Εργοδική Θεωρία

Πρόχειρες Σημειώσεις

Αγγελική Μενεγάκη

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΑΘΗΝΑ, 2017

Περιεχόμενα

Εισαγωγή

Η Εργοδική Θεωρία είναι η μελέτη της μακροχρόνιας συμπεριφοράς συστημάτων, κατά μέσο όρο. Όλες οι καταστάσεις ενός συστήματος διαμορφώνουν έναν χώρο X. Η εξέλιξη ενός συστήματος γίνεται φανερή μέσω ενός μετασχηματισμού $T\colon X \to X$, όπου $T(x)$ εκπροσωπεί την κατάσταση του συστήματος την χρονική στιγμή 1, ενώ την χρονική στιγμή μηδέν η κατάσταση είναι x.

Στην περίπτωση που το σύστημα εξελίσσεται με συνεχή τρόπο, μπορούμε να θεωρήσουμε την μονοπαραμετρική οικογένεια απεικονίσεων απ το X στο X, $\{T_t\colon t \in \mathbb{R}\}$. Επίσης, αν το σύστημα είναι αναλλοίωτο ως προς το χρόνο, τότε θα ισχύει $T_{t+s} = T_t \circ T_s$, δηλαδή η $\{T_t\colon t \in \mathbb{R}\}$ είναι ροή ή δράση κάποιας ομάδας παραμετρικοποιημένης από το $\mathbb{R}$ στο X.

Για να αναλύσουμε ένα σύστημα χρειαζόμαστε, φυσικά, μια δομή στο X και κάποιους περιορισμούς για τον μετασχηματισμό T. Υπάρχουν τρεις κύριες περιπτώσεις:

1. Αν X είναι χώρος μέτρου και $T\colon X \to X$ μετασχηματισμός που διατηρεί το μέτρο, έχουμε την περίπτωση της εργοδικής θεωρίας.
2. Αν X είναι τοπολογικός χώρος και $T\colon X \to X$ συνεχής απεικόνιση, έχουμε τα τοπολογικά δυναμικά συστήματα.
3. Αν X είναι διαφορίσιμη πολλαπλότητα και $T\colon X \to X$ διαφορίσιμη απεικόνιση, έχουμε λείο (smooth) δυναμικό σύστημα.

1 Μετρήσιμα Δυναμικά Συστήματα

1.1 Ορισμοί και παραδείγματα

Ορισμός 1.1.1 (μετρήσιμο δυναμικό σύστημα). Ένα *σύστημα που διατηρεί το μέτρο* (ή *μετρήσιμο δυναμικό σύστημα*) είναι μια τετράδα $(X, \mathcal{B}, \mu, T)$, όπου $(X, \mathcal{B}, \mu)$ είναι χώρος πιθανότητας και $T\colon X \to X$ είναι μετασχηματισμός που είναι μετρήσιμος, δηλαδή $T^{-1}\mathcal{B} \subset \mathcal{B}$, και διατηρεί το μέτρο, δηλαδή $\mu\big(T^{-1}(A)\big) = \mu(A) \quad \text{για κάθε} \quad A \in \mathcal{B}$.

Συμβολισμοί: $T^{-1}\mathcal{B} = \{T^{-1}B\colon B \in \mathcal{B}\}$.

Λήμμα. Αν $(X, \mathcal{B}, \mu)$ είναι χώρος πιθανότητας, $(Y, \mathcal{A})$ μετρήσιμος χώρος και $T\colon X \to Y$ μετρήσιμη απεικόνιση ($T^{-1}\mathcal{A} \subset \mathcal{B}$), τότε η σχέση

$$T_*\mu(A) = \mu\big(T^{-1}(A)\big), \quad A \in \mathcal{A}, \tag{1}$$

ορίζει μέτρο στον Y.

Με $T_*\mu$ θα συμβολίζουμε το μέτρο που ορίζεται από την (1), εικόνα του μ μέσω της T, και θα λέμε ότι ο μετασχηματισμός T διατηρεί το μέτρο, ή ότι το μέτρο είναι T-αναλλοίωτο, αν $T_*\mu = \mu$.

Ορισμός 1.1.2 (αντιστρέψιμο σύστημα που διατηρεί το μέτρο). Ένα *αντιστρέψιμο* σύστημα που διατηρεί το μέτρο είναι ένα σύστημα που διατηρεί το μέτρο $(X, \mathcal{B}, \mu, T)$ για το οποίο η T είναι αντιστρέψιμη και η T^{-1} είναι μετρήσιμη ($T\mathcal{B} \subset \mathcal{B}$).

Ορισμός 1.1.3 Έστω X σύνολο και $\mathcal{A} \subset \mathcal{P}(X)$[1]. Η σ-*άλγεβρα που παράγεται από την* $\mathcal{A}$ και συμβολίζεται με $\sigma(\mathcal{A})$, είναι η μικρότερη σ-άλγεβρα που περιέχει την $\mathcal{A}$, δηλαδή

$$\sigma(\mathcal{A}) = \bigcap_{\substack{\mathcal{A} \subset \mathcal{B} \\ \mathcal{B}\colon\ \text{σ-άλγεβρα}}} \mathcal{B}.$$

Ορισμός 1.1.4 (σ-άλγεβρα παραγόμενη από συναρτήσεις). Αν $\mathcal{F}$ είναι μια οικογένεια συναρτήσεων $f\colon X \to \mathbb{R}$ (ή $f\colon X \to \mathbb{C}$), η σ-*άλγεβρα που παράγουν* αυτές οι συναρτήσεις και συμβολίζεται με $\sigma(\mathcal{F})$, είναι η μικρότερη σ-άλγεβρα ως προς την οποία όλες οι $f \in \mathcal{F}$ είναι Borel-μετρήσιμες.

Ορισμός 1.1.5 (π-σύστημα). Έστω X σύνολο. Μια οικογένεια $\mathcal{P}$ υποσυνόλων του X είναι π-*σύστημα* αν είναι κλειστή ως προς πεπερασμένες τομές.

Παράδειγμα. Είναι σαφές ότι στον $\mathbb{R}$, το σύνολο $\{(a, b)\colon a, b \in \mathbb{R}\} \cup \{\emptyset\}$ είναι π-σύστημα.

Λήμμα 1.1.6. Έστω $(X, \mathcal{B}, \mu)$ χώρος πιθανότητας και $T\colon X \to X$ μετρήσιμος μετασχηματισμός. Ο T διατηρεί το μέτρο μ ανν $\mu\big(T^{-1}(A)\big) = \mu(A)$ για κάθε $A \in \mathcal{P}$, όπου $\mathcal{P}$ είναι π-σύστημα που παράγει την σ-άλγεβρα $\mathcal{B}$.

Απόδειξη. Εφόσον $T_*\mu$ και μ είναι μέτρα πιθανότητας και ταυτίζονται σ' ένα π-σύστημα που παράγει την σ-άλγεβρα, το συμπέρασμα είναι άμεσο πόρισμα του θεωρήματος $\pi - \lambda$. [2] $\square$

Λήμμα 1.1.7. Έστω $(X, \mathcal{B}, \mu)$ χώρος πιθανότητας και $T\colon X \to X$ μετρήσιμος μετασχηματισμός.

(i) αν $\int_X f \circ T \, d\mu = \int_X f \, d\mu$ για κάθε $f \in L^\infty(\mu)$, τότε ο T διατηρεί το μέτρο ($T_*\mu = \mu$).

(ii) αν $T_*\mu = \mu$, τότε $\int_X f \circ T \, d\mu = \int_X f \, d\mu$ για κάθε $f \in L^1(\mu)$ και για κάθε f θετική, μετρήσιμη.

[1] Με $\mathcal{P}(X)$ συμβολίζεται το δυναμοσύνολο ενός συνόλου X; χρησιμοποιείται και ο συμβολισμός $2^X = \mathcal{P}(X)$.

[2] Θεώρημα $\pi - \lambda$: Αν X σύνολο και $\mathcal{A} \subset \mathcal{P}(X)$ π-σύστημα, τότε η κλάση Dynkin που παράγεται απ την $\mathcal{A}$ είναι ίση με την $\sigma(\mathcal{A})$.

Απόδειξη. (i) Θέλουμε να δείξουμε ότι $\mu\big(T^{-1}(A)\big) = \mu(A)$ για κάθε $A \in \mathcal{B}$. Αν $A \in \mathcal{B}$, παίρνοντας $f = \mathbb{1}_A$, τότε $f \circ T = \mathbb{1}_{T^{-1}(A)}$, και άρα

$$\int_X f \circ T \, d\mu = \int_X f \, d\mu \quad \Leftrightarrow \quad \int_X \mathbb{1}_{T^{-1}(A)} \, d\mu = \int_X \mathbb{1}_A \, d\mu \quad \Leftrightarrow \quad \mu\big(T^{-1}(A)\big) = \mu(A).$$

(ii) Αν $f = \mathbb{1}_A$, όπου $A \in \mathcal{B}$, τότε όπως στο (i) έχουμε το ζητούμενο. Αν f απλή, $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$, $a_i \geq 0, A_i \in \mathcal{B}$, τότε λόγω γραμμικότητας του ολοκληρώματος, ισχύει η σχέση για την f. Αν $f \colon X \to [0, \infty]$ μετρήσιμη, τότε υπάρχει αύξουσα ακολουθία $(s_n)_n$ θετικών, μετρήσιμων έτσι ώστε $s_n \to f$. Το (ii) ισχύει για τις s_n και από το θεώρημα μονότονης σύγκλισης, ισχύει το ζητούμενο. Αν $f \in L^1(\mu)$, τότε $f = f^+ - f^-$ και είμαστε στην προηγούμενη περίπτωση. □

Παρατήρηση. Δείξαμε στο (ii) της παραπάνω απόδειξης ότι $\int f \, dT_*\mu = \int f \circ T \, d\mu$ για κάθε $f \in L^1(\mu)$ και για κάθε f θετική, μετρήσιμη.

$\mathbb{S}^1 = \{e^{2\pi i t} \colon t \in [0,1)\} = \{z \in \mathbb{C} \colon |z| = 1\}$ είναι η πολλαπλασιαστική ομάδα μιγαδικών αριθμών μέτρου ένα. Σαν τοπολογική ομάδα

$$\mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z} = \{x + \mathbb{Z} \colon x \in \mathbb{R}\} \cong \mathbb{T} = [0,1)$$

όπου τα σημεία $0, 1$ είναι ταυτοποιημένα. Η τοπολογία στο $[0,1)$ είναι αυτή που έχει σαν βάση περιοχών τα διαστήματα $(x - \varepsilon, x + \varepsilon) \cap [0,1)$, αν $x \in (0,1)$ και $[0, \varepsilon) \cup (1 - \varepsilon, \varepsilon)$, αν $x = 0$, όπου $\varepsilon > 0$. Μία μετρική που είναι συμβατή με αυτή την τοπολογία είναι η $d_{\mathbb{T}}(t,s) := \min\{|t-s|, 1 - |t-s|\}$, $t, s \in \mathbb{T} = [0,1)$. Η τοπολογία στον $\mathbb{R}/\mathbb{Z}$ είναι, εξ' ορισμού, η μικρότερη τοπολογία που κάνει την απεικόνιση $x \mapsto x + \mathbb{Z}$, $x \in \mathbb{R}$, συνεχή. Μία συμβατή μετρική είναι η $d_{\mathbb{R}/\mathbb{Z}}(x+\mathbb{Z}, y+\mathbb{Z}) := \min_{m \in \mathbb{Z}} |x - y + m|$. Τέλος, η τοπολογία στον κύκλο $\mathbb{S}^1$ είναι η συνήθης τοπολογία που έχει ως υποσύνολο του επιπέδου (σχετική τοπολογία). Ως συμβατή μετρική μπορούμε να πάρουμε είτε την συνηθισμένη ευκλείδεια μετρική στο επίπεδο είτε την μετρική «μήκος τόξου» $d_{\mathbb{S}^1}(z,w) := 2 \arcsin(|z - w|/2)$ για $z, w \in \mathbb{S}^1$.

Με αυτή τις τοπολογίες αυτές στα $[0,1)$, $\mathbb{R}/\mathbb{Z}$ και $\mathbb{S}^1$, οι απεικονίσεις $t \mapsto e^{2\pi i t}$ και $t \mapsto t + \mathbb{Z}$ είναι ομοιομορφισμοί μεταξύ των $[0,1)$ και $\mathbb{S}^1$ και $[0,1)$ και $\mathbb{R}/\mathbb{Z}$, αντίστοιχα.

Παράδειγμα 1.1.8 $X = [0,1)$, $\mathcal{B} = \mathcal{B}[0,1)$, λ το μέτρο Lebesgue στο $[0,1)$. Ορίζουμε

$$T_a(x) := x + a \ (\mathrm{mod}\ 1) = x + a - \lfloor x + a \rfloor \quad \text{ή}$$
$$T_a(x + \mathbb{Z}) = x + \mathbb{Z} + (a + \mathbb{Z}) = x + a + \mathbb{Z}.$$

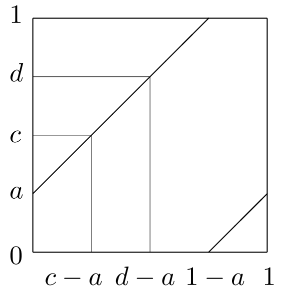
Ισοδύναμα, αν πάρουμε $X = \mathbb{S}^1$, $\mathcal{B} = \mathcal{B}(\mathbb{S}^1)$, λ το μέτρο Lebesgue στην $\mathbb{S}^1$,

$$T_a(z) = wz, \qquad \text{όπου} \quad w = e^{2\pi i a} \in \mathbb{S}^1.$$

Η T_a διατηρεί το μέτρο, για κάθε a. Πράγματι, αρκεί να δειχτεί ότι

$$\lambda\big(T_a^{-1}([c,d])\big) = \lambda([c,d])$$

λόγω του λήμματος 1.1.6, αφού τα $[c,d]$ είναι διαστήματα και τα διαστήματα αυτά, για $0 \leq c \leq d < 1$, μαζί με το $\varnothing$, αποτελούν π σύστημα που παράγει την σ-άλγεβρα. Τώρα, επειδή η ευθεία έχει κλίση 1, βλέπουμε από το παρακάτω σχήμα ότι μήκος$\big(T_a^{-1}([c,d])\big)$ = μήκος$([c,d])$.



Μπορεί να υπολογίσει κανείς ότι

$$T_a^{-1}([c,d]) = \begin{cases} [c-a, d-a] & \text{αν } a \leq c < d < 1 \\ [c-a+1, d-a+1] & \text{αν } 0 \leq c < d < a \\ [0, d-a] \cup [c-a+1, 1) & \text{αν } 0 \leq c < a \leq d < 1. \end{cases}$$

Ορισμός 1.1.9 *Τοπολογική ομάδα* είναι μια ομάδα G εφοδιασμένη με μια τοπολογία ως προς την οποία η πράξη $(x,y) \mapsto xy$ και η αντίστροφη $x \mapsto x^{-1}$ να είναι συνεχείς συναρτήσεις. (Για την συνέχεια της πράξης θεωρούμε την τοπολογία γινόμενο στην $G \times G$.)

Συμβολισμοί: Αν G ομάδα $A \subset G$ και $x \in G$, $xA := \{xa\colon a \in A\}$.

Θεώρημα 1.1.10 Αν G τοπικά συμπαγής ομάδα με Hausdorff τοπολογία, τότε υπάρχει ένα κανονικό μέτρο Borel στην G που είναι αναλλοίωτο ως προς όλες τις μεταφορές. Δηλαδή υπάρχει κανονικό[3] μέτρο Borel λ_G τέτοιο ώστε

$$\lambda_G(aB) = \lambda_G(B), \qquad \text{για κάθε } B \in \mathcal{B}(G) \text{ και } a \in G. \tag{2}$$

Αυτό το μέτρο ονομάζεται ένα αριστερό μέτρο Haar της ομάδας.

Το λ_G είναι ουσιαστικά μοναδικό, με την έννοια ότι αν λ' είναι ένα άλλο μέτρο που ικανοποιεί την (2), τότε υπάρχει σταθερά $c \in (0,\infty)$ τέτοια ώστε $\lambda_G = c\lambda'$.

Τέλος, $\lambda_G(U) > 0$, για κάθε U ανοικτό υποσύνολο της G.

Τα μέτρα Haar μιας ομάδας είναι πεπερασμένα ανν η ομάδα είναι συμπαγής. Στην περίπτωση αυτή, όταν θα λέμε μέτρο Haar θα εννοούμε το αριστερό μέτρο Haar για το οποίο $\lambda_G(G) = 1$.

Παραδείγματα.

- Στον $\mathbb{R}^n$, το μέτρο Haar είναι το μέτρο Lebesgue (ή οποιοδήποτε πολλαπλάσιο).
- Το ίδιο και για τον τόρο $\mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z} \cong \mathbb{T} = [0,1)$ και για τους n-διάστατους τόρους $\mathbb{S}^1 \times \mathbb{S}^1 \times \cdots \times \mathbb{S}^1 \cong \mathbb{R}^n/\mathbb{Z}^n \cong [0,1)^n$.

Παράδειγμα 1.1.11 Έστω $X = G$, όπου G τοπολογική ομάδα, συμπαγής, $\mathcal{B} = \mathcal{B}(G)$, λ_G το μέτρο Haar της ομάδας G. Για οποιοδήποτε $a \in G$, ο μετασχηματισμός

$$T_a\colon G \to G, \quad T_a(x) = ax \quad \text{για} \quad x \in G$$

διατηρεί το μέτρο Haar (εξ' ορισμού):

$$(T_a)_*\lambda_G(B) = \lambda_G(T_a^{-1}B) = \lambda_G(a^{-1}B) = \lambda_G(B) \quad \text{για κάθε} \quad B \in \mathcal{B}.$$

Παράδειγμα 1.1.12 Οι n-διάστατοι τόροι $X = [0,1)^n$, $\mathcal{B} = \mathcal{B}\big([0,1)^n\big)$, λ το n-διάστατο μέτρο Lebesgue στο $[0,1)^n$, $a = (a_1,\dots,a_n) \in [0,1)^n$

$$T_a(x_1,\dots,x_n) = \big(x_1 + a_1 \ (\mathrm{mod}\ 1), \dots, x_n + a_n \ (\mathrm{mod}\ 1)\big).$$

Παράδειγμα 1.1.13 (a) $X = \mathbb{T}$, $\mathcal{B} = \mathcal{B}(\mathbb{T})$, $\mu = \lambda$ το μέτρο Lebesgue στο $[0,1)$. Θεωρούμε τον μετασχηματισμό

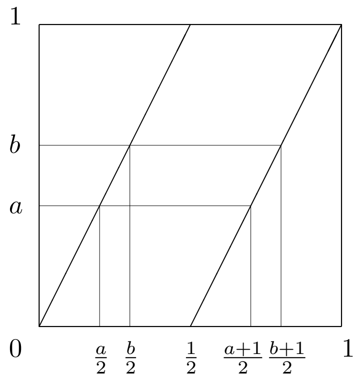
$$T\colon X \to X, \qquad T(x) = 2x \ (\mathrm{mod}\ 1) = 2x - \lfloor 2x \rfloor, \qquad x \in [0,1).$$

[3] Κανονικό εδώ θα πει ότι κάθε συμπαγές σύνολο έχει πεπερασμένο μέτρο, το μέτρο κάθε Borel υποσυνόλου προσεγγίζεται από έξω με τα μέτρα ανοικτών συνόλων που το περιέχουν και το μέτρο κάθε ανοικτού συνόλου προσεγγίζεται από μέσα από τα μέτρα συμπαγών συνόλων που περιέχονται στο ανοικτό.

Ο T διατηρεί το μέτρο Lebesgue. Πράγματι, αρκεί να δειχτεί ότι

$$\lambda\big(T^{-1}(A)\big) = \lambda(A), \quad \text{για κάθε} \quad A = (a,b),$$

που ισχύει, εφόσον οι ευθείες έχουν κλίση 2 και άρα το $T^{-1}(a,b)$ έχει το ίδιο μέτρο Lebesgue με το (a,b).



2ος τρόπος ότι ο T διατηρεί το μέτρο Lebesgue. Αν $f \in L^\infty(\lambda)$, θέλουμε

$$\int_0^1 f(T(x))\,dx = \int_0^1 f(x)\,dx.$$

Ορίζοντας $\tilde{f} \colon \mathbb{R} \to \mathbb{R}$, 1-περιοδική με $f = \tilde{f}\mid_{[0,1)}$, έχουμε

$$\begin{aligned}\int_0^1 f(T(x))\,dx &= \int_0^{1/2} f(2x)\,dx + \int_{1/2}^1 \tilde{f}(2x)\,dx \\ &= \frac{1}{2}\int_0^1 f(x)\,dx + \frac{1}{2}\int_1^2 \tilde{f}(x)\,dx = \int_0^1 f(x)\,dx.\end{aligned}$$

□

(b) Γενικότερα, αν $X = [0,1)$, $\mathcal{B} = \mathcal{B}\big([0,1)\big)$, $\mu = \lambda$ το μέτρο Lebesgue και, για όλα τα $k \in \mathbb{N}$, θεωρούμε τους μετασχηματισμούς

$$T_k(x) = kx \ (\mathrm{mod}\ 1) \qquad \big(\text{στην } \mathbb{S}^1 \colon T_k(z) = z^k\big).$$

Ισχύει ότι οι T_k διατηρούν το μέτρο Lebesgue.

Για την $T_2(x) = 2x$ (mod 1) ένα αναλλοίωτο μέτρο είναι το μέτρο Lebesgue, αλλά δεν είναι μοναδικό. Έστω

$$\mu = \tfrac{1}{2}\delta_{\{1/3\}} + \tfrac{1}{2}\delta_{\{2/3\}}.$$

Το μ είναι αναλλοίωτο ως προς τον T_2. Πράγματι,

$$\big(T_2\big)_*\delta_{\{1/3\}}(A) = \delta_{\{1/3\}}\big(T_2^{-1}(A)\big) = \mathbb{1}_{T_2^{-1}(A)}\big(\tfrac{1}{3}\big) = \mathbb{1}_A\big(T_2\big(\tfrac{1}{3}\big)\big) = \delta_{\{2/3\}}(A)$$

και όμοια δείχνουμε ότι

$$\big(T_2\big)_*\,\delta_{\{2/3\}} = \delta_{\{1/3\}}.$$

Εικασία Furstenberg: Αν μ είναι συνεχές μέτρο πιθανότητας στον $\mathbb{T}$ τέτοιο ώστε

$$\big(T_2\big)_*\,\mu = \mu \quad \text{και} \quad \big(T_3\big)_*\,\mu = \mu$$

τότε το μ είναι το μέτρο Lebesgue.

Παράδειγμα 1.1.14 Έστω $X = \mathbb{T}^n \cong \mathbb{R}^n/\mathbb{Z}^n$ και A ένας $n \times n$ πίνακας με $A_{ij} \in \mathbb{Z}$. Τότε ο A ορίζει μια απεικόνιση

$$T \colon \mathbb{T}^n \to \mathbb{T}^n, \qquad T\big(x + \mathbb{Z}^n\big) = Ax + \mathbb{Z}^n, \quad x \in \mathbb{R}^n.$$

Η παραπάνω απεικόνιση είναι καλά ορισμένη, αφού αν $x, y \in \mathbb{R}^n$ τέτοια ώστε $x - y \in \mathbb{Z}^n \Rightarrow Ax - Ay = A(x-y) \in \mathbb{Z}^n$.

Άσκηση. *Έστω G ομάδα, H υπο-ομάδα και $\varphi\colon G \to G$ ομομορφισμός τέτοιος ώστε $\varphi\big(H\big) \subset H$. Τότε*

$$\tilde{\varphi}\colon G/H \to G/H, \qquad \tilde{\varphi}\big(xh\big) := \varphi(x)H$$

ορίζει καλά μια απεικόνιση. Αν $H \lhd G$, τότε η $\tilde{\varphi}\colon G/H \to G/H$ είναι ομομορφισμός. Τέλος, αν η φ είναι επί, τότε και η $\tilde{\varphi}$ είναι επί.

Λόγω της παραπάνω άσκησης, ο $T\big(x + \mathbb{Z}^n\big) = Ax + \mathbb{Z}^n$ (ή $T(x) = Ax \pmod 1$, $x \in \mathbb{T}^n$) είναι επιμορφισμός της ομάδας $\mathbb{R}^n/\mathbb{Z}^n$ όταν $\det(A) \neq 0$.
Επίσης, ο μετασχηματισμός T διατηρεί το μέτρο Lebesgue, ως ειδική περίπτωση του επόμενου παραδείγματος.

Παράδειγμα. Έστω G συμπαγής ομάδα, Hausdorff και $T\colon G \to G$ Borel-μετρήσιμος επιμορφισμός. Τότε ο T διατηρεί το μέτρο Haar της ομάδας G.

Απόδειξη. Θέλουμε $T_*\lambda_G = \lambda_G$. Θέτουμε $\mu := T_*\lambda_G$, το οποίο είναι μέτρο πιθανότητας (αφού και το λ_G είναι μέτρο πιθανότητας).
Ισχυρισμός: Το μ είναι επίσης μέτρο Haar για την G. Τότε, εφόσον είναι και τα δύο μέτρα πιθανότητας, από μοναδικότητα του μέτρου Haar, έχουμε το ζητούμενο. Για την απόδειξη του ισχυρισμού, παίρνουμε ένα $B \in \mathcal{B}(G)$ και $y \in G$. Αφού ο T είναι επιμορφισμός, υπάρχει $x \in G$ τέτοιο ώστε $T(x) = y$, και τότε

$$\mu(yB) = \big(T_*\lambda_G\big)(yB) = \lambda_G\big(T^{-1}(yB)\big) = \lambda_G\big(x\,T^{-1}(B)\big) = \lambda_G\big(T^{-1}(B)\big) = T_*\lambda_G(B) = \mu(B),$$

η τέταρτη ισότητα επειδή το λ_G είναι μέτρο Haar. Για την τρίτη ισότητα εύκολα ελέγχει κανείς ότι $T^{-1}(yB) = x\,T^{-1}(B)$. $\square$

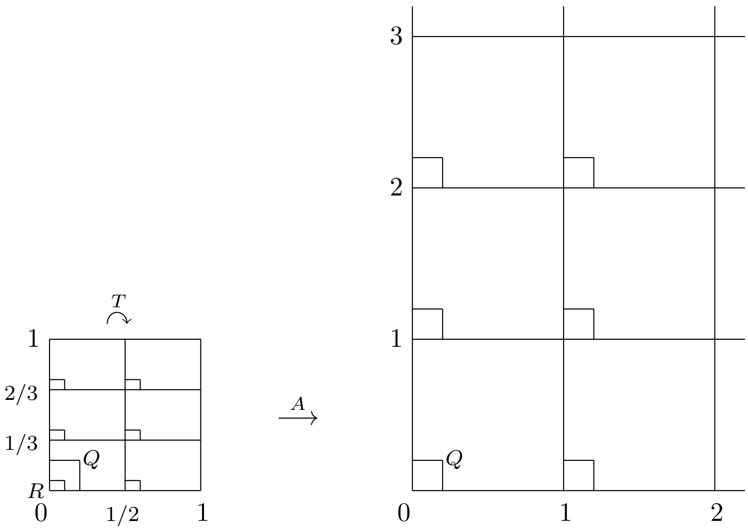
Για παράδειγμα, έστω

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Ο γραμμικός μετασχηματισμός που ορίζει ο πίνακας A στον $\mathbb{R}^2$ απεικονίζει το $[0,1)^2$ στο ορθογώνιο $A[0,1)^2$ με κορυφές

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Έστω Q ένα ορθογώνιο στο $[0,1)^2$. Η αντίστροφη εικόνα του μέσω του μετασχηματισμού T που επάγει ο A στον τόρο $\mathbb{T}^2$ είναι τα έξι ορθογώνια όμοια με το R (ο $x \mapsto Ax$, σαν μετασχηματισμός του R^2 στον εαυτό του, στέλνει τα έξι μικρά ορθογώνια όμοια με το R στο $[0,1)^2$ στα έξι τετράγωνα όμοια με το Q στο δεξί σχήμα, και όλα αυτά ταυτίζονται στο Q μέσα στον τόρο στο αριστερό σχήμα).

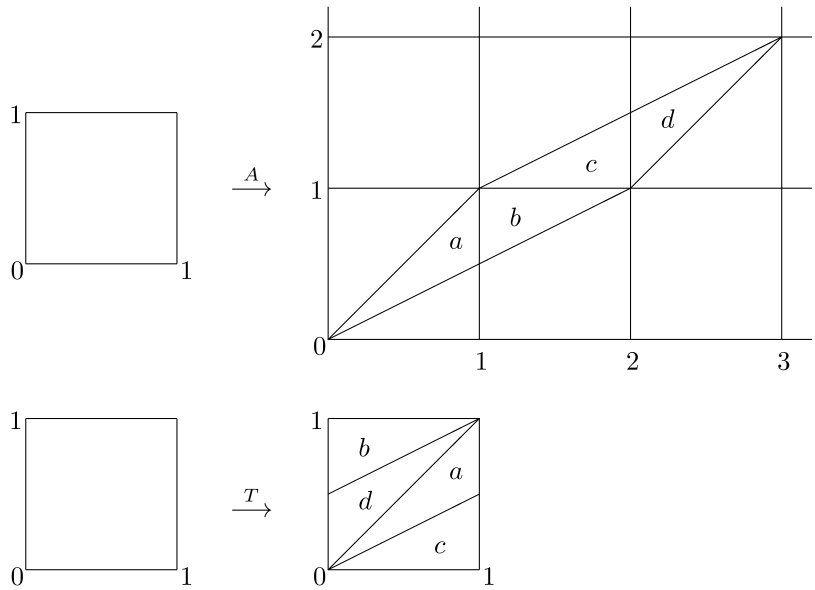


Βλέπουμε ότι το μέτρο διατηρείται σε κάθε τέτοιο ορθογώνιο Q. Αφού τα ορθογώνια είναι π-σύστημα που παράγουν την $\mathcal{B}\big([0,1)^2\big)$, έχουμε ότι ο T διατηρεί το μέτρο Lebesgue στον $[0,1)^2$.

Ακόμα ένα παράδειγμα, με τον A να μην είναι διαγώνιος πίνακας, είναι το εξής:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Ορίζεται αυτομορφισμός του τόρου, αφού $\det(A) = 1$. Ένας τέτοιος T διατηρεί το μέτρο Lebesgue. Το σχήμα θα έχει την εξής μορφή.



Παράδειγμα 1.1.15 *(Απεικόνιση Gauss)* Έστω $T\colon (0,1] \to [0,1]$ η απεικόνιση

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor.$$

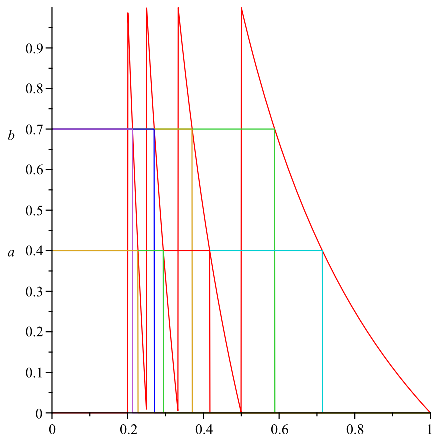
Η T είναι Borel μετρήσιμη, υπό την έννοια ότι $T^{-1}(A)$ είναι Borel υποσύνολο του $(0,1]$ για κάθε Borel υποσύνολο A του $[0,1]$.

Έχει κανείς ότι

$$n = \left\lfloor \frac{1}{x} \right\rfloor \quad \Leftrightarrow \quad n \leq \frac{1}{x} < n+1 \quad \Leftrightarrow \quad \frac{1}{n+1} < x \leq \frac{1}{n}.$$

Άρα

$$T(x) = \frac{1}{x} - n, \qquad \text{όταν} \quad \frac{1}{n+1} < x \leq \frac{1}{n}, \qquad n \in \mathbb{N}.$$



Ισχυρισμός: ο T διατηρεί το μέτρο
$$d\mu(x) = \frac{dx}{x+1}\frac{1}{\ln 2}.$$
Πράγματι, λόγω του λήμματος 1.1.6, αρκεί να δειχτεί ότι
$$\mu\big(T^{-1}(a,b)\big) = \mu\big((a,b)\big).$$
Για $a, b \in (0,1)$,
$$a < T(x) < b \quad \text{και} \quad \frac{1}{n+1} < x \leq \frac{1}{n} \quad \text{ανν} \quad a+n < \frac{1}{x} < b+n \quad \Leftrightarrow \quad \frac{1}{a+n} > x > \frac{1}{b+n}.$$
Άρα
$$T^{-1}(a,b) = \bigcup_{n=1}^{\infty}\left(\frac{1}{b+n}, \frac{1}{a+n}\right).$$
Άρα,
$$\begin{aligned}\mu\big(T^{-1}(a,b)\big) &= \sum_{n\geq 1}\int_{\frac{1}{b+n}}^{\frac{1}{a+n}} \frac{dx}{x+1}\frac{1}{\ln 2} = \sum_{n\geq 1}\left(\ln\Big(\frac{1}{a+n}+1\Big) - \ln\Big(\frac{1}{b+n}+1\Big)\right)\frac{1}{\ln 2}\\ &= \frac{1}{\ln 2}\ln\Big(\frac{b+1}{a+1}\Big) = \mu\big((a,b)\big).\end{aligned}$$
Δηλαδή, ο T διατηρεί το μέτρο μ. □

Η T δεν είναι της μορφής $T\colon X \to X$ για τον ίδιο χώρο X. Ορίζουμε $X = [0,1] \smallsetminus \mathbb{Q}$. Έχουμε ότι $T(x) \in \mathbb{Q}$ ανν $1/x - \lfloor 1/x \rfloor \in \mathbb{Q}$, και αφού ο $\lfloor 1/x \rfloor$ είναι ακέραιος, αυτό είναι ισοδύναμο με το $1/x \in \mathbb{Q} \Leftrightarrow x \in \mathbb{Q}$. Άρα $T(X) \subseteq X$ και ο περιορισμός της T στον X, τον οποίο συμβολίζουμε επίσης με T, είναι της μορφής $T\colon X \to X$. Ανάλογα παίρνουμε για μ τον περιορισμό του παραπάνω μέτρου στα Borel υποσύνολα του X και $\mathcal{B}$ να είναι η Borel σ-άλγεβρα του X (που αποτελείται τελικά από όλα τα σύνολα της μορφής $B \cap X$ με $B \in \mathcal{B}(\mathbb{R})$). Ο παραπάνω υπολογισμός για το ότι η T διατηρεί το μ δεν αλλάζει με τον περιορισμό στο X αφού το Lebesgue μέτρο του $\mathbb{Q}$ είναι μηδέν και η εξαίρεση των ρητών δεν μεταβάλλει τα μ-μέτρα των συνόλων που θεωρούμε.

Η απεικόνιση Gauss σχετίζεται με τα συνεχή κλάσματα. Αν $x \in (0,1] \smallsetminus \mathbb{Q}$, τότε ο x γράφεται ως
$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}, \qquad \text{με} \quad a_i \in \mathbb{N}, \quad i \in \mathbb{N}.$$
Σημείωση: Αποδεικνύεται ότι για άρρητους συγκλίνει.

Τώρα, $T(x) = 1/x - \lfloor 1/x \rfloor$,
$$x = \frac{1}{T(x) + \left\lfloor \frac{1}{x} \right\rfloor} = \cfrac{1}{\left\lfloor \frac{1}{x} \right\rfloor + \cfrac{1}{\left\lfloor \frac{1}{T(x)} \right\rfloor + T^2(x)}} = \cdots.$$
Αποδεικνύεται ότι
$$a_n = a_n(x) = \left\lfloor \frac{1}{T^{n-1}(x)} \right\rfloor, \qquad T^0(x) = x, \quad n \in \mathbb{N}.$$
Με τον συμβολισμό $x = [a_1, a_2, \dots]$, $T(x) = [a_2, a_3, \dots]$.

Χώροι Shift: (a) *(Μονόπλευρο Bernoulli Shift)* Έστω S ένα πεπερασμένο σύνολο, έστω $S = \{1, \dots, s\}$, και έστω $p = (p_i)_{i\in S}$ ένα διάνυσμα πιθανότητας, δηλαδή $p_i \geq 0\ \forall\, i \in S$ και $\sum_{i\in S} p_i = 1$. Θεωρούμε
$$X = S^{\mathbb{N}} = \{\underset{\sim}{x} = (x_1, x_2, \dots)\colon x_i \in S\ \forall\, i \in \mathbb{N}\}$$

και την σ-άλγεβρα που παράγεται απ' όλους τους πεπερασμένους κυλίνδρους

$$\mathcal{B} = \sigma\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = s_1, \dots, x_n = s_n\} : n \in \mathbb{N},\ s_1, s_2, \dots, s_n \in S\big).$$

Η σ-άλγεβρα $\mathcal{B}$ είναι η σ-άλγεβρα γινόμενο $\mathcal{B} = \otimes_{\mathbb{N}} 2^S$. Ορίζοντας τις προβολές $\pi_i : X \to S$, $\pi_i(\underset{\sim}{x}) = x_i$, $i \in \mathbb{N}$, για $\underset{\sim}{x} = (x_1, x_2, \dots) \in X$, έχουμε ότι $\mathcal{B} = \sigma(\pi_1, \pi_2, \dots)$.

Πρόταση 1.1.16 Υπάρχει μοναδικό μέτρο πιθανότητας μ ορισμένο στην $\mathcal{B}$ για το οποίο

$$\mu\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = s_1, x_2 = s_2, \dots, x_n = s_n\}\big) = p_{s_1} p_{s_2} \cdots p_{s_n}.$$

(Για την απόδειξη, βλ. §2 και το Θεώρημα 3.1 του *Probability and Measure* του P. Billingsley, για παράδειγμα.)

Για παράδειγμα, αν $S = \{0, 1\}$, τότε $X = \{\underset{\sim}{x} = (x_1, x_2, \dots) : x_i \in \{0, 1\}\ \forall\, i \in \mathbb{N}\}$. Έστω $0 < p_0 < 1$ και $p_1 = 1 - p_0$. Τότε

$$\mu\big(\{\underset{\sim}{x} = (x_1, x_2, x_3, \dots) \in X : x_1 = 0, x_2 = 0, x_3 = 1\}\big) = p_0 p_0 p_1 = p_0^2 p_1.$$

$T : X \to X$ είναι η απεικόνιση shift, δηλαδή

$$T(x_1, x_2, \dots) = (x_2, x_3, \dots).$$

Το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι ένα σύστημα που διατηρεί το μέτρο. Πράγματι, εφόσον τα σύνολα

$$\{\underset{\sim}{x} \in X : x_1 = s_1, \dots, x_n = s_n\}, \quad s_1, \dots, s_n \in S,\ n \in \mathbb{N},$$

μαζί με το κενό, αποτελούν π-σύστημα και παράγουν την $\mathcal{B}$, αρκεί να ελεγχθεί ότι

$$\mu\big(T^{-1}\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = s_1, \dots, x_n = s_n\}\big) = \mu\big(\underset{\sim}{x} \in X : x_1 = s_1, \dots, x_n = s_n\}\big),$$

για κάθε $n \in \mathbb{N}$ και $s_1, \dots, s_n \in S$.

Τώρα,

$$\begin{aligned}
T^{-1}&\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = s_1, \dots, x_n = s_n\}\big) \\
&= \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_2 = s_1, x_3 = s_2, \dots, x_{n+1} = s_n\} \\
&= \bigcup_{i \in S} \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = i, x_2 = s_1, \dots, x_{n+1} = s_n\}.
\end{aligned}$$

Άρα,

$$\begin{aligned}
\mu\big(T^{-1}&\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = s_1, \dots, x_n = s_n\}\big)\big) \\
&= \sum_{i \in S} \mu\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X : x_1 = i, x_2 = s_1, \dots, x_{n+1} = s_n\}\big) \\
&= \sum_{i \in S} p_i p_{s_1} \dots p_{s_n} \\
&= p_{s_1} \dots p_{s_n} \\
&= \mu\big(\underset{\sim}{x} \in X : x_1 = s_1, \dots, x_n = s_n\}\big). \qquad \square
\end{aligned}$$

(b) *(Αμφίπλευρο Bernoulli Shift)* Έστω $S = \{1, 2, \dots, s\}$ πεπερασμένο σύνολο και $p = (p_i)_{i \in S}$ διάνυσμα πιθανότητας όπως πριν. Τώρα

$$X = S^{\mathbb{Z}} = \{\underset{\sim}{x} = (\dots, x_{-1}, x_0, x_1, \dots) : x_i \in S\ \forall\, i \in \mathbb{Z}\}$$

με σ-άλγεβρα

$$\begin{aligned}
\mathcal{B} = \big(2^S\big)^{\mathbb{Z}} &= \sigma\big(\{\underset{\sim}{x} = (\dots, x_{-1}, x_0, x_1, \dots) \in X : x_{-n} = s_{-n}, \dots, x_m = s_m\} : n, m \in \mathbb{N}_0, s_{-n} \dots, s_m \in S\big) \\
&= \sigma\big(\pi_i : i \in \mathbb{Z}\big),
\end{aligned}$$

όπου $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, και $\pi_i\colon X \to S$ είναι πάλι οι προβολές $\pi_i(\dots, x_{-1}, x_0, x_1, \dots) = x_i$, $i \in \mathbb{Z}$. Όπως πριν, υπάρχει μοναδικό μέτρο πιθανότητας μ τέτοιο ώστε

$$\mu\big(\{\underset{\sim}{x} = (\dots, x_{-1}, x_0, x_1, \dots) \in X\colon x_{-n} = s_{-n}, \dots, x_m = s_m\}\big) = p_{s_{-n}} \cdots p_{s_m},$$

για κάθε $n, m \in \mathbb{N}_0$ και $s_{-n}, \dots, s_m \in S$.

Θεωρούμε τον μετασχηματισμό shift

$$T(\dots, x_{-1}, \underset{\wedge}{x_0}, x_1, \dots) = (\dots, x_0, \underset{\wedge}{x_1}, x_2, \dots),$$

όπου το σύμβολο $\wedge$ δηλώνει την θέση 0 της αμφίπλευρης ακολουθίας. Ο T διατηρεί και πάλι το μέτρο γινόμενο μ.

Το (αμφίπλευρο) σύστημα $(X, \mathcal{B}, \mu, T)$ είναι ένα αντιστρέψιμο σύστημα που διατηρεί το μέτρο.

(c) *(Μονόπλευρο Markov Shift)* Έστω $S = \{1, 2, \dots, s\}$ πεπερασμένο σύνολο και

$$X = S^{\mathbb{N}_0} = \{\underset{\sim}{x} = (x_0, x_1, \dots)\colon x_n \in S \quad \forall n \in \mathbb{N}_0\},$$

όπου $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Η σ-άλγεβρα παράγεται απ' όλους τους πεπερασμένους κυλίνδρους

$$\mathcal{B} = \sigma\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}, \quad n \in \mathbb{N}_0, \quad s_0, s_1, \dots, s_n \in S\big),$$

και έχουμε τον shift μετασχηματισμό

$$(x_0, x_1, x_2, \dots) \xmapsto{T} (x_1, x_2, \dots).$$

Τώρα έχουμε τα δεδομένα:

$P = (P_{ij})_{i,j \in S}$ στοχαστικός πίνακας, δηλαδή $P_{ij} \geq 0$ για κάθε $i, j \in S$ και $\sum_{j \in S} P_{ij} = 1$ για κάθε $i \in S$. Επίσης, έστω $p = (p_i)_{i \in S}$, διάνυσμα πιθανότητας, που να είναι το αριστερό ιδιοδιάνυσμα του P που αντιστοιχεί στην ιδιοτιμή 1, δηλαδή ικανοποιεί την σχέση: $\sum_{i \in S} p_i P_{ij} = p_j$, για κάθε $j \in S$, $((p_1, \dots, p_s)P = (p_1, \dots, p_s))$.

Σημείωση. Ένα τέτοιο διάνυσμα $p = (p_i)_{i \in S}$ υπάρχει πάντα για έναν στοχαστικό πίνακα, λόγω του θεωρήματος Perron–Frobenius.

Πρόταση 1.1.17 Με ορισμένα όπως παραπάνω τα $S, X, \mathcal{B}, p, P$, υπάρχει μοναδικό μέτρο πιθανότητας στον $(X, \mathcal{B})$ για το οποίο ισχύει

$$\mu\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big) = p_{s_0} P_{s_0 s_1} P_{s_1 s_2} \cdots P_{s_{n-1} s_n}, \tag{3}$$

όπου p_{s_0} είναι η s_0-συντεταγμένη του $p = (p_i)_{i \in S}$ και $P_{s_i s_{i+1}}$ είναι η πιθανότητα να μεταβούμε από την κατάσταση s_i στην κατάσταση s_{i+1}.

Απόδειξη. Έστω

$$\mathcal{B}_n = \sigma\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}, s_0, s_1, \dots, s_n \in S\big)$$

και ορίζουμε το μέτρο μ_n από την (3):

$$\mu_n\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big) = p_{s_0} P_{s_0 s_1} P_{s_1 s_2} \cdots P_{s_{n-1} s_n}.$$

Έχουμε,

$$\begin{aligned}
\mu_{n+1}\big(\{\underset{\sim}{x} &= (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big) \\
&= \sum_{s \in S} \mu_{n+1}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n, x_{n+1} = s\}\big) \\
&= \sum_{s \in S} p_{s_0} P_{s_0 s_1} P_{s_1 s_2} \cdots P_{s_{n-1} s_n} P_{s_n s} \\
&= p_{s_0} P_{s_0 s_1} P_{s_1 s_2} \cdots P_{s_{n-1} s_n} \sum_{s \in S} P_{s_n s} \\
&= p_{s_0} P_{s_0 s_1} P_{s_1 s_2} \cdots P_{s_{n-1} s_n} \\
&= \mu_n\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X\colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big),
\end{aligned}$$

όπου η πρώτη ισότητα ισχύει εφόσον έχουμε ξένη ένωση και η τέταρτη επειδή έχουμε στοχαστικό πίνακα.

Έπεται ότι τα μέτρα μ_{n+1} και μ_n συμφωνούν επί της $\mathcal{B}_n$ και άρα, επαγωγικά, για κάθε $n < m$, τα μ_n και μ_m συμφωνούν επί της $\mathcal{B}_n$. Άρα, μπορούμε να ορίσουμε (καλά)[4] ένα μέτρο μ στην άλγεβρα $\bigcup_{n=0}^{\infty} \mathcal{B}_n$, το οποίο είναι πεπερασμένα αθροιστικό. Αποδεικνύεται ότι το μ είναι, τελικά, αριθμήσιμα αθροιστικό στην $\bigcup_{n=0}^{\infty} \mathcal{B}_n$ και άρα, επεκτείνεται μονοσήμαντα στην σ-άλγεβρα $\mathcal{B} = \sigma\left(\bigcup_{n\geq 0} \mathcal{B}_n\right)$. Η απόδειξη είναι εντελώς όμοια με αυτήν για τα Bernoulli shifts στην §2 του βιβλίου *Probability and Measure* του P. Billingsley. □

Το μέτρο μ είναι shift-αναλλοίωτο. Πράγματι, αρχικά παρατηρούμε ότι οι κύλινδροι

$$\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}, \quad s_0, s_1, \dots, s_n \in S, n \in \mathbb{N}_0,$$

μαζί με το κενό σύνολο, αποτελούν π-σύστημα (που εξ' ορισμού παράγει την σ-άλγεβρα $\mathcal{B}$), και για αυτούς τους κυλίνδρους έχουμε ότι,

$$\begin{aligned}
\mu\big(T^{-1}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big)\big)& \\
&= \mu\big(\{\underset{\sim}{x} \in X \colon x_1 = s_0, x_2 = s_1, \dots, x_{n+1} = s_n\}\big)\\
&= \sum_{s\in S} \mu\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = s, x_1 = s_0, \dots, x_{n+1} = s_n\}\big)\\
&= \sum_{s\in S} p_s P_{ss_0} P_{s_0 s_1} \cdots P_{s_{n-1}s_n}\\
&= p_{s_0} P_{s_0 s_1} \cdots P_{s_{n-1}s_n}\\
&= \mu\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = s_0, x_1 = s_1, \dots, x_n = s_n\}\big),
\end{aligned}$$

όπου η προ-τελευταία ισότητα ισχύει επειδή το διάνυσμα p είναι αριστερό ιδιοδιάνυσμα του πίνακα P που αντιστοιχεί στην ιδιοτιμή 1. Τονίζεται ότι αυτό που κάνει το μέτρο μ shift-αναλλοίωτο είναι η συγκεκριμένη επιλογή αρχικής κατανομής p. Κάθε αρχική κατανομή $p = (p_i)_{i\in S}$ ορίζει, μέσω της (3), ένα μέτρο μ στον χώρο ακολουθιών $(X, \mathcal{B})$. Μόνο η συγκεκριμένη επιλογή αρχικής κατανομής όμως, δηλαδή με το διάνυσμα p να είναι αριστερό ιδιοδιάνυσμα για τον πίνακα P για την ιδιοτιμή 1, δίνει shift-αναλλοίωτο μέτρο μ.

Το σύστημα $(X, \mathcal{B}, \mu, T)$ λέγεται (μονόπλευρο) Markov shift με πίνακα μετάβασης P και αρχική κατανομή p. Ανάλογα ορίζεται και το αμφίπλευρο Markov shift. □

1.2 Θεώρημα Επαναφοράς του Poincaré

Θεώρημα 1.2.1 Έστω $(X, \mathcal{B}, \mu, T)$ ένα σύστημα που διατηρεί το μέτρο και $A \in \mathcal{B}$ με $\mu(A) > 0$. Τότε υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε

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

Απόδειξη. Θεωρούμε τα $A, T^{-1}(A), T^{-2}(A), \dots$. Έστω ότι $\mu\big(T^{-n}(A) \cap T^{-m}(A)\big) = 0$ για κάθε $n, m \in \mathbb{N}_0$ με $n \neq m$. Τότε

$$\mu\left(\bigcup_{n\geq 0} T^{-n}(A)\right) = \sum_{n\geq 0} \mu\big(T^{-n}(A)\big) = \sum_{n\geq 0} \mu(A),$$

που είναι άπειρο αφού $\mu(A) > 0$, άρα έχουμε άτοπο (είμαστε σε χώρο πεπερασμένου μέτρου). Έπεται ότι υπάρχουν $m, n \in \mathbb{N}_0$, $m \neq n$, τέτοια ώστε

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

Χωρίς βλάβη της γενικότητας θεωρούμε ότι $m > n$. Τότε,

$$\mu\big(T^{-n}(A) \cap T^{-m}(A)\big)\mu\big(T^{-n}\big(A \cap T^{-(m-n)}(A)\big)\big) = \mu\big(A \cap T^{-(m-n)}(A)\big),$$

[4] Ικανοποιούνται οι συνθήκες στο Θεώρημα Ύπαρξης/Συνέπειας του Kolmogorov (Kolmogorov existence theorem).

εφόσον το μ είναι T-αναλλοίωτο. Άρα

$$\mu\big(A \cap T^{-(m-n)}(A)\big) > 0, \quad \text{με} \quad m-n \in \mathbb{N}. \qquad \square$$

Παρατήρηση. Το ίδιο επιχείρημα δείχνει ότι ισχύει το συμπέρασμα $\mu\big(A \cap T^{-n}(A)\big) > 0$ για κάποιο $n \leq \lfloor 1/\mu(A) \rfloor$. Πράγματι, αν $\mu\big(T^{-n}(A) \cap T^{-m}(A)\big) = 0$ για κάθε $n, m \in \{1, \ldots, N\}$ με $n \neq m$, όπου $N := \lfloor 1/\mu(A) \rfloor$, τότε

$$\mu\left(\bigcup_{n=0}^{N} T^{-n}(A)\right) = \sum_{n=0}^{N} \mu\big(T^{-n}(A)\big) = \sum_{n=0}^{N} \mu(A) = (N+1)\mu(A) > 1,$$

άτοπο.

Θεώρημα 1.2.2 (Θεώρημα επαναφοράς του Poincaré) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $A \in \mathcal{B}$ με $\mu(A) > 0$. Τότε για σχεδόν κάθε $x \in A$, υπάρχει ακολουθία $n_1 < n_2 < \cdots$ τέτοια ώστε $T^{n_k}(x) \in A$ για κάθε k.

Παρατηρήσεις.

- Το θεώρημα επαναφοράς του Poincaré δεν ισχύει αν το μ είναι άπειρο μέτρο. Για παράδειγμα, έστω $X = \mathbb{Z}, \mathcal{B} = 2^{\mathbb{Z}}$, μ το μέτρο με $\mu(A) = |A|$ (πληθάριθμος του A) και $T(x) = x+1$. Είναι σαφές ότι ο T διατηρεί το μέτρο μ, αλλά αν $A = \{0\}$, τότε το A έχει θετικό μέτρο αλλά $\mu\big(A \cap T^{-n}(A)\big) = 0$ $\forall n \in \mathbb{N}$, αφού $A \cap T^{-n}(A) = \varnothing$ $\forall n \in \mathbb{N}$.

- Το θεώρημα επαναφοράς του Poincaré δεν ισχύει αν ο T δεν διατηρεί το μέτρο. Για παράδειγμα, έστω $X = [0,1]$, μ το μέτρο Lebesgue και $T(x) = x^2$. Εδώ ο T δεν διατηρεί το μέτρο Lebesgue και $T^n(x) \to 0$, καθώς $n \to \infty$, για κάθε $x \neq 1$. Άρα, ένα σύνολο θετικού μέτρου που δεν περιέχει τα $0, 1$ "δεν θα επανέλθει ποτέ". Ακριβέστερα, αν $a, b \in (0,1)$ με $a < b < \sqrt{a}$, τότε $T^{-1}\big((a,b)\big) = \big(\sqrt{a}, \sqrt{b}\big)$ και άρα $(a,b) \cap T^{-1}\big((a,b)\big) = \varnothing$. Επιπλέον $T^{-n}\big(a,b)\big) = \big(a^{2^{-n}}, b^{2^{-n}}\big)$, $n \in \mathbb{N}_0$, και τα διαστήματα αυτά είναι όλα ξένα αφού

$$0 < a < b < \sqrt{a} < \sqrt{b} < \cdots < a^{2^{-n}} < b^{2^{-n}} < a^{2^{-(n+1)}} < b^{2^{-(n+1)}} < \cdots$$

  (όπως μπορεί κανείς να δει επαγωγικά). Να σημειωθεί ωστόσο ότι δεν χρειάζεται κανείς διατήρηση του μέτρου ακριβώς για το θεώρημα επαναφοράς του Poincaré αλλά κάτι ασθενέστερο, και συγκεκριμένα μόνο ότι το σύστημα είναι ασυμπίεστο: η σχέση $\mu\big(T^{-1}(B)\big) = \mu(B)$ δεν ισχύει κατ' ανάγκη για όλα τα $B \in \mathcal{B}$, αλλά μόνο για εκείνα τα $B \in \mathcal{B}$ για τα οποία $T^{-1}(B) \subseteq B$.

Απόδειξη Θεωρήματος 1.2.2. Έστω $A \in \mathcal{B}$ με $\mu(A) > 0$. Θεωρούμε το σύνολο

$$B = A \cap T^{-1}(X \smallsetminus A) \cap T^{-2}(X \smallsetminus A) \cap \cdots .$$

Δηλαδή το B περιέχει όλα τα σημεία που είναι στο A και δεν επιστρέφουν ποτέ στο A. Αν $j \neq k$ ισχύει ότι

$$T^{-j}(B) \cap T^{-k}(B) = \varnothing.$$

Πράγματι, έστω ότι $j < k$. Αν $x \in T^{-j}(B)$ τότε $T^j(x) \in B$, δηλαδή $T^j(x) \in A$ και $T^{j+n}(x) \notin A$ για κάθε $n \geq 1$. Αν για το ίδιο x είχαμε ότι $x \in T^{-k}(B)$, τότε θα έπρεπε $T^k(x) \in B$, άρα $T^k(x) \in A$, ενώ $k > j$, και αυτό θα ήταν άτοπο. Επομένως,

$$1 \geq \mu\left(\bigcup_{j \geq 0} T^{-j}B\right) = \sum_{j \geq 0} \mu\big(T^{-j}B\big) = \sum_{j \geq 0} \mu(B).$$

Δηλαδή, θα πρέπει $\mu(B) = 0$ (αλλιώς το τελευταίο άθροισμα θα ήταν άπειρο). Τώρα, αφού $\mu(A) = \mu(A \cap B) + \mu(A \cap B^{\mathrm{C}})$ και $\mu(A \cap B) = 0$, έχουμε

$$\mu(A) = \mu\left(A \cap \bigcup_{n \geq 1} T^{-n}(A)\right).$$

Το δεξί μέλος είναι το μέτρο του συνόλου που περιέχει τα σημεία του A που επανέρχονται στο A τουλάχιστον μία φορά. Το ονομάζουμε A_1 και έχουμε $\mu(A_1) = \mu(A)$. Επαναλαμβάνουμε τώρα το επιχείρημα με T^k στην θέση του T. Ορίζουμε

$$A_k = A \cap \bigcup_{n\geq 1} T^{-kn}(A).$$

Τότε $A_k \subset A$ και $\mu(A_k) = \mu(A)$. Ορίζουμε και

$$A_\infty = \bigcap_{k\geq 1} A_k.$$

Για κάθε $x \in A_\infty$ έχει κανείς ότι

$$\forall\, k \in \mathbb{N} \quad \exists\, n_k \in \mathbb{N}: \quad T^{kn_k}(x) \in A$$

ή ισοδύναμα, για κάθε $x \in A_\infty$

$$\forall\, k \in \mathbb{N} \quad \exists\, m_k \geq k: \quad T^{m_k}(x) \in A.$$

Δηλαδή, για $x \in A_\infty$, $x \in A$ και άπειρες εικόνες $T^n(x)$ του x ανήκουν στο A. Αρκεί λοιπόν να δειχτεί ότι $\mu(A_\infty) = \mu(A)$. Όμως,

$$\mu(A \setminus A_\infty) = \mu\left(\bigcup_{k\geq 1} \mu(A \setminus A_k)\right) \leq \sum_{k\geq 1} \mu(A \setminus A_k) = 0. \qquad \square$$

Εφαρμογή 1.2.3 Αν $a \in [0,1)$, τότε η ακολουθία na, $n \in \mathbb{N}$, έρχεται ε-κοντά σε ακέραιους για άπειρα το πλήθος $n \in \mathbb{N}$, για κάθε $\varepsilon > 0$.

Απόδειξη. Έστω $X = \mathbb{R}/\mathbb{Z}$, $T_a(x+\mathbb{Z}) = x + a + \mathbb{Z}$, μ το μέτρο Lebesgue. Έστω $\varepsilon > 0$ και έστω $A = (-\varepsilon/2, \varepsilon/2) + \mathbb{Z}$. Τότε $\mu(A) = \varepsilon > 0$. Έχουμε

$$T_a^n(x+\mathbb{Z}) = x + na + \mathbb{Z}.$$

Από το θεώρημα επαναφοράς του Poincaré έχουμε ότι για σχεδόν κάθε $x + \mathbb{Z} \in A$, υπάρχει ακολουθία $n_1 < n_2 < \cdots$ τέτοια ώστε

$$T_a^{n_i}(x+\mathbb{Z}) = x + n_i a + \mathbb{Z} \in A,$$

για κάθε $i \in \mathbb{N}$. Σταθεροποιούμε ένα τέτοιο $x + \mathbb{Z} \in A$. Παρατηρείστε ότι

$$A = \{y + \mathbb{Z} \in \mathbb{R}/\mathbb{Z} \colon d_{\mathbb{R}/\mathbb{Z}}(y+\mathbb{Z}, \mathbb{Z}) < \varepsilon/2\},$$

όπου υπενθυμίζεται ότι η μετρική $d_{\mathbb{R}/\mathbb{Z}}$ δίνεται από την σχέση

$$d_{\mathbb{R}/\mathbb{Z}}(t+\mathbb{Z}, s+\mathbb{Z}) = \min_{m\in\mathbb{Z}} |t - s + m|$$

για $t, s \in \mathbb{R}$. Άρα έχουμε ότι για άπειρους δείκτες n_i,

$$d_{\mathbb{R}/\mathbb{Z}}(x + n_i a + \mathbb{Z}, \mathbb{Z}) < \varepsilon/2.$$

Εφόσον για $x + \mathbb{Z} \in A$ ισχύει ότι

$$d_{\mathbb{R}/\mathbb{Z}}(x + n_i a + \mathbb{Z}, n_i a + \mathbb{Z}) = \min_{m\in\mathbb{Z}} |x + n_i a - n_i a + m| = \min_{m\in\mathbb{Z}} |x + m| < \varepsilon/2,$$

από τριγωνική ανισότητα παίρνουμε ότι

$$d_{\mathbb{R}/\mathbb{Z}}(n_i a + \mathbb{Z}, \mathbb{Z}) < \varepsilon,$$

για κάθε $i \in \mathbb{N}$. Δηλαδή, η ακολουθία na έρχεται ε-κοντά σε ακέραιο, για κάθε $n = n_i$, $i \in \mathbb{N}$, δηλαδή για άπειρα $n \in \mathbb{N}$. $\square$

1.3 Στρεβλά γινόμενα (Skew Products)

Ορισμός 1.3.1 (Skew product) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $(Y, \mathcal{A}, \nu)$ χώρος πιθανότητας. Έστω επίσης $S\colon X \times Y \to Y$ μετρήσιμη και τέτοια ώστε $(S_x)_*\nu = \nu$ για κάθε $x \in X$, όπου για κάθε $x \in X$,

$$S_x\colon Y \to Y, \qquad S_x(y) = S(x, y).$$

Ορίζουμε την απεικόνιση (skew product) τ στον $X \times Y$ με

$$\tau(x, y) = \big(T(x), S(x, y)\big).$$

Πρόταση 1.3.2 Το $(X \times Y, \mathcal{B} \otimes \mathcal{A}, \mu \otimes \nu, \tau)$ είναι σύστημα που διατηρεί το μέτρο (σ.δ.μ.).

Απόδειξη. Άσκηση. □

Συγκεκριμένη περίπτωση. Έστω $X = Y = \mathbb{R}/\mathbb{Z}$, ή $X = Y = [0, 1)$, με $\mu = \nu$ το μέτρο Lebesgue στο $[0, 1)$. Επίσης,

$$T_a(x) = x + a \ (\mathrm{mod}\ 1) \qquad \text{ή} \qquad T_a(x + \mathbb{Z}) = x + a + \mathbb{Z},$$

και

$$S(x, y) = x + y \ (\mathrm{mod}\ 1) \qquad \text{ή} \qquad S\big((x, y) + \mathbb{Z}^2\big) = x + y + \mathbb{Z}.$$

Τότε

$$\tau_a(x, y) = (x + a, x + y) \ (\mathrm{mod}\ 1) \qquad \text{ή} \qquad \tau_a\big((x, y) + \mathbb{Z}^2\big) = (x + a, x + y) + \mathbb{Z}^2,$$

και

$$\tau_a^{n+1}\big((x, y) + \mathbb{Z}^2\big) = (x_n + a, x_n + y_n) + \mathbb{Z}^2,$$

όπου

$$\tau_a^n\big((x, y) + \mathbb{Z}^2\big) = (x_n, y_n) + \mathbb{Z}^2,$$

$n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $x_0 = x, y_0 = y$. Δηλαδή

$$x_{n+1} + \mathbb{Z} = x_n + a + \mathbb{Z} = \cdots = x + (n+1)a + \mathbb{Z} \qquad \text{και} \qquad y_{n+1} + \mathbb{Z} = y_n + x + na + \mathbb{Z},$$

οπότε

$$\sum_{k=0}^{n-1}(y_{k+1} - y_k) = nx + \frac{n(n-1)a}{2} \quad \Rightarrow \quad y_n = y_0 + nx + \frac{n(n-1)a}{2}$$
$$= y_0 + \Big(x - \frac{a}{2}\Big)n + \frac{n^2}{2}a,$$

και άρα μπορούμε να γράψουμε

$$\tau_a^n\big((x, y) + \mathbb{Z}^2\big) = \left(x + na, y + \Big(x - \frac{a}{2}\Big)n + \frac{n^2}{2}a\right) + \mathbb{Z}^2,$$

για κάθε $n \in \mathbb{N}_0$.

Δοθέντος $\alpha \in [0, 1)$ επιλέγουμε $a = 2\alpha$ στα παραπάνω και παίρνουμε ότι

$$\tau_{2\alpha}\big((x, y) + \mathbb{Z}^2\big) = \big(x + 2n\alpha, y + (x - \alpha)n + n^2\alpha\big) + \mathbb{Z}^2,$$

για κάθε $n \in \mathbb{N}_0$. Εφαρμόζουμε το θεώρημα του Poincaré για δοθέν $\varepsilon > 0$ και το σύνολο $A = \mathbb{R}/\mathbb{Z} \times \big((-\varepsilon/2, \varepsilon/2) + \mathbb{Z}\big)$, που ταυτίζεται με το σύνολο

$$A = \left\{(x, y) + \mathbb{Z}^2 \in \mathbb{R}^2/\mathbb{Z}^2 \colon d_{\mathbb{R}/\mathbb{Z}}(y + \mathbb{Z}, \mathbb{Z}) < \varepsilon/2\right\},$$

και παίρνουμε ότι για σχεδόν κάθε $(x, y) + \mathbb{Z}^2 \in A$, υπάρχουν δείκτες $n_1 < n_2 < \cdots$ τέτοιοι ώστε $\tau_{2\alpha}^{n_k}\big((x, y) + \mathbb{Z}^2\big) \in A$ για κάθε $k \in \mathbb{N}$, δηλαδή τέτοιοι ώστε

$$d_{\mathbb{R}/\mathbb{Z}}\big(y + (x - \alpha)n_k + n_k^2\alpha + \mathbb{Z}, \mathbb{Z}\big) < \varepsilon/2 \tag{4}$$

για κάθε $k \in \mathbb{N}$. (Σημειώνεται εδώ ότι οι δείκτες $n_1, n_2, \ldots$ εξαρτώνται, φυσικά, από το συγκεκριμένο ζεύγος $(x,y) \in \mathbb{T}^2$ ή, αλλιώς, το σύμπλοκο $(x,y)+\mathbb{Z}^2$.) Έστω τώρα

$$B := \big\{(x,y)+\mathbb{Z}^2 \in A\colon \tau_{2\alpha}^n\big((x,y)+\mathbb{Z}^2\big) \in A \text{ για άπειρα } n \in \mathbb{N}\big\} = A \cap \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} \tau_{2\alpha}^{-n}(A),$$

που, όπως ήδη αναφέρθηκε, έχει από το θεώρημα Poincaré δι-διάστατο μέτρο Lebesgue ε. Από το θεώρημα Fubini, πρέπει τότε για κάποιο $y \in (-\varepsilon/2, \varepsilon/2)$ (ή κάποιο $y+\mathbb{Z} \in (-\varepsilon/2,\varepsilon/2)+\mathbb{Z}$), το σύνολο των $x \in [0,1)$ (ή $x+\mathbb{Z} \in \mathbb{R}/\mathbb{Z}$) για τα οποία $(x,y)+\mathbb{Z}^2 \in B$ να έχει μέτρο ένα. Σταθεροποιούμε ένα τέτοιο y (ή $y+\mathbb{Z}$) και τότε έχουμε ότι για σχεδόν κάθε $x \in [0,1)$, υπάρχουν δείκτες $n_1 < n_2 < \cdots$ τέτοιοι ώστε

$$d_{\mathbb{R}/\mathbb{Z}}\big((x-\alpha)n_k + n_k^2\alpha + \mathbb{Z}, \mathbb{Z}\big) < \varepsilon$$

για κάθε $k \in \mathbb{N}$, έχοντας χρησιμοποιήσει την (4) και την

$$d_{\mathbb{R}/\mathbb{Z}}\big(y + (x-\alpha)n_k + n_k^2\alpha + \mathbb{Z}, (x-\alpha)n_k + n_k^2\alpha + \mathbb{Z}\big) = \min_{m\in\mathbb{Z}} |y+m| < \varepsilon/2$$

για κάθε $k \in \mathbb{N}$, αφού $y+\mathbb{Z} \in (-\varepsilon/2,\varepsilon/2)+\mathbb{Z}$.

Αυτό μπορεί να διατυπωθεί ισοδύναμα ως εξής. Δοθέντος $\alpha \in \mathbb{R}$, για σχεδόν κάθε $\beta \in \mathbb{R}$ η ακολουθία $\beta n + n^2\alpha$ έρχεται ε-κοντά σε ακέραιο για άπειρα $n \in \mathbb{N}$, για κάθε $\varepsilon > 0$. $\square$

2 Εργοδικότητα

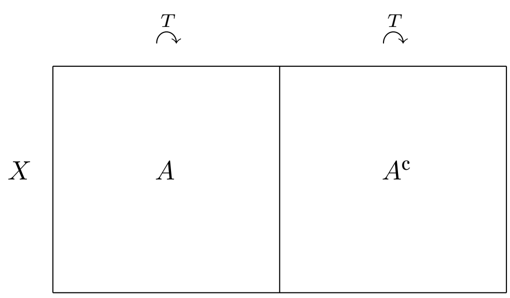
2.1 Ορισμός και παραδείγματα

Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. Ο χώρος X υποτίθεται πως είναι ο χώρος καταστάσεων ενός φυσικού συστήματος και η απεικόνιση T παριστάνει την εξέλιξη του συστήματος με το χρόνο: αν το σύστημα ξεκινήσει από μία αρχική κατάσταση $x \in X$ την χρονική στιγμή $n = 0$, και το παρατηρούμε σε διακριτό χρόνο $n = 0, 1, \ldots$, τότε $T(x)$ είναι η κατάσταση του συστήματος την χρονική στιγμή $n = 1$, $T^2(x) = T(T(x))$ την χρονική στιγμή $n = 2$, κ.ο.κ. Το ότι ο T διατηρεί το μέτρο μ, δηλαδή η $T_*\mu = \mu$, εκφράζει το γεγονός ότι οι νόμοι που διέπουν την εξέλιξη του συστήματος δεν αλλάζουν με τον χρόνο. Δηλαδή, η πιθανότητα το σύστημα να βρίσκεται σε μια κατάσταση $x \in A$ την χρονική στιγμή $n = 0$ είναι ίδια και για τις χρονικές στιγμές $n = 1, 2, \ldots$. Μία συνάρτηση $f\colon X \to \mathbb{R}$, ή $f\colon X \to \mathbb{C}$, παριστάνει ένα παρατηρήσιμο μέγεθος του συστήματος (μία μέτρηση ας πούμε, π.χ. $f(x)$ μπορεί να είναι η πίεση του συστήματος).

Εργοδική Υπόθεση (Boltzmann):

$$n^{-1}\sum_{i=0}^{n-1} f(T^i(x)) \xrightarrow{n\to\infty} \int_X f\,d\mu \qquad \text{για } \mu\text{-σχεδόν κάθε } x.$$

$$\Leftrightarrow \qquad \text{χρονικοί μέσοι όροι} = \text{χωρικοί μέσοι όροι.}$$



Παρατηρούμε τώρα ότι αν υπάρχει σύνολο $A \subset X$ τέτοιο ώστε $T^{-1}(A) = A$ με $\mu(A) \in (0,1)$, τότε η εργοδική υπόθεση δεν μπορεί να ισχύει. Για παράδειγμα, αν πάρουμε f να είναι η χαρακτηριστική συνάρτηση του συμπληρώματος A^c του A, έχουμε ότι $\int f\,d\mu > 0$ αφού $\mu(A^c) > 0$. Όμως

$$x \in A \quad \Rightarrow \quad T^n(x) \in A \quad \forall\, n \in \mathbb{N}_0 \quad \Rightarrow \quad f\big(T^n(x)\big) = 0 \quad \forall\, n \in \mathbb{N}_0,$$

άρα

$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(T^k(x)\big) = 0 \quad \forall\, n\in\mathbb{N} \quad \forall\, x\in A.$$

Επομένως

$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(T^k(x)\big) = 0 \nrightarrow \int_X f\, d\mu > 0$$

για $x\in A$, και $\mu(A)>0$.

Ορισμός 2.1.1 (αναλλοίωτο σύνολο) Έστω $(X,\mathcal{B},\mu,T)$ σ.δ.μ. Ένα $A\in\mathcal{B}$ λέγεται *αναλλοίωτο* αν $T^{-1}(A)=A$ και λέγεται *μ-σχεδόν αναλλοίωτο* (ή *αναλλοίωτο* mod μ) αν $\mu\big(A\triangle T^{-1}(A)\big)=0$. [5]

Ορισμός 2.1.2 (αναλλοίωτη συνάρτηση) Μια μετρήσιμη συνάρτηση $f\colon X\to\mathbb{C}$ λέγεται *αναλλοίωτη* αν $f=f\circ T$ και λέγεται *μ-σχεδόν αναλλοίωτη* (ή *αναλλοίωτη* mod μ) αν $f=f\circ T$ μ-σχεδόν παντού.

Παρατηρήσεις.

(i) $T^{-1}\big(A\triangle B\big)=T^{-1}(A)\triangle T^{-1}(B)$.
Πράγματι,

$$\begin{aligned} T^{-1}\big(A\triangle B\big) &= T^{-1}\big((A\cap B^{\mathrm c})\cup(A^{\mathrm c}\cap B)\big) = \big(T^{-1}(A)\cap T^{-1}(B)^{\mathrm c}\big)\cup\big(T^{-1}(A)^{\mathrm c}\cap T^{-1}(B)\big)\\ &= T^{-1}(A)\triangle T^{-1}(B). \end{aligned}$$

(ii) $|\mu(A)-\mu(B)|\le\mu(A\triangle B)$.
Πράγματι,

$$\begin{aligned} |\mu(A)-\mu(B)| &= \max\{\mu(A),\mu(B)\}-\min\{\mu(A),\mu(B)\}\le\mu(A\cup B)-\mu(A\cap B)\\ &= \mu\big((A\cup B)\setminus(A\cap B)\big)=\mu\big(A\triangle B\big). \end{aligned}$$

Ορισμός 2.1.3 (εργοδικό σύστημα) Ένα σ.δ.μ. $(X,\mathcal{B},\mu,T)$ λέγεται *εργοδικό* αν δεν έχει μη-τετριμμένα αναλλοίωτα υποσύνολα, δηλαδή αν $A\in\mathcal{B}$ και $A=T^{-1}(A)\Rightarrow\mu(A)\in\{0,1\}$.

Πρόταση 2.1.4 Έστω $(X,\mathcal{B},\mu,T)$ σ.δ.μ. Τότε τα εξής είναι ισοδύναμα.

(1) Το σύστημα είναι εργοδικό.

(2) Αν $A\in\mathcal{B}$ τέτοιο ώστε $\mu\big(A\triangle T^{-1}(A)\big)=0$ (δηλαδή αν A είναι μ-σχεδόν αναλλοίωτο), τότε $\mu(A)\in\{0,1\}$.

(3) Για κάθε $A\in\mathcal{B}$ με $\mu(A)>0$, έχουμε $\mu\left(\bigcup_{n\ge1}T^{-n}(A)\right)=1$.

(4) Για κάθε $A,B\in\mathcal{B}$ με $\mu(A)\mu(B)>0$, υπάρχει $n\in\mathbb{N}$ τέτοιο ώστε $\mu\big(A\cap T^{-n}(B)\big)>0$.

Απόδειξη. (1) $\Rightarrow$ (2). Έστω $A\in\mathcal{B}$ με $\mu\big(A\triangle T^{-1}(A)\big)=0$. Θα ορίσουμε ένα σύνολο A_∞ τέτοιο ώστε $\mu(A\triangle A_\infty)=0$ και $A_\infty=T^{-1}(A_\infty)$. Ορίζουμε

$$\begin{aligned} A_m &= \bigcup_{n\ge m}T^{-n}(A), \qquad m\in\mathbb{N}_0=\mathbb{N}\cup\{0\}\\ A_\infty &= \bigcap_{m\ge0}A_m=\limsup_{n\to\infty}T^{-n}(A) \end{aligned}$$

Ισχυρισμός 1: $A\triangle T^{-n}(A)\subset\bigcup_{k=0}^{n-1}\big(T^{-k}(A)\triangle T^{-(k+1)}(A)\big)$.
Πράγματι, αν $x\in A\triangle T^{-n}(A)$, τότε

$$\text{είτε}\quad (a)\quad x\in A\cap T^{-n}(A)^{\mathrm c}\qquad\text{είτε}\quad (b)\quad x\in A^{\mathrm c}\cap T^{-n}(A).$$

[5]Υπενθύμιση: $A\triangle B=(A\setminus B)\cup(B\setminus A)=(A\cup B)\setminus(A\cap B)=(A\cap B^{\mathrm c})\cup(A^{\mathrm c}\cap B)$

Στην περίπτωση (a) ορίζουμε k να είναι ο πρώτος θετικός ακέραιος για τον οποίο $T^k(x) \notin A$. Για τον k ισχύει ότι
$$T^k(x) \notin A \quad \text{και} \quad T^{k-1}(x) \in A.$$
Ξέρουμε ότι $T^0(x) \in A$ και $T^n(x) \notin A$, άρα τέτοιο k υπάρχει και ικανοποιεί την $1 \leq k \leq n$. Τώρα,
$$x \in T^{-(k-1)}(A) \bigtriangleup T^{-k}(A) \quad \Rightarrow \quad x \in \bigcup_{k=1}^{n} \Big(T^{-(k-1)}(A) \bigtriangleup T^{-k}(A)\Big).$$
Στην περίπτωση (b) ορίζουμε k να είναι η πρώτη φορά που $T^k(x) \in A$, δηλαδή k είναι ο μικρότερος θετικός ακέραιος για τον οποίο
$$T^{k-1}(x) \notin A \quad \text{και} \quad T^k(x) \in A.$$
Στην περίπτωση αυτή $T^0(x) \notin A$ και $T^n(x) \in A$, άρα τέτοιο k υπάρχει πάλι και ικανοποιεί την $1 \leq k \leq n$. Τότε, πάλι,
$$x \in T^{-k}(A) \cap T^{-(k-1)}(A)^{\mathrm{c}} \subset T^{-k}(A) \bigtriangleup T^{-(k-1)}(A) \subset \bigcup_{j=1}^{n} \Big(T^{-j}(A) \bigtriangleup T^{-(j-1)}(A)\Big).$$
Από τον Ισχυρισμό 1 έπεται ότι
$$\begin{aligned}
\mu\big(A \bigtriangleup T^{-n}(A)\big) &\leq \sum_{k=0}^{n-1} \mu\big(T^{-k}(A) \bigtriangleup T^{-(k+1)}(A)\big) \\
&= \sum_{k=0}^{n-1} \mu\big(T^{-k}\big(A \bigtriangleup T^{-1}(A)\big)\big) \\
&= n\mu\big(A \bigtriangleup T^{-1}(A)\big) = 0.
\end{aligned}$$
Επομένως,
$$\mu\big(A \bigtriangleup T^{-n}(A)\big) = 0.$$
Ισχυρισμός 2: $A \bigtriangleup A_m \subset \bigcup_{n \geq m} (A \bigtriangleup T^{-n}(A))$.
Πράγματι,
$$A \bigtriangleup A_m = A \bigtriangleup \bigcup_{n \geq m} T^{-n}(A).$$
Άρα, αν $x \in A \bigtriangleup A_m$, τότε
$$\text{είτε} \quad (a) \quad x \in A \quad \text{και} \quad x \in \left(\bigcup_{n \geq m} T^{-n}(A)\right)^{\mathrm{c}}, \qquad \text{είτε} \quad (b) \quad x \notin A \quad \text{και} \quad x \in \bigcup_{n \geq m} T^{-n}(A).$$
Στην (a) περίπτωση, $x \in A$ και για κάθε $n \geq m$, $x \in T^{-n}(A)^{\mathrm{c}}$. Άρα
$$x \in A \cap T^{-n}(A)^{\mathrm{c}} \quad \forall n \geq m \quad \Rightarrow \quad x \in A \bigtriangleup T^{-n}(A) \quad \forall n \geq m \quad \Rightarrow \quad x \in \bigcup_{n \geq m} \left(A \bigtriangleup T^{-n}(A)\right).$$
Στην (b) περίπτωση, $x \in A^{\mathrm{c}}$ και $x \in \bigcup_{n \geq m} T^{-n}(A)$. Άρα,
$$\exists\, n \geq m\colon \quad x \in A^{\mathrm{c}} \cap T^{-n}(A) \quad \Rightarrow \quad \exists\, n \geq m\colon \quad x \in A \bigtriangleup T^{-n}(A) \quad \Rightarrow \quad x \in \bigcup_{n \geq m} \left(A \bigtriangleup T^{-n}(A)\right).$$
Λόγω του Ισχυρισμού 2 έπεται ότι, για κάθε $m \in \mathbb{N}_0$,
$$\mu\big(A \bigtriangleup A_m\big) \leq \sum_{n \geq m} \mu\big(A \bigtriangleup T^{-n}(A)\big) = 0.$$
Όμως, $A_\infty = \bigcap_{m \geq 0} A_m$ και $A_m \downarrow A_\infty$. Άρα $A_m \cap A^{\mathrm{c}} \downarrow A_\infty \cap A^{\mathrm{c}}$ και $A_m^{\mathrm{c}} \cap A \uparrow A_\infty^{\mathrm{c}} \cap A$. Δηλαδή,
$$\mu\big(A_m \cap A^{\mathrm{c}}\big) \downarrow \mu\big(A_\infty \cap A^{\mathrm{c}}\big) \quad \text{και} \quad \mu\big(A_m^{\mathrm{c}} \cap A\big) \uparrow \mu\big(A_\infty^{\mathrm{c}} \cap A\big),$$

οπότε

$$\mu(A_m \triangle A) = \mu\big(A_m \cap A^{\mathrm{c}}\big) + \mu\big(A_m^{\mathrm{c}} \cap A\big) \longrightarrow \mu\big(A_\infty \cap A^{\mathrm{c}}\big) + \mu\big(A_\infty^{\mathrm{c}} \cap A\big) = \mu\big(A_\infty \triangle A\big).$$

Επομένως,

$$\mu\big(A_\infty \triangle A\big) = 0.$$

Τέλος, ισχύει ότι $T^{-1}(A_\infty) = A_\infty$. Πράγματι, $A_m = \bigcup_{n \geq m} T^{-n}(A)$ και είναι σαφές ότι $T^{-1}(A_m) = A_{m+1}$. Άρα

$$\begin{aligned} T^{-1}(A_\infty) = T^{-1}\left(\bigcap_{m \geq 0} A_m\right) &= \bigcap_{m \geq 1} A_m \\ &= \bigcap_{m \geq 0} A_m = A_\infty \end{aligned}$$

επειδή $\left(A_m\right)_m$ είναι φθίνουσα ακολουθία. Από υπόθεση (εργοδικότητα), έπεται ότι $\mu(A_\infty) \in \{0, 1\}$ και αφού $\mu\big(A_\infty \triangle A\big) = 0$, πρέπει $\mu(A) \in \{0, 1\}$.

(2) ⇒ (3). Έστω $A \in \mathcal{B}$ με $\mu(A) > 0$ και έστω $B = \bigcup_{n \geq 1} T^{-n}(A)$. Τότε

$$T^{-1}(B) = \bigcup_{n \geq 1} T^{-(n+1)}(A) = \bigcup_{n \geq 2} T^{-n}(A) \subset B.$$

Άρα

$$\mu\big(B \triangle T^{-1}(B)\big) = \mu\big(B \setminus T^{-1}(B)\big) = \mu(B) - \mu(T^{-1}(B)) = 0$$

εφόσον το μέτρο είναι αναλλοίωτο. Άρα το B είναι μ-σχεδόν αναλλοίωτο, επομένως, λόγω του (2), $\mu(B) \in \{0, 1\}$. Όμως $\mu(B) \geq \mu\big(T^{-1}(A)\big) = \mu(A) > 0$, άρα πρέπει $\mu(B) = 1$.

(3) ⇒ (4). Έστω $A, B \in \mathcal{B}$ με $\mu(A)\mu(B) > 0$. Τότε, από το (3), έχουμε ότι $\mu\left(\bigcup_{n \geq 1} T^{-n}(B)\right) = 1$, άρα

$$\mu\left(A \cap \bigcup_{n \geq 1} T^{-n}(B)\right) = \mu(A).$$

Επομένως,

$$0 < \mu(A) = \mu\left(A \cap \bigcup_{n \geq 1} T^{-n}(B)\right) = \mu\left(\bigcup_{n \geq 1} \big(A \cap T^{-n}(B)\big)\right) \leq \sum_{n \geq 1} \mu\big(A \cap T^{-n}(B)\big).$$

Άρα υπάρχει $n \in \mathbb{N}$ έτσι ώστε $\mu\big(A \cap T^{-n}(B)\big) > 0$.

(4) ⇒ (1). Έστω $A \in \mathcal{B}$ τέτοιο ώστε $A = T^{-1}(A)$ και έστω ότι $\mu(A) \in (0, 1)$. Παρατηρούμε ότι

$$T^{-2}(A) = T^{-1}\big(T^{-1}(A)\big) = T^{-1}(A) = A,$$

και επαγωγικά, $A = T^{-n}(A)$ για κάθε $n \in \mathbb{N}$. Θέτουμε $B = A^{\mathrm{c}}$ και τότε $\mu(A)\mu(B) > 0$. Άρα, από το (4), υπάρχει $n \in \mathbb{N}$ έτσι ώστε $\mu\big(A \cap T^{-n}(B)\big) > 0$, και τότε

$$0 = \mu\big(A \cap A^{\mathrm{c}}\big) = \mu\big(T^{-n}(A \cap A^{\mathrm{c}})\big) = \mu\big(A \cap T^{-n}(A)^{\mathrm{c}}\big) = \mu\big(A \cap T^{-n}(A^{\mathrm{c}})\big) = \mu\big(A \cap T^{-n}(B)\big) > 0,$$

που είναι φυσικά άτοπο. □

Παρατηρήσεις. Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $A \in \mathcal{B}$.

(i) Αν $T^{-1}(A) \subset A$, τότε

$$\mu\big(A \triangle T^{-1}(A)\big) = \mu\big(A \setminus T^{-1}(A)\big) = \mu(A) - \mu\big(T^{-1}(A)\big) = 0,$$

και όμοια αν $A \subset T^{-1}(A)$, τότε πάλι $\mu\big(A \triangle T^{-1}(A)\big) = 0$. Επίσης, αν $T(A) \subset A$, τότε $A \subset T^{-1}(A)$ και άρα πάλι $\mu\big(A \triangle T^{-1}(A)\big) = 0$.

(ii) Το (3) της Πρότασης ισχυροποιείται ως εξής. Αν $(X,\mathcal{B},\mu,T)$ είναι εργοδικό σύστημα που διατηρεί το μέτρο και $A\in\mathcal{B}$ με $\mu(A)>0$, τότε

$$\mu\left(\bigcap_{m\geq 1}\bigcup_{n\geq m}T^{-n}(A)\right)=1.$$

Απόδειξη της Παρατήρησης (ii). Η Πρόταση δίνει ότι $\mu\left(\bigcup_{n\geq 1}T^{-n}(A)\right)=1$. Έχουμε

$$\mu\left(\bigcup_{n\geq m}T^{-n}(A)\right)=\mu\left(T^{-(m-1)}\bigcup_{n\geq 1}T^{-n}(A)\right)=\mu\left(\bigcup_{n\geq 1}T^{-n}(A)\right)=1.$$

Άρα

$$\mu\left(\bigcap_{m\geq 1}\bigcup_{n\geq m}T^{-n}(A)\right)=\lim_{n\to\infty}\mu\left(\bigcup_{n\geq m}T^{-n}(A)\right)=1$$

αφού $\bigcup_{n\geq m}T^{-n}(A)\downarrow\bigcap_{m\geq 1}\bigcup_{n\geq m}T^{-n}(A)$. □

Πρόταση 2.1.5 Έστω $(X,\mathcal{B},\mu,T)$ σ.δ.μ.

(1) Αν το σύστημα είναι εργοδικό, τότε κάθε μ-σχεδόν αναλλοίωτη μετρήσιμη συνάρτηση $f\colon X\to\mathbb{C}$ είναι σταθερή μ-σχεδόν παντού.

(2) Αν κάθε μ-σχεδόν αναλλοίωτη μετρήσιμη συνάρτηση $f\in L^\infty(X,\mathcal{B},\mu)$ είναι σταθερή σχεδόν παντού, τότε το σύστημα είναι εργοδικό.

Παρατήρηση. Το (2) ισχυροποιείται ως εξής. Αν, για οποιοδήποτε $p\in[1,+\infty]$, κάθε μ-σχεδόν αναλλοίωτη μετρήσιμη συνάρτηση $f\in L^p(X,\mathcal{B},\mu)$ είναι σταθερή σχεδόν παντού, τότε το σύστημα είναι εργοδικό. Αυτό συνάγεται άμεσα από την Πρόταση επειδή, αν $f\in L^\infty$ και είναι μ-σχεδόν αναλλοίωτη, τότε $f\in L^p$ (για κάθε p, διότι το μέτρο είναι μέτρο πιθανότητας) και άρα πρέπει η f να είναι σταθερή μ-σχεδόν παντού.

Απόδειξη. (1) Κατ' αρχήν αρκεί να δείξουμε την περίπτωση που $f\colon X\to\mathbb{R}$ (δηλαδή f πραγματική). Στην μιγαδική περίπτωση απλά διασπούμε την f σε φανταστικό και πραγματικό μέρος. Έστω λοιπόν $f\colon X\to\mathbb{R}$ τέτοια ώστε $f=f\circ T$, μ-σχεδόν παντού. Θεωρούμε $F_t=\{x\in X\colon f(x)\leq t\}=f^{-1}\big((-\infty,t]\big)$, για $t\in\mathbb{R}$. Τα F_t είναι μ-σχεδόν αναλλοίωτα, αφού, για κάθε $t\in\mathbb{R}$,

$$F_t\bigtriangleup T^{-1}(F_t)\subset\big\{x\in X\colon f(x)\neq f\big(T(x)\big)\big\}.$$

Το σύστημα είναι εργοδικό, άρα $\mu\big(F_t\big)\in\{0,1\}\ \forall\, t\in\mathbb{R}$. Τώρα, αφού η συνάρτηση $t\mapsto\mu\big(F_t\big)$ είναι αύξουσα (μη φθίνουσα), και θέτοντας $c=\sup\{t\in\mathbb{R}\colon\mu(F_t)=0\}$, έχουμε ότι $\mu\big(\{x\in X\colon f(x)=c\}\big)=1$. Πράγματι,

$$\mu\big(\{x\in X\colon f(x)<c\}\big)=\lim_{n\to\infty}\mu\big(\{x\in X\colon f(x)\leq c-1/n\}\big)=\lim_{n\to\infty}\mu\big(F_{c-1/n}\big)=0,$$

αφού $F_{c-1/n}\uparrow\{x\in X\colon f(x)<c\}$ $\big($επειδή $(-\infty,c-1/n]\uparrow(-\infty,c)$ και επομένως και για τις αντίστροφες εικόνες $f^{-1}\big((-\infty,c-1/n]\big)\uparrow f^{-1}\big((-\infty,c)\big)\big)$, και επίσης

$$\begin{aligned}\mu\big(\{x\in X\colon f(x)>c\}\big)&=\lim_{n\to\infty}\mu\big(\{x\in X\colon f(x)>c+1/n\}\big)\\&=\lim_{n\to\infty}\big[1-\mu\big(\{x\in X\colon f(x)\leq c+1/n\}\big)\big]=0.\end{aligned}$$

Τελικά

$$\mu\big(\{x\in X\colon f(x)=c\}\big)=1.$$

Μένει να δειχθεί και ότι $c\in\mathbb{R}$. Αυτό είναι συνέπεια του ότι η f παίρνει τιμές στο $\mathbb{R}$. Αυτό έχει σαν συνέπεια την $X=\bigcup_{n\in\mathbb{N}}F_n$, με την ένωση αύξουσα (μη φθίνουσα) και επομένως πρέπει $1=\lim_{n\to\infty}\mu(F_n)$, οπότε $\mu(F_n) > 0$ για κάποιο n και άρα $c \leq n$. Όμοια $X = \bigcup_{n\in\mathbb{N}} F_{-n}^{\mathrm{c}}$, με $F_{-n}^{\mathrm{c}} = f^{-1}\big((-n,+\infty)\big) \uparrow X$ πάλι, οπότε πάλι $1 = \lim_{n\to\infty}[1-\mu(F_{-n})]$, δηλαδή $\mu(F_{-n}) \to 0$, και άρα $\mu(F_{-n}) = 0$ για κάποιο n, αφού $\mu(F_t) \in \{0,1\}\ \forall\, t \in \mathbb{R}$. Τότε $c \geq -n$ για αυτό το n.

(2) Έστω $A \in \mathcal{B}$ με $A = T^{-1}(A)$ και θέτουμε $f = \mathbb{1}_A$. Τότε $f \in L^\infty(X,\mathcal{B},\mu)$ και

$$f \circ T = \mathbb{1}_A \circ T = \mathbb{1}_{T^{-1}(A)} = \mathbb{1}_A = f.$$

Από την υπόθεση, η f είναι σταθερή μ-σχεδόν παντού και άρα $f = 1$ μ-σχεδόν παντού ή $f = 0$ μ-σχεδόν παντού. Ισοδύναμα $\mu(A) \in \{0,1\}$. Εξ' ορισμού, το σύστημα είναι εργοδικό. □

Παραδείγματα

1. *Στροφές του κύκλου ή τόρου.* Έστω $X = \mathbb{T} = [0,1)$, $\mathcal{B} = \mathcal{B}(\mathbb{T})$, μ το μέτρο Lebesgue στο $[0,1)$, $a \in [0,1)$ και $T_a(x) = x + a \pmod 1$.

Πρόταση 2.1.6 Το σύστημα $(X,\mathcal{B},\mu,T)$, ορισμένο όπως παραπάνω, είναι εργοδικό ανν $a \notin \mathbb{Q}$.

Απόδειξη. Έστω, αρχικά, ότι το σύστημα είναι εργοδικό. Αν $a \in \mathbb{Q}$, έστω ότι $a = m/n$, με $m \in \mathbb{Z}, n \in \mathbb{N}, (m,n) = 1$. Τότε,

$$T_a^n(x) = x + na \pmod 1 = x + m \pmod 1 = x, \qquad \text{αφού} \quad m \in \mathbb{Z}.$$

Άρα, ο T_a^n είναι ο ταυτοτικός μετασχηματισμός και επομένως κάθε τροχιά $x, T_a(x), T_a^2(x), \dots$ του T_a είναι περιοδική. Ορίζουμε

$$f(x) = e^{2\pi i[x+T_a(x)+T_a^2(x)+\cdots+T_a^{n-1}(x)]}$$

τότε

$$f \circ T_a(x) = e^{2\pi i[T_a(x)+T_a^2(x)+\cdots+T_a^{n-1}(x)+T_a^n(x)]} = e^{2\pi i[T_a(x)+T_a^2(x)+\cdots+T_a^{n-1}(x)+x]} = f(x),$$

και έχουμε βρει μια T_a-αναλλοίωτη συνάρτηση που δεν είναι όμως σταθερή (για παράδειγμα $f(0) = -f(1/2n)$ και $f(0) = e^{\pi i n(n-1)a} = e^{\pi i m(n-1)} \neq 0$, άρα $f(0) \neq f(1/2n)$). Το σύστημα δηλαδή δεν είναι εργοδικό, άτοπο.

Δεύτερος τρόπος απόδειξης αυτής της κατεύθυνσης. Αν $a = m/n$ με m και n όπως πριν ($m \in \mathbb{Z}, n \in \mathbb{N}, (m,n) = 1$), τότε είδαμε ότι ο T_a^n είναι ο ταυτοτικός. Έστω A μετρήσιμο, έτσι ώστε $0 < \mu(A) < 1/n$ και $B = A \cup T_a(A) \cup \cdots \cup T_a^{n-1}(A)$. Το B είναι μετρήσιμο σύνολο και θα δειχτεί ότι $T_a(B) = B$. Έχουμε

$$T_a(B) = T_a(A) \cup T_a^2(A) \cup \cdots \cup T_a^n(A) \overset{T^n(A)=A}{=\!=\!=} B$$

και αφού

$$1 > n\,\mu(A) \geq \mu(B) > \mu(A) > 0,$$

έχουμε βρει μη τετριμμένο, μετρήσιμο, αναλλοίωτο σύνολο. Άρα έχουμε άτοπο.

Παρατηρούμε εδώ ότι επειδή ο T_a είναι αντιστρέψιμος ($T_a^{-1} = T_{1-a}$), όλα τα παραπάνω σύνολα είναι μετρήσιμα και επιπλέον, η σχέση $T_a(B) = B$ συνεπάγεται και την $B = T_a^{-1}(B)$.

Αντίστροφα, αν $a \in [0,1) \setminus \mathbb{Q}$, έστω $f \in L^2(\mu)$ αναλλοίωτη μ-σχεδόν παντού, δηλαδή τέτοια ώστε $f \overset{L^2}{=} f \circ T$. Κάθε $f \in L^2(\mathbb{T})$ έχει ανάπτυγμα Fourier:

$$f \overset{L^2}{=} \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n,$$

όπου, για $n \in \mathbb{Z}$,

$$e_n(t) = e^{2\pi i n t}$$

και

$$\langle f, e_n\rangle = \int_0^1 f(t)\overline{e_n(t)}\,dt = \int_0^1 f(t)e^{-2\pi i n t}\,dt = \widehat{f}(n).$$

Επειδή $f \overset{L^2}{=} f \circ T_a$, πρέπει $\widehat{f}(n) = \widehat{f \circ T_a}(n)$, για κάθε $n \in \mathbb{Z}$. Υπολογίζουμε,

$$\begin{aligned}
\widehat{f \circ T_a}(n) &= \int_0^1 f\big(T_a(t)\big) e^{-2\pi i n t}\, dt = \int_0^1 f\big(T_a(t)\big) e^{-2\pi i n (t+a)}\, dt\, e^{2\pi i n a} \\
\text{(επειδή } e_n \text{ περιοδική)} \qquad &= \int_0^1 f\big(T_a(t)\big) e^{-2\pi i n T_a(t)}\, dt\, e^{2\pi i n a} = \int_0^1 f(t) e^{-2\pi i n t} d\,[(T_a)_*\lambda](t)\; e^{2\pi i n a} \\
\text{(επειδή } (T_a)_*\lambda = \lambda) \qquad &= \int_0^1 f(t) e^{-2\pi i n t}\, dt\, e^{2\pi i n a} = \widehat{f}(n)\, e^{2\pi i n a}.
\end{aligned}$$

Επομένως, για κάθε $n \in \mathbb{Z}$,

$$\widehat{f}(n) = \widehat{f \circ T_a}(n) \quad \Leftrightarrow \quad \widehat{f}(n)\big(e^{2\pi i n a} - 1\big) = 0.$$

Όμως $a \notin \mathbb{Q} \Rightarrow na \notin \mathbb{Z}\, \forall\, n \in \mathbb{Z} \setminus \{0\}$. Επομένως,

$$\widehat{f}(n) = 0 \quad \forall\, n \in \mathbb{Z} \setminus \{0\},$$

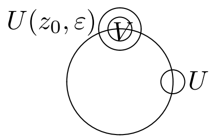
και άρα

$$f \overset{L^2}{=} \widehat{f}(0) e_0 = \widehat{f}(0) \mathbb{1}_{\mathbb{T}}.$$

Δηλαδή η f είναι σταθερή (και ίση με την σταθερά $\widehat{f}(0)$) μ-σχεδόν παντού. Δηλαδή, κάθε μ-σχεδόν αναλλοίωτη συνάρτηση στον $L^2(\mu)$ είναι σταθερή μ-σχεδόν παντού. Έπεται ότι το σύστημα είναι εργοδικό. □

Θεώρημα 2.1.7 (Kronecker) Αν $a \in [0,1) \setminus \mathbb{Q}$, τότε το σύνολο $\{e^{2\pi i n a} : n \in \mathbb{N}\}$ είναι πυκνό στον κύκλο. Ισοδύναμα, $\{na \;(\mathrm{mod}\; 1) : n \in \mathbb{N}\} = \{na - \lfloor na \rfloor : n \in \mathbb{N}\}$ είναι πυκνό στο $[0,1)$.

Απόδειξη. Έστω $\zeta = e^{2\pi i a}$ (μιγαδικός συμβολισμός) και για αυθαίρετο $z_0 \in \mathbb{S}^1$ και $\varepsilon > 0$, έστω $U(z_0, \varepsilon)$ η ανοικτή μπάλα κέντρου z_0 και ακτίνας ε στο επίπεδο. Για να δείξουμε ότι το $\{\zeta^n : n \in \mathbb{N}\}$ είναι πυκνό στον κύκλο $\mathbb{S}^1$ αρκεί να δείξουμε ότι $I \cap \{\zeta^n : n \in \mathbb{N}\} \neq \varnothing$, για κάθε ανοικτό διάστημα I στον κύκλο $\mathbb{S}^1$. Έστω λοιπόν $I = U(z_0, \varepsilon) \cap \mathbb{S}^1$ ένα ανοικτό διάστημα στον κύκλο $\mathbb{S}^1$ και έστω $U = U(1, \varepsilon/2)$ η ανοικτή μπάλα κέντρου $1 \in \mathbb{C}$ και ακτίνας $\varepsilon/2$ και V η ανοικτή μπάλα $V = U(z_0, \varepsilon/2)$. Τότε τα σύνολα



$U \cap \mathbb{S}^1$ και $V \cap \mathbb{S}^1$ έχουν θετικό $\lambda_{\mathbb{S}^1}$-μέτρο, δηλαδή θετικό μήκος τόξου. Αφού $a \in [0,1) \setminus \mathbb{Q}$, η T_a είναι εργοδική και άρα από το (4) της Πρότασης 2.1.4, υπάρχει $n \in \mathbb{N}$ ώστε $\lambda_{\mathbb{S}^1}((U \cap \mathbb{S}^1) \cap (T_a^{-n}(V \cap \mathbb{S}^1))) > 0$. Ειδικότερα, υπάρχει $z \in U \cap \mathbb{S}^1$ ώστε $T_a^n(z) \in V$, δηλαδή $\zeta^n z \in V$, για αυτό το n. Τότε όμως $|\zeta^n z - z_0| < \varepsilon/2$. Επιπλέον

$$z \in U \quad \Rightarrow \quad |z-1| < \varepsilon/2 \quad \Rightarrow \quad |\zeta^n z - \zeta^n| = |\zeta^n||z-1| = |z-1| < \varepsilon/2,$$

και άρα τελικά, από την τριγωνική ανισότητα, $|\zeta^n - z_0| < \varepsilon$, δηλαδή $\zeta^n \in I = U(z_0, \varepsilon) \cap \mathbb{S}^1$. □

Παρατήρηση. Το ίδιο επιχείρημα δίνει το εξής: αν G συμπαγής ομάδα, $a \in G$ και $T_a(x) = ax$, τότε το σύστημα με $X = G$, $\mathcal{B} = \mathcal{B}(G)$, μ το μέτρο Haar της G και μετασχηματισμό T_a, είναι εργοδικό ανν $\overline{\{a^n : n \in \mathbb{N}\}} = G$.

Ειδικότερα, για να είναι κάποια T_a εργοδική, πρέπει η G να είναι αβελιανή.

2. *Στροφές του n-διάστατου τόρου.* Έστω $X = \mathbb{T}^n$, μ το n-διάστατο μέτρο Lebesgue στο $[0,1)^n$ και

$$T_a(x) = (x_1 + a_1, \dots, x_n + a_n) \;(\mathrm{mod}\; 1), \qquad \text{όπου} \quad x = (x_1, \dots, x_n), \quad a = (a_1, \dots, a_n).$$

Ο μετασχηματισμός αυτός είναι εργοδικός ανν

$$k_1, k_2, \dots, k_n \in \mathbb{Z} \quad \text{και} \quad k_1 a_1 + k_2 a_2 + \cdots + k_n a_n \in \mathbb{Z} \qquad \Rightarrow \qquad k_1 = k_2 = \cdots = k_n = 0$$

$$\Leftrightarrow \qquad 1, a_1, a_2, \ldots, a_n \quad \text{είναι γραμμικά ανεξάρτητα πάνω από το } \mathbb{Q}.$$

3. *Επιμορφισμοί του τόρου.* Έστω $X = \mathbb{T}$, $\mathcal{B} = \mathcal{B}(\mathbb{T})$, μ το μέτρο Lebesgue στο $[0, 1)$ και

$$T_2(x) = 2x \ (\mathrm{mod}\ 1).$$

Το σύστημα αυτό είναι εργοδικό. Πράγματι, έστω $f \in L^2$ με $f \overset{L^2}{=} f \circ T_2$. Έστω

$$f \overset{L^2}{=} \sum_{n\in\mathbb{Z}} \widehat{f}(n) e_n$$

ο μετασχηματισμός Fourier της f. Τότε

$$\widehat{f \circ T_2}(n) = \int_0^1 f(T_2(x)) e^{-2\pi i n x}\, dx$$

για κάθε $n \in \mathbb{Z}$. Τώρα για άρτιους ακέραιους έχει κανείς ότι

$$\begin{aligned}
\widehat{f \circ T_2}(2n) = \int_0^1 f(T_2(x)) e^{-2\pi i 2nx}\, dx &= \int_0^{1/2} f(2x) e^{-2\pi i n 2x}\, dx + \int_{1/2}^1 f(2x-1) e^{-2\pi i n 2x}\, dx \\
&= \int_0^{1/2} f(2x) e^{-2\pi i n 2x}\, dx + \int_{1/2}^1 f(2x-1) e^{-2\pi i n (2x-1)}\, dx \\
&= \frac{1}{2}\int_0^1 f(x) e^{-2\pi i n x}\, dx + \frac{1}{2}\int_0^1 f(x) e^{-2\pi i n x}\, dx = \widehat{f}(n),
\end{aligned}$$

$n \in \mathbb{Z}$. Αφού $f \overset{L^2}{=} f \circ T_2$, πρέπει

$$\widehat{f \circ T_2}(2n) = \widehat{f}(2n) \quad \forall\, n \in \mathbb{Z} \qquad \Leftrightarrow \qquad \widehat{f}(n) = \widehat{f}(2n) \quad \forall\, n \in \mathbb{Z}.$$

Έστω $A_n := \{2^k n \colon k \in \mathbb{N}\}$ για $n \in \mathbb{Z}$. Από τα παραπάνω, $\widehat{f}(m) = \widehat{f}(n)$ για κάθε $m \in A_n$. Επιπλέον, αν $n \neq 0$ τότε το σύνολο A_n είναι άπειρο σύνολο. Από την ταυτότητα Parseval,

$$\|f\|_2 = \sum_{m\in\mathbb{Z}} |\widehat{f}(m)|^2 \geq \sum_{m\in A_n} |\widehat{f}(m)|^2 = \sum_{m\in A_n} |\widehat{f}(n)|^2$$

και το τελευταίο άθροισμα θα είναι άπειρο αν $\widehat{f}(n) \neq 0$ και $n \neq 0$ γιατί το σύνολο A_n είναι άπειρο για $n \neq 0$. Αφού $f \in L^2$ και άρα $\|f\| < \infty$, θα πρέπει

$$\widehat{f}(n) = 0 \quad \forall\, n \in \mathbb{Z} \setminus \{0\}.$$

Δηλαδή, η $f = \widehat{f}(0)$ θα είναι σταθερή μ-σχεδόν παντού. Άρα κάθε μ-σχεδόν αναλλοίωτη συνάρτηση στον $L^2(\mu)$ είναι σταθερή μ-σχεδόν παντού και άρα το σύστημα είναι εργοδικό. □

Παρατήρηση. Το παραπάνω επιχείρημα δουλεύει για κάθε επιμορφισμό $T_k(x) = kx \ (\mathrm{mod}\ 1)$, $x \in \mathbb{T}$, του τόρου $\mathbb{T}$, $k \in \mathbb{N}$, $k \geq 2$, και άρα όλοι αυτοί οι επιμορφισμοί του τόρου είναι εργοδικοί.

4. *Επιμορφισμοί του n-διάστατου τόρου.* Έστω $X = \mathbb{T}^n$, $\mathcal{B} = \mathcal{B}(\mathbb{T}^n)$, μ το μέτρο Lebesgue στο $[0, 1)^n$ και

$$T(x) = Ax \ (\mathrm{mod}\ 1),$$

όπου A είναι $n \times n$ πίνακας με $A_{ij} \in \mathbb{Z}$ για κάθε i, j και $\det A \neq 0$. Όπως ήδη γνωρίζουμε (Παράδειγμα 1.1.14), αυτό είναι ένα σύστημα που διατηρεί το μέτρο.

Πρόταση 2.1.8 Το παραπάνω σύστημα είναι εργοδικό ανν καμιά ιδιοτιμή του A δεν είναι ρίζα της μονάδας.

Απόδειξη. Θα δείξουμε ότι αν $f \in L^2$ είναι σχεδόν αναλλοίωτη, δηλαδή $f \overset{L^2}{=} f \circ T$, τότε είναι σταθερή μ-σχεδόν παντού. Θεωρούμε το ανάπτυγμα Fourier της f,

$$f \overset{L^2}{=} \sum_{m\in\mathbb{Z}^n} \langle f, e_m \rangle e_m$$

όπου $e_{(k_1,\dots,k_n)}(x_1,\dots,x_n) = e^{2\pi i\langle k,x\rangle}$, $k = (k_1,\dots,k_n)$, $x = (x_1,\dots,x_n)$ και $\langle x,k\rangle$ είναι το εσωτερικό γινόμενο $\langle x,k\rangle = \sum_{j=1}^n k_j x_j$, και

$$\langle f, e_m\rangle = \int_{\mathbb{T}^n} f(t)\overline{e_m(t)}\,dt = \widehat{f}(m).$$

Υπολογίζουμε

$$\widehat{(f\circ T)}(k) = \int_{\mathbb{T}^n} f\big(T(x)\big)e^{-2\pi i\langle k,x\rangle}\,dx.$$

Τώρα,

$$e^{-2\pi i\langle k,x\rangle} = e^{-2\pi i\langle k,A^{-1}Ax\rangle} = e^{-2\pi i\langle (A^{-1})^* k,Ax\rangle} = e^{-2\pi i\langle (A^{-1})^* k,T(x)\rangle},$$

όπου $(A^{-1})^*$ είναι ο συζυγής του πίνακα A^{-1}, δηλαδή εδώ ο ανάστροφος πίνακας[6]. Άρα, για κάθε $k \in \mathbb{Z}^n$,

$$\widehat{(f\circ T)}(A^*k) = \int_{\mathbb{T}^n} f\big(T(x)\big)e^{-2\pi i\langle k,T(x)\rangle}\,dx \overset{\text{σ.δ.μ.}}{=\!=} \int_{\mathbb{T}^n} f(x)e^{-2\pi i\langle k,x\rangle}\,dx = \widehat{f}(k).$$

Επομένως,

$$\begin{aligned} f \overset{L^2}{=} f\circ T \quad &\Rightarrow \quad \widehat{(f\circ T)}(m) = \widehat{f}(m) \quad \forall\, m\in\mathbb{Z}^n \quad \Rightarrow \quad \widehat{(f\circ T)}(A^*k) = \widehat{f}(A^*k) \quad \forall\, k\in\mathbb{Z}^n \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \Rightarrow \quad \widehat{f}(k) = \widehat{f}(A^*k) \quad \forall\, k\in\mathbb{Z}^n, \end{aligned}$$

και επαγωγικά έχει κανείς ότι,

$$\widehat{f}(k) = \widehat{f}\big((A^*)^j k\big) \quad \forall\, k\in\mathbb{Z}^n \quad \forall\, j\in\mathbb{N}.$$

Ορίζουμε τώρα $B_k := \{(A^*)^j k : j\in\mathbb{N}\cup\{0\}\}$, για $k\in\mathbb{Z}\setminus\{0\}$. Αν το B_k είναι άπειρο σύνολο, τότε επειδή

$$+\infty > \|f\|^2 = \sum_{m\in\mathbb{Z}^n} |\widehat{f}(m)|^2 \geq \sum_{m\in B_k} |\widehat{f}(m)|^2 = \sum_{m\in B_k} |\widehat{f}(k)|^2,$$

και το τελευταίο άθροισμα είναι άπειρο εάν $|\widehat{f}(k)| > 0$, θα πρέπει τελικά $\widehat{f}(k) = 0$. Τώρα αν το B_k είναι πεπερασμένο σύνολο, υπάρχουν $j_1, j_2\in\mathbb{Z}$ με $j_1\neq j_2$ τέτοια ώστε

$$(A^*)^{j_1}k = (A^*)^{j_2}k \quad \Leftrightarrow \quad (A^*)^{|j_1-j_2|}k = k. \tag{5}$$

Αν $k\neq 0$, αφού η (5) ισχύει για $j_1\neq j_2$, έπεται ότι ο πίνακας $(A^*)^{|j_1-j_2|}$ έχει ιδιοτιμή το 1, δηλαδή $(A^*)^j k = k$ για κάποιο $j\in\mathbb{N}$ και επομένως ο $(A^*)^j$ έχει ιδιοτιμή το 1, και άρα μια j-οστή ρίζα της μονάδας θα είναι ιδιοτιμή του A^*. Επειδή ο A είναι πραγματικός πίνακας, οι ιδιοτιμές του A και του A^* ταυτίζονται και άρα ο A έχει μία ιδιοτιμή που είναι j-στη ρίζα της μονάδας. Όμως αυτό είναι άτοπο από την υπόθεση. Άρα

$$j_1\neq j_2 \quad \Rightarrow \quad (A^*)^{j_1}k \neq (A^*)^{j_2}k \quad \text{για } k\neq 0.$$

Δηλαδή το B_k είναι άπειρο σύνολο $\forall\, k\in\mathbb{Z}^n\setminus\{0\}$, και επομένως $\widehat{f}(k) = 0$ $\forall\, k\in\mathbb{Z}^n\setminus\{0\}$. Άρα

$$f \overset{L^2}{=} \sum_{k\in\mathbb{Z}^n} \widehat{f}(k)e_k = \widehat{f}(0)e_0 = \widehat{f}(0)\mathbb{1}_X.$$

Δηλαδή η f είναι σταθερή σχεδόν παντού. Επομένως το σύστημα είναι εργοδικό όταν ο A δεν έχει ιδιοτιμή που να είναι ρίζα της μονάδας.

Αντίστροφα, έστω ότι ο A έχει ιδιοτιμή κάποια ρίζα της μονάδας. Τότε ο 1 είναι ιδιοτιμή του $(A^*)^m$ για κάποιο $m\in\mathbb{N}$. Έστω m ο ελάχιστος τέτοιος φυσικός αριθμός. Τότε υπάρχει[7] $k\in\mathbb{Z}^n\setminus\{0\}$ τέτοιος ώστε $(A^*)^m k = k$. Τότε θεωρώντας την συνάρτηση

$$f(x) = e^{2\pi i\langle k,x\rangle} + e^{2\pi i\langle k,T(x)\rangle} + \cdots + e^{2\pi i\langle k,T^{m-1}(x)\rangle},$$

[6]Διευκρίνηση: για οποιοδήποτε $x = (x_1,\dots,x_n)\in[0,1)^n$, $Ax - T(x)\in\mathbb{Z}^n$ και άρα $\langle k, Ax - T(x)\rangle\in\mathbb{Z}$ για $k\in\mathbb{Z}^n$, και επομένως $e^{-2\pi i\langle k,Ax-T(x)\rangle} = 1$, ισοδύναμα $e^{-2\pi i\langle k,Ax\rangle} = e^{-2\pi i\langle k,T(x)\rangle}$.

[7]Υπάρχει καταρχήν $v\in\mathbb{R}^n\setminus\{0\}$ τέτοιο ώστε $(A^*)^m v = v$, αλλά επειδή όλα τα στοιχεία του πίνακα A είναι ακέραιοι, υπάρχει τελικά $k\in\mathbb{Z}^n\setminus\{0\}$ τέτοιο ώστε $(A^*)^m k = k$.

έχουμε

$$\begin{aligned} f\big(T(x)\big) &= e^{2\pi i\langle k,T(x)\rangle} + e^{2\pi i\langle k,T^2(x)\rangle} + \dots + e^{2\pi i\langle k,T^{m-1}(x)\rangle} + e^{2\pi i\langle k,T^m(x)\rangle} \\ &= e^{2\pi i\langle k,T(x)\rangle} + e^{2\pi i\langle k,T^2(x)\rangle} + \dots + e^{2\pi i\langle k,T^{m-1}(x)\rangle} + e^{2\pi i\langle k,A^m x\rangle} \end{aligned}$$

και επειδή

$$e^{2\pi i\langle k,A^m x\rangle} = e^{2\pi i\langle (A^*)^m k,x\rangle} = e^{2\pi i\langle k,x\rangle},$$

έπεται ότι

$$f\big(T(x)\big) = f(x).$$

Αλλά η f δεν είναι σταθερή. Πράγματι,

$$f(x) = e^{2\pi i\langle k,x\rangle} + e^{2\pi i\langle A^*k,x\rangle} + \dots + e^{2\pi i\langle (A^*)^{m-1}k,x\rangle}$$

σημαίνει ότι

$$f = e_k + e_{A^*k} + \dots + e_{(A^*)^{m-1}k}.$$

Τα $k, A^*k, \dots, (A^*)^{m-1}k$ είναι διαφορετικά μεταξύ τους, γιατί το m είναι ο ελάχιστος φυσικός για τον οποίο $(A^*)^m$ έχει ιδιοτιμή τον 1. Άρα, οι συναρτήσεις $e_k, e_{A^*k}, \dots, e_{(A^*)^{m-1}k}$ είναι ορθογώνιες μεταξύ τους, επομένως και γραμμικά ανεξάρτητες. Μάλιστα, το σύνολο $\{e_0, e_k, e_{A^*k}, \dots, e_{(A^*)^{m-1}k}\}$ είναι γραμμικά ανεξάρτητο, όπου $e_0(x) = 1 \ \forall x \in [0,1)^n$, γιατί έχουμε ότι επίσης $(A^*)^j k \neq 0$ για κάθε $j \in \{0,1,\dots,m-1\}$, αφού αλλιώς θα είχαμε ότι $(A^*)^m k = 0 \neq k$. Τώρα, αν η f ήταν σταθερή, θα είχαμε έναν μη τετριμμένο γραμμικό συνδυασμό των $e_0, e_k, e_{A^*k}, \dots, e_{(A^*)^{m-1}k}$ που θα ήταν ταυτοτικά ίσος με 0. Συγκεκριμένα, αν $f = c\mathbb{1}_X = ce_0$ στον L^2 (δηλαδή αν η f ισούται με την σταθερά c μ-σχεδόν παντού), θα είχαμε ότι $-ce_0 + e_k + e_{A^*k} + \ldots + e_{(A^*)^{m-1}k} = 0$. Αυτό αντίκειται στην γραμμική ανεξαρτησία του συνόλου $\{e_0, e_k, e_{A^*k}, \dots, e_{(A^*)^{m-1}k}\}$. Άρα η f δεν είναι σταθερή και το σύστημα δεν είναι εργοδικό. □

5. *Μονόπλευρο Bernoulli Shift.* Έστω $S = \{1,\dots,s\}$ ένα πεπερασμένο σύνολο και $p = (p_i)_{i\in S}$ διάνυσμα πιθανότητας. Επίσης έστω $X = S^{\mathbb{N}} = \{\underset{\sim}{x} = (x_1,x_2,\dots) : x_i \in S \quad \forall i \in \mathbb{N}\}$, $\mathcal{B} = \sigma(\pi_1,\pi_2,\dots)$, όπου $\pi_i(x_1,x_2,\dots) = x_i$, $i \in \mathbb{N}$, $\mu\big(\{\underset{\sim}{x} \in X : x_1 = s_1,\dots,x_n = s_n\}\big) = p_{s_1}p_{s_2}\cdots p_{s_n}$ και ο μετασχηματισμός $T\colon X \to X$ να είναι ο shift $T(x_1,x_2,\dots) = (x_2,x_3,\dots)$. Ξέρουμε ότι το σύστημα $(X,\mathcal{B},\mu,T)$ είναι σύστημα που διατηρεί το μέτρο.

Πρόταση 2.1.9 Το παραπάνω σύστημα (Bernoulli shift) είναι εργοδικό.

Λήμμα 2.1.10 Έστω $(X,\mathcal{B},\mu)$ χώρος πιθανότητας και $\mathcal{A}$ μια άλγεβρα που παράγει την $\mathcal{B}$. Τότε για κάθε $B \in \mathcal{B}$ και για κάθε $\varepsilon > 0$, υπάρχει $A \in \mathcal{A}$ τέτοιο ώστε

$$\mu(A \triangle B) < \varepsilon.$$

Απόδειξη. Έστω $\mathcal{C} = \{C \in \mathcal{B} : \forall\, \varepsilon > 0\ \exists\, A \in \mathcal{A}$ τέτοιο ώστε $\mu(A \triangle C) < \varepsilon\}$. Αυτή η κλάση συνόλων είναι σ-άλγεβρα και περιέχει την $\mathcal{A}$, άρα $\mathcal{C} = \mathcal{B}$. □

Απόδειξη Πρότασης. Οι κύλινδροι

$$\{\underset{\sim}{x} = (x_1,x_2,\dots) \in X : (x_1,\dots,x_m) \in B\}, \quad m \in \mathbb{N}, \quad B \subset S^m,$$

μαζί με το κενό σύνολο, αποτελούν άλγεβρα, που εξ' ορισμού παράγει την $\mathcal{B}$.

Έστω $A \in \mathcal{B}$ με $A = T^{-1}(A)$. Θα δειχτεί ότι $\mu(A) \in \{0,1\}$.

Έστω $\varepsilon > 0$. Από το Λήμμα 2.1.10 υπάρχει κύλινδρος

$$C = \{\underset{\sim}{x} = (x_1,x_2,\dots) \in X : (x_1,\dots,x_m) \in B\},$$

όπου $B \subset S^m$, τέτοιος ώστε

$$\mu(A \triangle C) < \varepsilon.$$

Έστω $n \geq m$ και θεωρούμε το σύνολο

$$C \cap T^{-n}(C) = \{\underset{\sim}{x} = (x_1,x_2,\dots) \in X : (x_1,\dots,x_m) \in B, (x_{n+1},\dots,x_{n+m}) \in B\}.$$

Τότε, επειδή το μ είναι μέτρο γινόμενο[8],

$$\mu(C \cap T^{-n}(C)) = \mu(C)\mu(T^{-n}(C)) = \mu(C)^2.$$

Τώρα,

$$A \triangle \big(C \cap T^{-n}(C)\big) \subset \big(A \triangle C\big) \cup \big(A \triangle T^{-n}(C)\big) = \big(A \triangle\ C\big) \cup \big(T^{-n}(A) \triangle\ T^{-n}(C)\big),$$

αφού το A είναι αναλλοίωτο. Επομένως,

$$\mu\big(A \triangle \big(C \cap T^{-n}(C)\big)\big) \leq \mu(A \triangle C) + \mu(T^{-n}(A) \triangle T^{-n}(C)) < \varepsilon + \mu(T^{-n}(A \triangle\ C)) < 2\varepsilon.$$

Άρα

$$\big|\mu(A) - \mu\big(C \cap T^{-n}(C)\big)\big| \leq \mu\big(A \triangle \big(C \cap T^{-n}(C)\big)\big) < 2\varepsilon \quad \Rightarrow \quad \big|\mu(A) - \mu(C)^2\big| < 2\varepsilon.$$

Επομένως,

$$\begin{aligned}
\big|\mu(A) - \mu(A)^2\big| &\leq \big|\mu(A) - \mu(C)^2\big| + \big|\mu(C)^2 - \mu(A)^2\big| \\
&< 2\varepsilon + \big|\mu(A) - \mu(C)\big|[\mu(A) + \mu(C)] \\
&\leq 2\varepsilon + 2\mu(A \triangle C) \\
&< 4\varepsilon.
\end{aligned}$$

Το ε ήταν αυθαίρετο, άρα $\mu(A) = \mu(A)^2$ και επομένως $\mu(A) \in \{0, 1\}$. □

6. *Απεικόνιση Gauss.* Αν $X = (0, 1]$, $\mathcal{B} = \mathcal{B}(0, 1]$, ο μετασχηματισμός Gauss

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

διατηρεί το μέτρο

$$d\mu(x) = \frac{1}{\ln 2} \frac{dx}{1 + x}.$$

Πρόταση 2.1.10 Το παραπάνω σύστημα (Gauss) είναι εργοδικό.

2.2 Εργοδικά Θεωρήματα

Ο Τελεστής Koopman

Ορισμός 2.2.1 (τελεστής Koopman) Έστω $(X, \mathcal{B}, \mu, T)$ σύστημα που διατηρεί το μέτρο. Αν $f\colon X \to \mathbb{R}$ (ή $f\colon X \to \mathbb{C}$), ορίζεται ο *τελεστής Koopman*

$$U_T f := f \circ T.$$

Ο U_T είναι ισομετρία σε κάθε χώρο L^p, $p \in [1, \infty]$. Πράγματι, έστω $f \in L^p(X, \mathcal{B}, \mu)$, $1 \leq p < \infty$, τότε

$$\begin{aligned}
\|U_T f\|_p^p = \|f \circ T\|_p^p &= \int |f \circ T|^p \, d\mu = \int |f|^p \circ T \, d\mu \\
&= \int |f|^p \, dT_*\mu = \int |f|^p \, d\mu = \|f\|_p^p
\end{aligned}$$

Ειδικότερα, ο U_T είναι ισομετρία $U_T\colon L^2(\mu) \to L^2(\mu)$. Αυτό σημαίνει ότι

$$\langle U_T f, U_T g\rangle_{L^2(\mu)} = \langle f, g\rangle_{L^2(\mu)}, \quad \text{όπου} \quad \langle f, g\rangle_{L^2(\mu)} = \int f\bar{g} \, d\mu.$$

Για $p = \infty$,

$$\mu\big(\{x \in X : |f \circ T(x)| > t\}\big) = \mu\big(T^{-1}\big(\{x \in X : |f(x)| > t\}\big)\big) = \mu\big(\{x \in X : |f(x)| > t\}\big)$$

[8]δηλαδή οι προβολές $\pi_1, \pi_2, \ldots$ ανεξάρτητες τυχαίες μεταβλητές

για κάθε $t \in [0, \infty)$ και άρα $\|f \circ T\|_\infty = \|f\|_\infty$.

Όταν το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι αντιστρέψιμο, ο U_T είναι unitary (ορθομοναδιαίος), δηλαδή είναι και ισομορφισμός (και επί). Όταν όμως το σύστημα δεν είναι αντιστρέψιμο αυτό δεν ισχύει. Για παράδειγμα, αν $X = \mathbb{T}$, $\mathcal{B} = \mathcal{B}([0,1))$, μ το μέτρο Lebesgue και $T(x) = 2x \pmod 1$, και αν $f = \mathbb{1}_{[0,1/2)}$, τότε $f \in L^2(\mu)$ αλλά $U_T g \neq f \ \ \forall\, g \in L^2(\mu)$.

Σε κάθε σύστημα, αν $\mathbb{1}_X$ είναι η σταθερή συνάρτηση ίση με ένα, τότε αυτή είναι ιδιοδιάνυσμα του U_T για την ιδιοτιμή ένα, αφού

$$U_T \mathbb{1}_X = \mathbb{1}_X$$

και άρα ο U_T έχει ιδιοτιμή 1. Επειδή

$$f \circ T = f \quad \Leftrightarrow \quad U_T f = f,$$

οι αναλλοίωτες συναρτήσεις είναι ακριβώς τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή 1 για τον U_T. Άρα έχουμε την παρακάτω πρόταση.

Πρόταση 2.2.2 Το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι εργοδικό ανν η ιδιοτιμή 1 είναι απλή ιδιοτιμή [9] για τον τελεστή Koopman $U_T \colon L^2(\mu) \to L^2(\mu)$.

Μέσο Εργοδικό Θεώρημα Von Neumann

Θεώρημα 2.2.3 (Μέσο Εργοδικό Θεώρημα von Neumann) Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. Τότε για κάθε $f \in L^2(X, \mathcal{B}, \mu)$ υπάρχει $\tilde{f} \in L^2(X, \mathcal{B}, \mu)$ τέτοιο ώστε

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \xrightarrow[(n\to\infty)]{L^2} \tilde{f} \qquad \text{και} \qquad \tilde{f} \overset{L^2}{=} \tilde{f} \circ T.$$

Μέσο Εργοδικό Θεώρημα Von Neumann για Συστολές σε Χώρους Hilbert

Ορισμός 2.2.4 (συστολή) Ένας γραμμικός τελεστής $U \colon H \to H$, όπου H χώρος Hilbert, καλείται *συστολή* αν[10] $\|U\| \le 1$.

Το θεώρημα von Neumann είναι ειδική περίπτωση του παρακάτω θεωρήματος.

Θεώρημα 2.2.5 (Μέσο Εργοδικό Θεώρημα von Neumann για συστολή σε χώρο Hilbert) Έστω H χώρος Hilbert και $U \colon H \to H$ μια συστολή. Έστω $F = \{h \in H \colon Uh = h\}$ και P_F η ορθογώνια προβολή στον F. Τότε

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k h \to P_F(h) \quad \forall h \in H.$$

Σημείωση. Ο F είναι γραμμικός και κλειστός υπόχωρος του H.

Από το παραπάνω θεώρημα, αν $H = L^2(X, \mathcal{B}, \mu)$ και U ο τελεστής Koopman, παίρνουμε το μέσο εργοδικό θεώρημα von Neumann για σ.δ.μ. (Θεώρημα 2.2.3). Επίσης, το γεγονός ότι $P_F(h) \in F$ μας δίνει ότι το όριο $\tilde{h} = P_F(h)$ είναι T-αναλλοίωτη συνάρτηση.

Λήμμα 1. Έστω H χώρος Hilbert και $U \colon H \to H$ συστολή. Τότε

$$Uh = h \quad \Leftrightarrow \quad U^* h = h.$$

[9] Ο ιδιόχωρος που αντιστοιχεί σε αυτήν έχει διάσταση 1.

[10] Όπου $\|U\| := \sup\{\|Ux\| \colon x \in H,\ \|x\| \le 1\}$ είναι η νόρμα του U.

Απόδειξη. Το γεγονός ότι U είναι συστολή, μας δίνει ότι και U^* είναι συστολή. Έστω $h \in H$ τέτοιο ώστε $Uh = h$. Τότε

$$\begin{aligned}
\|U^*h - h\|^2 &= \|U^*h\|^2 + \|h\|^2 - \langle U^*h, h\rangle - \langle h, U^*h\rangle \\
&\leq \|h\|^2 + \|h\|^2 - \langle h, Uh\rangle - \langle Uh, h\rangle \\
&= 2\|h\|^2 - \langle h, h\rangle - \langle h, h\rangle = 0.
\end{aligned}$$

□

Λήμμα 2. Αν $U\colon H \to H$ συστολή, όπου H χώρος Hilbert, $F = \{f \in H : Uf = f\}$ και $N = \{Uh - h : h \in H\}$, τότε $N^\perp = F$ και άρα $F^\perp = \overline{N}$. Δηλαδή

$$H = F \oplus \overline{N}.$$

Απόδειξη. Για $f \in H$ έχει κανείς ότι

$$\begin{aligned}
&\langle Uh - h, f\rangle = 0 \quad \forall h \in H \\
\Leftrightarrow\quad &\langle (U - I)h, f\rangle = 0 \quad \forall h \in H \\
\Leftrightarrow\quad &\langle h, (U - I)^*f\rangle = 0 \quad \forall h \in H \\
\Leftrightarrow\quad &(U^* - I)f = 0 \\
\Leftrightarrow\quad &U^*f = f \\
\Leftrightarrow\quad &Uf = f,
\end{aligned}$$

όπου στην τελευταία ισοδυναμία χρησιμοποιήθηκε το Λήμμα 1. Άρα $N^\perp = F$. □

Απόδειξη του Θεωρήματος 2.2.5. Έστω $h \in H$. Τότε, λόγω του Λήμματος 2, το h γράφεται μονοσήμαντα στη μορφή

$$h = f + g, \qquad f \in F, \quad g \in \overline{N}.$$

Συγκεκριμένα, $f = P_F(h)$. Τώρα,

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k h = \frac{1}{n}\sum_{k=0}^{n-1} U^k f + \frac{1}{n}\sum_{k=0}^{n-1} U^k g$$

και επειδή $n^{-1}\sum_{k=0}^{n-1} U^k f = f \;\; \forall n \in \mathbb{N}$, αφού $f \in F$ και άρα $Uf = f$, αρκεί να δειχθεί ότι

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k g\right\| \to 0.$$

Αν $g \in N$, τότε υπάρχει $\varphi \in H$ τέτοια ώστε $g = U\varphi - \varphi$, οπότε,

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1} U^k g = \frac{1}{n}\sum_{k=1}^{n} U^k \varphi - \frac{1}{n}\sum_{k=0}^{n-1} U^k \varphi = \frac{1}{n}U^n\varphi - \frac{1}{n}\varphi \quad \Rightarrow \\
\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k g\right\| \leq \frac{1}{n}\big(\|U^n\varphi\| + \|\varphi\|\big) \leq \frac{1}{n}2\|\varphi\| \to 0,
\end{aligned}$$

όπου η τελευταία ανισότητα ισχύει επειδή U είναι συστολή. Τώρα με ένα επιχείρημα προσέγγισης ολοκληρώνεται η απόδειξη. Πιο συγκεκριμένα, έστω $\varepsilon > 0$. Για τυχούσα $g \in \overline{N}$, βρίσκουμε $\tilde{g} \in N$ τέτοια ώστε $\|g - \tilde{g}\| < \varepsilon/2$. Ξέρουμε ότι υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $\forall\, n \geq n_0$, $\left\|n^{-1}\sum_{k=0}^{n-1} U^k \tilde{g}\right\| < \varepsilon/2$ και τότε

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k g\right\| &\leq \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k \tilde{g}\right\| + \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k (g - \tilde{g})\right\| \\
&< \varepsilon/2 + \frac{1}{n}\sum_{k=0}^{n-1} \left\|U^k(g - \tilde{g})\right\| \\
&\leq \varepsilon/2 + \frac{1}{n}\sum_{k=0}^{n-1} \|g - \tilde{g}\| \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon,
\end{aligned}$$

η προτελευταία ανισότητα πάλι γιατί ο U είναι συστολή. $\square$

Παρατηρήσεις. Όπως ήδη αναφέρθηκε, η επιλογή $H = L^2(X, \mathcal{B}, \mu)$ σε ένα σύστημα που διατηρεί το μέτρο $L^2(X, \mathcal{B}, \mu, T)$ δίνει το μέσο εργοδικό θεώρημα von Neumann για συστήματα που διατηρούν το μέτρο (Θεώρημα 2.2.3).

1. Η $\tilde{f}$ είναι ορθογώνια προβολή στον χώρο των L^2 συναρτήσεων που είναι αναλλοίωτες μ- σχεδόν παντού, δηλαδή στον χώρο $L^2(X, \mathcal{A}, \mu)$ όπου $\mathcal{A}$ η σ-άλγεβρα των μ-σχεδόν αναλλοίωτων συνόλων στην $\mathcal{B}$.
2. Με αυτή την θεώρηση, τίποτε δεν αλλάζει αν θεωρήσουμε συστήματα $(X, \mathcal{B}, \mu, T)$ με το μέτρο μ να μην είναι μέτρο πιθανότητας αλλά αυθαίρετο μέτρο, πιθανώς και άπειρο. Το μέσο εργοδικό θεώρημα Von Neumann ισχύει δηλαδή και για συστήματα όπου ο χώρος $(X, \mathcal{B}, \mu)$ δεν είναι χώρος πιθανότητας. Επιτρέπεται και η περίπτωση $\mu(X) = \infty$.

Θεώρημα 2.2.6 (Θεώρημα Επαναφοράς Khintchine) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $B \in \mathcal{B}$ με $\mu(B) > 0$. Τότε, για κάθε $\varepsilon > 0$, το σύνολο

$$\left\{n \in \mathbb{N} \colon \mu(T^{-n}(B) \cap B) > \mu(B)^2 - \varepsilon\right\}$$

έχει φραγμένα κενά.

Απόδειξη. Έστω $\varepsilon > 0$. Εφαρμόζουμε το θεώρημα von Neumann για $f = \mathbb{1}_B$. Τότε υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_B \circ T^k - P_F \mathbb{1}_B \right\|_2 < \varepsilon \qquad (\text{όπου } P_F \mathbb{1}_B = \tilde{\mathbb{1}}_B).$$

Τώρα, για αυθαίρετο $j \in \mathbb{N}$,

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_B \circ T^{k+j} - P_F \mathbb{1}_B \circ T^j \right\|_2 = \left\| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_B \circ T^k - P_F \mathbb{1}_B \right\|_2,$$

επειδή η T διατηρεί το μέτρο. Ακόμα $(P_F \mathbb{1}_B) \circ T^j = P_F \mathbb{1}_B$, αφού $P_F \mathbb{1}_B \in F$, άρα

$$\begin{aligned} &\left\| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_B \circ T^{k+j} - P_F \mathbb{1}_B \right\|_2 < \varepsilon \quad \Rightarrow \\ &\left\| \frac{1}{n} \sum_{k=j}^{n+j-1} \mathbb{1}_B \circ T^k - P_F \mathbb{1}_B \right\|_2 < \varepsilon. \end{aligned}$$

Από Cauchy–Schwarz,

$$\left| \left\langle \frac{1}{n} \sum_{k=j}^{n+j-1} f \circ T^k - P_F \mathbb{1}_B, \mathbb{1}_B \right\rangle \right| < \varepsilon \sqrt{\mu(B)} \leq \varepsilon.$$

Άρα,

$$\begin{aligned} &\left\langle \frac{1}{n} \sum_{k=j}^{n+j-1} f \circ T^k, \mathbb{1}_B \right\rangle - \langle P_F \mathbb{1}_B, \mathbb{1}_B \rangle > -\varepsilon \quad \Rightarrow \\ &\frac{1}{n} \sum_{k=j}^{n+j-1} \int f \circ T^k \mathbb{1}_B \, d\mu > -\varepsilon + \langle P_F \mathbb{1}_B, \mathbb{1}_B \rangle \quad \Rightarrow \\ &\frac{1}{n} \sum_{k=j}^{n+j-1} \int \mathbb{1}_{T^{-k}(B) \cap B} \, d\mu > \langle P_F \mathbb{1}_B, \mathbb{1}_B \rangle - \varepsilon = \|P_F \mathbb{1}_B\|^2 - \varepsilon \quad \Rightarrow \\ &\frac{1}{n} \sum_{k=j}^{n+j-1} \mu(T^{-k}(B) \cap B) > \|P_F \mathbb{1}_B\|^2 - \varepsilon. \end{aligned}$$

Παρατηρούμε τώρα ότι

$$\langle \mathbb{1}_B, \mathbb{1}_X \rangle = \langle P_F \mathbb{1}_B, \mathbb{1}_X \rangle$$

αφού $\mathbb{1}_X \in F$ και $(\mathbb{1}_B - P_F \mathbb{1}_B) \perp F$ και άρα $\langle \mathbb{1}_B - P_F \mathbb{1}_B, \mathbb{1}_X \rangle = 0$. Επομένως,

$$\mu(B)^2 = \langle \mathbb{1}_B, \mathbb{1}_X \rangle^2 = \langle P_F \mathbb{1}_B, \mathbb{1}_X \rangle^2 \leq \|P_F \mathbb{1}_B\|^2.$$

Έπεται ότι

$$\frac{1}{n} \sum_{k=j}^{n+j-1} \mu(T^{-k}(B) \cap B) \geq \mu(B)^2 - \varepsilon.$$

Αφού το παραπάνω ισχύει για κάθε $j \in \mathbb{N}$, έπεται ότι $\forall\, j \in \mathbb{N}$, το σύνολο $\mathbb{N} \cap [j, j+n-1]$ περιέχει τουλάχιστον ένα k για το οποίο

$$\mu(T^{-k}(B) \cap B) \geq \mu(B)^2 - \varepsilon.$$

Άρα το σύνολο

$$\{j \in \mathbb{N} \colon \mu(T^{-j}B \cap B) \geq \mu(B)^2 - \varepsilon\}$$

έχει κενά που φράσσονται από το n. □

Πόρισμα 2.2.7 (L^p-Μέσο Εργοδικό Θεώρημα) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ., $1 \leq p < \infty$ και $f \in L^p(X, \mathcal{B}, \mu)$. Τότε, υπάρχει $\tilde{f} \in L^p(X, \mathcal{B}, \mu)$ τέτοια ώστε

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k - \tilde{f} \right\|_p \to 0, \qquad \text{και} \qquad \tilde{f} \overset{L^2}{=} \tilde{f} \circ T.$$

Σημείωση. Εδώ, τουλάχιστον όταν $p = 1$, πρέπει $\mu(X) < +\infty$. Αντιπαράδειγμα: $X = \mathbb{R}$, $\mathcal{B} = \mathcal{B}(\mathbb{R})$, μ το μέτρο Lebesgue στο $\mathbb{R}$, $T(x) = x + 1$ και $f = \mathbb{1}_{[0,1)} \in L^1(X, \mathcal{B}, \mu)$. Τότε $f + f \circ T + \cdots + f \circ T^{n-1} = \mathbb{1}_{[0,1)} + \mathbb{1}_{T^{-1}([0,1))} + \cdots + \mathbb{1}_{T^{-n+1}([0,1))} = \mathbb{1}_{[-n+1,1)}$. Επομένως $n^{-1} \sum_{k=0}^{n-1} f \circ T^k \to 0$ σημειακά αλλά $\left\| n^{-1} \sum_{k=0}^{n-1} f \circ T^k \right\| = 1 \nrightarrow 0$.

Το Πόρισμα 2.2.7 μπορεί να αποδειχθεί από το θεώρημα von Neumann, απευθείας. Εδώ θα δοθεί, παρακάτω, μία απόδειξη που στηρίζεται στο κατά σημείο εργοδικό θεώρημα του Birkhoff, η οποία είναι ελάχιστα απλούστερη.

Κατά σημείο Εργοδικό Θεώρημα Birkhoff

Θεώρημα 2.2.8 (Κατά σημείο Εργοδικό Θεώρημα Birkhoff) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $f \in L^1(X, \mathcal{B}, \mu)$. Τότε υπάρχει $\tilde{f} \in L^1(X, \mathcal{B}, \mu)$ τέτοια ώστε

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \to \tilde{f} \qquad \mu\text{-σχεδόν παντού.}$$

Επιπλέον, $\tilde{f} = \tilde{f} \circ T$ μ-σχεδόν παντού, $\|\tilde{f}\|_1 \leq \|f\|$ και $\int_A \tilde{f}\, d\mu = \int_A f\, d\mu$ για κάθε $A \in \mathcal{B}$ με $A = T^{-1}(A)$.

Παρατηρήσεις.

1. Αν επιπλέον το σύστημα είναι εργοδικό, τότε η $\tilde{f}$ είναι σταθερή μ-σχεδόν παντού, αφού είναι αναλλοίωτη. Και επειδή $\int \tilde{f}\, d\mu = \int f\, d\mu$, έπεται ότι $\tilde{f} = \int f\, d\mu$ μ-σχεδόν παντού.
2. Το θεώρημα Birkhoff ισχύει και για συστήματα με $\mu(X) = \infty$. Στην περίπτωση αυτή, όταν το σύστημα είναι εργοδικό[11], θα πρέπει το όριο $\tilde{f}$ μιας $f \in L^1(X, \mathcal{B}, \mu)$ να είναι πάντα 0, αφού θα πρέπει $\tilde{f} \in L^1(X, \mathcal{B}, \mu)$ και $\tilde{f}$ σταθερή σχεδόν παντού, ως αναλλοίωτη.

Συμβολισμοί. Για $f \colon X \to \mathbb{C}$ συμβολίζουμε

[11] Όταν δηλαδή $\mu(A) = 0$ ή $\mu(A^c) = 0$ για κάθε $A \in \mathcal{B}$ με $A = T^{-1}(A)$.

- $S_n f = f + f \circ T + \cdots + f \circ T^{n-1}, \quad n \in \mathbb{N}, \qquad S_0 f(x) = 0 \quad \forall\, x \in X.$
- $S_n^* f = \max\{S_0 f, S_1 f, \ldots, S_n f\} = \max\{0, S_1 f, \ldots, S_n f\} \geq 0, \quad \forall n \in \mathbb{N}_0.$
- $A_n f = n^{-1} S_n f, \quad n \in \mathbb{N}_0.$
- $A_n^* f = \sup_{n \in \mathbb{N}_0} n^{-1} S_n f = \sup_{n \in \mathbb{N}_0} A_n f.$

Παρατηρούμε ότι

$$S_n(f \circ T) = (S_n f) \circ T \qquad \forall n \in \mathbb{N}_0$$

και

$$S_n f \circ T = S_{n+1} f - f \qquad \forall n \in \mathbb{N}_0.$$

Ορίζουμε επίσης

$$B_\alpha^f := \left\{ x \in X \colon \sup_{n \in \mathbb{N}} \frac{1}{n} S_n f(x) > \alpha \right\} \qquad \alpha \in \mathbb{R}.$$

Λήμμα 2.2.9 (Μεγιστική Ανισότητα) Για $f \in L^1(X, \mathcal{B}, \mu)$ ισχύει ότι

$$\alpha \mu(B_\alpha^f) \leq \int_{B_\alpha^f} f \, d\mu, \qquad \forall\, \alpha \in \mathbb{R}.$$

Σημείωση. Ειδικότερα, για $\alpha > 0$ έχει κανείς ότι

$$\mu(B_\alpha^f) \leq \frac{1}{\alpha} \int_{B_\alpha^f} f \, d\mu \leq \frac{1}{\alpha} \|f\|_1.$$

Λήμμα 2.2.10 (Μεγιστικό Εργοδικό Θεώρημα) Για $f \in L^1(X, \mathcal{B}, \mu, T)$ έχουμε ότι

$$\int_{\{S_n^* f > 0\}} f \, d\mu \geq 0.$$

Απόδειξη (Garcia). Έχουμε

$$\begin{aligned} S_n^* f \circ T &= \max\{S_0 f \circ T, S_1 f \circ T, \ldots, S_n f \circ T\} \\ &= \max\{S_1 f - f, S_2 f - f, \ldots, S_{n+1} f - f\} \\ &= \max\{S_1 f, S_2 f, \ldots, S_{n+1} f\} - f \\ &\geq \max\{S_1 f, S_2 f, \ldots, S_n f\} - f \\ &= \max\{S_0 f, S_1 f, \ldots, S_n f\} - f \qquad \text{στο σύνολο } \{x \in X \colon S_n^* f(x) > 0\}. \end{aligned}$$

Επομένως,

$$\begin{aligned} \int_{\{S_n^* f > 0\}} f \, d\mu &\geq \int_{\{S_n^* f > 0\}} (S_n^* f - S_n^* f \circ T) \, d\mu \\ &= \int_{\{S_n^* f > 0\}} S_n^* f \, d\mu - \int_{\{S_n^* f > 0\}} S_n^* f \circ T \, d\mu \\ &\overset{S_n^* f \geq 0}{=\!=} \int_X S_n^* f \, d\mu - \int_{\{S_n^* f > 0\}} S_n^* f \circ T \, d\mu \\ &\geq \int_X S_n^* f \, d\mu - \int_X S_n^* f \circ T \, d\mu \\ &= 0 \qquad (\mu \text{ είναι } T\text{-αναλλοίωτο}). \qquad \square \end{aligned}$$

Απόδειξη Λήμματος 2.2.9. Θέλουμε να δείξουμε ότι, για κάθε $\alpha \in \mathbb{R}$, $\alpha\mu(B_\alpha^f) \leq \int_{B_\alpha^f} f\,d\mu$, όπου $B_\alpha^f = \{x \in X\colon \sup_n n^{-1}S_nf(x) > \alpha\}$. Θέτουμε $g := f - \alpha\mathbb{1}_X$. Τότε

$$S_ng = \sum_{k=0}^{n-1} g \circ T^k = S_nf - n\alpha\mathbb{1}_X \quad \Rightarrow$$
$$\frac{1}{n}S_ng = \frac{1}{n}S_nf - \alpha\mathbb{1}_X.$$

Άρα

$$\frac{1}{n}S_nf > \alpha \quad \Leftrightarrow \quad \frac{1}{n}S_ng > 0 \quad \Leftrightarrow \quad S_ng > 0,$$

και επομένως

$$B_\alpha^f = \bigcup_{n\geq 1}\{x \in X\colon S_n^*g(x) > 0\} \qquad \text{και} \qquad \{x \in X\colon S_n^*g > 0\} \uparrow B_\alpha^f.$$

Από το Λήμμα 2.2.10 έχουμε ότι

$$\int_{\{S_n^*g>0\}} g\,d\mu \geq 0 \qquad \forall n \in \mathbb{N},$$

οπότε, παίρνοντας όρια όταν $n \to \infty$, από το θεώρημα κυριαρχημένης σύγκλισης έχουμε ότι

$$\int_{B_\alpha^f} g\,d\mu \geq 0.$$

Χρησιμοποιήθηκε εδώ ότι $f \in L^1(X, \mathcal{B}, \mu)$, και ότι $\mu(X) < +\infty$, τα οποία μαζί δίνουν ότι $g = f - \alpha\mathbb{1}_X \in L^1(X, \mathcal{B}, \mu)$. Τελικά,

$$\int_{B_\alpha^f}(f - \alpha\mathbb{1}_X)\,d\mu \geq 0 \quad \Rightarrow \quad \int_{B_\alpha^f} f\,d\mu \geq \alpha\mu(B_\alpha^f). \qquad \square$$

Πόρισμα 2.2.11 (Μεγιστική Ανισότητα 2) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ., $f \in L^1(X, \mathcal{B}, \mu)$ και $\alpha \in \mathbb{R}$. Τότε

$$\alpha\mu(B_\alpha^f \cap A) \leq \int_{B_\alpha^f \cap A} f\,d\mu,$$

για κάθε $A \in \mathcal{B}$ με $A = T^{-1}(A)$.

Απόδειξη. Υποθέτουμε χωρίς βλάβη της γενικότητας ότι $\mu(A) > 0$ και εφαρμόζουμε την προηγούμενη ανισότητα στο σύστημα $(A, \mathcal{B}_A, \mu_A, T_A)$ όπου $\mathcal{B}_A = \{A \cap B\colon B \in \mathcal{B}\}$, $T_A = T|_A\colon A \to A$, $\mu_A(B) = \mu(B)/\mu(A)$ για κάθε $B \in \mathcal{B}_A$. $\square$

Απόδειξη Θεωρήματος 2.2.8. Ορίζουμε

$$f^*(x) = \limsup_{n\to\infty} \frac{1}{n}S_nf(x) \quad \text{και} \quad f_*(x) = \liminf_{n\to\infty} \frac{1}{n}S_nf(x).$$

Οι f^*, f_* είναι αναλλοίωτες. Πράγματι, έχουμε ότι $S_nf\big(T(x)\big) = S_{n+1}f(x) - f(x)$, άρα

$$\frac{1}{n}S_nf\big(T(x)\big) = \frac{n+1}{n}\frac{1}{n+1}S_{n+1}f(x) - \frac{f(x)}{n}. \tag{6}$$

Για σταθεροποιημένο $x \in X$, έστω υπακολουθία $n_1 < n_2 < \cdots$ τέτοια ώστε το αριστερό μέλος της (6) τείνει στο $f^*\big(T(x)\big)$. Τότε το δεξί μέλος έχει αναγκαστικά όριο και

$$\lim_{k\to\infty} \frac{n_k+1}{n_k}\frac{1}{n_k+1}S_{n_k+1}f(x) - \frac{1}{n_k}f(x) = f^*\big(T(x)\big) \quad \Rightarrow \quad \frac{1}{n_k+1}S_{n_k+1}f(x) \to f^*\big(T(x)\big)$$

και αφού

$$\lim_{k\to\infty} \frac{1}{n_k+1}S_{n_k+1}f(x) \leq f^*(x),$$

έπεται ότι

$$f^*\big(T(x)\big) \leq f^*(x).$$

Για την αντίστροφη ανισότητα παίρνουμε υπακολουθία $m_1 < m_2 < \cdots$ τέτοια ώστε το δεξί μέλος της (6), και συγκεκριμένα το

$$\frac{1}{m_k+1} S_{m_k+1} f(x)$$

να τείνει στο $f^*(x)$. Τότε το αριστερό μέλος τείνει αναγκαστικά στο ίδιο όριο $f^*(x)$:

$$\frac{1}{m_k} S_{m_k} f\big(T(x)\big) \xrightarrow{(k\to\infty)} f^*(x).$$

Αφού αναγκαστικά $\lim_{k\to\infty} m_k^{-1} S_{m_k} f\big(T(x)\big) \leq f^*\big(T(x)\big)$, παίρνει κανείς ότι $f^*(x) \leq f^*\big(T(x)\big)$. Ομοίως βλέπει κανείς και ότι η f_* είναι αναλλοίωτη.

Ορίζουμε για $a < \beta$, $a, \beta \in \mathbb{R}$, τα σύνολα

$$E^f_{a,\beta} = \{x \in X \colon f_*(x) < a < \beta < f^*(x)\}$$

και θα δείξουμε ότι $\mu(E^f_{a,\beta}) = 0 \,\forall\, a, \beta \in \mathbb{R}$ με $a < \beta$. Είναι σαφές ότι εφόσον οι f_*, f^* είναι αναλλοίωτες, το $E^f_{a,\beta}$ είναι T-αναλλοίωτο. Δηλαδή

(1) $T^{-1}(E^f_{a,\beta}) = E^f_{a,\beta}$.

(2) $E^{-f}_{-\beta,-a} = E^f_{a,\beta}$.

(3) $E^f_{a,\beta} \subset B^f_\beta$.

Η απόδειξη της (2) είναι άμεση αφού $(-f)_* = -f^*$ και $(-f)^* = -f_*$.

Η απόδειξη του (3): αν $x \in E^f_{a,\beta}$, τότε $\limsup_{n\to\infty} n^{-1} S_n f(x) > \beta$, άρα $\sup_{n\in\mathbb{N}} n^{-1} S_n f(x) > \beta$, επομένως $x \in B^f_\beta$, εξ' ορισμού.

Λόγω της (1), μπορούμε να εφαρμόσουμε την Μεγιστική Ανισότητα 2 και έχουμε ότι

$$\beta\mu(E^f_{a,\beta}) \overset{(3)}{=} \beta\mu(B^f_\beta \cap E^f_{a,\beta}) \leq \int_{B^f_\beta \cap E^f_{a,\beta}} f\, d\mu \overset{(3)}{=} \int_{E^f_{a,\beta}} f\, d\mu. \tag{7}$$

Εφαρμόζοντας την (7) για την $-f$ και με $-a, -\beta$ στις θέσεις των β, a αντίστοιχα, έχουμε λόγω της (2) ότι

$$\begin{aligned} & -a\mu(E^{-f}_{-\beta,-a}) \leq \int_{E^{-f}_{-\beta,-a}} (-f)\, d\mu \\ \Leftrightarrow \quad & -a\mu(E^f_{a,\beta}) \leq -\int_{E^f_{a,\beta}} f\, d\mu \end{aligned}$$

και επομένως,

$$a\mu(E^f_{a,\beta}) \geq \int_{E^f_{a,\beta}} f\, d\mu. \tag{8}$$

Συνδυάζοντας τις (7) και (8) παίρνουμε

$$\beta\mu(E^f_{a,\beta}) \leq a\mu(E^f_{a,\beta})$$

το οποίο, εάν $a < \beta$ και $\mu(E^f_{a,\beta}) > 0$, είναι άτοπο. Άρα θα πρέπει

$$\mu(E^f_{a,\beta}) = 0.$$

Τώρα, προφανώς ισχύει ότι $f_* \leq f^*$ και

$$\mu\big(\{x \in X \colon f_*(x) < f^*(x)\}\big) = \mu\left(\bigcup_{\substack{a,\beta\in\mathbb{Q}\\ a<\beta}} E^f_{a,\beta}\right) \leq \sum_{\substack{a,\beta\in\mathbb{Q}\\ a<\beta}} \mu(E^f_{a,\beta}) = 0,$$

άρα έπεται ότι το όριο $\lim_{n\to\infty} n^{-1}S_n f(x) = \tilde{f}(x)$ υπάρχει μ-σχεδόν παντού και $\tilde{f} = f^* = f_*$ μ-σχεδόν παντού. Επίσης, $\tilde{f} = \tilde{f} \circ T$ μ-σχεδόν παντού γιατί $\tilde{f} = f^*$ μ-σχεδόν παντού και $f^* = f^* \circ T$.

Ισχύει ότι $\tilde{f} \in L^1(X, \mathcal{B}, \mu)$ και $\|\tilde{f}\|_1 \leq \|f\|_1 < \infty$. Πράγματι,

$$
\begin{aligned}
\|\tilde{f}\|_1 = \int \liminf_{n\to\infty} |n^{-1}S_n f\| \, d\mu &\leq \liminf_{n\to\infty} \int n^{-1}|S_n f| \, d\mu \\
&\leq \liminf_{n\to\infty} n^{-1} \int S_n |f| \, d\mu \\
&= \int |f| \, d\mu \\
&= \|f\|_1,
\end{aligned}
$$

όπου στην πρώτη ανισότητα χρησιμοποιήθηκε το Λήμμα του Fatou, στην επόμενη η τριγωνική ανισότητα και το ότι $|f \circ T^k| = |f| \circ T^k$ και τελικά το γεγονός ότι ο T διατηρεί το μέτρο.

Τώρα θα δείξουμε ότι

$$\int_A f \, d\mu = \int_A \tilde{f} \, d\mu \qquad \forall A \in \mathcal{B} \quad \text{με} \quad A = T^{-1}(A).$$

Αυτό αρκεί να το δείξουμε για $A = X$. Για τυχόν A μετά, θεωρούμε τον περιορισμό του συστήματος στο A, δηλαδή το σύστημα $(A, \mathcal{B}_A, \mu_A, T_A)$ και την ίδια f περιορισμένη στο A.

Ορίζουμε

$$D^f_{k,n} = \left\{ x \in X \colon \frac{k}{n} \leq f^*(x) < \frac{k+1}{n} \right\}, \qquad k \in \mathbb{Z},\ n \in \mathbb{N}.$$

Παρατηρούμε ότι

(1) $D^f_{k,n} = T^{-1}(D^f_{k,n}) \quad \forall\, k \in \mathbb{Z},\ n \in \mathbb{N}$.

(2) $D^f_{k,n} \subset B^f_{\frac{k}{n}-\varepsilon} \quad \forall\, k \in \mathbb{Z},\ n \in \mathbb{N},\ \varepsilon > 0$.

Το (1) ισχύει επειδή η f^* είναι αναλλοίωτη ($f^* = f^* \circ T$) και το (2) επειδή αν $x \in D^f_{k,n}$ και $\varepsilon > 0$, έχουμε ότι $f^*(x) \geq k/n > k/n - \varepsilon \Rightarrow x \in B^f_{\frac{k}{n}-\varepsilon}$.

Εφαρμόζοντας την Μεγιστική Ανισότητα 2, παίρνει κανείς ότι

$$\int_{D^f_{k,n}} f \, d\mu = \int_{D^f_{k,n} \cap B^f_{\frac{k}{n}-\varepsilon}} f \, d\mu \geq \left(\frac{k}{n} - \varepsilon\right) \mu\big(D^f_{k,n} \cap B^f_{\frac{k}{n}-\varepsilon}\big) = \left(\frac{k}{n} - \varepsilon\right) \mu\big(D^f_{k,n}\big) \qquad \forall \varepsilon > 0.$$

Άρα,

$$\int_{D^f_{k,n}} f \, d\mu \geq \frac{k}{n}\mu(D^f_{k,n}) \geq \int_{D^f_{k,n}} \left(f^* - \frac{1}{n}\right) d\mu = \int_{D^f_{k,n}} f^* \, d\mu - \frac{1}{n}\mu(D^f_{k,n}) \qquad \forall\, k \in \mathbb{Z}\, \forall\, n \in \mathbb{N}.$$

Αθροίζοντας ως προς όλα τα $k \in \mathbb{Z}$, παίρνει κανείς ότι

$$
\begin{aligned}
\sum_{k\in\mathbb{Z}} \int_{D^f_{k,n}} f \, d\mu &\geq \sum_{k\in\mathbb{Z}} \int_{D^f_{k,n}} f^* \, d\mu - \frac{1}{n} \sum_{k\in\mathbb{Z}} \mu(D^f_{k,n}) \\
\int_X f \, d\mu &\geq \int_X f^* \, d\mu - \frac{1}{n}\mu\left(\bigcup_{k\in\mathbb{Z}} D^f_{k,n}\right) \qquad \forall\, n \in \mathbb{N},
\end{aligned}
$$

εφόσον $\bigcup_{k\in\mathbb{Z}} D^f_{k,n} = X$, επειδή $\tilde{f} \in L^1$, και τα $D^f_{k,n}$ είναι ξένα ανά δύο για διαφορετικά k και σταθερό n. Άρα $\mu\left(\bigcup_{k\in\mathbb{Z}} D^f_{k,n}\right) = 1$, οπότε τελικά

$$\int_X f \, d\mu \geq \int_X f^* \, d\mu = \int_X \tilde{f} \, d\mu. \tag{9}$$

Εφαρμόζοντας την ανισότητα (9) και στην $-f$, έχει κανείς ότι

$$\int_X -f\,d\mu \geq \int_X (-f)^*\,d\mu = -\int_X f_*\,d\mu = -\int_X \tilde{f}\,d\mu.$$

Τελικά

$$\int_X f\,d\mu = \int_X \tilde{f}\,d\mu. \qquad \square$$

Εφαρμογές του Εργοδικού Θεωρήματος

Πρόταση 2.2.12 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Το σύστημα είναι εργοδικό ανν

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu\big(T^{-k}(A) \cap B\big) \to \mu(A)\mu(B) \qquad \forall\, A, B \in \mathcal{B}.$$

Απόδειξη. Έστω ότι το σύστημα είναι εργοδικό και έστω $A, B \in \mathcal{B}$. Από το εργοδικό θεώρημα του Birkhoff, εφαρμοσμένο στην $\mathbb{1}_A$, έχουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A \circ T^k \to \int \mathbb{1}_A\,d\mu = \mu(A) \qquad \mu - \text{σχεδόν παντού}.$$

Επομένως, πολλαπλασιάζοντας με $\mathbb{1}_B$,

$$\frac{1}{n}\sum_{k=0}^{n-1} (\mathbb{1}_A \circ T^k)\,\mathbb{1}_B \to \mu(A)\mathbb{1}_B \qquad \mu - \text{σχεδόν παντού}.$$

Το αριστερό μέλος είναι φραγμένο από το 1, για κάθε $x \in X$. Άρα, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης, παίρνουμε ότι

$$\int \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{T^{-k}(A)}\,\mathbb{1}_B\,d\mu \to \mu(A)\int \mathbb{1}_B\,d\mu$$

και τελικά ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu\big(T^{-k}(A) \cap B\big) \to \mu(A)\mu(B).$$

Αντίστροφα, για να δείξουμε την εργοδικότητα, έστω $A \in \mathcal{B}$ με $A = T^{-1}(A)$. Εφαρμόζουμε την δοθείσα σχέση για $B = X \setminus A$. Τότε

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu\big(T^{-k}(A) \cap (X \setminus A)\big) \to \mu(A)\mu(X \setminus A)$$

Από την υπόθεση έχουμε ότι $T^{-k}(A) = A$ για κάθε k, επομένως το αριστερό μέλος είναι πάντα 0. Δηλαδή

$$\mu(A)\mu(X \setminus A) = 0 \quad \Rightarrow \quad \mu(A) \in \{0, 1\}. \qquad \square$$

2ος τρόπος απόδειξης της εργοδικότητας. Έστω $A, B \in \mathcal{B}$ με $\mu(A)\mu(B) > 0$. Εφαρμόζοντας την δοθείσα σχέση και εφόσον το όριο $\mu(A)\mu(B) > 0$, έχουμε ότι θα πρέπει για κάποιο n

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) > 0,$$

επομένως θα πρέπει για κάποιο $k \geq 1$

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

Άρα το σύστημα είναι εργοδικό από την Πρόταση 2.1.4. $\square$

Θεώρημα 2.2.13 (L^p-Εργοδικό Θεώρημα) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ., $1 \le p < \infty$ και $f \in L^p(X, \mathcal{B}, \mu)$. Τότε υπάρχει $\tilde{f} \in L^p$ τέτοια ώστε

$$\left\|n^{-1} S_n f - \tilde{f}\right\|_p \to 0 \qquad \text{και} \qquad \tilde{f} \overset{L^p}{=} \tilde{f} \circ T$$

Απόδειξη. Έστω $g \in L^\infty(X, \mathcal{B}, \mu)$. Τότε $g \in L^1(X, \mathcal{B}, \mu)$ και από θεώρημα Birkhoff έχουμε ότι υπάρχει $\tilde{g} \in L^1$ τέτοια ώστε

$$\frac{1}{n} S_n g \to \tilde{g} \qquad \mu - \text{σχεδόν παντού.}$$

Επειδή $g \in L^\infty$, έπεται ότι και $\tilde{g} \in L^\infty$ και επομένως

$$\left\|\frac{1}{n} S_n g - \tilde{g}\right\|_\infty \le 2\|g\|_\infty \qquad \text{(τριγωνική ανισότητα)}$$

και άρα εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης,

$$\int |n^{-1} S_n g - \tilde{g}|^p \, d\mu \to 0 \quad \Leftrightarrow \quad \|n^{-1} S_n g - \tilde{g}\|_p \to 0,$$

για αυθαίρετο $p \in [1, +\infty)$. Για την γενική περίπτωση, έστω $f \in L^p(X, \mathcal{B}, \mu)$, όπου $1 \le p < \infty$. Τότε υπάρχει[12] $g \in L^\infty(X, \mathcal{B}, \mu)$ τέτοια ώστε

$$\|f - g\|_p < \varepsilon/3,$$

για δοθέν $\varepsilon > 0$. Τώρα, για την g γνωρίζουμε ότι $n^{-1} S_n g$ συγκλίνουν στον L^p, άρα η ακολουθία $\left(n^{-1} S_n g\right)_{n \in \mathbb{N}}$ είναι βασική στον L^p. Επομένως υπάρχει $n_\varepsilon \in \mathbb{N}$ έτσι ώστε

$$\left\|\frac{1}{n} S_n g - \frac{1}{m} S_m g\right\|_p < \varepsilon/3 \qquad \forall\, n, m \ge n_\varepsilon.$$

Επίσης, για κάθε $n \in \mathbb{N}$, έχουμε ότι

$$\begin{aligned}
\left\|\frac{1}{n} S_n g - \frac{1}{n} S_n f\right\|_p &= \frac{1}{n} \left\|\sum_{k=0}^{n-1} (f - g) \circ T^k\right\|_p \\
&\le \frac{1}{n} \sum_{k=0}^{n-1} \left\|(f - g) \circ T^k\right\|_p \\
&= \frac{1}{n} \sum_{k=0}^{n-1} \left\|f - g\right\|_p \\
&= \left\|f - g\right\|_p,
\end{aligned}$$

όπου η προτελευταία ισότητα ισχύει επειδή ο τελεστής Koopman είναι ισομετρία. Επομένως,

$$\begin{aligned}
\left\|\frac{1}{n} S_n f - \frac{1}{m} S_m f\right\|_p &\le \left\|\frac{1}{n} S_n f - \frac{1}{n} S_n g\right\|_p + \left\|\frac{1}{n} S_n g - \frac{1}{m} S_m g\right\|_p + \left\|\frac{1}{m} S_m g - \frac{1}{m} S_m f\right\|_p \\
&\le 2\|f - g\|_p + \left\|\frac{1}{n} S_n g - \frac{1}{m} S_m g\right\|_p \\
&< 2\varepsilon/3 + \varepsilon/3 = \varepsilon,
\end{aligned}$$

για όλα τα $n, m \ge n_\varepsilon$. Αφού αυτό γίνεται για κάθε $\varepsilon > 0$, έπεται ότι η $\left(n^{-1} S_n f\right)_{n \in \mathbb{N}}$ είναι βασική στον $L^p(X, \mathcal{B}, \mu)$ και άρα συγκλίνει στον $L^p(X, \mathcal{B}, \mu)$. Δηλαδή υπάρχει $\tilde{f} \in L^p(X, \mathcal{B}, \mu)$ έτσι ώστε

$$\tilde{f} \overset{L^p}{=} \lim_n \frac{1}{n} S_n f.$$

Για την $\tilde{f}$ έχουμε ακόμα ότι $\tilde{f} \overset{L^p}{=} \tilde{f} \circ T$. Πράγματι,

$$\frac{1}{n} S_n f \circ T = \frac{n+1}{n} \frac{1}{n+1} S_{n+1} f - \frac{1}{n} f$$

και έχουμε το συμπέρασμα παίρνοντας όρια, εφόσον $n^{-1} f \to 0$. $\square$

[12] Αν ορίσουμε $f_n(x) = f(x)\, \mathbb{1}_{[-n,n]}\big(f(x)\big)$ για $n \in \mathbb{N}$, τότε $\|f_n\|_\infty \le n$ και $\int |f_n - f|^p \, d\mu = \int |f|^p \mathbb{1}_{(n,\infty)}\big(|f|\big) \, d\mu \to 0$, από θεώρημα φραγμένης σύγκλισης. Άρα διαλέγουμε για g μια f_n για αρκούντως μεγάλο n.

Εφαρμογή του Εργοδικού Θεωρήματος στους κανονικούς αριθμούς. Για $x \in [0,1)$, υπάρχει μοναδική ακολουθία $i_1(x), i_2(x), \dots$ τέτοια ώστε

$$x = \sum_{n\geq 1} \frac{i_n(x)}{2^n}, \qquad i_n(x) \in \{0,1\}\ \forall\, n \in \mathbb{N} \quad \text{και η ακολουθία } (i_n(x))_{n\in\mathbb{N}} \text{ δεν τελειώνει με όλο } 1.$$

Ορισμός 2.2.14 (κανονικός αριθμός) Κανονικό αριθμό λέμε έναν x για τον οποίο

$$\frac{1}{n}\big|\{k \leq n \colon i_k(x) = 0\}\big| \to \tfrac{1}{2}.$$

Συμβολισμός: $|A|$ συμβολίζει τον πληθάριθμο ενός συνόλου A.

Θεώρημα 2.2.15 (Θεώρημα κανονικών αριθμών Borel) Σχεδόν κάθε αριθμός στο $[0,1)$ είναι κανονικός.

Απόδειξη. Θεωρούμε το $(X, \mathcal{B}, \mu, T)$ όπου $X = [0,1)$, $\mathcal{B} = \mathcal{B}([0,1))$, μ είναι το μέτρο Lebesgue στο $[0,1)$ και $T_2(x) = 2x \pmod 1$. Παρατηρούμε ότι αν $i_1(x) = 1$, τότε $x = \sum_{n\geq 1} i_n(x)2^{-n} \geq \frac{1}{2}$, και αν $i_1(x) = 0$, τότε $x = \sum_{n\geq 1} i_n(x)2^{-n} < \frac{1}{2}$ (η τελευταία ανισότητα είναι γνήσια επειδή η ακολουθία $(i_n(x))_{n\in\mathbb{N}}$ δεν τελειώνει σε όλο 1). Άρα

$$i_1(x) = 0 \Leftrightarrow x \in [0, 1/2)$$

Τώρα,

$$T_2(x) = 2x = 2\sum_{n\geq 1} \frac{i_n(x)}{2^n} = \sum_{n\geq 2} \frac{i_n(x)}{2^{n-1}} = \sum_{n\geq 1} \frac{i_{n+1}(x)}{2^n},$$

αν $x \in [0,1/2) \Leftrightarrow i_1(x) = 0$, και

$$T_2(x) = 2x - 1 = 2\sum_{n\geq 1} \frac{i_n(x)}{2^n} - 1 = i_1(x) + \sum_{n\geq 2} \frac{i_n(x)}{2^{n-1}} - 1 = \sum_{n\geq 2} \frac{i_n(x)}{2^{n-1}} = \sum_{n\geq 1} \frac{i_{n+1}(x)}{2^n},$$

αν $x \in [1/2, 1) \Leftrightarrow i_1(x) = 1$. Άρα

$$i_n(T_2(x)) = i_{n+1}(x) \qquad \forall\, n \in \mathbb{N}\ \forall\, x \in [0,1)$$

και επαγωγικά,

$$i_n(T_2^k(x)) = i_{n+k}(x) \qquad \forall\, n, k \in \mathbb{N}\ \forall\, x \in [0,1).$$

Επομένως

$$\begin{aligned}
\frac{1}{n}\big|\{k \leq n \colon i_k(x) = 0\}\big| &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{0\}}\big(i_k(x)\big) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{0\}}\big(i_1\big(T_2^{k-1}(x)\big)\big) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{[0,1/2)}\big(T_2^k(x)\big) = \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{[0,1/2)} \circ T_2^k(x)
\end{aligned}$$

και από το εργοδικό θεώρημα Birkhoff και επειδή το σύστημα είναι εργοδικό, αυτό συγκλίνει στο

$$\int_{[0,1)} \mathbb{1}_{[0,1/2)}(x)\, dx = \tfrac{1}{2}$$

για σχεδόν κάθε $x \in [0,1)$. □

Το ίδιο επιχείρημα δίνει ανάλογα αποτελέσματα για αναπαράσταση σε κάθε k-δικό σύστημα.

Εφαρμογή στα συνεχή κλάσματα. Κάθε $x \in (0,1] \setminus \mathbb{Q}$ έχει μοναδική αναπαράσταση σε συνεχή κλάσματα:

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{a_3(x) + \cdots}}}, \qquad \text{με} \quad a_i(x) \in \mathbb{N} \quad \forall\, i \in \mathbb{N}.$$

Αν $T(x) = 1/x - \lfloor 1/x \rfloor$ η απεικόνιση Gauss, έχουμε

$$a_k(x) = \Big\lfloor \frac{1}{T^{k-1}(x)} \Big\rfloor \qquad \forall\, k \in \mathbb{N},$$

και επομένως

$$a_1\big(T^{k-1}(x)\big) = a_k(x) \qquad \forall\, k \in \mathbb{N}.$$

Τώρα,

$$\begin{aligned} \frac{1}{n}\big|\{k \le n \colon a_k(x) = j\}\big| &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{j\}}\big(a_k(x)\big) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{j\}}\big(a_1\big(T^{k-1}(x)\big)\big) \\ &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{(\frac{1}{j+1},\frac{1}{j}]\setminus\mathbb{Q}} \circ T^{k-1}(x), \end{aligned}$$

αφού $a_1(y) = \lfloor 1/y \rfloor$, άρα

$$a_1(y) = j \quad \Leftrightarrow \quad \frac{1}{j+1} < y \le \frac{1}{j}.$$

Από το εργοδικό θεώρημα αυτό έχει όριο το

$$\int_0^1 \mathbb{1}_{(\frac{1}{j+1},\frac{1}{j}]\setminus\mathbb{Q}}(x)\frac{dx}{x+1}\frac{1}{\ln 2} = \frac{1}{\ln 2}\ln\left(\frac{1/j+1}{1/(j+1)+1}\right) = \frac{1}{\ln 2}\ln\left(\frac{(j+1)^2}{j(j+2)}\right)$$

μ-σχεδόν παντού, όπου μ το μέτρο με παράγωγο Radon–Nikodym $d\mu(x)/dx = (x+1)^{-1}/\ln 2$, και άρα τελικά και Lebesgue-σχεδόν παντού.

Πρόταση 2.2.16 Για σχεδόν κάθε x, ισχύει ότι

$$[a_1(x)\cdots a_n(x)]^{1/n} \to \prod_{j=1}^{\infty}\left(\frac{(j+1)^2}{j(j+2)}\right)^{\log_2 j}$$

Απόδειξη. Θέλουμε να δείξουμε ότι

$$\frac{1}{n}\sum_{k=1}^{n}\ln a_k(x) \to \sum_{j\ge 1}\ln\left(\frac{(j+1)^2}{j(j+2)}\right)\frac{\ln j}{\ln 2},$$

για σχεδόν κάθε x στο $(0,1]\setminus\mathbb{Q}$. Παίρνουμε για $f(x) = \ln a_1(x)$ και έχουμε

$$\frac{1}{n}\sum_{k=1}^{n}\ln a_k(x) = \frac{1}{n}\sum_{k=1}^{n} f\circ T^{k-1}(x)$$

και από το εργοδικό θεώρημα αυτό έχει σχεδόν παντού όριο το

$$\int_0^1 f(x)\frac{dx}{x+1}\frac{1}{\ln 2} = \int_0^1 \ln\big(a_1(x)\big)\frac{dx}{x+1}\frac{1}{\ln 2} = \sum_{j\ge 1}\int_{1/(j+1)}^{1/j}\ln j\,\frac{dx}{x+1}\frac{1}{\ln 2} = \sum_{j\ge 1}\ln\left(\frac{(j+1)^2}{j(j+2)}\right)\frac{\ln j}{\ln 2}. \quad \square$$

Ισχυρός νόμος των μεγάλων αριθμών. Θεωρούμε μια στοχαστική διαδικασία $\xi = (\xi_1, \xi_2, \dots)$ σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$, όπου κάθε $\xi_i \colon \Omega \to \mathbb{R}$ είναι Borel μετρήσιμη συνάρτηση, δηλαδή τυχαία μεταβλητή. Απαιτούμε να ικανοποιούνται οι

$$P\big((\xi_1, \xi_2, \dots, \xi_n) \in B\big) = P\big((\xi_2, \xi_3, \dots, \xi_{n+1}) \in B\big) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^n)\ \forall\, n \in \mathbb{N}. \tag{10}$$

Δηλαδή να έχουμε μια στάσιμη στοχαστική διαδικασία, όπου η κατανομή της να είναι αμετάβλητη στον χρόνο. Θεωρούμε τον χώρο $X = \mathbb{R}^{\mathbb{N}}$ και $T \colon X \to X$ τον μετασχηματισμό shift. Η σ-άλγεβρα $\mathcal{B}$ θα είναι η σ-άλγεβρα γινόμενο, που παράγεται απ' όλους τους κυλίνδρους[13]. Αν θεωρήσουμε την απεικόνιση

$$\underset{\sim}{\xi} \colon \Omega \to X, \qquad \underset{\sim}{\xi}(\omega) = \big(\xi_1(\omega), \xi_2(\omega), \dots\big),$$

τότε η απεικόνιση αυτή είναι μετρήσιμη. Πράγματι,

$$\begin{aligned} &\underset{\sim}{\xi}^{-1}\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_{i_1} \in B_{i_1}, \dots, x_{i_n} \in B_{i_n}\}\big) \\ &\qquad\qquad = \{\omega \in \Omega \colon \xi_{i_1}(\omega) \in B_{i_1}, \dots, \xi_{i_n}(\omega) \in B_{i_n}\} = \bigcap_{j=1}^{n} \{\omega \in \Omega \colon \xi_{i_j}(\omega) \in B_{i_j}\} \in \mathcal{F}, \end{aligned}$$

για έναν κύλινδρο $\prod_{i \geq 1} B_i$, όπου $B_i \in \mathcal{B}(\mathbb{R})\ \forall\, i \in \mathbb{N}$ και $B_i = \mathbb{R}$ για $i \notin \{i_1, \dots, i_n\}$. Άρα ορίζεται το μέτρο $\mu = \xi_* P$ στον μετρήσιμο χώρο $(X, \mathcal{B})$, δηλαδή το μέτρο με

$$\mu(B) = P\big(\{\omega \in \Omega \colon \underset{\sim}{\xi}(\omega) \in B\}\big).$$

Από την (10) βλέπουμε ότι το μέτρο αυτό είναι shift-αναλλοίωτο αφού αρκεί να ισχύει η

$$\mu(T^{-1}(B)) = \mu(B) \tag{11}$$

για κυλίνδρους της μορφής

$$B = \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 \in B_1, x_2 \in B_2, \dots, x_n \in B_n\}$$

και για τέτοια σύνολα οι (10) και (11) είναι ίδιες.

Από τα παραπάνω έπεται ότι το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι σ.δ.μ. και από το εργοδικό θεώρημα παίρνουμε ότι, για κάθε $f \in L^1(X, \mathcal{B}, \mu)$, υπάρχει $\tilde{f} \in L^1(X, \mathcal{B}, \mu)$ τέτοια ώστε

$$\frac{1}{n} \sum_{k=1}^{n} f(\xi_k, \xi_{k+1}, \dots) \to \tilde{f}(\xi_1, \xi_2, \dots) \qquad P\text{-σχεδόν παντού.}$$

Αυτός είναι ο νόμος των μεγάλων αριθμών για στάσιμες στοχαστικές διαδικασίες.

Ειδική περίπτωση. Οι $\xi_1, \xi_2, \dots$ είναι ανεξάρτητες και ισόνομες και έστω

$$\nu(B) = P(\xi_i \in B) \quad \forall\, i \in \mathbb{N}.$$

Τότε το μ είναι μέτρο γινόμενο, δηλαδή $\mu = \nu \otimes \nu \otimes \cdots$, δηλαδή

$$\mu(\{x = (x_1, x_2, \dots) \in X \colon x_1 \in B_1, \dots, x_n \in B_n\}) = \nu(B_1) \cdots \nu(B_n)$$

για κάθε $n \in \mathbb{N}$ και $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$.

Πρόταση 2.2.17 Στην περίπτωση αυτή το σύστημα είναι εργοδικό.

Απόδειξη. Έστω

$$\mathcal{I} = \{A \in \mathcal{B} \colon A = T^{-1}(A)\}$$

η σ-άλγεβρα των αναλλοίωτων υποσυνόλων. Θεωρούμε και την τελική σ-άλγεβρα (tail σ-algebra):

$$\mathcal{T} = \bigcap_{n \in \mathbb{N}} T^{-n}\mathcal{B},$$

[13] Δηλαδή σύνολα $\prod_{n \geq 1} B_n$, όπου $B_n \in \mathcal{B}(\mathbb{R})\ \forall\, n \in \mathbb{N}$ και $B_n \neq \mathbb{R}$ μόνο για πεπερασμένο πλήθος $n \in \mathbb{N}$.

όπου $T^{-n}\mathcal{B} = \{T^{-n}(B) \colon B \in \mathcal{B}\}$. Αν $A \in \mathcal{T}$, $A \in T^{-n}\mathcal{B}$ για κάθε $n \in \mathbb{N}$, δηλαδή, για κάθε $n \in \mathbb{N}$ υπάρχει $A_n \in \mathcal{B}$ τέτοιο ώστε $A = T^{-n}(A_n)$ και τότε

$$\begin{aligned} \underset{\sim}{\xi}^{-1}(A) = \{\omega \in \Omega \colon \underset{\sim}{\xi}(\omega) \in A\} = \{\omega \in \Omega \colon \underset{\sim}{\xi}(\omega) \in T^{-n}(A_n)\} = \{\omega \in \Omega \colon T^n(\underset{\sim}{\xi}(\omega)) \in A_n\} \\ = \{\omega \in \Omega \colon (\xi_n(\omega), \xi_{n+1}(\omega) \dots) \in A_n\} \in \sigma(\xi_n, \xi_{n+1}, \dots) \end{aligned}$$

για κάθε $n \in \mathbb{N}$, οπότε $\underset{\sim}{\xi}^{-1}(A) \in \bigcap_{n\in\mathbb{N}} \sigma(\xi_n, \xi_{n+1}, \dots)$. Άρα για κάθε $A \in \mathcal{T}$,

$$\mu(A) = P\big(\underset{\sim}{\xi}^{-1}(A)\big) \in \{0, 1\}$$

από το νόμο 0-1 του Kolmogorov. Όμως είναι σαφές ότι $\mathcal{I} \subset \mathcal{T}$, άρα $\mu(A) \in \{0,1\}\ \forall A \in \mathcal{I}$. $\square$

Τώρα, το εργοδικό θεώρημα δίνει ότι, για $f \in L^1(X, \mathcal{B}, \mu)$, με P-πιθανότητα ίση με ένα,

$$\frac{1}{n} \sum_{k=1}^{n} f(\xi_k, \xi_{k+1}, \dots) \to \int f \, d\mu.$$

Για $f(\underset{\sim}{x}) = x_1$, όπου $\underset{\sim}{x} = (x_1, x_2, \dots)$, παίρνουμε τον γνωστό (ισχυρό) νόμο των μεγάλων αριθμών: αν $E(|\xi_1|) < +\infty$, δηλαδή αν $\xi_1 \in L^1(\Omega, \mathcal{F}, P)$, τότε επειδή

$$E(|\xi_1|) = \int_\Omega |\xi_1| \, dP = \int_{\mathbb{R}} |x| \, d\nu(x) = \int_X |x_1| \, d\mu(\underset{\sim}{x}) = \int_X |f| \, d\mu(\underset{\sim}{x}),$$

έχουμε ότι $f \in L^1(X, \mathcal{B}, \mu)$, και άρα

$$\frac{1}{n} \sum_{k=1}^{n} \xi_k \to \int x_1 \, d\mu(\underset{\sim}{x}) = E(\xi_1).$$

Εργοδικότητα των Markov Shifts. Έστω $S = \{1, 2, \dots, s\}$ πεπερασμένο σύνολο, $X = S^{\mathbb{N}_0}$, $\mathcal{B} = \otimes_{\mathbb{N}_0} \mathcal{P}(S)$, $T \colon X \to X$ ο μετασχηματισμός shift και για το μέτρο: αν $P = (P_{ij})_{i,j\in S}$ στοχαστικός πίνακας, $p = (p_i)_{i\in S}$ διάνυσμα πιθανότητας με $pP = p$, δηλαδή p είναι αριστερό ιδιοδιάνυσμα του P που αντιστοιχεί στην ιδιοτιμή 1, ορίζεται μονοσήμαντα ένα μέτρο

$$\mu(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i_0, x_1 = i_1, \dots, x_n = i_n\}) = p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n},$$

ώστε το σύστημα $(X, \mathcal{B}, \mu, T)$ να είναι ένα σ.δ.μ.

Ορισμός 2.2.18 Ο πίνακας P είναι *ανάγωγος* αν, για κάθε $i, j \in S$, υπάρχει $m_{ij} \in \mathbb{N}$, έτσι ώστε $P_{ij}^{m_{ij}} > 0$.

Θεώρημα Έστω ότι $p_i > 0\ \forall i \in S$. Τότε το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι εργοδικό ανν ο πίνακας μετάβασης P είναι ανάγωγος.

Πριν την απόδειξη του Θεωρήματος θα αποδείξουμε το εξής Λήμμα.

Λήμμα 2.2.20 Υποθέτουμε ότι για το διάνυσμα $p = (p_1, \dots, p_s)$ έχουμε $p_i > 0\ \forall i \in S$. Τότε

(i) $\lim_{n\to\infty} \sum_{k=0}^{n-1} P_{ij}^k = Q_{ij}$ υπάρχει για κάθε $i, j \in S$.

(ii) $QP = PQ = Q^2 = Q$.

(iii) Κάθε ιδιοδιάνυσμα του P για την ιδιοτιμή 1, είναι ιδιοδιάνυσμα του Q.

(iv) Ο πίνακας Q είναι επίσης στοχαστικός.

Απόδειξη. Συμβολισμός: για $i \in S$, $[i] = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X : x_0 = i\}$.

(i) Από το εργοδικό θεώρημα παίρνουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{[j]} \circ T^k \to \tilde{\mathbb{1}}_{[j]} \qquad \mu - \text{σχεδόν παντού},$$

πολλαπλασιάζοντας με $\mathbb{1}_{[i]}$ έχουμε ότι

$$\frac{1}{n}\mathbb{1}_{[i]}\sum_{k=0}^{n-1} \mathbb{1}_{[j]} \circ T^k \to \mathbb{1}_{[i]}\tilde{\mathbb{1}}_{[j]} \qquad \mu - \text{σχεδόν παντού}$$

και εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης παίρνουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu\big([i] \cap T^{-k}[j]\big) \to \int \mathbb{1}_{[i]}\tilde{\mathbb{1}}_{[j]}\, d\mu.$$

Παρατηρούμε ότι

$$\begin{aligned}
\mu\big([i] \cap T^{-k}[j]\big) &= \mu\big(\{\underset{\sim}{x} \in X : x_0 = i, x_k = j\}\big) \\
&= \sum_{i_1,\dots,i_{k-1}\in S} \mu\big(\{\underset{\sim}{x} \in X : x_0 = i, x_1 = i_1, x_2 = i_2, \dots, x_{k-1} = i_{k-1}, x_k = j\}\big) \\
&= \sum_{i_1,\dots,i_{k-1}\in S} p_i P_{ii_1} P_{i_1 i_2} \cdots P_{i_{k-1}j} \\
&= p_i P^k_{ij}
\end{aligned}$$

Επομένως,

$$\frac{1}{n}\sum_{k=0}^{n-1} p_i P^k_{ij} \to \int \mathbb{1}_{[i]}\tilde{\mathbb{1}}_{[j]}\, d\mu$$

και εφόσον $p_i > 0$ για κάθε $i \in S$,

$$\frac{1}{n}\sum_{k=0}^{n-1} P^k_{ij} \to \frac{1}{p_i}\int \mathbb{1}_{[i]}\tilde{\mathbb{1}}_{[j]}\, d\mu =: Q_{ij} \qquad \text{καθώς} \quad n \to \infty.$$

(ii)

$$QP = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} P^k P = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} P^{k+1} = \lim_{n\to\infty}\Big(\frac{1}{n}\sum_{k=0}^{n-1} P^k + \frac{1}{n}P^n - \frac{1}{n}P^0\Big) = Q + 0 - 0 = Q,$$

εφόσον $P^n_{ij} \in [0,1]\ \ \forall\, i,j \in S, n \in \mathbb{N}$, και άρα $n^{-1}P^n_{ij} \to 0$ καθώς $n \to \infty, \forall\, i,j \in S$.

(iii) Έστω $v = (v_1, \dots, v_s)$ ένα αριστερό ιδιοδιάνυσμα του P για την ιδιοτιμή 1. Τότε

$$vQ = v\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} P^k = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} vP^k = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} v = v.$$

(iv) Για κάθε γραμμή του Q, δηλαδή για κάθε $i \in S$,

$$\sum_{j\in S} Q_{ij} = \sum_{j\in S}\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} P^k_{ij} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\sum_{j\in S} P^k_{ij} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} 1 = 1,$$

αφού κάθε πίνακας P^k είναι στοχαστικός (όπως βλέπει κανείς επαγωγικά). □

Θεώρημα 2.2.19 Έστω ότι $p_i > 0\ \ \forall\, i \in S$. Τα εξής είναι ισοδύναμα.

(1) Το $(X, \mathcal{B}, \mu, T)$ είναι εργοδικό.

(2) Όλες οι γραμμές του Q είναι ίσες.

(3) $Q_{ij} > 0$ για κάθε $i, j \in S$.

(4) Ο P είναι ανάγωγος.

(5) Η ιδιοτιμή 1 είναι απλή ιδιοτιμή για τον P.

Απόδειξη. (1) ⇒ (2) Από το εργοδικό θεώρημα έχουμε τώρα ότι

$$
\begin{aligned}
& \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{[j]} \circ T^k \to \int \mathbb{1}_{[j]}\, d\mu = p_j \\
\Rightarrow \quad & \mathbb{1}_{[i]}\, \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{[j]} \circ T^k \to \mathbb{1}_{[i]}\, p_j .
\end{aligned}
$$

Εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης,

$$
\begin{aligned}
& \frac{1}{n}\sum_{k=0}^{n-1} \mu\big([i] \,\cap\, T^{-k}[j]\big) \to p_i\, p_j \\
\Leftrightarrow \quad & \frac{1}{n}\sum_{k=0}^{n-1} p_i P_{ij}^k \to p_i\, p_j
\end{aligned}
$$

και αφού $p_i > 0$, καταλήγουμε στο ότι

$$Q_{ij} = p_j .$$

Άρα κάθε γραμμή του Q ισούται με το διάνυσμα-γραμμή $p = (p_1, \ldots, p_s)$.

(2) ⇒ (3) Ξέρουμε ότι $Q_{ij} = q_j$ για κάθε $i \in S$ και για κάθε $j \in S$. Επίσης, από το Λήμμα 2.2.20 ξέρουμε ότι, εφόσον το p είναι ιδιοδιάνυσμα του P, θα είναι και ιδιοδιάνυσμα του Q, για την ιδιοτιμή 1. Δηλαδή,

$$pQ = p \quad \Rightarrow \quad p_j = \sum_{i\in S} p_i\, Q_{ij} = \sum_{i \in S} p_i\, q_j = q_j \qquad \forall\, j \in S.$$

Έπεται ότι κάθε γραμμή του Q ισούται με το διάνυσμα p. Άρα

$$Q_{ij} = p_j > 0 \quad \forall\, i, j \in S.$$

(3) ⇒ (4) Για κάθε $i, j \in S$ έχουμε ότι $Q_{ij} = \lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} P_{ij}^k$ και $Q_{ij} > 0$. Άρα πρέπει κάποιο $P_{ij}^k > 0$, δηλαδή ο P είναι ανάγωγος.

(4) ⇒ (3) Ορίζουμε $S_i = \{j \in S \colon Q_{ij} > 0\}$, για $i \in S$. Επειδή ο Q είναι στοχαστικός πίνακας, θα πρέπει $S_i \neq \varnothing$. Έστω $i \in S$ και $k \in S_i$, δηλαδή υποθέτουμε ότι $k \in S$ είναι τέτοιο ώστε

$$Q_{ik} > 0.$$

Τώρα, έστω $j \in S$, και $n \in \mathbb{N}$ έτσι ώστε

$$P_{kj}^n > 0.$$

Ισχύει ότι

$$Q = QP = \cdots = QP^n,$$

άρα

$$Q_{ij} = \sum_{m\in S} Q_{im} P_{mj}^n \geq Q_{ik} P_{kj}^n > 0.$$

Δηλαδή δείξαμε ότι, για τυχόν $i \in S$, $S_i = S$. Άρα $Q_{ij} > 0\ \forall\, i, j \in S$.

(3) $\Rightarrow$ (2) Αρκεί να δειχτεί ότι για δοθέν $j \in S$, το Q_{ij} δεν εξαρτάται από το i. Ορίζουμε

$$q_j = \max_i Q_{ij}.$$

Έστω ότι υπάρχει $i_* \in S$ με $Q_{i_*j} < q_j$. Τότε

$$Q_{ij} = (Q^2)_{ij} = \sum_{k \in S} Q_{ik}\, Q_{kj} < \sum_{k \in S} Q_{ik}\, q_j = q_j \qquad \forall\, i \in S,$$

η γνήσια ανισότητα επειδή $Q_{ik}Q_{kj} \leq Q_{ik}q_j$ για κάθε $k \in S$ και η ανισότητα είναι γνήσια για ένα τουλάχιστον k, το $k = i_*$, και $Q_{ii_*} > 0$, αφού όλα τα $Q_{lm} > 0$. Άρα

$$q_j = \max_i Q_{ij} < q_j,$$

που είναι άτοπο φυσικά, άρα έχουμε το ζητούμενο.

(2) $\Rightarrow$ (1) Υποθέτουμε ότι όλες οι γραμμές του Q είναι ίδιες. Τότε, έχουμε δείξει, ότι όλες οι γραμμές του Q θα είναι ίσες με το διάνυσμα $p = (p_1, \ldots, p_s)$. Από την Πρόταση 2.2.12, αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(A \,\cap\, T^{-k}B) \to \mu(A)\mu(B) \qquad \forall\, A, B \in \mathcal{B}. \tag{12}$$

Αντ' αυτού θα χρησιμοποιήσουμε το εξής Λήμμα.

Λήμμα 2.2.21 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Τότε το σύστημα είναι εργοδικό ανν ισχύει η (12) για κάθε $A, B \in \mathcal{S}$, όπου $\mathcal{S}$ είναι μια ημι-άλγεβρα [14] που παράγει την $\mathcal{B}$.

Απόδειξη. Άσκηση. $\square$

Οι κύλινδροι

$$\{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = i_0, \ldots, x_n = i_n\}, \quad n \in \mathbb{N}_0, \quad i_0, \ldots, i_n \in S,$$

μαζί με το $\varnothing$, αποτελούν μια ημι-άλγεβρα $\mathcal{S}$. Άρα για να δείξουμε την εργοδικότητα των Markov shifts, αρκεί να δείξουμε την (11) για κάθε $A, B \in \mathcal{S}$.

Έστω

$$A = \{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = i_0, \ldots, x_n = i_n\}$$

και

$$B = \{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = j_0, \ldots, x_m = j_m\},$$

με $m, n \in \mathbb{N}_0$ και $i_0, \ldots, i_n, j_0, \ldots, j_m \in S$. Τότε

$$A \cap T^{-k}(B) = \{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = i_0, \ldots, x_n = i_n, x_k = j_0, \ldots, x_{k+m} = j_m\}.$$

Αν $k > n$, τότε

$$\begin{aligned}
\mu(A \cap T^{-k}(B)) &= \mu\big(\{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = i_0, \ldots, x_n = i_n, x_k = j_0, \ldots, x_{k+m} = j_m\}\big) \\
&= \sum_{l_1, \ldots, l_{k-n-1} \in S} \mu\big(\{\underset{\sim}{x} = (x_0, x_1, \ldots) \in X \colon x_0 = i_0, \ldots, x_n = i_n, \\
&\qquad\qquad x_{n+1} = l_1, \ldots, x_{k-1} = l_{k-n-1}, x_k = j_0, \ldots, x_{k+m} = j_m\}\big) \\
&= \sum_{l_1, \ldots, l_{k-n-1} \in S} p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} P_{i_n l_1} \cdots P_{l_{k-n-1} j_0} P_{j_0 j_1} \cdots P_{j_{m-1} j_m} \\
&= p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} P^{k-n}_{i_n j_0} P_{j_0 j_1} \cdots P_{j_{m-1} j_m}.
\end{aligned}$$

[14]Έστω $(X, \mathcal{B})$ μετρήσιμος χώρος. Μια ημι-άλγεβρα μετρήσιμων υποσυνόλων είναι μια οικογένεια $\mathcal{S} \subset \mathcal{B}$ με τις εξής ιδιότητες: (i) $\varnothing \in \mathcal{S}$, (ii) η $\mathcal{S}$ είναι κλειστή ως προς πεπερασμένες τομές, (iii) αν $A \in \mathcal{S}$, τότε $X \setminus A$ είναι πεπερασμένη ένωση ξένων ανά δύο στοιχείων της $\mathcal{S}$.

Επομένως

$$\begin{aligned}
\frac{1}{N}\sum_{k=0}^{N-1}\mu(A\cap T^{-k}B) &= \frac{1}{N}\sum_{k=0}^{n}\mu(A\cap T^{-k}B)+\frac{1}{N}\sum_{k=n+1}^{N-1}p_{i_0}P_{i_0i_1}\cdots P_{i_{n-1}i_n}P^{k-n}_{i_nj_0}P_{j_0j_1}\cdots P_{j_{m-1}j_m}\\
&= \frac{1}{N}\sum_{k=0}^{n}\mu(A\cap T^{-k}B)+\mu(A)\frac{1}{N}\sum_{l=1}^{N-n-1}P^{l}_{i_nj_0}P_{j_0j_1}\cdots P_{j_{m-1}j_m}\\
&\xrightarrow{N\to\infty} 0+\mu(A)Q_{i_nj_0}P_{j_0j_1}\cdots P_{j_{m-1}j_m}\\
&= \mu(A)p_{j_0}P_{j_0j_1}\cdots P_{j_{m-1}j_m}\\
&= \mu(A)\mu(B),
\end{aligned}$$

όπου η προ-τελευταία ισότητα ισχύει λόγω της υπόθεσης ότι όλες οι γραμμές του Q είναι ίδιες. Για την σύγκλιση καθώς $N\to\infty$, χρησιμοποιήθηκε το γεγονός ότι, αν για μία ακολουθία $(a_n)_{n\in\mathbb{N}}$ αριθμών $a_n\to a$, τότε και $N^{-1}\sum_{l=1}^{N}a_l\to a$.

(2) ⇒ (5) Έστω ότι όλες οι γραμμές του Q είναι ίδιες. Τότε

$$Q_{ij}=p_j\quad \forall i,j\in S.$$

Έστω v αριστερό ιδιοδιάνυσμα του P για την ιδιοτιμή 1. Από το Λήμμα 2.2.20, το v θα είναι και αριστερό ιδιοδιάνυσμα του Q για την ιδιοτιμή 1, δηλαδή

$$vQ=v.$$

Τότε

$$v_j=\sum_{i\in S}v_iQ_{ij}=\sum_{i\in S}v_ip_j\quad \forall j\in S$$

δηλαδή

$$v=\left(\sum_{i\in S}v_i\right)p.$$

Άρα κάθε ιδιοδιάνυσμα του P για την ιδιοτιμή 1 είναι πολλαπλάσιο του p και άρα ο ιδιόχωρος που αντιστοιχεί στην ιδιοτιμή 1 είναι διάστασης 1.

(5) ⇒ (2) Έστω ότι το 1 είναι απλή ιδιοτιμή του P. Τότε ο αντίστοιχος ιδιόχωρος από αριστερά ιδιοδιανύσματα του P έχει διάσταση[15] 1, και αφού υποθέτουμε ότι το p είναι αριστερό ιδιοδιάνυσμα του P για την ιδιοτιμή 1, τα ιδιοδιανύσματα του P που αντιστοιχούν στην ιδιοτιμή 1 πρέπει να είναι όλα πολλαπλάσια του p. Από την σχέση $QP=Q$ παίρνουμε ότι κάθε γραμμή του Q είναι αριστερό ιδιοδιάνυσμα του P για την ιδιοτιμή 1, και άρα όλες οι γραμμές του Q είναι πολλαπλάσια του p. Επειδή ο Q είναι στοχαστικός πίνακας και το p διάνυσμα πιθανότητας, πρέπει κάθε γραμμή του Q να είναι ίση με το p. □

2.3 Mixing

Ορισμός 2.3.1 (ισχυρά mixing) Έστω $(X,\mathcal{B},\mu,T)$ σ.δ.μ. Το σύστημα είναι *ισχυρά mixing* (*strongly mixing*) αν

$$\mu(A\,\cap\,T^{-k}B)\to\mu(A)\mu(B)\qquad \forall\,A,B\in\mathcal{B}. \tag{13}$$

[15]Ο ιδιόχωρος του P που αποτελείται από δεξιά ιδιοδιανύσματα για την ιδιοτιμή 1 είναι ο πυρήνας $\ker(I-P)$, όπου I ο ταυτοτικός πίνακας; η διάστασή του είναι η διάσταση του υποκείμενου χώρου μείον την τάξη στηλών του $I-P$: είναι $|S|-\dim(\mathrm{Im}(I-P))=|S|-\mathrm{rank}(I-P)$, όπου $\mathrm{Im}(I-P)$ η εικόνα του $I-P$ και $\mathrm{rank}(I-P)$ η τάξη του $I-P$, που ισούται με την διάσταση της εικόνας του $I-P$. Ο ιδιόχωρος του P που αποτελείται από αριστερά ιδιοδιανύσματα για την ιδιοτιμή 1 είναι ο πυρήνας $\ker\big((I-P)^t\big)=\ker(I-P^t)$, όπου t δηλώνει ανάστροφο πίνακα; η διάστασή του είναι ίση με την διάσταση του υποκείμενου χώρου μείον την τάξη στηλών του $(I-P)^t$: είναι $|S|-\dim\big(\mathrm{Im}\big((I-P)^t\big)\big)=|S|-\mathrm{rank}\big((I-P)^t\big)$, όπου $\mathrm{Im}\big((I-P)^t\big)$ η εικόνα του $(I-P)^t$ και $\mathrm{rank}\big((I-P)^t\big)$ η τάξη του $(I-P)^t$; η τελευταία ισούται με την διάσταση της εικόνας του $(I-P)^t$, η οποία ισούται με το πλήθος των γραμμικά ανεξάρτητων στηλών του $(I-P)^t$, και άρα με το πλήθος των γραμμικά ανεξάρτητων γραμμών του $I-P$. Αυτό όμως είναι επίσης ίσο με την τάξη του πίνακα $I-P$, από Γραμμική Άλγεβρα, και άρα ο ιδιόχωρος του P που αποτελείται από τα αριστερά ιδιοδιανύσματα του P για την ιδιοτιμή 1 (ή οποιαδήποτε ιδιοτιμή εν τέλει) έχει την ίδια διάσταση με τον ιδιόχωρο του P που αποτελείται από τα δεξιά ιδιοδιανύσματα για την ιδιοτιμή 1.

Ορισμός 2.3.2 (ασθενές mixing) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Το σύστημα είναι *ασθενώς mixing* (*weakly mixing*) αν

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\mu\big(A \,\cap\, T^{-k}(B)\big) - \mu(A)\mu(B)\big| \to 0 \qquad \forall\, A, B \in \mathcal{B}. \tag{14}$$

Παρατήρηση. Ισχυρό mixing ⇒ ασθενές mixing ⇒ εργοδικότητα και δεν ισχύουν εν γένει οι αντίστροφες συνεπαγωγές.

Παράδειγμα. Έστω $X = \mathbb{T}, \mathcal{B} = \mathcal{B}(\mathbb{T})$, μ το μέτρο Lebesgue και $T_a(x) = x + a \pmod 1$. Από εργοδικότητα του συστήματος αν το a είναι άρρητος, και συγκεκριμένα από το Θεώρημα 2.1.7 του Kronecker, και από περιοδικότητα του συστήματος αν το a είναι ρητός, και συγκεκριμένα επειδή T^n είναι η ταυτοτική απεικόνιση αν $a = m/n$, υπάρχει ακολουθία φυσικών αριθμών $n_1 < n_2 < \cdots$ τέτοια ώστε

$$n_k a \pmod 1 \xrightarrow{(k\to\infty)} 0.$$

Τώρα αν $A = B = [0, \frac{1}{2})$, τότε

$$\mu\big(A \cap T_a^{n_k}(B)\big) \to \tfrac{1}{2}$$

ενώ $\mu(A)\mu(B) = \frac{1}{4}$. Επομένως σε αυτό το παράδειγμα, αν $a \in \mathbb{R} \setminus \mathbb{Q}$, το σύστημα θα είναι εργοδικό (Πρόταση 2.1.6) αλλά όχι ισχυρά mixing. Θα δούμε αμέσως παρακάτω ότι δεν είναι ούτε ασθενώς mixing.

Ορισμός 2.3.3 Ένα σύνολο φυσικών $J \subset \mathbb{N}$ έχει πυκνότητα 0 εάν

$$\frac{1}{n}\big|J \,\cap\, [1, n]\big| \xrightarrow{(n\to\infty)} 0,$$

όπου $|A|$ είναι ο πληθάριθμος του A.

Θεώρημα 2.3.4 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Τότε τα εξής είναι ισοδύναμα

(1) Το σύστημα είναι ασθενώς mixing.

(2) Το σύστημα $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \otimes \mu, T \times T)$[16] είναι εργοδικό.

(3) Το σύστημα $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \otimes \mu, T \times T)$ είναι ασθενώς mixing.

(4) Για οποιοδήποτε εργοδικό σ.δ.μ., $(Y, \mathcal{B}_Y, \nu, S)$, το σύστημα $(X \times Y, \mathcal{B} \otimes \mathcal{B}_Y, \mu \otimes \mu, T \times S)$ είναι επίσης εργοδικό.

(5) Οι μόνες μετρήσιμες ιδιοσυναρτήσεις του τελεστή Koopman είναι οι σταθερές. (Το σύστημα έχει συνεχές φάσμα.)

(6) Για κάθε $A, B \in \mathcal{B}$, υπάρχει $J_{A,B} \subset \mathbb{N}$ με πυκνότητα 0 τέτοιο ώστε

$$\mu\big(A \,\cap T^{-n}(B)\big) \xrightarrow[n \notin J_{A,B}]{(n\to\infty)} \mu(A)\mu(B).$$

(7) Για κάθε $A, B \in \mathcal{B}$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)\big|^2 \to 0.$$

Παράδειγμα. Δείξαμε πιο πάνω ότι το σύστημα $X = \mathbb{T}, \mathcal{B} = \mathcal{B}(\mathbb{T})$, μ το μέτρο Lebesgue και $T_a(x) = x + a \pmod 1$ δεν είναι ισχυρά mixing. Από το Θεώρημα 2.3.4 έπεται ότι δεν ούτε και ασθενώς mixing. Πράγματι, το σύστημα $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \times \mu, T_a \times T_a)$ δεν είναι εργοδικό αφού, για παράδειγμα, η συνάρτηση $f(x, y) = e^{2\pi i(x-y)}$ είναι σαφές ότι είναι $T_a \times T_a$-αναλλοίωτη αλλά δεν είναι σταθερή.

Λήμμα 2.3.5 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\mathcal{S} \subset \mathcal{B}$ μια ημι-άλγεβρα που παράγει την $\mathcal{B}$. Τότε το σύστημα είναι ισχυρά mixing ανν η (13) ισχύει για κάθε $A, B \in \mathcal{S}$. Όμοια, το σύστημα είναι ασθενώς mixing ανν η (14) ισχύει για κάθε $A, B \in \mathcal{S}$.

Λήμμα 2.3.6 Έστω $(a_n)_{n\in\mathbb{N}}$ φραγμένη ακολουθία με $a_n \geq 0$, για κάθε $n \in \mathbb{N}$. Τα εξής είναι ισοδύναμα.

[16] Αν $T \colon X \to X$ και $S \colon Y \to Y$ ορίζεται ο $T \times S \colon X \times Y \to X \times Y$ με τύπο $(T \times S)(x, y) = \big(T(x), S(y)\big)$.

(1) $n^{-1}\sum_{k=1}^{n} a_k \to 0$.

(2) Υπάρχει $J \subset \mathbb{N}$ που έχει πυκνότητα 0 τέτοιο ώστε

$$a_n \xrightarrow[n\notin J]{(n\to\infty)} 0.$$

(3) $n^{-1}\sum_{k=1}^{n} a_k^2 \to 0$.

Απόδειξη. Θα δειχτεί ότι (1) $\Leftrightarrow$ (2). Τότε με a_n^2 στην θέση του a_n θα έχουμε και το $(2) \Leftrightarrow (3)$.

(1) $\Rightarrow$ (2) Ορίζουμε για κάθε $k \in \mathbb{N}$,

$$J_k = \{n \in \mathbb{N} \colon a_n > 1/k\}.$$

Παρατηρούμε ότι $J_1 \subset J_2 \subset \cdots$ και ακόμα κάθε J_k έχει πυκνότητα 0. Πράγματι,

$$\frac{\big|J_k \ \cap \ [1,n]\big|}{k} \le \sum_{m\in J_k \ \cap \ [1,n]} a_m \le \sum_{m=1}^{n} a_m \qquad (a_m \ge 0),$$

όπου η πρώτη ανισότητα είναι η ανισότητα Markov. Άρα,

$$\frac{1}{n}\big|J_k \ \cap \ [1,n]\big| \le \frac{k}{n}\sum_{m=1}^{n} a_m \to 0 \qquad \text{(από υπόθεση)}.$$

Έπεται ότι για κάθε $k \in \mathbb{N}$, υπάρχει $m_k \in \mathbb{N}$ τέτοιο ώστε

$$\frac{1}{n}\big|J_k \ \cap \ [1,n]\big| < \frac{1}{k} \qquad \forall\, n \ge m_k.$$

Μπορούμε φυσικά να επιλέξουμε τα m_k έτσι ώστε $m_1 < m_2 < \cdots$. Ορίζουμε

$$J = \bigcup_{k\ge 1} J_k \ \cap \ [m_k, m_{k+1}).$$

Ισχυρισμός 1. Το J έχει πυκνότητα 0.

Πράγματι, έστω $n \in \mathbb{N}$. Υπάρχει $k \in \mathbb{N}$ τέτοιο ώστε $m_k \le n < m_{k+1}$ και τότε

$$J \cap [1,n] \subset J_k \cap [1,n].$$

Επομένως

$$\big|J \cap [1,n]\big| \le \big|J_k \cap [1,n]\big| < \frac{1}{k}n.$$

Αν $\varepsilon > 0$, επιλέγουμε $k_0 \in \mathbb{N}$ έτσι ώστε $1/k_0 < \varepsilon$ και τότε, για $n \ge m_{k_0}$, θα έχουμε ότι $m_k \le n < m_{k+1}$ για κάποιο $k \ge k_0$. Επομένως, για κάθε τέτοιο $n \ge m_{k_0}$,

$$\frac{1}{n}\big|J \cap [1,n]\big| < \frac{1}{k} \le \frac{1}{k_0} < \varepsilon.$$

Ισχυρισμός 2. $a_n \xrightarrow[n\notin J]{n\to\infty} 0$.

Έστω $n \notin J$ και επιλέγουμε $k \in \mathbb{N}$ τέτοιο ώστε $m_k \le n < m_{k+1}$. Αφού $n \notin J$ έχουμε ότι για κάθε $l \in \mathbb{N}$, $n \notin J_l \cap [m_l, m_{l+1})$. Τώρα για $l = k$, επειδή $n \in [m_k, m_{k+1})$, θα πρέπει $n \notin J_k$. Δηλαδή θα πρέπει $a_n \le 1/k$. Για δοθέν $\varepsilon > 0$, επιλέγουμε $k_0 \in \mathbb{N}$ τέτοιο ώστε $1/k_0 < \varepsilon$. Τότε, για $n \ge m_{k_0}$, θα έχουμε ότι $m_k \le n < m_{k+1}$ για κάποιο $k \ge k_0$ και άρα

$$a_n \le \frac{1}{k} \le \frac{1}{k_0} < \varepsilon.$$

(2) $\Rightarrow$ (1) Έστω $J \subset \mathbb{N}$ με πυκνότητα 0 τέτοιο ώστε $a_n \xrightarrow[n\notin J]{n\to\infty} 0$.

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} a_k &= \frac{1}{n}\sum_{k\in J\cap[1,n]} a_k + \frac{1}{n}\sum_{k\in J^c\cap[1,n]} a_k \\ &\leq a\frac{1}{n}\big|J\cap[1,n]\big| + \frac{1}{n}\sum_{k\in J^c\cap[1,n]} a_k\end{aligned}$$

όπου $a = \sup_{n\in\mathbb{N}} a_n < \infty$. Έστω $\varepsilon > 0$ και επιλέγουμε $n_1 \in \mathbb{N}$ τέτοιο ώστε

$$a_n < \varepsilon/3 \qquad \forall\, n \geq n_1,\ n \notin J,$$

και τότε

$$\begin{aligned}\frac{1}{n}\sum_{k\in J^c\cap[1,n]} a_k &= \frac{1}{n}\sum_{\substack{k=1\\ k\notin J}}^{n_0-1} a_k + \frac{1}{n}\sum_{\substack{k=n_0\\ k\notin J}}^{n} a_k \\ &< \frac{n_0-1}{n}a + \frac{\varepsilon}{3}\frac{1}{n}(n-n_0).\end{aligned}$$

Αν επιλέξουμε και n_2 έτσι ώστε $\forall\, n \geq n_2$,

$$a\frac{n_0-1}{n} < \varepsilon/3$$

και $n_3 \in \mathbb{N}$ τέτοιο ώστε

$$a\frac{1}{n}\big|J\cap[1,n]\big| < \varepsilon/3$$

$\forall\, n \geq n_3$, τότε για κάθε $n \geq \max\{n_1, n_2, n_3\}$,

$$\frac{1}{n}\sum_{k=1}^{n} a_k < \varepsilon. \qquad \square$$

Απόδειξη Θεωρήματος 2.3.4 Οι ισοδυναμίες (1) $\Leftrightarrow$ (6) $\Leftrightarrow$ (7) είναι άμεση εφαρμογή του Λήμματος 2.3.6 και το (3) $\Rightarrow$ (2) είναι άμεσο (βλ. Παρατήρηση μετά τον Ορισμό 2.3.2).

(6) $\Rightarrow$ (3) Αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\Big|(\mu\times\mu)\big(A\cap(T\times T)^{-k}B\big) - (\mu\times\mu)(A)(\mu\times\mu)(B)\Big| \to 0$$

όταν $A = A_1 \times A_2$ και $B = B_1 \times B_2$, εφόσον τα σύνολα της μορφής $C \times D$ με $C, D \in \mathcal{B}$ αποτελούν ημι-άλγεβρα που παράγει την $\mathcal{B} \otimes \mathcal{B}$. Τώρα,

$$A_1 \times A_2 \cap (T\times T)^{-k}(B_1 \times B_2) = A_1 \cap T^{-k}(B_1) \times A_2 \cap T^{-k}(B_2)$$

επομένως,

$$(\mu\times\mu)\big(A_1\times A_2 \cap (T\times T)^{-k}(B_1\times B_2)\big) = \mu\big(A_1\cap T^{-k}(B_1)\big)\mu\big(A_2\cap T^{-k}(B_2)\big)$$

Από το (6) υπάρχουν $J_1, J_2 \subset \mathbb{N}$ με πυκνότητα 0 το καθένα τέτοια ώστε

$$\mu(A_i\cap T^{-n}B_i) \xrightarrow[n\notin J_i]{n\to\infty} \mu(A_i)\mu(B_i) \quad i=1,2.$$

Εάν $J = J_1 \cup J_2$, τότε το J έχει πυκνότητα 0 και

$$\mu(A_i\cap T^{-n}B_i) \xrightarrow[n\notin J]{n\to\infty} \mu(A_i)\mu(B_i) \quad i=1,2$$

και άρα

$$(\mu\times\mu)\big(A_1\times A_2\cap(T\times T)^{-n}(B_1\times B_2)\big) \xrightarrow[n\notin J]{n\to\infty} \mu(A_1)\mu(B_1)\mu(A_2)\mu(B_2).$$

Εφαρμόζοντας το Λήμμα 2.3.6 παίρνουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\Big|(\mu\times\mu)\big(A_1\times A_2\ \cap\ (T\times T)^{-k}(B_1\times B_2)\big)-(\mu\times\mu)(A_1\times A_2)(\mu\times\mu)(B_1\times B_2)\Big|\to 0.$$

(3) $\Rightarrow$ (1) Έστω $A,B\in\mathcal{B}$. Από το (3) παίρνουμε ότι

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1}\Big|(\mu\times\mu)\big(A\times X\ \cap\ (T\times T)^{-k}(B\times X)\big)-(\mu\times\mu)(A\times X)(\mu\times\mu)(B\times X)\Big|\to 0\\
\Leftrightarrow\quad&\frac{1}{n}\sum_{k=0}^{n-1}\Big|(\mu\times\mu)\big(A\cap T^{-k}(B)\times X\cap T^{-k}(X)\big)-\mu(A)\mu(X)\mu(B)\mu(X)\Big|\to 0\\
\Leftrightarrow\quad&\frac{1}{n}\sum_{k=0}^{n-1}\Big|(\mu\times\mu)\big(A\cap T^{-k}(B)\times X\big)-\mu(A)\mu(B)\Big|\to 0\\
\Leftrightarrow\quad&\frac{1}{n}\sum_{k=0}^{n-1}\Big|\mu(A\cap T^{-k}(B))-\mu(A)\mu(B)\Big|\to 0
\end{aligned}$$

εφόσον $\mu(X)=1$. Άρα το σύστημα είναι ασθενώς mixing.

(1) $\Rightarrow$ (4) Έστω $(Y,\mathcal{B}_Y,\nu,S)$ ένα εργοδικό σύστημα που διατηρεί το μέτρο και θεωρούμε το ευθύ γινόμενο $(X\times Y,\mathcal{B}\otimes\mathcal{B}_Y,\mu\times\nu,T\times S)$. Για την εργοδικότητα, αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\nu)\big(A_1\times A_2\ \cap\ (T\times S)^{-k}(B_1\times B_2)\big)\to(\mu\times\nu)(A_1\times A_2)(\mu\times\nu)(B_1\times B_2),$$

για αυθαίρετα $A_1,B_1\in\mathcal{B}$ και $A_2,B_2\in\mathcal{B}_Y$. Το αριστερό μέλος ισούται με

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1}\mu(A_1\cap T^{-k}(B_1))\nu(A_2\cap S^{-k}(B_2))\\
&=\frac{1}{n}\sum_{k=0}^{n-1}\big[\mu(A_1\cap T^{-k}(B_1))-\mu(A_1)\mu(B_1)\big]\nu(A_2\cap S^{-k}(B_2))+\mu(A_1)\mu(B_1)\frac{1}{n}\sum_{k=0}^{n-1}\nu(A_2\cap S^{-k}(B_2))\\
&\to 0+\mu(A_1)\mu(B_1)\nu(A_2)\nu(B_2)=(\mu\times\nu)(A_1\times A_2)(\mu\times\nu)(B_1\times B_2),
\end{aligned}$$

όπου ο δεύτερος όρος της προτελευταίας σειράς συγκλίνει στο $\mu(A_1)\mu(B_1)\nu(A_2)\nu(B_2)$ επειδή το σύστημα $(Y,\mathcal{B}_Y,\nu,S)$ είναι εργοδικό, ενώ ο πρώτος όρος της της προτελευταίας σειράς συγκλίνει στο μηδέν επειδή κυριαρχείται από το

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{k=0}^{n-1}\big[\mu(A_1\cap T^{-k}(B_1))-\mu(A_1)\mu(B_1)\big]\nu(A_2\cap S^{-k}(B_2))\right|\\
&\qquad\qquad\le\frac{1}{n}\sum_{k=0}^{n-1}\big|\mu(A_1\cap T^{-k}(B_1))-\mu(A_1)\mu(B_1)\big|\nu(A_2\cap S^{-k}(B_2))\to 0,
\end{aligned}$$

αφού $\nu(A_2\cap S^{-k}(B_2))\le 1$ και το σύστημα $(X,\mathcal{B},\mu,T)$ είναι ασθενώς mixing.

(4) $\Rightarrow$ (2) Πρώτα δείχνουμε ότι το σύστημα $(X,\mathcal{B},\mu,T)$ είναι εργοδικό. Παίρνουμε για $Y=\{y\}$, $\mathcal{B}_Y=\{\varnothing,Y\}$, $\nu=\delta_{\{y\}}$ και $S(y)=y$. Το σύστημα $(Y,\mathcal{B}_Y,\nu,S)$ είναι εργοδικό (τετριμμένα). Τότε εφαρμόζοντας την υπόθεση, έχουμε ότι το το σύστημα $(X\times Y,\mathcal{B}\otimes\mathcal{B}_Y,\mu\times\nu,T\times S)$ είναι εργοδικό. Άρα για $A,B\in\mathcal{B}$,

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\nu)(A\times Y\ \cap\ (T\times S)^{-k}(B\times Y)\to(\mu\times\nu)(A\times Y)(\mu\times\nu)(B\times Y)\\
\Leftrightarrow\quad&\frac{1}{n}\sum_{k=0}^{n-1}\mu(A\ \cap\ T^{-k}B)\nu(Y)\to\mu(A)\nu(Y)\mu(B)\nu(Y)=\mu(A)\mu(B).
\end{aligned}$$

Αφού αυτό ισχύει για αυθαίρετα $A, B \in \mathcal{B}$, έπεται ότι το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι εργοδικό. Τώρα, εφαρμόζουμε το (4) για $(Y, \mathcal{B}_Y, \nu, S) = (X, \mathcal{B}, \mu, T)$ και έχουμε το ζητούμενο.

(2) $\Rightarrow$ (7) Έστω $A, B \in \mathcal{B}$. Θέλουμε

$$\frac{1}{n}\sum_{k=0}^{n-1}\left|\mu(A \cap T^{-k}(B)) - \mu(A)\mu(B)\right|^2 \to 0.$$

Το αριστερό μέλος είναι ίσο με

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(A \cap T^{-k}B)^2 + \mu(A)^2\mu(B)^2 - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu(A \cap T^{-k}B).$$

Έχουμε ότι

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\mu(A \cap T^{-k}B)^2 &= \frac{1}{n}\sum_{k=0}^{n-1}(\mu \times \mu)\big(A \times A \ \cap \ (T \times T)^{-k}(B \times B)\big)\\ &\to (\mu \times \mu)(A \times A)(\mu \times \mu)(B \times B) = \mu(A)^2\mu(B)^2,\end{aligned}$$

όπου η σύγκλιση ισχύει λόγω της υπόθεσης (2). Επίσης,

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\mu(A \cap T^{-k}B) &= \frac{1}{n}\sum_{k=0}^{n-1}(\mu \times \mu)\big(A \times X \ \cap \ (T \times T)^{-k}(B \times X)\big)\\ &\to (\mu \times \mu)(A \times X)(\mu \times \mu)(B \times X) = \mu(A)\mu(B),\end{aligned}$$

άρα αποδείχθηκε το ζητούμενο.

(2) $\Rightarrow$ (5) Έστω f ιδιοσυνάρτηση του τελεστή U_T. Τότε υπάρχει $\lambda \in \mathbb{C}$ τέτοιο ώστε

$$U_T f = \lambda f.$$

Επειδή ο U_T είναι ισομετρία σε κάθε χώρο L_p, πρέπει $|\lambda| = 1$. Πράγματι,

$$\|f\|_p = \|U_T f\|_p = |\lambda|\|f\|_p$$

και αφού $\|f\|_p \neq 0 \Rightarrow |\lambda| = 1$. Το επιχείρημα αυτό δείχνει ότι κάθε ιδιοτιμή του τελεστή Koopman U_T, σαν τελεστή $L^p(X, \mathcal{B}, \mu) \to L^p(X, \mathcal{B}, \mu)$, έχει μέτρο ένα. Όμως και αν θεωρήσουμε τον U_T σαν τελεστή πάνω στις μετρήσιμες συναρτήσεις, τότε πάλι κάθε ιδιοτιμή αυτού του τελεστή πρέπει να έχει μέτρο ένα (απόδειξη: άσκηση). Επομένως θεωρούμε ότι η f είναι μία μετρήσιμη ιδιοσυνάρτηση του τελεστή U_T, με αντίστοιχη ιδιοτιμή λ και γνωρίζουμε ότι $|\lambda| = 1$.

Ορίζουμε $g\colon X \times X \to \mathbb{C}$ με

$$g(x, y) = f(x)\overline{f(y)} \qquad (x, y) \in X \times Y.$$

Τότε

$$\begin{aligned}g \circ (T \times T)(x, y) = g\big(T(x), T(y)\big) &= f\big(T(x)\big)\overline{f\big(T(y)\big)} = U_T f(x)\,\overline{U_T f(y)}\\ &= \lambda f(x)\bar{\lambda}\overline{f(y)} = |\lambda|^2 g(x, y) = g(x, y).\end{aligned}$$

Επομένως

$$U_{T \times T}\, g = g,$$

άρα η g είναι ιδιοδιάνυσμα του $U_{T\times T}$ για την ιδιοτιμή 1 ή, ισοδύναμα, η g είναι $T \times T$ είναι αναλλοίωτη. Εξ' υποθέσεως, το σύστημα $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \otimes \mu, T \times T)$ είναι εργοδικό, άρα πρέπει η g να είναι σταθερή $\mu \times \mu$-σχεδόν παντού. Αυτό όμως συνεπάγεται ότι η f είναι σταθερή μ-σχεδόν παντού.

Στην συνέχεια θα αποδείξουμε ότι η (5) $\Rightarrow$ ασθενές mixing.

Ορισμός 2.3.7 (θετικά ορισμένη ακολουθία) Μια ακολουθία μιγαδικών αριθμών $(a_n)_{n\in\mathbb{Z}}$ λέγεται *θετικά ορισμένη* αν, για κάθε $n \in \mathbb{N}$ και για κάθε επιλογή $z_1, z_2, \dots, z_n \in \mathbb{C}$ ισχύει ότι

$$\sum_{k=1}^{n}\sum_{m=1}^{n} a_{k-m} z_k \overline{z_m} \geq 0.$$

Θεώρημα 2.3.8 (Herglotz) Μια ακολουθία μιγαδικών αριθμών $(a_n)_{n\in\mathbb{Z}}$ είναι θετικά ορισμένη ανν υπάρχει πεπερασμένο θετικό μέτρο Borel στον $\mathbb{T}$, ή ισοδύναμα στον $\mathbb{S}^1$, τέτοιο ώστε

$$\widehat{\mu}(n) = a_n \qquad \forall n \in \mathbb{Z}.$$

Λήμμα 2.3.9 Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. Τότε, για δοθείσα $f \in L^2(X, \mathcal{B}, \mu)$, η ακολουθία

$$a_n = \begin{cases} \langle U_T^n f, f\rangle & n \in \mathbb{N}_0 \\ \langle f, U_T^{-n} f\rangle & -n \in \mathbb{N} \end{cases}$$

είναι θετικά ορισμένη. Το αντίστοιχο μέτρο στον $\mathbb{T}$ του θεωρήματος Herglotz θα το συμβολίζουμε με μ_f και θα το καλούμε φασματικό μέτρο της f. Δηλαδή το μ_f είναι το πεπερασμένο θετικό μέτρο Borel στον $\mathbb{T}$ για το οποίο

$$\widehat{\mu_f}(n) = \begin{cases} \langle U_T^n f, f\rangle & n \in \mathbb{N}_0 \\ \langle f, U_T^{-n} f\rangle & -n \in \mathbb{N}. \end{cases}$$

Ισχύει ότι

$$\langle U_T^n f, U_T^m f\rangle = \widehat{\mu_f}(n-m) \qquad \forall\, n, m \in \mathbb{N}_0. \tag{15}$$

Απόδειξη. Θα δείξουμε ότι η $(a_n)_{n\in\mathbb{Z}}$, όπως ορίστηκε, είναι θετικά ορισμένη. Έστω $n \in \mathbb{N}$, $z_1, \dots, z_n \in \mathbb{C}$ και ζητάμε

$$\sum_{k=1}^{n}\sum_{m=1}^{n} a_{k-m} z_k \overline{z_m} \geq 0.$$

Εάν $k - m \geq 0$,

$$\begin{aligned} a_{k-m} = \langle U_T^{k-m} f, f\rangle &= \int_X (f\circ T^{k-m})\bar{f}\, d\mu = \int_X (f\circ T^{k-m})\bar{f}\, dT_*^m\mu \\ &= \int_X (f\circ T^k)\overline{(f\circ T^m)}\, d\mu = \langle U_T^k f, U_T^m f\rangle, \end{aligned}$$

όπου χρησιμοποιήθηκε ότι ο T, και άρα και ο T^m, διατηρεί το μέτρο μ. Ομοίως, αν $k - m < 0$,

$$\begin{aligned} a_{k-m} = \langle f, U_T^{m-k} f\rangle = \int_X f\,\overline{U_T^{m-k} f}\, d\mu = \int_X f\,\overline{(f\circ T^{m-k})}\, d\mu &= \int_X f\,\overline{(f\circ T^{m-k})}\, dT_*^k\mu \\ &= \int_X (f\circ T^k)\,\overline{(f\circ T^m)}\, d\mu = \langle U_T^k f, U_T^m f\rangle \end{aligned}$$

και πάλι. Άρα βλέπουμε ήδη ότι θα ισχύει η (15). Τώρα

$$\begin{aligned} \sum_{k=1}^{n}\sum_{m=1}^{n} a_{k-m} z_k \overline{z_m} = \sum_{k=1}^{n}\sum_{m=1}^{n} z_k\overline{z_m}\langle U_T^k f, U_T^m f\rangle &= \sum_{k=1}^{n}\sum_{m=1}^{n} \langle z_k U_T^k f, z_m U_T^m f\rangle \\ &= \left\langle \sum_{k=1}^{n} z_k U_T^k f, \sum_{m=1}^{n} z_m U_T^m f \right\rangle = \left\| \sum_{k=1}^{n} z_k U_T^k f \right\|_2^2 \geq 0. \end{aligned}$$

Άρα η $(a_n)_{n\in\mathbb{Z}}$ είναι θετικά ορισμένη, άρα υπάρχει θετικό πεπερασμένο μέτρο μ_f από το θεώρημα Herglotz, τέτοιο ώστε

$$\widehat{\mu_f}(n) = a_n \qquad \forall\, n \in \mathbb{Z}.$$

□

Λήμμα 2.3.10 Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. και $f \in L^2(X, \mathcal{B}, \mu)$. Τότε

(1) $n^{-1}\sum_{k=0}^{n-1}\langle U_T^k f, f\rangle e^{2\pi i t k} \to \mu_f\big(\{t\}\big)$ για κάθε $t \in \mathbb{T}$.

(2) $n^{-1}\sum_{k=0}^{n-1}\big|\langle U_T^k f, f\rangle\big|^2 \to \sum_{t\in\mathbb{T}}\big|\mu_f\big(\{t\}\big)\big|^2$.

Σημείωση: Το άθροισμα στο (2) είναι ένα αριθμήσιμο άθροισμα γιατί για ένα πεπερασμένο μέτρο, ν έστω, $\nu\left(\{u\}\right) > 0$ μπορεί να ισχύει μόνο για αριθμήσιμο το πλήθος σημείων u στον χώρο στον οποίο είναι ορισμένο το μέτρο.

Απόδειξη. (1)

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1}\langle U_T^k f, f\rangle e^{2\pi i t k} &= \frac{1}{n}\sum_{k=0}^{n-1}\widehat{\mu_f}(k)e^{2\pi i t k}\\
&= \frac{1}{n}\sum_{k=0}^{n-1}\int_{\mathbb{T}} e^{-2\pi i k s}\, d\mu_f(s) e^{2\pi i t k} = \frac{1}{n}\sum_{k=0}^{n-1}\int_{\mathbb{T}} e^{2\pi i k(t-s)}\, d\mu_f(s)\\
&= \frac{1}{n}\int_{\mathbb{T}\setminus\{t\}} \frac{1-e^{2\pi i n(t-s)}}{1-e^{2\pi i(t-s)}}\, d\mu_f(s) + \mu_f(\{t\}),
\end{aligned}$$

αφού για την γεωμετρική σειρά έχουμε ότι

$$\sum_{k=0}^{n-1} e^{2\pi i k(t-s)} = \begin{cases} \dfrac{1-e^{2\pi i n(t-s)}}{1-e^{2\pi i(t-s)}} & \text{αν } s \neq t\\ n & \text{αν } s = t. \end{cases}$$

Τώρα, εφόσον

$$\frac{1}{n}\left|\frac{1-e^{2\pi i n(t-s)}}{1-e^{2\pi i(t-s)}}\right| \leq \frac{1}{n}\frac{1+\left|e^{2\pi i n(t-s)}\right|}{\left|1-e^{2\pi i(t-s)}\right|} = \frac{1}{n}\frac{2}{\left|1-e^{2\pi i n(t-s)}\right|} \xrightarrow{n\to\infty} 0, \qquad \text{για } s \neq t,$$

και το αριστερό μέλος είναι φραγμένο από το 1, επειδή $\left|\sum_{k=0}^{n-1} e^{2\pi i k(t-s)}\right| \leq \sum_{k=0}^{n-1}\left|e^{2\pi i k(t-s)}\right| \leq n$, και είμαστε σε χώρο πεπερασμένου μέτρου, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης παίρνουμε ότι

$$\frac{1}{n}\int_{\mathbb{T}\setminus\{t\}} \frac{1-e^{2\pi i n(t-s)}}{1-e^{2\pi i(t-s)}}\, d\mu_f(s) \to 0.$$

Άρα έχουμε το ζητούμενο.

(2) Έχουμε διαδοχικά

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1}\big|\langle U_T^k f, f\rangle\big|^2 = \frac{1}{n}\sum_{k=0}^{n-1}|\widehat{\mu_f}(k)|^2 &= \frac{1}{n}\sum_{k=0}^{n-1}\int_{\mathbb{T}} e^{-2\pi i k s}\, d\mu_f(s)\overline{\int_{\mathbb{T}} e^{-2\pi i k t}\, d\mu_f(t)}\\
&= \frac{1}{n}\sum_{k=0}^{n-1}\int_{\mathbb{T}}\int_{\mathbb{T}} e^{2\pi i k(t-s)}\, d\mu_f(t)\, d\mu_f(s)\\
&= \iint_{\{(t,s)\in\mathbb{T}^2\,:\,t\neq s\}} \frac{1}{n}\frac{1-e^{2\pi i(t-s)n}}{1-e^{2\pi i(t-s)}}\, d\mu_f(t)\, d\mu_f(s)\\
&\qquad + \iint_{\{(t,s)\in\mathbb{T}^2\,:\,t=s\}} 1\, d\mu_f(t)\, d\mu_f(s)\\
&\xrightarrow{n\to\infty} \iint_{\{(t,s)\in\mathbb{T}^2\,:\,t=s\}} 1\, d\mu_f(t)\, d\mu_f(s)\\
&= \int_{\mathbb{T}} \mu_f(\{s\})\, d\mu_f(s)\\
&= \sum_{s\in\mathbb{T}}\big|\mu_f(\{s\})\big|^2,
\end{aligned}$$

έχοντας χρησιμοποιήσει και το θεώρημα Fubini στην προτελευταία ισότητα. Άρα αποδείχθηκε το ζητούμενο. $\square$

Για να δείξουμε ότι το σύστημα είναι ασθενώς mixing, δεδομένου του (5), αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\langle U_T^k f, g\rangle - \langle f, \mathbb{1}_X\rangle\langle \mathbb{1}_X, g\rangle\big| \to 0 \qquad \forall\, f, g \in L^2(\mu),$$

αφού τότε, για $f = \mathbb{1}_B$, $g = \mathbb{1}_A$, $A, B \in \mathcal{B}$, θα έχουμε την συνθήκη (14) του ορισμού του ασθενούς mixing. Επίσης, χρησιμοποιώντας την ταυτότητα

$$\langle U_T^n f, g\rangle = \langle U_T^n(f+g), f+g\rangle + i\langle U_T^n(f+ig), f+ig\rangle - \langle U_T^n(f-g), f-g\rangle - i\langle U_T^n(f-ig), f-ig\rangle,$$

για $f, g \in L^2(\mu)$, βλέπει κανείς ότι αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\langle U_T^k f, f\rangle - \langle f, \mathbb{1}_X\rangle\langle \mathbb{1}_X, f\rangle\big| \to 0 \qquad \forall\, f \in L^2(\mu), \tag{16}$$

για το οποίο, τέλος, αρκεί να δείξουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\langle U_T^k f, f\rangle\big| \to 0 \qquad \forall\, f \in L^2(\mu) \quad \text{με} \quad \int f\,d\mu = 0 \tag{17}$$

εφόσον, εφαρμόζοντας την (17) στην $f - \big(\int f\,d\mu\big)\mathbb{1}_X$ για μία αυθαίρετη $f \in L^2(\mu)$, θα έχουμε την ζητούμενη (16). Από το Λήμμα 2.3.6, αρκεί να δειχτεί ότι

$$\frac{1}{n}\sum_{k=0}^{n-1}\big|\langle U_T^k f, f\rangle\big|^2 \to 0 \qquad \forall\, f \in L^2(\mu) \quad \text{με} \quad \int f\,d\mu = 0.$$

Επομένως, λόγω του Λήμματος 2.3.10 (2), ζητάμε άμα έχουμε μια $f \in L^2(X, \mathcal{B}, \mu)$ με $\int f\,d\mu = 0$, το φασματικό μέτρο μ_f να μην έχει σημειακές μάζες:

$$\sum_{t\in\mathbb{T}} \mu_f\big(\{t\}\big)^2 = 0.$$

Από το Λήμμα 2.3.10 (1) έχουμε ότι

$$\mu_f(\{0\}) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\langle U_T^k f, f\rangle.$$

Τώρα γράφουμε

$$\frac{1}{n}\sum_{k=0}^{n-1}\langle U_T^k f, f\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle U_T^k f - P_F(f), f\rangle + \langle P_F(f), f\rangle,$$

όπου P_F είναι η προβολή στον χώρο $F = \{h \in L^2(\mu) \colon U_T h = h\}$ των T-αναλλοίωτων συναρτήσεων στον $L^2(\mu)$, ισοδύναμα, των ιδιοσυναρτήσεων του $U_T \colon L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{B}, \mu)$ για την ιδιοτιμή 1. Από το μέσο εργοδικό θεώρημα von Neumann παίρνουμε ότι

$$\left|\left\langle \frac{1}{n}\sum_{k=0}^{n-1} U_T^k f - P_F(f), f\right\rangle\right| \leq \left\|\frac{1}{n}\sum_{k=0}^{n-1} U_T^k f - P_F(f)\right\|_2 \|f\|_2 \to 0,$$

επομένως

$$\frac{1}{n}\sum_{k=0}^{n-1}\langle U_T^k f, f\rangle \to \langle P_F(f), f\rangle = \langle P_F(f), P_F(f)\rangle = \|P_F(f)\|_2^2$$

($\langle P_F(f), f - P_F(f)\rangle = 0$ για κάθε f, αφού $P_f(f)$ είναι ορθογώνια προβολή στον F και άρα $P_F(f) \in F$ και $f - P_F(f) \perp F$). Έχουμε υποθέσει ότι $\int f\,d\mu = 0$, δηλαδή η f είναι κάθετη σε κάθε σταθερή συνάρτηση και αν οι μόνες ιδιοσυναρτήσεις του U_T για την ιδιοτιμή 1 είναι οι σταθερές συναρτήσεις (σχεδόν παντού), με άλλα λόγια αν ο F αποτελείται από τις σχεδόν παντού σταθερές συναρτήσεις, τότε $P_F(f) = 0$. Άρα,

$$\mu_f(\{0\}) = \|P_F(f)\|_2^2 = 0.$$

Για τα $\mu_f(\{t\})$ για αυθαίρετα $t \in \mathbb{T}$ θεωρούμε τους τελεστές

$$V_t = e^{2\pi i t} U_T.$$

Κάθε V_t είναι ισομετρία, άρα είναι συστολή και εφαρμόζοντας το μέσο εργοδικό θεώρημα von Neumann για τον V_t, παίρνουμε ότι

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} V_t^k f - P_{F_{V_t}} f \right\|_2 \to 0,$$

όπου

$$F_{V_t} = \{h \in L^2(\mu) \colon V_t h = h\} = \{h \in L^2(\mu) \colon U_T h = e^{-2\pi i t} h\}.$$

Παρατηρεί κανείς ότι ο χώρος F_{V_t} είναι ο χώρος των ιδιοσυναρτήσεων (ιδιόχωρος) του τελεστή Koopman U_T για την ιδιοτιμή $e^{-2\pi i t}$, αν αυτή είναι ιδιοτιμή, διαφορετικά $F_{V_t} = \{0\}$ είναι ο τετριμμένος υπόχωρος. Όπως προηγουμένως, έχουμε ότι

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} e^{2\pi i k t} U_T^k f - P_{F_{V_t}} f \right\|_2 \to 0 \quad \Rightarrow \quad \frac{1}{n} \sum_{k=0}^{n-1} e^{2\pi i k t} \langle U_T^k f, f \rangle \to \|P_{F_{V_t}}(f)\|_2^2.$$

Άρα για κάθε $f \in L^2(\mu)$,

$$\mu_f(\{t\}) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} e^{2\pi i k t} \langle U_T^k f, f \rangle = \|P_{F_{V_t}}(f)\|_2^2,$$

από το Λήμμα 2.3.10 (1) και επομένως, από το Λήμμα 2.3.10 (2),

$$\frac{1}{n} \sum_{k=0}^{n-1} \left| \langle U_T^k f, f \rangle \right|^2 \to \sum_{t \in \mathbb{T}} \mu_f(\{t\})^2 = \sum_{t \in \mathbb{T}} \|P_{F_{V_t}}(f)\|_2^4.$$

Τώρα, αν $\int f \, d\mu = 0$ τότε η f είναι κάθετη σε κάθε σταθερή συνάρτηση. Επειδή οι μόνες ιδιοσυναρτήσεις του τελεστή U_T είναι οι σταθερές συναρτήσεις, κάθε F_{V_t} περιέχει μόνο σταθερές συναρτήσεις. Άρα η f είναι κάθετη στα F_{V_t} για κάθε $t \in \mathbb{T}$, δηλαδή $P_{F_{V_t}}(f) = 0$ για κάθε $t \in \mathbb{T}$, και άρα

$$\frac{1}{n} \sum_{k=0}^{n-1} \left| \langle U_T^k f, f \rangle \right|^2 \to 0 \qquad \forall f \in L^2(\mu) \quad \text{με} \quad \int f \, d\mu = 0.$$ □

2.4 Γενικευμένα Μέσα Εργοδικά Θεωρήματα

Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. και U_T ο αντίστοιχος τελεστής Koopman στον $L^2(X, \mathcal{B}, \mu)$. Έστω P_n οικογένεια πολυωνύμων,

$$P_n(z) = \sum_{k=0}^{d_n} a_{k,n} z^k,$$

όπου $d_n \in \mathbb{N}_0$ για κάθε $n \in \mathbb{N}$. Τα πολυώνυμα P_n ορίζουν τελεστές

$$P_n(U_T) = \sum_{k=0}^{d_n} a_{k,n} U_T^k$$

στον $L^2(X, \mathcal{B}, \mu)$:

$$P_n(U_T) f = \sum_{k=0}^{n-1} a_{k,n} U_T^k f = \sum_{k=0}^{n-1} a_{k,n} f \circ T^k \qquad f \in L^2(X, \mathcal{B}, \mu).$$

Για παράδειγμα, όταν $a_{k,n} = 1/n$ για κάθε $k \in \{0, 1, \ldots, d_n\}$, όπου $d_n = n - 1$, έχουμε τον τελεστή

$$P_n(U_T) f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k.$$

Για $f \in L^2(X, \mathcal{B}, \mu)$, υπολογίζουμε

$$\begin{aligned}
\|P_n(U_T)f\|_2^2 &= \langle P_n(U_T)f, P_n(U_T)f\rangle = \left\langle \sum_{k=0}^{d_n} a_{k,n} U_T^k f, \sum_{m=0}^{d_n} a_{m,n} U_T^m f \right\rangle \\
&= \sum_{k=0}^{d_n}\sum_{m=0}^{d_n} a_{k,n}\bar{a}_{m,n} \, \langle U_T^k f, U_T^m f\rangle =: \sum_{k=0}^{d_n}\sum_{m=0}^{d_n} a_{k,n}\bar{a}_{m,n} \, \widehat{\mu_f}(k-m) \\
&= \sum_{k=0}^{d_n}\sum_{m=0}^{d_n} a_{k,n}\bar{a}_{m,n} \int_{\mathbb{T}} e^{-2\pi i s(k-m)} \, d\mu_f(s) = \int_{\mathbb{T}} \sum_{k=0}^{d_n} a_{k,n} e^{-2\pi i s k} \overline{\left(\sum_{m=0}^{d_n} a_{m,n} e^{-2\pi i s m} \right)} \, d\mu_f(s) \\
&= \int_{\mathbb{T}} \left| \sum_{k=0}^{d_n} a_{k,n} e^{-2\pi i s k} \right|^2 d\mu_f(s) = \int_{\mathbb{T}} \left| P_n(e^{-2\pi i s}) \right|^2 d\mu_f(s),
\end{aligned}$$

όπου μ_f το φασματικό μέτρο της f, δηλαδή το θετικό πεπερασμένο μέτρο Borel στον $\mathbb{T}$ για το οποίο $\widehat{\mu_f}(k) = \langle U^k f, f\rangle$ για κάθε $k \in \mathbb{N}_0$. Δηλαδή, τελικά

$$\|P_n(U_T)f\|_2^2 = \int_{\mathbb{T}} \left| P_n(e^{-2\pi i s}) \right|^2 d\mu_f(s) \tag{18}$$

για $f \in L^2(X, \mathcal{B}, \mu)$.

Για $f \in L^2(X, \mathcal{B}, \mu)$, γράφουμε $f = g + h$, όπου g είναι η ορθογώνια προβολή της f στον χώρο

$$F_0 = \{u \in L^2(\mu) \colon U_T \, u = u\}$$

και $h \perp F_0$. Υποθέτουμε επιπλέον τώρα ότι $\sum_{k=0}^{d_n} a_{k,n} = P_n(1) = 1$. Τότε, αφού $g \in F_0$, έχουμε ότι $U_T^k g = g$ για κάθε $k \in \mathbb{N}_0$ και άρα $P_n(U_T)g = g$. Επομένως,

$$\|P_n(U_T)f - g\|_2^2 = \|P_n(U_T)f - P_n(U_T)g\|_2^2 = \|P_n(U_T)h\|_2^2 = \int_{\mathbb{T}} \left| P_n(e^{-2\pi i s}) \right|^2 d\mu_h(s),$$

από την (18). Αφού $h \perp F_0$, έχουμε επίσης ότι

$$\mu_h(\{0\}) = \|P_{F_0}(h)\|_2^2 = 0.$$

Επομένως,

$$\|P_n(U_T)f - g\|_2^2 \to 0 \qquad \text{αν} \qquad P_n(e^{-2\pi i s}) \to 0 \quad \forall\, s \in \mathbb{T} \setminus \{0\}$$

και αν επιπλέον τα πολυώνυμα P_n είναι ομοιόμορφα φραγμένα πάνω στην περιφέρεια του μοναδιαίου κύκλου, ώστε να μπορεί κανείς να εφαρμόσει το θεώρημα κυριαρχημένης σύγκλισης. Αυτό συμβαίνει για παράδειγμα, από την τριγωνική ανισότητα, αν $P_n(1) = 1$ και $a_{k,n} \geq 0$ για κάθε $k \in \{0, 1, \ldots, d_n\}$ και $n \in \mathbb{N}$. Αποδείξαμε το εξής.

Θεώρημα 2.4.1 (γενικευμένο μέσο εργοδικό θεώρημα) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $P_n(z) = \sum_{k=0}^{d_n} a_{k,n} z^k$ πολυώνυμα τέτοια ώστε

(1) $P_n(1) = 1 \ \forall\, n \in \mathbb{N}$,

(2) $a_{k,n} \geq 0 \ \forall\, k \in \{0, 1, \ldots, d_n\} \ \forall\, n \in \mathbb{N}$,

(3) $P_n(e^{-2\pi i s}) \to 0 \quad \forall\, s \in \mathbb{T} \setminus \{0\}$.

Τότε για κάθε $f \in L^2(\mu)$,

$$\|P_n(U_T)f - g\|_2 \to 0,$$

όπου g είναι η προβολή της f στον υπόχωρο του L^2 που αποτελείται από τις T-αναλλοίωτες συναρτήσεις.

Για παράδειγμα, όταν $P_n(z) = n^{-1} \sum_{k=0}^{n-1} z^k$, τότε ικανοποιούνται οι απαιτούμενες συνθήκες, δηλαδή

$$P_n(1) = 1 \qquad \text{και} \qquad a_{k,n} = \frac{1}{n} \geq 0 \qquad \forall\, k \in \{0, 1, \ldots, n-1\}, \ n \in \mathbb{N},$$

και
$$P_n(e^{-2\pi is}) = \frac{1}{n}\frac{1-e^{-2\pi isn}}{1-e^{2\pi is}} \to 0 \qquad \text{για } s \neq 0.$$
Επίσης
$$P_n(U_T)f = \frac{1}{n}\sum_{k=0}^{n-1} U_T^k f = \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k.$$
Δηλαδή το Θεώρημα 2.4.1 σε αυτήν την περίπτωση δίνει το κλασικό θεώρημα von Neumann[17].

Για ασθενώς mixing συστήματα έχει κανείς ένα ισχυρότερο θεώρημα.

Θεώρημα 2.4.2 Έστω $(X, \mathcal{B}, \mu, T)$ ένα ασθενώς mixing σύστημα που διατηρεί το μέτρο και έστω
$$P_n(z) = \sum_{k=0}^{d_n} a_{k,n} z^k$$
πολυώνυμα τέτοια ώστε $P_n(1) = 1$ και $a_{k,n} \geq 0$ για κάθε $k \in \{0, 1, \dots, d_n\}$ και $n \in \mathbb{N}$ και $P_n(e^{-2\pi is}) \to 0$ για κάθε s εκτός από ένα αριθμήσιμο σύνολο. Τότε
$$\|P_n(U_T)f - g\|_2 \to 0,$$
όπου g είναι η προβολή της f στον υπόχωρο του L^2 των T-αναλλοίωτων συναρτήσεων.

Απόδειξη. Έστω $f \in L^2(X, \mathcal{B}, \mu)$ και γράφουμε πάλι $f = g + h$ με g να είναι η ορθογώνια προβολή της f στον υπόχωρο του L^2 των T-αναλλοίωτων συναρτήσεων και $h = f - g$. Έχουμε πάλι, από την (18), ότι
$$\|P_n(U_T)f - g\|_2^2 = \|P_n(U_T)h\|_2^2 = \int_{\mathbb{T}} |P_n(e^{-2\pi is})|^2 \, d\mu_h(s).$$

Ισχυρισμός: Για κάθε αριθμήσιμο σύνολο $A \subseteq \mathbb{T}$, $\mu_h(A) = 0$.

Δοθέντος του Ισχυρισμού, το Θεώρημα έπεται άμεσα. Πράγματι, εφαρμόζοντας θεώρημα κυριαρχημένης σύγκλισης βλέπουμε ότι
$$\int_{\mathbb{T}} |P_n(e^{-2\pi is})|^2 \, d\mu_h(s) \to 0,$$
αφού $|P_n(e^{-2\pi is})| \leq \sum_{k=0}^{n-1} |a_{k,n}| = \sum_{k=0}^{n-1} a_{k,n} = 1 \ \forall n \in \mathbb{N}$ και $P_n(e^{-2\pi is}) \to 0$ για s εκτός από ένα αριθμήσιμο σύνολο, δηλαδή μ_h-σχεδόν παντού. $\square$

Απόδειξη Ισχυρισμού: Το μ_h δεν έχει σημειακές μάζες. Πράγματι, γνωρίζουμε ότι
$$\mu_h(\{t\}) = \|P_{F_t}(h)\|_2^2,$$
όπου $F_t = \{u \in L^2(\mu) \colon U_T h = e^{-2\pi it} h\}$. Επειδή το σύστημα είναι ασθενώς mixing έχουμε ότι $F_t = \{0\}$ $\forall t \neq 0$ επειδή όλες οι ιδιοσυναρτήσεις του U_T είναι σταθερές και αυτές είναι όλες ιδιοσυναρτήσεις για την ιδιοτιμή 1 και επομένως ο U_T δεν έχει άλλες ιδιοτιμές. Άρα
$$\mu_h(\{t\}) = \|P_{F_t}(h)\|_2^2 = 0 \qquad \forall t \neq 0.$$
Επιπλέον, αφού $h \perp F_0$,
$$\mu_h(\{0\}) = \|P_{F_0}(h)\|_2^2 = 0.$$
Επομένως το μ_h δεν έχει σημειακές μάζες. Αν τώρα το A είναι ένα αριθμήσιμο υποσύνολο του $\mathbb{T}$,
$$\mu_h(A) = \sum_{a\in A} \mu_h(\{a\}) = 0. \qquad \square$$

[17]Το οποίο, θεώρημα von Neumann, χρησιμοποιήθηκε όμως στην απόδειξη του Θεωρήματος 2.4.1.

2.5 Παραδείγματα και Εφαρμογές

Εφαρμογή 2.5.1 Έστω $(X, \mathcal{B}, \mu, T)$ ένα ασθενώς mixing σύστημα. Τότε για κάθε $f \in L^2(\mu)$

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^{2k} \xrightarrow{L^2} P_{F_0}(f),$$

όπου F_0 είναι ο υπόχωρος του L^2 των αναλλοίωτων συναρτήσεων και P_{F_0} η ορθογώνια προβολή πάνω στον F_0.

Πράγματι, θεωρούμε τα πολυώνυμα

$$P_n(z) = \frac{1}{n}\sum_{k=0}^{n-1} z^{2k} \qquad n \in \mathbb{N},$$

δηλαδή $a_{k,n} = 1/n \ \forall k \in \{0, 2, 4, \ldots, 2(n-1)\}$ και $a_{k,n} = 0 \ \forall k \in \{1, 3, 5, \ldots, 2n-3\}$. Τότε $a_{k,n} \geq 0$, $P_n(1) = 1$ και

$$P_n(z) = \frac{1}{n}\frac{1 - z^{2n}}{1 - z^2}, \qquad z \notin \{-1, 1\}.$$

Άρα

$$P_n(e^{-2\pi i s}) = \frac{1}{n}\frac{1 - e^{-4\pi i s n}}{1 - e^{-4\pi i s}} \to 0, \qquad s \notin \{0, 1/2\},$$

και επομένως από το Θεώρημα 2.4.2 παίρνουμε

$$P_n(U_T)f \xrightarrow{L^2} P_{F_0}(f),$$

που μας δίνει το ζητούμενο. □

Παράδειγμα 2.5.2 $X = \mathbb{T}$, $\mathcal{B} = \mathcal{B}(\mathbb{T})$, μ το μέτρο Lebesgue και $T_a(x) = x + a \pmod 1$. Το σύστημα αυτό δεν είναι ασθενώς mixing για κανένα $a \in [0, 1)$.

Παράδειγμα 2.5.3 Έστω $X = \mathbb{T}^n$, $\mathcal{B} = \mathcal{B}(\mathbb{T}^n)$, μ το n-διάστατο μέτρο Lebesgue και $T(x) = Ax \pmod 1$, όπου A είναι ένας $n \times n$ πίνακας με $A_{ij} \in \mathbb{Z} \ \forall i, j$ και $\det(A) \neq 0$. Το σύστημα αυτό είναι ισχυρά mixing ανν είναι εργοδικό.

Απόδειξη. Ισχυρό mixing συνεπάγεται εργοδικότητα, πάντα. Έστω, τώρα, ότι το σύστημα είναι εργοδικό. Δηλαδή καμιά ιδιοτιμή του A δεν είναι ρίζα της μονάδας. Θα δείξουμε ότι

$$\langle U_T^k f, g\rangle \xrightarrow{k\to\infty} \langle f, \mathbb{1}_X\rangle\langle \mathbb{1}_X, g\rangle \qquad \forall f, g \in L^2(\mu). \tag{19}$$

Τότε αφού $\langle U_T^k f, g\rangle = \int (f \circ T^k)\bar{g}\, d\mu$, $\langle f, \mathbb{1}_X\rangle = \int_X f\, d\mu$, $\langle \mathbb{1}_X, g\rangle = \overline{\int_X g\, d\mu}$, αν $f = \mathbb{1}_B$ και $g = \mathbb{1}_A$, έχουμε ότι

$$\langle U_T^k f, g\rangle = \mu(T^{-k}B \cap A), \quad \langle f, \mathbb{1}_X\rangle = \mu(B), \quad \langle \mathbb{1}_X, g\rangle = \mu(A)$$

και η (19) είναι ισοδύναμη με την συνθήκη (13) του ορισμού του ισχυρού mixing.

Για την (19), έστω, ως συνήθως,

$$e_m(x) = e^{2\pi i\langle m, x\rangle} \qquad m \in \mathbb{Z}^n,\ x \in \mathbb{R}^n.$$

Για κάθε $m \in \mathbb{Z}^n$, θεωρούμε τον

$$H_m = \{f \in L^2(\mu) \colon \langle U_T^k f, e_m\rangle \xrightarrow{k\to\infty} \langle f, \mathbb{1}_X\rangle\langle \mathbb{1}_X, e_m\rangle\}.$$

Ο H_m είναι γραμμικός υπόχωρος του $L^2(\mu)$. Όταν $m = (0, 0, \ldots, 0)$, $e_m(x) = 1 \ \forall x \in \mathbb{R}^n$ και

$$\langle U_T^k f, e_m\rangle = \langle U_T^k f, \mathbb{1}_X\rangle = \int U_T^k f\, d\mu = \int f \circ T^k\, d\mu = \int f\, d\mu = \langle f, \mathbb{1}_X\rangle,$$

αφού ο T διατηρεί το μέτρο. Επειδή δε, προφανώς $\langle \mathbb{1}_X, e_m\rangle = 1$, έπεται ότι $H_{(0,\ldots,0)} = L^2(\mu)$.

Για $m \neq (0, 0, \ldots, 0)$,

$$H_m = \{f \in L^2(\mu) \colon \langle U_T^k f, e_m\rangle \to 0\}$$

λόγω της ορθογωνιότητας των e_l, $l \in \mathbb{Z}^n$: $\langle \mathbb{1}_X, e_m\rangle = \langle e_0, e_m\rangle = 0$. Δείχνουμε ότι ο H_m είναι κλειστός υπόχωρος. Πράγματι, έστω $f_j \in H_m$, $j \in \mathbb{N}$, τέτοια ώστε

$$f_j \xrightarrow[j\to\infty]{L^2} f.$$

Έχουμε

$$\begin{aligned} \left|\langle U_T^k f, e_m\rangle\right| &\leq \left|\langle U_T^k (f - f_j), e_m\rangle\right| + \left|\langle U_T^k f_j, e_m\rangle\right| \\ &\leq \|U_T^k (f-f_j)\|_2 \|e_m\|_2 + \left|\langle U_T^k f_j, e_m\rangle\right| \\ &= \|f - f_j\|_2 + \left|\langle U_T^k f_j, e_m\rangle\right|, \end{aligned}$$

αφού U_T είναι ισομετρία και $\|e_m\|_2 = 1$. Άρα, δοθέντος $\varepsilon > 0$, επιλέγουμε j έτσι ώστε $\|f - f_j\|_2 < \varepsilon/2$ και για αυτό το j, επιλέγουμε k_0 τέτοιο ώστε $\left|\langle U_T^k f_j, e_m\rangle\right| < \varepsilon/2 \ \forall\, k \geq k_0$. Τότε $\left|\langle U_T^k f, e_m\rangle\right| < \varepsilon \ \forall\, k \geq k_0$. Δείξαμε δηλαδή ότι

$$\left|\langle U_T^k f, e_m\rangle\right| \xrightarrow{k\to\infty} 0.$$

Επομένως $f \in H_m$ και έπεται ότι ο H_m κλειστός υπόχωρος του $L^2(\mu)$.

Θα αποδείξουμε τώρα ότι $e_l \in H_m$ για κάθε $l \in \mathbb{Z}^n$ και τότε θα έπεται ότι $H_m = L^2(\mu)$. Υπολογίζουμε

$$\begin{aligned} \langle U_T^k e_l, e_m\rangle &= \int_X e_l \circ T^k(x)\, \overline{e_m(x)}\, dx \\ &= \int_X \exp(2\pi i \langle l, A^k x\rangle - 2\pi i \langle m, x\rangle)\, dx \\ &= \int_X \exp(2\pi i \langle (A^*)^k l - m, x\rangle)\, dx \\ &= \begin{cases} 1 & \text{αν } (A^*)^k l = m \\ 0 & \text{αλλιώς.} \end{cases} \end{aligned}$$

Όμως για δοθέν $l \in \mathbb{Z}^n$, η σχέση $(A^*)^k l = m$, όταν $m \neq (0, 0, \dots, 0)$, μπορεί να ισχύει για το πολύ ένα $k \in \mathbb{N}$. Διαφορετικά, θα υπήρχαν $k_1 \neq k_2$ στο $\mathbb{N}$ τέτοια ώστε

$$(A^*)^{k_1} l = m = (A^*)^{k_2} l,$$

δηλαδή το l θα ήταν ιδιοδιάνυσμα για τον πίνακα $(A^*)^{|k_1 - k_2|}$ με αντίστοιχη ιδιοτιμή 1, άρα ο $A^{|k_1-k_2|}$ θα είχε επίσης ιδιοτιμή το 1, που σημαίνει ότι ο A θα είχε ιδιοτιμή κάποια ρίζα της μονάδας, που είναι αδύνατον. Επομένως, για κάθε $l \in \mathbb{Z}^n$,

$$\langle e_l \circ T^k, e_m\rangle = 0$$

για όλα τα αρκούντως μεγάλα $k \in \mathbb{N}$ και επομένως, $e_l \in H_m$.

Τώρα, επειδή οι γραμμικοί συνδυασμοί των e_l, $l \in \mathbb{Z}$, είναι πυκνοί στον $L^2(\mu)$, έχουμε ότι $H_m = L^2(\mu)$. Έπεται ότι, για κάθε $f \in L^2(X, \mathcal{B}, \mu)$, το σύνολο

$$F_f = \{g \in L^2(\mu) \colon \langle U_T^k f, g\rangle \to \langle f, \mathbb{1}_X\rangle \langle \mathbb{1}_X, g\rangle\}$$

περιέχει όλα τα εκθετικά e_m, $m \in \mathbb{Z}^n$. Κάθε F_f είναι κλειστός υπόχωρος και άρα πρέπει $F_f = L^2(\mu)$, για κάθε $f \in L^2(\mu)$. Δηλαδή δείξαμε ότι

$$\langle U_T^k f, g\rangle \to \langle f, \mathbb{1}_X\rangle \langle \mathbb{1}_X, g\rangle \qquad \forall\, f, g \in L^2(\mu).$$

□

Παράδειγμα 2.5.4 (Bernoulli Shifts) Έστω $S = \{1, 2, \dots, s\}$ πεπερασμένο σύνολο, $X = S^{\mathbb{N}}$, $\mathcal{B} = \otimes_{\mathbb{N}} \mathcal{P}(S)$ η σ-άλγεβρα γινόμενο που παράγεται από όλα τα σύνολα της μορφής

$$\{\underset{\sim}{x} = (x_1, x_2, \dots) \colon x_1 = i_1, \dots, x_n = i_n\}, \qquad n \in \mathbb{N}, \quad i_1, \dots, i_n \in S, \tag{20}$$

$T \colon X \to X$ ο μετασχηματισμός shift. Επίσης, έστω $(p_i)_{i \in S}$ διάνυσμα πιθανότητας και μ το μέτρο πιθανότητας επί του $(X, \mathcal{B})$ για το οποίο

$$\mu(\{\underset{\sim}{x} = (x_1, x_2, \dots) \colon x_1 = i_1, \dots, x_n = i_n\}) = p_{i_1} p_{i_2} \cdots p_{i_n},$$

για κάθε $n \in \mathbb{N}$ και $i_1, \dots, i_n \in S$. Αυτό το σύστημα είναι πάντα ισχυρά mixing. Πράγματι, αρκεί να αποδειχθεί ότι

$$\mu(T^{-k}B \cap A) \to \mu(A)\mu(B)$$

για κάθε A, B στην ημιάλγεβρα που σχηματίζουν τα σύνολα της μορφής (20) μαζί με το κενό σύνολο. Αν

$$A = \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = i_1, \dots, x_n = i_n\}, \;\; B = \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = j_1, \dots, x_m = j_m\},$$

$n, m \in \mathbb{N}$, $i_1, \dots, i_n, j_1, \dots, j_m \in S$, δύο τέτοια σύνολα, τότε

$$T^{-k}B \cap A = \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = i_1, \dots, x_n = i_n, x_{k+1} = j_1, \dots, x_{k+m} = j_m\}$$

και αν $k \geq n$,

$$\mu(T^{-k}B \cap A) = p_{i_1} \cdots p_{i_n} p_{j_1} \cdots p_{j_m} = \mu(A)\mu(B).$$

Επομένως το σύστημα είναι ισχυρά mixing.

Παράδειγμα 2.5.5 (Markov Shifts) Έστω πάλι $S = \{1, 2, \dots, s\}$, $X = S^{\mathbb{N}_0}$, $\mathcal{B} = \otimes_{\mathbb{N}_0} \mathcal{P}(S)$, $T \colon X \to X$ ο μετασχηματισμός shift. Επίσης, έστω $P = (P_{ij})_{i,j \in S}$ στοχαστικός πίνακας, $p = (p_i)_{i \in S}$ διάνυσμα πιθανότητας που είναι αριστερό ιδιοδιάνυσμα του P για την ιδιοτιμή 1, δηλαδή $(p_1, \dots, p_s)P = (p_1, \dots, p_s)$ και μέτρο μ επί του μετρήσιμου χώρου $(X, \mathcal{B})$ τέτοιο ώστε

$$\mu(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i_0, x_1 = i_1, \dots, x_n = i_n\}) = p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n},$$

για κάθε $n \in \mathbb{N}_0$, $i_0, i_1, \dots, i_n \in S$.

Πρόταση 2.5.6 Υποθέτουμε ότι $p_i > 0$ για κάθε $i \in S$. Τα εξής είναι ισοδύναμα.

(i) Το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι ασθενώς mixing.

(ii) Το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι ισχυρά mixing.

(iii) Ο πίνακας P είναι ανάγωγος και απεριοδικός, δηλαδή υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε $P_{ij}^n > 0 \;\forall\, i, j \in S$.

(iv) $P_{ij}^n \to p_j$ για κάθε $i, j \in S$.

Απόδειξη. (ii) $\Rightarrow$ (i) Ισχύει πάντα.

(iii) $\Rightarrow$ (iv) Κλασικό θεώρημα στοχαστικών διαδικασιών.

(i) $\Rightarrow$ (iii) Έστω $i, j \in S$, $A = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i\}$ και $B = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = j\}$. Ξεκινάμε από την υπόθεση ότι το σύστημα είναι ασθενώς mixing, άρα

$$\begin{aligned} &\frac{1}{n} \sum_{k=0}^{n-1} \Big| \mu(T^{-k}B \,\cap\, A) - \mu(A)\mu(B) \Big| \to 0 \\ \Rightarrow \quad &\frac{1}{n} \sum_{k=0}^{n-1} \Big| p_i P_{ij}^k - p_i p_j \Big| \to 0. \end{aligned}$$

Για κάθε $(i, j) \in S \times S$, υπάρχει ένα σύνολο $J_{ij} \subset \mathbb{N}$ με πυκνότητα 0, τέτοιο ώστε

$$p_i P_{ij}^n \xrightarrow[n \notin J_{ij}]{n \to \infty} p_i p_j.$$

Επειδή $p_i > 0$, έπεται ότι

$$P_{ij}^n \xrightarrow[n \notin J_{ij}]{n \to \infty} p_j.$$

Θεωρούμε το $J = \bigcup_{(i,j) \in S \times S} J_{ij}$, το οποίο έχει πυκνότητα 0, αφού το $S \times S$ είναι πεπερασμένο και για το οποίο έχουμε ότι

$$P_{ij}^n \xrightarrow[n \notin J]{n \to \infty} p_j \qquad \forall\, i, j \in S.$$

Εφόσον αυτά τα όρια είναι όλα θετικά, θα υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε $P_{ij}^n > 0$ για κάθε $(i,j) \in S \times S$. Πράγματι, υπάρχουν $n_{ij} \in \mathbb{N}$ τέτοια ώστε $|P_{ij}^n - p_j| < \frac{1}{2}p_j$ για κάθε $n \notin J$ με $n \geq n_{ij}$, και τότε για οποιοδήποτε $n \notin J$ με $n \geq \max_{(i,j)\in S\times S} n_{ij}$ έχει κανείς ότι $P_{ij}^n > \frac{1}{2}p_j > 0$ για όλα τα $i,j \in S$.

(iv) $\Rightarrow$ (ii) Έστω

$$A = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i_0, \dots, x_n = i_n\}$$

και

$$B = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = j_0, \dots, x_m = j_m\},$$

$n, m \in \mathbb{N}_0, i_0, \dots, i_n, j_0, \dots, j_m \in S$. Αν $k > n$, τότε

$$\begin{aligned}\mu(T^{-k}B \cap A) &= \mu\left(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i_0, \dots, x_n = i_n, x_k = j_0, \dots, x_{k+m} = j_m\}\right) \\ &= p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} P_{i_n j_0}^{k-n} P_{j_0 j_1} \cdots P_{j_{m-1} j_m}.\end{aligned}$$

Από την υπόθεση έχουμε ότι

$$P_{i_n j_0}^{k-n} \xrightarrow{k\to\infty} p_{j_0}.$$

Επομένως

$$\mu(T^{-k}B \cap A) \xrightarrow{k\to\infty} p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} p_{j_0} P_{j_0 j_1} \cdots P_{j_{m-1} j_m} = \mu(A)\mu(B). \qquad \square$$

Σημείωση. Δεν αποτελεί βλάβη της γενικότητας να υποθέτει κανείς ότι $p_i > 0$ για κάθε $i \in S$ για μία αλυσίδα Markov (Markov shift) με χώρο καταστάσεων S και πίνακα μετάβασης P. Πράγματι, έστω $p = (p_i)_{i\in S}$ ένα αριστερό ιδιοδιάνυσμα για τον πίνακα P, που είναι και διάνυσμα πιθανότητας. Έστω και $S' := \{i \in S \colon p_i > 0\}$. Καταρχήν παρατηρεί κανείς ότι, με τον συμβολισμό

$$[i] = \{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i\},$$

$$\mu([i]) = p_i = 0 \qquad \forall\, i \in S \setminus S',$$

όπου μ το μέτρο στον χώρο ακολουθιών $X = S^{\mathbb{N}_0}$ που αντιστοιχεί στο διάνυσμα p. Επειδή το μέτρο μ είναι shift-αναλλοίωτο, έχει κανείς ότι

$$\mu\left(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_k = i\}\right) = \mu(T^{-k}[i]) = 0$$

για κάθε $k \in \mathbb{N}_0$, για $i \in S \setminus S'$, και επομένως

$$\mu\left(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_k = i \text{ για κάποιο } k \in \mathbb{N}_0\}\right) = \mu\left(\bigcup_{k=0}^{\infty} T^{-k}[i]\right) = 0,$$

για κάθε $i \in S \setminus S'$. Άρα και

$$\begin{aligned}\mu\Big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_k = i \text{ για κάποιο } k \in \mathbb{N}_0, \text{ για κάποιο } i \in S \setminus S'\}\Big) \\ = \mu\left(\bigcup_{i\in S\setminus S'} \bigcup_{k=0}^{\infty} T^{-k}[i]\right) = 0.\end{aligned}$$

Έπεται ότι το μέτρο μ είναι συγκεντρωμένο στον υπόχωρο $X' := (S')^{\mathbb{N}_0}$, δηλαδή $\mu(X') = 1$, και το αρχικό Markov shift είναι ισοδύναμο με το Markov shift που προκύπτει αν διαγράψουμε τις καταστάσεις $i \in S \setminus S'$ από τον χώρο καταστάσεων S, το σύστημα δηλαδή $(X', \mathcal{B}', \mu', T')$, όπου $X' = (S')^{\mathbb{N}_0}$, $\mathcal{B}' = \otimes_{\mathbb{N}_0} \mathcal{P}(S')$, $T' = T|_{X'}$ το shift περιορισμένο στο X' και μ' το μέτρο στον μετρήσιμο χώρο $(X', \mathcal{B}')$ που αντιστοιχεί στον πίνακα $P' = P|_{S'}$ (ο P περιορισμένος στις καταστάσεις στο S') και το διάνυσμα πιθανότητας $p' = (p_i)_{i\in S'}$ (το p περιορισμένο στις καταστάσεις στο S'). Σημειώνεται εδώ ότι ο πίνακας P' που προκύπτει με διαγραφή των καταστάσεων που δεν ανήκουν στο S' είναι στοχαστικός. Πράγματι, προφανώς $P'_{ij} = P_{ij} \geq 0 \ \forall\, i,j \in S'$, αφού αυτό ισχύει ήδη για όλα τα $i,j \in S$, και για κάθε $j \in S \setminus S'$,

$$0 = p_j = \sum_{i\in S} p_i P_{ij}$$

και άρα πρέπει $P_{ij} = 0$ για τα i για τα οποία $p_i \neq 0$, δηλαδή για όλα τα $i \in S'$. Έπεται ότι, για $i \in S'$, $P_{ij} = 0$ για κάθε $j \in S \setminus S'$ και άρα

$$1 = \sum_{j \in S} P_{ij} = \sum_{j \in S'} P_{ij} = \sum_{i \in S'} P'_{ij}.$$

Επίσης το p' αποτελεί αριστερό ιδιοδιάνυσμα του P': για $j \in S'$,

$$p_j = \sum_{i \in S} p_i P_{ij} = \sum_{i \in S'} p_i P_{ij} = \sum_{i \in S'} p'_i P'_{ij},$$

αφού $p_i = 0$ για $i \in S \setminus S'$.

2.6 Ιδιοτιμές και ιδιοδιανύσματα για εργοδικά συστήματα

Σε αυτή την παράγραφο θεωρούμε ένα σ.δ.μ. $(X, \mathcal{B}, \mu, T)$ και τον αντίστοιχο τελεστή Koopman στον $L^2(X, \mathcal{B}, \mu)$. Υπενθυμίζεται ότι ο U_T είναι ισομετρία.

Θεώρημα 2.6.1 Έστω $(X, \mathcal{B}, \mu, T)$ ένα εργοδικό σ.δ.μ. και U_T ο αντίστοιχος τελεστής Koopman στον $L^2(X, \mathcal{B}, \mu)$.

(i) Αν $U_T f = \lambda f$, $f \in L^2(\mu)$, τότε $|\lambda| = 1$ και $|f|$ είναι σταθερή μ-σχεδόν παντού.

(ii) Ιδιοσυναρτήσεις που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι ορθογώνιες μεταξύ τους.

(iii) Κάθε ιδιοτιμή είναι απλή, δηλαδή αν f και g είναι ιδιοσυναρτήσεις που αντιστοιχούν στην ίδια ιδιοτιμή λ, τότε υπάρχει σταθερά c τέτοια ώστε

$$f \overset{L^2}{=} cg.$$

(iv) Οι ιδιοτιμές του U_T αποτελούν υποομάδα της $\mathbb{S}^1$.

Απόδειξη. (i) Αν $U_T f = \lambda f$ για κάποιο $\lambda \in \mathbb{C}$ και $f \in L^2(\mu)$, μη μηδενική, τότε

$$\|f\|_2 = \|U_T f\|_2 = \|\lambda f\|_2 = |\lambda| \|f\|_2.$$

Εφόσον $\|f\|_2 \neq 0$ θα πρέπει $|\lambda| = 1$. Επίσης,

$$U_T \circ |f| = |f| \circ T = |f \circ T| = |U_T f| = |\lambda f| = |\lambda||f| = |f|.$$

Επειδή το σύστημα είναι εργοδικό, πρέπει

$$|f| = c \quad \mu\text{-σχεδόν παντού.}$$

(ii) Αν λ_1, λ_2 δύο διαφορετικές ιδιοτιμές και

$$U_T f_1 = \lambda_1 f_1, \quad U_T\, f_2 = \lambda_2 f_2, \qquad f_1, f_2 \in L^2(\mu) \quad \text{μη μηδενικές,}$$

αντίστοιχες ιδιοσυναρτήσεις, τότε

$$\langle f_1, f_2 \rangle = \langle U_T f_1, U_T f_2 \rangle = \langle \lambda_1 f_1, \lambda_2 f_2 \rangle = \lambda_1 \overline{\lambda_2} \langle f_1, f_2 \rangle.$$

Άρα πρέπει

$$\langle f_1, f_2 \rangle = 0 \qquad \text{ή} \qquad \lambda_1 \overline{\lambda_2} = 1.$$

Όμως, από το (i) έχουμε ότι $|\lambda_2| = 1$ και άρα $\overline{\lambda_2} = 1/\lambda_2$. Αν ήταν $\lambda_1 \overline{\lambda_2} = 1$, τότε θα έπρεπε $\lambda_1 = \lambda_2$, άτοπο. Επομένως

$$\langle f, g \rangle = 0.$$

(iii) Έστω λ ιδιοτιμή του U_T και έστω f, g αντίστοιχες ιδιοσυναρτήσεις. Από το (i) έχουμε ότι

$$|g| = c \quad \mu\text{-σχεδόν παντού}$$

και αφού η g δεν είναι ταυτοτικά 0 (είναι ιδιοσυνάρτηση), θα είναι και $c \neq 0$ και άρα θα ισχύει $g \neq 0$ μ-σχεδόν παντού. Έπεται ότι η

$$h = \frac{f}{g}$$

είναι καλά ορισμένη μ-σχεδόν παντού. Τότε

$$U_T h = h \circ T = \frac{f \circ T}{g \circ T} = \frac{\lambda f}{\lambda g} = \frac{f}{g} = h$$

και η h δεν είναι ταυτοτικά 0, αφού δεν είναι ούτε η f. Έπεται ότι η h είναι ιδιοσυνάρτηση του U_T για την ιδιοτιμή 1, ισοδύναμα η h είναι T-αναλλοίωτη και λόγω εργοδικότητας θα πρέπει να είναι σταθερή.

(iv) Έστω λ_1, λ_2 δύο ιδιοτιμές. Έστω και f_1, f_2 αντίστοιχες ιδιοσυναρτήσεις, δηλαδή $U_T f_1 = \lambda_1 f_1$, $U_T f_2 = \lambda_2 f_2$, $f_1, f_2 \in L^2(\mu)$ και όχι μηδενικές μ-σχεδόν παντού. Τότε

$$U_T \overline{f_2} = \overline{f_2} \circ T = \overline{f_2 \circ T} = \overline{U_T f_2} = \overline{\lambda_2 f_2} = \overline{\lambda_2}\,\overline{f_2}.$$

Άρα το $\lambda_2^{-1} = \overline{\lambda_2}$ είναι επίσης ιδιοτιμή, με ιδιοσυνάρτηση την $\overline{f_2}$. Τώρα

$$U_T\big(f_1 \overline{f_2}\big) = (f_1 \circ T)\ \big(\overline{f_2 \circ T}\big) = \lambda f_1\, \overline{\lambda_2 f_2} = \lambda_1 \overline{\lambda_2}\ f_1 \overline{f_2}.$$

Άρα το $\lambda_1 \overline{\lambda_2}$ είναι ιδιοτιμή με ιδιοσυνάρτηση την $f_1 \overline{f_2}$. Δηλαδή αν λ_1, λ_2 δύο ιδιοτιμές του U_T, τότε $|\lambda_1| = |\lambda_2| = 1$ και $\lambda_1 \lambda_2^{-1}$ είναι ιδιοτιμή. Άρα το σύνολο των ιδιοτιμών είναι υποομάδα της $\mathbb{S}^1$. $\square$

2.7 Επαγόμενος μετασχηματισμός — Kakutani Skyscraper

Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ. και έστω $B \in \mathcal{B}$ με $\mu(B) > 0$. Ορίζουμε

$$N_B(x) = \inf\{n \in \mathbb{N} \colon T^n(x) \in B\},$$

με την σύμβαση $\inf \varnothing = \infty$. Από το θεώρημα Poincaré έχουμε ότι $N_B(x) < \infty$ για σχεδόν κάθε $x \in B$. Ορίζουμε

$$A = \{x \in B \colon T^n(x) \in B \text{ για άπειρα } n \in \mathbb{N}\} = \bigcap_{n \geq 1} \bigcup_{m \geq n} T^{-m} B \,\cap\, B$$

Από το θεώρημα Poincaré έχουμε ότι $\mu(A) = \mu(B)$. Ορίζουμε

$$T_A \colon A \to A, \qquad T_A(x) = T^{N_B(x)}(x).$$

Έστω $\mathcal{B}_A = \{C \cap A \colon C \in \mathcal{B}\} = \{C \in \mathcal{B} \colon C \subset A\}$ και μ_A το μέτρο στον μετρήσιμο χώρο $(A, \mathcal{B}_A)$ με

$$\mu_A(C) = \frac{\mu(C \cap A)}{\mu(A)}.$$

Πρόταση 2.7.1

(i) $N_B \colon X \to \mathbb{N} \cup \{\infty\}$, $N_B \colon A \to \mathbb{N}$ είναι $\mathcal{B}$- και $\mathcal{B}_A$-μετρήσιμες, αντίστοιχα.

(ii) $T_A \colon A \to A$ είναι $\mathcal{B}_A$-μετρήσιμη.

(iii) $(A, \mathcal{B}_A, \mu_A, T_A)$ είναι σ.δ.μ.

Απόδειξη. (i) Γράφουμε $N_B^{-1}(\{1\}) = T^{-1}(B) \in \mathcal{B}$ και για $n \in \mathbb{N}$ με $n > 1$

$$N_B^{-1}(\{n\}) = \{x \in X \colon N_B(x) = n\} = T^{-1}(B^c) \cap T^{-2}(B^c) \cdots T^{-(n-1)}(B^c) \cap T^{-n}(B) \in \mathcal{B},$$

αφού $B \in \mathcal{B}$ και T μετρήσιμη ως προς $\mathcal{B}$. Επίσης, για $n = \infty$,

$$N_B^{-1}(\{\infty\}) = \bigcap_{n \geq 1} T^{-n}(B^c) \in \mathcal{B}.$$

Αν θεωρήσουμε την N_B σαν συνάρτηση στο A, οι αντίστροφες εικόνες είναι $N_B^{-1}(\{1\}) = A \cap T^{-1}(B) \in \mathcal{B}_A$ και για $n \in \mathbb{N}$ με $n > 1$

$$N_B^{-1}(\{n\}) \cap A = A \cap T^{-1}(B^c) \cap T^{-2}(B^c) \cdots T^{-(n-1)}(B^c) \cap T^{-n}(B) \in \mathcal{B}_A.$$

(ii) Έστω $C \in \mathcal{B}_A$. Τότε

$$\begin{aligned}
T_A^{-1}(C) &= \{x \in A \colon T_A(x) \in C\} = \bigcup_{n \geq 1} \{x \in A \colon T_A(x) \in C, N_B(x) = n\} \\
&= \bigcup_{n \geq 1} \{x \in A \colon T^n(x) \in C, N_B(x) = n\} = \bigcup_{n \geq 1} A \cap T^{-n}(C) \cap \{N_B(x) = n\} \\
&= A \cap \underbrace{T^{-1}(C)}_{\in \mathcal{B}} \cap \underbrace{T^{-1}(B)}_{\in \mathcal{B}} \cup \bigcup_{n \geq 2} A \cap \underbrace{T^{-n}(C)}_{\in \mathcal{B}} \cap \underbrace{\bigcap_{k=1}^{n-1} T^{-k}(B^c) \cap T^{-n}(B)}_{\in \mathcal{B}} \in \mathcal{B}_A.
\end{aligned}$$

Άρα η T_A είναι $\mathcal{B}_A$-μετρήσιμη.

(iii) Ορίζουμε

$$A_1 = \{x \in A \colon N_B(x) = 1\} = \{x \in A \colon T(x) \in A\}$$

και

$$B_1 = \{x \in X \setminus A \colon T(x) \in A\},$$

και για $n \in \mathbb{N}$ με $n > 1$,

$$A_n = \{x \in A \colon N_B(x) = n\} = \{x \in A \colon T(x) \notin A, \ldots, T^{n-1}(x) \notin A, T^n(x) \in A\},$$

η δεύτερη ισότητα επειδή, για οποιοδήποτε $i \in \mathbb{N}$, $T^i(x) \in A \Leftrightarrow T^i(x) \in B$ για $x \in A$, και

$$B_n = \{x \in X \setminus A \colon T(x) \notin A, \ldots, T^{n-1}(x) \notin A, T^n(x) \in A\}.$$

Είναι σαφές ότι

$$B_1 \cup A_1 = T^{-1}(A). \tag{21}$$

Επίσης,

$$A_{n+1} \cup B_{n+1} = T^{-1}(B_n) \qquad \forall\, n \in \mathbb{N}. \tag{22}$$

Πράγματι,

$$x \in A_{n+1} \cup B_{n+1} \quad \Leftrightarrow \quad T(x) \notin A, T^2(x) \notin A, \ldots, T^n(x) \notin A, T^{n+1}(x) \in A.$$

Όμως και

$$x \in T^{-1}(B_n) \quad \Leftrightarrow \quad T(x) \notin A, T^2(x) \notin A, \ldots, T^n(x) \notin A, T^{n+1}(x) \in A$$

και άρα έχουμε όντως ισότητα στην (22).

Έστω $C \in \mathcal{B}_A$. Θέλουμε να δείξουμε ότι

$$\mu_A\left(T_A^{-1}(C)\right) = \mu_A(C).$$

Αυτό είναι ισοδύναμο με

$$\mu\left(T_A^{-1}(C)\right) = \mu(C) \quad \Leftrightarrow \quad \mu\left(T_A^{-1}(C)\right) = \mu\left(T^{-1}(C)\right),$$

αφού ο T διατηρεί το μέτρο μ. Τώρα, εφόσον τα A_n, $n \in \mathbb{N}$, αποτελούν διαμέριση του A, γράφουμε

$$T_A^{-1}(C) = \bigcup_{n\geq 1} T_A^{-1}(C) \cap A_n = \bigcup_{n\geq 1} T^{-n}(C) \cap A_n$$

και η ένωση είναι ξένη αφού τα A_n, $n \in \mathbb{N}$ είναι ξένα ανά δύο. Επομένως,

$$\mu\left(T_A^{-1}(C)\right) = \sum_{n\geq 1} \mu\left(T^{-n}(C) \cap A_n\right).$$

Τώρα,

$$\begin{aligned}
\mu\left(T^{-1}(C)\right) &= \mu\left(T^{-1}(C) \cap A_1\right) + \mu\left(T^{-1}(C) \cap B_1\right) \\
&= \mu\left(T^{-1}(C) \cap A_1\right) + \mu\left(T^{-2}(C) \cap T^{-1}(B_1)\right) \\
&= \mu\left(T^{-1}(C) \cap A_1\right) + \mu\left(T^{-2}(C) \cap (A_2 \cup B_2)\right) \\
&= \mu\left(T^{-1}(C) \cap A_1\right) + \mu\left(T^{-2}(C) \cap A_2\right) + \mu\left(T^{-2}(C) \cap B_2\right),
\end{aligned}$$

όπου η πρώτη ισότητα ισχύει λόγω της (21) και αφού A_1, B_1 είναι ξένα μεταξύ τους, η δεύτερη επειδή ο T διατηρεί το μέτρο μ, η τρίτη από την (22) και η τέταρτη επειδή τα A_2 και B_2 είναι εξ' ορισμού ξένα. Επαγωγικά, μετά από n βήματα, θα έχουμε

$$\mu\left(T^{-1}(C)\right) = \sum_{j=1}^{n} \mu\left(T^{-j}(C) \cap A_j\right) + \mu\left(T^{-n}(C) \cap B_n\right) \xrightarrow{n\to\infty} \sum_{j\geq 1} \mu\left(T^{-j}(C) \cap A_j\right),$$

διότι $\mu\left(T^{-n}(C) \cap B_n\right) \leq \mu(B_n)$ και αφού τα B_n είναι ξένα ανά δύο, η σειρά $\sum_{n\geq 1}\mu(B_n) = \mu\left(\bigcup_{n\geq 1} B_n\right) \leq 1$ συγκλίνει και άρα $\mu(B_n) \to 0$. Επομένως,

$$\mu\left(T^{-1}(C)\right) = \mu\left(T_A^{-1}(C)\right)$$

και άρα $(A, \mathcal{B}_A, \mu_A, T_A)$ είναι σ.δ.μ. □

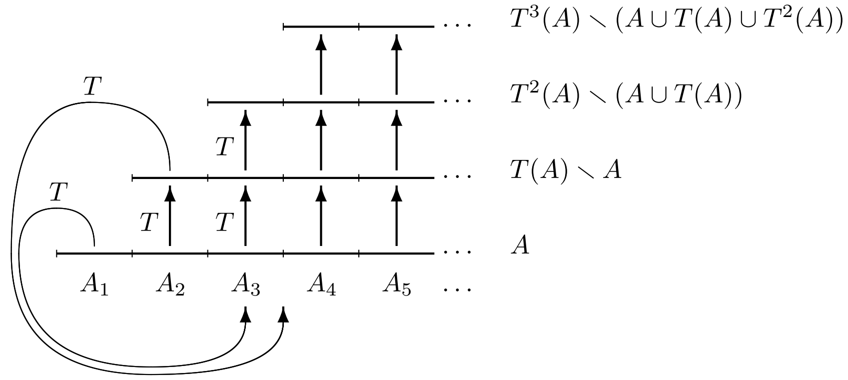
Από εδώ και στο εξής θα θεωρούμε δεδομένο ότι το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι αντιστρέψιμο σ.δ.μ. Δοθέντος $B \in \mathcal{B}$ με $\mu(B) > 0$, θεωρούμε αντί για το A όπως ορίστηκε για γενικά, όχι κατ' ανάγκη αντιστρέψιμα, συστήματα, το σύνολο

$$A = B \cap \bigcap_{n\geq 1} \bigcup_{m\geq n} T^{-m}(B) \cap \bigcap_{n\geq 1} \bigcup_{m\geq n} T^{m}(B) \tag{23}$$

ή αλλιώς

$$A = B \cap \limsup_n T^{-n}(B) \cap \limsup_n T^{n}(B).$$

Εφαρμόζοντας δύο φορές το θεώρημα Poincare, βλέπουμε ότι $\mu(A) = \mu(B)$.



Θεώρημα 2.7.2 Έστω $(X,\mathcal{B},\mu,T)$ ένα αντιστρέψιμο και εργοδικό σ.δ.μ. Έστω $B \in \mathcal{B}$ με $\mu(B) > 0$ και για αυτό το B έστω A το σύνολο που ορίζεται στην (23). Τότε το επαγώμενο σύστημα $(X_A,\mathcal{B}_A,\mu_A,T_A)$ είναι επίσης εργοδικό.

Απόδειξη Έστω ότι το $(X_A,\mathcal{B}_A,\mu_A,T_A)$ δεν είναι εργοδικό. Τότε θα υπάρχει $C \in \mathcal{B}_A$ τέτοιο ώστε $T_A(C) = C$ και $0 < \mu_A(C) < 1$. Χρησιμοποιούμε εδώ το γεγονός ότι το σύστημα $(X_A,\mathcal{B}_A,\mu_A,T_A)$ είναι αντιστρέψιμο όταν το αρχικό σύστημα $(X,\mathcal{B},\mu,T)$ είναι αντιστρέψιμο (άσκηση). Τότε το σύνολο

$$\tilde{C} = \bigcup_{n\geq 1}\bigcup_{k=0}^{n-1} T^k(A_n \cap C)$$

είναι T-αναλλοίωτο. Πράγματι,

$$\begin{aligned}
T(\tilde{C}) = \bigcup_{n\geq 1}\bigcup_{k=0}^{n} T^k(A_n \cap C) &= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, \bigcup_{n\geq 1} T^n(A_n \cap C) \\
&= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, \bigcup_{n\geq 1} T_A(A_n \cap C) \\
&= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, T_A\left(\bigcup_{n\geq 1} A_n \cap C\right) \\
&= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, T_A(A \cap C) \\
&= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, T_A(C) \\
&= \bigcup_{n\geq 1}\bigcup_{k=1}^{n-1} T^k(A_n \cap C) \,\cup\, C.
\end{aligned}$$

Άρα, αφού $C = A \cap C = \bigcup_{n\geq 1} A_n \cap C = \bigcup_{n\geq 1} T^0(A_n \cap C)$,

$$T(\tilde{C}) = \bigcup_{n\geq 1}\bigcup_{k=0}^{n-1} T^k(A_n \cap C) = \tilde{C},$$

δηλαδή το $\tilde{C}$ είναι όντως T-αναλλοίωτο. Επίσης

$$\mu(\tilde{C}) = \mu\left(\bigcup_{n\geq 1}\bigcup_{k=0}^{n-1} T^k(A_n \cap C)\right) \geq \mu\left(\bigcup_{n\geq 1}(A_n \cap C)\right) = \mu(C) > 0$$

και

$$\mu(X \setminus \tilde{C}) = \mu\left(\bigcap_{n\geq 1}\bigcap_{k=0}^{n-1}(X \setminus T^k(A_n \cap C))\right) \geq \mu(A \setminus C) > 0,$$

επειδή $A \setminus C \subseteq \bigcap_{n\geq 1}\bigcap_{k=0}^{n-1}(X \setminus T^k(A_n \cap C))$. Πράγματι, αν $x \in A \setminus C$, τότε

$$\begin{aligned}
x \in A \quad &\Rightarrow \quad x \notin T^k(A_n) \quad \forall\, k \in \{1,\ldots,n-1\} \quad \forall\, n \geq 2 \\
&\Rightarrow \quad x \notin T^k(A_n \cap C) \quad \forall\, k \in \{1,\ldots,n-1\} \quad \forall\, n \geq 2,
\end{aligned}$$

η πρώτη συνεπαγωγή επειδή $x \in A_n$ σημαίνει ότι $T(x) \notin A,\ldots,T^{n-1}(x) \notin A$ και η δεύτερη επειδή $T^k(A_n \cap C) \subset T^k(A_n)$ για κάθε k και n, και

$$x \notin C \quad \Rightarrow \quad x \notin A_n \cap C \quad \forall\, n \geq 1.$$

Επομένως

$$0 < \mu(\tilde{C}) < 1$$

και άρα βρήκαμε ένα μη τετριμμένο αναλλοίωτο $\tilde{C}$ στο εργοδικό σύστημα $(X, \mathcal{B}, \mu, T)$, που είναι άτοπο. Άρα πρέπει και το επαγόμενο σύστημα να είναι εργοδικό. $\square$

Θεώρημα 2.7.4 (Kac) Έστω $(X, \mathcal{B}, \mu, T)$ εργοδικό σ.δ.μ. Αν $A \in \mathcal{B}$ με $\mu(A) > 0$ τότε

$$\int_A N_A(x)\, d\mu(x) = 1.$$

Δηλαδή, $\int_A N_A(x)\, d\mu_A(x) = 1/\mu(A)$ και άρα ο μέσος χρόνος επανόδου στο A είναι $1/\mu(A)$.

Απόδειξη. Έστω πάλι

$$A_1 = \{x \in A \colon N(x) = 1\} = \{x \in A \colon T(x) \in A\}$$

και

$$B_1 = \{x \in X \setminus A \colon N(x) = 1\} = \{x \in X \setminus A \colon T(x) \in A\},$$

και για $n \in \mathbb{N}$ με $n > 1$,

$$A_n = \{x \in A \colon N_A(x) = n\} = \{x \in A \colon T(x) \notin A, \dots, T^{n-1}(x) \notin A, T^n(x) \in A\},$$

και

$$B_n = \{x \in X \setminus A \colon N(x) = n\} = \{x \in X \setminus A \colon T(x) \notin A, \dots, T^{n-1}(x) \notin A, T^n(x) \in A\}.$$

Όπως έχουμε δείξει στην (22),

$$A_{n+1} \cup B_{n+1} = T^{-1}(B_n) \qquad \forall\, n \in \mathbb{N}.$$

Επομένως

$$\mu(B_n) = \mu\big(T^{-1}(B_n)\big) = \mu(A_{n+1}) + \mu(B_{n+1})$$

για κάθε $n \in \mathbb{N}$, και επαγωγικά,

$$\begin{aligned} \mu(B_n) = \mu\big(T^{-1}(B_n)\big) = \mu(A_{n+1}) + \mu(B_{n+1}) &= \mu(A_{n+1}) + \mu\big(T^{-1}(B_{n+1})\big) \\ &= \mu(A_{n+1}) + \mu(A_{n+2}) + \mu(B_{n+1}) = \cdots \\ &= \sum_{j=n+1}^{m} \mu(A_j) + \mu(B_m), \end{aligned}$$

για κάθε $n, m \in \mathbb{N}$ με $m > n$. Επειδή δε τα B_m, $m \in \mathbb{N}$, είναι ξένα ανά δύο, έχουμε πάλι ότι $\sum_{m\in\mathbb{N}} \mu(B_m) = \mu\left(\bigcup_{m\in\mathbb{N}} B_m\right) \leq 1 < +\infty$ και επομένως $\mu(B_m) \to 0$ καθώς $m \to \infty$. Έπεται τώρα ότι

$$\mu(B_n) = \sum_{j=n+1}^{\infty} \mu(A_j),$$

για κάθε $n \in \mathbb{N}$.

Ορίζουμε τώρα και

$$A_\infty = \{x \in A \colon N_A(x) = \infty\} = A \cap \bigcap_{n=1}^{\infty} T^{-n}(A^c)$$

και

$$B_\infty = \{x \in X \setminus A \colon N_A(x) = \infty\} = \bigcap_{n=0}^{\infty} T^{-n}(A^c).$$

Τα σύνολα A_n, B_n, $n \in \mathbb{N} \cup \{\infty\}$, διαμερίζουν τον X. Επίσης, από το θεώρημα Poincare $\mu(A_\infty) = 0$. Έπεται ότι

$$\begin{aligned} 1 = \mu(X) = \mu(A\infty) + \mu(B_\infty) + \sum_{n=1}^{\infty}[\mu(A_n) + \mu(B_n)] &= \mu(B_\infty) + \sum_{n=1}^{\infty}[\mu(A_n) + \mu(B_n)] \\ &= \mu(B_\infty) + \sum_{n=1}^{\infty}\sum_{j=n}^{\infty} \mu(A_j)] \\ &= \mu(B_\infty) + \sum_{n=1}^{\infty} \mu\big(\{x \in A \colon N_A(x) \geq n\}\big) \\ &= \mu(B_\infty) + \int_A N_A\, d\mu. \end{aligned}$$

Επειδή όμως το σύστημα είναι εργοδικό, και $\mu(A) > 0$, έχουμε ότι $\mu(B_\infty^c) = \mu\left(\bigcup_{n=0}^{\infty} T^{-n}(A)\right) = 1$ και άρα $\mu(B_\infty) = 0$. $\square$

Σημείωση. Γενικά, $\int_A N_A(x)\,d\mu(x) = 1 - \mu(B_\infty) \leq 1$, για συστήματα όχι απαραίτητα εργοδικά.

Θεώρημα 2.7.5 (Kakutani–Rokhlin) Έστω $(X, \mathcal{B}, \mu, T)$ ένα αντιστρέψιμο εργοδικό σ.δ.μ. Υποθέτουμε ότι το μέτρο μ δεν έχει άτομα. Τότε για κάθε $n \in \mathbb{N}$ και $\varepsilon > 0$, υπάρχει $C \in \mathcal{B}$ τέτοιο ώστε

(1) $C, T(C), \ldots, T^{n-1}(C)$ να είναι ξένα ανά δύο.

(2) $\mu\big(C \cup T(C) \cup \cdots \cup T^{n-1}(C)\big) > 1 - \varepsilon$.

Απόδειξη. Επειδή το μέτρο μ δεν έχει άτομα, υπάρχει $B \in \mathcal{B}$ τέτοιο ώστε

$$0 < \mu(B) < \frac{\varepsilon}{n-1},$$

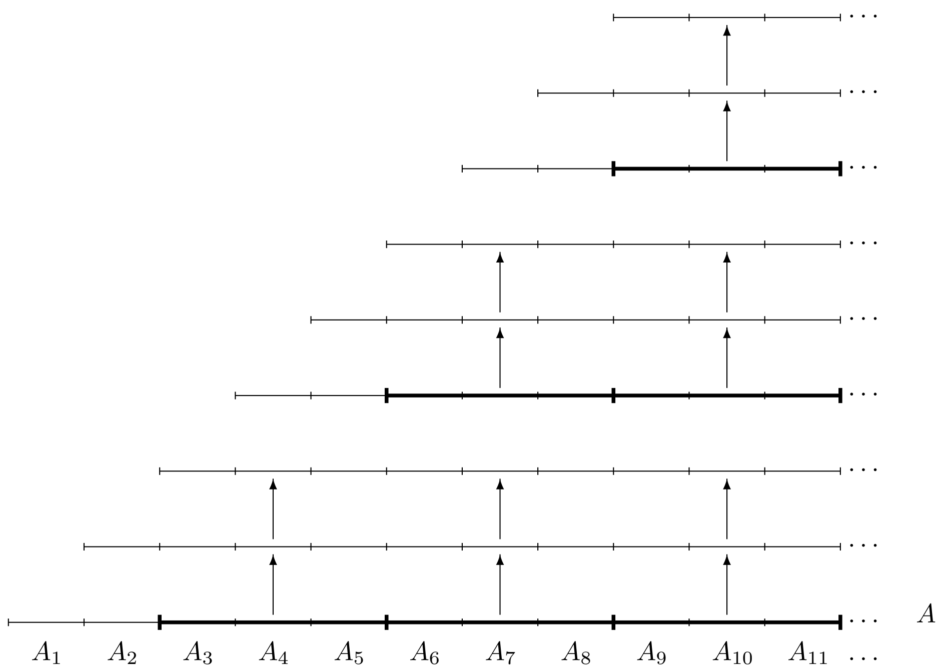
για δοθέντα $\varepsilon > 0$ και $n \in \mathbb{N}$. Θεωρούμε το σύνολο A όπως στην κατασκευή του πύργου Kakutani, δηλαδή το σύνολο που ορίζεται στην (23) για το συγκεκριμένο B. Τότε $\mu(A) = \mu(B)$. Ορίζουμε

$$C := \bigcup_{k \geq n} \bigcup_{j=0}^{\lfloor k/n \rfloor - 1} T^{jn}(A_k) \in \mathcal{B}.$$

Ισοδύναμα, και αφού $\lfloor k/n \rfloor = j \Leftrightarrow jn \leq k < (j+1)n$,

$$\begin{aligned} C = {} & A_n \cup \cdots \cup A_{2n-1} \\ & \cup A_{2n} \cup \cdots \cup A_{3n-1} \cup T^n(A_{2n}) \cup \cdots \cup T^n(A_{3n-1}) \\ & \cup A_{3n} \cup \cdots \cup A_{4n-1} \cup T^n(A_{3n}) \cup \cdots \cup T^n(A_{4n-1}) \cup T^{2n}(A_{3n}) \cup \cdots \cup T^{2n}(A_{4n-1}) \\ & \cup \cdots \end{aligned}$$

Στο σχήμα παρακάτω, $n = 3$ και το σύνολο C είναι η ένωση των έντονα μαύρων γραμμών.



Από την κατασκευή, τα $C, T(C), \ldots, T^{n-1}(C)$ είναι ξένα ανά δύο. Πράγματι, καταρχήν

$$T^i(C) := \bigcup_{k \geq n} \bigcup_{j=0}^{\lfloor k/n \rfloor - 1} T^{jn+i}(A_k) \qquad i \in \{0, 1, \ldots, n-1\}.$$

Ισχυρισμός 1: Αν $k_1, k_2 \in \mathbb{N}$, $l_1, l_2 \in \mathbb{N} \cup \{0\}$ και $0 \le l_1 < k_1$, $0 \le l_2 < k_2$ και $l_1 \ne l_2$, τότε $T_{l_1}(A_{k_1}) \cap T_{l_2}(A_{k_2}) = \varnothing$.

Απόδειξη Ισχυρισμού 1: Έστω $x \in T^{l_1}(A_{k_1}) \cap T^{l_2}(A_{k_2})$. Τότε υπάρχουν $x_1 \in A_{k_1}$ και $x_2 \in A_{k_2}$ τέτοια ώστε $T^{l_1}(x_1) = x = T^{l_2}(x_2)$. Έστω, χωρίς βλάβη της γενικότητας φυσικά, ότι $l_1 < l_2$. Τότε $T^{l_2 - l_1}(x_2) = x_1$, με τα $x_1 \in A$, $x_2 \in A_{k_2}$ και $l_2 - l_1 \le l_2 < k_2$. Αυτό όμως είναι άτοπο, από τον ορισμό του A_{k_2}. □

Αν $i_1, i_2 \in \{0, 1, \dots, n-1\}$ και $k_1, k_2 \ge n$, $j_1 \in \{0, 1, \dots, \lfloor k_1/n \rfloor - 1\}$, $j_2 \in \{0, 1, \dots, \lfloor k_2/n \rfloor - 1\}$, τότε $T^{j_1 n + i_1}(A_{k_1}) \cap T^{j_2 n + i_2}(A_{k_2}) = \varnothing$ αν $j_1 n + i_1 \ne j_2 n + i_2$, από τον ισχυρισμό και αφού $j_l n + i_l < k_l$ και για τα δύο $l \in \{1, 2\}$. Αν πάλι $j_1 n + i_1 = j_2 n + i_2$, τότε επειδή $|i_2 - i_1| < n$ και $|j_1 - j_2| n = |i_2 - i_1|$, πρέπει, καταρχήν $j_1 = j_2$ και κατόπιν και $i_1 = i_2$. Άρα $T^{j_1 n + i_1}(A_{k_1}) \cap T^{j_2 n + i_2}(A_{k_2}) = \varnothing$ αν $i_1, i_2 \in \{0, 1, \dots, n-1\}$ με $i_1 \ne i_2$, για κάθε $k_1, k_2 \ge n$, $j_1 \in \{0, 1, \dots, \lfloor k_1/n \rfloor - 1\}$, $j_2 \in \{0, 1, \dots, \lfloor k_2/n \rfloor - 1\}$, από όπου έπεται ότι $T^{i_1}(C) \cap T^{i_2}(C) = \varnothing$, για $i_1, i_2 \in \{0, 1, \dots, n-1\}$ με $i_1 \ne i_2$.

Επίσης,

$$\begin{aligned} C \cup T(C) \cup \dots \cup T^{n-1}(C) &= \bigcup_{i=0}^{n-1} T^i(C) = \bigcup_{i=0}^{n-1} \bigcup_{k \ge n} \bigcup_{j=0}^{\lfloor k/n \rfloor - 1} T^{jn+i}(A_k) \\ &= \bigcup_{k \ge n} \bigcup_{j=0}^{\lfloor k/n \rfloor - 1} \bigcup_{i=0}^{n-1} T^{jn+i}(A_k) = \bigcup_{k \ge n} \bigcup_{m=0}^{\lfloor k/n \rfloor n - 1} T^m(A_k), \end{aligned}$$

αφού $\bigcup_{j=0}^{\lfloor k/n \rfloor - 1} \{jn + i \colon i \in \{0, 1, \dots, n-1\}\} = \{0, 1, \dots, \lfloor k/n \rfloor n - 1\}$ και

$$\bigcup_{k \ge n} \bigcup_{m=0}^{\lfloor k/n \rfloor n - 1} T^m(A_k) = \bigcup_{j=1}^{\infty} \bigcup_{k=jn}^{(j+1)n-1} \bigcup_{m=0}^{jn-1} T^m(A_k).$$

Ισχυρισμός 2:

$$\bigcup_{m \ge 0} T^m(A) = \bigcup_{k=1}^{\infty} \bigcup_{m=0}^{k-1} T^m(A_k).$$

Απόδειξη Ισχυρισμού 2: Το δεξί μέλος της παραπάνω ισότητας προφανώς περιέχεται στο αριστερό. Έστω τώρα $y \in \bigcup_{m \ge 0} T^m(A)$. Τότε υπάρχει $x \in A$ με $T^m(x) = y$ για κάποιο $m \in \mathbb{N} \cup \{0\}$. Το σύνολο $\{n \in \mathbb{N} \cup \{0\} \colon n \le m,\ T^n(x) \in A\}$ είναι μη κενό (αφού περιέχει το 0) και έστω j το μέγιστο στοιχείο του. Αν $j = m$, τότε $y \in A = \bigcup_{k \in \mathbb{N}} A_k = \bigcup_{k \in \mathbb{N}} T^0(A_k) \subseteq \bigcup_{k \in \mathbb{N}} \bigcup_{i=0}^{k-1} T^i(A_k)$. Αν πάλι $j < m$, τότε $T^j(x) \in A, T(T^j(x)) \notin A, \dots T^{m-j}(T^j(x)) = T^m(x) \notin A$ και επομένως $T^j(x) \in A_k$ για κάποιο $k > m - j$. Τότε $y = T^m(x) = T^{m-j}(T^j(x)) \in T^{m-j}(A_k) \subseteq \bigcup_{i=0}^{k-1}(A_k)$ για αυτό το k και άρα $y \in \bigcup_{k \in \mathbb{N}} \bigcup_{i=0}^{k-1} T^i(A_k)$. Δηλαδή $y \in \bigcup_{k \in \mathbb{N}} \bigcup_{i=0}^{k-1} T^i(A_k)$ σε κάθε περίπτωση. □

Γράφουμε την παραπάνω ισότητα του Ισχυρισμού 2 σαν

$$\bigcup_{m \ge 0} T^m(A) = \bigcup_{k=1}^{n-1} \bigcup_{m=0}^{k-1} T^m(A_k) \cup \bigcup_{j=1}^{\infty} \bigcup_{k=jn}^{(j+1)n-1} \bigcup_{m=0}^{k-1} T^m(A_k).$$

Έπεται τώρα ότι

$$\bigcup_{m \ge 0} T^m(A) \setminus [C \cup T(C) \cup \dots \cup T^{n-1}(C)] = \bigcup_{k=1}^{n-1} \bigcup_{m=0}^{k-1} T^m(A_k) \cup \bigcup_{j=1}^{\infty} \bigcup_{k=jn}^{(j+1)n-1} \bigcup_{m=jn}^{k-1} T^m(A_k)$$

και επειδή τα σύνολα $T^m(A_k)$ είναι ξένα ανά δύο και το καθένα με μέτρο $\mu(A_k)$,

$$\begin{aligned}\mu\left(\bigcup_{m\geq 0} T^m(A)\setminus\left[C\cup T(C)\cup\cdots\cup T^{n-1}(C)\right]\right) &= \sum_{k=1}^{n-1}\sum_{m=0}^{k-1}\mu\big(T^m(A_k)\big)+\sum_{j=1}^{\infty}\sum_{k=jn}^{(j+1)n-1}\sum_{m=jn}^{k-1}\mu\big(T^m(A_k)\big)\\ &= \sum_{k=1}^{n-1}\sum_{m=0}^{k-1}\mu(A_k)+\sum_{j=1}^{\infty}\sum_{k=jn}^{(j+1)n-1}\sum_{m=jn}^{k-1}\mu(A_k)\\ &= \sum_{k=1}^{n-1}k\mu(A_k)+\sum_{j=0}^{\infty}\sum_{k=jn}^{(j+1)n-1}(k-jn)\mu(A_k)\\ &\leq (n-1)\sum_{k\geq 1}\mu(A_k)=(n-1)\mu(A)=(n-1)\mu(B)<\varepsilon.\end{aligned}$$

Τέλος, επειδή το σύστημα έχει υποτεθεί εργοδικό, και επομένως και το $(X,\mathcal{B},\mu,T^{-1})$ είναι εργοδικό, αφού ο T και ο T^{-1} έχουν τα ίδια αναλλοίωτα σύνολα, έχουμε από την Πρόταση 2.1.4 (3) ότι $\mu\left(\bigcup_{m=0}^{\infty}T^m(A)\right)=1$, αφού και $\mu(A)>0$. Έπεται τώρα και ότι

$$\mu\big(C\cup T(C)\cup\cdots\cup T^{n-1}(C)\big)>1-\varepsilon. \qquad \square$$

3 Τοπολογικά Δυναμικά Συστήματα

3.1 Ορισμός και Παραδείγματα

Ορισμός 3.1.1 (τοπολογικό δυναμικό σύστημα) Έστω X συμπαγής μετρικός χώρος και $T\colon X\to X$ συνεχής μετασχηματισμός. Τότε το ζεύγος (X,T) λέγεται *τοπολογικό δυναμικό σύστημα* (τ.δ.σ.).

Ένα τ.δ.σ. είναι αντιστρέψιμο αν ο $T\colon X\to X$ είναι ομοιομορφισμός.

Παράδειγμα 3.1.2 Έστω $X=\mathbb{T}\simeq\mathbb{R}/\mathbb{Z}\simeq\mathbb{S}^1$, $a\in[0,1)$ και $T(x)=(x+a)\pmod 1$. Μετρική στον $\mathbb{T}$ έχει οριστεί η

$$d_{\mathbb{T}}(x,y)=\min\{|x-y|,|x-y-1|\}.$$

Το παραπάνω είναι ένα τ.δ.σ. αφού ο T είναι συνεχής.

Παράδειγμα 3.1.3 (a) (Ενδομορφισμοί του τόρου) Έστω $X=\mathbb{T}$ και $T(x)=kx\pmod 1$, όπου $k\in\mathbb{N}$, $k>1$. Το (X,T) είναι τ.δ.σ. (με την μετρική του προηγούμενου παραδείγματος στον $X=\mathbb{T}$).

Παράδειγμα 3.1.3 (b) (Ενδομορφισμοί του d-διάστατου τόρου) Έστω $X=\mathbb{T}^d$, $T(x)=Ax\pmod 1$, όπου A ένας $d\times d$ πίνακας με $A_{ij}\in\mathbb{Z}$ για κάθε $i,j\in\{1,\dots,d\}$ και $\det(A)\neq 0$. Το (X,T) είναι τ.δ.σ.

Παράδειγμα 3.1.4 (a) Έστω $S=\{1,2,\dots,s\}$ ένα πεπερασμένο σύνολο, $X=S^{\mathbb{N}}$ και $T\colon X\to X$ ο μετασχηματισμός shift. Ορίζουμε

$$d\big((x_1,x_2,\dots),(y_1,y_2,\dots)\big)=2^{-n},\quad \text{όπου}\quad n=\inf\{k\in\mathbb{N}\colon x_k\neq y_k\},$$

με την σύμβαση $\inf\varnothing=0$, οπότε η απόσταση σε αυτή την περίπτωση, όπου $(x_1,x_2,\dots)=(y_1,y_2,\dots)$, θα είναι 0. Η d είναι μετρική, συμβατή με την τοπολογία γινόμενο στον X που επάγεται από την διακριτή τοπολογία σε κάθε παράγοντα S του γινομένου $X=S^{\mathbb{N}}$, και άρα ο μετρικός χώρος (X,d) είναι συμπαγής. Το shift T είναι συνεχής μετασχηματισμός ως προς αυτή την μετρική. Άρα το (X,T) είναι τ.δ.σ.

Παράδειγμα 3.1.4 (b) Έστω πάλι $S=\{1,2,\dots,s\}$ πεπερασμένο σύνολο και τώρα $X=S^{\mathbb{Z}}$. $T\colon X\to X$ είναι πάλι ο μετασχηματισμός shift. Ορίζουμε την μετρική

$$d\big((x_1,x_2,\dots),(y_1,y_2,\dots)\big)=2^{-n(\underset{\sim}{x},\underset{\sim}{y})}$$

όπου

$$n(\underset{\sim}{x},\underset{\sim}{y}) = \inf\{n \in \mathbb{N}_0 \colon x_i = y_i \quad \forall i \in \{-n+1,\dots,0,\dots,n-1\} \quad \text{και} \quad x_n \neq y_n \quad \text{ή} \quad x_{-n} \neq y_{-n}\}$$

Η d είναι μετρική, συμβατή πάλι με την τοπολογία γινόμενο στον X που επάγεται από την διακριτή τοπολογία σε κάθε παράγοντα S του γινομένου $X = S^{\mathbb{Z}}$, και άρα πάλι ο μετρικός χώρος (X,d) είναι συμπαγής. Ο T είναι τώρα ομοιομορφισμός ως προς αυτή την μετρική. Άρα το (X,T) είναι αντιστρέψιμο τ.δ.σ.

3.2 Ελαχιστικά Τοπολογικά Δυναμικά Συστήματα

Ορισμός 3.2.1 Έστω (X,T) τ.δ.σ. Ένα $A \subset X$ λέγεται *αναλλοίωτο* αν $T(A) \subset A$ ή $A \subset T^{-1}(A)$.

Αν υπάρχει $E \subset X$ κλειστό και τέτοιο ώστε $T(E) \subset E$, τότε το σύστημα $(E, T|_E)$ είναι τ.δ.σ.

Ορισμός 3.2.2 (ελαχιστικό τ.δ.σ.) Ένα τ.δ.σ. (X,T) λέγεται *ελαχιστικό* (*minimal*), αν για κάθε $E \subset X$ κλειστό και αναλλοίωτο, δηλαδή τέτοιο ώστε $T(E) \subset E$, έχουμε ότι $E = X$ ή $E = \varnothing$.

Ορισμός 3.2.3 Ένα σημείο $x \in X$ σε ένα τ.δ.σ. (X,T) λέγεται *περιοδικό*, αν υπάρχει $n \in \mathbb{N}$ με $T^n(x) = x$. (Ο μικρότερος φυσικός n με αυτήν την ιδιότητα λέγεται *περίοδος* του x.)

Παρατήρηση. Αν X είναι άπειρο σύνολο, τότε αν το σύστημα είναι ελαχιστικό δεν μπορεί να έχει περιοδικά σημεία. Πράγματι, αν $T^n(x) = x$ για κάποιο $n \in \mathbb{N}$, τότε $\{x, T(x), \dots, T^{n-1}(x)\}$ είναι κλειστό και αναλλοίωτο σύνολο. Αν X είναι άπειρο σύνολο, η τροχιά αυτή θα είναι ένα μη τετριμμένο, κλειστό, αναλλοίωτο υποσύνολο, άτοπο.

Πρόταση 3.2.4 Έστω (X,T) τ.δ.σ. Τα εξής είναι ισοδύναμα.

(1) (X,T) είναι ελαχιστικό.

(2) $\overline{\{T^n(x) \colon n \in \mathbb{N}_0\}} = X$ για κάθε $x \in X$.

(3) Για κάθε $U \subset X$ ανοικτό και μη κενό, $\bigcup_{n\geq 0} T^{-n}(U) = X$.

Απόδειξη. (1) $\Rightarrow$ (2) Έστω $x \in X$. Το $\{T^n(x) \colon n \in \mathbb{N}_0\}$ είναι αναλλοίωτο σύνολο και άρα το $E = \overline{\{T^n(x) \colon n \in \mathbb{N}_0\}}$ είναι κλειστό και αναλλοίωτο, εφόσον

$$T(E) \subset \overline{T(\{T^n(x) \colon n \in \mathbb{N}_0\})} = \overline{\{T^n(x) \colon n \in \mathbb{N}\}} \subset E$$

επειδή ο T είναι συνεχής. Επίσης, $E \neq \varnothing$ $(x \in E)$ άρα $E = X$. Επομένως η τροχιά ενός τυχόντος x είναι πυκνή στον X.

(2) $\Rightarrow$ (3) Έστω U ανοικτό και μη κενό υποσύνολο του X και έστω $x \in X$. Αφού $\overline{\{T^n(x) \colon n \in \mathbb{N}_0\}} = X$, δηλαδή η τροχιά είναι πυκνή, πρέπει να τέμνει κάθε μη κενό ανοικτό υποσύνολο του X και άρα

$$\{T^n(x) \colon n \in \mathbb{N}_0\} \cap U \neq \varnothing.$$

Επομένως υπάρχει $n \in \mathbb{N}_0$ τέτοιο ώστε $T^n(x) \in U$, ή ισοδύναμα $x \in T^{-n}(U)$ και άρα $x \in \bigcup_{n\geq 0} T^{-n}(U)$. Αφού το x ήταν τυχόν σημείο του X έπεται ότι $X = \bigcup_{n\geq 0} T^{-n}(U)$.

(3) $\Rightarrow$ (1) Έστω $E \subset X$, $E \neq X$, κλειστό και αναλλοίωτο. Τότε $E \subset T^{-1}(E)$. Αν $U = X \setminus E$, τότε $U \neq \varnothing$, το U είναι ανοικτό και $T^{-1}(U) \subset U$. Από την υπόθεση,

$$U \neq \varnothing \text{ ανοικτό} \quad \Rightarrow \quad X = \bigcup_{n\geq 0} T^{-n}(U)$$

και

$$T^{-1}(U) \subset U \quad \Rightarrow \quad \bigcup_{n\geq 0} T^{-n}(U) \subset U = X \setminus E.$$

Επομένως θα πρέπει $E = \varnothing$. $\square$

Σημείωση. Η ιδιότητα (3) ενός ελαχιστικού συστήματος στην παραπάνω Πρόταση είναι το τοπολογικό ανάλογο της ιδιότητας (3) στην Πρόταση 2.1.4 για ένα εργοδικό σύστημα.

Ορισμός 3.2.5′ Αν (X, T) ένα αντιστρέψιμο τ.δ.σ., τότε ορίζεται και η εξής έννοια *αμφίπλευρης ελαχιστικότητας*: για κάθε κλειστό $E \subset X$ τέτοιο ώστε $T(E) = E$, ισχύει ότι είτε $E = \varnothing$ είτε $E = X$.

Πρόταση 3.2.6 Έστω (X, T) αντιστρέψιμο τ.δ.σ. Τα εξής είναι ισοδύναμα

(1) (X, T) είναι αμφίπλευρα ελαχιστικό.

(2) $\overline{\{T^n(x) \colon n \in \mathbb{Z}\}} = X$ για κάθε $x \in X$.

(3) Για κάθε $U \subset X$ ανοικτό και μη κενό, $\bigcup_{n\in\mathbb{Z}} T^n(U) = X$.

Για αντιστρέψιμα συστήματα είναι προφανές ότι ελαχιστικότητα συνεπάγεται την αμφίπλευρη ελαχιστικότητα.

Πρόταση 3.2.7 Αν (X, T) είναι αντιστρέψιμο τ.δ.σ. ελαχιστικότητα και αμφίπλευρη ελαχιστικότητα είναι ισοδύναμα.

Λήμμα 3.2.8. Έστω (X, T) τ.δ.σ. και E κλειστό, μη κενό σύνολο με $T(E) \subset E$. Τότε υπάρχει $F \subset E$ κλειστό και $F \neq \varnothing$ τέτοιο ώστε
$$T(F) = F.$$

Απόδειξη Πρότασης 3.2.7 Με βάση το παραπάνω Λήμμα, αν $E \neq \varnothing$ με $T(E) \subset E$, υπάρχει $F \subset E$, $F \neq \varnothing$, με $T(F) = F$. Λόγω αμφίπλευρης ελαχιστικότητας πρέπει $F = X$ και επομένως και $E = X$. $\square$

Απόδειξη Λήμματος 3.2.8 Έστω $E \subset X$ κλειστό και μη κενό με $T(E) \subset E$. Τότε,
$$E \supset T(E) \supset T^2(E) \supset \cdots .$$
Επομένως η $\{T^n(E) \colon n \in \mathbb{N}_0\}$ είναι φθίνουσα ακολουθία συμπαγών, μη κενών συνόλων στον X. Αν
$$F = \bigcap_{n\in\mathbb{N}_0} T^n(E),$$
το F είναι μη κενό, κλειστό σύνολο και $T(F) \subset F$.

Για τον αντίστροφο εγκλεισμό, έστω $x \in F$. Τότε για κάθε $n \in \mathbb{N}_0$, $x \in T^n(E)$. Έπεται ότι
$$T^{-1}\big(\{x\}\big) \cap T^{n-1}(E) \neq \varnothing \qquad \forall\, n \in \mathbb{N}.$$
Πράγματι, αν $x \in T^n(E)$, υπάρχει $y \in E$ με $T^n(y) = x$ και τότε $T^{n-1}(y) \in T^{-1}\big(\{x\}\big) \cap T^{n-1}(E)$. Τώρα, η ακολουθία
$$T^{-1}\big(\{x\}\big) \cap T^{n-1}(E), \qquad n \in \mathbb{N},$$
είναι φθίνουσα ακολουθία, μη κενών, συμπαγών συνόλων στον συμπαγή χώρο X. Επομένως
$$\bigcap_{n\geq 1} T^{-1}\big(\{x\}\big) \cap T^{n-1}(E) \neq \varnothing.$$
Δηλαδή
$$\bigcap_{n\geq 1} T^{-1}\big(\{x\}\big) \cap T^{n-1}(E) = T^{-1}\big(\{x\}\big) \cap \bigcap_{n\geq 1} T^{n-1}(E) = T^{-1}\big(\{x\}\big) \cap F \neq \varnothing.$$
Άρα υπάρχει $y \in F$ τέτοιο ώστε $T(y) = x$ Άρα $x \in T(F)$. $\square$

Παρατήρηση. Αν (X, T) είναι ελαχιστικό τ.δ.σ., τότε η T είναι επί, γιατί διαφορετικά το $T(X)$ θα είναι γνήσιο και μη κενό, κλειστό και αναλλοίωτο υποσύνολο του X, που είναι αδύνατον.

Θεώρημα 3.2.9 (Birkhoff) Κάθε τ.δ.σ. (X,T) $(X \neq \varnothing)$ περιέχει υποσύστημα που είναι ελαχιστικό. Δηλαδή υπάρχει κλειστό $E \subset X$ τέτοιο ώστε $T(E) \subset E$ και το σύστημα $(E, T|_E)$ να είναι ελαχιστικό.

Απόδειξη. Έστω
$$\mathcal{A} = \{F \subset X \colon F \neq \varnothing, F \text{ κλειστό και } T(F) \subset F\}.$$
Τότε $\mathcal{A} \neq \varnothing$ $(X \in \mathcal{A})$. Εφοδιάζουμε το $\mathcal{A}$ με μερική διάταξη τον εγκλεισμό συνόλων: $F_1 \subset F_2$. Έστω $\mathcal{B}$ ολικά διατεταγμένο υποσύνολο του $\mathcal{A}$. Τότε το $\mathcal{B}$ έχει ελάχιστο στοιχείο, την τομή όλων των στοιχείων του $\mathcal{B}$. Πράγματι, η τομή όλων των στοιχείων ενός ολικά διατεταγμένου $\mathcal{B}$ θα είναι μη κενό και συμπαγές υποσύνολο του X, από την ιδιότητα πεπερασμένων τομών στον συμπαγή X, και θα είναι και αναλλοίωτο σύνολο. Επομένως, κάθε ολικά διατεταγμένο υποσύνολο του $\mathcal{A}$ έχει ελάχιστο στοιχείο και άρα από το Λήμμα του Zorn, το $\mathcal{A}$ έχει ελαχιστικό στοιχείο. Ένα οποιοδήποτε ελαχιστικό στοιχείο του $\mathcal{A}$ κάνει για το ζητούμενο E, γιατί ένα τέτοιο σύνολο θα είναι μη κενό κλειστό και αναλλοίωτο, αφού ανήκει στο $\mathcal{A}$ και δεν θα περιέχει γνήσια μικρότερο και μη κενό T-αναλλοίωτο υποσύνολο.. □

Παρατήρηση. Αν (X,T) τ.δ.σ. και F_1, F_2 είναι ελαχιστικά υποσυστήματα τότε είτε $F_1 \cap F_2 = \varnothing$ είτε $F_1 = F_2$. Πράγματι, αν $x \in F_1 \cap F_2$, τότε $\overline{\{T^n(x) \colon n \in \mathbb{N}_0\}} = F_1 = F_2$, επειδή καθένα από τα F_i, $i \in \{1,2\}$, είναι ελαχιστικό.

Σημείωση. Δεν μπορεί κάθε σύστημα (X,T) να γραφτεί ως ένωση ξένων ανά 2 ελαχιστικών υποσυστημάτων. Για παράδειγμα, αν $X = \{0,1\}^{\mathbb{Z}}$, T ο μετασχηματισμός shift και $x = (\dots, 0, 0, 1, 1, 1, \dots) \in X$, όπου το πρώτο 1 είναι στην θέση 0, τότε
$$\overline{\{T^n(x) \colon n \in \mathbb{Z}\}} = \{(\dots,0,0,0,\dots)\} \cup \{(\dots,1,1,1,\dots)\} \cup \{T^n(x) \colon n \in \mathbb{Z}\}$$
δεν διασπάται σε ένωση ξένων ανά δύο ελαχιστικών συστημάτων. Το παράδειγμα αυτό δείχνει επίσης ότι δεν περιέχεται κάθε σύστημα σε κάποιο ελαχιστικό σύστημα.

Πρόταση 3.2.10 Αν (X,T) ένα ελαχιστικό τ.δ.σ. τότε κάθε συνεχής αναλλοίωτη συνάρτηση είναι σταθερή.

Απόδειξη. Έστω $f \colon X \to \mathbb{C}$ συνεχής και αναλλοίωτη συνάρτηση, δηλαδή $f = f \circ T$. Επειδή $f \circ T^n = f$ για κάθε $n \in \mathbb{N}$, η f θα είναι σταθερή στο πυκνό σύνολο $\{T^n(x) \colon n \in \mathbb{N}\}$. Από την συνέχεια της f έπεται ότι θα πρέπει να είναι σταθερή παντού. □

Παρατήρηση. Το αντίστροφο της πρότασης 3.2.10 δεν ισχύει. Αυτό φαίνεται και από την απόδειξη, αφού αρκεί η ύπαρξη μίας και μόνο πυκνής τροχιάς για την εξαγωγή του συμπεράσματος. Για ένα συγκεκριμένο παράδειγμα, έστω (X,T) το τ.δ.σ. όπου $X = \mathbb{T}$ και $T(x) = 2x \pmod 1$. Το σύστημα αυτό δεν είναι ελαχιστικό γιατί, για παράδειγμα, έχει σταθερό σημείο, το $x = 0$ (βλ. Παρατήρηση μετά τον Ορισμό 3.2.3 πιο πάνω). Από την άλλη το σύστημα $(X, \mathcal{B}, \lambda, T)$, όπου $\mathcal{B} = \mathcal{B}(\mathbb{T})$ η Borel σ-άλγεβρα του $\mathbb{T}$ και λ το μέτρο Lebesgue, είναι εργοδικό, άρα κάθε μετρήσιμη αναλλοίωτη συνάρτηση είναι σταθερή σχεδόν παντού, επομένως κάθε συνεχής αναλλοίωτη συνάρτηση είναι σταθερή (παντού).

Παράδειγμα 3.2.11 Έστω $X = \mathbb{T} = [0,1)$, $a \in [0,1)$, $T_a(x) = x + a \pmod 1$. Αν $a \in \mathbb{R} \setminus \mathbb{Q}$, ξέρουμε ότι η τροχιά του 0 είναι πυκνή (συνέπεια εργοδικότητας, βλ. Θεώρημα 2.1.7 (Kronecker)) και από αυτό έπεται άμεσα ότι και κάθε τροχιά είναι πυκνή. Άρα αν ο a είναι άρρητος, το σύστημα είναι ελαχιστικό. Αν πάλι ο a είναι ρητός, έστω $a = m/n$, $m \in \mathbb{Z}$, $n \in \mathbb{N}$, $(m,n) = 1$. Τότε $T_a^n = \mathrm{id}_X$, όπου $\mathrm{id}_X \colon X \to X$ ο ταυτοτικός μετασχηματισμός, δηλαδή κάθε σημείο είναι περιοδικό και άρα το σύστημα δεν μπορεί να είναι ελαχιστικό (από προηγούμενη παρατήρηση).

Αντιπαράδειγμα 3.2.12 Αν $X = \mathbb{T}^d = [0,1)^d$ και $T(x) = Ax \pmod 1$, όπου A $d \times d$ πίνακας με όλα τα στοιχεία $A_{ij} \in \mathbb{Z}$ και $\det(A) \neq 0$, δηλαδή ένας ενδομορφισμός του d-διάστατου τόρου, τότε το σύστημα έχει σταθερό σημείο, το $x = 0$, και άρα δεν είναι ελαχιστικό. Ειδικότερα, για $d = 1$ όλα τα συστήματα $(\mathbb{T}, T_k)$ όπου $T_k(x) = kx \pmod 1$, $k \in \mathbb{N} \setminus \{1\}$, είναι μη ελαχιστικά.

Αντιπαράδειγμα 3.2.13 Όλοι οι χώροι shift, δηλαδή $X = \{1, \dots, s\}^{\mathbb{N}}$, $T \colon X \to X$ το shift, ή $X = \{1, \dots, s\}^{\mathbb{Z}}$ με το shift πάλι, όπου $s \in \mathbb{N} \setminus \{1\}$, είναι μη ελαχιστικά γιατί έχουν σταθερά σημεία.

3.3 Μεταβατικότητα σε Τοπολογικά Δυναμικά Συστήματα

Ορισμός 3.3.1 (μονόπλευρα μεταβατικό τ.δ.σ.) Ένα (X,T) τ.δ.σ. είναι *μονόπλευρα μεταβατικό*, αν υπάρχει $x \in X$ με πυκνή τροχιά, δηλαδή υπάρχει $x \in X$ τέτοιο ώστε

$$\overline{\{T^n(x) \colon n \in \mathbb{N}_0\}} = X.$$

Ορισμός 3.3.2 (αμφίπλευρα μεταβατικό τ.δ.σ.) Ένα (X,T) αντιστρέψιμο τ.δ.σ. είναι *αμφίπλευρα μεταβατικό*, αν υπάρχει $x \in X$ με πυκνή αμφίπλευρη τροχιά, δηλαδή υπάρχει $x \in X$ τέτοιο ώστε

$$\overline{\{T^n(x) \colon n \in \mathbb{Z}\}} = X.$$

Παρατήρηση. Προφανώς η ελαχιστικότητα συνεπάγεται την μεταβατικότητα. Επίσης, αν το τ.δ.σ. είναι αντιστρέψιμο, η μονόπλευρη μεταβατικότητα συνεπάγεται την αμφίπλευρη μεταβατικότητα. Το αντίστροφο δεν ισχύει, όπως θα δούμε παρακάτω, και άρα οι δύο έννοιες σε ένα αντιστρέψιμο τ.δ.σ. δεν είναι ισοδύναμες.

Πρόταση 3.3.3 Έστω (X,T) αντιστρέψιμο τ.δ.σ. Τα εξής είναι ισοδύναμα.

(1) Το σύστημα είναι αμφίπλευρα μεταβατικό.

(2) Για κάθε $E \subset X$ κλειστό με $T(E) = E$ ισχύει ότι είτε $E = X$ είτε E είναι πουθενά πυκνό.

(3) Για κάθε δύο ανοικτά, μη κενά $U, V \subset X$, υπάρχει $n \in \mathbb{Z}$ τ.ω. $T^n(U) \cap V \neq \varnothing$.

(4) $\left\{x \in X \colon \overline{\{T^n(x) \colon n \in \mathbb{Z}\}} = X\right\}$ είναι G_δ-πυκνό σύνολο.

Σημείωση. Η (2) μπορεί να εκφραστεί ισοδύναμα και ως εξής: για κάθε $U \subset X$ ανοικτό με $T(U) = U$ ισχύει $U = \varnothing$ ή U είναι πυκνό. Πράγματι, αν U ανοικτό, θέτουμε $E = X \setminus U$. Τότε το E είναι κλειστό και $T(E) = E$, άρα $E = X \Leftrightarrow U = \varnothing$ ή E είναι πουθενά πυκνό $\Leftrightarrow \overline{U} = X$.

Απόδειξη. (1) $\Rightarrow$ (2) Έστω $x_0 \in X$ με $\overline{\{T^n(x_0) \colon n \in \mathbb{Z}\}} = X$. Έστω $E \subset X$ τέτοιο ώστε $T^{-1}(E) = E$ και E κλειστό. Εάν υπάρχει ανοικτό, μη κενό σύνολο U τέτοιο ώστε $U \subset E$ τότε υπάρχει $n \in \mathbb{Z}$ με $T^n(x_0) \in U \subset E$ και αφού $T(E) = E$ έχουμε ότι $T^m(x_0) \in E$ για κάθε $m \in \mathbb{Z}$. Άρα

$$\{T^m(x_0) \colon m \in \mathbb{Z}\} \subset E.$$

Το E είναι κλειστό, επομένως

$$\overline{\{T^m(x_0) \colon m \in \mathbb{Z}\}} \subset E$$

και άρα $X = E$.

(2) $\Rightarrow$ (3) Έστω $U, V \subset X$ ανοικτά και μη κενά. Θεωρούμε το σύνολο $\bigcup_{n \in \mathbb{Z}} T^n(U)$ το οποίο είναι ανοικτό και T-αναλλοίωτο, άρα, λόγω υπόθεσης, είναι είτε κενό είτε πυκνό. Όμως $U \neq \varnothing \Rightarrow \bigcup_{n \in \mathbb{Z}} T^n(U) \neq \varnothing$ και άρα το $\bigcup_{n \in \mathbb{Z}} T^n(U)$ πρέπει να είναι πυκνό. Επομένως

$$\bigcup_{n \in \mathbb{Z}} T^n(U) \cap V \neq \varnothing,$$

αφού V ανοικτό και μη κενό και άρα

$$T^n(U) \cap V \neq \varnothing$$

για κάποιο $n \in \mathbb{Z}$.

(3) $\Rightarrow$ (4) Θέλουμε να δείξουμε ότι το σύνολο των πυκνών τροχιών είναι G_δ-πυκνό. Επειδή ο X είναι συμπαγής μετρικός χώρος, έχει αριθμήσιμη βάση[18], και έστω $\{U_n \colon n \in \mathbb{N}\}$ η αριθμήσιμη βάση για την

[18] Σαν συμπαγής μετρικός χώρος, ο X είναι διαχωρίσιμος και έστω $\{x_n \colon n \in \mathbb{N}\}$ αριθμήσιμο πυκνό υποσύνολο του X. Τότε η οικογένεια $U(x_n, 1/m)$, $n, m \in \mathbb{N}$, όπου $U(x, \varepsilon)$ συμβολίζει την ανοικτή μπάλα κέντρου $x \in X$ και ακτίνας $\varepsilon > 0$ στον X, είναι μία αριθμήσιμη βάση για την τοπολογία του.

τοπολογία του, όπου, χωρίς βλάβη της γενικότητας φυσικά, υποθέτουμε ότι $U_n \neq \varnothing \ \forall n \in \mathbb{N}$. Δηλαδή για κάθε $V \subset X$ ανοικτό και $x \in V$, υπάρχει $m \in \mathbb{N}$ τ.ω. $x \in U_m \subset V$. Τότε

$$\left\{x \in X\colon \overline{\{T^n(x)\colon n \in \mathbb{Z}\}} = X\right\} = \bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{Z}} T^m(U_n).$$

Πράγματι, αν $x \in X$ έχει πυκνή τροχιά, τότε για κάθε $n \in \mathbb{N}$, $T^m(x) \in U_n$ για κάποιο $m \in \mathbb{Z}$, αφού $U_n \neq \varnothing$ και ανοικτό, δηλαδή για κάθε $n \in \mathbb{N}$, $x \in T^{-m}(U_n)$ για κάποιο $m \in \mathbb{Z}$, δηλαδή για κάθε $n \in \mathbb{N}$, $x \in \bigcup_{m\in\mathbb{Z}} T^{-m}(U_n)$. Και αντίστροφα, αν $x \in \bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{Z}} T^m(U_n)$, και $U \subset X$ μη κενό και ανοικτό, τότε υπάρχει $n \in \mathbb{N}$ ώστε $U_n \subset U$, και αφού $x \in T^{-m}(U_n)$ για κάποιο $m \in \mathbb{Z}$ για αυτό το n, έπεται ότι $x \in T^{-m}(U_n) \subset T^{-m}(U)$ για αυτό το m, δηλαδή $T^m(x) \in U$, που αποδεικνύει ότι η τροχιά του x τέμνει κάθε μη κενό ανοικτό υποσύνολο του X και άρα είναι πυκνή.

Επειδή τώρα κάθε U_n είναι ανοικτό και μη κενό, το $G_n := \bigcup_{m\in\mathbb{Z}} T^m(U_n)$ θα είναι πυκνό, λόγω της υπόθεσής μας (3) (τέμνει κάθε μη κενό ανοικτό). Αφού κάθε G_n είναι και ανοικτό, έπεται από το Θεώρημα Baire ότι το ζητούμενο σύνολο

$$\left\{x \in X\colon \overline{\{T^n(x)\colon n \in \mathbb{Z}\}} = X\right\} = \bigcap_{n\in\mathbb{N}} G_n,$$

είναι πυκνό και G_δ.

(4) ⇒ (1) Προφανές. □

Ανάλογα ισχύουν και για μονόπλευρη μεταβατικότητα.

Πρόταση 3.3.4 Έστω (X,T) τ.δ.σ. τέτοιο ώστε $T(X) = X$. Τα εξής είναι ισοδύναμα.

(1) (X,T) είναι μονόπλευρα μεταβατικό.

(2) Για κάθε $U, V \subset X$ ανοικτά και μη κενά, υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε $T^n(U) \cap V \neq \varnothing$.

(3) Για κάθε $U, V \subset X$ ανοικτά και μη κενά, υπάρχει $n \in \mathbb{N}$ τέτοιο ώστε $T^{-n}(U) \cap V \neq \varnothing$.

(4) Για κάθε E κλειστό και αναλλοίωτο ($T(E) \subset E$) ισχύει ότι είτε $E = X$ είτε E είναι πουθενά πυκνό.

(5) Το σύνολο $\left\{x \in X\colon \overline{\{T^n(x)\colon n \in \mathbb{N}_0\}} = X\right\}$ είναι G_δ-πυκνό.

Σημείωση. Το (4) ισοδυναμεί με το εξής: για κάθε $V \subset X$ ανοικτό με $T^{-1}(U) \subset U$ θα ισχύει ότι είτε $U = \varnothing$ είτε U είναι πυκνό.

Απόδειξη. (2) ⇔ (3) Ισχύει ότι

$$T^n\big(T^{-n}(U) \cap V\big) = U \cap T^n(V) \qquad (n \in \mathbb{N}_0).$$

Πράγματι, έστω $y \in T^n\big(T^{-n}(U) \cap V\big)$. Τότε υπάρχει $x \in T^{-n}(U) \cap V$ τέτοιο ώστε $T^n(x) = y$. Άρα $y \in T^n(V)$ και $y = T^n(x) \in U$ επομένως,

$$y \in T^n(V) \cap U.$$

Αντίστροφα, αν $y \in U \cap T^n(V)$, υπάρχει $x \in V$ τέτοιο ώστε $T^n(x) = y$. Επειδή $y \in U$, $x \in T^{-n}(U)$, άρα

$$x \in T^{-n}(U) \cap V \quad \Rightarrow \quad y = T^n(x) \in T^n\big(T^{-n}(U) \cap V\big).$$

(3) ⇒ (5) Ομοίως με το (3) ⇒ (4) στην Πρόταση 3.3.3. Συγκεκριμένα, αν $\{U_n\colon n \in \mathbb{N}\}$ είναι μία αριθμήσιμη βάση για την τοπολογία του X, όπου πάλι, χωρίς βλάβη της γενικότητας, υποθέτουμε ότι $U_n \neq \varnothing$ $\forall n \in \mathbb{N}$, τότε

$$\left\{x \in X\colon \overline{\{T^n(x)\colon n \in \mathbb{N}_0\}} = X\right\} = \bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{N}_0} T^{-m}(U_n).$$

Επειδή, πάλι, κάθε U_n είναι ανοικτό και μη κενό, το $G_n := \bigcup_{m\in\mathbb{N}_0} T^{-m}(U_n)$ θα είναι πυκνό, λόγω της υπόθεσής μας (3). Επειδή κάθε G_n είναι και ανοικτό, έπεται από το Θεώρημα Baire ότι το ζητούμενο σύνολο είναι πυκνό και G_δ.

(5) $\Rightarrow$ (1) Προφανές.

(1) $\Rightarrow$ (4) Έστω $x_0 \in X$ τέτοιο ώστε $\overline{\{T^n(x_0)\colon n \in \mathbb{N}_0\}} = X$, E κλειστό με $T(E) \subset E$ και έστω $U \subset E, U \neq \varnothing$, ανοικτό. Υπάρχει $n_0 \in \mathbb{N}_0$ τέτοιο ώστε $T^{n_0}(x_0) \in U \subset E$ και επειδή $T(E) \subset E$ έπεται ότι

$$T^n(x_0) \in E \quad \forall n \geq n_0.$$

Αφού το E είναι κλειστό ισχύει ότι

$$\overline{\{T^n(x_0)\colon n \geq n_0\}} \subset E.$$

Όμως,

$$\overline{\{T^n(x_0)\colon n \in \mathbb{N}_0\}} = X \quad \Rightarrow \quad X = \overline{\{T^n(x_0)\colon n \geq n_0\}} \cup \{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}$$

και άρα

$$X = E \cup \{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}.$$

Άρα,

$$T^{n_0}(X) = T^{n_0}\big(E \cup \{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}\big) \subset E.$$

Τώρα, επειδή T είναι επί, και άρα $T^{n_0}(X) = X$, έπεται ότι $X = E$.

(4) $\Rightarrow$ (3) Έστω U, V ανοικτά και μη κενά. Θεωρούμε το σύνολο

$$\bigcup_{n\in\mathbb{N}} T^{-n}(U).$$

Αυτό είναι ανοικτό, μη κενό (επειδή T επί) και

$$T^{-1}\left(\bigcup_{n\in\mathbb{N}} T^{-n}(U)\right) = \bigcup_{n\geq 2} T^{-n}(U) \subset \bigcup_{n\in\mathbb{N}} T^{-n}(U).$$

Άρα από την υπόθεση το $\bigcup_{n\in\mathbb{N}} T^{-n}(U)$ είναι πυκνό. Επομένως, αν $V \neq \varnothing$ ανοικτό, πρέπει

$$V \cap \bigcup_{n\in\mathbb{N}} T^{-n}(U) \neq \varnothing$$

και τελικά

$$T^{-n}(U) \cap V \neq \varnothing$$

για κάποιο $n \in \mathbb{N}$. $\square$

Σημείωση. Η πρόταση αυτή ισχύει και αν αντί για την υπόθεση ότι $T(X) = X$, υποθέσουμε ότι ο X απλά δεν έχει μεμονωμένα σημεία. Πράγματι, η υπόθεση ότι ο T είναι επί χρησιμοποιήθηκε στην παραπάνω απόδειξη στα (1) $\Rightarrow$ (4) και (4) $\Rightarrow$ (3). Για το (1) $\Rightarrow$ (4) συμπεραίνουμε καταρχήν όπως στην παραπάνω απόδειξη ότι $X = E \cup \{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}$. Τότε όμως $X \setminus E \subset \{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}$, και αφού το $X \setminus E$ είναι ανοικτό πρέπει να περιέχεται και στο εσωτερικό του $\{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}$. Αν όμως το X δεν έχει μεμονωμένα σημεία, το εσωτερικό του $\{x_0, T(x_0), \ldots, T^{n_0-1}(x_0)\}$ είναι το κενό σύνολο και άρα πρέπει $X \setminus E = \varnothing$. Το άλλο σημείο όπου χρησιμοποιήθηκε η υπόθεση ότι ο T είναι επί στην παραπάνω απόδειξη είναι η συνεπαγωγή (4) $\Rightarrow$ (3). Όμως η υπόθεση (4) από μόνη της συνεπάγεται ότι ο T είναι επί (και άρα δεν χρειάζεται να το υποθέσουμε). Πράγματι, έστω $G := X \setminus T(X)$. Τότε $T^{-1}(G) = \varnothing$. Από συμπάγεια, το $T(X)$ είναι συμπαγές και άρα το G ανοικτό. Έστω ότι $G \neq \varnothing$. Η υπόθεση (4) συνεπάγεται τότε ότι το G είναι πυκνό, γιατί είναι μη κενό, ανοικτό και $T^{-1}(G) = \varnothing \subset G$. Έπεται ότι το G δεν μπορεί να είναι μονοσύνολο, γιατί είναι πυκνό και ο χώρος X έχει τουλάχιστον δύο σημεία[19]: $T(X) \neq \varnothing$ και $T(X) \cap G = \varnothing$. Άρα υπάρχουν $x, y \in G$ και έστω $U, V \subset G$ ανοικτές μπάλες ώστε $x \in U$, $y \in V$ και $U \cap V = \varnothing$. Όμως από την υπόθεση (4) πάλι, το U είναι πυκνό, αφού $U \subset G \Rightarrow T^{-1}(U) = \varnothing \subset U$ και U ανοικτό και μη κενό[20]. Άρα θα πρέπει να τέμνει το ανοικτό V, που είναι φυσικά άτοπο.

[19]Υποθέτουμε φυσικά ότι $X \neq \varnothing$.

[20]Όμοια και το V θα πρέπει να είναι πυκνό.

Παράδειγμα 3.3.5 Έστω $X = \{0\} \cup \{1/n\colon n \in \mathbb{N}\}$, με την σχετική τοπολογία ως υποσύνολο του $\mathbb{R}$, και $T\colon X \to X$ να δίνεται από τις

$$T(0) = 0 \qquad \text{και} \qquad T(1/n) = \frac{1}{n+1} \quad \text{για } n \in \mathbb{N}.$$

Εδώ $\overline{\{T^n(1)\colon n \in \mathbb{N}_0\}} = X$ άρα το σύστημα είναι μονόπλευρα μεταβατικό. Όμως καμία άλλη τροχιά δεν είναι πυκνή, άρα τα (1) και (5) της Πρότασης 3.3.4 δεν είναι ισοδύναμα. Επίσης, για τα ανοικτά, μη κενά σύνολα $U = \{1\}, V = \{1\}$ έχουμε ότι $T^{-n}(U) = \varnothing$ για κάθε $n \in \mathbb{N}$ και άρα

$$T^{-n}(U) \cap V = \varnothing \qquad \forall\, n \in \mathbb{N}.$$

Μάλιστα για τα ανοικτά και μη κενά $U = \{1\}, V = \{1/2\}$ έχουμε ότι $T^{-n}(U) \cap V = \varnothing\ \forall\, n \in \mathbb{N}_0$. Τέλος, το $E = X \setminus \{1\}$ είναι κλειστό, $T(E) = X \setminus \{1, 1/2\} \subset E$, όμως $E \neq X$ και το E έχει εσωτερικό το $E \setminus \{0,1\} \neq \varnothing$.

Παράδειγμα 3.3.6 Έστω $X = \{0\} \cup \{1\} \cup \{1/n\colon n \in \mathbb{N}\} \cup \{1 - 1/n\colon n \in \mathbb{N}\}$, με την σχετική τοπολογία ως υποσύνολο του $\mathbb{R}$ πάλι, και $T\colon X \to X$ να δίνεται από τις

$$T(0) = 0, \quad T(1) = 1, \quad T(1/n) = \frac{1}{n+1} \quad \text{για } n \in \mathbb{N} \setminus \{1\}, \quad T(1 - 1/n) = 1 - \frac{1}{n-1} \quad \text{για } n \in \mathbb{N} \setminus \{1, 2\},$$

δηλαδή το 0 και το 1 είναι σταθερά σημεία και κάθε άλλο σημείο του X πάει στο αμέσως επόμενο στα αριστερά του. Εδώ $\overline{\{T^n(x)\colon n \in \mathbb{Z}\}} = X$ για κάθε $x \in X \setminus \{0,1\}$, άρα το σύστημα είναι αμφίπλευρα μεταβατικό, όχι όμως μονόπλευρα αφού δεν υπάρχει μονόπλευρη πυκνή τροχιά.

Πρόταση 3.3.7 Έστω (X,T) μεταβατικό τ.δ.σ. και $f \in C(X)$ τέτοια ώστε $f = f \circ T$. Τότε η f είναι σταθερή.

Απόδειξη. Βλέπε Πρόταση 3.2.10. □

Σημείωση. Η παραπάνω πρόταση ισχύει και για μονόπλευρα και για αμφίπλευρα συστήματα. Το αντίστροφο δεν ισχύει.

Παράδειγμα 3.3.8 Έστω $X = \{0,1\}^{\mathbb{N}}$ και $T\colon X \to X$ ο μετασχηματισμός shift. Η μετρική δίνεται από την $d(\underset{\sim}{x}, \underset{\sim}{y}) = e^{-n}$, όπου $n = n(\underset{\sim}{x}, \underset{\sim}{y}) = \inf\{k \in \mathbb{N}\colon x_k \neq y_k\}$. Έστω $\underset{\sim}{x}$ η ακολουθία που προκύπτει αν παρατεθούν διαδοχικά τα στοιχεία του $\bigcup_{n \geq 1} \{0,1\}^n$, το οποίο είναι αριθμήσιμο σύνολο (και άρα μπορούμε να το αριθμήσουμε). Αυτή η ακολουθία έχει πυκνή τροχιά, άρα το (X,T) είναι μονόπλευρα μεταβατικό.

Παράδειγμα 3.3.9 Έστω $X = \mathbb{T}$, $T_a(x) = x + a \pmod 1$. Έχουμε δει ότι ο X είναι συμπαγής μετρικός χώρος και ξέρουμε ότι υπάρχει πυκνή τροχιά ανν $a \in \mathbb{R} \setminus \mathbb{Q}$, λόγω του θεωρήματος Kronecker.

Παράδειγμα 3.3.10 Έστω $X = \mathbb{T}$, $T(x) = kx \pmod 1$, ή γενικότερα έστω $X = \mathbb{T}^n$ και $T(x) = Ax \pmod 1$, με A έναν $n \times n$ πίνακα με $A_{ij} \in \mathbb{Z}$ για κάθε $i, j \in \{1, \ldots, n\}$ και $\det(A) \neq 0$. Το σύστημα (X,T) είναι μεταβατικό όταν είναι εργοδικό. Αυτό μπορεί να αποδειχθεί σαν συνέπεια της επόμενης Πρότασης.

Πρόταση 3.3.11 Έστω (X,T) τ.δ.σ. και έστω μ Borel μέτρο πιθανότητας στον X που είναι T-αναλλοίωτο και εργοδικό (δηλαδή το $(X, \mathcal{B}(X), \mu, T)$ είναι εργοδικό σ.δ.μ.). Αν $\mu(U) > 0$ για κάθε $U \subset X$ μη κενό, ανοικτό, το σύστημα (X,T) είναι τοπολογικά μονόπλευρα μεταβατικό.

Απόδειξη. Έστω $(U_n)_{n \in \mathbb{N}}$ αριθμήσιμη βάση του X για την τοπολογία του, με $U_n \neq \varnothing\ \forall\, n \in \mathbb{N}$. Τότε

$$\left\{x \in X\colon \overline{\{T^n(x)\colon n \in \mathbb{N}_0\}} = X\right\} = \bigcap_{n \geq 1} \bigcup_{k \geq 0} T^{-k}(U_n).$$

Όμως, για κάθε n,

$$\mu\left(\bigcup_{k \geq 0} T^{-k}(U_n)\right) = 1,$$

λόγω εργοδικότητας και επειδή $\mu(U_n) > 0$ (Πρόταση 2.1.4 (3)), και άρα

$$\mu\left(\bigcap_{n\geq 1}\bigcup_{k\geq 0} T^{-k}(U_n)\right) = 1.$$

Δηλαδή μ-σχεδόν κάθε $x \in X$ έχει πυκνή τροχιά. □

Εφαρμογή Αν μ είναι το μέτρο Lebesgue, από την παραπάνω Πρόταση παίρνουμε ότι τα συστήματα του παραδείγματος 3.3.10 είναι μεταβατικά όταν είναι εργοδικά.

3.4 Χώρος αναλλοίωτων μέτρων σε ένα Τοπολογικό Δυναμικό Σύστημα

Έστω (X, T) τ.δ.σ. δηλαδή X συμπαγής μ.χ. και $T\colon X \to X$ συνεχής. Με $\mathcal{B}(X)$ θα συμβολίζεται η Borel σ- άλγεβρα του X, που είναι η σ-άλγεβρα που παράγεται από τα ανοικτά υποσύνολα του X. Ένα μέτρο Borel είναι ένα μέτρο ορισμένο στον μετρήσιμο χώρο $(X, \mathcal{B}(X))$.

Πρόταση 3.4.1 Κάθε Borel μέτρο πιθανότητας σε έναν μετρικό χώρο είναι κανονικό. Δηλαδή για κάθε σύνολο Borel B και $\varepsilon > 0$, υπάρχουν ανοικτό U και κλειστό C έτσι ώστε $C \subset B \subset U$ και $\mu(U \setminus C) < \varepsilon$.

Απόδειξη. Ορίζουμε την κλάση όλων των Borel υποσυνόλων B που έχουν αυτήν την ιδιότητα, έστω

$$\mathcal{A} = \{B \in \mathcal{B}(X) \colon \forall\, \varepsilon > 0 \ \exists\, C \text{ κλειστό}, U \text{ ανοικτό, με } C \subset B \subset U \text{ και } \mu(U \setminus C) < \varepsilon\}.$$

Τότε η $\mathcal{A}$ είναι σ-άλγεβρα και περιέχει τα κλειστά σύνολα. Πράγματι, προφανώς $\varnothing, X \in \mathcal{A}$, αφού μπορούμε να πάρουμε, για κάθε ε, $U = C = \varnothing$ και $U = C = X$, αντίστοιχα. Επίσης, αν $B \in \mathcal{A}$, τότε και $X \setminus B \in \mathcal{A}$, αφού αν $C \subset B \subset U$ με C κλειστό και U ανοικτό και $\mu(U \setminus C) < \varepsilon$, τότε

$$X \setminus U \subset X \setminus B \subset X \setminus C,$$

με το $X \setminus U$ κλειστό και το $X \setminus C$ ανοικτό και

$$\mu\big((X \setminus C) \setminus (X \setminus U)\big) = \mu(X \setminus C) - \mu(X \setminus U) = 1 - \mu(C) - [1 - \mu(U)] = \mu(U \setminus C) < \varepsilon.$$

Και τέλος αν $B_n \in \mathcal{A}$ για κάθε $n \in \mathbb{N}$ είναι μία αριθμήσιμη οικογένεια στο $\mathcal{A}$, δοθέντος $\varepsilon > 0$ υπάρχουν C_n κλειστά και U_n ανοικτά, $n \in \mathbb{N}$, με $C_n \subset B_n \subset U_n$ και $\mu(U_n \setminus C_n) < \varepsilon 2^{-(n+1)}$ για κάθε $n \in \mathbb{N}$. Επίσης, επειδή $\bigcup_{m=1}^{n} U_m \nearrow \bigcup_{m=1}^{\infty} U_m$ και άρα $\mu\big(\bigcup_{m=1}^{n} U_m\big) \nearrow \mu\big(\bigcup_{m=1}^{\infty} U_m\big)$ καθώς $n \to \infty$, υπάρχει $n \in \mathbb{N}$ ώστε

$$\mu\left(\bigcup_{m=1}^{\infty} U_m \setminus \bigcup_{m=1}^{n} U_m\right) = \mu\left(\bigcup_{m=1}^{\infty} U_m\right) - \mu\left(\bigcup_{m=1}^{n} U_m\right) < \varepsilon/2.$$

Τότε το $C = \bigcup_{m=1}^{n} C_n$ είναι κλειστό, σαν πεπερασμένη ένωση κλειστών, $U = \bigcup_{m=1}^{\infty} U_m$ είναι ανοικτό, $C \subset \bigcup_{m\in\mathbb{N}} B_m \subset U$ και

$$U \setminus C \subset \bigcup_{m=1}^{n} (U_m \setminus C_m) \cup U \setminus \bigcup_{m=1}^{n} U_m,$$

οπότε

$$\mu(U \setminus C) \leq \sum_{m=1}^{n} \mu(U_n \setminus C_m) + \mu\left(U \setminus \bigcup_{m=1}^{n} U_m\right) < \varepsilon.$$

Αυτό αποδεικνύει ότι η $\mathcal{A}$ είναι σ-άλγεβρα.

Αν τώρα $C \subset X$ είναι κλειστό σύνολο, τότε αν θέσουμε $U_n = \{x \in X \colon \operatorname{dist}(x, C) < 1/n\}$, τα U_n είναι ανοικτά σύνολα και $C \subset U_n$ για κάθε $n \in \mathbb{N}$. Εφόσον $U_n \searrow C$ (αφού C κλειστό) και το μ είναι μέτρο πιθανότητας (άρα πεπερασμένο), $\mu(U_n) \searrow \mu(C)$. Επομένως, δοθέντος $\varepsilon > 0$ επιλέγουμε $n \in \mathbb{N}$ με $\mu(U_n \setminus C) < \varepsilon$ και προσεγγίσαμε το C από μέσα με το κλειστό C και από έξω με το ανοικτό U_n. Άρα κάθε κλειστό ανήκει στην $\mathcal{A}$. Επειδή η $\mathcal{B}(X)$ είναι η μικρότερη σ-άλγεβρα που περιέχει όλα τα κλειστά, πρέπει $\mathcal{A} = \mathcal{B}(X)$. □

Πόρισμα 3.4.2 Έστω X μετρικός χώρος και μ μέτρο Borel πιθανότητας στον X. Τότε

$$\mu(B) = \inf\{\mu(U) \colon U \supset B,\ U \text{ ανοικτό}\} = \sup\{\mu(C) \colon C \subset B,\ C \text{ κλειστό}\}.$$

Συμβολισμός: Αν X συμπαγής μετρικός χώρος, $C(X) = \{f \colon X \to \mathbb{C} \mid f \text{ συνεχής}\}$ και $C_{\mathbb{R}}(X) = \{f \colon X \to \mathbb{R} \mid f \text{ συνεχής}\}$. Η νόρμα και στους δύο χώρους είναι η sup-νόρμα: $\|f\| := \sup_{x\in X}|f(x)|$.

Πόρισμα 3.4.3 Έστω X συμπαγής μετρικός χώρος. Τότε δύο Borel μέτρα πιθανότητας μ, ν επί του X ταυτίζονται ανν

$$\int_X f\,d\mu = \int_X f\,d\nu \qquad \forall f \in C(X).$$

Απόδειξη. Η μία κατεύθυνση είναι προφανής. Για την άλλη, έστω ότι έχουμε $\int_X f\,d\mu = \int_X f\,d\nu\ \ \forall f \in C(X)$ και θέλουμε να δείξουμε ότι $\mu(B) = \nu(B)$ για κάθε $B \in \mathcal{B}(X)$. Από το Πόρισμα 3.4.2, αρκεί να δειχτεί ότι $\mu(C) = \nu(C)$ για κάθε C κλειστό. Έστω, λοιπόν, $C \subset X$ κλειστό και $\varepsilon > 0$. Υπάρχει $U \subset X$ ανοικτό τέτοιο ώστε $\mu(U \setminus C) < \varepsilon$ και $C \subset U$. Υπάρχει f συνεχής με $0 \le f \le 1$ και $f(x) = 1$ αν $x \in C$ και $f(x) = 0$ αν $x \in U^c$. Για παράδειγμα μπορούμε να πάρουμε

$$f(x) = \frac{\text{dist}(x, U^c)}{\text{dist}(x, C) + \text{dist}(x, U^c)}.$$

Τότε $\mathbb{1}_C \le f \le \mathbb{1}_U$ και επομένως

$$\nu(C) \le \int_X f\,d\nu = \int_X f\,d\mu \le \mu(U) \le \mu(C) + \varepsilon.$$

Αφού αυτό ισχύει για αυθαίρετο $\varepsilon > 0$, έπεται ότι $\nu(C) \le \mu(C)$. Λόγω συμμετρίας έχουμε και $\mu(C) \le \nu(C)$. Άρα $\mu(C) = \nu(C)$ για κάθε C κλειστό. $\square$

Άρα σε έναν συμπαγή μετρικό χώρο, ένα Borel μέτρο πιθανότητας καθορίζεται μονοσήμαντα από το πως ολοκληρώνει συνεχείς συναρτήσεις. Κάθε Borel μέτρο πιθανότητας σε έναν συμπαγή μετρικό χώρο ορίζει ένα γραμμικό συναρτησοειδές $L_\mu \colon C(X) \to \mathbb{C}$ στον $C(X)$, ως εξής:

$$L_\mu(f) = \int_X f\,d\mu,$$

που είναι φραγμένο από τον αριθμό 1 αφού $L_\mu(f) \le \|f\| = \sup_{x\in X}|f(x)|\ \ \forall f \in C(X)$. Επομένως ταυτίζουμε κάθε Borel μέτρο πιθανότητας στον συμπαγή μ.χ. X με ένα στοιχείο του δυϊκού $C(X)^*$ του $C(X)$.

Θεώρημα 3.4.4 (θεώρημα αναπαράστασης του Riesz) Αν X συμπαγής μ.χ., τότε κάθε φραγμένο, θετικό, γραμμικό συναρτησοειδές $\Lambda \in C(X)^*$ δίνεται από

$$\Lambda(f) = \int_X f\,d\mu \qquad f \in C(X),$$

για κάποιο (θετικό) πεπερασμένο μέτρο Borel στον X. Θετικό συναρτησοειδές εδώ σημαίνει ότι $L(f) \ge 0$ για κάθε $f \in C(X)$ με $f(x) \ge 0$ για κάθε $x \in X$.

Ορίζουμε

$$M(X) = \{\mu \colon \mu \text{ Borel μέτρο πιθανότητας στον } X\}.$$

Τότε το $M(X)$ ταυτίζεται με ένα υποσύνολο του $C(X)^*$. Το γεγονός αυτό μας επιτρέπει να θεωρήσουμε την ασθενή*-τοπολογία επί του $M(X)$. Η ασθενής*-τοπολογία επί του $M(X)$ είναι, εξ' ορισμού, η μικρότερη τοπολογία που κάνει όλες τις συναρτήσεις $\mu \mapsto \int_X f\,d\mu$ $(\mu \in M(X))$, $f \in C(X)$, συνεχείς. Στην ασθενή*-τοπολογία, ένα δίκτυο $\mu_i \in M(X)$, $i \in I$, συγκλίνει σε ένα $\mu \in M(X)$ ανν

$$\int_X f\,d\mu_i \to \int f\,d\mu \qquad \forall f \in C(X).$$

Μία βάση της τοπολογίας αποτελείται από τα σύνολα

$$\begin{aligned}V_\mu(f_1,\dots,f_n;\varepsilon) &= \left\{\nu\in M(X)\colon\ \left|\int_X f_i\,d\mu-\int_X f_i\,d\nu\right|<\varepsilon\ \ \forall\, i\in\{1,2,\dots,n\}\right\}\\ &= \bigcap_{i=1}^n\left\{\nu\in M(X)\colon\ \left|\int_X f_i\,d\mu-\int_X f_i\,d\nu\right|<\varepsilon\right\},\end{aligned}$$

με $\mu\in M(X)$, $n\in\mathbb{N}$, $f_1,\dots f_n\in C(X)$ και $\varepsilon>0$.

Πρόταση 3.4.5 Αν X συμπαγής μετρικός χώρος και $M(X)$ ο χώρος όλων των Borel μέτρων πιθανότητας στον X με την ασθενή*-τοπολογία, τότε ο $M(X)$ είναι μετρικοποιήσιμος. Μία μετρική συμβατή με την ασθενή*-τοπολογία είναι η

$$\operatorname{dist}(\mu,\nu)=\sum_{n\geq1}\frac{\left|\int_X f_n\,d\mu-\int_X f_n\,d\nu\right|}{2^n\|f_n\|}\qquad\qquad \mu,\nu\in M(X),$$

όπου $\{f_n\colon n\in\mathbb{N}\}$ ένα αριθμήσιμο πυκνό υποσύνολο του $C(X)$.

Υπενθυμίζεται το εξής βασικό αποτέλεσμα.

Θεώρημα 3.4.6 Έστω X συμπαγής χώρος Hausdorff. Τα εξής είναι ισοδύναμα

(1) Ο X είναι μετρικοποιήσιμος.

(2) Ο X είναι 2ος αριθμήσιμος.

(3) Ο $C(X)$ είναι διαχωρίσιμος.

Απόδειξη Πρότασης 3.4.5 Το $\operatorname{dist}(\mu,\nu)$ είναι μετρική (προφανές). Οι συναρτήσεις $\mu\mapsto\int f_i\,d\mu$, $i\in\mathbb{N}$, είναι συνεχείς ως προς την μετρική, διότι

$$\left|\int_X f_i\,d\mu-\int_X f_i\,d\nu\right|\leq\|f_i\|2^i\operatorname{dist}(\mu,\nu),$$

για κάθε i, από τον ορισμό της μετρικής. Από την πυκνότητα των f_i, $i\in\mathbb{N}$, στο $C(X)$, έπεται ότι η $\mu\mapsto\int f\,d\mu$ είναι συνεχής για κάθε $f\in C(X)$: για $f\in C(X)$, δοθέντος $\varepsilon>0$, επιλέγουμε $i\in\mathbb{N}$ έτσι ώστε $\|f-f_i\|\leq\varepsilon/4$, και τότε, αν

$$\operatorname{dist}(\mu,\nu)<\frac{\varepsilon}{2^{i+1}\|f_i\|},$$

έπεται ότι

$$\begin{aligned}\left|\int_X f\,d\mu-\int_X f\,d\nu\right| &\leq \left|\int_X f\,d\mu-\int_X f_i\,d\mu\right|+\left|\int_X f\,d\nu-\int_X f_i\,d\nu\right|+\left|\int_X f_i\,d\mu-\int_X f_i\,d\nu\right|\\ &= \left|\int_X (f-f_i)\,d\mu\right|+\left|\int_X (f-f_i)\,d\nu\right|+\left|\int_X f_i\,d\mu-\int_X f_i\,d\nu\right|\\ &\leq 2\|f-f_i\|+2^i\|f_i\|\operatorname{dist}(\mu,\nu)\\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

Η ασθενής*-τοπολογία είναι η μικρότερη τοπολογία που κάνει όλες τις $\mu\mapsto\int f\,d\mu$ συνεχείς για κάθε $f\in C(X)$. Έπεται ότι κάθε ασθενώς*-ανοικτό σύνολο είναι και ανοικτό ως προς την μετρική.

Για τον αντίστροφο εγκλεισμό, αρκεί να δειχτεί ότι κάθε μπάλα $U(\mu,\varepsilon):=\{\nu\in M(X)\colon \operatorname{dist}(\mu,\nu)<\varepsilon\}$, για την μετρική, είναι ασθενώς*-ανοικτό. Για αυτό αρκεί να δειχτεί ότι η μπάλα ως προς την μετρική περιέχει κάποιο

$$V_\mu(g_1,\dots,g_k;\delta)=\left\{\nu\in M(X)\colon\ \left|\int_X g_i\,d\mu-\int_X g_i\,d\nu\right|<\delta\quad\forall i\in\{1,\dots,k\}\right\}$$

για κάποιες $g_1, \ldots, g_k \in C(X)$ και $\delta > 0$, που είναι τα βασικά ανοικτά σύνολα της ασθενούς*-τοπολογίας. Δοθέντος $\varepsilon > 0$ και $\mu \in M(X)$, επιλέγουμε $N \in \mathbb{N}$ τέτοιο ώστε $\sum_{n>N} 2^{-n} < \varepsilon/4$. Θέτουμε

$$\delta = \frac{\varepsilon}{2} \left(\sum_{i=1}^{N} 2^{-i} \|f_i\|^{-1} \right)^{-1}.$$

Αν $\nu \in V_\mu(f_1, \ldots, f_N; \delta)$, τότε

$$\begin{aligned}
\text{dist}(\mu, \nu) &= \sum_{n=1}^{N} \frac{\left|\int_X f_n \, d\mu - \int_X f_n \, d\nu\right|}{2^n \|f_n\|} + \sum_{n>N} \frac{\left|\int f_n \, d\mu - \int_X f_n \, d\nu\right|}{2^n \|f_n\|} \\
&\leq \sum_{n=1}^{N} \frac{\delta}{2^n \|f_n\|} + \sum_{n>N} \frac{2}{2^n} \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Έπεται ότι

$$V_\mu(f_1, \ldots, f_N; \delta) \subset \{\nu \in M(X) \colon \text{dist}(\mu, \nu) < \varepsilon\} = U(\mu, \varepsilon).$$

Έπεται τώρα ότι κάθε ανοικτό ως προς την μετρική υποσύνολο του $M(X)$ είναι ανοικτό και ως προς την ασθενή*-τοπολογία. Άρα τελικά οι δύο τοπολογίες ταυτίζονται και ο $M(X)$ με την ασθενή*-τοπολογία είναι μετρικός χώρος. □

Αφού $M(X)$ με την ασθενή*-τοπολογία είναι μετρικός χώρος, αρκεί να θεωρούμε μόνο σύγκλιση ακολουθιών, αντί για δίκτυα που χρειάζονται σε πιο γενικούς τοπολογικούς χώρους. Υπενθυμίζεται ότι μια ακολουθία $(\mu_n)_{n\in\mathbb{N}} \in M(X)$ συγκλίνει ασθενώς* στο $\mu \in M(X)$ ανν

$$\int f \, d\mu_n \to \int f \, d\mu \qquad \forall \, f \in C(X).$$

Πρόταση 3.4.7 Αν X είναι συμπαγής μετρικός χώρος τότε ο $M(X)$ με την ασθενή*-τοπολογία είναι συμπαγής μετρικός χώρος.

Θεώρημα 3.4.8 (Alaoglou) Η κλειστή, μοναδιαία μπάλα ως προς την νόρμα του $C(X)^*$ είναι συμπαγές σύνολο ως προς την ασθενή*-τοπολογία.

Απόδειξη πρότασης 3.4.7 Γνωρίζουμε ότι $M(X)$ ταυτίζεται με υποσύνολο του $C(X)^*$. Αν $\mu \in M(X)$ και $f \in C(X)$, τότε

$$\left| \int f \, d\mu \right| \leq \|f\|.$$

Άρα για τη νόρμα του μ, σαν γραμμικό συναρτησοειδές του $C(X)$, δηλαδή ως στοιχείου του $C(X)^*$, ισχύει ότι $\|\mu\| \leq 1$. Επίσης, για $f = \mathbb{1}_X$ βλέπουμε ότι $\|\mu\| = 1$. Επομένως $M(X)$ περιέχεται στην μοναδιαία μπάλα του $C(X)^*$. Αν δειχτεί ότι το σύνολο $M(X)$ είναι κλειστό, εφαρμόζοντας το θεώρημα Alaoglou θα έχουμε ότι είναι και συμπαγές ως προς την ασθενή*-τοπολογία. Για την κλειστότητα, έστω $L_{\mu_n} \to L \in C(X)^*$ ως προς την ασθενή*-τοπολογία, όπου $\mu_n \in M(X)$ για κάθε $n \in \mathbb{N}$ και L_{μ_n} το γραμμικό συναρτησοειδές στον $C(X)^*$ που επάγει το μέτρο μ_n, δηλαδή το $L_{\mu_n}(f) = \int f \, d\mu_n$, $f \in C(X)$. Τότε το L ορίζει μέτρο μ (θεώρημα αναπαράστασης Riesz) και αυτό το μέτρο είναι θετικό. Πράγματι, αν $f \in C(X)$, $f \geq 0$, έχουμε

$$L(f) = \lim_{n\to\infty} L_{\mu_n}(f) = \lim_{n\to\infty} \int f \, d\mu_n \geq 0.$$

Επίσης

$$L(\mathbb{1}_X) = \lim_{n\to\infty} L_{\mu_n}(\mathbb{1}_X) = \lim_{n\to\infty} \mu_n(X) = 1.$$

Έπεται ότι το μέτρο μ είναι μέτρο πιθανότητας. Τελικά το $M(X)$ ταυτίζεται μ' ένα κλειστό, και άρα συμπαγές, υποσύνολο της μοναδιαίας μπάλας του $C(X)^*$. □

2η απόδειξη (χωρίς το Θεώρημα Alaoglou, διαγώνιο επιχείρημα Cantor). Ο $C(X)$ είναι διαχωρίσιμος και έστω $\{f_n : n \in \mathbb{N}\}$ αριθμήσιμο πυκνό υποσύνολο στον $C(X)$. Έστω $(\mu_n)_{n\in\mathbb{N}}$ ακολουθία στον $M(X)$. Η ακολουθία
$$\left(\int f_1\,d\mu_n\right)_{n\in\mathbb{N}}$$
αποτελεί φραγμένη ακολουθία μιγαδικών αριθμών, αφού $f_1 \in C(X)$ και άρα φραγμένη. Άρα έχει συγκλίνουσα υπακολουθία. Δηλαδή υπάρχει γνησίως αύξουσα συνάρτηση $k_1 : \mathbb{N} \to \mathbb{N}$ τέτοια ώστε το όριο $\lim_{n\to\infty} \int f_1\,d\mu_{k_1(n)}$ να υπάρχει και έστω $\mu_n^{(1)} := \mu_{k_1(n)}$ για $n \in \mathbb{N}$. Τότε
$$\int f_1\,d\mu_n^{(1)} \to L(f_1)$$
για κάποιο $L(f_1) \in \mathbb{C}$. Η υπακολουθία τώρα
$$\left(\int f_2\,d\mu_n^{(1)}\right)_{n\in\mathbb{N}}$$
της $\left(\int f_2\,d\mu_n\right)_{n\in\mathbb{N}}$ είναι επίσης φραγμένη ακολουθία και άρα υπάρχει συγκλίνουσα υπακολουθία της, δηλαδή υπάρχει γνησίως αύξουσα συνάρτηση $k_2 : \mathbb{N} \to \mathbb{N}$ τέτοια ώστε το όριο $\lim_{n\to\infty} \int f_2\,d\mu^{(1)}_{k_2(n)}$ να συγκλίνει. Έστω $\mu_n^{(2)} := \mu^{(1)}_{k_2(n)} = \mu_{k_1\circ k_2(n)}$, για κάθε $n \in \mathbb{N}$. Τότε
$$\int f_2\,d\mu_n^{(2)} \to L(f_2)$$
για κάποιο $L(f_2) \in \mathbb{C}$ και επειδή η $(\mu_n^{(2)})_{n\in\mathbb{N}}$ είναι υπακολουθία της $(\mu_n^{(1)})_{n\in\mathbb{N}}$, έπεται ότι επίσης
$$\int f_1\,d\mu_n^{(2)} \to L(f_1).$$
Συνεχίζοντας, βρίσκουμε για κάθε $m \in \mathbb{N}$, γνησίως αύξουσα συνάρτηση $k_m : \mathbb{N} \to \mathbb{N}$ τέτοια ώστε
$$\int f_m\,d\mu^{(m-1)}_{k_m(n)} \xrightarrow{n\to\infty} L(f_m)$$
για κάποιο $L(f_m) \in \mathbb{C}$ και θέτουμε $\mu_n^{(m)} := \mu^{(m-1)}_{k_m(n)} = \mu_{k_1\circ\cdots\circ k_m(n)}$, για κάθε $n \in \mathbb{N}$, όπου $\left(\mu_n^{(0)}\right)_{n\in\mathbb{N}}$ είναι η αρχική ακολουθία $(\mu_n)_{n\in\mathbb{N}}$. Τότε η $\left(\mu_n^{(m)}\right)_{n\in\mathbb{N}}$ είναι υπακολουθία της $\left(\mu_n^{(m-1)}\right)_{n\in\mathbb{N}}$ και
$$\int f_m\,d\mu_n^{(m)} \xrightarrow{n\to\infty} L(f_m).$$
Ορίζουμε $k(n) := k_1 \circ \cdots \circ k_n(n)$ για κάθε $n \in \mathbb{N}$. Τότε
$$k(n+1) = k_1 \circ \cdots \circ k_n \circ k_{n+1}(n+1) \geq k_1 \circ \cdots \circ k_n(n+1) > k_1 \circ \cdots \circ k_n(n) = k(n),$$
η πρώτη ανισότητα επειδή $h(m) \geq m$ για κάθε $m \in \mathbb{N}$ για κάθε γνησίως αύξουσα συνάρτηση $h : \mathbb{N} \to \mathbb{N}$ και η δεύτερη, γνήσια ανισότητα, επειδή όλες οι $k_i : \mathbb{N} \to \mathbb{N}$, $i \in \{1, 2, \ldots, n\}$, είναι γνησίως αύξουσες. Άρα $k(1) < k(2) < \cdots$ και η $\mu_{k(n)} = \mu_{k_1\circ\cdots\circ k_n(n)} = \mu_n^{(n)}$, $n \in \mathbb{N}$, είναι υπακολουθία της $(\mu_n)_{n\in\mathbb{N}}$. Επιπλέον
$$\int f_j\,d\mu_{k(n)} \xrightarrow{n\to\infty} L(f_j) \qquad \forall\, j \in \mathbb{N}. \tag{24}$$
Πράγματι, έστω $j \in \mathbb{N}$. Για τυχόν $\varepsilon > 0$ θεωρούμε $n_\varepsilon \in \mathbb{N}$ τέτοιο ώστε $\left|\int f_j\,d\mu_n^{(j)} - L(f_j)\right| < \varepsilon$ για κάθε $n \geq n_\varepsilon$. Τότε, για κάθε $n > j$, $\mu_{k(n)} = \mu_{k_1\circ\cdots\circ k_n(n)} = \mu^{(j)}_{k_{j+1}\circ\cdots\circ k_n(n)}$ και $\mu_{k(n)} = \mu_j^{(j)}$ για $n = j$, και άρα αν $n \geq \max\{n_\varepsilon, j\}$, τότε
$$\left|\int f_j\,d\mu_{k(n)} - L(f_j)\right| = \left|\int f_j\,d\mu^{(j)}_{k_{j+1}\circ\cdots\circ k_n(n)} - L(f_j)\right| < \varepsilon,$$
με την σύμβαση ότι το $k_{j+1} \circ \cdots \circ k_n(n)$ είναι n όταν $n = j$, αφού τότε $k_{j+1} \circ \cdots \circ k_n(n) \geq n \geq n_\varepsilon$, επειδή η $k_{j+1} \circ \cdots \circ k_n : \mathbb{N} \to \mathbb{N}$ είναι γνήσιως αύξουσα, όταν $j > n$.

Από την (24) και λόγω της πυκνότητας των $\{f_n : n \in \mathbb{N}\}$, έπεται ότι η ακολουθία

$$\left(\int f \, d\mu_{k(n)}\right)_{n\in\mathbb{N}}$$

συγκλίνει για κάθε $f \in C(X)$. Δηλαδή, για κάθε $f \in C(X)$,

$$\int f \, d\mu_{k(n)} \to L(f)$$

για κάποιο $L(f) \in \mathbb{C}$. Πράγματι, αν $f \in C(X)$ και το $\varepsilon > 0$ έχει δοθεί, έστω $j \in \mathbb{N}$ τέτοιο ώστε

$$\|f - f_j\| = \sup_{x\in X} |f(x) - f_j(x)| < \tfrac{1}{4}\varepsilon,$$

και για αυτό το j, έστω $n_\varepsilon \in \mathbb{N}$ τέτοιο ώστε

$$\left| \int f_j \, d\mu_{k(n)} - \int f_j \, d\mu_{k(m)} \right| < \tfrac{1}{2}\varepsilon$$

για όλα τα ζεύγη $m, n \in \mathbb{N}$ με $m, n \geq n_\varepsilon$. Τέτοιο n_ε υπάρχει επειδή η ακολουθία $\left(\int f_j \, d\mu_{k(n)}\right)_{n\in\mathbb{N}}$ συγκλίνει και άρα είναι βασική. Τότε

$$\begin{aligned}
&\left| \int f \, d\mu_{k(n)} - \int f \, d\mu_{k(m)} \right| \\
&\qquad \leq \left| \int f \, d\mu_{k(n)} - \int f_j \, d\mu_{k(n)} \right| + \left| \int f_j \, d\mu_{k(n)} - \int f_j \, d\mu_{k(m)} \right| + \left| \int f_j \, d\mu_{k(m)} - \int f \, d\mu_{k(m)} \right| \\
&\qquad \leq \int |f - f_j| \, d\mu_{k(n)} + \left| \int f_j \, d\mu_{k(n)} - \int f_j \, d\mu_{k(m)} \right| + \int |f_j - f| \, d\mu_{k(m)} \\
&\qquad \leq \tfrac{1}{4}\varepsilon + \tfrac{1}{2}\varepsilon + \tfrac{1}{4}\varepsilon = \varepsilon
\end{aligned}$$

για $m, n \geq n_\varepsilon$ και αφού το ε ήταν αυθαίρετο, αυτό δείχνει ότι η ακολουθία $\left(\int f \, d\mu_{k(n)}\right)_{n\in\mathbb{N}}$ είναι βασική, και άρα συγκλίνει.

Το L είναι γραμμικός τελεστής, αφού για $f, g \in C(X)$ και $a, b \in \mathbb{C}$ έχουμε ότι

$$L(af + bg) = \lim_{n\to\infty} \int (af + bg) \, d\mu_{k(n)} = \lim_{n\to\infty} \left(a \int f \, d\mu_{k(n)} + b \int g \, d\mu_{k(n)} \right) = aL(f) + bL(g),$$

και ισχύει ότι $f \geq 0 \Rightarrow L(f) \geq 0$. Το L είναι επίσης φραγμένο, αφού

$$|L(f)| = \left| \lim_{n\to\infty} \int f \, d\mu_{k(n)} \right| = \lim_{n\to\infty} \left| \int f \, d\mu_{k(n)} \right|$$

και $\left|\int f \, d\mu_{k(n)}\right| \leq \int |f| \, d\mu_{k(n)} \leq \|f\|$ για κάθε $f \in C(X)$. Άρα το L είναι φραγμένο θετικό γραμμικό συναρτησοειδές, και άρα από το Θεώρημα του Riesz υπάρχει θετικό Borel μέτρο μ τέτοιο ώστε

$$L(f) = \int f \, d\mu \qquad \forall\, f \in C(X).$$

Επιπλέον $\mu(X) = \int \mathbb{1}_X \, d\mu = L(\mathbb{1}_X) = \lim_{n\to\infty} \int \mathbb{1}_X \, d\mu_{k(n)} = 1$ και άρα το μ είναι μέτρο πιθανότητας. Τότε

$$\int f \, d\mu_{k(n)} \to \int f \, d\mu \qquad \forall\, f \in C(X),$$

δηλαδή $\mu_{k(n)} \to \mu$ στην ασθενή*-τοπολογία. Εφόσον η $(\mu_n)_{n\in\mathbb{N}}$ έχει συγκλίνουσα υπακολουθία, την $\left(\mu_{k(n)}\right)_{n\in\mathbb{N}}$, ο $M(X)$ είναι συμπαγής. □

Πρόταση 3.4.9 Έστω (X, T) τ.δ.σ. Ορίζουμε $T_* : M(X) \to M(X)$ με $T_*\mu(A) = \mu\big(T^{-1}(A)\big)$ για κάθε $A \in \mathcal{B}(X)$. Η T_* είναι συνεχής.

Απόδειξη. Έστω $(\mu_n)_{n\in\mathbb{N}}$ ακολουθία στον $M(X)$ τέτοια ώστε $\mu_n \xrightarrow{w^*} \mu$ (ασθενής*-σύγκλιση). Τότε, αφού για κάθε $f \in C(X)$ η $f \circ T$ είναι συνεχής συνάρτηση έχουμε,

$$\int f \, dT_*\mu_n = \int f \circ T \, d\mu_n \xrightarrow{n\to\infty} \int f \circ T \, d\mu = \int f \, dT_*\mu \qquad \forall\, f \in C(X).$$

Επομένως,

$$T_*\mu_n \to T_*\mu. \qquad \square$$

Σημείωση. Χρησιμοποιήσαμε ότι $\int f \, d\, T_*\mu = \int f \circ T \, d\mu$. Αυτό ισχύει για $f \in L^1(\mu)$, βλ. Παρατήρηση μετά το Λήμμα 1.1.7, και άρα για κάθε συνεχή f.

Θεωρούμε τον χώρο των T- αναλλοίωτων μέτρων πιθανότητας στον X,

$$M_T(X) = \{\mu \in M(X) \colon T_*\mu = \mu\}$$

και επίσης τον

$$E_T(X) = \{\mu \in M_T(X) \colon \mu \text{ είναι εργοδικό}\}.$$

Θεώρημα 3.4.10 Έστω (X, T) τ.δ.σ. Ισχύουν τα εξής.

(1) Το $M_T(X)$ είναι μη-κενό, συμπαγές και κυρτό υποσύνολο του $M(X)$.

(2) $\mathrm{Ext}\big(M_T(X)\big) = E_T(X)$, όπου $\mathrm{Ext}\big(M_T(X)\big)$ είναι τα ακραία σημεία[21] του $M_T(X)$.

(3) Αν $\mu, \nu \in E_T(X)$ τότε είτε $\mu = \nu$ είτε $\mu \perp \nu$.

Παρατήρηση. $\mu \in M_T(X) \Leftrightarrow T_*\mu = \mu \Leftrightarrow \int f \, d\mu = \int f \circ T \, d\mu \; \forall\, f \in C(X)$.

Λήμμα 3.4.11 Αν $\mu \in E_T(X)$, $\nu \in M_T(X)$ και $\nu \ll \mu$, τότε $\mu = \nu$.

Απόδειξη Θεωρήματος. (1) Για την κυρτότητα, έστω $\mu_1, \mu_2 \in M_T(X)$ και $\lambda \in [0,1]$. Αν $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$, τότε προφανώς $\mu \in M(X)$ και

$$\int f \, d\mu = \lambda \int f \, d\mu_1 + (1-\lambda) \int f \, d\mu_2 = \lambda \int f \circ T \, d\mu_1 + (1-\lambda) \int f \circ T \, d\mu_2 = \int f \circ T \, d\mu$$

άρα $\mu \in M_T(X)$.

Για την συμπάγεια αρκεί να δειχτεί ότι το $M_T(X)$ είναι κλειστό, αφού $M(X)$ είναι συμπαγής. Έστω $(\mu_n)_{n\in\mathbb{N}}$ ακολουθία στον $M_T(X)$ με $\mu_n \xrightarrow{w^*} \mu$. Έχουμε για $f \in C(X)$ ότι $f \circ T \in C(X)$ και άρα

$$\int f \circ T \, d\mu = \lim_{n\to\infty} \int f \circ T \, d\mu_n = \lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu \qquad \forall\, f \in C(X).$$

Άρα $\mu \in M_T(X)$.

Θα δείξουμε ότι $M_T(X) \neq \varnothing$ (Θεώρημα Krylov–Bogolyubov). Έστω $x \in X$ και θεωρούμε τα μέτρα

$$\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k(x)} \qquad n \in \mathbb{N}$$

όπου για $y \in X$,

$$\delta_y(A) = \begin{cases} 1, & y \in A \\ 0, & y \notin A. \end{cases}$$

[21] Αν C κυρτό σύνολο, σε έναν γραμμικό χώρο, ένα σημείο x λέγεται ακραίο σημείο του C αν δεν υπάρχουν $y, z \in C$ με $y \neq z$ και $\lambda \in (0,1)$ τέτοια ώστε $x = \lambda y + (1-\lambda)z$.

Ισχύει ότι $\mu_n \in M(X)$ για κάθε $n \in \mathbb{N}$. Λόγω του ότι ο $M(X)$ είναι συμπαγής μετρικός χώρος, υπάρχει υπακολουθία $(\mu_{n_k})_{k\in\mathbb{N}}$ της $(\mu_n)_{n\in\mathbb{N}}$ τέτοια ώστε $\mu_{n_k} \xrightarrow{w^*} \mu$ για κάποιο $\mu \in M(X)$. Υπολογίζουμε

$$\begin{aligned}
\int f \circ T\, d\mu &= \lim_{k\to\infty} \int f \circ T\, d\mu_{n_k} = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} \int f \circ T\, d\delta_{T^j(x)} \\
&= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f \circ T(T^j(x)) = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=1}^{n_k} f(T^j(x)) \\
&= \lim_{k\to\infty} \frac{1}{n_k} \left\{ \sum_{j=0}^{n_k-1} f(T^j(x)) - f(x) + f(T^{n_k}(x)) \right\} \\
&= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} \int f\, d\delta_{T^j(x)} = \int f\, d\mu,
\end{aligned}$$

όπου η προτελευταία ισότητα ισχύει επειδή η f είναι φραγμένη. Επομένως $M_T(X) \neq \varnothing$, αφού $\mu \in M_T(X)$.

(2) Έστω $\mu \in M_T(X)$ που δεν είναι εργοδικό. Τότε υπάρχει αναλλοίωτο σύνολο A, δηλαδή $T^{-1}(A) = A$, τέτοιο ώστε $\mu(A) \in (0,1)$. Θεωρούμε τα μέτρα

$$\mu_1(B) = \frac{\mu(A \cap B)}{\mu(A)} \qquad \text{και} \qquad \mu_2(B) = \frac{\mu(A^c \cap B)}{\mu(A^c)}.$$

Τότε

$$\mu = \mu(A)\, \mu_1 + (1 - \mu(A))\, \mu_2.$$

Έχουμε

$$\mu_1(T^{-1}(B)) = \frac{\mu(T^{-1}(B) \cap A)}{\mu(A)} = \frac{\mu(T^{-1}(B) \cap T^{-1}(A))}{\mu(A)} = \frac{\mu(T^{-1}(B \cap A))}{\mu(A)} = \frac{\mu(B \cap A)}{\mu(A)} = \mu_1(B)$$

Άρα το μ_1 είναι T- αναλλοίωτο. Ομοίως και το μ_2. Επομένως, το μ γράφεται σαν μη-τετριμμένος κυρτός συνδυασμός στοιχείων του $M_T(X)$ και άρα $\mu \notin \mathrm{Ext}\big(M_T(X)\big)$. Τελικά,

$$\mathrm{Ext}\big(M_T(X)\big) \subset E_T(X).$$

Αντίστροφα, έστω $\mu \in E_T(X)$ και έστω

$$\mu = \lambda \mu_1 + (1-\lambda)\mu_2 \qquad \text{με} \quad \mu_1, \mu_2 \in M_T(X).$$

Τότε $\mu_1 \ll \mu$ και $\mu_2 \ll \mu$. Εφαρμόζοντας το Λήμμα 3.4.11 παίρνουμε ότι $\mu = \mu_1 = \mu_2$ και αφού ο κυρτός συνδυασμός δεν μπορεί να είναι γνήσιος, παίρνουμε ότι $\mu \in \mathrm{Ext}(M_T(X))$.

(3) Έστω $\mu, \nu \in E_T(X)$ και έστω ότι $\mu \neq \nu$. Τότε υπάρχει $f \in C(X)$ τέτοια ώστε

$$\int f\, d\mu \neq \int f\, d\nu.$$

Από το εργοδικό θεώρημα ισχύει ότι

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) \to \int f d\mu, \qquad \mu\text{-σχεδόν παντού}$$

και

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) \to \int f\, d\nu, \qquad \nu\text{-σχεδόν παντού.}$$

Το σύνολο των $x \in X$ για τα οποία ισχύει η πρώτη σύγκλιση έχει μ-μέτρο 1 και ν-μέτρο 0, αφού $\int f\, d\mu \neq \int f\, d\nu$, άρα τα μ και ν είναι ορθογώνια. □

Απόδειξη Λήμματος 3.4.11 Έστω $f = d\nu/d\mu$ και θεωρούμε $A = \{x \in X\colon f(x) < 1\}$. Τότε

$$\begin{aligned}\int_{A \smallsetminus T^{-1}(A)} f\, d\mu &= \nu\big(A \smallsetminus T^{-1}(A)\big)\\ &= \nu(A) - \nu\big(A \cap T^{-1}(A)\big)\\ &= \nu\big(T^{-1}(A)\big) - \nu\big(A \cap T^{-1}(A)\big)\\ &= \nu\big(T^{-1}(A) \smallsetminus A\big)\\ &= \int_{T^{-1}(A) \smallsetminus A} f\, d\mu,\end{aligned}$$

εφόσον $\nu(A) = \nu\big(T^{-1}(A)\big)$. Όμοια,

$$\mu\big(A \smallsetminus T^{-1}(A)\big) = \mu(A) - \mu\big(A \cap T^{-1}(A)\big) = \mu\big(T^{-1}(A)\big) - \mu\big(A \cap T^{-1}(A)\big) = \mu\big(T^{-1}(A) \smallsetminus A\big),$$

αφού $\mu(A) = \mu\big(T^{-1}(A)\big)$. Η f πάνω στο $A \smallsetminus T^{-1}(A)$ είναι < 1 και πάνω στο $T^{-1}(A) \smallsetminus A$ είναι ≥ 1. Επομένως

$$\int_{A \smallsetminus T^{-1}(A)} f\, d\mu \leq \mu\big(A \smallsetminus T^{-1}(A)\big) = \mu\big(T^{-1}(A) \smallsetminus A\big) \leq \int_{T^{-1}(A) \smallsetminus A} f\, d\mu,$$

με την πρώτη ανισότητα γνήσια αν $\mu\big(A \smallsetminus T^{-1}(A)\big) > 0$. Άρα, για να έχουμε την ισότητα στην $(*)$, πρέπει

$$\mu\big(A \smallsetminus T^{-1}(A)\big) = \mu\big(T^{-1}(A) \smallsetminus A\big) = 0 \quad \Rightarrow \quad \mu\big(A \triangle T^{-1}(A)\big) = 0.$$

Από την εργοδικότητα του μ θα πρέπει

$$\mu(A) = \{0, 1\}.$$

Αν $\mu(A) = 1$, τότε $\nu(A) = \int_A f\, d\mu < 1$ που είναι αδύνατον, αφού λόγω της απόλυτης συνέχειας πρέπει $\nu(A) = 1$ ($\mu(A^c) = 0 \;\Rightarrow\; \nu(A^c) = 0$). Άρα $\mu(A) = 0$ και επομένως $f \geq 1$, μ-σχεδόν παντού. Όμοια δείχνουμε και ότι $\mu(\{x \in X\colon f(x) > 1\}) = 0$, οπότε τελικά $f = 1$, μ-σχεδόν παντού. □

Παράδειγμα 3.4.12 Έστω $S = \{1, 2, \ldots, s\}$, $X = S^{\mathbb{N}}$, T ο μετασχηματισμός shift. Ο $M_T(X)$ περιέχει όλα τα Bernoulli shifts

$$\mu\big(\{\underset{\sim}{x} \in X\colon x_1 = i_1, \ldots, x_n = i_n\}\big) = p_{i_1} p_{i_2} \cdots p_{i_n},$$

$n \in \mathbb{N}$, $i_1, \ldots, i_n \in S$, όπου $(p_i)_{i \in S}$ διάνυσμα πιθανότητας.

Επίσης, τα Markov shifts (p, P), όπου P στοχαστικός πίνακας και p διάνυσμα πιθανότητας που είναι αριστερό ιδιοδιάνυσμα του P, ορίζουν μέτρα στον $M_T(X)$, δηλαδή τα μέτρα

$$\mu\big(\{\underset{\sim}{x} \in X\colon x_1 = i_1, \ldots, x_n = i_n\}\big) = p_{i_1} P_{i_1\, i_2} \cdots P_{i_{n-1}\, i_n},$$

$n \in \mathbb{N}$, $i_1, \ldots, i_n \in S$, ανήκουν στον $M_T(X)$.

Σημειωτέον ότι τα δύο παραπάνω ήδη μέτρων, δηλαδή τα Bernoulli shifts και τα Markov shifts, σε καμία περίπτωση δεν εξαντλούν τον $M_T(X)$. Υπάρχουν πάρα πολλά άλλα ήδη αναλλοίωτων μέτρων για τον χώρο shift (X, T).

Παράδειγμα 3.4.13 (a) (Ενδομορφισμοί του τόρου) Έστω $X = \mathbb{T}$ και $T_2(x) = 2x \pmod 1$. Εδώ το μέτρο Lebesgue λ είναι αναλλοίωτο. Επίσης ισχύει ότι και $\delta_{\{0\}} \in M_{T_2}(X)$ αφού $T_2(0) = 0$. Ακόμα, κάθε περιοδική τροχιά ορίζει αναλλοίωτο μέτρο[22]. Ο μετασχηματισμός T_2, ισοδύναμα, γράφεται και ως $T_2(z) = z^2$, $z \in \{w \in \mathbb{C}\colon |w| = 1\}$ και

$$T_2^n(z) = z \quad \Leftrightarrow \quad z^{2^n} = z \quad \Leftrightarrow \quad z^{2^n - 1} = 1.$$

Άρα κάθε $(2^n - 1)$-ρίζα της μονάδας, $n \in \mathbb{N}$, ορίζει περιοδική τροχιά.

Παράδειγμα 3.4.13 (b) (Ενδομορφισμοί του τόρου) Έστω $X = \mathbb{T}$, $T_k(x) = kx \pmod 1$, $k \in \mathbb{N}$, $k > 1$. Εδώ ομοίως, αναλλοίωτα μέτρα είναι το Lebesgue όπως και αυτά που ορίζονται από περιοδικές τροχιές.

[22] Αν $\{x_0, x_1, \ldots, x_{n-1}\}$ είναι περιοδική τροχιά, τότε το μέτρο $\mu = n^{-1} \sum_{j=0}^{n-1} \delta_{x_j}$ είναι αναλλοίωτο μέτρο πιθανότητας.

Παρατηρεί κανείς πάλι ότι, αν θεωρήσει τον T_k σαν τον μετασχηματισμό $T_k(z) = z^k$ του κύκλου $\mathbb{S}^1$, τότε $T_k^n(z) = z$ ανν $z^{k^n-1} = 1$, δηλαδή όλες οι $(k^n - 1)$-ρίζες της μονάδας, $n \in \mathbb{N}$, ορίζουν περιοδικές τροχιές.

Παράδειγμα 3.4.13 (c) (Ενδομορφισμοί του d-διάστατου τόρου) $X = \mathbb{T}^d$, $T(x) = Ax \pmod 1$, όπου $A = (A_{ij})$ πίνακας $d \times d$ με $A_{ij} \in \mathbb{Z}$ για κάθε i, j και $\det(A) \neq 0$. Και εδώ, το μέτρο Lebesgue και τα μέτρα που ορίζονται από περιοδικές τροχιές ανήκουν στον $M_T(X)$.

Παράδειγμα 3.4.14 (Στροφές του κύκλου ή, ισοδύναμα, μεταφορές της ομάδας $\mathbb{T} \simeq \mathbb{S}^1$) Έστω $X = \mathbb{T}$, $T_a(x) = x + a \pmod 1$. Το μέτρο Lebesgue είναι στον $M_{T_a}(X)$, για κάθε $a \in \mathbb{R}$. Αν $a \in \mathbb{Q}$, τότε υπάρχουν περιοδικές τροχιές (κάθε τροχιά είναι περιοδική για την ακρίβεια) και αυτές ορίζουν και άλλα αναλλοίωτα μέτρα. Αν $a \in \mathbb{R} \setminus \mathbb{Q}$, τότε $M_{T_a}(X) = \{\lambda\}$, όπου λ το μέτρο Lebesgue (που είναι τότε και εργοδικό, αφού αναγκαστικά είναι ακραίο σημείο όντας μοναδικό).

3.5 Εργοδική ανάλυση ενός αναλλοίωτου μέτρου

Πρόταση 3.5.1 Έστω X ένα μετρικοποιήσιμο, συμπαγές και κυρτό σύνολο σε έναν τοπολογικό διανυσματικό χώρο. Τότε $\mathrm{Ext}(X)$ είναι G_δ- σύνολο κα άρα Borel.

Θεώρημα 3.5.2 (εργοδική ανάλυση αναλλοίωτου μέτρου) Έστω (X, T) τ.δ.σ. και $\mu \in M_T(X)$. Τότε υπάρχει μοναδικό Borel μέτρο πιθανότητας λ_μ στον συμπαγή μετρικό χώρο $M_T(X)$, τέτοιο ώστε $\lambda_\mu(E_T(X)) = 1$ και

$$\mu = \int_{M_T(X)} \nu \, d\lambda_\mu(\nu) = \int_{E_T(X)} \nu \, d\lambda_\mu(\nu).$$

Συγκεκριμένα, αυτό σημαίνει ότι

$$\int_X f \, d\mu = \int_{E_T(X)} \left(\int_X f \, d\nu \right) d\lambda_\mu(\nu) \qquad \forall f \in C(X).$$

Απόδειξη. Θεώρημα Choquet. □

Παράδειγμα 3.5.3 Έστω $S = \{1, 2, 3, 4\}$, $X = S^{\mathbb{N}_0}$,

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Τότε $P^n = P$ για κάθε $n \in \mathbb{N}$. Έστω $p = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ και έστω μ το μέτρο Markov που αντιστοιχεί στο ζεύγος (p, P). Το $(X, \mathcal{B}, \mu, T)$, όπου T ο shift μετασχηματισμός, είναι σ.δ.μ. Ο P δεν είναι ανάγωγος, άρα το σύστημα δεν είναι εργοδικό. Τώρα εάν $p^{(1)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$ και $p^{(2)} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, θεωρώντας τα μέτρα Markov $\mu^{(i)}$ που αντιστοιχούν στα ζεύγη $(p^{(i)}, P)$, $i \in \{1, 2\}$, έχουμε ότι καθένα από τα $\mu^{(i)}$ είναι εργοδικό, αφού το καθένα ουσιαστικά αντιστοιχεί στον πίνακα $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, που είναι ανάγωγος και εύκολα ελέγχεται ότι ισχύει

$$\mu = \tfrac{1}{2}\mu^{(1)} + \tfrac{1}{2}\mu^{(2)}.$$

Αυτή είναι η εργοδική ανάλυση του μ. Συγκεκριμένα εδώ το μέτρο λ_μ είναι συγκεντρωμένο στο δισύνολο $\{\mu^{(1)}, \mu^{(2)}\} \subset E_T(X)$, με $\lambda_\mu(\{\mu^{(1)}\}) = \lambda_\mu(\{\mu^{(2)}\}) = \frac{1}{2}$.

Για το ότι κάθε $\mu^{(i)}$, $i \in \{1, 2\}$, είναι εργοδικό, πιο συγκεκριμένα έχει κανείς ότι, για κάθε ένα από τα υποσύνολα $X_1 := \{1, 2\}^{\mathbb{N}}$ και $X_2 := \{3, 4\}^{\mathbb{N}}$ του X ισχύει ότι $T(X_i) \subset X_i$, $i \in \{1, 2\}$. Αν $A \subset X$ είναι αναλλοίωτο, δηλαδή $T^{-1}(A) = A$, τότε και το $X_i \cap A$ είναι T_i-αναλλοίωτο, όπου $T_i \colon X_i \to X_i$, $T_i = T|_{X_i}$ είναι ο περιορισμός του T στο X_i, $i \in \{1, 2\}$: $T_i^{-1}(A \cap X_i) = T^{-1}(A) \cap X_i = A \cap X_i$, για $i \in \{1, 2\}$, η πρώτη ισότητα επειδή $x \in T_i^{-1}(A \cap X_i)$ ανν $x \in X_i$ και $T(x) \in A$. Αν $\tilde{\mu}^{(1)}$ ο περιορισμός του μέτρου $\mu^{(1)}$

στον X_1 με την σ-άλγεβρα $\mathcal{B}_1 := \{B \cap X_1 \colon B \in \mathcal{B}\}$, τότε $\tilde{\mu}^{(1)}$ είναι το Markov μέτρο που αντιστοιχεί στο ζεύγος $(\tilde{p}^{(1)}, P^{(1)})$, όπου $P^{(1)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ και $\tilde{p}^{(1)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, αφού

$$\begin{aligned}
&\tilde{\mu}^{(1)}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X_1 \colon x_0 = i_0, \dots, x_n = i_n\}\big) \\
&\quad = \mu^{(1)}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_0 = i_0, \dots, x_n = i_n\}\big) = p_{i_0} P_{i_0 i_1} \cdots P_{i_{n-1} i_n} = \tilde{p}^{(1)}_{i_0} P^{(1)}_{i_0 i_1} \cdots P^{(1)}_{i_{n-1} i_n}
\end{aligned}$$

για $i_0, \dots, i_n \in \{1, 2\}$, $n \in \mathbb{N}_0$, και είναι εργοδικό γιατί ο πίνακας $P^{(1)}$ είναι ανάγωγος. Έπεται ότι, αφού το $X_1 \cap A$ είναι T_1-αναλλοίωτο, πρέπει να έχει μέτρο $\tilde{\mu}^{(1)}(A \cap X_1) \in \{0, 1\}$. Τότε όμως $\mu^{(1)}(A) = \mu^{(1)}(X_1 \cap A) = \tilde{\mu}^{(1)}(X_1 \cap A) \in \{0, 1\}$, η πρώτη ισότητα επειδή

$$\begin{aligned}
\mu^{(1)}(X \setminus X_1) &= \mu^{(1)}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon \exists\, n \in \mathbb{N}_0 \text{ τέτοιο ώστε } x_n \in \{3, 4\}\}\big) \\
&\le \sum_{n \in \mathbb{N}_0} \mu^{(1)}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_n \in \{3, 4\}\}\big) = 0,
\end{aligned}$$

επειδή από το αναλλοίωτο του μέτρου $\mu^{(1)}$,

$$\mu^{(1)}\big(\{\underset{\sim}{x} = (x_0, x_1, \dots) \in X \colon x_n \in \{3, 4\}\}\big) = \mu^{(1)}\big(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_0 \in \{3, 4\}\}\big) = \sum_{i \in \{3,4\}} p^{(1)}_i = 0$$

για κάθε $n \in \mathbb{N}_0$. Έπεται ότι το μέτρο $\mu^{(1)}$ είναι εργοδικό. Όμοια για το $\mu^{(2)}$.

Παράδειγμα 3.5.4 Έστω $T_a(x) = x + a \pmod 1$ στον τόρο $\mathbb{T}$, όπου $a \in \mathbb{R} \setminus \mathbb{Q}$, που αντιστοιχεί στον μετασχηματισμό $T_a(z) = e^{2\pi i a} z$, $z \in \mathbb{S}^1$, στον κύκλο. Θεωρούμε $T \colon B_1(0) \to B_1(0)$ όπου $B_1(0) = \{z \in \mathbb{C} \colon |z| \le 1\}$ και έστω $X = B_1(0)$. Το (X, T) είναι τ.δ.σ. και το κανονικοποιημένο μέτρο Lebesgue μ είναι T-αναλλοίωτο. Εδώ το μέτρο Lebesgue δεν είναι εργοδικό γιατί κάθε δακτύλιος (με μη κενό εσωτερικό) είναι αναλλοίωτο σύνολο με μέτρο στο $(0, 1)$. Έστω μ_r το κανονικοποιημένο μέτρο Lebesgue στην περιφέρεια ακτίνας r, δηλαδή, για $r \in (0, 1]$, το μ_r είναι (μήκος τόξου)$/2\pi r$ και $\mu_0 = \delta_{\{0\}}$. Κάθε μ_r είναι εργοδικό μέτρο πιθανότητας. Για $f \colon B_1(0) \to \mathbb{C}$ συνεχή έχουμε

$$\int_0^1 \int_0^1 f(x, y)\, d\mu(x, y) = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} f(r \cos t, r \sin t)\, r\, dt dr$$

και

$$\int f\, d\mu_r = \frac{1}{2\pi} \int_0^{2\pi} f(r \cos t, r \sin t)\, dt$$

και επομένως καταλήγουμε στο ότι

$$\mu = \int_0^1 2r\, \mu_r\, dr.$$

Αυτή είναι η εργοδική αναπαράσταση του κανονικοποιημένου Lebesgue μ. Δηλαδή εδώ λ_μ είναι το μέτρο $d\lambda_\mu = 2r\, dr$ συγκεντρωμένο πάνω στο σύνολο $\{\mu_r \colon r \in [0, 1]\} \subset E_T(X)$.

3.6 Μονοσήμαντα εργοδικά συστήματα

Ορισμός 3.6.1 Ένα τ.δ.σ. (X, T) λέγεται μονοσήμαντα εργοδικό αν το $M_T(X)$ είναι μονοσύνολο. Τότε $M_T(X) = E_T(X)$ και το μοναδικό αναλλοίωτο μέτρο είναι εργοδικό.

Όταν ο X είναι συμπαγής μετρικός χώρος, ο $C(X)$ είναι διαχωρίσιμος και έστω $\{f_n \colon n \in \mathbb{N}\}$ αριθμήσιμο, πυκνό υποσύνολο του $C(X)$ και μ εργοδικό μέτρο. Από το θεώρημα του Birkhoff παίρνουμε ότι για κάθε $m \in \mathbb{N}$ υπάρχει X_m με $\mu(X_m) = 1$ τέτοιο ώστε

$$\frac{1}{n} \sum_{k=0}^{n-1} f_m\big(T^k(x)\big) \to \int f_m\, d\mu \qquad \forall\, x \in X_m.$$

Ορίζουμε $Y = \bigcap_{m\in\mathbb{N}} X_m$. Τότε $\mu(Y) = 1$ και

$$\frac{1}{n}\sum_{k=0}^{n-1} f_m\big(T^k(x)\big) \to \int f_m\, d\mu \qquad \forall\, x \in Y,\ \forall\, m \in \mathbb{N}.$$

Επειδή κάθε $f \in C(X)$ προσεγγίζεται ομοιόμορφα από μια f_m, έπεται ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(T^k(x)\big) \to \int f\, d\mu \qquad \forall\, f \in C(X),\ \forall\, x \in Y.$$

Για μονοσήμαντα εργοδικά συστήματα ισχύει κάτι πιο ισχυρό.

Πρόταση 3.6.2 Έστω (X, T) τ.δ.σ. Τότε τα εξής είναι ισοδύναμα.

(1) Για κάθε $f \in C(X)$, $n^{-1}\sum_{k=0}^{n-1} f\circ T^k \to$ σταθερά(f), ομοιόμορφα ως προς x.

(2) Για κάθε $f \in C(X)$, $n^{-1}\sum_{k=0}^{n-1} f\circ T^k \to$ σταθερά(f), σημειακά για κάθε $x \in X$.

(3) Υπάρχει $\mu \in M_T(X)$ τέτοιο ώστε $n^{-1}\sum_{k=0}^{n-1} f\circ T^k(x) \to \int f d\mu$ για κάθε $f \in C(X)$ και $x \in X$.

(4) Το (X, T) είναι μονοσήμαντα εργοδικό.

Απόδειξη. (1) $\Rightarrow$ (2) προφανές.

(2) $\Rightarrow$ (3) Για κάθε $f \in C(X)$ ορίζουμε

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\big(T^k(x)\big) =: L(f).$$

Από υπόθεση, αυτό είμαι ανεξάρτητο του x. Το L είναι γραμμικό και φραγμένο, αφού $|L(f)| \le \|f\|$ (όπου $\|f\|$ είναι η sup-νόρμα της f), ισχύει ότι $L(f) \ge 0$ για κάθε $f \in C(X)$ με $f \ge 0$ και $L(\mathbb{1}_X) = 1$. Άρα από το Θεώρημα του Riesz, το συνεχές γραμμικό συναρτησοειδές L αναπαρίσταται από ένα μέτρο πιθανότητας μ, δηλαδή υπάρχει $\mu \in M(X)$ τέτοιο ώστε

$$L(f) = \int f\, d\mu \qquad \forall\, f \in C(X).$$

Τέλος $\mu \in M_T(X)$. Πράγματι, έχει κανείς ότι

$$\mu \in M_T(X) \quad \Leftrightarrow \quad L(f\circ T) = L(f) \quad \forall\, f \in C(X) \quad \Leftrightarrow \quad \int f\circ T\, d\mu = \int f\, d\mu \quad \forall\, f \in C(X).$$

Όμως για $f \in C(X)$,

$$L(f\circ T) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\circ T\big(T^k(x)\big) = \lim_{n\to\infty}\frac{1}{n}\left[\sum_{k=0}^{n-1} f\circ T^k(x) - f(x) + f\big(T^n(x)\big)\right] = L(f) - 0 + 0 = L(f).$$

(3) $\Rightarrow$ (4) Έστω $\nu \in M_T(X)$ ένα μέτρο διαφορετικό από το μ. Ξέρουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x) \to \int f d\mu \qquad \forall\, f \in C(X),\ \forall\, x \in X.$$

Εφόσον το αριστερό μέλος είναι μικρότερο από $\|f\|$, για κάθε $x \in X$, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης έχουμε ότι

$$\int \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x)\, d\nu(x) \to \int f\, d\mu,$$

ισοδύναμα

$$\frac{1}{n}\sum_{k=0}^{n-1}\int f\circ T^k d\nu \to \int f\,d\mu.$$

Όμως $\nu \in M_T(X)$, άρα

$$\frac{1}{n}\sum_{k=0}^{n-1}\int f\circ T^k\,d\nu = \frac{1}{n}\sum_{k=0}^{n-1}\int f\,d\nu = \int f\,d\nu.$$

Αφού αυτό ισχύει για κάθε $f \in C(X)$, έπεται ότι $\nu = \mu$ και άρα πρέπει τελικά $M_T(X) = \{\mu\}$.

(4) $\Rightarrow$ (1) Έστω ότι $M_T(X) = \{\mu\}$. Αν δεν ισχύει το (1) θα πρέπει να υπάρχει $f \in C(X)$ με

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k - \int f\,d\mu\right\| \nrightarrow 0,$$

όπου $\|\ \|$ η sup-νόρμα. Άρα θα υπάρχουν $\varepsilon > 0$ και $n_1 < n_2 < \cdots$ τέτοια ώστε

$$\left\|\frac{1}{n_j}\sum_{k=0}^{n_j-1} f\circ T^k - \int f\,d\mu\right\| > \varepsilon \qquad \forall\, j \in \mathbb{N}.$$

Δηλαδή, για κάθε $j \in \mathbb{N}$, υπάρχει $x_j \in X$ τέτοιο ώστε

$$\left|\frac{1}{n_j}\sum_{k=0}^{n_j-1} f\circ T^k(x_j) - \int f\,d\mu\right| > \varepsilon. \tag{25}$$

Ορίζουμε

$$\mu_j = \frac{1}{n_j}\sum_{k=0}^{n_j-1}\delta_{T^k(x_j)} \in M(X) \qquad j \in \mathbb{N}.$$

Τότε

$$\int f\,d\mu_j = \frac{1}{n_j}\sum_{k=0}^{n_j-1} f\big(T^k(x_j)\big) \qquad j \in \mathbb{N}.$$

Επειδή ο $M(X)$ είναι συμπαγής μετρικός χώρος, υπάρχει υπακολουθία $j_1 < j_2 < \cdots$ τέτοια ώστε

$$\mu_{j_m} \xrightarrow{w^*} \nu$$

για κάποιο $\nu \in M(X)$. Θα δειχτεί ότι $\nu \in M_T(X)$ και τότε θα πρέπει $\mu = \nu$. Πράγματι, έστω $g \in C(X)$. Τότε,

$$\begin{aligned}\int g\circ T\,d\nu &= \lim_{m\to\infty}\frac{1}{n_{j_m}}\sum_{k=0}^{n_{j_m}-1} g\circ T\big(T^k(x_{j_m})\big)\\ &= \lim_{m\to\infty}\frac{1}{n_{j_m}}\left[\sum_{k=0}^{n_{j_m}-1} g\circ T^k(x_{j_m}) - g(x_{j_m}) + g\big(T^{n_{j_m}}(x_{j_m})\big)\right] = \int g\,d\nu,\end{aligned}$$

αφού η g είναι φραγμένη. Επομένως

$$\int f\,d\nu = \int f\,d\mu = \lim_{m\to\infty}\frac{1}{n_{j_m}}\sum_{k=0}^{n_{j_m}-1} f\big(T^k(x_{j_m})\big)$$

και αυτή η ισότητα έρχεται σε αντίφαση με την (25). Άρα θα πρέπει να έχουμε ομοιόμορφη σύγκλιση, δηλαδή

$$\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k \xrightarrow{n\to\infty} \int f\,d\mu \quad \text{ομοιόμορφα} \quad \forall\, f \in C(X). \qquad \square$$

Πρόταση 3.6.3 Έστω $X = \mathbb{T}$ και $T_a \colon X \to X$ με $T_a(x) = x + a \pmod 1$, $a \in \mathbb{R} \setminus \mathbb{Q}$. Το σύστημα αυτό είναι μονοσήμαντα εργοδικό.

Απόδειξη. Έστω $\mu \in M_{T_a}(X)$. Θα δειχτεί ότι το μ είναι το μέτρο Lebesgue. Υπολογίζουμε

$$\begin{aligned}\big(\widehat{(T_a)_*\mu}\big)(n) &= \int_{[0,1)} e^{-2\pi i x n}\, d(T_a)_*\mu(x) = \int_{[0,1)} e^{-2\pi i n T_a(x)}\, d\mu(x)\\ &= \int_{[0,1)} e^{-2\pi i (x+a)n}\, d\mu(x) = e^{-2\pi i a n}\widehat{\mu}(n)\end{aligned}$$

για$n \in \mathbb{Z}$. Τώρα,

$$\begin{aligned}(T_a)_*\mu = \mu \ &\Leftrightarrow\ \big(\widehat{(T_a)_*\mu}\big)(n) = \widehat{\mu}(n) \quad \forall\, n \in \mathbb{Z}\\ &\Leftrightarrow\ e^{-2\pi i a n}\widehat{\mu}(n) = \widehat{\mu}(n) \quad \forall\, n \in \mathbb{Z}\\ &\Leftrightarrow\ \widehat{\mu}(n)(e^{-2\pi i a n} - 1) = 0 \quad \forall\, n \in \mathbb{Z}\\ &\Rightarrow\ \widehat{\mu}(n) = 0 \quad \forall\, n \in \mathbb{Z} \setminus \{0\}\end{aligned}$$

εφόσον $a \in \mathbb{R} \setminus \mathbb{Q}$. Άρα το μ είναι το μέτρο Lebesgue. □

Πρόταση 3.6.4 Έστω $(x_n)_{n\in\mathbb{N}}$ ακολουθία στον $\mathbb{T}$ και έστω λ το μέτρο Lebesgue. Τα εξής είναι ισοδύναμα.

(1) $n^{-1}\sum_{k=1}^{n} f(x_k) \to \int f(x)\,dx, \quad \forall\, f \in C(X)$.

(2) $n^{-1}\sum_{k=1}^{n} \delta_{\{x_k\}} \xrightarrow{\text{ασθενώς}^*} \lambda$.

(3) $n^{-1}\sum_{k=1}^{n} e^{2\pi i x_k m} \to 0, \quad \forall\, m \in \mathbb{Z} \setminus \{0\}$.

(4) Για κάθε διάστημα $I \subset \mathbb{T}$ έχουμε ότι $n^{-1}|\{k \le n \colon x_k \in I\}| \to \lambda(I)$, όπου για ένα πεπερασμένο υπούνολο A των φυσικών αριθμών $|A|$ συμβολίζει τον πληθάριθμο του A.

Ορισμός 3.6.5 Κάθε τέτοια ακολουθία $(x_n)_{n\in\mathbb{N}}$ που ικανοποιεί μία από τις παραπάνω συνθήκες, λέγεται ισοκατανεμημένη στον $\mathbb{T}$.

Απόδειξη. (1) ⇔ (2) Άμεσο, αφού είναι ο ορισμός της ασθενώς*-σύγκλισης.

(2) ⇔ (3) Είναι άμεσο από το θεώρημα συνέχειας του Levy αφού η ασθενής*-σύγκλιση είναι η σύγκλιση κατά κατανομή της θεωρίας πιθανοτήτων, η οποία ισοδυναμεί με σύγκλιση μετασχηματισμών Fourier των μέτρων (οι μετασχηματισμοί Fourier μέτρων είναι οι χαρακτηριστικές συναρτήσεις των αντίστοιχων κατανομών πιθανότητας).

(1) ⇒ (3) Είναι άμεσο, άμα θεωρήσουμε $f(x) = e^{2\pi i x m}$, $x \in \mathbb{T}$.

(3) ⇒ (1) Έπεται από την πυκνότητα[23] (ως προς την sup-νόρμα) των τριγωνομετρικών πολυωνύμων, δηλαδή γραμμικών συνδυασμών εκθετικών, στον $C(X)$.

(1) ⇒ (4) Έστω $I \subset \mathbb{T} = [0,1)$ ένα διάστημα. Αν $\inf I = a$ και $\sup I = b$, με $a < b$ και $b - a < 1$, ορίζουμε

$$g_n(x) = \begin{cases} 1, & \text{αν } x \in \left[a + \dfrac{b-a}{2n}, b - \dfrac{b-a}{2n}\right],\\ 0, & \text{αν } x \notin I,\\ \dfrac{2n(x-a)}{b-a}, & \text{αν } x \in \left[a, a + \dfrac{b-a}{2n}\right],\\ -\dfrac{2n(x-b)}{b-a}, & \text{αν } x \in \left[b - \dfrac{b-a}{2n}, b\right] \end{cases}$$

[23]συνέπεια του θεωρήματος Stone–Weierstrass

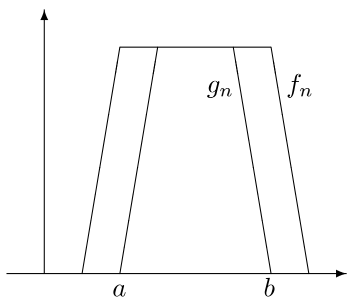
και

$$f_n(x) = \begin{cases} 1, & \text{αν } x \in I, \\ \dfrac{2n(x-a)}{1+a-b} + 1, & \text{αν } x \in \left[a - \dfrac{1+a-b}{2n}, a\right], \\ -\dfrac{2n(x-b)}{1+a-b} + 1, & \text{αν } x \in [0,1) \cap \left[b, b + \dfrac{1+a-b}{2n}\right], \\ -\dfrac{2n(x-b+1)}{1+a-b} + 1, & \text{αν } x \in \left(0, b - 1 + \dfrac{1+a-b}{2n}\right), \\ 0, & \text{αλλιώς,} \end{cases}$$

αν $2na \geq 1 + a - b$ και

$$f_n(x) = \begin{cases} 1, & \text{αν } x \in I \\ \dfrac{2n(x-a)}{1+a-b} + 1, & \text{αν } x \in [0,a] \\ \dfrac{2n(x-a-1)}{1+a-b} + 1, & \text{αν } x \in \left[a + 1 - \dfrac{1+a-b}{2n}, 1\right) \\ -\dfrac{2n(x-b)}{1+a-b} + 1, & \text{αν } x \in \left[b, b + \dfrac{1+a-b}{2n}\right] \\ 0, & \text{αλλιώς} \end{cases}$$

αν $2na < 1 + a - b$, για κάθε $n \in \mathbb{N}$.



Τότε $f_n, g_n \in C(\mathbb{T})$ και $g_n \leq \mathbb{1}_I \leq f_n$, για κάθε $n \in \mathbb{N}$. Άρα,

$$\frac{1}{n}\sum_{k=N}^{n} g_m(x_k) \leq \frac{1}{n}\sum_{k=N}^{n} \mathbb{1}_I(x_k) \leq \frac{1}{n}\sum_{k=N}^{n} f_m(x_k),$$

για κάθε $n \in \mathbb{N}$ και κάθε $k \in \mathbb{N}$. Τώρα, καθώς $n \to \infty$, λόγω της υπόθεσης ότι η $(x_k)_{k\in\mathbb{N}}$ είναι ισοκατανεμημένη, το αριστερό μέλος της ανισότητας τείνει στο $\int_{\mathbb{T}} g_m(x)\,dx$ ενώ το δεξί στο $\int_{\mathbb{T}} f_m(x)\,dx$. Επομένως,

$$\int_{\mathbb{T}} g_m(x)\,dx \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) \leq \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) \leq \int_{\mathbb{T}} f_m(x)\,dx. \tag{26}$$

Επειδή

$$\left|\int f_m(x)\,dx - \int \mathbb{1}_I(x)dx\right| \leq \frac{1+a-b}{m} \qquad \text{και} \qquad \left|\int g_m(x)dx - \int \mathbb{1}_I(x)dx\right| \leq \frac{b-a}{m},$$

από την (26) παίρνουμε ότι

$$\lambda(I) - \frac{b-a}{m} \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) \leq \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) \leq \lambda(I) + \frac{1+a-b}{m} \qquad \forall\, m \in \mathbb{N}.$$

Οπότε αφήνοντας και $m \to \infty$, παίρνει κανείς ότι

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) = \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_I(x_k) = \lambda(I).$$

Αν $a = b$, δηλαδή το I είναι τετριμμένο διάστημα, το παραπάνω επιχείρημα δουλεύει με τις f_n ορισμένες όπως παραπάνω και με $g_n(x) = 0$ για κάθε $n \in \mathbb{N}$ και $x \in \mathbb{T}$. Τέλος αν $b-a = 1$, τότε παίρνουμε $f_n(x) = 1$ για κάθε $n \in \mathbb{N}$ και $x \in \mathbb{T}$ και g_n όπως παραπάνω στο παραπάνω επιχείρημα.

(4) ⇒ (1) Έχουμε ότι $n^{-1} \sum_{k=0}^{n-1} f(x_k) \to \int f(x)\,dx$ για $f = \mathbb{1}_I$, όπου I διάστημα. Κάθε συνεχής συνάρτηση προσεγγίζεται ομοιόμορφα από γραμμικούς συνδυασμούς χαρακτηριστικών συναρτήσεων διαστημάτων και έπεται ότι η (1) ισχύει, αφού η σύγκλιση θα ισχύει για κάθε $f \in C(X)$. □

Σημείωση. Η απόδειξη της ισοδυναμίας (2) ⇔ (3) στην παραπάνω απόδειξη, που χρησιμοποιεί το θεώρημα συνέχειας του Levy, δεν χρειάζεται λογικά αφού αποδεικνύονται και οι (1) ⇔ (2) και (1) ⇔ (3).

Εφαρμογή 3.6.6 (Θεώρημα Weyl) Για κάθε $a \in \mathbb{R} \setminus \mathbb{Q}$, η ακολουθία $(na)_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη mod 1, δηλαδή η ακολουθία $(na \pmod 1)_{n\in\mathbb{N}} = \big(na - \lfloor na \rfloor\big)_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη στο $[0, 1)$.

Απόδειξη. Έστω $T_a : \mathbb{T} \to \mathbb{T}$ με $T_a(x) = x + a \pmod 1$. Ξέρουμε ότι το σύστημα $(\mathbb{T}, T_a)$ είναι μονοσήμαντα εργοδικό (Πρόταση 3.6.3), άρα

$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(T_a^k(x)\big) \to \int f(x)\,dx \qquad \forall\, x \in X \ \ \forall\, f \in C(X).$$

Επομένως ισχύει η (1) της Πρότασης 3.6.4. για την ακολουθία $x_k = T_a^k(x)$, όπου $x \in \mathbb{T}$ αυθαίρετο. Δηλαδή η ακολουθία $(T_a^k(x))_{k\in\mathbb{N}}$ είναι ισοκατανεμημένη για κάθε $x \in \mathbb{T}$ και ειδικότερα αυτό ισχύει για $x = 0$. Αφού $T_a^n(0) = na - \lfloor na \rfloor$, έχουμε το ζητούμενο. □

Γενικότερα, έχει έννοια ο εξής ορισμός.

Ορισμός 3.6.7 Έστω X συμπαγής μετρικός χώρος και μ ένα Borel μέτρο πιθανότητας στον X. Μία ακολουθία $(x_n)_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη ως προς το μέτρο μ αν

$$\frac{1}{n}\sum_{k=1}^{n} \delta_{x_k} \xrightarrow{\text{ασθενώς}^*} \mu$$

ή ισοδύναμα αν

$$\frac{1}{n}\sum_{k=1}^{n} f(x_k) \to \int f\,d\mu \qquad \forall\, f \in C(X).$$

Ορισμός 3.6.8 (generic σημείο) Έστω (X, T) τ.δ.σ. και $\mu \in M_T(X)$. Ένα $x \in X$ λέγεται generic για το μ αν η ακολουθία $\big(T^k(x)\big)_{k\in\mathbb{N}_0}$ είναι ισοκατανεμημένη ως προς μ. Ισοδύναμα, αν

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) \to \int f\,d\mu \qquad \forall\, f \in C(X).$$

Παρατήρηση. Έστω (X, T) τ.δ.σ. Ένα $x \in X$ δεν μπορεί να είναι generic για δύο διαφορετικά μέτρα $\mu, \nu \in M_T(X)$, αφού αν το x είναι generic και για το μ και για το ν πρέπει

$$\int f\,d\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x)) = \int f\,d\nu \qquad \forall\, f \in C(X),$$

η πρώτη ισότητα επειδή το x είναι generic για το μ και η δεύτερη επειδή το x είναι generic για το ν. Άρα πρέπει $\mu = \nu$, από το Πόρισμα 3.4.3.

Πρόταση 3.6.9

(1) Έστω (X, T) τ.δ.σ. και $\mu \in M_T(X)$. Το σύνολο των generic σημείων του μ είναι μετρήσιμο.

(2) Εάν το μ είναι εργοδικό, το σύνολο των μ-generic σημείων έχει μέτρο 1.

Απόδειξη. (1) Έχουμε ότι για κάθε μία $f \in C(X)$, το σύνολο

$$\left\{x \in X : \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) \to \int f\,d\mu\right\} = \bigcap_{N \geq 1}\bigcup_{m \geq 1}\bigcap_{n \geq m}\left\{x \in X : \left|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) - \int f\,d\mu\right| < \frac{1}{N}\right\}$$

είναι μετρήσιμο, αφού κάθε σύνολο $\left\{x \in X : \left|n^{-1}\sum_{k=0}^{n-1} f \circ T^k(x) - \int f\,d\mu\right| < N^{-1}\right\}$ στο δεξί μέλος είναι αντίστροφη εικόνα του $(-N^{-1}, N^{-1})$ μέσω της μετρήσιμης συνάρτησης $\sum_{k=0}^{n-1} f \circ T^k - \left(\int f\,d\mu\right)\mathbb{1}_X$. Επειδή ο X είναι συμπαγής μετρικός χώρος, ο $C(X)$ είναι διαχωρίσιμος και έστω $\{f_m : m \in \mathbb{N}\}$ πυκνό (ως προς την sup νόρμα) υποσύνολο. Εάν

$$X_m = \left\{x \in X : \frac{1}{n}\sum_{k=0}^{n-1} f_m \circ T^k(x) \to \int f_m\,d\mu\right\} \qquad m \in \mathbb{N},$$

τότε $X_m \in \mathcal{B}(X)$ για κάθε m και αν $Y = \bigcap_{m \in \mathbb{N}} X_m$, το Y είναι το σύνολο των generic σημείων του μ και είναι μετρήσιμο.

(2) Αν το μ είναι εργοδικό, τότε κάθε X_m έχει μέτρο 1 λόγω του θεωρήματος Birkhoff. Άρα και $Y = \bigcap_{m \in \mathbb{N}} X_m$ έχει μέτρο 1. □

Θεώρημα 3.6.10 (Furstenberg) Έστω (X, T) μονοσήμαντα εργοδικό τ.δ.σ. Έστω και $c\colon X \to \mathbb{T}^d$ συνεχής. Θεωρούμε το σύστημα $Y = X \times \mathbb{T}^d$, $\mathcal{B}_Y = \mathcal{B}(X) \otimes \mathcal{B}(\mathbb{T}^d)$, $\nu = \mu \times \lambda_d$, όπου μ το μοναδικό αναλλοίωτο μέτρο του (X, T) και λ_d το d-διάστατο μέτρο Lebesgue στον $\mathbb{T}^d$. Έστω $S\colon Y \to Y$ με $S(x, t) = (T(x), c(x) + t)$, $(x, t) \in Y$. Το σύστημα $(Y, \mathcal{B}_Y, \nu, S)$ είναι στρεβλό γινόμενο και άρα είναι σ.δ.μ. και αν είναι εργοδικό τότε είναι μονοσήμαντα εργοδικό.

Για παράδειγμα, αν $X = \mathbb{T}$, $c(x) = x$, $T(x) = x + a \pmod 1$, $a \in \mathbb{R} \setminus \mathbb{Q}$, τότε $S(x, y) = (x + a, x + y)$. Το τ.δ.σ. $(\mathbb{T}^2, S)$ θα είναι μονοσήμαντα εργοδικό αν το στρεβλό γινόμενο $(X \times \mathbb{T}, \mathcal{B}(\mathbb{T}) \otimes \mathcal{B}(\mathbb{T}), \lambda \times \lambda, S)$, δηλαδή το σ.δ.μ. $(\mathbb{T}^2, \mathcal{B}(\mathbb{T}^2), \lambda_2, S)$ είναι εργοδικό, αφού το $(\mathbb{T}, T)$ είναι μονοσήμαντα εργοδικό, με το μέτρο Lebesgue ως μοναδικό αναλλοίωτο μέτρο.

Απόδειξη. Το $(Y, \mathcal{B}_Y, \nu, S)$ είναι στρεβλό γινόμενο και άρα είναι γνωστό ότι είναι σ.δ.μ. Έστω

$$A = \{(x, t) \in Y : (x, t) \text{ generic για το } \mu \times \lambda_d\}.$$

Γνωρίζουμε ότι το A είναι μετρήσιμο (από Πρόταση 3.6.9).

Ισχυρισμός 1. Αν $(x, t) \in A$ και $s \in \mathbb{T}^d$, τότε $(x, s + t) \in A$.

Πράγματι, έστω $(x, t) \in A$ Τότε

$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(S^k(x, t)\big) \to \int f\,d(\mu \times \lambda_d) \qquad \forall\, f \in C(Y).$$

Έστω $s \in \mathbb{T}^d$. Για $f \in C(Y)$, θεωρούμε την συνάρτηση $f_s(y, \theta) = f(y, \theta + s)$, $y \in X$, $s, \theta \in \mathbb{T}$. Αυτή είναι συνεχής ως σύνθεση της συνεχούς f με τον συνεχή μετασχηματισμό $(y, \theta) \mapsto (y, \theta + s)$ του Y. Επομένως έχουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} f_s\big(S^k(x, t)\big) \to \int f_s\,d(\mu \times \lambda_d). \tag{27}$$

Όμως

$$\begin{aligned} f_s\big(S^k(x, t)\big) &= f_s\big(T^k(x), c(x) + c(T(x)) + \cdots + c(T^{k-1}(x)) + t\big) \\ &= f\big(T^k(x), c(x) + c(T(x)) + \cdots + c(T^{k-1}(x)) + t + s\big) = f\big(S^k(x, t + s)\big) \end{aligned}$$

για κάθε k και

$$\int f_s\,d(\mu \times \lambda_d) = \iint f(y, \theta + s)\,d\lambda_d(\theta)d\mu(y) = \iint f(y, \theta)\,d\lambda_d(\theta)d\mu(y) = \int f\,d(\mu \times \lambda_d),$$

επειδή το μέτρο Lebesgue λ_d είναι το μέτρο Haar της ομάδας $\mathbb{T}^d$ και άρα είναι αναλλοίωτο ως προς μεταφορές και άρα είναι αναλλοίωτο ως προς την μεταφορά $\theta \mapsto \theta + s$ ειδικότερα. Επομένως η (27) είναι ισοδυναμεί με την
$$\frac{1}{n}\sum_{k=0}^{n-1} f\big(S^k(x,t+s)\big) \to \int f\,d(\mu\times\lambda_d).$$
Αφού αυτό ισχύει για αυθαίρετη $f \in C(Y)$, έπεται ότι $(x,t+s) \in A$.

Ισχυρισμός 2. Υπάρχει μετρήσιμο $B \in \mathcal{B}(X)$ τέτοιο ώστε $A = B \times \mathbb{T}^d$ με $\mu(B) = 1$.

Πράγματι, σταθεροποιούμε $(x,t) \in A$, το οποίο είναι μη κενό αφού υποθέτουμε το μέτρο $\mu\times\lambda_d$ εργοδικό και άρα $\mu\times\lambda_d(A) = 1$, και ορίζουμε
$$B = \{y \in X \colon (y,t) \in A\}.$$
Τότε $B \in \mathcal{B}(X)$ από Fubini, αφού $A \in \mathcal{B}(X)\otimes\mathcal{B}(\mathbb{T}^d)$. Αν $y \in B$, $s \in \mathbb{T}^d$, τότε $y \in B \Rightarrow (y,t) \in A \Rightarrow (y,s) = (y, t+(s-t)) \in A$ (Ισχυρισμός 1), άρα $B \times \mathbb{T}^d \subset A$. Αντίστροφα, αν $(y,s) \in A$ τότε $(y,t) = (y, s+(t-s)) \in A$ (Ισχυρισμός 1) και άρα $y \in B$ και επομένως $(y,s) \in B\times\mathbb{T}^d$. Άρα $A \subset B\times\mathbb{T}^d$ και τελικά $A = B\times\mathbb{T}^d$. Επίσης
$$\mu(B) = \mu(B)\lambda_d(\mathbb{T}^d) = \mu\times\lambda_d(B\times\mathbb{T}^d) = \mu\times\lambda_d(A) = 1,$$
η τελευταία ισότητα επειδή υποθέτουμε το $\mu\times\lambda_d$ εργοδικό.

Έστω $\nu \in M_S(Y)$. Μπορούμε να υποθέσουμε ότι το ν είναι εργοδικό, αφού αν δείξουμε ότι κάθε εργοδικό ν ισούται με $\mu\times\lambda_d$, τότε κάθε αναλλοίωτο ν θα ισούται με $\mu\times\lambda_d$ από εργοδική αναπαράσταση του ν. Θεωρούμε $\pi\colon Y \to X$, $\pi(x,t) = x$. Τότε $\pi\circ S = T\circ\pi$. Έστω επίσης $\mu' = \pi_*\nu$. Τότε
$$\mu'\big(T^{-1}(E)\big) = \nu\big(\pi^{-1}(T^{-1}(E))\big) = \nu\big(S^{-1}(\pi^{-1}(E))\big) = \nu\big(\pi^{-1}(E)\big) = \mu'(E).$$
Άρα το μ' είναι T-αναλλοίωτο και άρα $\mu' = \mu$. Επομένως,
$$\nu(A) = \nu(B\times\mathbb{T}^d) = \mu'(B) = \mu(B) = 1.$$
Όμως τα generic σημεία του ν έχουν επίσης ν-μέτρο 1, επειδή υποθέτουμε το ν εργοδικό. Άρα πρέπει να υπάρχει $(x,t) \in A$ το οποίο να είναι επίσης generic για το ν. Όμως ένα σημείο δεν μπορεί να είναι generic για δύο διαφορετικά αναλλοίωτα μέτρα, από τον ορισμό του generic ως προς ένα μέτρο, άρα πρέπει $\nu = \mu\times\lambda_d$. $\square$

Πόρισμα 3.6.11 Το σύστημα $(\mathbb{T}^d, S)$, όπου
$$S(t_1,t_2,\dots,t_d) = (t_1+a, t_1+t_2, \dots, t_d+t_{d-1}) \ (\text{mod } 1)$$
με $a \in \mathbb{R}\setminus\mathbb{Q}$, είναι μονοσήμαντα εργοδικό.

Απόδειξη. Έστω $X_1 = \mathbb{T}$, $S_1(x) = x + a$ (mod 1) και έστω $X_k = X_{k-1}\times\mathbb{T} = \mathbb{T}^k$ και $S_k\colon X_k \to X_k$ με
$$S_k(t_1,t_2,\dots,t_k) = (S_{k-1}(t_1,\dots,t_{k-1}), t_{k-1}+t_k) \ (\text{mod } 1) = (t_1+a, t_1+t_2,\dots,t_k+t_{k-1}) \ (\text{mod } 1)$$
για $k \in \mathbb{N}$ με $k > 1$. Τότε $X_d = \mathbb{T}^d$ και $S_d = S$. Επειδή, για $d = 1$, το σύστημα $(\mathbb{T}, S_1)$ είναι μονοσήμαντα εργοδικό όταν $a \in \mathbb{R}\setminus\mathbb{Q}$, με το μέτρο Lesbesgue λ_1 ως το μοναδικό αναλλοίωτο μέτρο, αρκεί, από το προηγούμενο Θεώρημα, να δείξουμε ότι το σύστημα $(X_k, \mathcal{B}(X_k), \lambda_k, S_k)$ είναι εργοδικό για κάθε $k \in \mathbb{N}$, $k > 1$, όπου λ_k το μέτρο Lebesgue στον $X_k = \mathbb{T}^k$ και $\mathcal{B}(X_k)$ η Borel σ-άλγεβρα του $X_k = \mathbb{T}^k$. Γιατί τότε, εφαρμόζοντας το προηγούμενο Θεώρημα 3.6.10 με $X = X_{k-1}$, $T = S_{k-1}$ και $c(t_1,\dots,t_{k-1}) = t_{k-1}$ για $(t_1,\dots,t_{k-1}) \in X$, θα παίρνει κανείς την μονοσήμαντη εργοδικότητα του συστήματος (X_k, S_k) από την μονοσήμαντη εργοδικότητα του (X_{k-1}, S_{k-1}), αφού θα γνωρίζει ότι το
$$(X\times\mathbb{T}, \mathcal{B}(X)\otimes\mathcal{B}(\mathbb{T}), \lambda_{k-1}\times\lambda_1, S_k) = (X_k, \mathcal{B}(X_k), \lambda_k, S_k)$$
είναι εργοδικό.

Έστω $f \in L^2(\mathbb{T}^k, \mathcal{B}(\mathbb{T}^k), \lambda_k)$ με $f\circ S_k \stackrel{L^2}{=} f$. Θεωρούμε το ανάπτυγμα Fourier της f
$$f \stackrel{L^2}{=} \sum_{n\in\mathbb{Z}^k} \widehat{f}(n)\, e_n, \qquad e_n(t) = e^{2\pi i\langle n,t\rangle},\ n \in \mathbb{Z}^k,\ t \in \mathbb{T}^k.$$

Για τους συντελεστές της $f\circ S_k$ παρατηρεί κανείς ότι, αν $\tilde S\colon \mathbb{R}^k\to\mathbb{R}^k$ η (γραμμική) απεικόνιση

$$\tilde S\begin{bmatrix} t_1\\ t_2\\ t_3\\ \vdots\\ t_{k-1}\\ t_k\end{bmatrix}=\begin{bmatrix}1&0&0&\dots&0&0&0\\1&1&0&\dots&0&0&0\\0&1&1&\dots&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\\0&0&0&\dots&1&1&0\\0&0&0&\dots&0&1&1\end{bmatrix}\begin{bmatrix} t_1\\ t_2\\ t_3\\ \vdots\\ t_{k-1}\\ t_k\end{bmatrix}=\begin{bmatrix} t_1\\ t_1+t_2\\ t_2+t_3\\ \vdots\\ t_{k-2}+t_{k-1}\\ t_{k-1}+t_k\end{bmatrix}$$

και $\tilde S^*$ ο συζυγής (δηλαδή εδώ ο ανάστροφος) του $\tilde S$, τότε

$$\begin{aligned}(\widehat{f\circ S_k})(\tilde S^* n)&=\int_{\mathbb{T}^k}\exp\Big(-2\pi i\langle \tilde S^* n,t\rangle\Big)\,f(S_k(t))\,d\lambda_{\mathbb{T}^k}(t)\\&=\int_{\mathbb{T}^k}\exp\Big(-2\pi i\langle n,\tilde S t\rangle\Big)\,f(S_k(t))\,d\lambda_{\mathbb{T}^k}(t)\\&=e^{2\pi i n_1 a}\int_{\mathbb{T}^k}\exp\Big(-2\pi i(\langle n,\tilde S t\rangle+n_1 a)\Big)\,f(S_k(t))\,d\lambda_{\mathbb{T}^k}(t)\\&=e^{2\pi i n_1 a}\int_{\mathbb{T}^k}\exp\left(-2\pi i\langle n,S_k(t)\rangle\right)f(S_k(t))\,d\lambda_{\mathbb{T}^k}(t)\\&=e^{2\pi i n_1 a}\int_{\mathbb{T}^k}\exp\left(-2\pi i(\langle n,t\rangle)\right)f(t)\,d\lambda_{\mathbb{T}^k}(t)\\&=e^{2\pi i n_1 a}\widehat f(n),\end{aligned}\tag{28}$$

όπου $n_1,\dots,n_k$ είναι οι συντεταγμένες του $n\in\mathbb{Z}^k$ και στην προτελευταία ισότητα χρησιμοποιήθηκε το γεγονός ότι το $\lambda_{\mathbb{T}^k}$ είναι S_k-αναλλοίωτο. Επομένως, αφού $f\overset{L^2}{=}f\circ S_k$, και άρα $\widehat f(m)=(\widehat{f\circ S_k})(m)$ $\forall\, m\in\mathbb{Z}^k$, πρέπει

$$\widehat f\big(\tilde S^*(n)\big)=\widehat f(n)e^{2\pi i n_1 a}\qquad \forall\, n\in\mathbb{Z}^k.$$

Άρα και

$$\big|\widehat f\big(\tilde S^*(n)\big)\big|=\big|\widehat f(n)\big|\qquad \forall\, n\in\mathbb{Z}^k,$$

επομένως και

$$\big|\widehat f\big((\tilde S^*)^p(n)\big)\big|=\big|\widehat f(n)\big|\qquad \forall\, n\in\mathbb{Z}^k\ \ \forall\, p\in\mathbb{N}_0.$$

Αν για κάποιο $n\in\mathbb{Z}^k$ τα $(\tilde S^*)^p n$, $p\in\mathbb{N}_0$, είναι όλα διαφορετικά μεταξύ τους, το σύνολο

$$\big\{(\tilde S^*)^p n\colon p\in\mathbb{N}_0\big\}$$

είναι άπειρο και η ισότητα Parseval

$$\infty>\|f\|_2^2=\sum_{m\in\mathbb{Z}^k}\big|\widehat f(m)\big|^2\geq\sum_{m\in\big\{(\tilde S^*)^p n\colon p\in\mathbb{N}_0\big\}}\big|\widehat f(m)\big|^2=\sum_{p\in\mathbb{N}_0}\big|\widehat f(n)\big|^2$$

μας δίνει ότι $\big|\widehat f(n)\big|=0$.

Αν πάλι $(\tilde S^*)^{p_1}n=(\tilde S^*)^{p_2}n$ για κάποια $p_1,p_2\in\mathbb{N}_0$ με $p_1\neq p_2$, για κάποιο $n\in\mathbb{Z}^k$, τότε $(\tilde S^*)^p n=n$ για κάποιο $p\in\mathbb{N}$ (το $p=|p_1-p_2|$), γι αυτό το $n\in\mathbb{Z}^k$.

Ισχυρισμός. Αν για κάποιο $n\in\mathbb{Z}^k$, $(\tilde S^*)^p n=n$ για κάποιο $p\in\mathbb{N}$, τότε $n_2=\dots=n_k=0$.

Δοθέντος του Ισχυρισμού, έχουμε ότι για τα $n\in\mathbb{Z}^k$ με $(\tilde S^*)^{p_1}n=(\tilde S^*)^{p_2}n$ για κάποια $p_1,p_2\in\mathbb{N}_0$ με $p_1\neq p_2$, θα έχουμε $n_2=\dots=n_k=0$ και η σχέση (28) μας δίνει ότι

$$\widehat f(n)(e^{2\pi i n_1 a}-1)=0,$$

αφού

$$\tilde S^* n=\tilde S^*\begin{bmatrix}n_1\\0\\ \vdots\\0\end{bmatrix}=\begin{bmatrix}n_1\\0\\ \vdots\\0\end{bmatrix}=n.$$

Τώρα, επειδή $a \in \mathbb{R} \setminus \mathbb{Q}$ και άρα $e^{2\pi i n_1 a} - 1 \neq 0$ αν $n_1 \neq 0$, έπεται ότι $\widehat{f}(n) = 0$ αν $n = (n_1, 0, \dots, 0)$ με $n_1 \neq 0$. Επομένως, έχει δειχτεί ότι $\widehat{f}(n) = 0 \ \forall n \in \mathbb{Z}^k \setminus \{0\}$ και άρα

$$f \overset{L^2}{=} \widehat{f}(0),$$

δηλαδή η f είναι σταθερή σχεδόν παντού. Έπεται ότι το σύστημα $(\mathbb{T}^k, \mathcal{B}(\mathbb{T}^k), \lambda_k, S_k)$ είναι εργοδικό.

Για την απόδειξη του Ισχυρισμού θα χρησιμοποιήσουμε το επόμενο Λήμμα.

Λήμμα. Για $1 \leq n < k-1$,

$$\tilde{S}^n = \begin{bmatrix} 1 &&&&&&& \\ n & 1 &&&&&& \\ \binom{n}{2} & n & 1 &&&&& \\ \vdots & \vdots & \ddots & \ddots &&&& \\ \binom{n}{n} & \binom{n}{n-1} & \dots & n & 1 &&& \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots && \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots && \\ 0 & \dots & 0 & \binom{n}{n} & \binom{n}{n-1} & \dots & n & 1 \end{bmatrix}$$

και για $n \geq k-1$,

$$\tilde{S}^n = \begin{bmatrix} 1 &&&& \\ n & 1 &&& \\ \binom{n}{2} & n & 1 && \\ \vdots & \ddots & \ddots & \ddots & \\ \binom{n}{k-1} & \dots & \dots & n & 1 \end{bmatrix}$$

Απόδειξη. Με επαγωγή στο n. Έχουμε ότι $\tilde{S}_{ij} = 1$ αν $j \in \{i-1, i\}$, $i \in \{2, \dots, k\}$, $\tilde{S}_{11} = 1$ και $\tilde{S}_{ij} = 0$ για όλα τα άλλα i, j. Ο ισχυρισμός είναι ότι, για αυθαίρετο $n \in \mathbb{N}$,

$$\tilde{S}^n_{ij} = \binom{n}{i-j} \quad \text{για } i, j \in \{1, \dots, k\} \quad \text{με } i \geq j \geq i-n$$

και $\tilde{S}^n_{ij} = 0$ για όλα τα άλλα i, j. Τότε για τον $\tilde{S}^{n+1}$ έχουμε

$$\tilde{S}^{n+1}_{ij} = \sum_{m=1}^{k} \tilde{S}_{im} \tilde{S}^n_{mj} = \tilde{S}^n_{ij} + \tilde{S}^n_{i-1\ j} = \binom{n}{i-j} + \binom{n}{i-1-j} = \binom{n+1}{i-j} \tag{29}$$

αν $i \in \{2, \dots, k\}$ και $j \in \{1, \dots, k\}$ με $i \geq j \geq i-n$ και $i-1 \geq j \geq i-1-n$, ισοδύναμα αν $i \in \{2, \dots, k\}$ και $j \in \{1, \dots, k\}$ με $i > j \geq i-n$.

Για $i = 1$,

$$\tilde{S}^{n+1}_{1j} = \sum_{m=1}^{k} \tilde{S}_{1m} \tilde{S}^n_{mj} = \tilde{S}_{11}\, \tilde{S}^n_{1j} = \tilde{S}^n_{1j} = \delta_{1j} = \begin{cases} 1, & \text{αν } j = 1 \\ 0, & \text{αν } j \neq 1. \end{cases}$$

Επίσης, για $i \in \{2, \dots, k\}$ και $j \in \{1, \dots, k\}$ με $i < j$, έχουμε $\tilde{S}^n_{ij} = 0$ και $\tilde{S}^n_{i-1\ j} = 0$ άρα η (29) δίνει ότι $\tilde{S}^{n+1}_{ij} = 0$. Ομοίως, για $i \in \{2, \dots, k\}$ και $j \in \{1, \dots, k\}$ με $j < i-1-n = i-(n+1)$, $\tilde{S}^n_{ij} = 0$ και $\tilde{S}^n_{i-1\ j} = 0$ πάλι, οπότε, από την (29), έχουμε ότι $\tilde{S}^{n+1}_{ij} = 0$. Τέλος, για $i = j,\ i \in \{2, \dots, k\}$,

$$\tilde{S}^{n+1}_{ij} = \sum_{m=1}^{k} \tilde{S}_{im} \tilde{S}^n_{mj} = \tilde{S}^n_{ii} + \tilde{S}^n_{i-1\ i} = \tilde{S}^n_{ii} = 1 = \binom{n+1}{i-j}$$

και για $j = i-1-n = i-(n+1),\ i \in \{2, \dots, k\}$,

$$\tilde{S}^{n+1}_{ij} = \tilde{S}^n_{ij} + \tilde{S}^n_{i-1\ j} = \tilde{S}^n_{i-1\ j} = \binom{n}{i-j-1} = \binom{n}{n} = \binom{n+1}{n+1} = \binom{n+1}{i-j}.$$

Άρα τελικά,

$$\tilde{S}^{n+1}_{ij} = \begin{cases} \binom{n+1}{i-j} & \text{για } i \in \{2,\dots,k\},\ j \in \{1,\dots,k\} \text{ με } i \geq j \geq i-(n+1) \\ 0 & \text{αλλιώς.} \end{cases}$$ □

2η απόδειξη. Θεωρούμε τον πίνακα

$$M = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & & \\ & & & 1 & 0 \end{bmatrix}$$

δηλαδή

$$M_{ij} = \delta_{i-1,j} = \begin{cases} 1, & \text{αν } j = i-1 \\ 0, & \text{αν } j \neq i-1 \end{cases} \qquad i,j \in \{1,\dots,k\}.$$

Τότε $\tilde{S} = I + M$, όπου I ο $k \times k$ μοναδιαίος πίνακας. Από το διωνυμικό θεώρημα

$$\tilde{S}^n = (I+M)^n = \sum_{m=0}^{n} \binom{n}{m} M^m.$$

Παρατηρούμε τώρα ότι

$$M^2_{ij} = \sum_{l=1}^{k} M_{il} M_{lj} = \sum_{l=1}^{k} \delta_{i-1,l}\, \delta_{l-1,j} = \delta_{i-2,j}$$

και επαγωγικά,

$$M^m_{ij} = \delta_{i-m,j} \qquad \forall\, i,j \in \{1,\dots,k\},\ m \in \mathbb{N}_0.$$

Επομένως,

$$\tilde{S}^n_{ij} = \sum_{m=0}^{n} \binom{n}{m} \delta_{i-m,j} = \binom{n}{i-j}$$

εφόσον $0 \leq i-j \leq n \Leftrightarrow i-n \leq j \leq i$ και $\tilde{S}^n_{ij} = 0$ αλλιώς. □

Απόδειξη Ισχυρισμού. Το Λήμμα δίνει ότι η i-συντεταγμένη του διανύσματος $(\tilde{S}^*)^p n$, όπου n το διάνυσμα με συντεταγμένες $n_1,\dots,n_k$, είναι

$$\sum_{j=1}^{k} (\tilde{S}^*)^p_{ij} n_j = \sum_{j=1}^{k} \tilde{S}^p_{ji}\, n_j = \sum_{\substack{j=1 \\ j \geq i \geq j-p}}^{k} \binom{p}{j-i} n_j = \sum_{j=i}^{\min\{k,i+p\}} \binom{p}{j-i} n_j.$$

Αν $(\tilde{S}^*)^p n = n$, για $i = k$ η παραπάνω σχέση δίνει

$$\sum_{j=k}^{k} \binom{p}{j-k} n_j = n_k \quad \Leftrightarrow \quad n_k = n_k,$$

που είναι ταυτολογία. Για $i = k-1$ δίνει (αφού $p \geq 1$)

$$\sum_{j=k-1}^{k} \binom{p}{j-k+1} n_j = n_{k-1} \quad \Leftrightarrow \quad n_{k-1} + p n_k = n_{k-1} \quad \Leftrightarrow \quad n_k = 0.$$

Επαγωγικά, αν έχει δειχθεί ότι $n_{i+1} = n_{i+2} = \cdots = n_k = 0$, η εξίσωση $(\tilde{S}^*)^p n = n$ για την $i-1$ συντεταγμένη δίνει

$$\sum_{j=i-1}^{\min\{k,p+i-1\}} \binom{p}{j-i+1} n_j = n_{i-1} \quad \Leftrightarrow \quad n_{i-1} + p n_i = n_{i-1} \quad \Leftrightarrow \quad n_i = 0.$$

Έπεται ότι $n_k = 0,\ n_{k-1} = 0, \dots, n_2 = 0$ (χρησιμοποιώντας την εξίσωση μέχρι $i - 1 = 1$). $\square$

Θεώρημα 3.6.12 (Weyl) Αν $p(z) = a_d z^d + \cdots + a_1 z + a_0$ πολυώνυμο με τουλάχιστον ένα συντελεστή, εκτός του σταθερού όρου, άρρητο, τότε η ακολουθία $(p(n))_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη $(\mathrm{mod}\ 1)$.

1η Απόδειξη. Καταρχήν υποθέτουμε ότι ο συντελεστής του μεγιστοβάθμιου όρου a_d είναι άρρητος. Θεωρούμε το μονοσήμαντα εργοδικό σύστημα $(\mathbb{T}^d, S)$ όπου

$$S(t_1, t_2, \dots, t_d) = (t_1 + a, t_1 + t_2, \dots, t_{d-1} + t_d)\ (\mathrm{mod}\ 1),$$

με $a := a_d d! \in \mathbb{R} \setminus \mathbb{Q}$. Ορίζουμε πολυώνυμα $p_0, p_1, \dots, p_d$ ορίζοντας $p_d := p$ και, για κάθε $i \in \{0, 1, \dots, d-1\}$,

$$p_i(z) := p_{i+1}(z+1) - p_{i+1}(z).$$

Ισχυρισμός: Για $i \in \{0, 1, \dots, d\}$, το p_i είναι πολυώνυμο βαθμού i με συντελεστή μεγιστοβάθμιου όρου $a/i!$.

Απόδειξη Ισχυρισμού: Για $i = d$ ο Ισχυρισμός ισχύει από τους ορισμούς των p_d και a. Υποθέτουμε τώρα ότι ο Ισχυρισμός ισχύει για κάποιο $i \in \{2, \dots, d\}$. Τότε, αν $p_i(z) = \sum_{j=0}^{i} c_j z^j$, τότε

$$\begin{aligned}
p_{i-1}(z) &= p_i(z+1) - p_i(z) \\
&= \sum_{j=0}^{i} c_j (z+1)^j - \sum_{j=0}^{i} c_j z^j \\
&= \frac{a}{i!}[(z+1)^i - z^i] + c_{i-1}[(z+1)^{i-1} - z^{i-1}] + q_{i-2}(z) \\
&= \frac{a}{i!}\left[\sum_{j=0}^{i} \binom{i}{j} z^j - z^i\right] + c_{i-1}\left[\sum_{j=0}^{i-1} \binom{i}{j} z^j - z^{i-1}\right] + q_{i-2}(z) \\
&= \frac{a}{i!}\sum_{j=0}^{i-1} \binom{i}{j} z^j + c_{i-1}\sum_{j=0}^{i-2} \binom{i}{j} z^j + q_{i-2}(z) \\
&= \frac{a}{i!}\binom{i}{i-1} z^{i-1} + \frac{a}{i!}\sum_{j=0}^{i-2} \binom{i}{j} z^j + c_{i-1}\sum_{j=0}^{i-2} \binom{i}{j} z^j + q_{i-2}(z) \\
&= \frac{a}{(i-1)!} z^{i-1} + Q_{i-2}(z),
\end{aligned}$$

όπου τα q_{i-2} και Q_{i-2} είναι πολυώνυμα βαθμού το πολύ $i-2$. Επομένως, επαγωγικά, ο Ισχυρισμός ισχύει για τα $p_1, p_2, \dots, p_d$. Τέλος για $i = 0$ έχουμε ότι $p_0(z) = p_1(z+1) - p_1(z) = a(z+1) - az = a$, και ο Ισχυρισμός ισχύει και για $i = 0$. $\square$

Παρατηρούμε τώρα ότι, για αυθαίρετο $n \in \mathbb{N}$,

$$\begin{aligned}
S((p_1(n), p_2(n), \dots, p_d(n))\ (\mathrm{mod}\ 1)) &= S(p_1(n) - \lfloor p_1(n) \rfloor, p_2(n) - \lfloor p_2(n) \rfloor, \dots, p_d(n) - \lfloor p_d(n) \rfloor) \\
&= (p_1(n) + p_0(n), p_2(n) + p_1(n), \dots, p_d(n) + p_{d-1}(n))\ (\mathrm{mod}\ 1) \\
&= (p_1(n+1), p_2(n+1), \dots, p_d(n+1))\ (\mathrm{mod}\ 1)
\end{aligned}$$

από τον ορισμό των πολυωνύμων $p_0, p_1, \dots, p_d$ και τον Ισχυρισμό. Επομένως επαγωγικά,

$$S^n((p_1(0), p_2(0), \dots, p_d(0))\ (\mathrm{mod}\ 1)) = (p_1(n), p_2(n), \dots, p_d(n))\ (\mathrm{mod}\ 1).$$

Αφού το σύστημα $(\mathbb{T}^d, S)$ είναι μονοσήμαντα εργοδικό, έχουμε ότι

$$\frac{1}{n}\sum_{k=0}^{n-1} F(S^k(p_1(0), p_2(0), \dots, p_d(0))\ (\mathrm{mod}\ 1)) \to \int_{\mathbb{T}^d} F\, d\lambda_{\mathbb{T}^d}$$

και άρα

$$\frac{1}{n}\sum_{k=0}^{n-1} F((p_1(k), p_2(k), \dots, p_d(k))\ (\mathrm{mod}\ 1) \to \int_{\mathbb{T}^d} F\, d\lambda_{\mathbb{T}^d},$$

για κάθε $F \in C(\mathbb{T}^d)$. Ειδικότερα, αν $f \in C(\mathbb{T})$, τότε η $F\colon \mathbb{T}^d \to \mathbb{C}$ με $F(t_1, \ldots, t_d) = f(t_d)$, $(t_1, \ldots, t_d) \in \mathbb{T}^d$, είναι συνεχής[24] άρα

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1} f(p_d(k) \pmod 1) =& \frac{1}{n}\sum_{k=0}^{n-1} F((p_1(k), p_2(k), \ldots, p_d(k)) \pmod 1)\\ &\to \int_{\mathbb{T}^d} F\, d\lambda_{\mathbb{T}^d}\\ &= \int_{\mathbb{T}} \cdots \int_{\mathbb{T}}\int_{\mathbb{T}} f(t_d)\, d\lambda_{\mathbb{T}}(t_1) \cdots d\lambda_{\mathbb{T}}(t_{d-1}) d\lambda_{\mathbb{T}}(t_d)\\ &= \int f\, d\lambda_{\mathbb{T}}.\end{aligned}$$

Αυτό δείχνει ότι η ακολουθία $(p(n))_{n\in\mathbb{N}} = (p_d(n))_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη $\pmod 1$.

Στην γενική περίπτωση, ο συντελεστής του μεγιστοβάθμιου όρου ενδέχεται να μην είναι άρρητος. Η απόδειξη για την γενική περίπτωση μπορεί να γίνει με επαγωγή. Καταρχήν, για $d = 1$ ο άρρητος συντελεστής είναι αναγκαστικά ο συντελεστής του μεγιστοβάθμιου όρου, οπότε το αποτέλεσμα ισχύει, από την περίπτωση που εξετάσαμε παραπάνω. Έστω τώρα $d \in \mathbb{N}$ με $d > 1$, και έστω ότι το συμπέρασμα του Θεωρήματος ισχύει για όλα τα πολυώνυμα βαθμού $< d$ με έναν τουλάχιστον συντελεστή εκτός του σταθερού όρου άρρητο. Έστω $p(x) = a_d x^d + \cdots + a_0$ πολυώνυμο βαθμού d με έναν τουλάχιστον συντελεστή εκτός του σταθερού όρου άρρητο. Αν a_d είναι άρρητος, η απόδειξη έχει δοθεί παραπάνω. Έστω λοιπόν ότι $a_d \in \mathbb{Q}$ και έστω $m \in \mathbb{N}$ τέτοιος ώστε $ma_d \in \mathbb{Z}$. Σταθεροποιούμε, καταρχήν, και έναν ακέραιο $\upsilon \in \{0, 1, \ldots, m-1\}$. Τότε, για $k \in \mathbb{N}_0$,

$$p(km+\upsilon) = \sum_{i=0}^{d} a_i(km+\upsilon)^i = \sum_{i=0}^{d} a_i \sum_{j=0}^{i}\binom{i}{j} k^j m^j \upsilon^{i-j} = \sum_{j=0}^{d} k^j m^j \sum_{i=j}^{d}\binom{i}{j} a_i \upsilon^{i-j} = k^d m^d a_d + p_\upsilon(k), \tag{30}$$

όπου p_υ το πολυώνυμο $p_\upsilon(x) = c_{d-1}^{(\upsilon)} x^{d-1} + \cdots + c_1^{(\upsilon)} x + c_0^{(\upsilon)}$ με συντελεστές

$$c_j^{(\upsilon)} = m^j \sum_{i=j}^{d}\binom{i}{j} a_i \upsilon^{i-j} = m^j \left[\binom{d}{j} a_d \upsilon^{d-j} + \binom{d-1}{j} a_{d-1}\upsilon^{d-j-1} + \cdots + (j+1)a_{j+1}\upsilon + a_j\right],$$

$j \in \{0, 1, \ldots, d-1\}$, βαθμού $\leq d-1$ και με έναν τουλάχιστον συντελεστή άρρητο: αν $j \in \{0, 1, \ldots, d-1\}$ είναι ο μεγαλύτερος δείκτης με τον συντελεστή a_j του αρχικού πολυωνύμου p άρρητο, τότε και ο $c_j^{(\upsilon)}$ είναι άρρητος, επειδή όλοι οι άλλοι όροι στο άθροισμα για τον $c_j^{(\upsilon)}$ είναι ρητοί. Έπεται από την επαγωγική υπόθεση ότι η ακολουθία $\big(p_\upsilon(k)\big)_{k\in\mathbb{N}_0}$ είναι ισοκατανεμημένη mod 1, ως προς το μέτρο Lebesgue. Όμως από την (30),

$$p(km+\upsilon) \pmod 1 = p_\upsilon(k) \pmod 1$$

για κάθε $k \in \mathbb{N}_0$, επειδή $k^d m^d a_d \in \mathbb{Z}\ \forall\, k \in \mathbb{N}_0$, λόγω της επιλογής του m. Έπεται λοιπόν ότι η ακολουθία $\big(p(km+\upsilon)\big)_{k\in\mathbb{N}_0}$ είναι ισοκατανεμημένη mod 1, για κάθε $\upsilon \in \{0, 1, \ldots, m-1\}$. Από αυτό έπεται άμεσα η ισοκατανομή της ακολουθίας $\big(p(n)\big)_{n\in\mathbb{N}}$. Πράγματι, αν $f \in C(\mathbb{T})$, τότε

$$\frac{1}{N}\sum_{n=0}^{N-1} f(p(n)) = \sum_{\upsilon=0}^{m-1}\frac{1}{N}\sum_{k=0}^{\lfloor N/m\rfloor} f(p(km+\upsilon)) - \frac{1}{N}\sum_{n=N}^{m\lfloor N/m\rfloor+m-1} f(p(n)).$$

Ο δεύτερος όρος κυριαρχείται από

$$\frac{1}{N}\sum_{n=N}^{m\lfloor N/m\rfloor+m-1} f(p(n)) \leq \frac{m\lfloor N/m\rfloor + m - N}{N}\sup_{t\in\mathbb{T}}|f(t)| \leq \frac{m}{N}\sup_{t\in\mathbb{T}}|f(t)| \xrightarrow{(N\to\infty)} 0,$$

[24] Αν $\epsilon > 0$ και $\delta > 0$ τέτοιο ώστε $|f(t)-f(s)| < \epsilon$ για $t, s \in \mathbb{T}$ με $d_{\mathbb{T}}(t,s) < \delta$, όπου $d_{\mathbb{T}}$ η μετρική $d_{\mathbb{T}}(t,s) = \min\{|t-s|, 1-|t-s|\}$ στον $\mathbb{T}$, τότε $d_{\mathbb{T}^d}((t_1,\ldots,t_d),(s_1,\ldots,s_d)) < \delta \Rightarrow d_{\mathbb{T}}(t_d,s_d) < \delta \Rightarrow |F(t_1,\ldots,t_d) - F(s_1,\ldots,s_d)| = |f(t_d)-f(s_d)| < \epsilon$, όπου $d_{\mathbb{T}^d}$ η μετρική $d_{\mathbb{T}^d}((t_1,\ldots,t_d),(s_1,\ldots,s_d)) = \max_{i\in\{1,\ldots,d\}} d_{\mathbb{T}}(t_i,s_i)$ στον $\mathbb{T}^d$, που είναι μία μετρική γινόμενο συμβατή με την τοπολογία γινόμενο του $\mathbb{T}^d$.

ενώ για τον πρώτο όρο έχει κανείς ότι

$$\sum_{\upsilon=0}^{m-1}\frac{1}{N}\sum_{k=0}^{\lfloor N/m\rfloor} f(p(km+\upsilon)) = \sum_{\upsilon=0}^{m-1}\frac{\lfloor N/m\rfloor}{N}\frac{1}{\lfloor N/m\rfloor}\sum_{k=0}^{\lfloor N/m\rfloor} f(p(km+\upsilon)) \xrightarrow{(N\to\infty)} \sum_{\upsilon=0}^{m-1}\frac{1}{m}\int f\,d\lambda = \int f\,d\lambda,$$

από την ισοκατανομή mod 1 των ακολουθιών $\big(p(km+\upsilon)\big)_{k\in\mathbb{N}_0}$, $\upsilon\in\{0,1,\dots,m-1\}$. Έπεται ότι

$$\frac{1}{N}\sum_{n=0}^{N-1} f(p(n)) \xrightarrow{(N\to\infty)} \int f\,d\lambda,$$

και άρα και

$$\frac{1}{N}\sum_{n=1}^{N} f(p(n)) \xrightarrow{(N\to\infty)} \int f\,d\lambda,$$

και αφού αυτό ισχύει για αυθαίρετη $f\in C(\mathbb{T})$, η ακολουθία $\big(p(n)\big)_{n\in\mathbb{N}}$ είναι ισοκατανεμημένη mod 1, ως προς το μέτρο Lebesgue. $\square$.

2η Απόδειξη. Καταρχήν υποθέτουμε ότι ο συντελεστής του μεγιστοβάθμιου όρου a_d είναι άρρητος. Έστω $a\in\mathbb{R}$, που θα προσδιορισθεί παρακάτω. Θεωρούμε το μονοσήμαντα εργοδικό σύστημα $(\mathbb{T}^d, S)$ όπου

$$S(t_1,t_2,\dots,t_d) = (t_1+a, t_1+t_2,\dots,t_{d-1}+t_d)\ (\text{mod }1).$$

Έχουμε δει ότι το σύστημα αυτό περιγράφεται ως

$$S(t_1,\dots,t_d) = \tilde{S}\begin{bmatrix} t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} + \begin{bmatrix} a\\ 0\\ \vdots\\ 0\end{bmatrix}\ (\text{mod }1),$$

όπου

$$\tilde{S} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & & \ddots & \\ & & & 1 & 1\end{bmatrix}_{(d\times d)}.$$

Περιγράφεται όμως και από το σύστημα $(\{a\}\times\mathbb{T}^d, \tilde{\tilde{S}})$ με

$$\tilde{\tilde{S}}\begin{bmatrix} a\\ t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & & \ddots & \\ & & & 1 & 1\end{bmatrix}_{(d+1)\times(d+1)}\begin{bmatrix} a\\ t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} = \begin{bmatrix} a\\ a+t_1\\ t_1+t_2\\ \vdots\\ t_{d-1}+t_d\end{bmatrix}.$$

Επομένως,

$$\tilde{\tilde{S}}^n\begin{bmatrix} a\\ t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ n & 1 & & & & \\ \binom{n}{2} & n & 1 & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \binom{n}{d} & \binom{n}{d-1} & \binom{n}{d-2} & \dots & n & 1\end{bmatrix}\begin{bmatrix} a\\ t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} = \begin{bmatrix} a\\ na+t_1\\ \binom{n}{2}a+nt_1+t_2\\ \vdots\\ \binom{n}{d}a+\binom{n}{d-1}t_1+\dots+nt_{d-1}+t_d\end{bmatrix},$$

από το προηγούμενο Λήμμα, τουλάχιστον για $n\geq d$. Έπεται ότι

$$S^n\begin{bmatrix} t_1\\ t_2\\ \vdots\\ t_d\end{bmatrix} = \begin{bmatrix} na+t_1\\ \binom{n}{2}a+nt_1+t_2\\ \vdots\\ \binom{n}{d}a+\binom{n}{d-1}t_1+\dots+nt_{d-1}+t_d\end{bmatrix}\ (\text{mod }1)$$

για $n \geq d$ και $(t_1, \ldots, t_d) \in \mathbb{T}^d$. Πράγματι, αν θεωρήσουμε την απεικόνιση $\pi \colon \{a\} \times \mathbb{R}^d \to \mathbb{R}^d$ που δίνεται από την $\pi(a, t_1, \ldots, t_d) = (t_1, \ldots, t_d)$, τότε η π είναι αντιστρέψιμη ($\pi^{-1}(t_1, \ldots, t_d) = (a, t_1, \ldots, t_d)$) και επιπλέον

$$\pi \circ \tilde{\tilde{S}} = S' \circ \pi$$

όπου $S' \colon \mathbb{R}^d \to \mathbb{R}^d$ η απεικόνιση

$$S'(x_1, \ldots, x_d) = (x_1 + a, x_1 + x_2, \ldots, x_{d-1} + x_d).$$

Οπότε

$$S' = \pi \circ \tilde{\tilde{S}} \circ \pi^{-1}$$

και άρα

$$S'^n = \pi \circ \tilde{\tilde{S}}^n \circ \pi^{-1}, \qquad \forall\, n \in \mathbb{N}.$$

Η απεικόνιση S' επάγει από την άλλη την S αφού

$$S((t_1, \ldots, t_d) + \mathbb{Z}^d) = S'(t_1, \ldots, t_d) + \mathbb{Z}^d$$

και επομένως

$$S^n((t_1, \ldots, t_d) + \mathbb{Z}^d) = S^{n-1}(S((t_1, \ldots, t_d) + \mathbb{Z}^d)) = S^{n-1}(S'(t_1, \ldots, t_d) + \mathbb{Z}^d) = \cdots = (S')^n(t_1, \ldots, t_d) + \mathbb{Z}^d$$

ή ισοδύναμα

$$S^n((t_1, \ldots, t_d)(\text{mod } 1)) = S'^n(t_1, \ldots, t_d)(\text{mod } 1).$$

Επιλέγουμε τώρα $a = d!\, a_d$. Τότε $a \in \mathbb{R} \setminus \mathbb{Q}$ και κατόπιν επιλέγουμε $(x_1, \ldots, x_d) \in \mathbb{R}^d$ έτσι ώστε

$$p(n) = \binom{n}{d} a + \binom{n}{d-1} x_1 + \cdots + n x_{d-1} + x_d \qquad \forall\, n \in \mathbb{N}. \tag{31}$$

(Η επιλογή των a και $(x_1, \ldots, x_d)$ είναι ανεξάρτητη του n.) Αυτό είναι εφικτό, αφού το δεξί μέλος της (31) είναι πολυώνυμο βαθμού d ως προς n:

$$\begin{aligned} &\binom{n}{d} a + \binom{n}{d-1} x_1 + \cdots + n x_{d-1} + x_d \\ &\qquad = \frac{n(n-1)\cdots(n-d+1)}{d!} d!\, a_d + \frac{n(n-1)\cdots(n-d+2)}{(d-1)!} x_1 + \cdots + n x_{d-1} + x_d \end{aligned}$$

και ο πρώτος όρος είναι πολυώνυμο βαθμού d ως προς n, ο δεύτερος πολυώνυμο βαθμού $d-1$ κ.ο.κ., ο προτελευταίος βαθμού 1 και ο τελευταίος βαθμού 0 ως προς n. Άρα συνολικά το δεξί μέλος είναι πολυώνυμο βαθμού d ως προς n, όπως και το αριστερό μέλος της (31). Ο συντελεστής του n^d και στα δύο μέλη είναι a_d. Επομένως, τα $x_1, \ldots, x_d$ προκύπτουν εξισώνοντας συντελεστές όρων ίδιου βαθμού στα δύο πολυώνυμα. Με την επιλογή του a εξισώσαμε ήδη τους συντελεστές των μεγιστοβάθμιων όρων n^d. Η εξίσωση των συντελεστών των υπολοίπων d όρων $n^{d-1}, n^{d-2}, \ldots, n, 1$ δίνει d εξισώσεις από όπου προκύπτουν τα $x_1, \ldots, x_d$ (γραμμικό σύστημα ως προς $x_1, \ldots, x_d$ με d- εξισώσεις). Μάλιστα, ο όρος n^{d-1} εμφανίζεται μόνο στους όρους

$$n(n-1)\cdots(n-d+1)a_d + \frac{n(n-1)\cdots(n-d+2)}{(d-1)!} x_1,$$

άρα εξισώνοντας τον συντελεστή του n^{d-1}, στο δεξί μέλος της (31), με το a_{d-1} προκύπτει μια εξίσωση με άγνωστο μόνο το x_1, γραμμική ως προς x_1 και με συντελεστή του x_1 το $1/(d-1)! \neq 0$. Έτσι βρίσκουμε το x_1. Ο όρος n^{d-2} εμφανίζεται στους 3 πρώτους όρους, του δεξιού μέλους της (31),

$$n(n-1)\cdots(n-d+1)a_d + \frac{n(n-1)\cdots(n-d+2)}{(d-1)!} x_1 + \frac{n(n-1)\cdots(n-d+3)}{(d-2)!} x_2$$

και εξισώνοντας τον συντελεστή του n^{d-2} με το a_{d-2} και γνωρίζοντας πλέον το x_1, προκύπτει πρωτοβάθμια εξίσωση ως προς x_2 με συντελεστή του x_2 το $1/(d-2)! \neq 0$. Έτσι βρίσκουμε το x_2 κ.ο.κ.

Η παραπάνω διαδικασία δίνει ένα $(x_1, \dots, x_d) \in \mathbb{R}^d$ που ικανοποιεί την (31). Θέτοντας $t_i := x_i - \lfloor x_i \rfloor$, $i \in \{1, \dots, d\}$, και επειδή οι συντελεστές $\binom{n}{d-i}$ των x_i στο δεξί μέλος της (31) είναι όλοι ακέραιοι, προκύπτει ότι το δεξί μέλος της (31) mod 1 ισούται με

$$\binom{n}{d}a + \binom{n}{d-1}x_1 + \cdots + nx_{d-1} + x_d \ (\mathrm{mod}\ 1) = \binom{n}{d}a + \binom{n}{d-1}t_1 + \cdots + nt_{d-1} + t_d \ (\mathrm{mod}\ 1)$$

και άρα τελικά και

$$p(n) \ (\mathrm{mod}\ 1) = \binom{n}{d}a + \binom{n}{d-1}t_1 + \cdots + nt_{d-1} + t_d \ (\mathrm{mod}\ 1)$$

για κάθε $n \in \mathbb{N}$ με $(t_1, \dots, t_d) \in \mathbb{T}^d$. Δηλαδή, ισοδύναμα, η (31) γράφεται και ως

$$p(n) + \mathbb{Z}^d = \binom{n}{d}a + \binom{n}{d-1}x_1 + \cdots + nx_{d-1} + x_d + \mathbb{Z}^d \qquad \forall\, n \in \mathbb{N}.$$

Αφού $a = d!\, a_d \in \mathbb{R} \setminus \mathbb{Q}$, το σύστημα $(\mathbb{T}^d, S)$ είναι μονοσήμαντα εργοδικό και άρα κάθε τροχιά είναι ισοκατανεμημένη ως προς το μοναδικό αναλλοίωτο μέτρο λ_d (= το μέτρο Lebesgue στον $\mathbb{T}^d$). Άρα και η τροχιά του επιλεχθέντος $(t_1, \dots, t_d)$ είναι ισοκατανεμημένη στον $\mathbb{T}^d$ ως προς λ_d. Αυτό έχει ως συνέπεια και η ακολουθία που προκύπτει παίρνοντας την τελευταία (d-οστή) συντεταγμένη κάθε σημείου της τροχιάς $(S^n(t_1, \dots, t_d))_{n \in \mathbb{N}_0}$ να είναι ισοκατανεμημένη. Πράγματι, αν $\pi_d \colon \mathbb{T}^d \to \mathbb{T}$ είναι προβολή

$$\pi(s_1, \dots, s_d) = s_d \in \mathbb{T},$$

τότε αυτή η συνάρτηση είναι συνεχής. Αν $f \in C(\mathbb{T})$, τότε η συνάρτηση $g = f \circ \pi_d \colon \mathbb{T}^d \to \mathbb{C}$, δηλαδή η $g(s_1, \dots, s_d) = f(s_d)$ ανήκει στον $C(\mathbb{T}^d)$. Εφόσον η $(S^n(t_1, \dots, t_d))_{n \in \mathbb{N}_0}$ είναι ισοκατανεμημένη, θα πρέπει

$$\begin{aligned} & \frac{1}{n}\sum_{k=0}^{n-1} g\big(S^k(t_1, \dots, t_d)\big) \to \int_{\mathbb{T}^d} g\, d\lambda_d \\ \Leftrightarrow \quad & \frac{1}{n}\sum_{k=0}^{n-1} f\big(\pi_d\big(S^k(t_1, \dots, t_d)\big)\big) \to \int_{\mathbb{T}} \cdots \int_{\mathbb{T}} f(s_d)\, ds_1 \dots ds_d = \int_{\mathbb{T}} f d\, \lambda_1. \end{aligned}$$

Αφού αυτό ισχύει για κάθε $f \in C(\mathbb{T})$, η ακολουθία $(\pi_d(S^k(t_1, \dots, t_d)))_{k \in \mathbb{N}_0}$ είναι ισοκατανεμημένη (ως προς το μέτρο Lebesgue) στον $\mathbb{T}^d$. Όμως

$$\pi_d(S^n(t_1, \dots, t_d)) = \binom{n}{d}a + \binom{n}{d-1}t_1 + \cdots + nt_{d-1} + t_d \ (\mathrm{mod}\ 1) = p(n) \ (\mathrm{mod}\ 1)$$

για $n \geq d$, και έπεται ότι η ακολουθία $(p(n))_{n \in \mathbb{N}_0}$ είναι Lebesgue-ισοκατανεμημένη $(\mathrm{mod}\ 1)$.

Η απόδειξη για την γενική περίπτωση στην οποία ο συντελεστής του μεγιστοβάθμιου όρου δεν είναι κατ' ανάγκην άρρητος, γίνεται όπως και στην 1η Απόδειξη του Θεωρήματος. □

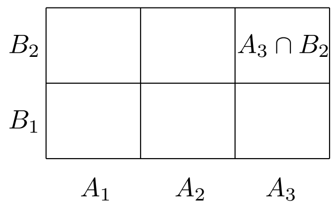
4 Μετροθεωρητική Εντροπία

Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Με τον όρο *μετρήσιμη διαμέριση* θα εννοείται μία διαμέριση ξ του X, πεπερασμένη ή άπειρη αριθμήσιμη, που αποτελείται από μετρήσιμα σύνολα, δηλαδή $\xi = \{A_1, \dots, A_n\}$, $n \in \mathbb{N}$, ή $\xi = \{A_1, A_2, \dots\}$, με $A_i \in \mathcal{B}$ για κάθε i. Ενίοτε το επίθετο μετρήσιμη θα παραλείπεται, θα εννοείται όμως πάντα ότι τα στοιχεία μιας διαμέρισης είναι μετρήσιμα σύνολα, δηλαδή στοιχεία της σ-άλγεβρας $\mathcal{B}$, δηλαδή θα εννοείται ότι $\xi \subset \mathcal{B}$ για κάθε διαμέριση ξ.

Τα στοιχεία της διαμέρισης καλούνται *άτομα*. Για μια διαμέριση ξ, $\sigma(\xi)$ είναι η μικρότερη σ-άλγεβρα που περιέχει όλα τα άτομα της διαμέρισης. Αν η ξ είναι πεπερασμένη, η $\sigma(\xi)$ αποτελείται από όλες τις (αναγκαστικά πεπερασμένες) ενώσεις στοιχείων της διαμέρισης και τα στοιχεία της διαμέρισης είναι άτομα της σ-άλγεβρας $\sigma(\xi)$.

Αν ξ και η διαμερίσεις λέμε ότι η η είναι *εκλέπτυνση* της ξ και γράφουμε $\xi \leqslant \eta$, αν κάθε στοιχείο της ξ γράφεται σαν ένωση στοιχείων της η. Ισοδύναμα, $\xi \leqslant \eta$ ανν για κάθε $A \in \xi$ και $B \in \eta$, ή $A \cap B = \varnothing$ ή $B \subset A$.

Αν ξ και η διαμερίσεις του X, η κοινή τους εκλέπτυνση είναι η διαμέριση $\xi \vee \eta$ που αποτελείται από όλες τις τομές $A \cap B$, $A \in \xi$, $B \in \eta$.



Αν για σ-άλγεβρες $\mathcal{A}_1, \mathcal{A}_2$ συμβολίσουμε με $\mathcal{A}_1 \vee \mathcal{A}_2$ την μικρότερη σ-άλγεβρα που περιέχει την ένωση τους $\mathcal{A}_1 \cup \mathcal{A}_2$, τότε
$$\sigma(\xi \vee \eta) = \sigma(\xi) \vee \sigma(\eta).$$

Παρατηρήσεις. (1) Αν $\xi \leqslant \eta$, τότε $\xi \vee \eta = \eta$.

(2) Επίσης αν $\xi \subset \mathcal{B}$ μετρήσιμη διαμέριση, τότε $T^{-1}\xi = \{T^{-1}(A) \colon A \in \xi\}$ είναι επίσης μετρήσιμη διαμέριση.

Πρόταση 4.0.1 Έστω ξ, ζ, η και $\xi_1, \xi_2, \eta_1, \eta_2$ διαμερίσεις.

(1) $\xi \vee \eta = \eta \vee \xi$.

(2) $(\xi \vee \eta) \vee \zeta = \xi \vee (\eta \vee \zeta)$ και άρα μπορεί κανείς να γράφει $\xi \vee \zeta \vee \eta$.

(3) $T^{-1}\xi \vee T^{-1}\eta = T^{-1}(\xi \vee \eta)$.

(4) $\xi \leqslant \eta \ \Rightarrow\ T^{-1}\xi \leqslant T^{-1}\eta$.

(5) Αν $\xi_1 \leqslant \xi_2$ και $\eta_1 \leqslant \eta_2$, τότε $\xi_1 \vee \eta_1 \leqslant \xi_2 \vee \eta_2$.

Απόδειξη. (1) Προφανώς $\xi \vee \eta = \{A \cap B \colon A \in \xi,\ B \in \eta\} = \{B \cap A \colon A \in \xi,\ B \in \eta\} = \eta \vee \xi$.

(2) Εξ' ορισμού,
$$(\xi \vee \eta) \vee \zeta = \{D \cap C \colon D \in \xi \vee \eta,\ C \in \zeta\} = \{A \cap B \cap C \colon A \in \xi,\ B \in \eta,\ C \in \zeta\},$$
αφού το γενικό στοιχείο της $\xi \vee \eta$ είναι της μορφής $D = A \cap B$ με $A \in \xi$ και $B \in \eta$. Όμοια
$$\xi \vee (\eta \vee \zeta) = \{A \cap D \colon A \in \xi,\ D \in \eta \vee \zeta\} = \{A \cap B \cap C \colon A \in \xi,\ B \in \eta,\ C \in \zeta\}$$
επίσης, αφού το γενικό στοιχείο της $\eta \vee \zeta$ είναι της μορφής $B \cap C$ με $B \in \eta$ και $C \in \zeta$.

(3) Επειδή οι αντίστροφες εικόνες διατηρούν τις συνολοθεωρητικές πράξεις και συγκεκριμένα επειδή $T^{-1}(A) \cap T^{-1}(B) = T^{-1}(A \cap B)$ για αυθαίρετα σύνολα $A, B \subset X$, έπεται ότι
$$\begin{aligned} T^{-1}\xi \vee T^{-1}\eta &= \big\{C \cap D \colon C \in T^{-1}\xi,\ D \in T^{-1}\eta\big\} = \big\{T^{-1}(A) \cap T^{-1}(B) \colon A \in \xi,\ B \in \eta\big\} \\ &= \big\{T^{-1}(A \cap B) \colon A \in \xi,\ B \in \eta\big\} = \big\{T^{-1}(C) \colon C \in \xi \vee \eta\big\} = T^{-1}(\xi \vee \eta). \end{aligned}$$

(4) Αν $\xi \leqslant \eta$, τότε για κάθε $A \in \xi$ και $B \in \eta$, ή $A \cap B = \varnothing$ ή $B \subset A$. Τότε όμως, για κάθε $A \in \xi$ και $B \in \eta$, ή $T^{-1}(A) \cap T^{-1}(B) = T^{-1}(A \cap B) = \varnothing$, αν $A \cap B = \varnothing$, ή $T^{-1}(B) \subset T^{-1}(A)$, αν $B \subset A$. Αυτό δείχνει ότι $T^{-1}\xi \leqslant T^{-1}\eta$.

(5) Έστω $A_i \in \xi_i$ και $B_i \in \eta_i$ για $i \in \{1, 2\}$. Τότε, επειδή $\xi_1 \leqslant \xi_2$, ή $A_1 \cap A_2 = \varnothing$ ή $A_2 \subset A_1$. Όμοια, επειδή $\eta_1 \leqslant \eta_2$, ή $B_1 \cap B_2 = \varnothing$ ή $B_2 \subset B_1$. Επομένως, για τα $A_1 \cap B_1 \in \xi_1 \vee \eta_1$ και $A_2 \cap B_2 \in \xi_2 \vee \eta_2$ ισχύει ότι ή $(A_1 \cap B_1) \cap (A_2 \cap B_2) = \varnothing$, αν $A_1 \cap A_2 = \varnothing$ ή $B_1 \cap B_2 = \varnothing$, ή $(A_2 \cap B_2) \subset (A_1 \cap B_1)$, αν $A_1 \cap A_2 \neq \varnothing$ και $B_1 \cap B_2 \neq \varnothing$, οπότε $A_2 \subset A_1$ και $B_2 \subset B_1$. Δείξαμε δηλαδή ότι για τα γενικά στοιχεία $A_1 \cap B_1$ της $\xi_1 \vee \eta_1$ και $A_2 \cap B_2$ της $\xi_2 \vee \eta_2$ ισχύει ότι ή $(A_1 \cap B_1) \cap (A_2 \cap B_2) = \varnothing$ ή $(A_2 \cap B_2) \subset (A_1 \cap B_1)$. Έπεται ότι $\xi_1 \vee \eta_1 \leqslant \xi_2 \vee \eta_2$. $\square$

4.1 Εντροπία διαμέρισης

Για $n \in \mathbb{N}$ θέτουμε $\Sigma_n = \{(p_1, \dots, p_n) \in [0,1]^n \colon \sum_{i=1}^n p_i = 1\}$, $\Sigma = \bigcup_{n \geq 1} \Sigma_n$ και $H \colon \Sigma \to [0, \infty)$ με τύπο

$$H(p_1, \dots, p_n) = -\sum_{i=1}^{n} p_i \ln p_i.$$

Η συνάρτηση H ορίζεται επίσης και για άπειρα διανύσματα $(p_i)_{i \in \mathbb{N}}$ με $p_i \geq 0 \ \forall\, i \in \mathbb{N}$ και $\sum_{i \geq 1} p_i = 1$:

$$H(p_1, p_2, \dots) = -\sum_{i \geq 1} p_i \ln p_i \in [0, +\infty].$$

Ορισμός 4.1.1 Αν $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\xi \subset \mathcal{B}$ μετρήσιμη διαμέριση με $\xi = \{A_1, A_2, \dots\}$, πεπερασμένη ή άπειρη, η εντροπία της ξ ορίζεται ως

$$H_\mu(\xi) = H(\mu(A_1), \mu(A_2), \dots) = -\sum_{A \in \xi} \mu(A) \ln \mu(A).$$

Επίσης, αν $\eta \subset \mathcal{B}$ είναι μια άλλη διαμέριση, η δεσμευμένη εντροπία της ξ δεδομένης της η είναι

$$\begin{aligned} H_\mu(\xi \mid \eta) &= \sum_{B \in \eta} \mu(B)\, H\left(\frac{\mu(A_1 \cap B)}{\mu(B)}, \frac{\mu(A_2 \cap B)}{\mu(B)}, \dots\right) \\ &= \sum_{B \in \eta} \mu(B)\, H_{\mu(B)^{-1}\mu|_B}(\xi) \\ &= -\sum_{A \in \xi} \sum_{B \in \eta} \mu(A \cap B) \ln \frac{\mu(A \cap B)}{\mu(B)}. \end{aligned}$$

Παρατηρεί κανείς εδώ ότι τα αθροίσματα μπορούν να περιοριστούν στα $B \in \eta$ με $\mu(B) > 0$, αφού για τα $B \in \eta$ με $\mu(B) = 0$ οι αντίστοιχοι όροι είναι μηδέν. Επίσης, για $B \in \mathcal{B}$ με $\mu(B) > 0$, $\mu(B)^{-1}\mu|_B$ συμβολίζει το μέτρο πιθανότητας $\mu(B)^{-1}\mu|_B(A) = \mu(A \cap B)/\mu(B)$, $A \in \mathcal{B}$.

Πρόταση 4.1.2 Ισχύουν τα εξής.

(1) $H(p_1, p_2, \dots, p_n) \geq 0$ με ισότητα ανν $p_i = 1$ για κάποιο i.

(2) Η $H|_{\Sigma_n}$ είναι συνεχής, συμμετρική συνάρτηση (αναλλοίωτη ως προς τις μεταθέσεις των p_i) και έχει μέγιστο ανν $(p_1, \dots, p_n) = (1/n, \dots, 1/n)$.

(3) $H(p_1, \dots, p_n, 0) = H(p_1, \dots, p_n) \quad \forall\, (p_1, \dots, p_n) \in \Sigma_n\ \forall\, n \in \mathbb{N}$.

(4) Για οποιεσδήποτε πεπερασμένες διαμερίσεις $\xi \subset \mathcal{B}$ και $\eta \subset \mathcal{B}$ ενός χώρου πιθανότητας $(X, \mathcal{B}, \mu)$, ισχύει ότι $H_\mu(\xi \vee \eta) = H_\mu(\eta) + H_\mu(\xi \mid \eta)$.

Θεώρημα 4.1.3 (Khintchine) Μια συνάρτηση η οποία ικανοποιεί τις ιδιότητες (1)–(4) της Πρότασης 4.1.2 δίνεται αναγκαστικά από τον τύπο

$$H(p_1, \dots, p_n) = -c \sum_{i=1}^{n} p_i \ln p_i,$$

για κάποια θετική σταθερά c.

Λήμμα 4.1.4 Η συνάρτηση $\phi \colon [0, \infty) \to \mathbb{R}$ με $\phi(x) = x \ln x$ είναι γνήσια κυρτή, δηλαδή

$$\phi(\lambda x + (1-\lambda) y) \leq \lambda \phi(x) + (1-\lambda)\phi(y) \qquad \forall\, \lambda \in [0,1],\ \forall\, x, y \in [0, \infty),$$

με ισότητα μόνο αν $\lambda \in \{0, 1\}$ ή $x = y$.

Απόδειξη. Για $x > 0$ έχει κανείς ότι

$$\phi'(x) = 1 + \ln x$$
$$\phi''(x) = \frac{1}{x} > 0.$$

Για $\lambda \in (0,1),\ x \neq y,$

$$\phi(y) - \phi(\lambda x + (1-\lambda)y) = \phi'(t)\lambda(y-x)$$

για κάποιο t γνήσια μεταξύ των $y, \lambda x + (1-\lambda)y$ και

$$\phi(\lambda x + (1-\lambda)y) - \phi(x) = \phi'(s)(1-\lambda)(y-x)$$

για κάποιο s γνήσια μεταξύ των $x, \lambda x + (1-\lambda)y$. Υποθέτουμε ότι $y > x$. Τότε $t > \lambda x + (1-\lambda)y > s$ και επειδή $\phi'' > 0$ έχουμε ότι $\phi'(t) > \phi'(s)$. Άρα

$$\begin{aligned}\big[\phi(y) - \phi(\lambda x + (1-\lambda)y)\big](1-\lambda) = \phi'(t)\lambda(1-\lambda)(y-x) &> \phi'(s)\lambda(1-\lambda)(y-x) \\ &= \lambda\big[\phi(\lambda x + (1-\lambda)y) - \phi(x)\big].\end{aligned}$$

Επομένως

$$\phi(\lambda x + (1-\lambda)y) < \lambda\phi(x) + (1-\lambda)\phi(y),$$

για όλα τα $x, y \in [0,\infty)$ με $x \neq y$ και $\lambda \in (0,1)$. Η ανισότητα $\phi(\lambda x + (1-\lambda)y) \leq \lambda\phi(x) + (1-\lambda)\phi(y)$ ισχύει επίσης, ως ισότητα, όταν $x = y$ ή $\lambda \in \{0,1\}$ και άρα ισχύει τελικά για όλα τα $x, y \in [0,\infty)$ και $\lambda \in [0,1]$. □

Πόρισμα 4.1.5 Έστω $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ με $\sum_{i=1}^n \lambda_i = 1$ και $x_1, \dots, x_n \geq 0$. Τότε

$$\phi\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i \phi(x_i)$$

με ισότητα ανν όλα τα x_i που αντιστοιχούν σε μη-μηδενικά λ_i είναι ίσα μεταξύ τους.

Απόδειξη. Με επαγωγή στο n. Για $n = 2$ αποδείχθηκε στο Λήμμα 4.1.4. Έστω ότι ισχύει για κάποιο $n \in \mathbb{N}$ και έστω $x_1, \dots, x_{n+1} \geq 0$ και $\lambda_1, \lambda_2, \dots, \lambda_{n+1} \geq 0$ με $\sum_{i=1}^{n+1} \lambda_i = 1$. Έστω i_* τέτοιο ώστε

$$x_{i_*} = \max\{x_1, \dots, x_{n+1}\}.$$

Αν $\lambda_{i_*} \neq 1$,

$$\begin{aligned}\phi\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) &= \phi\left(\lambda_{i_*} x_{i_*} + \sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \frac{\lambda_i}{1-\lambda_{i_*}} x_i \,(1-\lambda_{i_*})\right) \\ &\leq \lambda_{i_*}\phi(x_{i_*}) + (1-\lambda_{i_*})\,\phi\left(\sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \frac{\lambda_i}{1-\lambda_{i_*}} x_i\right) \\ &\leq \lambda_{i_*}\phi(x_{i_*}) + (1-\lambda_{i_*}) \sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \frac{\lambda_i}{1-\lambda_{i_*}} \phi(x_i) \\ &= \sum_{i=1}^{n+1} \lambda_i \phi(x_i)\end{aligned}$$

λόγω της επαγωγικής υπόθεσης. Αν $\lambda_{i_*} = 1$ έχουμε ισότητα. Αν $\lambda_{i_*} \neq 1$, ισότητα έχουμε ανν και οι δύο παραπάνω ανισότητες είναι ισότητες. Η δεύτερη ανισότητα είναι ισότητα ανν όλα τα x_i με $i \in \{1, \dots, n+1\} \setminus \{i_*\}$ που αντιστοιχούν σε μη μηδενικά λ_i είναι ίσα μεταξύ τους, από την επαγωγική υπόθεση. Ισότητα στην πρώτη ανισότητα έχει κανείς είτε όταν $\lambda_{i_*} = 0$, οπότε τότε συνολικά έχουμε ισότητα ανν όλα τα x_i με $i \in \{1, \dots, n+1\}$ που αντιστοιχούν σε μη μηδενικά λ_i είναι ίσα μεταξύ τους, από την προηγούμενη πρόταση, είτε όταν

$$\begin{aligned}x_{i_*} = \sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \frac{\lambda_i}{1-\lambda_{i_*}} x_i \quad &\Leftrightarrow \quad (1-\lambda_{i_*}) x_{i_*} = \sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \lambda_i x_i \\ \Leftrightarrow \quad \sum_{i \in \{1,\dots,n+1\} \setminus \{i_*\}} \lambda_i (x_i - x_{i_*}) = 0 \quad &\Leftrightarrow \quad x_i = x_{i_*} \quad \text{ή} \quad \lambda_i = 0 \quad \forall\, i \in \{1, \dots, n+1\} \setminus \{i_*\},\end{aligned}$$

αφού, λόγω της επιλογής του i_*, $x_i - x_{i_*} \leq 0 \;\; \forall i \in \{1, \ldots, n+1\}$. □

Πόρισμα 4.1.6 Η H έχει μέγιστο στο Σ_n ανν $(p_1, \ldots, p_n) = (1/n, \ldots, 1/n)$.

Απόδειξη. Έχουμε ότι

$$\begin{aligned}
-\frac{1}{n}\ln n = \phi\left(\frac{1}{n}\right) &= \phi\left(\frac{1}{n}\sum_{i=1}^{n} p_i\right) \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\phi(p_i) = \frac{1}{n}\sum_{i=1}^{n} p_i \ln p_i \\
&= -\frac{1}{n}H(p_1, \ldots, p_n).
\end{aligned}$$

Άρα $\ln n \geq H(p_1, \ldots, p_n)$, με ισότητα ανν $p_1 = p_2 = \cdots = p_n$ και αφού $\sum_{i=1}^{n} p_i = 1$ πρέπει $p_i = 1/n$ για κάθε $i \in \{1, \ldots, n\}$. □

Πρόταση 4.1.7 (ιδιότητες εντροπίας και δεσμευμένης εντροπίας διαμέρισης) Έστω $\xi \subset \mathcal{B}$, $\eta \subset \mathcal{B}$ και $\zeta \subset \mathcal{B}$ διαμερίσεις του X με πεπερασμένη εντροπία: $H_\mu(\xi) < \infty$, $H_\mu(\eta) < \infty$, $H_\mu(\zeta) < \infty$. Ισχύουν τα εξής.

(1) $H_\mu(\xi \vee \eta) = H_\mu(\xi \mid \eta) + H_\mu(\eta)$.

(2) $H_\mu(\xi \mid \eta) \leq H_\mu(\xi)$.

(3) $H_\mu(\xi \vee \eta) \leq H_\mu(\xi) + H_\mu(\eta)$.

(4) $H_\mu(\xi \vee \eta \mid \zeta) = H_\mu(\xi \mid \eta \vee \zeta) + H_\mu(\eta \mid \zeta)$.

(5) $H_\mu(\xi \mid \zeta \vee \eta) \leq H_\mu(\xi \mid \zeta)$.

(6) $H_\mu(\xi \vee \eta \mid \zeta) \leq H_\mu(\xi \mid \zeta) + H_\mu(\eta \mid \zeta)$.

Απόδειξη. (1) Έχουμε

$$\begin{aligned}
H_\mu(\xi \vee \eta) &= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(A\cap B)\ln\mu(A\cap B) \\
&= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) - \sum_{A\in\xi}\sum_{B\in\eta}\mu(A\cap B)\ln\mu(B) \\
&= H_\mu(\xi \mid \eta) - \sum_{B\in\eta}\left(\sum_{A\in\xi}\mu(A\cap B)\right)\ln\mu(B) \\
&= H_\mu(\xi \mid \eta) - \sum_{B\in\eta}\mu(B)\ln\mu(B) \\
&= H_\mu(\xi \mid \eta) + H_\mu(\eta).
\end{aligned}$$

όπου η προτελευταία ισότητα ισχύει επειδή η ξ είναι διαμέριση.

(2) Έχει κανείς διαδοχικά

$$\begin{aligned}
H_\mu(\xi \mid \eta) &= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(B)\frac{\mu(A\cap B)}{\mu(B)}\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) \\
&= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(B)\,\phi\left(\frac{\mu(A\cap B)}{\mu(B)}\right) \\
&\leq -\sum_{A\in\xi}\phi\left(\sum_{B\in\eta}\mu(A\cap B)\right) \\
&= -\sum_{A\in\xi}\phi(\mu(A)) \\
&= H_\mu(\xi)
\end{aligned}$$

όπου η ανισότητα αιτιολογείται από την κυρτότητα της ϕ και η προτελευταία ισότητα από το γεγονός, πάλι, ότι η η αποτελεί διαμέριση.

(3) Άμεση συνέπεια των ιδιοτήτων (1) και (2).

(4) Έχει κανείς από την ιδιότητα (1) ότι

$$\begin{aligned} H_\mu(\xi \vee \eta \mid \zeta) &\overset{(1)}{=} H_\mu(\xi \vee \eta \vee \zeta) - H_\mu(\zeta) \\ &= H_\mu(\xi \vee \eta \vee \zeta) - H_\mu(\eta \vee \zeta) + H_\mu(\eta \vee \zeta) - H_\mu(\zeta) \\ &\overset{(1)}{=} H_\mu(\xi \mid \eta \vee \zeta) + H_\mu(\eta \mid \zeta). \end{aligned}$$

(5) Η ιδιότητα αυτή γενικεύει την (2) και για την απόδειξή της μπορεί να χρησιμοποιηθεί η (2) για τα μέτρα $\mu(C)^{-1}\mu|_C$, $C \in \zeta$, $\mu(C) > 0$. Συγκεκριμένα,

$$\begin{aligned} H_\mu(\xi \mid \zeta \vee \eta) &\overset{(4)}{=} H_\mu(\xi \vee \eta \mid \zeta) - H_\mu(\eta \mid \zeta) \\ &= \sum_{C \in \zeta} \mu(C) \left[H_{\mu(C)^{-1}\mu|_C}(\xi \vee \eta) - H_{\mu(C)^{-1}\mu|_C}(\eta) \right] \\ &= \sum_{C \in \zeta} \mu(C) H_{\mu(C)^{-1}\mu|_C}(\xi \mid \eta) \\ &\leq \sum_{C \in \zeta} \mu(C) H_{\mu(C)^{-1}\mu|_C}(\xi) \\ &= H_\mu(\xi \mid \zeta). \end{aligned}$$

(6) Άμεση συνέπεια των ιδιοτήτων (4) και (5). □

Πρόταση 4.1.8 (ιδιότητες μονοτονίας) Έστω ξ, η και ζ μετρήσιμες διαμερίσεις με πεπερασμένη εντροπία.

(1) Αν $\xi \leqslant \eta$, τότε $H_\mu(\xi) \leq H_\mu(\eta)$.

(2) Αν $\xi \leqslant \eta$, τότε $H_\mu(\xi \mid \zeta) \leq H_\mu(\eta \mid \zeta)$.

(3) Αν $\xi \leqslant \eta$, τότε $H_\mu(\zeta \mid \xi) \geq H_\mu(\zeta \mid \eta)$.

(4) $H_\mu(T^{-1}\xi \mid T^{-1}\eta) = H_\mu(\xi \mid \eta)$ και $H_\mu(T^{-1}\xi) = H_\mu(\xi)$.

Απόδειξη. (1) Αν $\xi \leqslant \eta$, τότε $\xi \vee \eta = \eta$. Άρα

$$H_\mu(\eta) = H_\mu(\xi \vee \eta) = H_\mu(\eta \mid \xi) + H_\mu(\xi) \geq H_\mu(\xi)$$

από την ιδιότητα (1) της Πρότασης 4.1.7 και το γεγονός ότι η εντροπία είναι μη αρνητική ποσότητα, εξ' ορισμού.

(2) Όμοια,

$$H_\mu(\eta \mid \zeta) = H_\mu(\eta \vee \xi \mid \zeta) = H_\mu(\eta \mid \xi \vee \zeta) + H_\mu(\xi \mid \zeta) \geq H_\mu(\xi \mid \zeta),$$

αυτή τη φορά από την ιδιότητα (4) της Πρότασης 4.1.7 (και το γεγονός ότι και η δεσμευμένη εντροπία είναι μη αρνητική ποσότητα, εξ' ορισμού).

(3) Τώρα

$$H_\mu(\zeta \mid \eta) = H_\mu(\zeta \mid \eta \vee \xi) \leq H_\mu(\zeta \mid \xi),$$

λόγω της ιδιότητας (5) της Πρότασης 4.1.7.

(4) Έχουμε

$$\begin{aligned}
H_\mu(T^{-1}\xi \mid T^{-1}\eta) &= -\sum_{A\in T^{-1}\xi}\ \sum_{B\in T^{-1}\eta} \mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right)\\
&= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(T^{-1}A\ \cap\ T^{-1}B)\ln\left(\frac{\mu(T^{-1}A\ \cap\ T^{-1}B)}{\mu(T^{-1}B)}\right)\\
&= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(T^{-1}(A\cap B))\ln\left(\frac{\mu(T^{-1}(A\cap B))}{\mu(T^{-1}B)}\right)\\
&= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right)\\
&= H_\mu(\xi\mid\eta).
\end{aligned}$$

Η δεύτερη σχέση αποδεικνύεται παρομοίως. □

Παρατήρηση. Παρατηρεί κανείς ότι για την τετριμμένη διαμέριση $\eta = \{X\}$, έχει κανείς ότι $H_\mu(\xi) = H_\mu(\xi \mid \eta) = H_\mu(\xi \mid \{X\})$, για οποιαδήποτε διαμέριση $\xi \subset \mathcal{B}$. Άρα όλες οι ιδιότητες παραπάνω που αναφέρονται σε αδέσμευτη εντροπία αποτελούν ειδική περίπτωση των αντίστοιχων ιδιοτήτων για την δεσμευμένη εντροπία. Για παράδειγμα, στην τελευταία ιδιότητα (4) της Πρότασης 4.1.8, η ιδιότητα $H_\mu(T^{-1}\xi) = H_\mu(\xi)$, είναι ειδική περίπτωση της ιδιότητας $H_\mu(T^{-1}\xi \mid T^{-1}\eta) = H_\mu(\xi \mid \eta)$, αφού για $\eta = \{X\}$ έχει κανείς ότι και $T^{-1}\eta = \{X\}$, και άρα

$$H_\mu(T^{-1}\xi) = H_\mu(T^{-1}\xi \mid \{X\}) = H_\mu(T^{-1}\xi \mid T^{-1}\{X\}) = H_\mu(\xi \mid \{X\}) = H_\mu(\xi).$$

Πρόταση 4.1.9 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\xi \subset \mathcal{B}$ και $\eta \subset \mathcal{B}$ διαμερίσεις του X με πεπερασμένη εντροπία.

(1) $H_\mu(\xi \mid \eta) = 0$ ανν $\eta \geqslant \xi \bmod \mu$.

(2) $H_\mu(\xi \mid \eta) = H_\mu(\xi)$ ανν οι ξ και η είναι ανεξάρτητες, δηλαδή αν $\mu(A\cap B) = \mu(A)\mu(B)$ για κάθε $A \in \xi$, $B \in \eta$.

Απόδειξη. (1) Αν $\xi \leqslant \eta \bmod \mu$ τότε για κάθε $A \in \xi$ και $B \in \eta$ είναι $\mu(A\cap B) = 0$ ή $\mu(A\cap B) = \mu(B)$. Επομένως

$$\begin{aligned}
H_\mu(\xi\mid\eta) &= -\sum_{A\in\xi}\sum_{B\in\eta}\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right)\\
&= -\sum_{\substack{A\in\xi,B\in\eta\\ \mu(A\cap B)>0}}\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right)\\
&= -\sum_{\substack{A\in\xi,B\in\eta\\ \mu(A\cap B)>0}}\mu(A\cap B)\ln 1 = 0.
\end{aligned}$$

Αντίστροφα, έστω ότι $H_\mu(\xi \mid \eta) = 0$. Επειδή για κάθε $A \in \xi$, $B \in \eta$ έχουμε

$$\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) \le 0,$$

πρέπει

$$\mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) = 0 \qquad \forall\ A \in \xi,\ B \in \eta.$$

Άρα πρέπει $\forall\ A \in \xi,\ B \in \eta$ να ισχύει $\mu(A\cap B) \in \{0, \mu(B)\}$, δηλαδή $\xi \leqslant \eta \bmod \mu$.

(2) Έστω ότι ξ και η είναι ανεξάρτητες διαμερίσεις. Τότε

$$\begin{aligned}
H_\mu(\xi \mid \eta) &= -\sum_{A\in\xi}\sum_{B\in\eta} \mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) \\
&= -\sum_{A\in\xi}\sum_{B\in\eta} \mu(A\cap B)\ln \mu(A) \\
&= -\sum_{A\in\xi} \mu(A)\ln \mu(A) \\
&= H_\mu(\xi),
\end{aligned}$$

η τρίτη ισότητα επειδή η η είναι διαμέριση.

Αντίστροφα, έστω ότι $H_\mu(\xi \mid \eta) = H_\mu(\xi)$. Τότε

$$\sum_{A\in\xi}\sum_{B\in\eta} \mu(A\cap B)\ln\left(\frac{\mu(A\cap B)}{\mu(B)}\right) = \sum_{A\in\xi} \mu(A)\ln\mu(A).$$

Δηλαδή,

$$\sum_{A\in\xi}\sum_{B\in\eta} \mu(B)\,\phi\left(\frac{\mu(A\cap B)}{\mu(B)}\right) = \sum_{A\in\xi} \phi(\mu(A))$$

ή ισοδύναμα

$$\sum_{A\in\xi}\left[\sum_{B\in\eta} \mu(B)\,\phi\left(\frac{\mu(A\cap B)}{\mu(B)}\right) - \phi(\mu(A))\right] = 0. \tag{32}$$

Όμως λόγω της κυρτότητας της ϕ και του ότι η η είναι διαμέριση, έχουμε ότι

$$\sum_{B\in\eta} \mu(B)\,\phi\left(\frac{\mu(A\cap B)}{\mu(B)}\right) \geq \phi\left(\sum_{B\in\eta}\mu(A\cap B)\right) = \phi(\mu(A)) \qquad \forall\, A\in\xi.$$

Επομένως όλοι οι όροι του αθροίσματος της σχέσης (32) είναι θετικοί και άρα κάθε όρος θα είναι μηδέν. Δηλαδή

$$\sum_{B\in\eta} \mu(B)\,\phi\left(\frac{\mu(A\cap B)}{\mu(B)}\right) = \phi(\mu(A)) \qquad \forall\, A\in\xi.$$

Από το Πόρισμα 4.1.5, η ισότητα στην ανισότητα Jensen ισχύει μόνο αν όλα τα $\mu(A\cap B)/\mu(B)$ που αντιστοιχούν σε μη μηδενικά $\mu(B)$ είναι ίσα μεταξύ τους. Επομένως υπάρχει c τέτοιο ώστε

$$\mu(B) = 0 \quad \text{ή} \quad \frac{\mu(A\cap B)}{\mu(B)} = c \qquad \forall\, A\in\xi,\ B\in\eta.$$

Τώρα προσθέτοντας ως προς $B\in\eta$ έχουμε

$$\mu(A) = \sum_{B\in\eta}\mu(A\cap B) = c\sum_{B\in\eta}\mu(B) = c,$$

αφού η η είναι διαμέριση, και άρα

$$\mu(B) = 0 \quad \text{ή} \quad \frac{\mu(A\cap B)}{\mu(B)} = \mu(A) \qquad \forall\, A\in\xi,\ B\in\eta.$$

Επομένως οι ξ και η είναι ανεξάρτητες διαμερίσεις, αφού για κάθε $A\in\xi$ και $B\in\eta$ ισχύει ότι $\mu(A\cap B) = \mu(A)\mu(B)$. $\square$

4.2 Εντροπία Συστήματος

Λήμμα 4.2.1 (Fekete) Αν $(a_n)_{n\in\mathbb{N}}$ ακολουθία πραγματικών αριθμών τέτοια ώστε $a_{n+m} \le a_n + a_m$ για κάθε $n, m \in \mathbb{N}$, $a_0 = 0$, τότε

$$\frac{a_n}{n} \to \inf_{k\in\mathbb{N}} \frac{a_k}{k}$$

Απόδειξη. Έστω $k \in \mathbb{N}$. Για τυχόν $n \in \mathbb{N}$ γράφουμε

$$n = km_n + \upsilon_n \qquad \upsilon_n \in \{0, \dots, k-1\}$$

και από υποπροσθετικότητα παίρνουμε

$$a_n \le a_{km_n} + a_{\upsilon_n} \le a_k + a_{k(m_n-1)} + a_{\upsilon_n} \le \cdots \le a_k m_n + \upsilon_n a_1.$$

Αν $a_k \ge 0$ και $a_1 \ge 0$ τότε

$$\frac{a_n}{n} \le \frac{a_k m_n}{km_n + \upsilon_n} + \frac{\upsilon_n a_1}{km_n + \upsilon_n} \le \frac{a_k m_n}{km_n} + \frac{\upsilon_n a_1}{km_n} \le \frac{a_k}{k} + \frac{a_1}{m_n}.$$

Άρα

$$\limsup_{n\to\infty} \frac{a_n}{n} \le \frac{a_k}{k},$$

επειδή $m_n \to \infty$ καθώς $n \to \infty$.

Αν $a_k < 0$ και $a_1 \ge 0$ τότε

$$\frac{a_n}{n} \le \frac{a_k m_n}{km_n + \upsilon_n} + \frac{\upsilon_n a_1}{km_n + \upsilon_n} < \frac{m_n a_k}{(m_n+1)k} + \frac{a_1}{m_n}$$

εφόσον $km_n + \upsilon_n < km_n + k$. Άρα πάλι

$$\limsup_{n\to\infty} \frac{a_n}{n} \le \frac{a_k}{k}.$$

Αν, τέλος, $a_k < 0$ και $a_1 < 0$ τότε

$$\frac{a_n}{n} \le \frac{m_n a_k}{km_n + \upsilon_n} < \frac{m_n a_k}{(m_n+1)k}$$

και άρα και πάλι

$$\limsup_{n\to\infty} \frac{a_n}{n} \le \frac{a_k}{k}.$$

Επομένως,

$$\limsup_{n\to\infty} \frac{a_n}{n} \le \inf_{k\in\mathbb{N}} \frac{a_k}{k}.$$

Όμως προφανώς

$$\liminf_{n\to\infty} \frac{a_n}{n} \ge \inf_{k\in\mathbb{N}} \frac{a_k}{k}$$

και άρα τελικά υπάρχει το όριο και είναι $\inf_{k\in\mathbb{N}} k^{-1} a_k$. □

Πόρισμα 4.2.2 Αν $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\xi \subset \mathcal{B}$ διαμέριση, τότε το όριο

$$\lim_{n\to\infty} \frac{1}{n} H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right)$$

υπάρχει.

Σημείωση. Τα στοιχεία της διαμέρισης $\bigvee_{k=0}^{n-1} T^{-k}\xi$ είναι σύνολα της μορφής $\bigcap_{i=0}^{n-1} T^{-i} A_{j_i}$, με $A_{j_0}, A_{j_1}, \dots, A_{j_{n-1}} \in \xi$.

Απόδειξη. Έστω
$$a_n = H_\mu\left(\bigvee_{k=0}^{n-1} T^{-k}\xi\right), \qquad n \in \mathbb{N}.$$
Τότε
$$\begin{aligned} a_{n+m} &= H_\mu\left(\bigvee_{k=0}^{n+m-1} T^{-k}\xi\right) = H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\xi \vee \bigvee_{k=m}^{n+m-1} T^{-k}\xi\right) \\ &\le H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\xi\right) + H_\mu\left(\bigvee_{k=m}^{n+m-1} T^{-k}\xi\right) \\ &= H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\xi\right) + H_\mu\left(T^{-m}\bigvee_{k=0}^{n-1} T^{-k}\xi\right) \\ &= H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\xi\right) + H_\mu\left(\bigvee_{k=0}^{n-1} T^{-k}\xi\right) \\ &= a_m + a_n \end{aligned}$$
όπου χρησιμοποιήθηκαν διαδοχικά οι ιδιότητες 4.0.1 (2), για την δεύτερη ισότητα, 4.1.7 (3), για την ανισότητα, 4.0.1 (3), για την τρίτη ισότητα και 4.1.8 (4), για την τέταρτη ισότητα. Από το Λήμμα 4.2.1 έπεται ότι το ζητούμενο όριο υπάρχει. □

Ορισμός 4.2.3 (εντροπία μετασχηματισμού) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\xi \subset \mathcal{B}$ μετρήσιμη διαμέριση. Η εντροπία του T ως προς την διαμέριση ξ είναι
$$h_\mu(T,\xi) = \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{k=0}^{n-1} T^{-k}\xi\right).$$

Εντροπία του συστήματος ορίζουμε
$$h_\mu(T) = \sup_{\substack{\xi \subset \mathcal{B} \text{ πεπερασμένη} \\ \text{διαμέριση}}} h_\mu(T,\xi) = \sup_{\substack{H_\mu(\xi)<\infty \\ \xi \subset \mathcal{B} \text{ αριθμήσιμη}}} h_\mu(T,\xi). \tag{33}$$

Λήμμα 4.2.4 Ισχύει ότι
$$H_\mu\left(\bigvee_{k=0}^{n-1} T^{-k}\xi\right) = H_\mu(\xi) + \sum_{k=1}^{n-1} H_\mu\left(\xi \;\middle|\; \bigvee_{j=1}^{k} T^{-j}\xi\right).$$

Απόδειξη. Με επαγωγή. Για $n = 1$ η ισότητα είναι ταυτολογία, αφού και τα δύο μέλη ισούνται με $H_\mu(\xi)$. Για τυχόν n έχουμε
$$\begin{aligned} H_\mu\left(\bigvee_{k=0}^{n} T^{-k}\xi\right) &= H_\mu\left(\xi \vee \bigvee_{k=1}^{n} T^{-k}\xi\right) \\ &= H_\mu\left(\xi \;\middle|\; \bigvee_{k=1}^{n} T^{-k}\xi\right) + H_\mu\left(\bigvee_{k=1}^{n} T^{-k}\xi\right) \\ &= H_\mu\left(\xi \;\middle|\; \bigvee_{k=1}^{n} T^{-k}\xi\right) + H_\mu\left(\bigvee_{k=0}^{n-1} T^{-k}\xi\right) \\ &= H_\mu\left(\xi \;\middle|\; \bigvee_{k=1}^{n} T^{-k}\xi\right) + H_\mu(\xi) + \sum_{k=1}^{n-1} H_\mu\left(\xi \;\middle|\; \bigvee_{j=1}^{k} T^{-j}\xi\right) \\ &= H_\mu(\xi) + \sum_{k=1}^{n} H_\mu\left(\xi \;\middle|\; \bigvee_{j=1}^{k} T^{-j}\xi\right), \end{aligned}$$

λόγω των ιδιοτήτων 4.0.1 (2), για την πρώτη ισότητα, 4.1.7 (1), για την δεύτερη ισότητα, 4.1.8 (4), για την τρίτη και της επαγωγικής υπόθεσης, για την τέταρτη ισότητα. □

Πόρισμα 4.2.5 Η ακολουθία

$$\frac{1}{n} H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right) \searrow h_\mu(T,\xi).$$

Απόδειξη. Επειδή για $k \leq n$,

$$\bigvee_{j=1}^{k} T^{-j}\xi \leqslant \bigvee_{j=1}^{n} T^{-j}\xi,$$

χρησιμοποιώντας τον τύπο του προηγούμενου Λήμματος και την ιδιότητα 4.1.8 (3) παίρνουμε ότι

$$H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right) = H_\mu(\xi) + \sum_{k=1}^{n-1} H_\mu \left(\xi \,\middle|\, \bigvee_{j=1}^{k} T^{-j}\xi \right) \geq n\, H_\mu \left(\xi \,\middle|\, \bigvee_{j=1}^{n} T^{-j}\xi \right). \tag{34}$$

Τώρα, χρησιμοποιώντας ξανά τις ιδιότητες 4.0.1 (2), 4.1.7 (1) και 4.1.8 (4), ακριβώς όπως στην απόδειξη του προηγούμενου Λήμματος, παίρνει κανείς ότι

$$\begin{aligned}
n\, H_\mu \left(\bigvee_{k=0}^{n} T^{-k}\xi \right) &= n\, H_\mu \left(\xi \vee \bigvee_{k=1}^{n} T^{-k}\xi \right) \\
&= n\, H_\mu \left(\bigvee_{k=1}^{n} T^{-k}\xi \right) + n\, H_\mu \left(\xi \,\middle|\, \bigvee_{k=1}^{n} T^{-k}\xi \right) \\
&= n\, H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right) + n\, H_\mu \left(\xi \,\middle|\, \bigvee_{k=1}^{n} T^{-k}\xi \right) \\
&\overset{(34)}{\leq} n\, H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right) + H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right) \\
&= (n+1)\, H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right).
\end{aligned}$$

Επομένως

$$\frac{1}{n+1} H_\mu \left(\bigvee_{k=0}^{n} T^{-k}\xi \right) \leq \frac{1}{n} H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}\xi \right).$$ □

Πρόταση 4.2.6 (Ιδιότητες εντροπίας συστήματος) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Έστω και $\xi \subset \mathcal{B}$ και $\eta \subset \mathcal{B}$ δύο μετρήσιμες διαμερίσεις με πεπερασμένη εντροπία. Τότε

(1) $h_\mu(T,\xi) \leq H_\mu(\xi)$.

(2) $h_\mu(T,\xi \vee \eta) \leq h_\mu(T,\xi) + h_\mu(T,\eta)$.

(3) Αν $\xi \leqslant \eta$ τότε $h_\mu(T,\xi) \leq h_\mu(T,\eta)$.

(4) $h_\mu(T,\xi) \leq h_\mu(T,\eta) + H_\mu(\xi \mid \eta)$.

(5) $h_\mu(T,\xi) = h_\mu(T,T^{-1}\xi)$.

(6) $h_\mu(T,\xi) = h_\mu\left(T, \bigvee_{i=0}^{n-1} T^{-i}\xi\right)$.

(7) Αν το σύστημα $(X, \mathcal{B}, \mu, T)$ είναι αντιστρέψιμο, τότε

$$h_\mu(T,\xi) = h_\mu \left(T, \bigvee_{i=-n}^{n} T^{i}\xi \right).$$

Απόδειξη. (1) Για κάθε $n \in \mathbb{N}$,

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) \le \sum_{i=0}^{n-1} H_\mu(T^{-i}\xi) = \sum_{i=0}^{n-1} H_\mu(\xi) = nH_\mu(\xi).$$

Άρα

$$\frac{1}{n}\,H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) \le H_\mu(\xi)$$

και αφήνοντας το n να τείνει στο άπειρο, παίρνουμε το ζητούμενο.

(2) Χρησιμοποιώντας διαδοχικά τις ιδιότητες 4.0.1 (3), 4.0.1 (1), 4.0.1 (2) και κατόπιν την 4.1.7 (3), έχουμε

$$\begin{aligned}
H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(\xi\vee\eta)\right) &= H_\mu\left(\bigvee_{i=0}^{n-1} (T^{-i}\xi\vee T^{-i}\eta)\right)\\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\vee\bigvee_{i=0}^{n-1} T^{-i}\eta\right)\\
&\le H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) + H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right).
\end{aligned}$$

Διαιρώντας με n και παίρνοντας όριο έχουμε το ζητούμενο.

(3) Χρησιμοποιώντας τις ιδιότητες 4.0.1 (4) και 4.0.1 (5), από την $\xi \leqslant \eta$ παίρνει κανείς ότι

$$\bigvee_{i=0}^{n-1} T^{-i}\xi \le \bigvee_{i=0}^{n-1} T^{-i}\eta.$$

Επομένως

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) \le H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right)$$

και διαιρώντας με n και παίρνοντας όριο έχουμε το ζητούμενο.

(4) Έχει κανείς διαδοχικά

$$\begin{aligned}
H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) &\le H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\vee\bigvee_{i=0}^{n-1} T^{-i}\eta\right) && \text{από 4.1.8 (1)}\\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) + H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi \,\middle|\, \bigvee_{i=0}^{n-1} T^{-i}\eta\right) && \text{από 4.1.7 (1)}\\
&\le H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) + \sum_{i=0}^{n-1} H_\mu\left(T^{-i}\xi \,\middle|\, \bigvee_{i=0}^{n-1} T^{-i}\eta\right) && \text{από 4.1.7 (6)}\\
&\le H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) + \sum_{i=0}^{n-1} H_\mu\left(T^{-i}\xi \mid T^{-i}\eta\right) && \text{από 4.1.8 (3)}\\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) + nH_\mu(\xi\mid\eta) && \text{από 4.1.8 (4).}
\end{aligned}$$

Άρα

$$\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) \le \frac{1}{n}\,H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\eta\right) + H_\mu(\xi|\eta)$$

και αφήνοντας το n να τείνει στο άπειρο παίρνουμε το ζητούμενο.

(5) Έχει κανείς ότι

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(T^{-1}\xi)\right) = H_\mu\left(T^{-1}\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right).$$

Διαιρώντας με n και παίρνοντας όριο έχουμε

$$h_\mu(T, T^{-1}\xi) = h_\mu(T,\xi).$$

(6) Αφού προφανώς $\zeta \vee \zeta = \zeta$ για κάθε διαμέριση ζ, έχει κανείς από τις ιδιότητες της Πρότασης 4.0.1 ότι

$$\begin{aligned}
H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right)\right) &= H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\left(\xi \vee T^{-1}\xi \vee \cdots \vee T^{-(n-1)}\xi\right)\right)\\
&= H_\mu\left(\bigvee_{k=0}^{m-1}\left(T^{-k}\xi \vee T^{-(k+1)}\xi \vee \cdots \vee T^{-(k+n-1)}\xi\right)\right)\\
&= H_\mu\Big(\xi \vee T^{-1}\xi \vee \cdots T^{-(n-1)}\xi\\
&\qquad \vee T^{-1}\xi \vee T^{-2}\xi \vee \cdots \vee T^{-n}\xi\\
&\qquad\qquad \vee \cdots \vee T^{-(m-1)}\xi \vee \cdots \vee T^{-m-n+2}\xi\Big)\\
&= H_\mu(\xi \vee T^{-1}\xi \vee \cdots \vee T^{-m-n+2}\xi).
\end{aligned}$$

Άρα

$$\begin{aligned}
h_\mu\left(T, \bigvee_{i=0}^{n-1} T^{-i}\xi\right) &= \lim_{m\to\infty}\frac{1}{m} H_\mu\left(\bigvee_{k=0}^{m-1} T^{-k}\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right)\right)\\
&= \lim_{m\to\infty}\frac{m+n-2}{m}\frac{1}{m+n-2} H_\mu\left(\bigvee_{k=0}^{m+n-2} T^{-k}\xi\right)\\
&= h_\mu(T,\xi).
\end{aligned}$$

(7) Για ένα αντιστρέψιμο σύστημα έχει κανείς ότι

$$\begin{aligned}
h_\mu\left(T, \bigvee_{i=-n}^{n} T^{i}\xi\right) &= h_\mu\left(T, \bigvee_{i=0}^{2n} T^{n-i}\xi\right)\\
&= h_\mu\left(T, T^n\bigvee_{i=0}^{2n} T^{-i}\xi\right)\\
&= h_\mu\left(T, \bigvee_{i=0}^{2n} T^{-i}\xi\right) \qquad \text{από (5)}\\
&= h_\mu(T,\xi) \qquad \text{από (6)},
\end{aligned}$$

όπου στην προτελευταία ισότητα, η ιδιότητα (5) εφαρμόστηκε διαδοχικά n φορές για την διαμέριση $\eta := T^n \bigvee_{i=0}^{2n} T^{-i}\xi$, για να προκύψει η $h_\mu(T,\eta) = h_\mu(T, T^{-n}\eta)$. □

Θεώρημα 4.2.7 Έστω $(X, \mathcal{B}, \mu, T)$ ένα σ.δ.μ.

(1) $h_\mu(T^k) = k h_\mu(T) \quad \forall\, k \in \mathbb{N}$.

(2) Όταν $(X, \mathcal{B}, \mu, T)$ είναι αντιστρέψιμο, τότε $h_\mu(T^k) = |k| h_\mu(T) \quad \forall\, k \in \mathbb{Z}$.

Απόδειξη. (1) Έστω $k \in \mathbb{N}$.

Ισχυρισμός. Αν $\xi \subset \mathcal{B}$ διαμέριση με πεπερασμένη εντροπία, τότε, για κάθε $k \in \mathbb{N}$,

$$h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\xi\right) = k h_\mu(T, \xi).$$

Πράγματι,

$$\begin{aligned}
H_\mu\left(\bigvee_{j=0}^{n-1} T^{-kj}\left(\bigvee_{i=0}^{k-1} T^{-i}\xi\right)\right) &= H_\mu\left(\bigvee_{j=0}^{n-1}\left(T^{-kj}\xi \vee T^{-(kj+1)}\xi \vee \cdots \vee T^{-(kj+k-1)}\xi\right)\right)\\
&= H_\mu(\xi \vee \cdots \vee T^{-(k-1)}\xi \vee T^{-k}\xi \vee {}^{-k-1}\xi \vee \cdots \vee T^{-2k-1}\xi \vee \cdots \vee T^{-(n-1)k-(k-1)}\xi)\\
&= H_\mu\left(\bigvee_{i=0}^{-nk+1} T^{-i}\xi\right)
\end{aligned}$$

και επομένως

$$\frac{1}{n} H_\mu\left(\bigvee_{j=0}^{n-1} T^{-kj}\left(\bigvee_{i=0}^{k-1} T^{-i}\xi\right)\right) = k\,\frac{1}{nk}\,H_\mu\left(\bigvee_{i=0}^{-nk+1} T^{-i}\xi\right).$$

Αφήνοντας το n να τείνει στο άπειρο παίρνουμε

$$h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\xi\right) = k h_\mu(T, \xi).$$

Τώρα

$$k h_\mu(T) = k \sup_\xi h_\mu(T, \xi) = \sup_\xi h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\xi\right) \le \sup_\eta h_\mu(T^k, \eta) = h_\mu(T^k).$$

Από την άλλη, για τυχούσα διαμέριση ξ έχουμε

$$h_\mu(T^k, \xi) \overset{4.2.6\,(3)}{\le} h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\xi\right) = k h_\mu(T, \xi).$$

Επομένως

$$h_\mu(T^k) = \sup_\xi h_\mu(T^k, \xi) \le k \sup_\xi h_\mu(T, \xi) = k h_\mu(T).$$

(2) Έστω αντιστρέψιμο σύστημα $(X, \mathcal{B}, \mu, T)$. Αρκεί να δειχτεί ότι $h_\mu(T^{-1}) = h_\mu(T)$. Έστω ξ διαμέριση. Τότε

$$\frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{i}\xi\right) = \frac{1}{n} H_\mu\left(T^{n-1}\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right),$$

όπου για την δεύτερη ισότητα εφαρμόστηκε η ιδιότητα 4.1.7 (4), $n-1$ φορές, για τον μετασχηματισμό T^{-1}, που διατηρεί το μέτρο. Παίρνοντας τώρα όρια καθώς n τείνει στο άπειρο παίρνουμε ότι

$$h_\mu(T^{-1}, \xi) = h_\mu(T, \xi). \qquad \square$$

Θεώρημα 4.2.8 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Αν $\xi \subset \mathcal{B}$ μετρήσιμη διαμέριση με πεπερασμένη εντροπία,

$$h_\mu(T, \xi) = \lim_{n\to\infty} H_\mu\left(\xi \;\middle|\; \bigvee_{i=1}^{n} T^{-i}\xi\right) = H_\mu\left(\xi \;\middle|\; \bigvee_{i=1}^{\infty} T^{-i}\xi\right),$$

όπου $\bigvee_{i=1}^{\infty} T^{-i}\xi$ είναι η τομή όλων των σ-αλγεβρών που περιέχουν όλες τις $T^{-i}\xi$, $i \in \mathbb{N}$.

Απόδειξη. Το όριο υπάρχει γιατί η ακολουθία $H_\mu\left(\xi \mid \bigvee_{i=1}^n T^{-i}\xi\right)$, $n \in \mathbb{N}$, είναι φθίνουσα, από την 4.1.8 (3). Από το Λήμμα 4.2.4 έχουμε ότι

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = H_\mu(\xi) + \sum_{k=1}^{n-1} H_\mu\left(\xi \;\middle|\; \bigvee_{j=1}^{k} T^{-j}\xi\right)$$

και επομένως

$$\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = \frac{1}{n}H_\mu(\xi) + \frac{1}{n}\sum_{k=1}^{n-1} H_\mu\left(\xi \;\middle|\; \bigvee_{j=1}^{k} T^{-j}\xi\right).$$

Αφήνοντας το n να τείνει στο άπειρο και αφού το Cesaro-όριο μιας ακολουθίας που έχει όριο είναι το όριο της ακολουθίας[25], παίρνουμε ότι

$$h_\mu(T,\xi) = \lim_{n\to\infty} H_\mu\left(\xi \;\middle|\; \bigvee_{i=1}^{n} T^{-i}\xi\right). \qquad \square$$

Κλείνουμε αυτή την παράγραφο αποδεικνύοντας τον ισχυρισμό στην (33) του Ορισμού 4.2.3.

Πρόταση 4.2.9 Έστω $(X,\mathcal{B},\mu,T)$ σ.δ.μ. Τότε

$$\sup_{\substack{\xi\subset\mathcal{B}\text{ πεπερασμένη}\\ \text{διαμέριση}}} h_\mu(T,\xi) = \sup_{\substack{H_\mu(\xi)<\infty\\ \xi\subset\mathcal{B}\text{ αριθμήσιμη}}} h_\mu(T,\xi).$$

Απόδειξη. Το αριστερό μέλος είναι πάντα μικρότερο ή ίσο από το δεξιό επειδή κάθε πεπερασμένη διαμέριση έχει πεπερασμένη εντροπία, και άρα αρκεί να δείξουμε ότι και το δεξιό μέλος είναι μικρότερο ή ίσο από το αριστερό. Από την ιδιότητα (4) του Θεωρήματος 4.2.6, για να το δείξουμε αυτό αρκεί να δείξουμε ότι για κάθε διαμέριση $\xi \subset \mathcal{B}$ με πεπερασμένη εντροπία και κάθε $\epsilon > 0$, υπάρχει πεπερασμένη διαμέριση $\eta \subset \mathcal{B}$ τέτοια ώστε $H_\mu(\xi \mid \eta) < \epsilon$. Μπορούμε να θεωρήσουμε την ξ άπειρη αριθμήσιμη, φυσικά, αφού αλλιώς παίρνουμε $\eta = \xi$, και έστω ότι $\xi = \{A_1, A_2, \dots\}$. Έστω λοιπόν $\epsilon > 0$ και έστω $n \in \mathbb{N}$ που θα επιλεγεί κατάλληλα παρακάτω. Θέτουμε $\eta = \{A_1, A_2, \dots A_n, B\}$, όπου $B = \bigcup_{k>n} A_k = X \setminus (A_1 \cup \dots \cup A_n)$. Τότε

$$\begin{aligned}
H_\mu(\xi \mid \eta) &= -\sum_{i=1}^{\infty}\sum_{j=1}^{n} \mu(A_i\cap A_j)\ln\frac{\mu(A_i\cap A_j)}{\mu(A_j)} - \sum_{i=1}^{\infty}\mu(A_i\cap B)\ln\frac{\mu(A_i\cap B)}{\mu(B)}\\
&= -\sum_{i=1}^{n}\sum_{j=1}^{n} \mu(A_i\cap A_j)\ln\frac{\mu(A_i\cap A_j)}{\mu(A_j)} - \sum_{i=n+1}^{\infty}\mu(A_i\cap B)\ln\frac{\mu(A_i\cap B)}{\mu(B)}\\
&= \sum_{i=1}^{n} \mu(A_i\cap A_i)\ln\frac{\mu(A_i\cap A_i)}{\mu(A_i)} - \sum_{i=n+1}^{\infty}\mu(A_i)\ln\frac{\mu(A_i)}{\mu(B)}\\
&= -\sum_{i=1}^{n} \mu(A_i)\ln 1 - \sum_{i=n+1}^{\infty}\mu(A_i)\ln\frac{\mu(A_i)}{\mu(B)}\\
&= -\sum_{i=n+1}^{\infty} \mu(A_i)\ln\mu(A_i) + \sum_{i=n+1}^{\infty}\mu(A_i)\ln\mu(B)\\
&= -\sum_{i=n+1}^{\infty} \mu(A_i)\ln\mu(A_i) + \mu(B)\ln\mu(B)\\
&\le -\sum_{i=n+1}^{\infty} \mu(A_i)\ln\mu(A_i).
\end{aligned}$$

Αφού $H_\mu(\xi) = -\sum_{i=1}^{\infty}\mu(A_i)\ln\mu(A_i) < +\infty$, υπάρχει $n \in \mathbb{N}$ ώστε $-\sum_{i=n+1}^{\infty}\mu(A_i)\ln\mu(A_i) < \epsilon$, και αν επιλέξουμε αυτό το n για τον ορισμό της η, τότε έχουμε ότι $H_\mu(\xi \mid \eta) < \epsilon$. $\square$

[25] Δηλαδή, αν $(a_n)_{n\in\mathbb{N}}$ ακολουθία αριθμών τέτοια ώστε $a_n \to a \in [-\infty,+\infty]$, τότε και $n^{-1}\sum_{k=1}^{n} a_k \to a$ καθώς $n \to \infty$.

4.3 Γεννήτορες και το Θεώρημα Kolmogorov–Sinai

Ορισμός 4.3.1 (μονόπλευρος γεννήτορας) Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. Μια μετρήσιμη πεπερασμένη διαμέριση $\xi \subset \mathcal{B}$ του X λέγεται πεπερασμένος, μονόπλευρος γεννήτορας του συστήματος αν

$$\bigvee_{k=0}^{\infty} T^{-k}\xi = \mathcal{B} \mod \mu,$$

όπου $\bigvee_{k=0}^{\infty} T^{-k}\xi = \sigma(\xi, T^{-1}\xi, T^{-2}\xi, \cdots)$ η μικρότερη σ-άλγεβρα που περιέχει όλες τις $T^{-k}\xi$, $k \in \mathbb{N}_0$.

Αν $(X, \mathcal{B}, \mu, T)$ είναι αντιστρέψιμο, μια τέτοια ξ λέγεται αμφίπλευρος (πεπερασμένος) γεννήτορας αν

$$\bigvee_{k=-\infty}^{\infty} T^{-k}\xi = \mathcal{B} \mod \mu.$$

Θεώρημα 4.3.2 (Kolmogorov–Sinai) Αν $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και ξ μονόπλευρος (πεπερασμένος) γεννήτορας, τότε

$$h_\mu(T) = h_\mu(T, \xi).$$

Αν το σύστημα είναι αντιστρέψιμο, τότε αν ξ αμφίπλευρος γεννήτορας τότε

$$h_\mu(T) = h_\mu(T, \xi).$$

Λήμμα 4.3.3 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $n \in \mathbb{N}$. Τότε για κάθε $\epsilon > 0$, υπάρχει $\delta > 0$, τέτοιο ώστε αν $\xi = \{A_1, \ldots, A_n\}$ και $\eta = \{B_1, \ldots, B_n\}$ είναι δύο μετρήσιμες διαμερίσεις με $\sum_{i=1}^{n} \mu(A_i \triangle B_i) < \delta$ τότε

$$H_\mu(\xi \mid \eta) + H_\mu(\eta \mid \xi) < \epsilon.$$

Απόδειξη. Έστω $\epsilon > 0$. Επιλέγουμε $\delta > 0$ με $\delta < 1/e$ και

$$-n(n-1)\delta \ln \delta - (1-\delta)\ln(1-\delta) < \frac{\epsilon}{2}.$$

Έστω $\xi = \{A_1, \ldots, A_n\}$ και $\eta = \{B_1, \ldots, B_n\}$ δύο μετρήσιμες διαμερίσεις με $\sum_{i=1}^{n} \mu(A_i \triangle B_j) < \delta$. Θεωρούμε την διαμέριση

$$\zeta = \left\{A_i \cap B_j : i, j \in \{1, \ldots, n\},\ i \neq j\right\} \cup \left\{\bigcup_{i=1}^{n} (A_i \cap B_i)\right\}.$$

Τότε για $i, j \in \{1, \ldots, n\}$ με $i \neq j$ έχει κανείς ότι

$$\mu(A_i \cap B_j) \leq \mu\left(\bigcup_{i=1}^{n} (A_i \triangle B_i)\right) < \delta,$$

αφού $A_i \cap B_j \subset A_i \triangle B_i$ (όπως επίσης και $A_i \cap B_j \subset A_j \triangle B_j$), όταν $i \neq j$. Επίσης

$$\mu\left(\bigcup_{i=1}^{n} (A_i \cap B_i)\right) = \sum_{i=1}^{n} \mu(A_i \cap B_i) = \sum_{i=1}^{n} [\mu(A_i \cup B_i) - \mu(A_i \triangle B_i)] > \sum_{i=1}^{n} \mu(A_i \cup B_i) - \delta \geq 1 - \delta,$$

αφού τα $A_i \cap B_i$, $i \in \{1, \ldots, n\}$, είναι ξένα ανά δύο, επειδή τα A_i είναι ξένα ανά δύο, όπως και τα B_i, και η ένωση $\bigcup_{i=1}^{n} (A_i \cup B_i)$ περιέχει κάθε μία από τις $\bigcup_{i=1}^{n} A_i$ και $\bigcup_{i=1}^{n} B_i$, κάθε μία από τις οποίες είναι όλος ο χώρος X.

Η $\phi(x) = x \ln x$ είναι κυρτή συνάρτηση με ελάχιστο στο $x = 1/e$. Αφού το δ επιλέχθηκε έτσι ώστε $\delta < 1/e$, οπότε επίσης $1 - \delta > 1/e$,

$$\begin{aligned} H_\mu(\zeta) &= -\sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} \mu(A_i \cap B_j) \ln \mu(A_i \cap B_j) - \mu\left(\bigcup_{i=1}^{n} (A_i \cap B_i)\right) \ln \mu\left(\bigcup_{i=1}^{n} (A_i \cap B_i)\right) \\ &< -n(n-1)\delta \ln \delta - (1-\delta)\ln(1-\delta) < \frac{\epsilon}{2}. \end{aligned}$$

Έπεται τώρα ότι

$$H_\mu(\eta) + H_\mu(\xi \mid \eta) = H_\mu(\xi \vee \eta) = H_\mu(\zeta \vee \eta) \leq H_\mu(\zeta) + H_\mu(\eta) < \frac{\epsilon}{2} + H_\mu(\eta),$$

οπότε

$$H_\mu(\xi \mid \eta) < \frac{\epsilon}{2}.$$

Παρομοίως, λόγω συμμετρίας έχουμε και ότι $H_\mu(\eta \mid \xi) < \epsilon/2$ και άρα τελικά το ζητούμενο. □

Λήμμα 4.3.4 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ., $\mathcal{A}$ μια άλγεβρα τέτοια ώστε

$$\sigma(\mathcal{A}) = \mathcal{B} \mod \mu$$

και έστω $\xi \subset \mathcal{B}$ μια πεπερασμένη διαμέριση. Τότε για κάθε $\epsilon > 0$, υπάρχει πεπερασμένη διαμέριση $\eta \subset \mathcal{A}$, τέτοια ώστε

$$H_\mu(\xi \mid \eta) < \epsilon.$$

Απόδειξη. Έστω $\epsilon > 0$, πεπερασμένη διαμέριση $\xi = \{A_1, \ldots, A_n\}$ και έστω δ το δ από το προηγούμενο Λήμμα (4.3.3) που αντιστοιχεί σε αυτό το n και το δοθέν ϵ. Από το Λήμμα 4.3.3. αρκεί να δειχτεί ότι υπάρχει $\eta \subset \mathcal{A}$ πεπερασμένη διαμέριση, έστω $\eta = \{B_1, \ldots, B_n\}$, τέτοια ώστε

$$\sum_{i=1}^{n} \mu(A_i \triangle B_i) < \delta.$$

Από το Λήμμα 2.1.10, δοθέντος $\gamma > 0$, για κάθε $i \in \{1, \ldots, n\}$ υπάρχει $C_i \in \mathcal{A}$ τέτοιο ώστε

$$\mu(A_i \triangle C_i) < \gamma.$$

Μέσω των C_i θα ορίσουμε μια διαμέριση. Έστω

$$C = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} (C_i \cap C_j).$$

Τότε, για $i \neq j$, είναι σαφές ότι

$$C_i \cap C_j \subset (C_i \triangle A_i) \cup (C_j \triangle A_j)$$

και άρα

$$\mu(C_i \cap C_j) \leq \mu(C_i \triangle A_i) + \mu(C_j \triangle A_j) < 2\gamma.$$

Έπεται ότι

$$\mu(C) \leq \frac{n(n-1)}{2} 2\gamma = n(n-1)\gamma. \tag{35}$$

Ορίζουμε

$$B_i = C_i \setminus C, \qquad i \in \{1, \ldots, n-1\},$$
$$B_n = X \setminus (B_1 \cup \cdots \cup B_{n-1}).$$

Τότε τα B_i, $i \in \{1, \ldots, n\}$, αποτελούν διαμέριση. Προφανώς $B_i \in \mathcal{A}$, εφόσον $C_i \in \mathcal{A}$ για κάθε $i \in \{1, \ldots, n\}$ και η $\mathcal{A}$ είναι άλγεβρα. Τώρα για $i \in \{1, \ldots, n-1\}$,

$$A_i \triangle B_i = A_i \triangle (C_i \cap C^{c}) \subset (A_i \triangle C_i) \cup (A_i \cap C).$$

Επομένως αν $i \in \{1, \ldots, n-1\}$,

$$\mu(A_i \triangle B_i) < \gamma + \mu(A_i \cap C).$$

Για $i = n$ ισχύει ότι

$$A_n \triangle B_n \subset \bigcup_{i=1}^{n-1} (A_i \triangle B_i),$$

επειδή και τα A_i και τα B_i, $i \in \{1, \ldots, n\}$, αποτελούν διαμέριση. Έπεται ότι

$$\sum_{i=1}^{n} \mu(A_i \triangle B_i) \leq 2 \sum_{i=1}^{n-1} \mu(A_i \triangle B_i) < 2(n-1)\gamma + 2 \sum_{i=1}^{n-1} \mu(A_i \cap C) \leq 2(n-1)\gamma + 2\mu(C) \leq 2(n^2-1)\gamma,$$

η προτελευταία ανισότητα πάλι επειδή τα $A_1, \ldots, A_n$ αποτελούν διαμέριση του X και η τελευταία από την (35). Άρα παίρνοντας το γ αρκούντως μικρό ώστε $2(n^2-1)\gamma < \delta$, μπορούμε να εφαρμόσουμε το Λήμμα 4.3.3 και να συμπεράνουμε ότι $H_\mu(\xi \mid \eta) < \epsilon$. □

Πόρισμα 4.3.5 Έστω $(X, \mathcal{B}, \mu, T)$ σ.δ.μ. και $\xi_1 \leqslant \xi_2 \leqslant \cdots$ πεπερασμένες μετρήσιμες διαμερίσεις τέτοιες ώστε

$$\bigvee_{n=1}^{\infty} \xi_n = \mathcal{B} \mod \mu.$$

Αν $\xi \subset \mathcal{B}$ πεπερασμένη διαμέριση, τότε

$$H_\mu(\xi \mid \xi_n) \xrightarrow{(n\to\infty)} 0.$$

Απόδειξη. Η $\mathcal{A} = \bigcup_{n=1}^{\infty} \sigma(\xi_n)$ είναι άλγεβρα γιατί η ακολουθία $\sigma(\xi_n)$, $n \in \mathbb{N}$, είναι αύξουσα. Από την υπόθεση, η άλγεβρα αυτή, $\mathcal{A}$, παράγει την $\mathcal{B}$, mod μ. Επομένως, από το Λήμμα 4.3.4 έχουμε ότι υπάρχει $\eta \subset \mathcal{A}$ πεπερασμένη διαμέριση με

$$H_\mu(\xi \mid \eta) < \epsilon.$$

Επειδή $\sigma(\xi_n) \subset \sigma(\xi_{n+1})$ $\forall n \in \mathbb{N}$ και η η είναι πεπερασμένη, ο εγκλεισμός $\eta \subset \mathcal{A}$ συνεπάγεται ότι $\eta \subset \sigma(\xi_n)$ για κάποιο $n \in \mathbb{N}$: πράγματι, κάθε στοιχείο A της η ανήκει σε κάποια $\sigma(\xi_{n_A})$ για κάποιο $n_A \in \mathbb{N}$, και επειδή $\sigma(\xi_n) \subset \sigma(\xi_{n+1})$ για κάθε $n \in \mathbb{N}$, έπεται ότι $\eta \subset \sigma(\xi_n)$, όπου n είναι το $\max_{A\in\eta} n_A$. Τότε όμως $\eta \leqslant \xi_n$, αφού η $\sigma(\xi_n)$ αποτελείται ακριβώς από πεπερασμένες ενώσεις στοιχείων της ξ_n και άρα $\eta \subset \sigma(\xi_n)$ συνεπάγεται ότι κάθε στοιχείο της η είναι ένωση στοιχείων της ξ_n. Για $m \geq n$ έχει τότε κανείς ότι

$$H_\mu(\xi \mid \xi_m) \leq H_\mu(\xi \mid \xi_n) \leq H_\mu(\xi \mid \eta) < \varepsilon,$$

από την ιδιότητα 4.1.8 (3) και την μονοτονία των ξ_n. □

Απόδειξη Θεωρήματος 4.3.2 Έστω ξ μονόπλευρος γεννήτορας. Αρκεί να δειχτεί ότι

$$h_\mu(T, \eta) \leq h_\mu(T, \xi) \tag{36}$$

για κάθε πεπερασμένη διαμέριση η. Για μία αυθαίρετη πεπερασμένη διαμέριση η έχει κανείς ότι

$$\begin{aligned} h_\mu(T, \eta) &\leq h_\mu\left(T, \bigvee_{i=0}^{n-1} T^{-i}\xi\right) + H_\mu\left(\eta \;\middle|\; \bigvee_{i=0}^{n-1} T^{-i}\xi\right) && \text{από 4.2.6 (4)} \\ &= h_\mu(T, \xi) + H_\mu\left(\eta \;\middle|\; \bigvee_{i=0}^{n-1} T^{-i}\xi\right) && \text{από 4.2.6 (6).} \end{aligned} \tag{37}$$

Τώρα εάν $\xi_n = \bigvee_{i=0}^{n-1} T^{-i}\xi$, $n \in \mathbb{N}$, τότε η ξ_n είναι αύξουσα ακολουθία και αφού $\bigvee_{n\geq 0} T^{-n}\xi = \mathcal{B} \mod \mu$ (ξ είναι γεννήτορας), από το Πόρισμα 4.3.5 παίρνουμε ότι

$$H_\mu\left(\eta \;\middle|\; \bigvee_{i=0}^{n-1} T^{-i}\xi\right) \xrightarrow{(n\to\infty)} 0$$

και η (37) δίνει την ζητούμενη (36).

Όταν το σύστημα είναι αντιστρέψιμο και η ξ αμφίπλευρος γεννήτορας,

$$\begin{aligned} h_\mu(T, \eta) &\leq h_\mu\left(T, \bigvee_{i=-n+1}^{n-1} T^{i}\xi\right) + H_\mu\left(\eta \;\middle|\; \bigvee_{i=-n+1}^{n-1} T^{i}\xi\right) && \text{από 4.2.6 (4)} \\ &= h_\mu(T, \xi) + H_\mu\left(\eta \;\middle|\; \bigvee_{i=-n+1}^{n-1} T^{i}\xi\right) && \text{από 4.2.6 (7),} \end{aligned}$$

Αν τώρα θέσουμε $\xi_n = \bigvee_{i=-n+1}^{n-1} T^i\xi$, $n \in \mathbb{N}$, τότε η ξ_n είναι πάλι αύξουσα ακολουθία και αφού $\bigvee_{n\in\mathbb{Z}} T^n\xi = \mathcal{B} \bmod \mu$ (ξ είναι αμφίπλευρος γεννήτορας), από το Πόρισμα 4.3.5 παίρνουμε πάλι ότι

$$H_\mu\left(\eta \;\middle|\; \bigvee_{i=-n+1}^{n-1} T^{-i}\xi\right) \xrightarrow{(n\to\infty)} 0.$$

Έπεται συνεπώς πάλι ότι $h_\mu(T,\eta) \leq h_\mu(T,\xi)$ για αυθαίρετη πεπερασμένη διαμέριση $\eta \subset \mathcal{B}$ του X. $\square$

Πόρισμα 4.3.6 Αν $(X,\mathcal{B},\mu,T)$ είναι αντιστρέψιμο σ.δ.μ. και έχει μονόπλευρο γεννήτορα $\xi \subset \mathcal{B}$, τότε

$$h_\mu(T) = 0.$$

Απόδειξη. Έχουμε

$$\begin{aligned}
h_\mu(T) &= h_\mu(T,\xi) && \text{από Θεώρημα 4.3.2}\\
&= \lim_{n\to\infty} H_\mu\left(\xi \;\middle|\; \bigvee_{i=1}^{n} T^{-i}\xi\right) && \text{από Θεώρημα 4.2.8}\\
&= \lim_{n\to\infty} H_\mu\left(T\xi \;\middle|\; T\bigvee_{i=1}^{n} T^{-i}\xi\right) && \text{από 4.1.8 (4)}\\
&= \lim_{n\to\infty} H_\mu\left(T\xi \;\middle|\; \bigvee_{i=0}^{n-1} T^{-i}\xi\right) &&\\
&= 0 && \text{από Πόρισμα 4.3.5,}
\end{aligned}$$

εφόσον $\xi_n = \bigvee_{i=0}^{n-1} T^{-i}\xi$, $n \in \mathbb{N}$, είναι αύξουσα ακολουθία και $\bigvee_{n\geq 1}\xi_n = \mathcal{B} \mod \mu$. $\square$

Εφαρμογή 4.3.7 Έστω $S = \{1,\dots,s\}$ πεπερασμένο σύνολο, $X = S^{\mathbb{N}}$, $T\colon X \to X$ ο μετασχηματισμός shift, δηλαδή

$$T(x_1,x_2,\dots) = (x_2,x_3,\dots)$$

και $\mathcal{B}$ η σ-άλγεβρα γινόμενο $\mathcal{B} = \otimes_{\mathbb{N}}\mathcal{P}(S)$, η οποία παράγεται από τους κυλίνδρους

$$\{\underset{\sim}{x} = (x_1,x_2,\dots) \in X \colon x_1 = i_1, x_2 = i_2, \dots, x_n = i_n\}, \qquad n \in \mathbb{N},\ i_1,\dots,i_n \in S.$$

Αν $p = (p_1,\dots,p_s)$ διάνυσμα πιθανότητας και μ το μέτρο στον $(X,\mathcal{B})$ που ικανοποιεί την

$$\mu(\{\underset{\sim}{x} = (x_1,x_2,\dots) \in X \colon x_1 = i_1, x_2 = i_2, \dots, x_n = i_n\}) = p_{i_1}p_{i_2}\cdots p_{i_n},$$

τότε

$$h_\mu(T) = -\sum_{i=1}^{s} p_i \ln p_i.$$

Απόδειξη. Ορίζουμε διαμέριση $\xi = \{[1],[2],\dots,[s]\}$, όπου $[i] = \{\underset{\sim}{x} = (x_1,x_2,\dots) \in X \colon x_1 = i\}$. Τότε

$$T^{-1}\xi = \{T^{-1}[1],\dots,T^{-1}[s]\}$$

κ.ο.κ και, για κάθε $n \in \mathbb{N}$,

$$\bigvee_{i=0}^{n-1} T^{-i}\xi = \{[i_1] \cap T^{-1}[i_2] \cap \dots \cap T^{-(n-1)}[i_n] \colon i_1,i_2,\dots,i_n \in S\} = \{\text{οι κύλινδροι μήκους } n\}.$$

Οι κύλινδροι παράγουν την $\mathcal{B}$, εξ' ορισμού, άρα

$$h_\mu(T) = h_\mu(T,\xi).$$

Υπολογίζουμε

$$\begin{aligned}
H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) &= -\sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} \mu\big([i_1]\cap T^{-1}[i_2]\cap\cdots\cap T^{-(n-1)}[i_n]\big)\ln\mu\big([i_1]\cap T^{-1}[i_2]\cap\cdots\cap T^{-(n-1)}[i_n]\big)\\
&= -\sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} p_{i_1}p_{i_2}\cdots p_{i_n}\ln(p_{i_1}p_{i_2}\cdots p_{i_n})\\
&= -\sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} p_{i_1}p_{i_2}\cdots p_{i_n}\,(\ln p_{i_1}+\ln p_{i_2}+\cdots+\ln p_{i_n})\\
&= -\sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} p_{i_1}p_{i_2}\cdots p_{i_n}\ln p_{i_1} - \sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} p_{i_1}p_{i_2}\cdots p_{i_n}\ln p_{i_2} - \cdots\\
&\qquad\qquad - \sum_{i_1=1}^{s}\cdots\sum_{i_n=1}^{s} p_{i_1}p_{i_2}\cdots p_{i_n}\ln p_{i_n}\\
&= -\sum_{i_1=1}^{s} p_{i_1}\ln p_{i_1} - \sum_{i_2=1}^{s} p_{i_2}\ln p_{i_2} - \cdots - \sum_{i_n=1}^{s} p_{i_n}\ln p_{i_n}\\
&= -n\sum_{i=1}^{s} p_i\ln p_i.
\end{aligned}$$

Άρα

$$\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = -\sum_{i=1}^{s} p_i\ln p_i \qquad \forall\, n\in\mathbb{N}$$

επομένως και το όριο καθώς n τείνει στο άπειρο είναι το ίδιο και τελικά

$$h_\mu(T) = -\sum_{i=1}^{s} p_i\ln p_i. \qquad \square$$

Εφαρμογή 4.3.8 (Markov Shifts) Έστω πάλι $S=\{1,\dots,s\}$ πεπερασμένο σύνολο, $X=S^{\mathbb{N}_0}$, $\mathcal{B}$ η σ-άλγεβρα γινόμενο $\mathcal{B}=\otimes_{\mathbb{N}_0}\mathcal{P}(S)$, η οποία παράγεται από τους κυλίνδρους, και $T\colon X\to X$ ο μετασχηματισμός shift. Έστω P στοχαστικός πίνακας, p διάνυσμα πιθανότητας που είναι αριστερό ιδιοδιάνυσμα του πίνακα P και μέτρο μ τέτοιο ώστε

$$\mu(\{\underset{\sim}{x}=(x_0,x_1,\dots)\in X\colon x_0=i_0,x_1=i_1,\dots,x_n=i_n\}) = p_{i_0}P_{i_0\,i_1}P_{i_1\,i_2}\cdots P_{i_{n-1}\,i_n}.$$

Τότε

$$h_\mu(T) = -\sum_{i,j=1}^{s} p_iP_{ij}\ln P_{ij}.$$

Η απόδειξη είναι όπως και στην περίπτωση των Bernoulli shifts της Εφαρμογής 4.3.7.

4.4 Το Θεώρημα Shannon–MacMillan–Breiman

Κλείνουμε με ένα σημαντικό θεώρημα με πολλές εφαρμογές, το οποίο παρατίθεται χωρίς απόδειξη.

Θεώρημα 4.4.1 (Shannon–McMillan–Breiman) Έστω $(X,\mathcal{B},\mu,T)$ ένα εργοδικό σ.δ.μ. Για μία πεπερασμένη (μετρήσιμη) διαμέριση ξ, έστω $\xi_n(x)$ το (μοναδικό) στοιχείο της διαμέρισης $\bigvee_{i=0}^{n-1}T^{-i}\xi$ που περιέχει το $x\in X$. Τότε

$$-\frac{1}{n}\ln\mu(\xi_n(x)) \xrightarrow{(n\to\infty)} h_\mu(T,\xi) \qquad \text{για } \mu\text{-σχεδόν κάθε } x\in X$$

και στον $L^1(X\mathcal{B},\mu)$.

Για παράδειγμα, έστω $S = \{1, 2\}$, $X = S^{\mathbb{N}}$, $T\colon X \to X$ ο μετασχηματισμός shift, δηλαδή

$$T(x_1, x_2, \dots) = (x_2, x_3, \dots)$$

και $\mathcal{B}$ η σ-άλγεβρα γινόμενο $\mathcal{B} = \otimes_{\mathbb{N}} \mathcal{P}(S)$, η οποία παράγεται από τους κυλίνδρους

$$\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = i_1, x_2 = i_2, \dots, x_n = i_n\}, \qquad n \in \mathbb{N},\ i_1, \dots, i_n \in S.$$

Αν $q \in (0, 1)$, έστω μ το μέτρο γινόμενο στον $(X, \mathcal{B})$ που αντιστοιχεί στο διάνυσμα πιθανότητας $p = (p_1, p_2) = (q, 1 - q)$. Το μ είναι το μέτρο που ικανοποιεί την

$$\begin{aligned}\mu(\{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = i_1, x_2 = i_2, \dots, x_n = i_n\}) &= p_{i_1} p_{i_2} \cdots p_{i_n} \\ &= q^{|\{k \in \mathbb{N} \cap [1,n] \colon i_k = 1\}|} (1 - q)^{|\{k \in \mathbb{N} \cap [1,n] \colon i_k = 2\}|}\end{aligned}$$

για κάθε $n \in \mathbb{N}$ και $i_1, \dots, i_n \in \{1, 2\}$. Έστω και ξ η διαμέριση $\xi = \{[1], [2]\}$, όπου, ως συνήθως,

$$[i] = \{\underset{\sim}{x} = (x_1, x_2, \dots) \in X \colon x_1 = i\} \qquad i \in \{1, 2\}.$$

Τότε, για κάθε ακολουθία $\underset{\sim}{x} = (x_1, x_2, \dots) \in X$, $\xi_n(\underset{\sim}{x})$ είναι ο μοναδικός κύλινδρος μήκους n που περιέχει το $\underset{\sim}{x}$. Για $q = \frac{1}{2}$, έχει κανείς ότι κάθε κύλινδρος μήκους n έχει μέτρο 2^{-n}, οπότε έχει κανείς ότι

$$-\frac{1}{n} \ln \mu(\xi_n(\underset{\sim}{x})) = \ln 2 = h_\mu(T, \xi) \qquad \forall\, n \in \mathbb{N}\ \forall\, \underset{\sim}{x} \in X.$$

Για άλλες τιμές του $q \in (0, 1)$, το Θεώρημα Shannon–McMillan–Breiman δίνει ότι

$$-\frac{1}{n} \ln \mu(\xi_n(\underset{\sim}{x})) \xrightarrow{(n \to \infty)} -q \ln q - (1 - q) \ln(1 - q),$$

για μ-σχεδόν κάθε ακολουθία $\underset{\sim}{x} = (x_1, x_2, \dots)$. Δηλαδή,

$$\mu(\xi_n(\underset{\sim}{x})) \approx e^{n[q \ln q + (1-q) \ln(1-q)]} = q^{nq} (1 - q)^{n(1-q)}$$

για μεγάλα n, για μ-σχεδόν κάθε ακολουθία $\underset{\sim}{x} = (x_1, x_2, \dots)$.