Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Triennale in Matematica

Tesi di Laurea Triennale

Il Teorema di Sárközy-Fürstenberg: una dimostrazione ergodica

Candidato
Alex Cardelli

Relatore
Prof. Mauro Di Nasso

Anno Accademico 2008/2009

Indice

Introduzione

La teoria ergodica di Ramsey nacque nel 1978 quando H. Fürstenberg fornì una dimostrazione (si veda [Für77]) del seguente risultato combinatorio

Teorema (Szemerédi)**.** Ogni $E \subset \mathbb{Z}$ con densità asinotica superiore positiva contiene progressioni aritmetiche di lunghezza arbitraria.

Il teorema era già stato dimostrato ([Sze75]) con metodi combinatori, ma la strategia utilizzata da Fürstenberg fu di natura diversa, utilizzava infatti strumenti di teoria della misura. Introducendo il Principio di Corrispondenza, Fürstenberg mostrò come il teorema seguisse da questo enunciato riguardante gli spazi probabilizzati
Teorema (Fürstenberg)**.** Sia T una trasformazione che conserva la misura su uno spazio misurato $(X, \mathcal{A}, \mu)$ con $\mu(X) < \infty$. Sia $A \in \mathcal{A}$ con $\mu(A) > 0$ e $k \geq 2$, allora esiste $n \in \mathbb{N}$ tale che

$$\mu(A \cap T^{-n}A \cap T^{-2n}A \cap \cdots \cap T^{-kn}A) > 0.$$

Lo scopo principale di questo lavoro è di provare il seguente risultato di natura simile a quello di Szemerédi, dimostrato indipendentemente da Sárközy con metodi combinatori ([Sár78]) e da Fürstenberg utilizzando i nuovi strumenti di teoria ergodica ([Für81]).

Teorema (Sárközy-Fürstenberg)**.** Dato $E \subset \mathbb{Z}$ con densità superiore di Banach positiva e $p(x) \in \mathbb{Z}[x]$ con $p(0) = 0$. Allora $(E-E) \cap \{p(n) : n \in \mathbb{Z}\} \neq \emptyset$.

La densità superiore di Banach è una generalizzazione della densità asintotica superiore. Questo risultato indica un'interessante proprietà degli insiemi differenza. Dato $Y \subset \mathbb{Z}$ non è difficile verificare che l'insieme $Y - Y$ è più "grande" dell'insieme di partenza Y. Questo teorema dice che, sotto un'ipotesi di densità, gli insiemi differenza sono "ben sparsi" nel senso che intersecano tutti gli insiemi polinomiali della forma $\{p(n) : n \in \mathbb{Z}\}$, dove $p(x)$ è un qualunque polinomio con $p(0) = 0$. La dimostrazione che noi daremo di questo teorema (si veda [McC99],[Ber06]) utilizza gli strumenti della teoria ergodica, ma è in parte differente da quella originale di Fürstenberg. Quest'ultima infatti utilizzava il Teorema Spettrale per spazi vettoriali di dimensione infinita, noi invece utilizzeremo un Lemma detto "trucco" di Van der Corput.

La prima sezione di questo lavoro è dedicata al Principio di Corrispondenza di Fürstenberg, uno dei risultati che più hanno contribuito alla nascita della cosiddetta teoria ergodica di Ramsey; questo, nel nostro caso, ci permetterà di "trasferire" l'enunciato del Teorema di Sárközy-Fürstenberg nel linguaggio degli spazi probabilizzati. Nella seconda sezione daremo la dimostrazione del teorema: questa consta di due parti. Nella prima verificheremo che è equivalente ad un risultato analogo di teoria ergodica, nella seconda dimostreremo quest'ultimo.

1 Principio di Corrispondenza di Fürstenberg

In questa sezione dimostreremo uno strumento fondamentale della teoria ergodica di Ramsey. Questo ci permette di "tradurre" un problema di densità sui numeri interi in uno riguardante gli spazi probabilizzati. L'idea dietro questo risutato è di sfruttare l'analogia tra i concetti di densità asintotica di insiemi di interi e di misura in uno spazio probabilizzato, costruendo uno spazio che esprime al meglio questa analogia.
Prima di tutto dobbiamo introdurre alcuni concetti, diamo quindi alcune definizioni:

Definizione 1.1 (Sistema che conserva la misura)**.** $(X, \mathcal{A}, \mu, T)$ è detto *sistema che conserva la misura* se X è un insieme, $\mathcal{A}$ è una σ-algebra di sottoinsiemi di X, μ è una misura di probabilità su $\mathcal{A}$ e $T : X \to X$ è una applicazione invertibile μ-invariante, cioè tale che $\mu(T^{-1}(A)) = \mu(A)$.

Definizione 1.2 (Densità asintotica superiore)**.** Dato $E \subset \mathbb{Z}$ il numero

$$\bar{d}(E) = \limsup_{n\to\infty} \frac{|E \cap \{-n, -n+1, ..., n-1, n\}|}{2n+1}$$

è detto *densità asintotica superiore* di E.

Generalizzando il concetto di densità asintotica superiore otteniamo un'altra nozione di grandezza per i sottoinsiemi di $\mathbb{Z}$, più debole di quella precedentemente definita.

Definizione 1.3 (Densità superiore di Banach)**.** Dato $E \subset \mathbb{Z}$ il numero

$$d^*(E) = \lim_{n\to\infty} \left(\max_{k\in\mathbb{Z}} \frac{|E \cap \{k, ..., k+n-1\}|}{n} \right)$$

è detto *densità superiore di Banach* di E.

Verificando che la successione $\{x_n\}_{n\in\mathbb{N}}$ con

$$x_n = \max_{k\in\mathbb{N}} |E \cap \{k, ..., k+n-1\}|$$

è subadditiva (cioè $x_{n+m} \leq x_n + x_m \quad \forall n, m \in \mathbb{N}$) abbiamo che

$$\exists \lim_{n\to\infty} \frac{x_n}{n} = \inf_{n\in\mathbb{N}} \left\{ \frac{x_n}{n} \right\}$$

quindi la definizione è buona. Inoltre vale

$$d^*(E) = \limsup_{N-M\to\infty} \frac{|E \cap \{M, M+1, ..., N-1\}|}{N-M}.$$

Di seguito useremo quest'ultima come definizione di densità superiore di Banach, perchè più comoda. Osserviamo che per ogni $E \subset \mathbb{Z}$ si ha $d^*(E) \geq \bar{d}(E)$ in quanto la densità asintotica superiore è un caso particolare di quella superiore di Banach. Possiamo adesso dare il seguente enunciato che dimostreremo in seguito.

Teorema (Principio di Corrispondenza di Fürstenberg)**.** Sia $E \subset \mathbb{Z}$ tale che $d^*(E) > 0$, allora esiste un sistema che conserva la misura $(X, \mathcal{A}, \mu, T)$ ed esiste $A \in \mathcal{A}$ con $\mu(A) = d^*(E)$ tale che $\forall F \in \text{Fin}(\mathbb{Z})$ (sottoinsiemi finiti di $\mathbb{Z}$) vale:

$$d^*(\bigcap_{n \in F}(E - n)) \geq \mu(\bigcap_{n \in F} T^{-n}A).$$

1.1 Costruzione di un sistema che conserva la misura

Consideriamo $X = \{0,1\}^{\mathbb{Z}}$ l'insieme delle funzioni $x : \mathbb{Z} \to \{0,1\}$. Possiamo osservare che una base della topologia prodotto su X è data da insiemi che chiameremo *cilindrici*, che sono della forma

$$C = \{x \in X \,:\, x(n_i) = \epsilon_i,\, 1 \leq i \leq t\}$$

con $n_i \in \mathbb{Z}$ e $\epsilon_i \in \{0,1\} \quad \forall i = 1, ..., t$. Chiamiamo $\mathcal{C}$ la famiglia delle unioni finite di insiemi cilindrici. Notiamo che $\mathcal{C}$ è un algebra di sottoinsiemi di X (chiusa rispetto a complementazione, unioni e intersezioni finite) e che tutti i suoi elementi sono insiemi aperti e chiusi.

Gli insiemi cilindrici godono inoltre di questa interessante proprietà: presi $C_1, C_2 \in \mathcal{C}$

$$\begin{aligned} C_1 &= \{x \in X \,:\, x(n_i) = \epsilon_i,\, 1 \leq i \leq t\} \\ C_2 &= \{x \in X \,:\, x(m_i) = \eta_i,\, 1 \leq i \leq s\} \end{aligned}$$

allora

$$C_1 \cup C_2 = \{x \in X \,:\, x(q_i) = \xi_i,\, 1 \leq i \leq r\} \in \mathcal{C},$$

dove $r \leq \min\{s,t\}$ e per ogni q_i, esistono n_j e m_k tali che $q_i = n_j = m_k$ e $\xi_i = \epsilon_j = \eta_k$; cioè facendo l'unione di due insiemi cilindrici otteniamo l'insieme delle funzioni (anch'esso cilindrico) che fissano i valori fissati dalle funzioni di entrambi gli insiemi.

Possiamo inoltre osservare che X dotato della topologia prodotto è anche uno spazio metrico con distanza

$$d(x,y) = 2^{-n} \qquad \text{dove} \qquad n = \min\{|m| \,:\, x(m) \neq y(m)\}.$$

Fissiamo adesso alcune notazioni. Dato $E \subset \mathbb{Z}$ denoteremo con χ_E la sua funzione caratteristica, chiameremo invece 1_C la funzione caratteristica di un $C \subset X$. Per semplicità indicheremo con $\{I_t\}_{t\in\mathbb{N}}$ le successioni di intervalli di $\mathbb{Z}$ di lunghezza crescente. Siano inoltre:

- T la funzione *shift* definita su X in questo modo:
  $Tx(n) = x(n+1) \quad \forall n \in \mathbb{Z}$;
- A l'insieme $\{x \in X \,:\, x(0) = 1\} \in \mathcal{C}$.

A questo punto, preso $E \subset \mathbb{Z}$ con $d^*(E) > 0$, procederemo definendo una funzione $p : \mathcal{C} \to [0,1]$ che sia (finitamente) additiva. Successivamente estenderemo $\mathcal{C}$ ad una σ-algebra (chiusa per unioni e intersezioni numerabili) e la funzione p a una misura σ-additiva definita su quest'ultima.
Per definizione esiste una successione di intervalli $\{I_t\}_{t\in\mathbb{N}}$ di lunghezza crescente tale che

$$\lim_{t\to\infty} \frac{|E \cap I_t|}{|I_t|} = d^*(E) > 0.$$

Possiamo adesso osservare che

$$n \in E \;\Leftrightarrow\; T^n(\chi_E) \in A$$

e quindi

$$d^*(E) = \lim_{t\to\infty} \frac{1}{|I_t|} \sum_{n\in I_t} \chi_E(n) = \lim_{t\to\infty} \frac{1}{|I_t|} \sum_{n\in I_t} 1_A(T^n\chi_E).$$

Definiamo adesso la p utilizzando una opportuna $\{J_s\}_{s\in\mathbb{N}}$ sottosuccessione di $\{I_t\}_{t\in\mathbb{N}}$, in questo modo: dato $C \in \mathcal{C}$

$$p(C) = \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_C(T^n\chi_E).$$

Verifichiamo che una tale sottosuccessione esiste:

Osservazione 1.1. *Sia $E \subset \mathbb{Z}$ con $d^*(E) > 0$. Per ogni $\{I_t\}_{t\in\mathbb{N}}$ successione di intervalli di lunghezza crescente, esiste $\{I_{t_s}\}_{s\in\mathbb{N}}$ sottosuccessione tale che il limite:*

$$\lim_{s\to\infty} \frac{1}{|I_{t_s}|} \sum_{n\in I_{t_s}} \chi_C(T^n\chi_E)$$

esiste $\quad \forall C \in \mathcal{C}$.

Dimostrazione. La cardinalità di $\mathcal{C}$ è numerabile, quindi $\mathcal{C} = \{C_1, ..., C_n, ...\}$. Possiamo quindi verificare la tesi con un metodo detto diagonale. Dato C_1 possiamo costruire una sottosuccessione $\{I_{t_s^{(1)}}\}_{s\in\mathbb{N}}$ di $\{I_t\}$ tale che esiste

$$\lim_{s\to\infty} \frac{1}{|I_{t_s^{(1)}}|} \sum_{n\in I_{t_s^{(1)}}} 1_{C_1}(T^n\chi_E).$$

A questo punto possiamo trovare $\{I_{t_s^{(2)}}\}_{s\in\mathbb{N}}$ sottosuccessione di $\{I_{t_s^{(1)}}\}$ tale che esiste

$$\lim_{s\to\infty} \frac{1}{|I_{t_s^{(2)}}|} \sum_{n\in I_{t_s^{(2)}}} 1_{C_2}(T^n\chi_E).$$

Procedendo induttivamente costruiamo una sequenza di successioni di intervalli di lunghezza crescente. La successione $\{I_{t_s^{(s)}}\}_{s\in\mathbb{N}}$ (che ha all's-esimo termine l's-esimo elemento della s-esima sottosuccessione trovata), verifica la tesi. $\square$

Abbiamo quindi che a meno di passare a sottosuccessioni, la p è ben definita. Diamo altre due definizioni:

Definizione 1.4 (Premisura)**.** Sia X uno insieme e $\mathcal{A}$ un algebra di sottoinsiemi su X. Una funzione $p : \mathcal{A} \to [0, \infty]$ è detta *premisura* su X se:

1. $p(\emptyset) = 0$.

2. Per ogni successione $(A_i)_{i=1}^{\infty}$ di elementi di $\mathcal{A}$ a due a due disgiunti, tali che $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, si ha $p(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} p(A_i)$.

Definizione 1.5 (Misura esterna)**.** Sia X un insieme, una funzione $\mu^* : \mathcal{P}(X) \to [0, \infty]$ sarà detta *misura esterna* su X se verifica le seguenti proprietà:

1. $\mu^*(\emptyset) = 0$

2. $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$

3. $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ con $A_i \subset X \quad \forall i$.

Osseviamo che p è additiva, infatti presi $C_1, ..., C_k \in \mathcal{C}$ a due a due disgiunti

$$1_{\bigcup_{i=1}^{k} C_i}(T^n\chi_E) = \sum_{i=1}^{k}(1_{C_i}(T^n\chi_E))$$

quindi

$$\begin{aligned}
\sum_{i=1}^{k} p(C_i) &= \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} \sum_{i=1}^{k} (1_{C_i}(T^n \chi_E)) \\
&= \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_{\bigcup_{i=1}^{k} C_i}(T^n \chi_E) \\
&= p(\bigcup_{i=1}^{k} C_i)
\end{aligned}$$

Inoltre verifica le ipotesi di premisura: $p(\emptyset) = 0$; inoltre se $(C_i)_{i=1}^{\infty} \subset \mathcal{C}$ é una sequenza di insiemi cilindrici a due a due disgiunti, per la proprietà precedentemente espressa $C = \bigcup_{i=1}^{\infty} C_i$ è ancora un insieme cilindrico ed é unione di un numero finito di C_i. Dall'additività segue che p verifica la seconda ipotesi di premisura. La precedente proprietà può essere stabilita anche osservando che lo spazio X è compatto.
Vorremmo adesso estendere la premisura trovata ad una vera e propria misura sullo spazio X, per far questo utilizzeremo un noto risultato di teoria della misura (di cui non diamo dimostrazione)

Teorema 1.1 (Carathéodory)**.** Sia X un insieme, $\mathcal{A}$ un algebra di sottoinsiemi su X e p una premisura su $\mathcal{A}$ tale che $p(X) = 1$. Per ogni $B \subset X$ sia

$$\mu^*(B) = \inf\Big\{ \sum_{i=1}^{\infty} p(A_i) \,|\, (A_i)_{i=1}^{\infty} \subset \mathcal{A}, B \subset \bigcup_{i=1}^{\infty} A_i \Big\}.$$

Sia adesso $\mathcal{B} = \{B \subset X \,|\, \mu^*(B) + \mu^*(B^c) = 1\}$. Allora μ^* è una misura esterna su X tale che $\mu^*_{|\mathcal{A}} = p$, inoltre $\mathcal{B}$ è σ-algebra su X contenente $\mathcal{A}$ e la restrizione di μ^* a $\mathcal{B}$ è una misura di probabilità.

Sia quindi μ la misura su X costruita estendendo p. Osserviamo che la definizione di p (quindi anche di μ) dipende dall'insieme E. Abbiamo che per come è stata definita

$$\mu(A) = d^*(E).$$

Inoltre se p è T-invariante anche μ lo è: infatti per definizione di μ, se $B \in \mathcal{B}$ allora $\forall\, \epsilon > 0\ \exists\, (C_i)_{i=0}^{\infty} \subset \mathcal{C}$ tale che $B \subset \bigcup_{i=1}^{\infty} C_i$ e vale

$$\mu(B) \geq \sum_{i=1}^{\infty} p(C_i) - \epsilon = \sum_{i=1}^{\infty} p(T^{-1}C_i) - \epsilon \geq \mu(T^{-1}B) - \epsilon.$$

Per l'arbitrarietà di ϵ otteniamo una disuguaglianza. L'altra si ottiene allo stesso modo partendo da $\mu(T^{-1}B)$. Quindi $\mu(B) = \mu(T^{-1}B)$. Dobbiamo verificare che p è effettivamente T-invariante. Preso $C \in \mathcal{C}$ abbiamo che

$$p(C) = \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_C(T^n \chi_E)$$

e d'altra parte

$$p(T(C)) = \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_{T(C)}(T^n \chi_E).$$

Abbiamo inoltre che per ogni n

$$T^n(\chi_E) \in T(C) \ \Leftrightarrow\ T^{n-1}(\chi_E) \in C$$

perciò

$$p(T(C)) = \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_C(T^{n-1}\chi_E) = \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n+1\in J_s} 1_C(T^n \chi_E)$$

che per la dimensione crescente degli intervalli è uguale a $p(C)$.

1.2 Dimostrazione del principio di corrispondenza

Dopo aver costruito un sistema che conserva la misura, la verifica del principio è semplice.

Teorema 1.2 (Principio di Corrispondenza di Fürstenberg)**.** Sia $E \subset \mathbb{Z}$ tale che $d^*(E) > 0$, allora esiste un sistema che conserva la misura $(X, \mathcal{A}, \mu, T)$ ed esiste $A \in \mathcal{A}$ con $\mu(A) = d^*(E)$ tale che $\forall F \in \mathrm{Fin}(\mathbb{Z})$ (sottoinsiemi finiti di $\mathbb{Z}$) vale:

$$d^*(\bigcap_{n\in F}(E-n)) \geq \mu(\bigcap_{n\in F} T^{-n}A).$$

Dimostrazione. Siano $(X, \mathcal{B}, \mu, T)$ e $A \in \mathcal{B}$ come costruiti precedentemente, allora $d^*(E) = \mu(A)$. Inoltre preso $F \in \mathrm{Fin}(\mathbb{Z})$ con $F = \{n_1, ..., n_k\}$ abbiamo che

$$\begin{aligned} n \in \bigcap_{n_i\in F}(E-n_i) \quad &\Leftrightarrow \quad n+n_i \in E \quad \forall n_i \in F \\ &\Leftrightarrow \quad T^n\chi_E \in \bigcap_{n_i \in F} T^{-n_i}A \end{aligned}$$

quindi

$$\begin{aligned}
\mu(T^{-n_1}A \cap ... \cap T^{-n_k}A) &= \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} 1_{(T^{-n_1}A\cap...\cap T^{-n_k}A)}(T^n\chi_E) \\
&= \lim_{s\to\infty} \frac{1}{|J_s|} \sum_{n\in J_s} \chi_{(E-n_1)\cap...\cap(E-n_k)}(n) \\
&\leq d^*((E-n_1)\cap ... \cap (E-n_k))
\end{aligned}$$

dove nell'ultima disuguaglianza abbiamo applicato la definizione di densità superiore di Banach. □

2 Teorema di Sárközy-Fürstenberg

In questa sezione andremo a dimostrare il seguente Teorema:

Teorema 2.1 (Teorema di Sárközy-Fürstenberg)**.** Sia $E \subset \mathbb{Z}$ con $d^*(E) > 0$ e $p(x) \in \mathbb{Z}[x]$ con $p(0) = 0$. Allora $(E - E) \cap \{p(n) : n \in \mathbb{Z}\} \neq \emptyset$.

2.1 Teorema di Ricorrenza di Poincaré: un risultato preliminare

Prima di passare alla versione ergodica del Teorema di Sárközy-Fürstenberg enunciamo un famoso risultato di Poincaré in una delle sue forme. Questo ci aiuterà a comprendere una definizione che sarà data successivamente, oltre a indicarci un metodo classico per affrontare delle questioni simili al teorema che vogliamo trattare in questo lavoro.

Teorema 2.2 (Teorema di Ricorrenza di Poincaré)**.** Sia $(X, \mathcal{A}, \mu, T)$ un sistema che conserva la misura e sia $A \in \mathcal{A}$ tale che $\mu(A) = a > 0$. Dato $Y \subset \mathbb{Z}$ infinito, allora $\exists\, n \in Y - Y$ tale che $\mu(T^{-n}A \cap A) > 0$

Dimostrazione. Procediamo per assurdo. Negando la tesi abbiamo che per ogni $n \in Y - Y$ vale $\mu(T^{-n}A \cap A) = 0$. Da questo segue che per ogni $i \neq j \in Y$ $\mu(T^{-i}A \cap T^{-j}A) = 0$. Sia $N = \lceil \frac{1}{a} \rceil + 1$. Prendendo $i_1, ..., i_N$ elementi distinti in Y, sia allora $U = \bigcup_{j=1}^{N} T^{-i_j}A$. Abbiamo che $\mu(U) = \mu(\bigcup_{j=1}^{N} T^{-i_j}A) = \sum_{j=1}^{N} \mu(T^{-j}A) = N \cdot a > 1$. Assurdo. □

Osserviamo che quello che abbiamo dimostrato è un risultato più forte di quello espresso nell'enunciato; infatti nella dimostrazione abbiamo utilizzato solo il fatto che la misura μ fosse additiva e non σ-additiva.

2.2 Versione ergodica

In questa parte enunceremo una versione equivalente del Teorema di Sárközy-Fürstenberg, espressa nel linguaggio degli spazi probabilizzati. Verificheremo poi, utilizzando il Principio di Corrispondenza di Fürstenberg, che i due risultati sono effettivamente equivalenti. Introduciamo adesso il concetto di sequenza di Poincaré.

Definizione 2.1 (Sequenza di Poincaré)**.** $R \subset \mathbb{Z}$ è detta *sequenza di Poincaré* se $\forall\, (X, \mathcal{A}, \mu, T)$ sistema che conserva la misura e $\forall A \in \mathcal{A}$ con $\mu(A) > 0$ $\exists n \in R$ tale che

$$\mu(A \cap T^{-n}A) > 0.$$

Dal Teorema 2.2 segue che ogni $Y - Y$ con Y infinito è una sequenza di Poincaré. Possiamo adesso enunciare il seguente teorema:

Teorema 2.3 (Sárközy-Fürstenberg versione ergodica)**.** Se $p(n) \in \mathbb{Z}$ con $p(0) = 0$ allora $\{p(n) : n \in \mathbb{Z}\}$ è una sequenza di Poincaré.

Osserviamo preliminarmente che il precedente enunciato è un'estensione del Teorema di ricorrenza di Poincaré a insiemi detti "polinomiali".
Per dimostrare l'equivalenza utilizzeremo il Teorema 1.2 e la Proposizione seguente di cui daremo dimostrazione in seguito:

Proposizione 2.1. Sia $(X, \mathcal{A}, \mu, T)$ sistema che conserva la misura e $\{A_n\}_{n\in\mathbb{Z}}$ successione di elementi di $\mathcal{A}$ tali che $\mu(A_n) = a > 0 \quad \forall n \in \mathbb{Z}$, allora esiste $\Gamma = \{n_i\}_{i\in\mathbb{Z}}$ sottoinsieme di indici tale che $\bar{d}(\Gamma) \geq a$ e per ogni $F = \{n_{i_1}, ..., n_{i_k}\}$ sottoinsieme finito di indici in Γ vale

$$\mu(\bigcap_{n_{i_j}\in F} A_{n_{i_j}}) > 0.$$

Abbiamo quindi che

Proposizione 2.2. I teoremi 2.1 e 2.3 sono equivalenti.

Dimostrazione. Teorema 2.3 $\Rightarrow$ Teorema 2.1. Dato $E \subset \mathbb{Z}$ e $p(n)$ come nelle ipotesi, applicando il Principio di Corrispondenza di Fürstenberg abbiamo che $\exists\, (X, \mathcal{A}, \mu, T)$ e A con $\mu(A) = d^*(E) > 0$. Allora utilizzando il Teorema 2.3 abbiamo che $\exists\, n \in \mathbb{Z}$ tale che $\mu(A \cap T^{-p(n)}A) > 0$. Ponendo $F = \{0, p(n)\}$ otteniamo che

$$d^*(E \cap (E - p(n))) \geq \mu(A \cap T^{-p(n)}A) > 0$$

quindi $\exists\, x, y \in E$ tali che $x = y - p(n)$ quindi $x - y = p(n)$.

Teorema 2.1 $\Rightarrow$ Teorema 2.3. Dato $p(n)$ come nelle ipotesi abbiamo che, preso $(X, \mathcal{A}, \mu, T)$ un sistema che conserva la misura, l'insieme $\{T^n A \mid n \in \mathbb{Z}\}$ verifica le ipotesi della Proposizione 2.1, perciò esiste sottoinsieme di indici $\Gamma = \{n_i\}_{i\in\mathbb{Z}}$ con $\bar{d}(\Gamma) > 0$ (quindi anche $d^*(\Gamma) > 0$) tale che per ogni F sottoinsieme finito di Γ si ha $\mu(\bigcap_{n_i\in F} A_{n_i}) > 0$. Ponendo $E = \Gamma$, abbiamo che per il Teorema 2.1, $\exists\, n_i, n_j \in \Gamma$ tali che $n_i - n_j = p(n)$ per qualche $n \in \mathbb{Z}$; quindi

$$\mu(T^{-p(n)}A \cap A) = \mu(A \cap T^{p(n)}A) = \mu(A \cap T^{n_i - n_j}A) = \mu(T^{n_j}A \cap T^{n_i}A) > 0.$$

□

Anche qua abbiamo dimostrato più di quello che volevamo, infatti partendo dal Teorema 2.3 abbiamo verificato non solo che $(E-E)\cap\{p(n) : n\in\mathbb{Z}\}\neq\emptyset$, ma anche che per questi n si ha $d^*(E\cap(E-p(n)))>0$. Diamo adesso una dimostrazione della Proposizione 2.1

Dimostrazione Proposizione 2.1. In questa dimostrazione adotteremo la seguente notazione: dato F sottoinsieme finito di $\mathbb{Z}$ chiameremo $A_F=\bigcap_{n\in F}A_n$. Innanzitutto verificheremo che esiste $B\in\mathcal{A}$ di misura nulla tale che per ogni $F\subset\mathbb{Z}$ finito vale

$$(X\setminus B)\cap A_F\neq\emptyset\ \Rightarrow\mu(A_F)>0.$$

Consideriamo $\mathcal{F}=\{F: F \text{ sottoinsieme finito di } \mathbb{Z} \text{ tale che } \mu(A_F)=0\}$ e $B=\bigcup_{F\in\mathcal{F}}A_F$. Poichè $\mathcal{F}\subset\mathrm{Fin}(\mathbb{Z})$ e $|\mathrm{Fin}(\mathbb{Z})|=\aleph_0$, allora $\mu(B)\leq\sum_{F\in\mathcal{F}}\mu(A_F)=0$. Abbiamo poi che, dato A_F, se $(X\setminus B)\cap A_F\neq\emptyset$, A_F non è sottoinsieme di B e quindi $\mu(A_F)>0$. Possiamo quindi supporre, a meno di togliere un insieme di misura nulla, che se $A_F\neq\emptyset$ allora $\mu(A_F)>0$.
Sia

$$f_n(x)=\frac{1}{2n+1}\sum_{|k|\leq n}1_{A_k}(x)$$

abbiamo che $f_n: X\to[0,1]$ e che $\int f_n d\mu=\frac{1}{2n+1}\sum_{|k|\leq n}\int 1_{A_k}d\mu=a$. Poniamo $f(x)=\limsup_{n\to\infty}f_n(x)$, allora per il Lemma di Fatou[1] vale

$$\int f d\mu=\int\limsup_{n\to\infty}f_n d\mu\geq\limsup_{n\to\infty}\int f_n d\mu=a.$$

Adesso $\mu(X)=1$ e quindi esiste $x_0\in X$ tale che $f(x_0)=\limsup f_n(x_0)\geq a$. Esiste quindi una sottosuccessione $\{f_{n_i}(x_0)\}_{i\in\mathbb{Z}}$ di $\{f_n(x_0)\}_{n\in\mathbb{Z}}$ tale che

$$\lim_{i\to\infty}f_{n_i}(x_0)=f(x_0)\geq a. \tag{1}$$

Prendiamo quindi $\Gamma=\{k\in\mathbb{Z}\,|\,x_0\in A_k\}$. Allora abbiamo dalla (1) che

$$\bar{d}(\Gamma)=\limsup_{n\to\infty}\frac{|\Gamma\cap[-n,n]|}{2n+1}=f(x_0)\geq a.$$

Inoltre, per definizione di Γ abbiamo che $\forall F\subset\Gamma$ finito, A_F è non vuoto. Perciò possiamo dire senza perdita di generalità che $\mu(A_F)>0$. $\square$

Osserviamo che per la dimostrazione dell'equivalenza sarebbe sufficiente che $d^*(\Gamma)>0$, il risultato dimostrato è leggermente più forte.

[1]Ricordiamo che il Lemma di Fatou afferma che, data $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni integrabili su uno spazio misurato $(X,\mathcal{A},\mu)$, allora $\int\limsup_{n\to\infty}f_n d\mu\geq\limsup_{n\to\infty}\int f_n d\mu$

2.3 Richiami su spazi di Hilbert e teoria ergodica

Per dimostrare la versione ergodica del Teorema di Sárközy-Fürstenberg abbiamo prima bisogno di alcuni risultati preliminari riguardanti gli spazi di Hilbert e la teoria ergodica. Ricordiamo il concetto di spazio di Hilbert:

Definizione 2.2 (Spazio di Hilbert reale)**.** Sia $\mathcal{H}$ spazio vettoriale su $\mathbb{R}$ dotato di prodotto scalare $\langle\cdot,\cdot\rangle$. Allora $\mathcal{H}$ è detto *spazio di Hilbert* se è completo rispetto alla norma indotta da $\langle\cdot,\cdot\rangle$ (cioè $\|x\| = \sqrt{\langle x, x\rangle}\ \forall\, x \in \mathcal{H}$).

Innanzitutto abbiamo che

Teorema 2.4. Sia $(X, \mathcal{A}, \mu)$ uno spazio probabilizzato allora

$$L^2(X, \mathcal{A}, \mu) = \{f : X \to (-\infty, +\infty)\ :\ \int f^2 d\mu < \infty\}$$

è uno spazio di Hilbert, dove $\langle f, g\rangle = \int f \cdot g\ d\mu$.

Verifichiamo adesso alcune proprietà che ci serviranno poi nella dimostrazione. Cominciamo con un risultato classico riguardante gli spazi di Hilbert di cui non daremo dimostrazione.

Teorema 2.5. Sia $\mathcal{H}$ uno spazio di Hilbert e sia $\mathcal{M}$ un suo sottospazio lineare chiuso. Preso $z \in \mathcal{H}$, questo può essere espresso in modo unico come somma $z = x+y$ con $x \in \mathcal{M}$ e $y \in \mathcal{M}^\perp$ (dove $\mathcal{M}^\perp = \{z \in \mathcal{H} : \langle z, x\rangle = 0 \quad \forall x \in \mathcal{M}\}$ è detto spazio ortogonale a $\mathcal{M}$). Dunque $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$.

Prima di vedere qualche altro risultato abbiamo bisogno di un'altra definizione:

Definizione 2.3 (Operatore unitario su spazio di Hilbert)**.** Sia $\mathcal{H}$ uno spazio di Hilbert, una isometria invertibile U definita su $\mathcal{H}$ (cioè tale che $\|Ux\| = \|x\| \quad \forall\, x \in \mathcal{H}$), sarà detta *operatore unitario* su $\mathcal{H}$.

Nella nostra dimostrazione, dato $(X, \mathcal{A}, \mu, T)$ un sistema che conserva la misura, utilizzeremo l'operatore unitario indotto dall'applicazione T.

Teorema 2.6. Sia $\mathcal{H}$ uno spazio di Hilbert e sia U un operatore unitario su $\mathcal{H}$. Siano $\mathcal{M}_1 = \{x \in \mathcal{H} : Ux = x\}$ e $\mathcal{M}_2 = \overline{\{y - Uy : y \in \mathcal{H}\}}$, allora $\mathcal{H} = \mathcal{M}_1 \oplus \mathcal{M}_2$.

Dimostrazione. Basterà verificare che $\mathcal{M}_2 = \mathcal{M}_1^\perp$, la tesi seguirà dal Teorema 2.5. Presi $y \in \mathcal{H}$ e $x \in \mathcal{M}_1$ allora $\langle y - Uy, x\rangle = \langle y, x\rangle - \langle Uy, x\rangle = \langle y, x\rangle - \langle Uy, Ux\rangle = 0$ quindi $\{y - Uy : y \in \mathcal{H}\} \subset \mathcal{M}_1^\perp$. Dimostriamo adesso che

$\mathcal{M}_1^\perp$ è chiuso. Data $\{x_n\}_{n=0}^\infty$ una successione di elementi di $\mathcal{M}_1^\perp$, cioè tali che $\langle x_n, y\rangle = 0$ per ogni $n \in \mathbb{N}$ e $y \in \mathcal{M}_1$, sia, se esiste, $x = \lim_{n\to\infty} x_n$. Fissato $y \in \mathcal{M}_1$, prendiamo $\epsilon > 0$ piccolo a piacere e $\bar{n}$ tale che $\|x - x_{\bar{n}}\| < \frac{\epsilon}{\|y\|}$; abbiamo che

$$|\langle x, y\rangle| \leq |\langle x - x_{\bar{n}}, y\rangle| + |\langle x_{\bar{n}}, y\rangle| \leq \|x - x_{\bar{n}}\| \cdot \|y\| < \epsilon.$$

Quindi $\langle x, y\rangle = 0$ per ogni $y \in \mathcal{M}_1$, perciò $\mathcal{M}_2 \subset \mathcal{M}_1^\perp$.
D'altra parte sia $z \in \mathcal{M}_1^\perp$ e $w = z - P_{\mathcal{M}_2} z$ dove $P_{\mathcal{M}_2}$ è la proiezione ortogonale sul sottospazio $\mathcal{M}_2$. Abbiamo che $w \in \mathcal{M}_2^\perp$ e in particolare $\forall\, y \in \mathcal{H}$ si ha che $0 = \langle w, y - Uy\rangle$. Perciò per ogni y vale $\langle w, y\rangle = \langle w, Uy\rangle = \langle U^{-1}w, y\rangle$ che implica $w = U^{-1}w$ cioè $w = Uw$ e quindi $w \in \mathcal{M}_1$. Ma abbiamo già verificato che $\mathcal{M}_2 \subset \mathcal{M}_1^\perp$ e perciò che $w = z - P_{\mathcal{M}_2} z \in \mathcal{M}_1^\perp$. Quindi $w = 0$ e quindi $z \in \mathcal{M}_2$. □

Teorema 2.7 (teorema della media ergodica unitaria)**.** Sia $\mathcal{H}$ uno spazio di Hilbert e U un operatore unitario su $\mathcal{H}$, allora, con le notazioni del Teorema 2.6, per ogni $x \in \mathcal{H}$ vale

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U^n x = P_{\mathcal{M}_1} x.$$

Dimostrazione. Utilizzando il Teorema 2.6 possiamo scrivere $x \in \mathcal{H}$ come $x = x_1 + x_2$ con $x_1 = P_{\mathcal{M}_1} x \in \mathcal{M}_1$ e $x_2 \in \mathcal{M}_2$, possiamo quindi trattare separatamente i limiti delle due componenti di x. Abbiamo innanzitutto che

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U^n x_1 = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_1 = x_1 = P_{\mathcal{M}_1} x. \tag{2}$$

Sia poi $W = \{w \in \mathcal{H} \,:\, \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U^n w = 0\}$. Vogliamo dimostrare che $\mathcal{M}_2 \subset W$. Verifichiamo innanzitutto che W è chiuso. Presa $\{w_n\}_{n\in\mathbb{N}}$ una successione di elementi di W, dunque

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} U^k w_n = 0 \quad \forall\, n \in \mathbb{N},$$

sia $w = \lim_{n\to\infty} w_n$. Fissato $\epsilon > 0$, esiste $\bar{n}$ tale che $\|w - w_{\bar{n}}\| < \epsilon$. Abbiamo quindi che

$$\begin{aligned}
\lim_{N\to\infty}\Big\|\frac{1}{N}\sum_{k=1}^{N}U^k w\Big\| &= \lim_{N\to\infty}\frac{1}{N}\Big\|\sum_{k=1}^{N}(U^k w - U^k w_{\bar n} + U^k w_{\bar n})\Big\| \\
&\leq \lim_{N\to\infty}\frac{1}{N}\Big\|\sum_{k=1}^{N}U^k(w-w_{\bar n})\Big\| + \frac{1}{N}\Big\|\sum_{k=1}^{N}U^k w_{\bar n}\Big\| \\
&\leq \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\|U^k(w-w_{\bar n})\| \\
&= \frac{1}{N}\sum_{k=1}^{N}\|w-w_{\bar n}\| < \epsilon.
\end{aligned}$$

Verifichiamo ora che $\{y - Uy : y \in \mathcal{H}\} \subset W$: per ogni $y \in \mathcal{H}$ abbiamo

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}U^n(y-Uy) &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}U^n y - \frac{1}{N}\sum_{n=1}^{N}U^{n+1}y \\
&= \lim_{N\to\infty}\frac{1}{N}\Big(Uy + U^{N+1}y + \sum_{n=2}^{N}U^n y - \sum_{n=1}^{N-1}U^{n+1}y\Big) \\
&= \lim_{N\to\infty}\frac{1}{N}(Uy - U^{N+1}y) = 0.
\end{aligned}$$

Infatti $\frac{1}{N}\|Uy - U^{N+1}y\| = \frac{1}{N}\|y - U^N y\| \leq \frac{2\|y\|}{N}$. Quindi $\mathcal{M}_2 \subset W$ e perciò

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}U^n x_2 = 0. \tag{3}$$

Usando (2) e (3) otteniamo la tesi. □

Prima di enunciare il prossimo risultato che seguirà banalmente dal Teorema precedente dobbiamo però osservare che se $(X, \mathcal{A}, \mu, T)$ è un sistema che conserva la misura, allora l'applicazione μ-invariante T induce naturalmente un operatore unitario (che chiameremo U) sullo spazio $L^2(X, \mathcal{A}, \mu)$ definito da $Uf(x) = f(Tx)$.

Teorema 2.8 (teorema della media ergodica)**.** *Sia $(X, \mathcal{A}, \mu, T)$ un sistema che conserva la misura, U un operatore unitario indotto da T e sia P la proiezione sullo spazio delle funzioni U-invarianti. Allora*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}U^n f = Pf$$

per ogni $f \in L^2(X, \mathcal{A}, \mu)$.

2.4 Dimostrazione del Teorema di Sárközy-Fürstenberg

Per affrontare la vera e propria dimostrazione abbiamo innanzitutto bisogno del seguente Lemma, detto "trucco" di Van der Corput.

Lemma 2.1 (Van der Corput trick)**.** Sia $\mathcal{H}$ spazio di Hilbert reale e $\{x_n\}_{n\in\mathbb{Z}}$ successione limitata di vettori. Se $\forall h \in \mathbb{Z}$ con $h \neq 0$ vale

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_n, x_{n+h} \rangle = 0$$

allora

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\| = 0$$

Dimostrazione. Fissiamo $\epsilon > 0$ e sia $L > \|x_n\| \quad \forall n \in \mathbb{Z}$, prendiamo poi H tale che $\frac{L^2}{H} < \epsilon$ e definiamo

$$\Psi_N = \frac{1}{N} \sum_{n=1}^{N} \Big(\frac{1}{H} \sum_{h=1}^{H} x_{n+h} \Big),$$

allora valgono le due seguenti osservazioni:

Osservazione 2.1.

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n - \Psi_N \right\| = 0$$

Dimostrazione.

$$\begin{aligned}
\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} x_n - \Psi_N \right\| &= \lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \Big(x_n - \frac{1}{H} \sum_{h=1}^{H} x_{n+h} \Big) \right\| \\
&= \lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \Big(\frac{1}{H} \sum_{h=1}^{H} (x_n - x_{n+h}) \Big) \right\| \\
&= \lim_{N\to\infty} \frac{1}{NH} \left\| \sum_{n=1}^{N} \Big(H \cdot x_n - \sum_{h=1}^{H} x_{n+h} \Big) \right\| \\
&= \lim_{N\to\infty} \frac{1}{NH} \left\| H \sum_{n=1}^{N} x_n - \sum_{n=1}^{N} \sum_{h=1}^{H} x_{n+h} \right\| \\
&= \lim_{N\to\infty} \frac{1}{NH} \left\| \sum_{i=1}^{H} (H - i + 1) x_i + H \sum_{i=1}^{H} x_{N+i} \right\| \\
&\leq \lim_{N\to\infty} \frac{1}{N} LKH = 0.
\end{aligned}$$

Dove $K < 2$. $\square$

Osservazione 2.2. *Siano* $y_1, ..., y_n \in \mathbb{R}$ *allora* $\left(\frac{1}{N}\sum_{n=1}^{N} y_n\right)^2 \leq \frac{1}{N}\sum_{n=1}^{N} y_n^2$

Dimostrazione. Ponendo $x = (1, ..., 1)$ e $y = (y_1, ..., y_n)$ e applicando la disuguaglianza di Schwarz si ottiene

$$\left(\frac{1}{N}\sum_{n=1}^{N} y_n\right)^2 = \langle \frac{y}{N}, x\rangle \leq \left\|\frac{y}{N}\right\|^2 \|x\|^2 = \frac{1}{N}\sum_{n=1}^{N} y_n^2$$

$\square$

Abbiamo inoltre che applicando prima la disuguaglianza triangolare, poi la Osservazione 2.2, si ottiene:

$$\begin{aligned}
\|\Psi_N\|^2 &= \left\|\frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{H}\sum_{h=1}^{H} x_{n+h}\right)\right\|^2 \leq \left(\frac{1}{N}\sum_{n=1}^{N}\left\|\frac{1}{H}\sum_{h=1}^{H} x_{n+h}\right\|\right)^2 \\
&\leq \frac{1}{N}\sum_{n=1}^{N}\left\|\frac{1}{H}\sum_{h=1}^{H} x_{n+h}\right\|^2 = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{H^2}\sum_{h,k=1}^{H}\langle x_{n+h}, x_{n+k}\rangle \\
&= \frac{1}{H^2}\sum_{h=1}^{H}\left(\frac{1}{N}\sum_{n=1}^{N}\|x_{n+h}\|^2 + \sum_{k=1,k\neq h}^{H}\left(\frac{1}{N}\sum_{n=1}^{N} 2\langle x_{n+h}, x_{n+k}\rangle\right)\right) \\
&= \frac{1}{H^2}\sum_{h=1}^{H}\left(\frac{1}{N}\sum_{n=1}^{N}\|x_{n+h}\|^2\right) + \frac{2}{H^2}\sum_{h=1}^{H}\sum_{k=1,k\neq h}^{H}\left(\frac{1}{N}\sum_{n=1}^{N}\langle x_{n+h}, x_{n+k}\rangle\right) \\
&\leq \frac{L^2}{H} + \Psi'_N < \epsilon + \Psi'_N.
\end{aligned}$$

Dove $\Psi'_N = \frac{2}{H^2}\sum_{h=1}^{H}\sum_{k=1,k\neq h}^{H}(\frac{1}{N}\sum_{n=1}^{N}\langle x_{n+h}, x_{n+k}\rangle) \to 0$ per $N \to \infty$. $\square$

Diamo adesso una definizione

Definizione 2.4 (Spettro Razionale)**.** Sia $(X, \mathcal{A}, \mu)$ uno spazio probabilizzato e $T : X \to X$ μ-invariante. Sia U l'operatore unitario su $L^2(X, \mathcal{A}, \mu)$ indotto da T. Definiamo *spettro razionale* di U l'insieme

$$\mathcal{H}_{rat} = \overline{\{f \in L^2(X, \mathcal{A}, \mu) \,:\, U^k f = f \textit{ per qualche } k \in \mathbb{N}\}}.$$

Prima di cominciare con la dimostrazione facciamo un paio di osservazioni preliminari:

Osservazione 2.3. $\mathcal{H}_{rat}$ *è sottospazio lineare chiuso di* $L^2(X, \mathcal{A}, \mu)$.

Osservazione 2.4. *Sia* $(X, \mathcal{A}, \mu)$ *uno spazio probabilizzato,* U *un operatore unitario su* $L^2(X, \mathcal{A}, \mu)$ *e* $A \in \mathcal{A}$ *tale che* $\mu(A) > 0$. *Indicando con* 1_A *la funzione caratteristica di* A *e con* P_{rat} *la proiezione ortogonale su* $\mathcal{H}_{rat}$, *abbiamo:*

1. $1_A = g + h$ *con* $g = P_{rat} 1_A \in \mathcal{H}_{rat}$ *e* $h = 1_A - g \in \mathcal{H}_{rat}^{\perp}$
2. $\langle g, 1_A \rangle \geq \mu(A)^2$.

Dimostrazione. Per dimostrare la prima parte ci basta applicare il Teorema 2.5 ottenendo che $L^2(X, \mathcal{A}, \mu) = \mathcal{H}_{rat} \oplus \mathcal{H}_{rat}^{\perp}$ e quindi la tesi.
Per la seconda parte osserviamo che le funzioni costanti c con $c \in \mathbb{R}$ appartengono a $\mathcal{H}_{rat}$. Infatti per definizione $Uc(x) = cT(x) = c$. Prendendo quindi $c = \mu(A)$ abbiamo che, per la disuguaglianza di Schwarz,

$$\mu(A)^2 \;=\; \langle \mu(A), 1_A \rangle \;=\; \langle \mu(A), g \rangle \;\leq\; \mu(A) \cdot \|g\|.$$

Perciò

$$\langle g, 1_A \rangle \;=\; \langle g, g \rangle \;+\; \langle g, h \rangle \;=\; \|g\|^2 \geq \mu(A)^2$$

□

Possiamo adesso passare alla dimostrazione vera e propria del Teorema 2.3 che rienunciamo di seguito

Teorema (Särkozy-Fürstenberg versione ergodica)**.** Sia $p(x) \in \mathbb{Z}[x]$ con $p(0) = 0$ allora $R = \{p(n) \;:\; n \in \mathbb{Z}\}$ è sequenza di Poincaré; cioè dato $(X, \mathcal{A}, \mu, T)$ sistema che conserva la misura e $A \in \mathcal{A}$ con $\mu(A) > 0$ allora chiamando $\Gamma = \{n \,:\, \mu(A \cap T^{-n}A) > 0\}$ si ha $R \cap \Gamma \neq \emptyset$.

Dimostrazione. Sia $(X, \mathcal{A}, \mu, T)$ un sistema che conserva la misura, A un elemento di $\mathcal{A}$ di misura positiva e U l'operatore unitario su $L^2(X, \mathcal{A}, \mu)$ indotto da T. Per l'Osservazione 2.3 e il Teorema 2.5 abbiamo che $L^2(X, \mathcal{A}, \mu) = \mathcal{H}_{rat} \oplus \mathcal{H}_{rat}^{\perp}$ e $1_A = g + h$.
Fissiamo $\epsilon > 0$ arbitrario. Abbiamo che $g \in \mathcal{H}_{rat}$ che è la chiusura di un insieme. Perciò esiste $f \in L^2(X, \mathcal{A}, \mu)$ e $k \in \mathbb{N}$ tali che $U^k f = f$ e $\|f - g\| < \epsilon$. Dato p, per ogni $n \in \mathbb{N}$ abbiamo che $k | p(kn)$ e quindi

$$\|U^{p(kn)}g - g\| \leq \|U^{p(kn)}g - U^{p(kn)}f\| + \|U^{p(kn)}f - f\| + \|f - g\| < 2\epsilon.$$

Abbiamo quindi anche

$$\Big\| \frac{1}{N} \sum_{n=1}^{N} U^{p(kn)} g - g \Big\| < 2\epsilon. \tag{4}$$

Vogliamo dimostrare, che al crescere di N, 1_A in media si avvicina arbitrariamente a g. Perciò proveremo che

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}h\right\|=0 \tag{5}$$

per induzione sul grado di p.
Per verificare il caso base ($p=an$) abbiamo che dal Teorema della media ergodica per $N\to\infty$ $\frac{1}{N}\sum_{n=1}^{N}U^{an}h\to P_a h$ in norma, dove P_a è la proiezione sullo spazio delle funzioni U^a-invarianti, che però è contenuto in $\mathcal{H}_{rat}$, perciò $P_a h=0$.
Per il caso induttivo utilizzeremo il Lemma 2.1 ponendo $x_n=U^{p(kn)}h$. Lo spazio X è probabilizzato ($\mu(X)=1$) e $x_n\in L^2(X,\mathcal{A},\mu)$ $\forall\, n$, quindi la successione $\{x_n\}_{n\in\mathbb{Z}}$ è limitata. Inoltre abbiamo che per ogni $r\in\mathbb{Z}$

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle x_n,x_{n+r}\rangle &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle U^{p(kn)}h,U^{p(kn+kr)}h\rangle\\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle h,U^{p(kn+kr)-p(kn)}h\rangle\\
&= \lim_{N\to\infty}\langle U^{-p(kr)}h,\frac{1}{N}\sum_{n=1}^{N}U^{p(kn+kr)-p(kn)-p(kr)}h\rangle\\
&\le \lim_{N\to\infty}\|h\|\cdot\left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn+kr)-p(kn)-p(kr)}h\right\|=0.
\end{aligned}$$

Dove la disuguaglianza si ottiene applicando la disuguaglianza di Schwarz. Per giustificare l'ultima uguaglianza dobbiamo osservare che, detto $q(kn)=p(kn+kr)-p(kn)-p(kr)$ abbiamo che $\deg q<\deg p$ e che $q(0)=0$. Infatti se

$$p(x)=a_s x^s+...+a_1 x$$

allora $p(kn+kr)=a_s(kn+kr)^s+...+a_1(kn+kr)$, il suo termine di grado maggiore è $a_s(kn)^s$ e il suo termine noto è $a_s(kr)^s+...+a_1(kr)=p(kr)$. Abbiamo quindi che togliendo $p(kn)$ da $p(kn+kr)$ abbassiamo di grado quest'ultimo senza però toccare il termine noto ($p(kn)$ non ha termini noti), togliendo poi anche $p(kr)$ otteniamo un $q(kn)$ a cui possiamo applicare l'ipotesi induttiva, da cui l'uguaglianza con 0. Adesso utilizzando (4) e (5) abbiamo che per N abbastanza grande

$$\left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}1_A-g\right\| \;\le\; \left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}g-g\right\|+\left\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}h\right\|\le 2\epsilon+\epsilon.$$

Quindi (sempre utilizzando la disuguaglianza di Schwarz)

$$\begin{aligned}
\Big|\frac{1}{N}\sum_{n=1}^{N}\Big(\langle U^{p(kn)}1_A,1_A\rangle-\langle g,1_A\rangle\Big)\Big| &= |\langle\ \Big(\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}1_A\Big)-g,1_A\rangle| \\
&\le \Big\|\frac{1}{N}\sum_{n=1}^{N}U^{p(kn)}1_A-g\Big\|\cdot\|1_A\|\le 3\epsilon.
\end{aligned}$$

In particolare, esisterà un $\bar{n}\le N$ tale che $|\langle U^{p(k\bar{n})}1_A,1_A\rangle-\langle g,1_A\rangle|\le 3\epsilon$ e quindi

$$\mu(A\cap T^{-p(k\bar{n})}A)=\langle U^{p(k\bar{n})}1_A,1_A\rangle\ge\langle g,1_A\rangle-3\epsilon\ge\mu(A)^2-3\epsilon.$$

Dove nell'ultima disuguaglianza abbiamo utilizzato l'Osservazione 2.4. Prendendo ϵ sufficentemente piccolo ($\epsilon<\frac{\mu(A)^2}{3}$) abbiamo la tesi. $\square$

Osserviamo infine che anche qua il risultato provato è più forte di quello espresso nell'enunciato, infatti abbiamo che $\forall\epsilon>0$ possiamo trovare un $\bar{n}$ e un k tali che $\mu(A\cap T^{-p(k\bar{n})}A)\ge\mu(A)^2-\epsilon$.

Riferimenti bibliografici

[Ber96] Vitaly Bergelson. Ergodic ramsey theory—an update. *London Math. Soc. Lecture Note Ser.*, 228:1–61, 1996. From: Ergodic theory of Z^d actions (Warwick, 1993–1994).

[Ber06] Vitaly Bergelson. Combinatorial and diophantine applications of ergodic theory. In *Handbook of dynamical systems*, volume 1B, pages 745–869. Elsevier B. V., Amsterdam, 2006.

[Für77] Harry Fürstenberg. Ergodic behavior of diagonal measures and a theorem of szemerédi on arithmetic progressions. *Journal d'Analyse Mathématique*, 31:204–256, 1977.

[Für81] Harry Fürstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, 1981.

[McC99] Randall McCutcheon. *Elemental methods in ergodic Ramsey theory.* Springer-Verlag, Berlin, 1999. Lecture Notes in Mathematics,1722.

[Sze75] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[Sár78] András Sárközy. On difference sets of sequences of integers iii. *Acta Math. Acad. Sci. Hungar.*, 31(3-4):355–386, 1978.