

se $n \to +\infty$ *e* $\forall f \in \tilde{C}[2\pi]$.

Dimostrazione.
Sia $\varepsilon > 0$ fissato.

Poiché $f \in \tilde{C}[2\pi]$, f è uniformemente continua. Allora è possibile scegliere $\delta > 0$ in modo che $|f(x) - f(x+t)| < \varepsilon$ per ogni x con $|t| < \delta$.
Per la proprietà (b) possiamo scrivere

$$\begin{aligned}\left| f(x) - \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) k_n(t)\, dt \right| =& \frac{1}{\pi} \left| \int_{-\pi}^{\pi} [f(x) - f(x+t)]\, k_n(t)\, dt \right| \\ \leq& \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x) - f(x+t)|\, k_n(t)\, dt \\ \leq& \frac{\varepsilon}{\pi} \int_{|t|<\delta} k_n(t)\, dt + \frac{2\|f\|_\infty}{\pi} \int_{\delta \leq |t| \leq \pi} k_n(t)\, dt \\ \leq& \varepsilon + \varepsilon = 2\varepsilon\end{aligned}$$

per ogni n sufficientemente grande e per la proprietà (c). □

Si verifica facilmente che il nucleo di Fejér $K_n(t)$ soddisfa le ipotesi del teorema. Poiché in particolare σ_n è un polinomio trigonometrico, abbiamo trovato un modo per approssimare ogni funzione continua $2\pi-$periodica con un polinomio trigonometrico. Dunque il teorema 1.4.9 implica il teorema 1.3.4.

1.5 Dimostrazione probabilistica

Vediamo ora una dimostrazione del teorema 1.2.1 che fa uso di concetti probabilistici, in particolare della legge dei grandi numeri. Vedremo come definendo un apposito polinomio sia possibile dimostrare un teorema con argomenti apparentemente così scorrelati.

Detta X una variabile aleatoria, indicheremo con $\mathbb{E}(X) = \sum x\mathbb{P}(X{=}x)$ il suo valore atteso, e con $Var(X) = \mathbb{E}(|X - \mathbb{E}(X)|^2)$ la sua varianza.

Definizione 1.5.1. X si dice variabile aleatoria di Bernouilli (di parametro $p \in [0,1]$) se $X \in \{0,1\}$ e

$$\mathbb{P}(X = 0) = p \text{ e } \mathbb{P}(X = 1) = (1-p) \quad \text{con } p \in [0;1]$$

Data una variabile aleatoria di Bernouilli X, è molto semplice calcolarne il valore atteso e la varianza. Infatti

$$\mathbb{E}(X) = \sum_{x=0,1} x\mathbb{P}(X = x) = p$$
$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X^2) = \sum_{x=0,1} x\mathbb{P}(X^2 = x) - p^2 = p(1-p)$$

Definizione 1.5.2. S_n si dice variabile aleatoria binomiale se

$$S_n = \sum_{i=1}^{n} X_i$$

con X_i variabili aleatorie di Bernouilli indipendenti identicamente distribuite. Per indipendenza delle n variabili di Bernouilli è immediato osservare che:

- S_n assume solo valori interi compresi tra 0 e n.
- $\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $\mathbb{E}(S_n) = \mathbb{E}(X_1 + ... + X_n) = np$
- $Var(S_n) = Var(X_1 + ... + X_n) = np(1-p)$

Per dimostrare il teorema di Weierstrass occorrono alcuni risultati sulle variabile aleatorie. Li ricordiamo, riportando una breve dimostrazione per i più importanti.

Lemma 1.5.3. *Sia Y una variabile aleatoria, $Y \geq 0$.*
Allora $\forall \delta > 0$

$$\mathbb{P}(Y > \delta) < \frac{1}{\delta}\mathbb{E}(Y)$$

Teorema 1.5.4 (Disuguaglianza di Chebyshev)**.**
Sia X una variabile aleatoria con $\mathbb{E}(X^2) < +\infty$
Allora $\forall \delta > 0$

$$\mathbb{P}(|X - \mathbb{E}(X)|) > \delta) < \frac{1}{\delta^2}Var(X)$$

Dimostrazione. Definiamo una nuova variabile aleatoria $Y = |X - \mathbb{E}(X)|$. È immediato osservare che $\mathbb{E}(Y) = \mathbb{E}(|X - \mathbb{E}(X)|) = |\mathbb{E}(X) - \mathbb{E}(X)| = 0$. Poichè Y è non negativa,

inoltre, $\mathbb{P}(Y > \delta) = \mathbb{P}(Y^2 > \delta^2)$.
Applicando il lemma 1.5.3
$$\mathbb{P}(Y^2 > \delta^2) \leq \frac{1}{\delta^2}\mathbb{E}(Y^2)$$
ovvero
$$\mathbb{P}(|X - \mathbb{E}(X)|^2 > \delta^2) \leq \frac{1}{\delta^2}\mathbb{E}(|X - \mathbb{E}(X)|^2) = \frac{1}{\delta^2}Var(X)$$

$\square$

Teorema 1.5.5 (Legge debole dei grandi numeri).
Siano $X_1, ..., X_n, ...$ *variabili aleatorie con* $\mathbb{E}(X_i^2) < +\infty$, *indipendenti e identicamente distribuite. Sia inoltre* $S_n = \sum_{i=1}^{n} X_i$.
Allora $\frac{S_n}{n}$ *converge in probabilità a* $\mathbb{E}(X_i)$, *ovvero*
$$\forall\delta > 0 \quad \lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X_i)\right| > \delta\right) = 0$$

Dimostrazione.
Per indipendenza e identica distribuzione delle X_i:

- $Var(S_n) = \sum_{i=1}^{n} Var(X_i) = nVar(X_i)$
- $\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbb{E}(S_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \mathbb{E}(X_i)$
- $Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2}Var(S_n) = \frac{1}{n}Var(X_i)$
- $\lim_{n\to+\infty} Var\left(\frac{S_n}{n}\right) = 0$

Dalla disuguaglianza 1.5.4 segue allora la tesi
$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| > \delta\right) \leq \frac{1}{\delta^2}Var\left(\frac{S_n}{n}\right) \xrightarrow{n\to+\infty} 0$$

$\square$

Teorema 1.5.6 (Teorema di Weierstrass)**.**
Sia $f : [0,1] \longrightarrow \mathbb{R}$ continua. Allora esiste un polinomio $a_n(x)$ a coefficienti reali tale che

$$\lim_{n\to+\infty} \max_{p\in[0,1]} |f(p) - a_n(p)| = 0$$

Dimostrazione.
Sia $a_n(p) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}$. Sia inoltre S_n la variabile aleatoria binomiale ottenuta come somma di n variabili X_i di Bernouilli di parametro p. Per quanto mostrato in 1.5.2 :

- $\mathbb{P}(S_n = k) = \binom{n}{k}p^k(1-p)^{n-k}$

- $\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = 1$ (è la probabilità dell'evento certo)

- $Var(S_n) = np(1-p)$

Allora:

$$\begin{aligned}|f(p) - a_n(p)| &= \left| f(p)\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} - \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}\right| \\ &\le \sum_{k=0}^{n}\left|f(p) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}\end{aligned}$$

Poichè f è continua in [0,1], intervallo chiuso e limitato (perciò compatto), per il teorema di Heine-Cantor f è uniformemente continua, ovvero:

$$\forall\varepsilon > 0 \quad \exists\delta_\varepsilon > 0 : \forall p, p' \in [0,1], |p - p'| < \delta_\varepsilon \Rightarrow |f(p) - f(p')| < \varepsilon$$

Fissato $\varepsilon > 0$, sia $\delta = \delta_\varepsilon$.

$$\sum_{k:\left|\frac{k}{n}-p\right|\le\delta}\left|f(p) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k} \le \varepsilon\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = \varepsilon$$

inoltre

$$\begin{aligned}
&\sum_{k:\left|\frac{k}{n}-p\right|>\delta}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}\\
&\leq (2\max_{p\in[0,1]}|f(p)|)\cdot\mathbb{P}\left(\left|\frac{S_n}{n}-p\right|>\delta\right)=\\
&=(2\max_{p\in[0,1]}|f(p)|)\cdot\mathbb{P}\left(\left|\frac{S_n}{n}-\mathbb{E}(X_i)\right|>\delta\right)
\end{aligned}$$

per la legge debole dei grandi numeri (1.5.5) tale quantità è infinitesima in n, ovvero esiste un n_ε dopo il quale $\mathbb{P}\left(\left|\frac{S_n}{n}-p\right|>\delta\right)<\varepsilon$. Sia $n>n_\varepsilon$, allora

$$\sum_{k:\left|\frac{k}{n}-p\right|>\delta}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}\leq (2\max_{p\in[0,1]}|f(p)|)\varepsilon=2M\varepsilon.$$

Dunque per $n>n_\varepsilon$

$$\sum_{k=0}^{n}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k:\left|\frac{k}{n}-p\right|\leq\delta}(...)\ +\sum_{k:\left|\frac{k}{n}-p\right|>\delta}(...)\leq(2M+1)\varepsilon$$

che conclude la dimostrazione.

□

La successione di polinomi definita nel dimostrare il teorema, apparentemente creata *ad hoc* per poter utilizzare la legge dei grandi numeri, ha in realtà molte altre applicazioni. Riportiamo perciò la sua definizione esplicita

Definizione 1.5.7. Data una funzione f continua su $[0,1]$

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

si dice n-esimo polinomio di Bernstein di f.

Abbiamo appena dimostrato che $\lim\limits_{n\to+\infty}B_n(f)=f$ nel senso della convergenza uniforme su $C[0,1]$, qualunque sia f. Vedremo nel prossimo capitolo una condizione per identificare altre successioni come questa.

1.6 Generalizzazioni

Il teorema di Weierstrass (1.2.1), di certo abbastanza generale, ha tuttavia due ipotesi abbastanza forti: l'intervallo $[a, b]$ deve essere chiuso e limitato, e le funzioni devono essere continue. In questo paragrafo vogliamo esaminare cosa succede indebolendo tali ipotesi, per estendere eventualmente il teorema in un ambito più generale.

Un risultato classico che generalizza il teorema di Weierstrass, presente su ogni testo di analisi funzionale, è il seguente:

Teorema 1.6.1 (di Stone-Weierstrass)**.**
Sia X è uno spazio metrico compatto, e A una sottoalgebra di $C(X, \mathbb{R})$. Se A contiene una funzione costante non nulla, allora A è densa in $C(X, \mathbb{R})$ se e solo se separa i punti in X (i.e. $\forall x, y \in X : x \neq y$, esiste f in A con $f(x) \neq f(y)$)

È facile osservare come nel caso $X = [a, b] \subset \mathbb{R}$, che per il teorema di Heine-Borel è l'unico esempio di sottoinsiemi (connessi) compatti di $\mathbb{R}$, questo teorema affermi che la sottoalgebra dei polinomi è densa in $C[a, b]$, ovvero il teorema di Weierstrass.

Come abbiamo visto, una delle conseguenze più importanti del teorema di Weierstrass è la separabilità dello spazio di funzioni considerato. Perciò tale teorema non può valere in spazi di funzioni non separabili, e quest'implicazione inversa sarà ampiamente sfruttata negli esempi che seguiranno.
Consideriamo anzitutto l'insieme $C_b(\mathbb{R})$ delle funzioni continue limitate su tutto $\mathbb{R}$.

Proposizione 1.6.2. *$C_b(\mathbb{R})$ non è separabile*

Dimostrazione. Sia K il sottospazio di $C_b(\mathbb{R})$ formato dalle funzioni f che assumono solo i valori 0 e 1 sugli interi. Esiste un sottoinsieme non numerabile $S \subset K$ con la proprietà

$$\|f - g\|_\infty \geq 1 \quad \forall f, g \in S, f \neq g$$

Un sottoinsieme di questo tipo si può ottenere con la seguente costruzione: per ogni successione $\{x_n\}$ formata da 0 e 1, si definisca $f(n) = x_n$ per ogni $n \in \mathbb{N}$, e f tra n e $n+1$ coincidente con il segmento che connette (n, x_n) a $(n+1, x_{n+1})$. Essendo le sequenze binarie non numerabili, anche questo sottoinsieme è non numerabile. Inoltre due elementi distinti di questo insieme avranno almeno un intero n in cui assumono valori diversi, da

cui segue $\|f-g\|_\infty \geq 1$ non appena $f \neq g$.
Sia ora B un sottoinsieme numerabile di $C_b(\mathbb{R})$. Se B fosse denso in $C_b(\mathbb{R})$, in particolare per ogni $s \in S$ ci sarebbe un $b \in B$ vicino a piacere, ad esempio con $\|s-b\|_\infty < \frac{1}{2}$.
Sia $\{s_n\}$ una successione di elementi di S a distanza minore di $\frac{1}{2}$ da $\{b_n\}$. Poichè S è non numerabile, tale successione è strettamente contenuta in S. Allora per la proprietà di S esiste almeno un $s \in S$ a distanza almeno $\frac{1}{2}$ da tutti gli elementi di B. Dunque B non può essere denso in $C_b(\mathbb{R})$.

□

Dalla proposizione appena dimostrata si deduce che il teorema di Weierstrass non può valere in $C_b(\mathbb{R})$.

Osservazione 1.6.3. La tecnica di trovare un sottoinsieme non numerabile di aperti disgiunti con lo stesso diametro (nel caso di $C_b(\mathbb{R})$ le palle di raggio $\frac{1}{2}$ centrate sugli elementi di S) è tra le più utilizzate per dimostrare la non separabilità.
Se consideriamo $C_c(\mathbb{R})$, cioé l'insieme delle funzioni continue a supporto compatto su $\mathbb{R}$, e proviamo a ripetere la costruzione di S, otteniamo un insieme in corrispondenza uno a uno con le successioni finite di 0 e 1 (infatti al di fuori di un certo compatto ogni funzione in $C_c(\mathbb{R})$ è identicamente nulla, e dunque la successione corrispondente è identicamente 0 tranne che per un numero finito di indici), che costituiscono un insieme di cardinalità numerabile. Ciò costituisce un indizio per la separabilità di $C_c(\mathbb{R})$, che sussiste ma che non dimostriamo in questa sede, non essendo i polinomi le funzioni dense cercate.

Un'altra questione interessante sorge indebolendo l'ipotesi di continuità del teorema di Weierstrass, ovvero considerando gli spazi $L^p[a,b]$ delle funzioni integrabili secondo Lebesgue con la norma p. Essendo infatti le funzioni continue dense in questi spazi ([Rud66, §3]), si ha che i polinomi costituiscono un sottoinsieme di funzioni denso anche per gli spazi L^p. Il teorema di Weierstrass invece non vale in $L^\infty[a,b]$ [d], in quanto $C[a,b]$ non ne costituisce un sottoinsieme denso. Tale spazio in effetti non è neppure separabile, e per dimostrarlo si procede in modo analogo a quanto fatto nella proposizione 1.6.2, considerando come sottoinsieme non numerabile a distanza 1 le funzioni caratteristiche $\{\chi_{[x,y]}\}$ al variare di $x < y$ in $[a,b]$.

[d] l'insieme delle funzioni limitate in $[a,b]$

Capitolo 2

Densità tramite funzionali

2.1 Teoremi di Riesz e Hahn-Banach

Fino ad ora abbiamo sempre cercato successioni di polinomi che convergessero alla funzione richiesta. Un punto di vista diverso sulla questione è quello introdotto da Riesz nel 1910, che propose lo studio non delle funzioni approssimanti in sé, ma di particolari funzionali a immagine densa in $C[a,b]$ che agendo su f restituissero una successione approssimante.
Per comprendere perché tale approccio sia, in certi contesti, naturale, si consideri il seguente teorema.

Teorema 2.1.1 (di Lerch). *Se $h \in C[0,1]$ ha la proprietà*

$$\int_0^1 x^n h(x)\, dx = 0 \qquad \forall n = 0, 1, ...$$

allora $h \equiv 0$.

Dimostrazione. Sia p_k la successione polinomiale che approssima uniformemente h in $[0,1]$, che esiste per il teorema di Weierstrass. Allora

$$\lim_{k \to +\infty} \int_0^1 p_k(x)h(x)\, dx = \int_0^1 [h(x)]^2\, dx$$

Ma, per linearità dell'integrale, per ogni k

$$\int_0^1 p_k(x)h(x)\, dx = 0$$

e dunque

$$\int_0^1 [h(x)]^2 \ dx = 0$$

che essendo h continua[a] implica $h \equiv 0$. □

Il teorema appena dimostrato era noto sin dal 1892, ma fino alla pubblicazione dei risultati di Riesz i tentativi per generalizzarlo a un sottospazio arbitrario diverso dai polinomi erano falliti. Nel 1910 in un solo articolo Riesz pubblicò il teorema che segue e un risultato analogo ad esso per gli spazi L^p

Teorema 2.1.2. *Sia M un sottospazio vettoriale di $C[a,b]$, e sia $f \in C[a,b]$. Allora f può essere uniformemente approssimata da elementi di M, i.e. f è nella chiusura di M, se e solo se ogni funzionale lineare che si annulla su M è nullo su tutto $C[a,b]$.*

Premettiamo un lemma alla dimostrazione, che ci limitiamo a enunciare. Per la dimostrazione si veda [Rie11].

Lemma 2.1.3. *Sia $\{u_k\}_{k\in K} \subseteq C[a,b]$, con K insieme di indici qualsiasi. Allora esiste un funzionale lineare continuo su $C[a,b]$ che soddisfa*

$$F(u_k) = c_k, \quad \forall k \in K$$

se e solo se per ogni $L \in \mathbb{R}$ tale che $\|F\| \le L$, per ogni $K' \subseteq K$ sottoinsieme finito di indici e per ogni $\{a_k\}_{k\in K} \subseteq \mathbb{R}$ si ha

$$\left| \sum_{k\in K'} a_k c_k \right| \le L \left\| \sum_{k\in K'} a_k u_k \right\|$$

Dimostrazione del teorema 2.1.2. Sia f nella chiusura di M, e sia F un funzionale lineare continuo che si annulla su M. Allora

$$F(f) = F(f-g) \quad \forall g \in M$$

[a] per il teorema di permanenza del segno se esistesse $y \in [0,1]$ tale che $h(y) \neq 0$ allora ci sarebbe un intorno di y in cui $h \neq 0$: in tal caso l'integrale di h^2 sarebbe positivo.

Essendo f nella chiusura di M, esiste $\{g_n\}$ tale che $\lim g_n = f$, cioè per ogni $\varepsilon > 0$ esiste $g \in M$ con $\|f - g\|_\infty < \varepsilon$ Allora

$$|F(f)| = |F(f - g)| \leq \|F\|\|f - g\|_\infty \leq \varepsilon\|F\|$$

da cui segue $F(f) = 0$.
Rimane da mostrare l'implicazione inversa. Supponiamo che f non sia nella chiusura di M: allora esiste $d > 0$ tale che $\|f - g\|_\infty \geq d$ per ogni g in M. Ma allora per il lemma esiste un funzionale lineare continuo F che soddisfa $F(g) = 0, F(f) = 1$ e $\|F\| \leq L$ per ogni $L \geq \frac{1}{d}$. Infatti

$$|a| \leq Ld|a| \leq L\|af - g\|_\infty$$

per ogni $g \in M, a \in R$

□

Un teorema assai più generale, che contiene come caso particolare i risultati ottenuti da Riesz, fu pubblicato nel 1927 da Hahn e Banach. Esso permette di applicare tecniche molto simili a quelle utilizzate su $C[a, b]$ a spazi normati generici, tanto che Banach lo introdusse affermando:
"Nous allons établir à présent quelques théorèmes qui jouent dans la théorie des espaces normes un rôle analogue à celui que le théorème de Weierstrass sur l'approximation des fonctions continues par les polynômes joue dans la théorie des fonctions de variable réelle"[b] [Ban32]
Ci limitiamo a enunciare questi due risultati.

Lemma 2.1.4. *Sia (E,‖.‖) uno spazio vettoriale normato, e sia M un suo sottospazio vettoriale. Allora se*

$$f \in E, \quad \|f - g\| \geq d > 0 \text{ per qualche } d \in \mathbb{R} \text{ e per ogni } g \in M$$

esiste un funzionale lineare continuo F che soddisfa $F(g) = 0, F(f) = 1$ e $\|F\| \leq 1/d$ per ogni $g \in M$.

[b] "Ora dimostreremo alcuni teoremi che svolgono nella teoria degli spazi normati lo stesso ruolo a quello che svolge il teorema di Weierstrass sull'approssimazione di funzioni continue con polinomi nella teoria delle funzioni a una variabile reale"

Questo lemma svolge un ruolo analogo a quello di 2.1.3 per il teorema 2.1.2 nel dimostrare il seguente

Teorema 2.1.5 (Hahn-Banach).
Sia (E,$\|.\|$) uno spazio vettoriale normato, e sia M un suo sottospazio vettoriale. Allora $f \in E$ è nella chiusura di M se e solo se ogni funzionale lineare continuo che si annulla su M si annulla anche in f.

2.2 Il teorema di Bohman-Korovkin

Nelle dimostrazioni del teorema di Weierstrass del capitolo uno abbiamo implicitamente utilizzato alcuni funzionali: ad esempio se nella dimostrazione di Fejér definiamo

$$\begin{aligned} \sigma_n : C[a,b] &\longrightarrow T_n \\ f &\longmapsto \frac{1}{n}\sum_{k=0}^{n} s_k \end{aligned} \tag{2.1}$$

dove, ricordiamo, T_n indica l'insieme dei polinomi trigonometrici di grado al più n, e s_k è la k-esima somma parziale della serie di Fourier di f, otteniamo un funzionale σ_n noto come l'n-esimo operatore di Fejér: abbiamo dimostrato che $\sigma_n(f)$ converge uniformemente a f per ogni f continua $2\pi-$periodica.
Un altro esempio si ha dalla dimostrazione probabilistica, definendo

$$\begin{aligned} B_n : C[0,1] &\longrightarrow \mathbb{R}[x]_{\leq n} \\ f &\longmapsto \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k} \end{aligned} \tag{2.2}$$

$B_n(f)$ si dice polinomio di Bernstein di f e, ancora una volta, si ha la convergenza uniforme dei $B_n(f)$ a f per ogni f in $C[0,1]$.

Avendo osservato come questo nuovo punto di vista permetta di sintetizzare il precedente, è lecito aspettarsi molti altri esempi che non derivano esplicitamente da altre teorie. Sarebbe ottimale dare una qualche condizione che, data una successione di funzionali lineari continui $\{L_n\}$, assicuri

$$\lim_{n\to+\infty} L_n(f) = f \quad \text{uniformemente}$$

La dimostrazione non probabilistica della convergenza dei polinomi di Bernstein $B_n(f)$, che non riproduciamo, dimostra con calcoli espliciti la convergenza uniforme dei polinomi di Bernstein sulle funzioni $1, x, x^2$ in $C[0,1]$

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

$$\begin{aligned}
B_n(x) =& \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} \\
=& \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k} \\
=& x \sum_{k=1}^{n} \binom{n-1}{k-1} x^{k-1} (1-x)^{n-k} = x \\
B_n(x^2) =& \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
=& \left(1 - \frac{1}{n}\right) \sum_{k=2}^{n} \binom{n-2}{k-2} x^k (1-x)^{n-k} + \frac{1}{n} \sum_{k=1}^{n} \frac{k-1}{n-1} \binom{n}{k} x^k (1-x)^{n-k} \\
=& \left(1 - \frac{1}{n}\right) x^2 + \frac{1}{n} x \xrightarrow{n \to +\infty} x^2 \quad \text{uniformemente}
\end{aligned}$$

e fa poi seguire la convergenza uniforme su tutte le funzioni continue da questo semplice fatto. Tale artificio, che può sembrare creato *ad hoc* per questa dimostrazione, è in realtà una conseguenza di un fatto molto più generale.

Infatti, analogamente, data la convergenza uniforme degli operatori di Fejér σ_n sulle tre funzioni test $1, \sin x, \cos x$

$$\begin{aligned}
\sigma_n(1) =& 1 \\
\sigma_n(\sin x) =& \frac{n}{n+1} \sin x \xrightarrow{n \to +\infty} \sin x \\
\sigma_n(\cos x) =& \frac{n}{n+1} \cos x \xrightarrow{n \to +\infty} \cos x
\end{aligned}$$

si può mostrare la convergenza uniforme sulle funzioni in $\tilde{C}[2\pi]$.

Queste osservazioni sono riassunte nel seguente teorema, che dà una condizione sufficiente affinché una successione di operatori lineari positivi generica fornisca un'approssimazione delle funzioni continue.

Teorema 2.2.1 (Bohman-Korovkin).
Sia $\{L_n\}$ una successione di operatori lineari e positivi[c] *da $C[a,b]$ in se stesso. Se*

$$\lim_{n\to+\infty} L_n(x^i) = x^i \quad \textit{uniformemente su } [a,b],\ \textit{per } i = 0,1,2$$

allora per ogni $f \in C[a,b]$

$$\lim_{n\to+\infty} L_n(f) = f \quad \textit{uniformemente su } [a,b]$$

Dimostrazione.
Sia $\varepsilon > 0$. Per il teorema di Heine-Cantor f è uniformemente continua, dunque esiste δ_ε tale che $|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$.
Per ogni $y \in [a,b]$, siano

$$\begin{aligned} p_l(x) &= f(y) - \varepsilon - \frac{2\|f\|(x-y)^2}{\delta^2} \\ p_u(x) &= f(y) + \varepsilon + \frac{2\|f\|(x-y)^2}{\delta^2} \end{aligned}$$

Poiché

$$\begin{cases} |f(x) - f(y)| < \varepsilon & \text{se } |x-y| < \delta \\ |f(x) - f(y)| < \dfrac{2\|f\|(x-y)^2}{\delta^2} & \text{se } |x-y| > \delta \end{cases}$$

allora $\forall x \in [a,b]$

$$p_l(x) \le f(x) \le p_u(x)$$

Ma allora per linearità e positività degli L_n

$$L_n(p_l)(x) \le L_n(f)(x) \le L_n(p_u)(x)$$

Fissata f, p_l e p_u sono polinomi di secondo grado nella variabile x dipendenti dal parametro y: esplicitamente

$$p_u(x) = \left(f(y) + \varepsilon + \frac{2\|f\|y^2}{\delta^2}\right) - \left(\frac{4\|f\|y}{\delta^2}\right)x + \left(\frac{2\|f\|}{\delta^2}\right)x^2$$

[c] un operatore lineare L si dice positivo se $f \ge 0 \Rightarrow L(f) \ge 0$

$$p_l(x) = \left(f(y) - \varepsilon - \frac{2\|f\|y^2}{\delta^2}\right) + \left(\frac{4\|f\|y}{\delta^2}\right)x - \left(\frac{2\|f\|}{\delta^2}\right)x^2$$

Poichè per linearità la convergenza uniforme su $[a, b]$ degli L_n su $1, x, x^2$ implica la convergenza uniforme su $[a, b]$ di ogni polinomio di secondo grado in x, esiste $N \in \mathbb{N}$ tale che per ogni scelta di y in $[a, b]$ si abbia

$$\begin{cases} |L_n(p_u)(x) - p_u(x)| < \varepsilon \\ |L_n(p_l)(x) - p_l(x)| < \varepsilon \end{cases}$$

Valutando i due polinomi in x=y si ha

$$L_n(p_u)(y) < p_u(y) + \varepsilon = f(y) + 2\varepsilon$$
$$L_n(p_l)(y) > p_l(y) - \varepsilon = f(y) - 2\varepsilon$$

per ogni $n > N$. Ma allora

$$f(y) - 2\varepsilon < L_n(f)(y) < f(y) + 2\varepsilon$$

che per arbitrarietà di ε dimostra il teorema. □

Esiste anche una versione del teorema per $f \in \tilde{C}[2\pi]$, dove $\{1, x, x^2\}$ vengono sostituite da $\{1, \sin x, \cos x\}$, che fu dimostrata da Korovkin in maniera sostanzialmente analoga.

Il teorema di Bohman-Korovkin, che utilizzando solo la convergenza su tre particolari funzioni assicura la convergenza uniforme su tutte le funzioni continue, fa sorgere alcuni interrogativi. Infatti ci si può domandare se è possibile rimpiazzare il sistema $\{1, x, x^2\}$ con altre funzioni, e se ad esempio possano bastarne solo due. A tali questioni si è dedicato principalmente Korovkin, che arrivò alle seguenti conclusioni

Teorema 2.2.2 (Terzo teorema di Korovkin [d]).
Se un sistema di funzioni $\{f_1, ..., f_n\} \subset C[0, 1]$ *svolge un ruolo analogo a quello di* $\{1, x, x^2\}$ *nel teorema 2.2.1, allora* $n > 2$

[d] in questa nomenclatura il primo è quello da noi denominato “Bohman-Korovkin”, e il secondo è la versione trigonometrica di quest’ultimo

Teorema 2.2.3. *Il sistema* $\{f_0, f_1, f_2\} \subset C[0,1]$ *svolge un ruolo analogo a quello di* $\left\{1, x, x^2\right\}$ *nel teorema 2.2.1 se e solo se è un sistema di Chebyshev di ordine due, ovvero se span* $\{f_0, f_1, f_2\}$ *è un sottospazio di dimensione tre su* $C[0,1]$ *i cui elementi hanno al più due zeri distinti nell'intervallo.*

Entrambi questi risultati (con le relative dimostrazioni) si possono trovare in [Kor59] così come pubblicati dall'autore e in [Bau78, AC94] riformulati all'interno di una teoria più moderna su questo tipo di approssimazione.

Capitolo 3

Il teorema di Müntz

3.1 Enunciato e dimostrazione classica

Una volta dimostrato che i polinomi approssimano le funzioni continue, è lecito chiedersi se non esistano insiemi di funzioni ancora più particolari con la stessa proprietà. Una risposta può essere trovata cercando sottospazi densi nei polinomi.

Esempio 3.1.1. Data una funzione $f(x)$ in $C[0,1]$, sia $x = t^\alpha$ con α numero reale positivo. Poiché l'immagine di $[0,1]$ tramite $\tau(t) = t^\alpha$ è ancora $[0,1]$, la funzione $F(t) = f(t^\alpha)$ è una funzione in $C[0,1]$ rispetto alla variabile t. Per il teorema di Weierstrass allora, per ogni $\varepsilon > 0$ dato, esiste un polinomio $P(t)$ tale che $|f(t) - p(t)| < \varepsilon \quad \forall x \in [0,1]$.
Esiste dunque una funzione $p(x) = P(x^{1/\alpha})$ in $C[0,1]$ che approssima $f(x)$ con la precisione richiesta; in altre parole, è possibile approssimare ogni funzione continua in $[0,1]$ con elementi dello spazio $\mathbb{R}[x^{1/\alpha}]$.
In particolare nel caso $\alpha = 1/3$ si ha che $\mathbb{R}[x^3]$ è un sottospazio dei polinomi con la proprietà cercata.

Più in generale, identificando $\mathbb{R}[x]$ con $span\{1, x, x^2, x^3, ...\}$, si può analogamente identificare ogni $A \subseteq \mathbb{R}[x]$ con $span\{x^{\lambda_0}, x^{\lambda_1}, x^{\lambda_2}, ...\}$, associando dunque a ogni successione di numeri naturali un sottospazio dei polinomi.
Si può dunque equivalentemente formulare un'altra domanda: quali condizioni deve soddisfare una successione $\{\lambda_i\}$ affinché gli elementi del corrispondente spazio approssimino le funzioni continue in $[0,1]$? La risposta, sorprendentemente elegante, è fornita dal teorema di Müntz.

Teorema 3.1.2 (Müntz classico)**.**
Sia $\{\lambda_i\}$ *una successione crescente di numeri reali non negativi.*
Lo spazio vettoriale $\Lambda = span\{x^{\lambda_0}, x^{\lambda_1}, x^{\lambda_2}, ...\}$ *è denso in* $C[0;1]$ *se e solo se*

- $\lambda_0 = 0$
- $\displaystyle\sum_{k=1}^{+\infty} \frac{1}{\lambda_k} = \infty$

Premettiamo alla dimostrazione un lemma tecnico sui prodotti infiniti, che sarà utilizzato diverse volte.

Lemma 3.1.3. *Sia* $\{a_k\}$ *una successione infinitesima di numeri reali non negativi. Allora*

- $\displaystyle\prod_{k=0}^{+\infty}(1+a_k)$ *converge/diverge* $\iff \displaystyle\sum_{k=0}^{+\infty} a_k$ *converge/diverge*
- *Se* $a_k \neq 1\ \forall i,\quad \displaystyle\prod_{k=0}^{+\infty}(1-a_k) = 0 \iff \sum_{k=0}^{+\infty} a_k = +\infty$

Dimostrazione.

$$\prod_{k=0}^{+\infty}(1+a_k) = \lim_{n\to+\infty}\prod_{k=0}^{n}(1+a_k) = \lim_{n\to+\infty}\exp\left(\log\left(\prod_{k=0}^{n}(1+a_k)\right)\right)$$

cioè

$$\prod_{k=0}^{+\infty}(1+a_k) = \exp\left(\lim_{n\to+\infty}\sum_{k=0}^{n}\log(1+a_k)\right)$$

Allora per confronto asintotico[a] la convergenza/divergenza della produttoria è equivalente alla convergenza/divergenza di $\displaystyle\sum_{k=0}^{n} a_k$.
Dimostriamo ora la seconda parte dell'enunciato.
Essendo $\{a_k\}$ una successione infinitesima, possiamo assumere senza perdita di generalità $a_k \leq \frac{1}{2}$ scartando al più un numero finito di termini. Allora ogni fattore $(1-a_k) \geq \frac{1}{2}$.
Definendo

$$p_n = \prod_{k=0}^{n}(1-a_k) \quad \text{e} \quad q_n = \prod_{k=0}^{n}(1+a_k)$$

[a] essendo $\displaystyle\lim_{x\to 0}\frac{\log(1+x)}{x} = 1$

e essendo $(1-a_j)(1+a_j) = 1-a_j^2 \leq 1$, si ha

$$p_n \leq \frac{1}{q_n}$$

Per la prima parte del lemma $q_n \xrightarrow{n\to\infty} +\infty$ se e solo se $\sum a_k$ è una serie divergente. Allora

$$\sum_{k=0}^{+\infty} a_k = +\infty \quad \Rightarrow \quad \lim_{n\to+\infty} p_n = 0$$

Per l'altro verso dell'implicazione è sufficiente considerare che se $\sum a_k$ converge allora $\prod(1-a_k)$ converge assolutamente, e dunque $\prod(1-a_k)$ converge semplicemente. Per una dimostrazione del legame tra convergenza assoluta e convergenza semplice nei prodotti infiniti si veda [Apo74, th. 8.54] □

La prima parte della dimostrazione che segue è quella dell'articolo originale in cui Müntz pubblicò il teorema. Per dimostrare che la condizione è anche sufficiente utilizziamo invece la dimostrazione di Szàzs, ben più elegante e diretta di quella originale.

Dimostrazione classica (di Müntz-Szàzs).
Sia $\Lambda_n = span\{x^{\lambda_0}, x^{\lambda_1}, ..., x^{\lambda_n}\}$, e sia $E(f, \Lambda_n)_\infty = \min_{p\in\Lambda_n} \|f-p\|_\infty$.
Una condizione equivalente alla tesi è

$$\lim_{n\to+\infty} E(x^m, \Lambda_n)_\infty = 0 \quad \forall m \in \mathbb{N} \tag{3.1}$$

Infatti è sufficiente che gli elementi di Λ_n siano densi nei polinomi, a loro volta densi in C[0,1] per il teorema di Weierstrass. Allo stesso modo, essendo le potenze di x funzioni continue, tale condizione è anche necessaria.
Osserviamo inoltre che $\lambda_0 = 0$ è una condizione indispensabile: senza di essa ogni funzione in Λ si annullerebbe per $x = 0$, e Λ non potrebbe dunque essere denso in $C[0,1]$. Rimane da mostrare che la seconda condizione è sia necessaria che sufficiente.
Definiamo

$$E(f, \Lambda_n)_2 = \min_{p\in\Lambda_n} \|f-p\|_2$$

dove $\|.\|_2$ è la norma di $L^2[0,1]$. Poiché $E(f, \Lambda_n)_2 \leq E(f, \Lambda_n)_\infty$, una condizione necessaria affinché valga 3.1 è

$$\lim_{n\to+\infty} E(x^m, \Lambda_n)_2 = 0 \quad \forall m \in \mathbb{N} \tag{3.2}$$

Vogliamo anzitutto capire quali $\{\lambda_i\}$ soddisfano la relazione 3.2.

Poiché $E(f, \Lambda_n)_2$ è la proiezione ortogonale di f sul sottospazio Λ_n

$$E(f, \Lambda_n)_2^2 = \frac{G(x^{\lambda_0}, ..., x^{\lambda_n}, f)}{G(x^{\lambda_0}, ..., x^{\lambda_n})}$$

dove

$$G(f_1, ..., f_k) = \det \begin{pmatrix} \langle f_1, f_1 \rangle & \cdots & \langle f_1, f_k \rangle \\ \vdots & \ddots & \vdots \\ \langle f_k, f_1 \rangle & \cdots & \langle f_k, f_k \rangle \end{pmatrix}$$

Ricordando

$$\langle x^a, x^b \rangle = \int_0^1 x^a x^b dx = \frac{1}{a+b+1}$$

e

$$\det \begin{pmatrix} \frac{1}{a_1+b_1+1} & \cdots & \frac{1}{a_1+b_k+1} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_k+b_1+1} & \cdots & \frac{1}{a_k+b_k+1} \end{pmatrix} = \frac{\prod\limits_{1 \le j < i \le k} (a_i - a_j)(b_i - b_j)}{\prod\limits_{i,j=1}^{k} (a_i + b_j + 1)}$$

si ha

$$E(x^m, \Lambda_n)_2^2 = \frac{\prod\limits_{k=0}^{n} (m - \lambda_k)^2}{(2m+1) \prod\limits_{k=0}^{n} (m + \lambda_k + 1)^2}$$

Grazie a quest'ultima uguaglianza possiamo concludere

$$\lim_{n \to +\infty} E(x^m, \Lambda_n)_2 = 0 \iff \lim_{n \to +\infty} \prod_{k=0}^{n} \frac{m - \lambda_k}{m + \lambda_k + 1} = 0$$

che è equivalente a richiedere

$$\prod_{k=0}^{\infty} \frac{m - \lambda_k}{m + \lambda_k + 1} = 0$$

che possiamo riscrivere come

$$\prod_{k=0}^{\infty} \left(1 - \frac{2m+1}{m + \lambda_k + 1} \right) = 0 \tag{3.3}$$

Supponendo senza perdita di generalità $m \neq \lambda_k$ $\forall k$ (altrimenti 3.2 è automaticamente soddisfatta, essendo $x^m \in \Lambda_n$ per qualche n), si ottiene

$$1 \neq \frac{2m+1}{m+\lambda_k+1} \quad , \quad \lim_{k\to+\infty} \frac{2m+1}{m+\lambda_k+1} = 0$$

Allora per il lemma 3.1.3, 3.3 è equivalente a

$$\sum_{k=0}^{\infty} \frac{2m+1}{m+\lambda_k+1} = \infty$$

che per confronto asintotico è a sua volta equivalente a

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty \tag{3.4}$$

Abbiamo dunque ottenuto una condizione necessaria e sufficiente per la densità con la norma di $L^2[0;1]$ che non dipende da m.
Supponiamo ora che la serie $\sum_{k=1}^{\infty} \frac{1}{\lambda_k}$ diverga, e che $\lambda_0 = 0$.
Mostriamo come ogni x^m si possa approssimare con la norma infinito. Per $x \in [0;1]$

$$\begin{aligned} \left| x^m - \sum_{k=1}^{n} a_k x^{\lambda_k} \right| &= \left| \int_0^x \left(m t^{m-1} - \sum_{k=1}^{n} a_k \lambda_k t^{\lambda_k - 1} \right) dt \right| \\ &\leq \int_0^1 \left| m t^{m-1} - \sum_{k=1}^{n} a_k \lambda_k t^{\lambda_k - 1} \right| dt \\ &\leq \left(\int_0^1 \left| m t^{m-1} - \sum_{k=1}^{n} a_k \lambda_k t^{\lambda_k - 1} \right|^2 dt \right)^{1/2} \end{aligned}$$

ovvero possiamo approssimare x^m nella norma infinito con combinazioni lineari di $\{x^{\lambda_1}, x^{\lambda_2}, ...\}$ se riusciamo a approssimare x^{m-1} nella norma di $L^2[0;1]$ con combinazioni lineari di $\{x^{\lambda_1-1}, x^{\lambda_2-1}, ...\}$. Per quanto mostrato in precedenza, questo è vero se e solo se

$$\sum_{k\geq k_0} \frac{1}{\lambda_k - 1} = \infty$$

con k_0 tale che $\lambda_{k_0} - 1 > 0$.
Poichè $\{\lambda_i\}$ è una successione crescente non limitata, quest'ultima condizione è automaticamente implicata da 3.4. Allora la condizione necessaria 3.4 è anche sufficiente. $\square$

3.2 Dimostrazione di Rudin

Il percorso più breve tra due verità nel dominio reale
passa per il dominio complesso.
JACQUES HADAMARD

Una seconda dimostrazione del Teorema di Müntz, dovuta a Rudin, necessita di alcuni strumenti dell'analisi complessa. Infatti, come si vedrà, il problema della densità degli x^{λ_i} può essere collegato all'esistenza di una funzione olomorfa nel disco unitario avente zeri fissati. Per una trattazione più approfondita di questi argomenti si fa riferimento a [Rud66] o [Rem98].
Tutte le funzioni in questo paragrafo si intendono definite in $\mathbb{C}$ o in suoi sottoinsiemi.

Definizione 3.2.1. Sia $U \subset \mathbb{C}$ un aperto. $f : U \to \mathbb{C}$ si dice *differenziabile in senso complesso* in un punto z_0 di U se esiste

$$\lim_{z \to z_0} \frac{f(z) - f(z_o)}{z - z_0}$$

f si dice *olomorfa* in U se è differenziabile in senso complesso in tutti i punti di U.
f si dice *meromorfa* in U se è olomorfa in tutto U ad eccezione di un insieme di punti isolati, nei quali f ha dei poli (z_0 si dice *polo* per f se $\lim_{z \to z_0} |f(z)| = \infty$)

Teorema 3.2.2 (Formula integrale di Cauchy)**.** *Sia $f(z)$ una funzione olomorfa in un dominio semplicemente connesso $U \subset \mathbb{C}$. Allora per ogni curva semplice chiusa percorsa in senso antiorario γ contenuta in U, detto S l'insieme tale che $\partial S = \gamma$, vale*

$$f(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{z - z_0}\, dz \quad \forall z_0 \in S$$

Lemma 3.2.3 (Formula di Jensen)**.**
Sia f una funzione olomorfa su $D_R(0)$ con $f(0) \neq 0$.
Sia inoltre $0 < r < R$, e siano $\alpha_1, ...\alpha_N$ gli zeri di f in $\overline{D}_r(0)$ elencati secondo la loro molteplicità. Allora

$$|f(0)| \prod_{k=1}^{N} \frac{r}{\alpha_k} = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log |f(re^{i\theta})| d\theta \right)$$

Dimostrazione.
Senza perdita di generalità possiamo supporre che gli α_k siano ordinati in modo che

$\alpha_1, ..., \alpha_m$ siano in $D_r(0)$, e $\alpha_{m+1}, ..., \alpha_N$ siano sulla frontiera di $D_r(0)$ [b]. Sia

$$g(z) = f(z) \prod_{k=1}^{m} \frac{r^2 - \overline{\alpha_k} z}{r(\alpha_k - z)} \prod_{k=m+1}^{N} \frac{\alpha_k}{\alpha_k - z} \tag{3.5}$$

g è olomorfa in $D_{r+\varepsilon}(0)$ per qualche $\varepsilon > 0$, e inoltre g non si annulla mai in $D_{r+\varepsilon}(0)$. Allora $\log |g|$ è armonica in $D_{r+\varepsilon}(0)$ ([Rud66, th. 13.12]), perciò

$$\log |g(0)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |g(re^{i\theta})| d\theta \tag{3.6}$$

Poichè

$$|g(0)| = |f(0)| \prod_{k=1}^{m} \frac{r}{\alpha_k}$$

Per $1 \leq i \leq m$ i fattori in 3.5 hanno modulo 1 se $|z| = r$; se $n > m$, sia $\alpha_i = re^{i\theta_i}$

$$\log |g(re^{i\theta})| = \log |f(re^{i\theta})| - \sum_{k=m+1}^{N} \log \left| 1 - e^{i(\theta - \theta_k)} \right|$$

Poiché $\int_0^{2\pi} \ln \left| 1 - e^{it} \right| dt = 0$ (si veda [Rud66, th.15.17]), allora

$$\log |g(0)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |f(re^{i\theta})| d\theta$$

ovvero

$$\log \left(|f(0)| \prod_{k=1}^{m} \frac{r}{\alpha_k} \right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |f(re^{i\theta})| d\theta$$

ovvero la tesi. □

Teorema 3.2.4.
Sia f una funzione olomorfa limitata nel disco unitario, e siano $\{\alpha_1, ..., \alpha_k, ...\}$ gli zeri di f elencati secondo la loro molteplicità. Allora, se f non è identicamente nulla,

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

[b] ovviamente, nulla vieta m=0 oppure m=N

Dimostrazione.
Se f ha un numero finito di zeri nel disco unitario, non c'è nulla da dimostrare. Supponiamo pertanto che f abbia un numero infinito di zeri, ordinati in modo che valga $|\alpha_k| \leq |\alpha_{k+1}|$ per ogni k. Se f ha uno zero di ordine m nell'origine, possiamo considerare $g(z) = z^{-m} f(z)$, che ha lo stesso numero di zeri di f al di fuori dell'origine e una singolarità eliminabile nell'origine. Senza perdita di generalità possiamo dunque supporre $f(0) \neq 0$. Sia $n(r)$ il numero di zeri di f in $\overline{D}_r(0)$. Fissato $w > 0$, sia r tale che $w < n(r)$. Allora per la formula di Jensen

$$|f(0)| \prod_{k=1}^{w} \frac{r}{|\alpha_k|} \leq \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|f(re^{i\theta})\right| \, d\theta\right)$$

Essendo f limitata, esiste una costante $C \in \mathbb{R}$ per cui

$$\prod_{k=1}^{n} |\alpha_k| \geq C^{-1} |f(0)| \, r^w$$

Poiché tale disuguaglianza resta vera per ogni w quando $r \to 1$, si ha

$$\prod_{k=1}^{+\infty} |\alpha_k| \geq C^{-1} |f(0)| > 0$$

Essendo dunque il prodotto infinito non nullo, osservando che $|\alpha_k| = 1 - (1 - |\alpha_k|)$, per il lemma 3.1.3 si ha la tesi. □

Teorema 3.2.5 (sui prodotti di Blaschke)**.**
Sia $\{\alpha_k\}$ una successione di numeri complessi nel disco unitario con $\alpha_k \neq 0 \; \forall n$ e

$$\sum_{k=1}^{+\infty} (1 - |\alpha_k|) < \infty,$$

Sia inoltre $w \geq 0$, z nel disco unitario. Allora

$$B(z) = z^w \prod_{k=1}^{+\infty} \frac{\alpha_k - z}{1 - \overline{\alpha_k} z} \frac{|\alpha_k|}{\alpha_k}$$

è una funzione olomorfa e limitata nel disco unitario, e si annulla solo in $z = \alpha_k$ oppure in $z = 0$ (se $w > 0$)

B(z) si dice prodotto di Blaschke *della successione* $\{\alpha_k\}$.

Dimostrazione.
Premettiamo due osservazioni:

- alcuni degli α_k potrebbero essere ripetuti: in tal caso B(z) ha zeri di molteplicità superiore a 1 in quei punti.
- ciascun fattore nella produttoria che definisce B(z) ha modulo 1 sulla circonferenza unitaria

L'n-esimo termine della serie

$$\sum_{k=1}^{+\infty} \left| 1 - \frac{\alpha_k - z}{1 - \overline{\alpha_k} z} \frac{|\alpha_k|}{\alpha_k} \right|$$

si può riscrivere come

$$\left| \frac{\alpha_k + |\alpha_k|\, z}{(1 - \overline{\alpha_k} z)\alpha_k} \right| (1 - |\alpha_k|) \leq \frac{1+r}{1-r} (1 - |\alpha_k|)$$

se $|z| \leq r$, dunque la serie converge per ipotesi. Allora, utilizzando vale un risultato analogo al lemma 3.1.3 per le successioni di funzioni ([Rud66, th. 15.4, 15.6]), si ha che $B(z)$ è una funzione olomorfa nel disco unitario e si annulla solo in $z = \alpha_k$ o $z = 0$.
Poiché inoltre ciascun fattore del prodotto infinito ha modulo minore di 1, risulta anche $|B(z)| < 1$. □

Osservazione 3.2.6. Gli ultimi due teoremi permettono di affermare che, data una successione di numeri complessi $\{\alpha_k\}_{k \in \mathbb{N}}$ nel disco unitario, una condizione necessaria e sufficiente all'esistenza di una funzione olomorfa e limitata nel disco unitario avente come unici zeri gli α_k è

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

Prima di esibire la dimostrazione di Rudin, infine, ricordiamo un risultato classico dell'analisi funzionale (si veda [Car00, §14])

Teorema 3.2.7 (Teorema di rappresentazione di Riesz). *Sia X uno spazio di Hausdorff localmente compatto. Allora ad ogni funzionale lineare limitato Φ su $C(X)$ corrisponde una e una sola misura complessa di Borel regolare μ tale che*

$$\Phi(f) = \int_X f d\mu \quad \forall f \in C(X)$$

Dimostrazione di Rudin del Teorema di Müntz.
Supponiamo che $\Lambda = span\{x^{\lambda_0}, x^{\lambda_1}, x^{\lambda_2}, ...\}$ non sia denso in C[0,1].
La condizione $\lambda_0 = 0$ è indispensabile, altrimenti ogni funzione in Λ si annullerebbe in 0, rendendo impossibile approssimare le funzioni costanti.
Per il teorema 2.1.5 esiste una funzione $F \in C[0,1]$, $F \notin \overline{\Lambda}$ se e solo se esistono un funzionale lineare limitato non nullo Φ su $C[0,1]$ e una funzione F in $C[0,1]$ tale che

$$\Phi(F) \neq 0 \quad , \quad \Phi(f) = 0 \quad \forall f \in \Lambda$$

Poiché per il teorema 3.2.7 ogni funzionale lineare limitato su $C[0,1]$ si ottiene come integrale di una certa misura complessa su $[0,1]$, allora esiste una misura complessa di Borel regolare (avente variazione totale limitata, i.e. $|\mu| < \infty$) su $[0,1]$ μ non nulla con la proprietà :

$$\int_0^1 x^{\lambda_n} d\mu(x) = 0 \quad \text{per } n = 1, 2, ...$$

Poiché le funzioni integrande si annullano in 0, è lecito assumere che μ sia concentrata in $(0,1]$. Sia ora

$$f(z) = \int_0^1 x^z d\mu(x)$$

Per $x \in [0,1)$ e $\operatorname{Re} z > 0$ si ha $x^z = e^{z \ln(x)}$ e $|x^z| = x^{\operatorname{Re} z} \leq 1$
Inoltre poiché f è continua, applicando il teorema di Morera[c] si ricava che f è olomorfa e limitata in $\{z \in \mathbb{C}, z = x + iy : x > 0\} \subseteq \mathbb{C}$, cioè nel semipiano destro. Inoltre soddisfa

$$f(\lambda_n) = 0 \quad \text{per } n = 1, 2, ...$$

[c] una conseguenza della formula di Cauchy. Si veda [Rud66, th.10.17]

Sia ora

$$g(z) = f\left(\frac{1+z}{1-z}\right), \quad z \in D_1(0)$$

Poichè la trasformazione $t : z \to \frac{1+z}{1-z}$ mappa conformemente il disco unitario nel semipiano destro, anche $g = f \circ t$ è olomorfa e limitata nel suo dominio $D_1(0)$.
Osserviamo inoltre che

$$g\left(\frac{\lambda_n - 1}{\lambda_n + 1}\right) = 0 \quad \text{per } n = 1, 2, ...$$

Allora, detto $\alpha_n = \dfrac{\lambda_n - 1}{\lambda_n + 1}$, per l'osservazione 3.2.6 gli $\{\alpha_n\}$ possono essere gli zeri nel disco unitario di una g olomorfa e limitata non banale se e solo se

$$\sum_{n=1}^{+\infty}(1 - |\alpha_n|) < +\infty \tag{3.7}$$

Detto N il primo indice per cui $\lambda_N \geq 1$, si ha

$$\sum_{n=1}^{+\infty}\left(1 - \left|\frac{\lambda_n - 1}{\lambda_n + 1}\right|\right) = \sum_{n=1}^{N}\frac{2\lambda_n}{\lambda_n + 1} + \sum_{n=N+1}^{+\infty}\frac{2}{\lambda_n + 1} \geq (\lambda_1 - \lfloor\lambda_1\rfloor)\sum_{n=1}^{+\infty}\frac{1}{\lambda_n + 1}$$

da cui si ha per confronto asintotico che 3.7 non può essere verificata, essendo la serie dei reciproci dei λ_n divergente.
Allora necessariamente $g(z) = 0 \quad \forall z \in D_1(0)$, e quindi $f = 0$.
Essendo

$$0 = f(k) = \int_0^1 x^k d\mu(x) \quad \text{per ogni } k = 1, 2, ...$$

per il teorema di Weierstrass $\mu = 0$. Poiché inizialmente avevamo supposto μ non nulla, siamo giunti a un assurdo: allora Λ deve essere denso in $C[0, 1]$.
Abbiamo dunque dimostrato che la condizione 3.4 è sufficiente.

Rimane da mostrare che se $\displaystyle\sum_{k=1}^{+\infty}\frac{1}{\lambda_k} < \infty$ allora Λ non è denso in $C[0, 1]$, ovvero che 3.4 è anche necessaria.
Definiamo

$$f(z) = \frac{1}{(z+2)^2}\prod_{n=0}^{+\infty}\frac{\lambda_n - z}{2 + \lambda_n + z}$$

f è ben definita perchè la convergenza di quel prodotto è equivalente all'ipotesi (sempre

per l'osservazione 3.2.6).
Inoltre f è una funzione meromorfa che si annulla in $z = \lambda_n$. con poli in $z = -2$ e $z = -\lambda_n - 2$, e f è anche limitata in $\mathrm{Re}(z) > -1$, essendo ogni fattore minore di 1 in modulo.
Allora per la formula integrale di Cauchy, se $\mathrm{Re}(z) > -1$

$$f(z) = \frac{1}{2\pi i}\int_{\Gamma_R}\frac{f(w)}{w-z}dw$$

dove Γ_R è la frontiera di un dominio, ed è composta dall'unione della semicirconferenza destra centrata in -1 avente raggio $R > 1 + |z|$ e del segmento $[-1 - iR, -1 + iR]$.
Si verifica facilmente che se $R \to +\infty$ l'integrale sulla semicirconferenza tende a 0; rimane perciò solo quello lungo il segmento

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{f(-1+is)}{1-is+z}ds$$

che, essendo $\displaystyle\int_0^1 x^{z-is}dx = \frac{1}{1+z-is}$, si può riscrivere come

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(-1+is)\left(\int_0^1 x^{z-is}dx\right)ds$$
$$f(z) = \int_0^1 x^z\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} f(-1+is)e^{-is\ln x}ds\right)dx$$

dove lo scambio dell'ordine di integrazione è lecito, essendo la funzione integranda in L^1.[d]
Definiamo allora

$$d\mu(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(-1+is)e^{-is\ln x}ds$$

Questa misura è limitata e continua su $(0;1]$, poiché la misura μ è definita dalla trasformata di Fourier[e] della funzione $f(-1+iz)$ in $z = \ln x$
Essendo

$$f(z) = \int_0^1 x^z d\mu(x)$$

la misura μ si annulla su tutti gli x^{λ_n}, ma non su x^t se $t \neq \lambda_n$.
Per il teorema 2.1.5 allora Λ non è denso in $C[0,1]$. La condizione 3.4, assieme a $\lambda_0 = 0$,

[d] il che è garantito dalla presenza del fattore $\frac{1}{(z-2)^2}$ nella definizione di f
[e] si veda [Rud66, cap. 9]

è quindi necessaria e sufficiente.

□

3.3 Sviluppi del teorema di Müntz

Abbiamo dimostrato il teorema di Müntz in $C[0,1]$, con $\lambda_0 = 0$ e $\{\lambda_i\}$ successione positiva, monotona e divergente. Come per il teorema di Weierstrass, è opportuno chiedersi se e con quali ipotesi aggiuntive sia possibile generalizzare il teorema a spazi di funzioni diversi e/o con successioni che non rispettano quelle ipotesi.
Le generalizzazioni del teorema di Müntz, tuttavia, risultano molto spesso ben più complicate. Potremmo supporre, ad esempio, che aver dimostrato il teorema in $[0,1]$ ci permetta una facile generalizzazione a un intervallo $[a,b]$: ciò è falso.
Infatti in questo caso l'isometria definita in 1.2.2 non può essere utilizzata, in quanto se $x \longmapsto \frac{x-a}{b-a}$, allora ad esempio

$$x^2 \longmapsto \left(\frac{x-a}{b-a}\right)^2 = \frac{1}{(b-a)^2}x^2 - \frac{2a}{(b-a)^2}x + \frac{a^2}{(b-a)^2}$$

Se però la successione $\{\lambda_i\}$ non conteneva 1, l'immagine di x^2 tramite l'isometria non è più un elemento dello spazio vettoriale generato dagli x^{λ_i}.
Poiché per la formula del binomio di Newton $(x-a)^n$ è un polinomio completo di grado n per ogni a non nullo, ne deduciamo che l'isometria non permette di estendere il teorema di Müntz a un intervallo generico per nessuna successione di $\{\lambda_i\}$ che non comprenda tutti i numeri naturali (ovvero vale solo quando il teorema è quello di Weierstrass).
L'isometria può tuttavia essere applicata per mostrare che il teorema vale in ogni intervallo del tipo $[0,b]$ con $b > 0$. Infatti in tal caso essa si limita a modificare i coefficienti di un fattore $1/b^{\lambda_n}$, e dunque l'immagine di un elemento di Λ rimane in Λ.
Ovviamente le considerazioni appena fatte non escludono in alcun modo che il teorema possa comunque essere vero in $[a,b]$ qualsiasi. Un altro facile esempio, tuttavia, mostra che non è così.

Esempio 3.3.1. Sia $[a,b] = [-1,1]$, $\lambda_k = 2k$. Si verifica immediatamente che questa successione soddisfa le ipotesi del teorema di Müntz, essendo

$$\begin{cases} \lambda_0 = 1 \\ \displaystyle\sum_{k=1}^{+\infty} \frac{1}{2k} = +\infty \end{cases}$$

Tuttavia ogni elemento di Λ è della forma $p(x) = \sum_{k=0}^{n} a_k x^{2k}$ per qualche n, e perciò ha la proprietà $p(x) = p(-x)$ per ogni $x \in [-1,1]$, cioè è pari. Ma allora risulta impossibile trovare un modo per approssimare una qualsiasi funzione dispari, ad esempio $f(x) = x$, e questo costituisce un controesempio al teorema di Müntz in $[-1,1]$.

Il caso appena esaminato ci permette di comprendere che per estendere il teorema di Müntz su un generico $[a,b]$ reale è indispensabile integrare o modificare le ipotesi sulla successione dei $\{\lambda_i\}$. La prima osservazione è che il punto 0 è speciale, poiché in esso tutte le funzioni x^λ si annullano qualunque sia $\lambda \neq 0$. La presenza del punto 0 nell'intervallo $[a,b]$ equivale allora alla presenza di un k tale che $\lambda_k = 0$ nella successione dei $\{\lambda_i\}$.
Il modo corretto di generalizzare il teorema a un intervallo $[-a,b]$ con $a,b > 0$ è il seguente

Teorema 3.3.2 (Müntz in $[a,b], a < 0 < b$)**.**
*Sia $a < 0 < b$, e sia $0 = \lambda_0 < \lambda_1 < \ldots$ una sequenza crescente di numeri naturali.
Allora $\Lambda = span\left\{1, x^{\lambda_1}, x^{\lambda_2}, \ldots\right\}$ è denso in $C[a,b]$ se e solo se*

$$\sum_{\substack{i=1 \\ \lambda_i \ pari}}^{\infty} \frac{1}{\lambda_i} = \sum_{\substack{i=1 \\ \lambda_i \ dispari}}^{\infty} \frac{1}{\lambda_i} = \infty$$

Dimostrazione. Sia $\{\lambda_i\} = \{\pi_j\} \cup \{\delta_k\}$, dove i π_j sono gli esponenti pari presenti nella successione e i δ_k i dispari. Siano inoltre $\Pi_n = span\{1, x^{\pi_1}, \ldots, x^{\pi_n}\}$ e $\Delta_m = span\left\{x^{\delta_1}, \ldots, x^{\delta_m}\right\}$.
Si supponga senza perdita di generalità $|a| < |b|$, e sia $I = [-b,b]$.
Ogni funzione si può decomporre nella somma di una funzione pari e una dispari[f], essendo

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} = p(x) + d(x)$$

[f]si può dimostrare facilmente che tale decomposizione di $C(\mathbb{R})$ è in realtà una decomposizione in somma diretta, essendo f(x)=0 l'unica funzione contemporaneamente pari e dispari

Sia $\varepsilon > 0$ fissato.
Applicando il teorema di Müntz all'intervallo $[0, b]$ si ha che esiste un n opportuno e un polinomio $P(x) \in \Pi_n$ tale che

$$\max_{x\in[0,b]} |p(x) - P(x)| < \varepsilon$$

se e solo se $\sum 1/\pi_j = \infty$.
Poiché $P(x)$ e $p(x)$ sono entrambe funzioni pari

$$\max_{x\in[0,b]} |p(x) - P(x)| = \max_{x\in[-b,b]} |p(x) - P(x)|$$

Con un ragionamento analogo, osservando che $d(0) = 0$ per costruzione e $\forall D(x) \in \Delta_m$ $D(0) = 0$, si dimostra che esiste un m opportuno e un polinomio $D(x) \in \Delta_n$ tale che

$$\max_{x\in[-b,b]} |d(x) - D(x)| < \varepsilon$$

se e solo se $\sum 1/\delta_k = \infty$.
Allora, detto $P(x) + D(x) = F(x) \in \Lambda_{n+m} = \Pi_n + \Delta_m$, vale

$$\max_{x\in[-b,b]} |f(x) - F(x)| < 2\varepsilon$$

cioè la tesi. □

Per ulteriori approfondimenti e per il caso $0 < a < b$ si veda [Bor95], che contiene anche molti lemmi tecnici che consentono di considerare una classe di successioni più vasta di quella da noi considerata.

Un'altra possibilità di generalizzazione consiste nel cercare sotto quali condizioni una successione crescente di esponenti di x reali possa generare un sottospazio denso negli spazi $L^p[0, 1]$. Osserviamo che nella dimostrazione classica del teorema di Müntz abbiamo dimostrato che la condizione 3.4 è sufficiente per la densità in $L^2[0, 1]$. Come vedremo a breve, però, essa è un caso particolare di una condizione necessaria leggermente più generale.
Diamo la dimostrazione del teorema di Müntz in $L^1[0, 1]$, dovuta a Borwein [Bor91].

Teorema 3.3.3. *Sia $\{\lambda_i\}$ una successione di numeri reali positivi distinti maggiori di -1,* *e sia* $\Lambda = span\left\{x^{\lambda_0}, x^{\lambda_1}, ...\right\}$*. Allora* Λ *è denso in* $L^1[0,1]$ *se e solo se*

$$\sum_{i=0}^{\infty} \frac{\lambda_i + 1}{(\lambda_i + 1)^2 + 1}$$

Dimostrazione.
Si assuma Λ denso in $L^1[0,1]$. Sia $m \geq 0$ un intero, e sia $\varepsilon > 0$. Allora è possibile scegliere $p \in \Lambda$ con la proprietà

$$\|x^m - p(x)\|_\infty < \varepsilon$$

Sia ora

$$q(x) := \int_0^x p(t)\ dt$$

Necessariamente $q \in \Lambda^* = span\left\{1, x^{\lambda_0+1}, x^{\lambda_1+1}, ...\right\}$, da cui segue

$$\left\| \frac{x^{m+1}}{m+1} - q(x) \right\|_\infty < \varepsilon$$

Allora si riesce a approssimare ogni polinomio con elementi di Λ^*, perciò per il teorema di Weierstrass Λ^* è denso in $C[0,1]$. Per il teorema di Müntz allora deve essere [g]

$$\sum_{i=0}^{\infty} \frac{\lambda_i + 1}{(\lambda_i + 1)^2 + 1} = \infty \tag{3.8}$$

Si assuma ora la divergenza della serie. Allora per il teorema di Hahn-Banach e per il teorema di Riesz (ripetendo il ragionamento effettuato all'inizio della dimostrazione del teorema di Müntz in 3.2) Λ non è denso in $L^1[0,1]$ se e solo se esiste $0 \neq h \in L^\infty[0,1]$ (il duale di L^1) che soddisfa

$$\int_0^1 t^{\lambda_i} h(t)\ dt = 0 \qquad i = 0, 1, ...$$

Supponendo che tale h esista (altrimenti il teorema è dimostrato), sia

$$f(z) := \int_0^1 t^z h(t) dt$$

[g] la condizione $\sum 1/\lambda_i = \infty$ diventa $\sum \frac{\lambda_i}{\lambda_i^2+1} = \infty$ quando i λ_i non sono tutti positivi. Si veda a tal proposito [Bor91]

Con un ragionamento analogo al precedente, si ha che

$$g(z) := f\left(\frac{1+z}{1-z} - 1\right)$$

è una funzione olomorfa limitata sul disco unitario, che soddisfa

$$g\left(\frac{\lambda_n}{\lambda_n + 2}\right) = 0 \qquad \text{per } n = 0, 1, 2, ...$$

Poiché la divergenza di 3.8 implica $\sum_{n=1}^{\infty}\left(1 - \left|\frac{\lambda_n}{\lambda_n+2}\right|\right) = \infty$, allora per il lemma sui prodotti di Blaschke $g(z) = 0$ sull'intero disco unitario, ovvero $f(z) = 0$ non appena $\text{Re}(z) > -1$, da cui

$$\int_0^1 t^n h(t)\, dt = 0 \qquad \text{per } n = 0, 1, ...$$

che per il teorema di Lerch (2.1.1) implica $h = 0$, contrariamente all'ipotesi iniziale $h \neq 0$. Ne deduciamo che nelle ipotesi del teorema Λ è denso in $L^1[0,1]$.

□

Con un procedimento analogo si può dimostrare che se $\{\lambda_i\}$ è una successione di numeri reali distinti e maggiori di $-1/p$, $p \in [1, +\infty)$, che soddisfa

$$\sum_{i=0}^{+\infty} \frac{\lambda_i + 1/p}{(\lambda_i + 1/p)^2 + 1} = \infty$$

allora $\Lambda = span\left\{x^{\lambda_0}, x^{\lambda_1}, ...\right\}$ è denso in $L^p[0,1]$. Dimostrare che tale condizione è anche sufficiente per la densità di Λ è ben più complicato; inoltre tali dimostrazioni non aggiungono molto alla comprensione del risultato, dunque non le riportiamo. Esse sono reperibili, ad esempio, nel già citato [Bor95] o in [Alm07].

Bibliografia

[AC94] Francesco Altomare and Michele Campiti. *Korovkin-type Approximation Theory and its Applications.* Walter de Gruyter, 1994.

[Alm07] J.M. Almira. Müntz Type Theorems I. *Surveys in Approximation Theory*, 3:152–194, 2007.

[Apo74] Tom Apostol. *Mathematical Analysis.* Addison-Wesley Publishing Company, 1974.

[Ban32] S. Banach. *Théorie des opérations linéaires.* Hafner, 1932.

[Bau78] Heinz Bauer. Approximation and abstract boundaries. *The American Mathematical Monthly*, 85:632–647, 1978.

[Bor91] Borwein, Peter and Erdélyi, Tamás. The Full Müntz Theorem In $C[0,1]$ and $L^p[0,1]$. *J.London Math. Soc.*, 1991.

[Bor95] Borwein, Peter and Erdélyi, Támas. *Polynomials and polynomial inequalities.* Springer, 1995.

[Car00] N.L. Carothers. *Real analysis.* Cambridge University Press, 2000.

[Kor59] P.P. Korovkin. *Linear Operators and Approximation Theory.* Hindustan Publ. Corp, 1959.

[Pin00] Allan Pinkus. Weierstrass and Approximation Theory. *J. Approx. Theory*, 107:1–66, 2000.

[Pin05] Allan Pinkus. Density in approximation theory. *Surveys in Approximation Theory*, 1:1–45, 2005.

[Rem98] R. Remmert. *Classical topics in complex function theory.* Springer-Verlag, 1998.

[Rie11] F. Riesz. Sur certains systèmes singuliers d'équations intégrales. *Euvres*, 2, 1911.

[Rud66] Walter Rudin. *Real and complex analysis.* McGraw Hill, 1966.

[Sch] Anton R. Schep. Weierstrass' proof of the Weierstrass approximation theorem.

Ringraziamenti

Ringrazio la mia famiglia, Sergio, Laura e Eleonora, per avermi sempre lasciato libero di scegliere la mia strada sostenendomi se necessario con consigli, parole di conforto e fiducia, mai con imposizioni.
Ringrazio Ludovica, per essere rimasta sempre al mio fianco in ogni situazione, sopportandomi e condividendo con me gioie e dolori.
Ringrazio il prof. Troianiello per avermi seguito in quest'ultimo anno con fiducia, pazienza e disponibilità, supportandomi sempre con preziosi consigli.
Ringrazio Matteo, Lorenzo, Livia, Tommaso e tutti i miei compagni di studio e amici, per questi tre anni di condivisione, supporto e aiuto reciproco quotidiani.
Infine ringrazio Riccardo, Alex, Roberto, Raffaele e gli altri miei amici non del Castelnuovo, per essere stati il mio solido collegamento con il mondo "di fuori".