ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΡΕΤΗΡΙΟ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΕΡΓΟΔΙΚΗ ΘΕΩΡΙΑ

Σημειώσεις
των Κ. Αθανασόπουλου και Δ. Γατζούρα

ΗΡΑΚΛΕΙΟ 1994

I. ΜΕΤΡΗΣΙΜΑ ΔΥΝΑΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

1. Βασικοί ορισμοί και παραδείγματα

Έστω $(X,\mathcal{A},\mu)$ ένας χώρος πιθανότητας. Ένας αυτομορφισμός του είναι μία 1-1 και επί απεικόνιση $T: X \to X$ για την οποία

$T(A), T^{-1}(A) \in \mathcal{A}$ και $\mu(A) = \mu(T(A)) = \mu(T^{-1}(A))$ για κάθε $A \in \mathcal{A}$.

Ο T ορίζει την διακριτή μονοπαραμετρική ομάδα αυτομορφισμών $(T^k)_{k\in\mathbb{Z}}$.

Μία μετρήσιμη ροή στον X είναι μία μονοπαραμετρική ομάδα αυτομορφισμών $(\varphi_t)_{t\in\mathbb{R}}$ του X ώστε η απεικόνιση εκτίμησης $\varphi: \mathbb{R}\times X \to X$ να είναι μετρήσιμη. Έτσι για κάθε μετρήσιμη συνάρτηση $f: X\to\mathbb{R}$, η $f\circ\varphi_t : X \to \mathbb{R}$ είναι μετρήσιμη για κάθε $t\in\mathbb{R}$.

Ένας ενδομορφισμός του X είναι μία επί απεικόνιση $T: X\to X$ ώστε $T^{-1}(A)\in\mathcal{A}$ και $\mu(T^{-1}(A)) = \mu(A)$ για κάθε $A\in\mathcal{A}$. Αν στον ορισμό της μετρήσιμης ροής αντικαταστήσουμε το $\mathbb{R}$ με το $\mathbb{R}^+$, το αντικείμενο που προκύπτει λέγεται μετρήσιμη ημιροή. Με τον γενικό όρο μετρήσιμο δυναμικό σύστημα θα εννοούμε ένα οποιοδήποτε από τα τέσσερα αντικείμενα που ορίστηκαν προηγουμένως.

Δύο μετρήσιμα δυναμικά συστήματα $(\varphi^i_t)_{t\in K}$ στους χώρους $(X,\mathcal{A}_i,\mu_i)$ $i=1,2$, αντίστοιχα, όπου $K=\mathbb{Z},\mathbb{Z}^+,\mathbb{R}$ ή $\mathbb{R}^+$, λέγονται (πιθανοθεωρητικά) ισόμορφα, αν υπάρχουν δύο αμετάβλητα σύνολα $Y_i\in\mathcal{A}_i$ με $\mu_i(Y_i)=1$, $i=1,2$ και ένας ισομορφισμός $h: (Y_1,\mathcal{A}_1,\mu_1) \to (Y_2,\mathcal{A}_2,\mu_2)$ ώστε

$\varphi^2_t \circ h = h\circ\varphi^1_t$ στο Y_1 και $\varphi^1_t\circ h^{-1} = h^{-1}\circ\varphi^2_t$ στο Y_2 για κάθε $t\in K$.

1.1. Παράδειγμα. Έστω G μία συμπαγής τοπολογική ομάδα. Τότε υπάρχει ένα μονοσήμαντα ορισμένο μέτρο πιθανότητας Borel μ στην G ώστε $\mu(gA)=\mu(A)$ για κάθε σύνολο Borel $A\subset G$. Το μ λέγεται μέτρο Haar της G. Για παράδειγμα αν $G=S^1$, τότε το μ είναι το κανονικοποιημένο μέτρο Lebesgue.

Έστω $T: G\to G$ ένας επιμορφισμός τοπολογικών ομάδων. Αν $\lambda(A) = \mu(T^{-1}(A))$ για κάθε σύνολο Borel $A\subset G$, έχουμε

$\lambda(T(x).A) = \mu(T^{-1}(T(x).A)) = \mu(x.T^{-1}(A)) = \mu(T^{-1}(A)) = \lambda(A)$ για κάθε $x\in G$.

Επειδή η T είναι επί και το μ είναι μονοσήμαντα ορισμένο, προκύπτει ότι $\lambda=\mu$.

Συνεπώς η T διατηρεί το μέτρο Haar. Ειδικά για $G=S^1$, η $T_n: S^1 \to S^1$ με $T_n(z)=z^n$, διατηρεί το κανονικοποιημένο μέτρο Lebesgue, $n\in\mathbb{Z}$.

<u>1.2 Παράδειγμα</u> Έστω $k\in\mathbb{N}$, $k\geq 2$, και $X_n=\{0,1,\dots,k-1\}$ για κάθε $n\in\mathbb{Z}$. Στο X_n θεωρούμε το μέτρο μ_0 με $\mu_0(j)=p_j$, όπου $p_j\geq 0$ και $p_0+\dots+p_{k-1}=1$. Αν στο X_n θεωρήσουμε την διακριτή τοπολογία τότε ο χώρος γινόμενο $X=\prod_{n\in\mathbb{Z}}X_n=\{0,1,\dots,k-1\}^{\mathbb{Z}}$ είναι μετρικός, συμπαγής και ολικά-μη-συνεκτικός. Η τοπολογία του X έχει βάση όλα τα σύνολα της μορφής

$$C=\{x=(x_n)_{n\in\mathbb{Z}}\in X:\ x_{n_1}\in A_{n_1},\dots,x_{n_\lambda}\in A_{n_\lambda}\}$$

όπου $n_1,\dots,n_\lambda\in\mathbb{Z}$, και $A_{n_i}\subset X_{n_i}$, $1\leq i\leq\lambda$, που λέγονται κύλινδροι.

Έστω $\mathcal{A}$ η σ-άλγεβρα του X που παράγεται από τους κυλίνδρους. Η $\mathcal{A}$ περιέχει όλα τα σύνολα Borel στον X. Στον X ορίζεται το μέτρο πιθανότητας Borel μ με τύπο

$$\mu(C)=\mu_0(A_{n_1})\cdots\mu_0(A_{n_\lambda}).$$

Ο μετασχηματισμός $\sigma: X\to X$ με τύπο $\sigma((x_n)_{n\in\mathbb{Z}})=(x_{n+1})_{n\in\mathbb{Z}}$ λέγεται shift. Για κάθε κύλινδρο C έχουμε

$$\sigma^{-1}(C)=\{x=(x_n)_{n\in\mathbb{Z}}\in X:\ x_{n_1-1}\in A_{n_1},\dots,x_{n_\lambda-1}\in A_{n_\lambda}\}$$

Συνεπώς $\mu(\sigma^{-1}(C))=\mu(C)$. Όμοια $\mu(\sigma(C))=\mu(C)$. Άρα το shift είναι αυτομορφισμός του χώρου πιθανότητας $(X,\mathcal{A},\mu)$.

Ειδικά όταν $p_j=\frac{1}{k}$, $0\leq j\leq k-1$, τότε το παράδειγμα γίνεται ειδική περίπτωση του παραδείγματος 1.1. Στο X_n ορίζεται δομή αβελιανής ομάδας που το κάνει ισόμορφο με την πεπερασμένη κυκλική ομάδα $\mathbb{Z}_k$. Στον X ορίζεται τότε η δομή ευθύ γινόμενο, δηλαδή

$$(x_n)_{n\in\mathbb{Z}}+(y_n)_{n\in\mathbb{Z}}=\left((x_n+y_n)\ (\mathrm{mod}\ k)\right)_{n\in\mathbb{Z}}$$

Ο X γίνεται έτσι μετρικοποιήσιμη, συμπαγής, αβελιανή ομάδα (ολικά-μη-συνεκτική) και το shift $\sigma: X\to X$ αυτομορφισμός ομάδων. Όταν $p_j=\frac{1}{k}$, $0\leq j\leq k-1$ το μ_0 είναι το μέτρο Haar στο X_n, γιατί το μέτρο κάθε συνόλου στο X_n εξαρτάται μόνο από τον πληθάριθμό του και συνεπώς το μ_0 είναι αναλλοίωτο από τις μεταφορές του X_n, που είναι βέβαια 1-1 και επί απεικονίσεις. Το μ είναι τότε το μέτρο Haar της X.

2. Το Θεώρημα της επαναφοράς του Poincaré

Είναι σαφές ότι η έννοια του μετρήσιμου δυναμικού συστήματος όπως ορίστηκε στην παράγραφο 1 είναι πολύ γενική και ότι για την εξαγωγή αξιοσημείωτων αποτελεσμάτων θα χρειαστούν επιπλέον υποθέσεις για την φύση του συστήματος. Υπάρχει όμως ένα γενικό αξιοσημείωτο θεώρημα ποιοτικού χαρακτήρα. Δίνουμε πρώτα την πιθανοθεωρητική του διατύπωση.

2.1. Θεώρημα (Poincaré - Gibbs) Έστω T ένας ενδομορφισμός ενός χώρου πιθανότητας $(X,\mathcal{A},\mu)$. Έστω $A\in\mathcal{A}$ και $A_0=\{x\in A:\ T^n(x)\in A$ για άπειρα $n\geq 0\}$. Τότε $A_0\in\mathcal{A}$ και $\mu(A_0)=\mu(A)$.

Απόδειξη. Θέτουμε $C_n=\{x\in A:\ T^m(x)\notin A$ για κάθε $m\geq n\}$. Τότε

$$A_0=A\setminus\bigcup_{n=1}^{\infty}C_n$$

Αρκεί λοιπόν να δείξουμε ότι $C_n\in\mathcal{A}$ και $\mu(C_n)=0$ για κάθε $n\in\mathbb{N}$.
Παρατηρούμε κατ' αρχήν ότι

$$C_n=A\setminus\bigcup_{m\geq n}T^{-m}(A)=A\setminus T^{-n}\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)$$

Αφού ο T είναι ενδομορφισμός, $T^{-m}(A)\in\mathcal{A}$ για κάθε $m\geq n$ και κατά συνέπεια $C_n\in\mathcal{A}$. Επίσης έχουμε:

$$C_n=A\setminus\bigcup_{m\geq n}T^{-m}(A)\subset\bigcup_{m\geq 0}T^{-m}(A)\setminus\bigcup_{m\geq n}T^{-m}(A)$$

και άρα

$$\begin{aligned}\mu(C_n)&\leq\mu\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)-\mu\Big(\bigcup_{m\geq n}T^{-m}(A)\Big)\\&=\mu\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)-\mu\Big(T^{-n}\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)\Big)\\&=\mu\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)-\mu\Big(\bigcup_{m\geq 0}T^{-m}(A)\Big)=0\qquad\text{οεδ.}\end{aligned}$$

Στην συνέχεια θα παρουσιάσουμε την τοπολογική άποψη του θεωρήματος του Poincaré για τοπολογικά συνεχή δυναμικά συστήματα. Μια τοπολογική ροή σ' ένα μετρικό χώρο X είναι μια μονοπαραμετρική ομάδα ομοιομορφισμών $(\varphi_t)_{t\in\mathbb{R}}$ του X, ώστε η απεικόνιση εκτίμησης $\varphi:\mathbb{R}\times X\to X$ να είναι συνεχής. Για κάθε $x\in X$, το σύνολο $L^{\pm}(x)=\{y\in X:\ \varphi(t_n,x)\to y$, για κάποια $t_n\to\pm\infty\}$ λέγεται θετικό (αντίστοιχα αρνητικό) οριακό σύνολο του x. Θέτουμε $P^{\pm}=\{x\in X:\ x\in L^{\pm}(x)\}$ και $P=P^+\cap P^-=\{x\in X:\ x\in L^+(x)\cap L^-(x)\}$.

2.2 Λήμμα. $x\in P^+$ τότε και μόνον τότε όταν για κάθε περιοχή V του x στον X υπάρχει $t\geq 1$ ώστε $\varphi_t(x)\in V$.

Απόδειξη. Το ευθύ είναι προφανές. Για το αντίστροφο, θεωρούμε μία βάση περιοχών $\{V_n : n\in\mathbb{N}\}$ του x. Σύμφωνα με την υπόθεση υπάρχουν $t_n\geq 1$ ώστε $\varphi_{t_n}(x)\in V_n$, $n\in\mathbb{N}$. Προφανώς $\varphi_{t_n}(x)\to x$. Η ακολουθία $(t_n)_{n\in\mathbb{N}}$ είτε έχει συγκλίνουσα υπακολουθία είτε $t_n\to+\infty$, οπότε $x\in P^+$. Στην πρώτη περίπτωση, λόγω της συνέχειας της φ, υπάρχει $t\geq 1$ ώστε $\varphi(t,x)=x$ και συνεπώς $\varphi(nt,x)=x$ για κάθε $n\in\mathbb{N}$. Άρα $x\in P^+$.

2.3. Θεώρημα Έστω $(\varphi_t)_{t\in\mathbb{R}}$ μία τοπολογική ροή σ' έναν διαχωρίσιμο μετρικό χώρο X, που διατηρεί ένα μέτρο πιθανότητας Borel μ στον X. Τότε $\mu(P)=1$.

Απόδειξη Έστω $A\subset X$ ένα σύνολο Borel. Τα σύνολα

$$A^+ = A\setminus \bigcup_{n=0}^{\infty} A\cap\varphi_n(A), \quad A^- = A\setminus\bigcup_{n=1}^{\infty} A\cap\varphi_{-n}(A)$$

είναι επίσης Borel και $\mu(A^\pm)=\mu(\varphi_t(A^\pm))$ για κάθε $t\in\mathbb{R}$. Προφανώς $(\varphi_t(A))^\pm=\varphi_t(A^\pm)$ για κάθε $t\in\mathbb{R}$. Επίσης αν $k>\lambda\geq 0$, έχουμε $\varphi_k(A^+)\cap\varphi_\lambda(A^+)=\varphi_{k-\lambda}(A^+)\cap A^+=\emptyset$. Προκύπτει τώρα ότι

$$1\geq\mu\Big(\bigcup_{k=0}^{\infty}\varphi_k(A^+)\Big)=\sum_{k=0}^{\infty}\mu(\varphi_k(A^+))=\sum_{k=0}^{\infty}\mu(A^+).$$

Αυτό μπορεί να συμβαίνει μόνο όταν $\mu(A^+)=0$. Ανάλογα αποδεικνύεται ότι $\mu(A^-)=0$.

Έστω τώρα $\{A_n : n\in\mathbb{N}\}$ μία αριθμήσιμη βάση της τοπολογίας του X. Θέτουμε $B^\pm=\bigcup_{n=1}^{\infty}A_n^\pm$ και $B=B^+\cup B^-$. Σύμφωνα με τα προηγούμενα, $\mu(B^\pm)=\mu(B)=0$ και συνεπώς $\mu(X\setminus B)=1$. Αρκεί λοιπόν να δείξουμε ότι $X\setminus B\subset P$. Έχουμε $X\setminus B=(X\setminus B^+)\cap(X\setminus B^-)$. Έστω $x\in X\setminus B$ και A_λ ένα βασικό ανοικτό σύνολο με $x\in A_\lambda$. Τότε $x\in(X\setminus A_\lambda^+)\cap(X\setminus A_\lambda^-)$. Συνεπώς υπάρχουν $m,n>0$ ώστε $x\in A_\lambda\cap\varphi_m(A_\lambda)\cap\varphi_{-n}(A_\lambda)$, δηλαδή $\varphi_{-m}(x)\in A_\lambda$ και $\varphi_n(x)\in A_\lambda$. Από το Λήμμα 2.2 προκύπτει αμέσως ότι $x\in P$. ο.ε.δ.

2.4 Πόρισμα Έστω $(\varphi_t)_{t\in\mathbb{R}}$ μία τοπολογική ροή σ' έναν διαχωρίσιμο μετρικό χώρο X. Αν μ είναι ένα μέτρο πιθανότητας Borel στον X που παραμένει αναλλοίωτο από την ροή, τότε ο φορέας του μ περιέχεται στο $\overline{P}$. Το $\overline{P}$ λέγεται κέντρο Birkhoff του συστήματος.

3. Αμετάβλητα μέτρα διαφορίσιμων διανυσματικών πεδίων.

Έστω M μία C^{∞} n-πολλαπλότητα. Αν η M είναι προσανατολίσιμη, τότε κάθε C^{1} προσανατολισμός ω της M με $\int_M \omega > 0$, ορίζει μία θετική γραμμική μορφή $\mu_\omega : C_0(M) \to \mathbb{R}$ με τύπο

$$\mu_\omega(f) = \int_M f\omega$$

Από το θεώρημα της αναπαράστασης του Riesz προκύπτει ότι η μ_ω είναι ένα μέτρο Borel της M. Όλα τα μέτρα μ_ω που παίρνονται μ' αυτόν τον τρόπο είναι ισοδύναμα μεταξύ τους και γι αυτό συνήθως το μ_ω λέγεται μέτρο Lebesgue της M.
Αν $h: M \to M$ είναι μία C^1 αμφιδιαφόριση που διατηρεί τον προσανατολισμό της M, τότε από τον τύπο αλλαγής της μεταβλητής κατά την ολοκλήρωση έχουμε για κάθε $f \in C_0(M)$

$$\mu_\omega(f) = \int_M f\omega = \int_M h^*(f\omega) = \int_M (f\circ h)\cdot h^*\omega = \mu_{h^*\omega}(f\circ h)$$

και επομένως για κάθε σύνολο Borel $A \subset M$,

$$\mu_\omega(h(A)) = \mu_\omega(\chi_{h(A)}) = \mu_{h^*\omega}(\chi_{h(A)}\circ h) = \mu_{h^*\omega}(\chi_A) = \mu_{h^*\omega}(A)$$

Η $h^*\omega$ είναι και αυτή ένας προσανατολισμός της M και $\int_M h^*\omega > 0$, επειδή η h διατηρεί τον προσανατολισμό της M. Για κάθε $x \in M$ υπάρχει ένας μοναδικά ορισμένος πραγματικός αριθμός που συμβολίζουμε με $\det_\omega Dh(x)$ ώστε $(h^*\omega)_x = (\det_\omega Dh(x))\,\omega_{h(x)}$. Η συνάρτηση $\det_\omega Dh : M \to \mathbb{R}$ είναι συνεχής αφού η h είναι C^1. Επιπλέον $\det_\omega Dh > 0$, γιατί η h διατηρεί τον προσανατολισμό. Έτσι έχουμε για κάθε $f \in C_0(M)$

$$\mu_\omega(f) = \int_M (f\circ h).(\det_\omega Dh)\,\omega = \mu_{h^*\omega}(f\circ h)$$

Από αυτό προκύπτει ότι $\mu_\omega(f\circ h) = \mu_\omega(f)$ για κάθε $f \in C_0(M)$ ακριβώς τότε όταν $\det_\omega Dh = 1$.

Έστω τώρα ξ ένα C^r διανυσματικό πεδίο, $r \geq 2$, στην προσανατολισμένη C^{∞} n-πολλαπλότητα (M,ω). Θα εξετάσουμε, πότε η ροή του ξ διατηρεί το μέτρο Lebesgue της M.

3.1. Ορισμός Η απόκλιση του ξ ως προς ω είναι η μοναδικά ορισμένη συνεχής συνάρτηση $div_\omega \xi$ με την ιδιότητα $d(i_\xi\omega) = (div_\omega \xi)\omega$.

Επειδή η ω είναι n-μορφή, $d\omega = 0$ και συνεπώς

$$L_\xi \omega = d(i_\xi \omega) + i_\xi(d\omega) = d(i_\xi \omega) = (div_\omega \xi)\omega.$$

3.2. Παράδειγμα. Αν $M = \mathbb{R}^n$ και $\omega = \rho\, dx_1 \wedge \dots \wedge dx_n$, όπου $\rho : \mathbb{R}^n \to \mathbb{R}$ είναι μία C^2 συνάρτηση με $\rho(x) > 0$ για κάθε $x \in \mathbb{R}^n$, τότε για ένα C^2 διανυσματικό πεδίο $\xi = (\xi_1, \dots, \xi_n)$ έχουμε, επειδή η i_ξ είναι αντιπαραγώγιση:

$$i_\xi \omega = \sum_{i=1}^n (-1)^{i-1} \rho\, dx_1 \wedge \dots \wedge i_\xi(dx_i) \wedge \dots \wedge dx_n$$

$$= \sum_{i=1}^n (-1)^{i-1} \rho \xi_i\, dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

αφού $i_\xi(dx_i) = dx_i(\xi) = (0, \dots, 1, \dots 0) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \xi_i$. Συνεπώς

$$d(i_\xi\omega) = \sum_{i=1}^n (-1)^{i-1} d(\rho\xi_i) \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

$$= \sum_{i=1}^n (-1)^{i-1} \Big(\sum_{j=1}^n \frac{\partial(\rho\xi_i)}{\partial x_j} dx_j \Big) \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

$$= \sum_{i=1}^n (-1)^{i-1} \frac{\partial(\rho\xi_i)}{\partial x_i} dx_i \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

$$= \Big(\sum_{i=1}^n \frac{\partial(\rho\xi_i)}{\partial x_i} \Big) dx_1 \wedge \dots \wedge dx_n .$$

Άρα $div_\omega \xi = \dfrac{1}{\rho} \Big(\displaystyle\sum_{i=1}^n \frac{\partial(\rho\xi_i)}{\partial x_i} \Big)$

3.3. Θεώρημα (Τύπος του Liouville). Αν ξ είναι ένα C^2 διανυσματικό πεδίο στην προσανατολισμένη πολλαπλότητα (M,ω) με ροή $(\varphi_t)_{t\in\mathbb{R}}$, τότε

$$\det_\omega D\varphi_t(x) = \exp \int_0^t div_\omega \xi\, (\varphi_s(x))\, ds \quad \text{για κάθε } x \in M$$

Απόδειξη. Για κάθε $t_0 \in \mathbb{R}$ έχουμε:

$$\Big(\frac{d}{dt} \det_\omega D\varphi_t(x)\Big|_{t=t_0}\Big) \omega_{\varphi_{t_0}(x)} = \lim_{h\to 0} \frac{1}{h}\Big[\det_\omega D\varphi_{t_0+h}(x) - \det_\omega D\varphi_{t_0}(x)\Big]\omega_{\varphi_{t_0}(x)}$$

$$= \lim_{h\to 0} \frac{1}{h}\Big[\det_\omega D(\varphi_{t_0}\circ\varphi_h)(x) - \det_\omega D\varphi_{t_0}(x)\Big]\omega_{\varphi_{t_0}(x)}$$

$$= \lim_{h\to 0} \frac{1}{h}\Big[\det_\omega \big(D\varphi_h(\varphi_{t_0}(x)) \cdot D\varphi_{t_0}(x)\big) - \det_\omega D\varphi_{t_0}(x)\Big]\omega_{\varphi_{t_0}(x)}$$

$$= \lim_{h\to 0} \frac{1}{h}\left[\det_\omega D\varphi_h(\varphi_{t_0}(x)) - 1\right] \det_\omega D\varphi_{t_0}(x)\cdot \omega_{\varphi_{t_0}(x)}$$

$$= \det_\omega D\varphi_{t_0}(x)\cdot \lim_{h\to 0}\frac{1}{h}\left[\det_\omega D\varphi_h(\varphi_{t_0}(x))\,\omega_{\varphi_{t_0}(x)} - \omega_{\varphi_{t_0}(x)}\right]$$

$$= \det_\omega D\varphi_{t_0}(x)\cdot \lim_{h\to 0}\frac{1}{h}\left[\left(D\varphi_h(\varphi_{t_0}(x))\right)^*\omega_{\varphi_{t_0+h}(x)} - \omega_{\varphi_{t_0}(x)}\right]$$

$$= \det_\omega D\varphi_{t_0}(x)\cdot (L_\xi\omega)_{\varphi_{t_0}(x)} = \det_\omega D\varphi_{t_0}(x)\cdot \operatorname{div}_\omega \xi(\varphi_{t_0}(x))\cdot \omega_{\varphi_{t_0}(x)}$$

Αρα $\displaystyle \frac{d}{dt}\det_\omega D\varphi_t(x)\Big|_{t=t_0} = \operatorname{div}_\omega\xi(\varphi_{t_0}(x))\cdot \det_\omega D\varphi_{t_0}(x)$ για κάθε $t_0\in\mathbb{R}$

Αυτό σημαίνει ότι η συνάρτηση $\psi(t) = \det_\omega D\varphi_t(x)$, $t\in\mathbb{R}$ είναι λύση της διαφορικής εξίσωσης χωριζομένων μεταβλητών

$$\psi'(t) = \operatorname{div}_\omega\xi(\varphi_t(x))\cdot\psi(t)$$

Δεδομένου ότι $\psi(0) = \det_\omega I_n = 1$, έχουμε $\psi(t) = \exp\int_0^t \operatorname{div}_\omega\xi(\varphi_s(x))ds$.

3.3. Πόρισμα Η ροή ενός C^1 διανυσματικού πεδίου ξ στην προσανατολισμένη πολλαπλότητα (M,ω) διατηρεί το μέτρο μ_ω τότε και μόνον τότε όταν

$$\operatorname{div}_\omega\xi = 0$$

Απόδειξη Από τα προηγούμενα, η ροή του ξ διατηρεί το μέτρο μ_ω τότε και μόνον τότε όταν $\det_\omega D\varphi_t(x) = 1$ για κάθε $t\in\mathbb{R}$ και $x\in M$. Από τον τύπο του Liouville αυτό συμβαίνει τότε και μόνον τότε όταν

$$\int_0^t \operatorname{div}_\omega\xi(\varphi_s(x))ds = 0 \quad \text{για κάθε } t\in\mathbb{R} \text{ και } x\in M$$

Ισοδύναμα $\operatorname{div}_\omega\xi = 0$.

3.4. Παράδειγμα Εστω σωματίδιο μάζας m και φορτίου q κινείται στον $\mathbb{R}^3$ υπό την επίδραση ηλεκτρομαγνητικού πεδίου ηλεκτρικής έντασης E και μαγνητικής έντασης B. Η κίνηση τότε περιγράφεται από την εξίσωση Lorentz

$$m\frac{dv}{dt} = q\,(E + v\times B)$$

όπου v είναι η ταχύτητα, που είναι ισοδύναμη με την 1ης τάξεως διαφορική εξίσωση στον $\mathbb{R}^6$, που είναι ο χώρος φάσεων,

$$x' = v$$
$$v' = \frac{q}{m}(E + v \times B)$$

Στο $\mathbb{R}^6$ θεωρούμε τον κανονικό προσανατολισμό $dx_1 \wedge \dots \wedge dx_6$, που ορίζει το μέτρο Lebesgue. Το αντίστοιχο διανυσματικό πεδίο της εξίσωσης έχει συντεταγμένες $\xi_i = v_i$, $1 \le i \le 3$ και $\xi_j = \frac{q}{m}(E_j + (v \times B)_{j-3})$, $4 \le j \le 6$, όπου $E = (E_1, E_2, E_3)$, $B = (B_1, B_2, B_3)$, $v = (v_1, v_2, v_3)$ και

$$v \times B = (v_2 \cdot B_3 - v_3 \cdot B_2,\ v_3 \cdot B_1 - v_1 B_3,\ v_1 B_2 - v_2 B_1)$$

Συνεπώς

$$\frac{m}{q} \operatorname{div} \xi = \frac{\partial}{\partial v_1}(E_1 + v_2 B_3 - v_3 B_2) + \frac{\partial}{\partial v_2}(E_2 + v_3 B_1 - v_1 B_3) + \frac{\partial}{\partial v_3}(E_3 + v_1 B_2 - v_2 B_1) = 0$$

Άρα η ροή της διαφορικής εξίσωσης κίνησης διατηρεί το μέτρο Lebesgue στο $\mathbb{R}^6$.

<u>3.5. Συστήματα του Hamilton.</u> Έστω (M, ω) μία συμπλεκτική $2n$-πολλαπλότητα δηλαδή η ω είναι μία 2-μορφή στην M, μη-εκφυλισμένη και κλειστή. Το πρώτο σημαίνει ότι αν $x \in M$ και $\omega_x(u, v) = 0$ για κάθε $v \in T_x M$, τότε $u = 0$.
Για παράδειγμα η $(\mathbb{R}^{2n}, \omega)$, όπου $\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$, είναι συμπλεκτική, όπου θεωρούμε συντεταγμένες $(q_1, \dots, q_n, p_1, \dots, p_n)$.

Σύμφωνα με το Θεώρημα του Darboux κάθε $x \in M$ έχει μία ανοικτή περιοχή V για την οποία υπάρχει μία C^∞ αμφιδιαφόριση $h: V \to \mathbb{R}^{2n}$, ώστε $h^*\left(\sum_{i=1}^{n} dq_i \wedge dp_i\right) = \omega|V$. Απ' αυτό προκύπτει ότι η $\Omega = \omega \wedge \dots \wedge \omega = \omega^n$ δεν μηδενίζεται πουθενά στο M και συνεπώς ορίζει έναν προσανατολισμό στην M. Επειδή η ω είναι μη-εκφυλισμένη, για κάθε C^1 συνάρτηση $H: M \to \mathbb{R}$ υπάρχει ένα μοναδικά ορισμένο συνεχές διανυσματικό πεδίο (συμπλεκτική gradient) $\nabla^\# H$, ώστε $dH(x) v = \omega_x(\nabla^\# H(x), v)$ για κάθε $v \in T_x M$, $x \in M$. Ισοδύναμα, $dH = i_{\nabla^\# H}\, \omega$. Σε τοπικές συντεταγμένες Darboux το $\nabla^\# H$ έχει τύπο

$$\nabla^\# H = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n}\right)$$

Πράγματι για κάθε $(v_1, \dots, v_n, u_1, \dots, u_n) \in \mathbb{R}^{2n}$ έχουμε

$$\sum_{i=1}^{n} dq_i\wedge dp_i\left(\nabla^\# H, \begin{pmatrix}v_1\\ \vdots\\ u_n\end{pmatrix}\right) = \sum_{i=1}^{n} dq_i\begin{pmatrix}\frac{\partial H}{\partial p_1}\\ \vdots\\ -\frac{\partial H}{\partial q_n}\end{pmatrix} dp_i\begin{pmatrix}v_1\\ \vdots\\ u_n\end{pmatrix} - dq_i\begin{pmatrix}v_1\\ \vdots\\ u_n\end{pmatrix} dp_i\begin{pmatrix}\frac{\partial H}{\partial p_1}\\ \vdots\\ -\frac{\partial H}{\partial q_n}\end{pmatrix}$$

$$= \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\cdot u_i - \left(-\frac{\partial H}{\partial q_i}\right)\cdot v_i = dH\begin{pmatrix}v_1\\ \vdots\\ u_n\end{pmatrix}.$$

Η H είναι ένα πρώτο ολοκλήρωμα του $\nabla^\# H$, αφού

$$dH(x)(\nabla^\# H(x)) = \omega_x(\nabla^\# H(x), \nabla^\# H(x)) = 0$$

και το $\nabla^\# H$ διατηρεί αναλλοίωτη την ω γιατί

$$L_{\nabla^\# H}\omega = (d\circ i_{\nabla^\# H})\omega + (i_{\nabla^\# H}\circ d)\omega = d(i_{\nabla^\# H} H) = d(dH) = 0$$

Έτσι αν $(\varphi_t)_{t\in\mathbb{R}}$ είναι η ροή του $\nabla^\# H$ έχουμε $\varphi_0^*\omega=\omega$, αφού $\varphi_0 = id_M$ και για κάθε $t_0\in\mathbb{R}$, $\frac{d}{dt}\varphi_t^*\omega\big|_{t=t_0} = \varphi_{t_0}^*(L_{\nabla^\# H}\omega) = 0$. Επομένως $\varphi_t^*\omega=\omega$ για κάθε $t\in\mathbb{R}$. Προκύπτει ότι $\varphi_t^*\underline{\Omega} = \underline{\Omega}$ για κάθε $t\in\mathbb{R}$.

Ένα C^1 διανυσματικό πεδίο ξ στην συμπλεκτική πολλαπλότητα (M,ω) λέγεται <u>τοπικά Hamiltonian</u> αν κάθε $x\in M$ έχει μία ανοιχτή περιοχή V για την οποία υπάρχει μία C^2 συνάρτηση $H: V\to\mathbb{R}$ ώστε $\nabla^\# H = \xi|V$, δηλαδή $i_\xi\omega|V = dH$.

<u>3.6. Πρόταση</u> Ένα C^1 διανυσματικό πεδίο ξ είναι τοπικά Hamiltonian τότε και μόνο τότε όταν $d\,i_\xi\omega = 0$

<u>Απόδειξη</u> Αν $i_\xi\omega|V = dH$, τότε $(d\,i_\xi\omega)|V = d(i_\xi\omega|V) = d(dH) = 0$.
Αντίστροφα αν $d\,i_\xi\omega = 0$, τότε από το Λήμμα του Poincaré, κάθε σημείο $x\in M$ έχει μία ανοιχτή περιοχή V για την οποία υπάρχει μία C^2 συνάρτηση $H: V\to\mathbb{R}$ ώστε $i_\xi\omega|V = dH$.

Επειδή η H είναι πρώτο ολοκλήρωμα του $\nabla^\# H$, το σύνολο $H^{-1}(c)$, $c\in\mathbb{R}$ είναι αναλλοίωτο από την ροή του $\nabla^\# H$. Συνεπώς η μελέτη του $\nabla^\# H$ περιορίζεται στην μελέτη του περιορισμένου δυναμικού συστήματος στο $H^{-1}(c)$ για κάθε $c\in\mathbb{R}$. Αν το c είναι κανονική τιμή της H, τότε το $H^{-1}(c)$ είναι C^1 υποπολλαπλότητα της M, από το Θεώρημα των πεπλεγμένων συναρτήσεων. Ο προσανατολισμός $\underline{\Omega}$ της M ορίζει έναν φυσικό προσανατολισμό $\underline{\tilde{\Omega}}$ στο $H^{-1}(c)$ με τύπο

$$\underline{\tilde{\Omega}}_x(u_1,\dots,u_{n-1}) = \underline{\Omega}_x(u, u_1,\dots,u_{n-1}), \quad x\in H^{-1}(c).$$

για κάθε $u_1,\dots,u_{n-1}\in T_xH^{-1}(c)$, όπου $u\in T_xM$ με $dH(x)u = 1$.
Ο ορισμός δεν εξαρτάται από το u, γιατί αν θεωρούσαμε ένα δεύτερο $u'\in T_xM$

με $dH(x)u'=1$, τότε $u-u'\in \operatorname{Ker} dH(x) = T_x H^{-1}(c)$ και συνεπώς

$$\Omega_x(u,u_1,\dots,u_{n-1}) - \Omega_x(u',u_1,\dots,u_{n-1}) = \Omega_x(u-u',u_1,\dots,u_{n-1}) = 0 .$$

Η $\tilde{\Omega}$ παραμένει αναλλοίωτη από το περιορισμένο δυναμικό σύστημα στο $H^{-1}(c)$.

Αν $x\in H^{-1}(c)$ και $u_1,\dots,u_{n-1}\in T_x H^{-1}(c)$, έχουμε:

$$(\varphi_t^*\tilde{\Omega})_x(u_1,\dots,u_{n-1}) = \tilde{\Omega}_{\varphi_t(x)}(D\varphi_t(x)u_1,\dots,D\varphi_t(x)u_{n-1})$$

$$= \Omega_{\varphi_t(x)}(u, D\varphi_t(x)u_1,\dots,D\varphi_t(x)u_{n-1})$$

όπου $u\in T_{\varphi_t(x)}M$ με $DH(\varphi_t(x))u=1$. Επειδή η φ_t είναι C^1 αμφιδιαφόριση, υπάρχει μοναδικά ορισμένο $u_0\in T_x M$ με $D\varphi_t(x)u_0 = u$. Παραγωγίζοντας την ισότητα $H\circ\varphi_t = H$ προκύπτει

$$DH(\varphi_t(x))\cdot D\varphi_t(x) = DH(x)$$

και συνεπώς $DH(x)u_0 = DH(\varphi_t(x))\cdot D\varphi_t(x)u_0 = DH(\varphi_t(x))u = 1$.

Άρα $(\varphi_t^*\tilde{\Omega})_x(u_1,\dots,u_{n-1}) = \Omega_{\varphi_t(x)}(D\varphi_t(x)u_0, D\varphi_t(x)u_1,\dots,D\varphi_t(x)u_{n-1})$

$$= (\varphi_t^*\Omega)_x(u_0,u_1,\dots,u_{n-1})$$

$$= \Omega_x(u_0,u_1,\dots,u_{n-1}) = \tilde{\Omega}_x(u_1,\dots,u_{n-1})$$

αφού η Ω είναι αναλλοίωτη από την φ_t.

4. ΕΡΓΟΔΙΚΟΤΗΤΑ

Εστω $(X, \mathcal{A}, \mu)$ ενας χωρος πιθανοτητας κ' $T: X \to X$ ενας ενδομορφισμός. Αν υπαρχει $A \in \mathcal{A}$ τετοιο ωστε $T^{-1}A = A$ τοτε ισχύει και $T^{-1}A^c = A^c$ (οπου $A^c = X - A$) κ' η μελετη του T μπορει να αναχθει στη μελέτη των $T|_A$ κ' $T|_{A^c}$ ξεχωριστά. Απο πιθανοθεωρητικης σκοπιας η μελετη και των δύο $T|_A$ κ' $T|_{A^c}$ ειναι απαραίτητη μονο αν $\mu(A) > 0$ κ' $\mu(A^c) > 0$. Αντίθετα, αν $\mu(A) \in \{0,1\}$ (δηλ. $\mu(A) = 0$ ή $\mu(A^c) = 0$) τοτε αρκεί να μελετήσουμε μονο μια εκ των $T|_A$, $T|_{A^c}$. Συστήματα για τα οποία: $A \in \mathcal{A}$, $T^{-1}A = A \Rightarrow \mu(A) \in \{0,1\}$ ειναι λοιπόν αναγωγα υπο την παραπάνω εννοια και ονομάζονται εργοδικά.

4.1 Ορισμος: Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας. Ενας ενδομορφισμος $T: X \to X$ λεγεται εργοδικός αν

$$A \in \mathcal{A}, \ \mu(A \triangle T^{-1}A) = 0 \Rightarrow \mu(A) \in \{0,1\}.$$

Μια μετρήσιμη ημιροή $\varphi: \mathbb{R}_+ \times X \to X$ λεγεται εργοδική αν

$$A \in \mathcal{A}, \ \mu(A \triangle \varphi_t^{-1}(A)) = 0 \ \ \forall t \in \mathbb{R}_+ \Rightarrow \mu(A) \in \{0,1\}.$$

Παρατήρηση: Αν ο $T: X \to X$ ειναι ενας αυτομορφισμος τοτε ο T ειναι εργοδικός αν ειναι εργοδικός σαν ενδομορφισμος του X. Αυτο παλι ειναι ισοδύναμο με το οτι και ο T και ο T^{-1} ειναι εργοδικοί σαν ενδομορφισμοί του X. Πραγματι εστω οτι T ειναι αυτομορφισμός κ' T ειναι εργοδικος σαν ενδομορφισμός του X. Εστω $A \in \mathcal{A}$ με $\mu(TA \triangle A) = 0$. Τοτε επειδή $A \triangle T^{-1}A = T^{-1}(TA \triangle A)$ επεται οτι $\mu(A \triangle T^{-1}A) = 0$ κ' ετσι $\mu(A) \in \{0,1\}$.

Ομοια, αν $\varphi: \mathbb{R} \times X \to X$ ειναι μια μετρήσιμη ροή, τοτε η φ ειναι εργοδική αν η ημιροή $\varphi|_{\mathbb{R}_+ \times X}$ ειναι εργοδική. Αυτο παλι ισοδυναμει με

$$A \in \mathcal{A}, \ \mu(\varphi_t(A) \triangle A) = 0 \ \ \forall t \in \mathbb{R} \Rightarrow \mu(A) \in \{0,1\}.$$

4.2 Προταση: Εστω $(X, \mathcal{A}, \mu)$ ενας χωρος πιθανοτητας κ' $T: X \to X$ ενας ενδομορφισμός. Τα εξης ειναι ισοδύναμα:

(i) T εργοδικός

(ii) $A \in \mathcal{A}$, $T^{-1}A = A \Rightarrow \mu(A) \in \{0,1\}$.

(iii) Αν $A, B \in \mathcal{A}$ και $\mu(A) > 0$, $\mu(B) > 0$ τοτε υπαρχει $n \in \mathbb{N}$ ωστε $\mu(T^{-n}A \cap B) > 0$.

ΑΠΟΔΕΙΞΗ: (i) $\Rightarrow$ (ii) Προφανές.

(ii) $\Rightarrow$ (i) : Εστω $A \in \mathcal{A}$ με $\mu(A \triangle T^{-1}A) = 0$. Θα βρουμε $B \in \mathcal{A}$ με $T^{-1}B = B$ κ' $\mu(A \triangle B) = 0$. Τοτε απο το (ii) θα εχουμε $\mu(B) \in \{0,1\}$ κ' ετσι $\mu(A) \in \{0,1\}$ κ' θα εχουμε τελειώσει. ($\mu(A \triangle B) = 0 \Rightarrow \mu(A) = \mu(B)$.)

Θετομε $B = \bigcap_{n=0}^{\infty} \bigcup_{k \geq n}^{\infty} T^{-k}A$. Το B αποτελείται απο ακριβώς εκεινα τα $x \in X$ για τα οποία το σύνολο $\{k \in \mathbb{N}_0 : T^k x \in A\}$ ειναι απειρο. Αυτο δείχνει αμεσως οτι $T^{-1}B = B$.

Τωρα

$$T^{-k}A \triangle A \subset \bigcup_{i=0}^{k-1} T^{-(i+1)}A \triangle T^{-i}A = \bigcup_{i=0}^{k-1} T^{-i}(T^{-1}A \triangle A)$$

κ' αρα

$$\mu(T^{-k}A \triangle A) \leq \sum_{i=0}^{k-1} \mu(T^{-i}(T^{-1}A \triangle A)) = k\,\mu(T^{-1}A \triangle A) = 0 ,$$

για ολα τα $k \in \mathbb{N}$. Επισης

$$\left(\bigcup_{k \geq n}^{\infty} T^{-k}A\right) \triangle A \subset \bigcup_{k \geq n}^{\infty} T^{-k}A \triangle A ,$$

αρα

$$\mu\left(\left(\bigcup_{k \geq n}^{\infty} T^{-k}A\right) \triangle A\right) = 0 \qquad \forall n \in \mathbb{N}.$$

Αφου $\bigcup_{k \geq n} T^{-k}A \searrow B$ επεται οτι

$$\mu(B \triangle A) = \lim_{n \to \infty} \mu\left(\left(\bigcup_{k \geq n}^{\infty} T^{-k}A\right) \triangle A\right) = 0.$$

(i) $\Rightarrow$ (iii) Εστω $A, B \in \mathcal{A}$ με $\mu(A) > 0$, $\mu(B) > 0$ αλλα $\mu(T^{-n}A \triangle B) = 0$ $\forall n \in \mathbb{N}$. Θετουμε $C = \bigcup_{n=1}^{\infty} T^{-n}A$ και εχουμε $\mu(C \cap B) \leq \sum_{n=1}^{\infty} \mu(T^{-n}A \triangle B) = 0$. Επισης $T^{-1}C = \bigcup_{n=2}^{\infty} T^{-n}A \subset \bigcup_{n=1}^{\infty} T^{-n}A = C$ και ετσι $\mu(C) \in \{0,1\}$ διοτι $\mu(C \triangle T^{-1}C) = \mu(C) - \mu(T^{-1}C) = 0$. Αφου ομως $\mu(C) \geq \mu(T^{-1}A) = \mu(A) > 0$ επεται οτι $\mu(C) = 1$. Τοτε ομως πρεπει κ' $\mu(C \cup B) = 1$ και αρα

$$1 = \mu(C \cup B) = \mu(C) + \mu(B) - \mu(C \cap B) = 1 + \mu(B)$$

που ομως αντικρούει το $\mu(B) > 0$. Αρα πρεπει $\mu(T^{-n}A \triangle B) > 0$ για καποιο n.

(iii) $\Rightarrow$ (ii) Εστω $A \in \mathcal{A}$ αναλλοίωτο, δηλ. $T^{-1}A = A$ με $0 < \mu(A) < 1$. Απο το (iii) κ' αφου $\mu(A) > 0$, $\mu(A^c) = 1 - \mu(A) > 0$, υπαρχει $n \in \mathbb{N}$ ωστε $\mu(T^{-n}A \cap A^c) > 0$. Αυτο ομως ειναι αδύνατο αφου $T^{-n}A = A$ για ολα τα n. ∎

4.3 Ορισμος: Εσω $(X,\mathcal{A},\mu)$ ενας χωρος πιθανότητας και $f: X\to\mathbb{R}$ μετρησιμη συναρτηση. Η f λεγεται αναλλοίωτη για εναν ενδομορφισμό $T: X\to X$ αν $f=f\circ T$. Αν $f=f\circ T$ μ-σ.π. (μ σχεδον παντου) τοτε η f λεγεται μ σχεδον αναλλοίωτη.

Η f ειναι αναλλοίωτη ως προς μια ημιροή $\{\phi_t : t\in\mathbb{R}_+\}$ του X αν $f=f\circ T$ $\forall t\geq 0$. Αν $\mu(x\in X: f(x)=f(\varphi_t(x)))=1$ $\forall t\in\mathbb{R}_+$ τοτε η f ειναι μ-σχεδον αναλλοίωτη.

4.4 Προταση: Εσω $(X,\mathcal{A},\mu)$ ενας χωρος πιθανότητας. Αν $T: X\to X$ ειναι ενδομορφισμος ή $\{\phi_t\}_{t\geq 0}$ μια ημιροή τοτε τα εξής ειναι ισοδύναμα:

(i) Το σύστημα ειναι εργοδικό.

(ii) Καθε μετρήσιμη μ-σχεδον αναλλοίωτη συναρτηση ειναι σταθερή μ-σ.π.

(iii) Καθε μετρήσιμη μ-σχεδον αναλλοίωτη συναρτηση στον $L^2(\mu)$ ειναι σταθερή μ-σ.

ΑΠΟΔΕΙΞΗ: (i)$\Rightarrow$(ii): Εστω οτι η $f: X\to\mathbb{R}$ ειναι μετρήσιμη κ' μ-σχεδον αναλλοίωτη. Ορίζουμε

$$A(k,n):=\{x\in X:\ k2^{-n}\leq f(x)<(k+1)2^{-n}\}\qquad (k\in\mathbb{Z},\ n\in\mathbb{N}).$$

Αφου f ειναι μ σχεδον αναλλοιωτη εχουμε οτι για καθε k και n

$$\mu(A(k,n)\,\triangle\, T^{-1}A(k,n))=0\qquad \text{(για ενδομορφισμό)}$$

$$\mu(A(k,n)\,\triangle\,\phi_t^{-1}(A(k,n)))=0\quad \forall t\in\mathbb{R}_+\qquad \text{(για ημιροή)}$$

Απο την εργοδικότητα του συστήματος $\mu(A(k,n))\in\{0,1\}$ $\forall (k,n)\in\mathbb{Z}\times\mathbb{N}$. Αφου για καθε $n\in\mathbb{N}$ η οικογενεια $\{A(k,n)\}_{k\in\mathbb{Z}}$ ειναι μια διαμέριση του X πρεπει να υπαρχει $k_n\in\mathbb{Z}$ ωστε $\mu(A(k_n,n))=1$. Θετουμε $Y=\bigcap_{n=1}^{\infty}A(k_n,n)$. Τοτε $\mu(Y)=1$, ενω η f ειναι σταθερή πανω στον Y αφου για οποιαδήποτε $x,y\in Y$ πρεπει $|f(x)-f(y)|\leq 2^{-n}$ για ολα τα n.

(ii)$\Rightarrow$(iii): Προφανες.

(iii)$\Rightarrow$(i): Εσω $A\in\mathcal{A}$ με $\mu(T^{-1}A\,\triangle\, A)=0$ (ή $\mu(\phi_t^{-1}(A)\,\triangle\, A)=0$ $\forall t\in\mathbb{R}_+$). Αν $\mathbb{1}_A$ ειναι η χαρακτηριστική συναρτηση του A τότε $\mathbb{1}_A\in L^2(\mu)$ κ' η $\mathbb{1}_A$ ειναι μ-σχεδον αναλλοίωτη. Επεται οτι η $\mathbb{1}_A$ ειναι σταθερή μ-σ.π. που σημαίνει ακριβώς οτι $\mu(A)\in\{0,1\}$. ■

Παρατηρήσεις: (i) Το (i)$\Rightarrow$(ii) αποδεικνύεται και ως εξής:

$$\{x\in X: f(x)\leq a\}\,\triangle\,\{x\in X: f(T(x))\leq a\}\subset\{x\in X: f(x)\neq f(T(x))\}\quad\forall a\in\mathbb{R}$$

και αφού το δεύτερο έχει μέτρο 0 και το πρώτο έχει μέτρο 0. Απο την εργοδικότητα $\mu(x\in X: f(x)\leq a) \in \{0,1\}$, $\forall a\in\mathbb{R}$. Τωρα η συναρτηση $g:\mathbb{R}\to[0,1]$ με τύπο $g(a):=\mu(x\in X: f(x)\leq a)$, ειναι μη φθίνουσα, συνεχής απο δεξιά κ' $\lim_{a\to-\infty} g(a)=0$, $\lim_{a\to+\infty} g(a)=1$. Επίσης $\mu(x\in X: f(x)=a) = g(a)-g(a-)$. Αφου $g(a)\in\{0,1\}$ επεται οτι για καποια $a\in\mathbb{R}$, $\mu(x\in X: f(x)=a)=1$.

Ομοια αποδεικνύεται το (i) $\Rightarrow$ (ii) και για ημιροές.

(ii) Για ενδομορφισμούς T, το σύστημα $(X,\mathcal{A},\mu,T)$ ειναι εργοδικο αν κ' μονο αν καθε αναλλοίωτη (παντού) συναρτηση ειναι σταθερή μ-σ.π. αν κ' μόνο αν καθε αναλλοίωτη συναρτηση στον $L^2(\mu)$ ειναι σταθερή μ-σ.π. Η αποδειξη ειναι ιδια με την απόδειξη της Προτασης 4.4 χρησιμοποιόντας την ισοδυναμία (i) $\Leftrightarrow$ (ii) της Πρότασης 4.2.

Γενικά για ενδομορφισμους ισχύει το εξής: για καθε μετρήσιμη μ-σχεδον αναλλοίωτη $f: X\to\mathbb{R}$ υπαρχει αναλλοίωτη (παντου) $g: X\to\mathbb{R}$ με $\mu(x\in X: f(x)=g(x))=1$ κ' g φυσικά μετρήσιμη. (Απλως θετουμε $g=f$ στο $\bigcap_{k=0}^{\infty} T^{-k}\{x\in X: f(x)=f\circ T(x)\}$ – τομή απο $-\infty$ εως $+\infty$ για αυτομορφισμούς – κ' $g=0$ αλλού.) Τετοια συναρτηση υπαρχει κ' στην περιπτωση της ημιροής με τη διαφορα οτι η g αυτή τη φορά ειναι μονο μετρήσιμη στην πληρωση του χωρου κ' οχι κατ' αναγκη μετρήσιμη ως προς τη σ-αλγεβρα $\mathcal{A}$. (Τουλαχιστον εμεις μονο τετοια g ξερουμε να φτιαχνουμε.)

4.5 Παραδειγμα: (Αρρητες στροφές στον κύκλο) Εστω $K=\{z\in\mathbb{C}: |z|=1\}$ κ' m κανονικοποιημενο μέτρο Lebesgue στον K ($m(K)=1$). Το m ειναι το μετρο Haar της αβελιανης συμπαγούς ομαδας K. Αρα καθε στροφή $T_a(z)=az$ (οπου $a\in K$ σταθερό) διατηρει το m. Θα αποδείξουμε οτι το σύστημα $(K,\mathcal{B}(K),m,T_a)$ ειναι εργοδικό αν κ' μονο αν το a δεν ειναι ρίζα της μονάδας, δηλ. $a^n\neq 1$ $\forall n\in\mathbb{Z}-\{0\}$. ($\mathcal{B}(K)$ τα Borel υποσύνολα του K). Παρατηρουμε οτι $a^n\neq 1$ $\forall n\neq 0$ αν κ' μονο αν $a=e^{2\pi i\theta}$ για θ αρρητο· ετσι κ' το ονομα "αρρητες στροφές".

Αν $a^p=1$ για κάποιο $p\neq 0$ τοτε η συναρτηση $f(z)=z^p$ ειναι αναλλοίωτη ($f\circ T_a(z)=f(az)=a^pz^p=z^p=f(z)$) αλλα προφανώς οχι σταθερή.

Αντίστροφα εστω οτι $a^p\neq 1$ $\forall p\neq 0$ κ' εστω $\theta\in[0,2\pi)$ ωστε $a=e^{i\theta}$. Εστω $f\in L^1(\mu)$ αναλλοίωτη. Εστω $\hat{f}_n=(2\pi)^{-1}\int_0^{2\pi} f(e^{it})e^{-int}dt$ $(n\in\mathbb{Z})$ οι συντελεστές Fourier της f. Τοτε οι συντελεστές της $f\circ T_a$ ειναι

$$\widehat{(f\circ T_a)}_n = \frac{1}{2\pi}\int_0^{2\pi} f(T_a(e^{it}))e^{-int}dt = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i(t+\theta)})e^{-int}dt =$$

$$= (2\pi)^{-1}\int_0^{2\pi-\theta} f(e^{i(t+\theta)})\,e^{-int}dt + (2\pi)^{-1}\int_{2\pi-\theta}^{2\pi} f(e^{i(t+\theta)})\,e^{-int}dt =$$

$$= (2\pi)^{-1}\int_\theta^{2\pi} f(e^{it})\,e^{-int}dt\; e^{in\theta} + (2\pi)^{-1}\int_0^{\theta} f(e^{it})\,e^{-int}dt\; e^{in(\theta-2\pi)} =$$

$$= e^{in\theta}\hat{f}_n = \alpha^n \hat{f}_n .$$

Αφου $f = f\circ T_\alpha$ πρεπει $(\widehat{f\circ T_\alpha})_n = \hat{f}_n$ δηλ. $\hat{f}_n(1-\alpha^n)=0$ $\forall n\in\mathbb{Z}$. Αρα $\hat{f}_n = 0$ $\forall n\in\mathbb{Z}-\{0\}$ κ' ετσι $f=\hat{f}_0$ (δηλ. σταθερή) m - σ.π.

Παρατηρήσεις: (1) Ενας ισοδύναμος τροπος να πουμε οτι το α δεν ειναι ρίζα της μοναδας ειναι να πούμε οτι το $\{\alpha^n : n\in\mathbb{Z}\}$ ειναι πυκνό στον K. Πραγματι αν $\alpha^n = 1$ για $n\neq 0$ τοτε το $\{\alpha^n : n\in\mathbb{Z}\}$ ειναι πεπερασμενο κ' ετσι δεν γίνεται να ειναι πυκνό. Αντίθετα αν $\alpha^n\neq 1$ $\forall n\neq 0$ τοτε το $\{\alpha^n : n\in\mathbb{Z}\}$ ειναι απειρο κ' αρα υπάρχουν p, q ωστε $0<|\alpha^p-\alpha^q|<\varepsilon$ για οποιοδήποτε δοσμενο $\varepsilon>0$. Τοτε ομως $0<|1-\alpha^{p-q}|<\varepsilon$ κ' το $\{\alpha^{n(p-q)} : n\in\mathbb{Z}\}$ ειναι ε - πυκνό.

Αυτη η διατύπωση μας επιτρεπει να διατυπώσουμε την εξης προταση για συμπαγής ομαδες γενικά:

Προταση: Μια στροφή $x\mapsto \alpha x$ μιας συμπαγούς ομαδας G ειναι εργοδική ως προς το μετρο Haar της ομαδας αν και μονο αν το σύνολο $\{\alpha^n\}_{n=-\infty}^{\infty}$ ειναι πυκνό στο G. (Ετσι αν $x\mapsto\alpha x$ εργοδικός τοτε αναγκαστικά η G ειναι αβελιανή.)

Η αποδειξη μπορει να γίνει οπως κ' στον κύκλο χρησιμοποιοντας θεωρία χαρακτήρων αντι για συνήθη Fourier αναλυση.

(2) Μια απ' ευθείας αποδειξη του οτι το κανονικοποιημένο μετρο Lebesgue ειναι αναλλοίωτο για αρρητες στροφές στον κύκλο μπορεί να δοθει ως εξής: Εστω $\alpha\in K$ με $\alpha^n\neq 1$ $\forall n\in\mathbb{Z}-\{0\}$ και εστω οτι $\alpha=e^{i\theta}$ οπου $\theta\in[0,2\pi)$. Κατ' αρχήν ενα Borel μετρο πιθανότητας μ στον K χαρακτηρίζεται πλήρως απο την ακολουθία $\{\hat{\mu}_n\}$ των συντελεστών Fourier του μετρου : $\hat{\mu}_n := \int_{[0,2\pi)} e^{-int}\tilde{\mu}(dt)$, $n\in\mathbb{Z}$ (οπου $\tilde{\mu}$ το αντιστοιχο μετρο στο $[0,2\pi)$).
Τωρα το μ ειναι αναλλοίωτο αν κ' μόνο αν $\mu = \mu\circ T_\alpha^{-1}$ (το μετρο $\mu\circ T_\alpha^{-1}$ ειναι το $\mu\circ T_\alpha^{-1}(A) = \mu(T^{-1}A)$, $\forall A\in\mathcal{B}(K)$). Ομως $\mu = \mu\circ T_\alpha^{-1} \Leftrightarrow \hat{\mu}_n = (\widehat{\mu\circ T_\alpha^{-1}})_n$ για ολα τα n. Εχουμε

$$(\widehat{\mu\circ T_\alpha})_n = \int_{[0,2\pi)} e^{-int}\,d\tilde{\mu}((t+\theta)\bmod 2\pi) = e^{in\theta}\int_{[0,2\pi)} e^{-int}d\tilde{\mu}(t) = \alpha^n\hat{\mu}_n .$$

Αρα $$\mu\circ T_\alpha^{-1} = \mu \Leftrightarrow (\alpha^n - 1)\hat{\mu}_n = 0 \quad \forall n\in\mathbb{Z}.$$

Αρα το μετρο Lebesgue ειναι αναλλοιωτο. Απο την αλλη αν μ ειναι αναλλοιωτο τοτε πρεπει $\hat{\mu}_n = 0$ $\forall n \in \mathbb{Z}-\{0\}$ και αρα πρεπει $\mu =$ Lebesgue.

Δηλαδη αν α δεν ειναι ριζα της μοναδας τοτε το μετρο Lebesgue ειναι το μονο μετρο που διατηρει η στροφη T_α και αρα το μονο μετρο ως προς το οποιο η T_α ειναι εργοδικη. Τετοια συστηματα ονομαζονται uniquely ergodic.

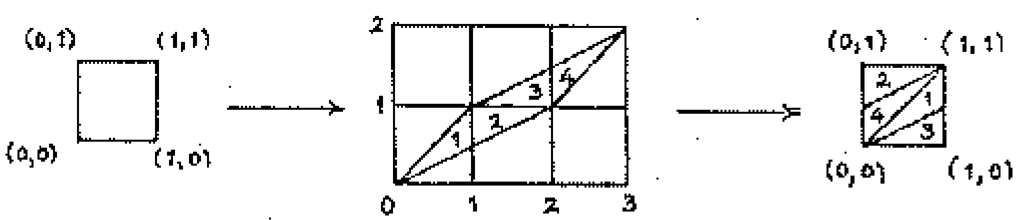
4.6 Παραδειγμα: (Ενδομορφισμοί του n-torus)

Για καθε p, ο μετασχηματισμος $T(z) = z^p$ ειναι ενας ενδομορφισμός του κύκλου K (1-torus) σαν τοπολογική ομαδα. (Δηλ. ο T ειναι ενας αλγεβρικος ομομορφισμος του K στον K, που ειναι συνεχης.) Αφου ο T ειναι επι πρεπει να διατηρεί το μετρο Lebesgue λ_1, που ειναι το μετρο Haar της τοπολογικης ομαδας K. Αποδεικνύεται οτι αυτοί ειναι οι μονοι ενδομορφισμοι του K σαν τοπολογική ομάδα, δηλαδη οι μονοι συνέχεις (αλγέβρικοί) ομομορφισμοί του K στον K. Αν $|p| > 1$ τοτε ο $T(z) = z^p$ ειναι εργοδικός. Πραγματι, εστω $f \in L^2(\lambda_1)$ με αναπτυγμα Fourier $f(z) = \sum_{n=-\infty}^{\infty} \hat{f}_n z^n$. Εστω οτι η f ειναι αναλλοίωτη δηλ. $f = f \circ T$. Τοτε το αναπτυγμα της $f \circ T$ ειναι $f \circ T(z) = f(z^p) = \sum_{n=-\infty}^{\infty} \hat{f}_n z^{np}$ και αρα $(\widehat{f \circ T})_{np} = \hat{f}_n$. Απο την αλλη αφου η f και η $f \circ T$ ταυτίζονται πρεπει να εχουν το ιδιο αναπτυγμα και ετσι $\hat{f}_{np} = (\widehat{f \circ T})_{np} = \hat{f}_n$. Επαναλαμβάνοντας για $f \circ T^k$ παιρνουμε τελικά οτι $\hat{f}_n = \hat{f}_{np} = \hat{f}_{np^2} = \dots$ για καθε $n \in \mathbb{Z}$. Ομως $\sum_{n=-\infty}^{\infty} |\hat{f}_n|^2 = \| f \|_2^2 < +\infty$ και αρα πρεπει $\hat{f}_n = 0$ $\forall n \neq 0$ δηλ. $f = \hat{f}_0$ σχεδον παντού.

Ο n torus μπορει να θεωρηθεί, πολλαπλασιαστικά, σαν το ευθύ γινομενο K^n, η αθροιστικά, σαν $\mathbb{R}^n / \mathbb{Z}^n$. Αν A ειναι ενας πίνακας με στοιχεία $a_{ij} \in \mathbb{Z}$ τοτε ο A επαγει εναν ενδομορφισμο του $\mathbb{R}^n/\mathbb{Z}^n$ (ενδομορφισμο τοπολογικης ομαδας):

$$[0,1)^n \ni x \mapsto T(x) = Ax \pmod 1 = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \pmod 1 \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \pmod 1 \end{pmatrix}$$

Για παραδειγμα το "Cat Map" του Arnold: $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ στον 2-torus $\mathbb{R}^2/\mathbb{Z}^2$:



(Αυτος μαλιστα ειναι αυτομορφισμός και ο λογος ειναι οτι $|\det A| = 1$.)

Αν $\det A \neq 0$ τοτε ο T ειναι επιμορφισμος (αφου η απεικονιση $x \mapsto Ax$ στον $\mathbb{R}^n$ ειναι επι σ'αυτη την περίπτωση) ενω αν $|\det A| = 1$ τοτε ο T ειναι αυτομορφισμός.

(Οι συνθήκες αυτές είναι ικανές και αναγκαίες ώστε ο T να είναι επιμορφισμός ή αυτομορφισμός αντίστοιχα.) Αποδεικνύεται ότι όλοι οι ενδομορφισμοί της τοπολογικής ομάδας $\mathbb{R}^n/\mathbb{Z}^n$ (δηλ. όλοι οι συνεχείς αλγεβρικοί ομομορφισμοί του $\mathbb{R}^n/\mathbb{Z}$ στον εαυτό του) είναι αυτής της μορφής για κάποιον $n\times n$ πίνακα A με στοιχεία στο $\mathbb{Z}$.

Τώρα $\det A \neq 0 \Rightarrow T(\mathbb{R}^n/\mathbb{Z}^n)=\mathbb{R}^n/\mathbb{Z}^n \Rightarrow T$ διατηρεί το μέτρο Haar της ομάδας, που φυσικά είναι το μέτρο Lebesgue λ_n στο $[0,1)^n$. Δηλ. αν $\det A\neq 0$ τότε ο T είναι ένας ενδομορφισμός του χώρου πιθανότητας $(\mathbb{R}^n/\mathbb{Z}^n, \mathcal{B}(\mathbb{R}^n/\mathbb{Z}^n), \lambda_n)$, ενδομορφισμός τώρα με την μετροθεωρητική έννοια της Σελίδας 1 αυτών των σημειώσεων.

Πρόταση: Ο T είναι εργοδικός αν και μόνο αν καμμία ιδιοτιμή του A δεν είναι ρίζα της μονάδας.

(Η πρόταση αυτή φυσικά εμπεριέχει την περίπτωση $n=1$, όπου ήδη είδαμε ότι όλοι οι ενδομορφισμοί $z\mapsto z^p$, για $|p|>1$, είναι εργοδικοί. Σε προσθετικό συμβολισμό ο 1-torus είναι ο $\mathbb{R}/\mathbb{Z}$ και ο ενδομορφισμός $z\mapsto z^p$ του K αντιστοιχεί στον $x\mapsto px \pmod 1$ του $\mathbb{R}/\mathbb{Z}$.)

ΑΠΟΔΕΙΞΗ: Πρώτα έστω ότι ο A έχει μια ιδιοτιμή l με $l^m=1$ για κάποιο $m\neq 0$. Τότε το 1 είναι ιδιοτιμή του A^m. Έστω k ο μικρότερος θετικός ακέραιος ώστε ο A^k να έχει ιδιοτιμή το 1. Τότε επειδή το 1 είναι και ιδιοτιμή του $(A^k)^T$ υπάρχει διάνυσμα $u\in\mathbb{R}^n$ ώστε $u^TA^k=u^T$ και επειδή ο A έχει στοιχεία στο $\mathbb{Z}$ μπορούμε να επιλέξουμε το $u\in\mathbb{Z}^n$. Τότε ο τύπος

$$f(x)=\exp(2\pi i u^Tx)+\exp(2\pi i u^TAx)+\dots+\exp(2\pi i u^TA^{k-1}x)$$

ορίζει μια (μιγαδική) συνάρτηση του $\mathbb{R}^n/\mathbb{Z}^n$. Η συνάρτηση αυτή είναι προφανώς αναλλοίωτη αλλά όχι σταθερή. Δεν είναι σταθερή γιατί το σύνολο $\{\exp(2\pi i u^TA^jx) : 0\leq j<k\}$ αποτελείται από ορθογώνιες στον $L^2(\lambda_1)$ συναρτήσεις (όπου $\langle h,g\rangle=\int h\bar{g}$) και άρα το σύνολο αυτό επαυξημένο κατά μια σταθερή συνάρτηση θα αποτελείται επίσης από ορθογώνιες ανά δύο συναρτήσεις και θα πρέπει να είναι γραμμικά ανεξάρτητο. Άρα από την Πρόταση 4.4 (που ισχύει προφανώς και για μιγαδικές f) ο T δεν μπορεί να είναι εργοδικός.

Για το αντίστροφο προχωρούμε όπως και στην περίπτωση $n=1$. Αν $f\in L^2(\lambda_1)$ είναι αναλλοίωτη ($f=f\circ T$), με ανάπτυγμα Fourier $f(x)=\sum_{a\in\mathbb{Z}^n}\hat{f}(a)\exp(2\pi i a^Tx)$ τότε $f\circ T(x)=\sum_{a\in\mathbb{Z}^n}\hat{f}(a)\exp(2\pi i a^TAx)$ και άρα πρέπει $\hat{f}(a^T)=\hat{f}(a^TA)=\hat{f}(a^TA^2)=\dots$. Όμως το σύνολο $\{a^TA^k\}_k$ είναι άπειρο (διαφορετικά ο A θα είχε για ιδιοτιμή μια ρίζα της μονάδας) και αφού $\sum_{a\in\mathbb{Z}^n}|\hat{f}(a)|^2=\|f\|_2^2<+\infty$ πρέπει $\hat{f}(a)=0$ για κάθε μη μηδενικό $a\in\mathbb{Z}^n$. Επομένως $f=\hat{f}(0,\dots,0)=$ σταθερά σχεδόν παντού.

4.7 Παράδειγμα: (Shift) Έστω $Y=\{0,1,\dots,k-1\}$ και $X=\prod_{-\infty}^{\infty}Y=\{(x_i)_{i\in\mathbb{Z}} : x_i\in Y\ \forall i\in\mathbb{Z}\}$. Έστω $(X,\mathcal{B},\mu,T)$ ο χώρος shift. Δηλαδή $\mathcal{B}$ η Borel σ-αλγεβρα στον X (οπου ο X εχει την τοπολογια γινομενο προερχομενη απο την διακριτη τοπολογια στον Y), $T : X \to X$ η απεικονιση shift, δηλ. αν $x=(x_i)$ τοτε η i συντεταγμενη του Tx ειναι η $(Tx)_i = x_{i+1}$ και $\mu=\mu(p_0,\dots,p_{k-1})$ το μετρο που δινει στον κυλινδρο $\{(x_i)\in X : x_j=l_1,\dots,x_{j+m}=l_m\}$ μετρο $p_{l_1} p_{l_1} \cdots p_{l_m}$ (οπου $p_i \geq 0$, $\sum_{i=0}^{k-1} p_i =1$). Δηλαδη $(X,\mathcal{B},\mu) = \prod_{-\infty}^{\infty} (Y, 2^Y, P)$ (μετροθεωρητικο γινομενο) οπου P το μετρο πιθανοτητας στον Y με $P(\{i\})=p_i$ (η' $P(A)=\sum_{i\in A} p_i$).

Το συστημα $(X,\mathcal{B},\mu,T)$ ειναι εργοδικο.

Πραγματι εστω $\mathcal{A}$ η αλγεβρα των υποσυνολων του X που αποτελειται απο πεπερασμενες ενωσεις κυλινδρων. Τοτε η $\mathcal{A}$ παραγει την σ-αλγεβρα $\mathcal{B}$. Εστω $B\in\mathcal{B}$ με $T^{-1}B=B$. Αφου $\mathcal{A}$ ειναι αλγεβρα που παραγει την $\mathcal{B}$, για καθε $\varepsilon>0$ υπαρχει $A\in\mathcal{A}$ ωστε $\mu(A \triangle B)<\varepsilon$. Ετσι

$$(*) \qquad |\mu(A)-\mu(B)| = |\mu(A\cap B)+\mu(A\setminus B) - \mu(A\cap B) - \mu(B\setminus A)| \leq$$

$$\leq \mu(A\setminus B)+\mu(B\setminus A) = \mu(A\triangle B) < \varepsilon.$$

Τα A, $T^{-n}A$ ειναι ενωσεις κυλινδρων για ολα τα n. Εστω λοιπον n αρκουντως μεγαλο ωστε A και $T^{-n}A$ να "δεσμευουν" διαφορετικες συντεταγμενες. Τοτε επειδη το μ ειναι μετρο γινομενο $\mu(A\cap T^{-n}A)=\mu(A)\mu(T^{-n}A)=\mu(A)^2$ (μ = αναλλοιωτο). Ετσι

$$|\mu(B)-\mu(A)^2| = |\mu(B)-\mu(A\cap T^{-n}A)| \leq \qquad \text{(οπως και στην (*))}$$

$$\leq \mu(B\triangle(A\cap T^{-n}A)) \leq \qquad (B\triangle(A\cap T^{-n}A) \subset (B\triangle A)\cup(B\triangle T^{-n}A))$$

$$\leq \mu(B\triangle A)+\mu(B\triangle T^{-n}A) = \qquad (\text{αφου } B=T^{-1}B)$$

$$= \mu(B\triangle A)+\mu(T^{-n}B\triangle T^{-n}A) =$$

$$= \mu(B\triangle A)+\mu(T^{-n}(B\triangle A)) = 2\mu(B\triangle A) < 2\varepsilon.$$

Επομενως

$$|\mu(B)-\mu(B)^2| \leq |\mu(A)^2-\mu(B)| + |\mu(A)^2-\mu(B)^2| <$$

$$< 2\varepsilon + [\mu(A)+\mu(B)]\,|\mu(A)-\mu(B)| \leq$$

$$\leq 2\varepsilon + 2|\mu(A)-\mu(B)| < 4\varepsilon.$$

Αφου το ε ηταν αυθαιρετο, $\mu(B)=\mu(B)^2$ και αρα $\mu(B)\in\{0,1\}$.

5. ΕΡΓΟΔΙΚΟ ΘΕΩΡΗΜΑ

5.1 Θεωρημα (Εργοδικο Θεωρημα Birkhoff) Εστω $(X,\mathcal{A},\mu)$ ενας χωρος πιθανοτητας και $T: X\to X$ ενας ενδομορφισμός. Τοτε για καθε $f\in L^1(\mu)$

(i) $f^*(x) = \lim\limits_{N\to\infty} \frac{1}{N}\sum\limits_{n=0}^{N-1} f\circ T^n(x)$ υπαρχει για μ-σχεδον καθε x.

(ii) $f^*\circ T = f^*$ $(\mu$-σ.π.$)$.

(iii) $f^*\in L^1(\mu)$.

(iv) $\int_X f^* d\mu = \int_X f\, d\mu$.

5.2 Πορισμα: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενας ενδομορφισμός. Αν το συστημα $(X,\mathcal{A},\mu,T)$ ειναι εργοδικο τοτε για καθε $f\in L^1(\mu)$

$$\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n(x) \to \int_X f\,d\mu \qquad (\mu\text{-σ.π.}).$$

ΑΠΟΔΕΙΞΗ: Το οριο f^* του αριστερου μελους παντα υπαρχει και ειναι σχεδον αναλλοιωτη συναρτηση. Απο την εργοδικοτητα του συστηματος και την Προταση 4.4 το οριο f^* ειναι μια σταθερη συναρτηση μ-σ.π. Απο το (iv) του Θεωρηματος 5.1 πρεπει $f^* = \int_X f\,d\mu$ μ-σ.π. ■

Παρατηρησεις: (i) Τα (i)-(iii) του Θεωρηματος 5.1 ισχυουν και για σ-πεπερασμενους χωρους $(X,\mathcal{A},\mu)$.

(ii) Εστω $\mathcal{A}_I$ και $\mathcal{A}_{AI}$ οι σ-αλγεβρες $\{A\in\mathcal{A}: T^{-1}A = A\}$ και $\{A\in\mathcal{A}: \mu(A\vartriangle T^{-1}A) = 0\}$ αντιστοιχα, δηλ. οι σ-αλγεβρες των αναλλοιωτων και σχεδον αναλλοιωτων συνολων. Τοτε απο την αποδειξη της Προτασης 4.2, για καθε $A\in\mathcal{A}_{AI}$ υπαρχει $A'\in\mathcal{A}_I$ με $\mu(A\vartriangle A') = 0$ και αρα εχουμε για δεσμευμενες μεσες τιμες $E(f\,|\,\mathcal{A}_I) = E(f\,|\,\mathcal{A}_{AI})$ $(\mu$-σ.π.$)$. Η f^* ειναι ακριβως αυτη η δεσμευμενη μεση τιμη. Πραγματι: απο την παραγραφο ακριβως πριν απο το Παραδειγμα 4.5 (σελ. 14) υπαρχει μετρησιμη g με $g = f^*$ μ-σ.π. και $g = g\circ T$ παντου. Τοτε η g ειναι $\mathcal{A}_I$ μετρησιμη. Εφαρμοζοντας το εργοδικο θεωρημα για τον $T|_A$ παιρνουμε $\int_A f^* d\mu = \int_A f\,d\mu$ και ετσι $\int_A g\,d\mu = \int_A f\,d\mu$ για καθε $A\in\mathcal{A}_I$. Αφου $g = f^*\in L^1(\mu)$ επεται οτι $f^* = g = E(f\,|\,\mathcal{A}_I) = E(f\,|\,\mathcal{A}_{AI})$ μ-σ.π.

Συμβολισμός: Για $f: X \to \mathbb{R}$ θα γραφουμε $S_n f = \sum_{i=0}^{n-1} f \circ T^i$, για $n \in \mathbb{N}$, $S_0 f = 0$ και $S_n^* f = \max_{0 \le k \le n} S_n f \geqslant 0$.

5.3 Λημμα (Maximal Εργοδικο Θεωρημα Yosida Kakutani) Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας, $T: X \to X$ ενδομορφισμος και $f \in L^1(\mu)$. Τοτε

$$\int_{\{S_n^* f > 0\}} f \, d\mu \geqslant 0 \qquad \forall n \in \mathbb{N}.$$

ΑΠΟΔΕΙΞΗ: Προφανως $g \in L^1(\mu) \Rightarrow \|g \circ T\|_1 = \|g\| < +\infty$ και $S_n^* f$, $S_n f \in L^1(\mu)$. Για καθε $0 \le k \le n$ εχουμε οτι $S_n^* f \geqslant S_k f$ και αρα $S_n^* f(T(x)) \geqslant S_k f(T(x)) = S_{k+1} f(x) - f(x)$. Ετσι

$$(S_n^* f) \circ T(x) + f(x) \geqslant \max_{1 \le k \le n} S_k f(x) = \qquad \text{οταν } S_n^* f(x) > 0$$

$$= \max_{0 \le k \le n} S_k f(x) = S_n^* f(x).$$

Δηλαδη $(S_n^* f) \circ T + f \geqslant S_n^* f$ στο $A := \{x \in X : S_n^* f(x) > 0\}$ και αρα

$$\int_A f \, d\mu \geqslant \int_A S_n^* f \, d\mu - \int_A (S_n^* f) \circ T \, d\mu =$$

$$= \int_X S_n^* f \, d\mu - \int_A (S_n^* f) \circ T \, d\mu \geqslant$$

$$\geqslant \int_X S_n^* f \, d\mu - \int_X (S_n^* f) \circ T \, d\mu = 0 \qquad \blacksquare$$

(Η παραπανω αποδειξη ειναι του **A. Garsia**.)

5.4 Πορισμα: Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας, $T: X \to X$ ενδομορφισμος και $f \in L^1(\mu)$. Αν

$$B_a := \left\{ x \in X : \sup_{n \geqslant 1} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) > a \right\}$$

τοτε

$$a \, \mu(A \cap B_a) \leqslant \int_{A \cap B_a} f \, d\mu$$

για καθε $A \in \mathcal{A}$ με $A = T^{-1}A$.

ΑΠΟΔΕΙΞΗ: Εστω $A = X$ πρωτα. Θετουμε $g = f - a$ και παρατηρουμε οτι αν $G_n := \{x \in X : S_n^* g(x) > 0\}$ τοτε $G_1 \subset G_2 \subset \dots$ και $B_a = \bigcup_{n=1}^{\infty} G_n$. Απο το Λημμα 5.3 $\int_{G_n} g \, d\mu \geqslant 0$ για καθε n και αρα $\int_{B_a} g \, d\mu = \lim_{n \to \infty} \int_{G_n} g \, d\mu \geqslant 0$, που ειναι το ζητουμενο για $A = X$. Στη γενικη περιπτωση εφαρμοζουμε τα παραπανω στον $T|_A$.

ΑΠΟΔΕΙΞΗ ΕΡΓΟΔΙΚΟΥ ΘΕΩΡΗΜΑΤΟΣ: (i) Ορίζουμε

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

Εστω $E = \{ x\in X : f^*(x) \neq f_*(x)\}$. Θα αποδείξουμε οτι $\mu(E) = 0$. Εχουμε οτι

$$E = \bigcup_{\substack{\alpha,\beta\in\mathbb{Q}\\ \alpha>\beta}} E(\alpha,\beta) \quad \text{οπου} \quad E(\alpha,\beta) = \{ x\in X : f_*(x)<\beta \ , \ f^*(x)>\alpha\}$$

και αρα αρκεί να δείξουμε οτι $\mu(E(\alpha,\beta)) = 0$ οταν $\alpha>\beta$.

Επειδή

$$\frac{1}{n} \left| S_{n+1} f(x) - S_n f(T(x))\right| = \frac{|f(x)|}{n}$$

οι ακολουθίες $\{S_n f(x)\}_{n\in\mathbb{N}}$ και $\{S_n f(T(x))\}_{n\in\mathbb{N}}$ εχουν τα ιδια υπακολουθιακα ορια και αρα

$$f^*\circ T = f^* \quad , \quad f_*\circ T = f_* .$$

Επεται οτι $T^{-1}E(\alpha,\beta) = E(\alpha,\beta)$. Απο το Πορισμα 5.4

$$\alpha\,\mu(E(\alpha,\beta)\cap B_\alpha) \leq \int_{E(\alpha,\beta)\cap B_\alpha} f\, d\mu$$

(B_α οπως στο Πορισμα 5.4) και αφου $E(\alpha,\beta)\cap B_\alpha = E(\alpha,\beta)$

$$\alpha\,\mu(E(\alpha,\beta)) \leq \int_{E(\alpha,\beta)} f\, d\mu .$$

Επαναλαμβανοντας για $-f$ στη θεση της f και $-\alpha$, $-\beta$ στις θεσεις των β,α αντιστοιχα παιρνουμε

$$-\beta\,\mu(E(\alpha,\beta)) \leq \int_{E(\alpha,\beta)} (-f)\, d\mu .$$

Ετσι $(\alpha-\beta)\,\mu(E(\alpha,\beta)) \leq 0$ και αφου $\alpha>\beta$ πρεπει $\mu(E(\alpha,\beta)) = 0$.

(ii) Αποδείχτηκε με το (i).

(iii) Χρησιμοποιουμε το Λημμα Fatou:

$$\int_X |f^*|\, d\mu = \int_X |f_*|\, d\mu = \int_X \underline{\lim}\ n^{-1} |S_n f|\, d\mu \leq$$

$$\leq \underline{\lim}\ n^{-1}\int_X |S_n f|\, d\mu \leq \underline{\lim}\ n^{-1}\sum_{i=0}^{n-1}\int_X |f\circ T^i|\, d\mu = \|f\|_1 < +\infty.$$

(iv) Θα χρησιμοποιησουμε το ορισμα 5.4 παλι. Οριζουμε

$$A(k,n) = \{ x\in X : \frac{k}{n} \leq f^*(x) < \frac{k+1}{n} \} \qquad (k\in\mathbb{Z},\ n\in\mathbb{N}).$$

Τοτε για καθε $\varepsilon > 0$ εχουμε οτι $A(k,n)\cap B_{\frac{k}{n}-\varepsilon} = A(k,n)$ και απο το Πορισμα 5.4

$$\int_{A(k,n)} f\, d\mu \geq \left(\frac{k}{n}-\varepsilon\right)\mu(A(k,n)) \qquad \forall \varepsilon>0$$

που δινει οτι

$$\int_{A(k,n)} f\, d\mu \geq \frac{k}{n}\,\mu(A(k,n)).$$

Απο την αλλη πλευρα (ορισμος των $A(k,n)$)

$$\int_{A(k,n)} f^*\, d\mu \leq \frac{k+1}{n}\,\mu(A(k,n)) \leq \int_{A(k,n)} f\, d\mu + \frac{1}{n}\,\mu(A(k,n)).$$

Αθροιζοντας ως προς ολα τα $k\in\mathbb{Z}$ παιρνουμε $\int_X f^* d\mu \leq \int_X f\, d\mu + \frac{1}{n}$ $(\mu(X)=1)$ για καθε $n\in\mathbb{N}$ και αρα

$$\int_X f^*\, d\mu \leq \int_X f\, d\mu$$

Επαναλαμβανουμε για την $-f$ στη θεση της f και παιρνουμε

$$\int_X (-f_*)\, d\mu = \int_X (-f)^* d\mu \leq \int_X (-f)\, d\mu.$$

Αρα $\int_X f^* d\mu \leq \int_X f\, d\mu \leq \int f_*\, d\mu$ και αφου $f^* = f_*$ (μ-σ.π.) πρεπει $\int_X f\, d\mu = \int_X f^*\, d\mu$. ∎

5.5 Πορισμα (L^p - Εργοδικο Θεωρημα von Neumann)

Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας, $T: X\to X$ ενδομορφισμος και $f\in L^p(\mu)$, $1\leq p<\infty$. Τοτε υπαρχει $f^*\in L^p(\mu)$ με $f^* = f^*\circ T$ (μ-σ.π.) τετοια ωστε

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{i=0}^{n-1} f\circ T^i - f^* \right\|_p = 0.$$

Παρατήρηση: Αφου το μ ειναι μετρο πιθανοτητας, $f\in L^p(\mu)$ για $p\geq 1$, συνεπαγεται $f\in L^1(\mu)$ και αρα απο το θεωρημα Birkhoff υπαρχει και το σχεδον παντου οριο $f^*_{\sigma.\pi.}$ των $n^{-1}\sum_{i=0}^{n-1} f\circ T^i$. Προφανως $f^*_{\sigma.\pi.} = f^*_{L^p}$ (μ-σ.π.) αφου και τα δυο ειδη συγκλισης συνεπαγονται συγκλιση κατα μετρο και κατα μετρο ορια ειναι μονοσημαντα (μ-σ.π.). Αρα το θεωρημα του von Neumann λεει οτι η συγκλιση στο θεωρημα Birkhoff ειναι και στον $L^p(\mu)$ αν $f\in L^p(\mu)$ και $1\leq p<+\infty$. Απο το θεωρημα von Neumann για $p=1$ θα μπορουσαμε να παρουμε και το (iv) του θεωρηματος Birkhoff.

ΑΠΟΔΕΙΞΗ ΘΕΩΡΗΜΑΤΟΣ VON NEUMANN: Πρωτα θεωρουμε $f\in L^\infty(\mu)$. Τοτε $f\in L^p(\mu)$ για καθε $p\geq 1$ και απο το θεωρημα Birkhoff υπαρχει $f^*\in L^1(\mu)$ ωστε $n^{-1}S_nf\to f^*$ (σ.π.). Προφανως $f^*\in L^\infty(\mu)$ και αρα $f^*\in L^p(\mu)$. Ακομη $|n^{-1}S_nf(x)-f^*(x)|^p\to 0$ για σχεδον καθε x και αρα απο το θεωρημα φραγμενης συγκλισης του Lebesgue, $\|n^{-1}S_nf-f^*\|_p\to 0$. Αρα για καθε $f\in L^\infty(\mu)$ και $\varepsilon>0$ υπαρχει $N(\varepsilon,f)\in\mathbb{N}$ ωστε

$$n\geq N(\varepsilon,f)\,,\quad k\in\mathbb{N}_0 \implies \left\|\frac{1}{n}S_nf-\frac{1}{n+k}S_{n+k}f\right\|_p<\varepsilon$$

Εστω τωρα $f\in L^p(\mu)$ και $\varepsilon>0$. Εστω $g\in L^\infty(\mu)$ με $\|f-g\|_p<\varepsilon$. Τοτε για $n\geq N(\frac{\varepsilon}{3},g)$ και $k\in\mathbb{N}_0$

$$\left\|\frac{1}{n}S_nf-\frac{1}{n+k}S_{n+k}f\right\|_p\leq$$

$$\leq\left\|\frac{1}{n}S_ng-\frac{1}{n+k}S_{n+k}g\right\|_p+\frac{1}{n}\left\|S_ng-S_nf\right\|_p+\frac{1}{n+k}\left\|S_{n+k}f-S_{n+k}g\right\|_p$$

$$\leq\frac{\varepsilon}{3}+\frac{1}{n}\|f-g\|_p+\|f-g\|_p<\varepsilon.$$

Αρα η $\{n^{-1}S_nf\}_{n\in\mathbb{N}}$ ειναι Cauchy στον (πληρη) $L^p(\mu)$ και ετσι πρεπει να συγκλινει προς καποια $f^*\in L^p(\mu)$.

Το γεγονος οτι $f^*=f^*\circ T$ βγαινει ειτε απο την παραπανω παρατηρηση και το (ii) του θεωρηματος Birkhoff ειτε απο το γεγονος οτι

$$\frac{1}{n}\left\|S_{n+1}f-S_nf\circ T\right\|_p=\frac{1}{n}\|f\|_p$$

και αρα οι ακολουθιες $\{S_nf\}$ και $\{S_nf\circ T\}$ πρεπει να εχουν τα ιδια L^p-ορια. ∎

5.6 Μερικες Εφαρμογες : (i) Ισχυρος Νομος Μεγαλων Αριθμων (Kolmogorov): Εστω $(\Omega, \mathcal{F}, P)$ ενας χωρος πιθανοτητας και $\{X_i\}_{i\in\mathbb{N}}$ μια ακολουθια τυχαιων μεταβλητων, με $X_i : \Omega \to \{0,1,\dots,k-1\}$ για ολα τα $i\in\mathbb{N}$ $(k>1)$. Αν $f:\{0,1,\dots,k-1\}^{\mathbb{N}} \to \mathbb{R}$ ειναι μια οποιαδηποτε συναρτηση με $E[|f(X_1,X_2,\dots)|] < +\infty$ (f μετρησιμη) τοτε μια εφαρμογη του Θεωρηματος 5.1 δινει

$$(*) \qquad \frac{1}{n}\sum_{i=1}^{n} f(X_i, X_{i+1},\dots) \to E[f(X_1,X_2,\dots)] \quad (\text{με πιθανοτητα } 1).$$

Το Θεωρημα 5.1 το εφαρμοζουμε στο χωρο $(X,\mathcal{B},\mu,T)$ οπου: $X=\{0,1,\dots,k-1\}^{\mathbb{N}}$, $\mathcal{B}$ η σ-αλγεβρα των Borel υποσυνολων του X (τοπολογια γινομενο στον X απο διακριτη στον $\{0,1,\dots,k-1\}$), T το (αριστερο) shift $(Tx)_i = x_{i+1}$ στον X κ' μ η κατανομη της ακολουθιας $\{X_i\}$, δηλ. $\mu(A) = P((X_i)_{i\in\mathbb{N}} \in A)$ για $A\in\mathcal{B}$. Επειδη οι X_i ειναι ανεξαρτητες το μετρο μ ειναι μετρο γινομενο και επειδη οι X_i ειναι ισονομες το μ διατηρειται απο το shift T, η με αλλα λογια ο χωρος $(X,\mathcal{B},\mu,T)$ ειναι ο χωρος shift του Παραδειγματος 4.7, που ηδη ξερουμε οτι ειναι εργοδικος.

Για $f(x_1,x_2,\dots) = x_1$ παιρνουμε το γνωστο νομο των μεγαλων αριθμων του Kolmogorov:

$$(**) \qquad n^{-1}\sum_{i=1}^{n} X_i \to E[X_1] \quad (\text{με πιθανοτητα } 1).$$

Φυσικα η $(*)$ ειναι ισχυροτερη της $(**)$ και <u>δεν</u> προκυπτει απο το νομο του Kolmogorov (τουλαχιστον οχι αμεσα), αφου οι $f(X_1,X_2,\dots)$, $f(X_2,X_3,\dots)$, ... δεν ειναι ανεξαρτητες.

(ii) Θεωρημα Κανονικων Αριθμων (Borel): Εστω $X=[0,1)$, $T: X\to X$ η απεικονιση $T(x) = 2x \pmod 1$ και λ το μετρο Lebesgue στο $[0,1)$. Το συστημα $(X,\mathcal{B}(X),\lambda,T)$ ειναι, οπως ειδαμε, εργοδικο, αφου η $T(x)=2x \pmod 1$ ειναι ουσιαστικα ιδια με την $z\to z^2$ στον κυκλο (1-torus) $K \cong \mathbb{R}/\mathbb{Z}$. Εφαρμοζοντας το Θεωρημα 5.1 για τη συναρτηση $f = \mathbb{1}_{[1/2,1)}$ παιρνουμε το Θεωρημα του Borel:

$$\text{σχετικη συχνοτητα } 1 \text{ στη δυαδικη αναπτυξη του } x \to \frac{1}{2} \quad \text{για σχεδον καθε } x.$$

Πραγματι αν $x=(x_1,x_2,\dots)$ ειναι το δυαδικο αναπτυγμα του x (δηλ.: $x=\sum_{n=1}^{\infty} x_n 2^{-n}$) τοτε $T(x) = T(\frac{x_1}{2}+\frac{x_2}{2^2}+\dots) = \frac{x_2}{2}+\frac{x_3}{2^2}+\dots$ και ετσι

$$f\circ T^i(x) = f\left(\frac{x_{i+1}}{2}+\frac{x_{i+2}}{2^2}+\dots\right) = \begin{cases} 1 & \text{αν } x_{i+1}=1 \\ 0 & \text{αν } x_{i+1}=0 \end{cases}$$

και αρα $\frac{1}{2} = \lambda([1/2,1)) = \int_{[0,1)} f\,d\lambda = \lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} f(T^i x) = \lim_{n\to\infty}$ (σχετικη συχνοτητα των 1 στο $x_1,\dots,x_n$).

(iii) Θεωρημα επαναφορας του Poincaré για Εργοδικα Συστηματα:

Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενδομορφισμος. Εστω $A\in\mathcal{A}$ με $\mu(A)>0$. Τοτε απο το Θεωρημα Birkhoff υπαρχει $f^*\in L^1(\mu)$ ωστε $n^{-1}\sum_{i=0}^{n-1}\mathbb{1}_A\circ T^i(x)\to f^*(x)$ για σχεδον καθε x. Αν $B=\{x\in X: f^*(x)>0\}$ τοτε $\mu(B)>0$ επειδη $\int f^* d\mu=\int\mathbb{1}_A\, d\mu=\mu(A)$ ειναι θετικο. Απο την αλλη για καθε $x\in B$ η τροχια $\{T^i(x): i\in\mathbb{N}_0\}$ επισκεπτεται το A απειρες φορες (δηλ. $T^i(x)\in A$ για απειρα i) αφου για να ειναι το οριο $\lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1}\mathbb{1}_A\circ T^i(x)$ θετικο πρεπει $\mathbb{1}_A\circ T^i(x)>0$ για απειρα i. Δηλαδη το Θεωρημα Birkhoff δινει το εξης αποτελεσμα για οποιοδηποτε συστημα: "για x σ' ενα συνολο θετικου μετρου $T^i(x)\in A$ για απειρα i." (Το Θεωρημα Poincaré ειναι ισχυροτερο: για σχεδον <u>καθε</u> $x\in A$ η τροχια $\{T^i(x)\}$ επιστρεφει στο A απειρες φορες.) Αν ομως ο T ειναι εργοδικος

$$\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{1}_A\circ T^i(x)\longrightarrow\mu(A)>0\qquad\text{για } \mu\text{-σχεδον καθε } x$$

και αρα "$T^i(x)\in A$ για απειρα $i\in\mathbb{N}_0$" ισχυει για σχεδον καθε x.

5.7 Θεωρημα: (Εργοδικο Θεωρημα για Ημιροες) Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $\{\varphi_t\}_{t\in\mathbb{R}_+}$ μια ημιροη. Τοτε για καθε $f\in L^1(\mu)$:

(i) $f^*(x)=\lim\limits_{T\to\infty}\frac{1}{T}\int_0^T f\circ\varphi_t(x)\,dt$ υπαρκει για μ-σχεδον καθε x.

(ii) $f^*\circ\varphi_t=f^*$, μ-σχεδον παντου, για καθε $t\in\mathbb{R}_+$.

(iii) $f^*\in L^1(\mu)$.

(iv) $\int_X f^*\,d\mu=\int_X f\,d\mu$.

Επιπλεον αν $f\in L^p(\mu)$ τοτε η συγκλιση (i) ειναι και στον $L^p(\mu)$ $(1\le p<\infty)$. Τελος αν η ημιροη $\{\varphi_t\}$ ειναι εργοδικη τοτε $f^*(x)=\int_X f\,d\mu$ για μ-σχεδον καθε $x\in X$.

ΑΠΟΔΕΙΞΗ: Κατ' αρχην η συναρτηση $X\times\mathbb{R}_+\ni(x,t)\mapsto\varphi_t(x)\in X$ ειναι μετρησιμη ως προς τις σ-αλγεβρες $\mathcal{A}\times\mathcal{R}_+$, $\mathcal{A}$ ($\mathcal{R}_+$ η σ-αλγεβρα στον $\mathbb{R}_+$) και η f ειναι μετρησιμη ως προς την $\mathcal{A}$ αρα η συναρτηση $X\times\mathbb{R}_+\ni(x,t)\longmapsto f(\varphi_t(x))$ ειναι μετρησιμη ως προς την $\mathcal{A}\times\mathcal{R}_+$. Απο το θεωρημα του Fubini η $t\mapsto f(\varphi_t(x))$ ειναι μετρησιμη για καθε x. Τωρα $\int_X\int_0^T|f\circ\varphi_t(x)|\,dt\,d\mu(x)=T\|f\|_1<+\infty$. Ετσι για καθε $n\in\mathbb{N}$ υπαρχει $A_n\in\mathcal{A}$ με $\mu(A_n)=1$ ωστε $\int_0^n|f\circ\varphi_t(x)|dt<+\infty$ $\forall x\in A_n$. Αρα για $x\in A=\bigcap_{n=1}^\infty A_n$, $\int_0^T|f\circ\varphi_t(x)|\,dt\le\int_0^{[T]}|f\circ\varphi_t(x)|\,dt<+\infty$ και το $\int_0^T f\circ\varphi_t(x)\,dt$

ειναι καλα ορισμενο για ολα τα $T \in \mathbb{R}_+$.

Εστω $F(x) = \int_0^1 f \circ \varphi_s(x)\, ds$ (για $x \in A$ και εστω $F(x) = 0$ για $x \in A^c$). Τοτε

$$\int_X |F(x)|\, \mu(dx) \leq \int_0^1 \int_X |f(\varphi_s(x))|\, \mu(dx)\, ds = \int_0^1 \|f\|_1\, ds = \|f\|_1$$

και ετσι $F \in L^1(\mu)$. Απο το Θεωρημα 5.1 το

$$f^*(x) := \lim_{\substack{N\to\infty \\ (N \in \mathbb{N})}} \frac{1}{N} \sum_{n=0}^{N-1} F \circ \varphi_1^n(x)$$

υπαρχει για μ-σχεδον καθε x και $f^* \in L^1(\mu)$. Ομως για $N \in \mathbb{N}$

$$\frac{1}{N}\int_0^N f \circ \varphi_t(x)\, dt = \frac{1}{N}\sum_{n=0}^{N-1} \int_n^{n+1} f \circ \varphi_t(x)\, dt = \frac{1}{N}\sum_{n=0}^{N-1} \int_n^{n+1} f \circ \varphi_{t-n}(\varphi_n(x))\, dt =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \int_0^1 f \circ \varphi_s(\varphi_n(x))\, ds = \frac{1}{N}\sum_{n=0}^{N-1} \int_0^1 f \circ \varphi_s(\varphi_1^n(x))\, ds =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} F \circ \varphi_1^n(x)$$

και αρα

$$\lim_{\substack{N\to\infty \\ (N \in \mathbb{N})}} \frac{1}{N}\int_0^N f \circ \varphi_t(x)\, dt = f^*(x) \qquad \text{για } \mu\text{-σχεδον καθε } x,$$

και αφου $f^* \in L^1(\mu)$ το οριο $f^*(x)$ ειναι πεπερασμενο μ-σ.π. Επαναλαμβανοντας για την $|f|$ στη θεση της f παιρνουμε οτι και το οριο

$$|f|^*(x) = \lim_{\substack{N\to\infty \\ (N \in \mathbb{N})}} \frac{1}{N}\int_0^N |f \circ \varphi_t(x)|\, dt$$

υπαρχει και ειναι πεπερασμενο μ-σχεδον παντου. Αρα για καθε $T \in \mathbb{R}_+$ τωρα

$$\frac{1}{T}\left|\int_0^T f \circ \varphi_t(x)\, dt - \int_0^{[T]} f \circ \varphi_t(x)\, dt\right| \leq \frac{1}{T}\int_{[T]}^T |f \circ \varphi_t(x)|\, dt \leq$$

$$\leq \frac{1}{T}\int_{[T]}^{[T]+1} |f \circ \varphi_t(x)|\, dt = \frac{1}{T}\left(\int_0^{[T]+1} |f \circ \varphi_t(x)|\, dt - \int_0^{[T]} |f \circ \varphi_t(x)|\, dt\right)$$

και καθως $T \to \infty$ η τελευταια παρασταση τεινει στο $|f|^*(x) - |f|^*(x) = 0$, μ-σ.π. Ετσι

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f \circ \varphi_t(x)\, dt = \lim_{\substack{N\to\infty \\ (N \in \mathbb{N})}} \frac{1}{N}\int_0^N f \circ \varphi_t(x)\, dt = f^*(x) \qquad (\mu\text{-σ.π.})$$

και ηδη γνωριζουμε οτι $f^* \in L^1(\mu)$.

Τωρα αν $s \in \mathbb{R}_+$

$$\frac{1}{T}\left|\int_0^T f\circ\varphi_t(\varphi_s(x))\,dt - \int_0^{T+s} f\circ\varphi_t(x)\,dt\right| =$$

$$= \frac{1}{T}\left|\int_s^{T+s} f\circ\varphi_t(x)\,dt - \int_0^{T+s} f\circ\varphi_t(x)\,dt\right| =$$

$$= \frac{1}{T}\left|\int_0^s f\circ\varphi_t(x)\,dt\right| \leqslant \frac{1}{T}\int_0^s |f\circ\varphi_t(x)|\,dt \longrightarrow 0$$

καθως $T\to\infty$, αφου για μ-σχεδον καθε x, $\int_0^s |f\circ\varphi_t(x)|\,dt < +\infty$. Ετσι

$$f^*\circ\varphi_s(x) = \lim_{T\to\infty}\int_0^T f\circ\varphi_t(\varphi_s(x))\,dt = \lim_{T\to\infty}\frac{1}{T}\int_0^{T+s} f\circ\varphi_t(x)\,dt = f^*(x)$$

για μ-σχεδον καθε x. Δηλαδη η f^* ειναι μ-σχεδον αναλλοιωτη.

Μεχρι εδω αποδειξαμε τα (i), (ii) και (iii). Τωρα θα αποδειξουμε οτι η συγκλιση $T^{-1}\int_0^T f\circ\varphi_t\,dt \to f^*$ ειναι και στον $L^p(\mu)$ αν $f\in L^p(\mu)$. Απο αυτο (για $p=1$) παιρνουμε αμεσως και το (iv) αφου

$$\int_X f^*\,d\mu = \int_X \lim_{T\to\infty}\frac{1}{T}\int_0^T f\circ\varphi_t(x)\,dt\;\mu(dx) = \quad (\text{λογω συγκλισης-}L^1)$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_X\int_0^T f\circ\varphi_t(x)\,dt\;\mu(dx) =$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T\int_X f\circ\varphi_t(x)\,\mu(dx)\,dt =$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T \int_X f\,d\mu\;dt = \int_X f\,d\mu .$$

Εστω λοιπον οτι $f\in L^p(\mu)$, οπου $1\leqslant p<\infty$. Τοτε απο την γενικευμενη ανισοτητα Minkowski

$$\|F\|_p = \left(\int_X\left|\int_0^1 f\circ\varphi_t(x)\,dt\right|^p\mu(dx)\right)^{1/p} \leqslant$$

$$\leqslant \int_0^1\left(\int_X |f\circ\varphi_t(x)|^p d\mu(x)\right)^{1/p} dt = \int_0^1 \|f\|_p\,dt = \|f\|_p$$

και αρα $F\in L^p(\mu)$. Ετσι απο το L^p- Εργοδικο Θεωρημα του von Neumann για ενδομορφισμους (Πορισμα 5.5)

$$\frac{1}{N}\int_0^N f\circ\varphi_t\quad dt = \frac{1}{N}\sum_{n=0}^{N-1} F\circ\varphi_1^n \xrightarrow{L^p} f^* .$$

καθως $N\to\infty$ με $N\in\mathbb{N}$.

Τωρα για $T \in \mathbb{R}_+$

$$\left| \frac{1}{T} \int_0^T f\circ\varphi_t(x)\,dt - \frac{1}{[T]} \int_0^{[T]} f\circ\varphi_t(x)\,dt \right| \leq$$

$$\leq \frac{1}{T} \left| \int_0^T f\circ\varphi_t(x)\,dt - \int_0^{[T]} f\circ\varphi_t(x)\,dt \right| + \left(\frac{1}{[T]} - \frac{1}{T}\right) \left| \int_0^{[T]} f\circ\varphi_t(x)\,dt \right| =$$

$$= \frac{1}{T} \left| \int_{[T]}^T f\circ\varphi_t(x)\,dt \right| + \left(\frac{1}{[T]} - \frac{1}{T}\right) \left| \int_0^{[T]} f\circ\varphi_t(x)\,dt \right| \leq$$

$$\leq \frac{1}{T} \int_{[T]}^T |f\circ\varphi_t(x)|\,dt + \left(\frac{1}{[T]} - \frac{1}{T}\right) \int_0^{[T]} |f\circ\varphi_t(x)|\,dt \leq$$

$$\leq \frac{1}{T} \int_{[T]}^{[T]+1} |f\circ\varphi_t(x)|\,dt + \left(\frac{1}{[T]} - \frac{1}{T}\right) \int_0^{[T]} |f\circ\varphi_t(x)|\,dt =$$

$$= \frac{[T]+1}{T} \cdot \frac{1}{[T]+1} \int_0^{[T]+1} |f\circ\varphi_t(x)|\,dt - \frac{[T]}{T} \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t(x)|\,dt +$$

$$+ \left(1 - \frac{[T]}{T}\right) \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t(x)|\,dt =$$

$$= \frac{[T]}{T} \left(\frac{1}{[T]+1} \int_0^{[T]+1} |f\circ\varphi_t(x)|\,dt - \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t(x)|\,dt \right) +$$

$$+ \frac{1}{T} \frac{1}{[T]+1} \int_0^{[T]+1} |f\circ\varphi_t(x)|\,dt + \left(1 - \frac{[T]}{T}\right) \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t(x)|\,dt .$$

Ετσι, χρησιμοποιοντας και την ανισοτητα Minkowski

$$\left\| \frac{1}{T} \int_0^T f\circ\varphi_t\,dt - \frac{1}{[T]} \int_0^{[T]} f\circ\varphi_t\,dt \right\|_p \leq$$

$$\leq \frac{[T]}{T} \left\| \frac{1}{[T]+1} \int_0^{[T]+1} |f\circ\varphi_t|\,dt - \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t|\,dt \right\|_p +$$

$$+ \frac{1}{T} \left\| \frac{1}{[T]+1} \int_0^{[T]+1} |f\circ\varphi_t|\,dt \right\|_p + \left(1 - \frac{[T]}{T}\right) \left\| \frac{1}{[T]} \int_0^{[T]} |f\circ\varphi_t|\,dt \right\|_p .$$

Ομως η ακολουθια $\{ N^{-1} \int_0^N |f\circ\varphi_t|\,dt : N \in \mathbb{N} \}$ συγκλινει στον $L^p(\mu)$ και αρα ειναι Cauchy. Επομενως ο πρωτος ορος στο δεξιο μελος της παραπανω ανισοτητας τεινει στο 0 καθως $T\to\infty$. Επισης, αφου η ακολουθια $\{ N^{-1} \int_0^N |f\circ\varphi_t|\,dt : N \in \mathbb{N} \}$ συγκλινει στον $L^p(\mu)$ ειναι και φραγμενη στον $L^p(\mu)$, δηλ.: $\sup_N N^{-1} \int_0^N |f\circ\varphi_t|\,dt < +\infty$. Επομενως και οι δυο τελευταιοι οροι στο δεξιο μελος της ανισοτητας τεινουν στο 0 καθως $T\to\infty$. Ετσι

$$\left\| \frac{1}{T} \int_0^T f\circ\varphi_t\,dt - \frac{1}{[T]} \int_0^{[T]} f\circ\varphi_t\,dt \right\|_p \to 0$$

και αφου ηδη γνωριζουμε οτι $f^* = L^p\text{-}\lim_{\substack{N\to\infty \\ (N\in\mathbb{N})}} N^{-1} \int_0^N f\circ\varphi_t\,dt$ επεται οτι

$$\frac{1}{T}\int_0^T f\circ\varphi_t\, dt \xrightarrow{L^p} f^* .$$

Τελος μενει να αποδειξουμε οτι αν η ημιροη $\{\varphi_t\}$ ειναι εργοδικη τοτε $f^* \equiv \int_X f\, d\mu$ (μ-σ.π.). Ομως αν η $\{\varphi_t\}$ ειναι εργοδικη τοτε αφου f^* ειναι μ-σχεδον αναλλοιωτη πρεπει να ειναι σταθερη μ-σ.π. (Προταση 4.4). Απο το (iv) η τιμη της σταθερας πρεπει να ειναι $\int_X f\, d\mu$. ■

6. MIXING

6.1 Θεωρημα: Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας και $T: X \to X$ ενδομορφισμος. Το συστημα $(X, \mathcal{A}, \mu, T)$ ειναι εργοδικο αν και μονο αν

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}A\cap B) \longrightarrow \mu(A)\mu(B) \qquad \forall A, B \in \mathcal{A}. \tag{6.1}$$

ΑΠΟΔΕΙΞΗ: Εστω οτι το συστημα ειναι εργοδικο και εστω $A, B \in \mathcal{A}$. Απο το εργοδικο θεωρημα

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A\circ T^i(x) \to \mu(A) \qquad (\mu\text{-σχεδον καθε } x)$$

και αρα

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A\circ T^i(x)\,.\,\mathbb{1}_B(x) \to \mu(A)\,\mathbb{1}_B(x) \qquad (\mu\text{-σ. καθε } x).$$

Απο το θεωρημα φραγμενης συγκλισης του Lebesgue

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}A\cap B) = \int_X \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A\circ T^i\cdot\mathbb{1}_B\, d\mu \to \int_X \mu(A)\,\mathbb{1}_B\, d\mu = \mu(A)\mu(B).$$

Αντιστροφα, εστω οτι ισχυει η (6.1) και εστω $A, B \in \mathcal{A}$ με $\mu(A)>0$, $\mu(B)>0$. Τοτε απο την (6.1)

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

Τοτε ομως απειροι οροι $\mu(T^{-i}A\cap B)$ πρεπει να ειναι θετικοι και επομενως υπαρχει $i \geq 1$ ωστε $\mu(T^{-i}A\cap B) > 0$. Η εργοδικοτητα του T επεται απο την Προταση 4.2. ■

Αλλαζοντας τον τροπο συγκλισης στην (6.1) (ισχυροποιοντας τον) παιρνουμε τις ακολουθες εννοιες:

6.2 Ορισμος: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενδομορφισμος. Το συστημα $(X,\mathcal{A},\mu,T)$ ειναι ασθενως mixing αν

$$\text{(6.2)} \qquad \frac{1}{n}\sum_{i=0}^{n-1} \left|\mu(T^{-i}A\cap B)-\mu(A)\mu(B)\right| \to 0 \qquad \forall A,B\in\mathcal{A}.$$

Το συστημα ειναι ισχυρα mixing αν

$$\text{(6.3)} \qquad \mu(T^{-n}A\cap B)\to\mu(A)\mu(B) \qquad \forall A,B\in\mathcal{A}.$$

Η επομενη προταση ειναι προφανης:

6.3 Προταση: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενδομορφισμος. Τοτε

$$T \text{ ισχυρα mixing} \Rightarrow T \text{ ασθενως mixing} \Rightarrow T \text{ εργοδικος.} \qquad \blacksquare$$

Παρατηρηση: Τα αντιστροφα δεν ισχυουν. Θα δουμε παρακατω οτι καμμια στροφη του κυκλου (και γενικοτερα μιας συμπαγους ομαδας) δεν ειναι ασθενως mixing ενω ηδη ξερουμε οτι οι αρρητες στροφες του κυκλου ειναι εργοδικες. Επισης υπαρχουν παραδειγματα ασθενως αλλα οχι ισχυρως mixing συστηματων. Δυο τετοια παραδειγματα (παρομοια) δινει ο Kakutani στο Examples of Ergodic Measure Preserving Transformations which are Weakly Mixing but not Strongly Mixing. (Springer Lecture Notes in Math. 318, 143-149 (1973).)

6.4 Προταση: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενας ενδομορφισμος. Τοτε ο T ειναι εργοδικος, ασθενως mixing ή ισχυρα mixing αν και μονο αν οι (6.1), (6.2) ή (6.3) αντιστοιχα ισχυουν για ολα τα A, B σε μια ημιαλγεβρα που παραγει την $\mathcal{A}$.

Παρατηρηση: Μια ημιαλγεβρα $\mathcal{Y}$ ειναι μια οικογενεια υποσυνολων του X τετοια ωστε: $\emptyset\in\mathcal{Y}$, $A,B\in\mathcal{Y}\Rightarrow A\cap B\in\mathcal{Y}$ και $A\in\mathcal{Y}\Rightarrow A^c=\bigcup_{i=1}^{n}E_i$ για καποια ξενα ανα δυο στοιχεια $E_1,\dots,E_n\in\mathcal{Y}$. Παραδειγματα: τα υποδιαστηματα ενος διαστηματος $[a,b]$ ειναι μια ημιαλγεβρα που μαλιστα γεννα την Borel σ-αλγεβρα του $[a,b]$. Οι κυλινδροι σ' ενα χωρο ακολουθιων ειναι μια ημιαλγεβρα που παραγει την Borel σ-αλγεβρα.

Μια αλγεβρα $\mathcal{A}$ ειναι οπως και μια σ-αλγεβρα με τη μονη διαφορα οτι απαιτουμε να ειναι κλειστη σε πεπερασμενες μονο ενωσεις στοιχειων της. Δηλ.

η $\mathcal{A}$ ειναι αλγεβρα αν $\emptyset\in\mathcal{A}$, $A\in\mathcal{A}\Rightarrow A^c\in\mathcal{A}$ και $A_1,\dots,A_n\in\mathcal{A}\Rightarrow \bigcup_{i=1}^{n} A_i\in\mathcal{A}$.

ΑΠΟΔΕΙΞΗ ΠΡΟΤΑΣΗΣ 6.4: Εστω $\mathcal{Y}$ μια ημιαλγεβρα που παραγει την $\mathcal{A}$ και εστω $\alpha(\mathcal{Y})$ η αλγεβρα που παραγει η $\mathcal{Y}$. Τοτε η αλγεβρα $\alpha(\mathcal{Y})$ παραγει την σ-αλγεβρα $\mathcal{A}$. Επιπλεον καθε στοιχειο $A\in\alpha(\mathcal{Y})$ ειναι πεπερασμενη ενωση ξενων ανα δύο στοιχείων της $\mathcal{Y}$. Αρα αν καποια απο τις (6.1) – (6.3) ισχύει για ολα τα στοιχεια της $\mathcal{Y}$ τοτε ισχύει και για ολα τα στοιχεια της $\alpha(\mathcal{Y})$.

Εστω τωρα $A,B\in\mathcal{A}$. Τοτε αν $\varepsilon>0$ υπαρχουν $A_0,B_0\in\alpha(\mathcal{Y})$ με $\mu(A\triangle A_0)<\varepsilon$ και $\mu(B\triangle B_0)<\varepsilon$ (επειδή η $\alpha(\mathcal{Y})$ παραγει την $\mathcal{A}$). Τοτε

$$
\begin{aligned}
|\mu(T^{-i}A\cap B)-\mu(A)\mu(B)| &\leq |\mu(T^{-i}A\cap B)-\mu(T^{-i}A_0\cap B_0)| + \\
&+ |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)| + |\mu(A_0)\mu(B_0)-\mu(A_0)\mu(B)| + \\
&+ |\mu(A_0)\mu(B)-\mu(A)\mu(B)| \leq \\
&\leq \mu\big((T^{-i}A\cap B)\triangle(T^{-i}A_0\cap B_0)\big) + \mu(A_0)|\mu(B_0)-\mu(B)| + \\
&+ \mu(B)|\mu(A_0)-\mu(B)| + |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)| \leq \\
&\leq \mu(T^{-i}A\triangle T^{-i}A_0)+\mu(B\triangle B_0)+\mu(A_0)\mu(B\triangle B_0)+\mu(B)\mu(A\triangle A_0)+ \\
&+ |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)| = \\
&= \mu(T^{-i}(A\triangle A_0)) + (1+\mu(A_0))\mu(B\triangle B_0)+\mu(B)\mu(A\triangle A_0)+ \\
&+ |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)| = \\
&= (1+\mu(B))\mu(A\triangle A_0) + (1+\mu(A_0))\mu(B\triangle B_0) + |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)| < \\
&< 4\varepsilon + |\mu(T^{-i}A_0\cap B_0)-\mu(A_0)\mu(B_0)|
\end{aligned}
$$

που δειχνει οτι αν καποια απο τις (6.2), (6.3) ισχύει για ολα τα στοιχεια της $\alpha(\mathcal{Y})$ τοτε ισχυει και για ολα τα στοιχεια της $\mathcal{A}$. Για την (6.1) έχουμε με ομοιο τροπο

$$
\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}A\cap B)-\mu(A)\mu(B)\right| < 4\varepsilon + \left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}A_0\cap B_0) - \right.
$$
$$
\left. -\mu(A_0)\mu(B_0)\right| . \blacksquare
$$

6.5 Προταση: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας και $T: X\to X$ ενδομορφισμος. Τοτε

(i) T εργοδικος $\iff \frac{1}{n}\sum_{i=0}^{n-1}(f\circ T^i, g) \to (f,1)(1,g) \quad \forall f,g\in L^2(\mu)$.

(ii) T ασθενως mixing $\iff \frac{1}{n}\sum_{i=0}^{n-1}\left|(f\circ T^i, g)-(f,1)(1,g)\right| \to 0 \quad \forall f,g\in L^2(\mu)$.

(iii) T ισχυρα mixing $\iff (f\circ T^n, g)\to (f,1)(1,g)\to 0 \quad \forall f,g\in L^2(\mu)$.

ΑΠΟΔΕΙΞΗ : Οι αποδειξεις των (i), (ii) κ' (iii) ειναι εντελως ομοιες και θα αποδειξουμε ενδεικτικα μονο την (iii).

Αν $A, B\in\mathcal{A}$ τοτε παιρνοντας για $f=\mathbb{1}_A$, $g=\mathbb{1}_B$ παιρνουμε οτι

$$\mu(T^{-n}A\cap B) = \int_X \mathbb{1}_A\circ T^n\cdot \mathbb{1}_B\, d\mu = (f\circ T^n, g)\to (f,1)(1,g) =$$

$$= \int_X f\,d\mu \int_X g\,d\mu = \mu(A)\mu(B).$$

Αρα ο T ειναι ισχυρα mixing.

Αντιστροφα, εστω οτι ο T ειναι ισχυρα mixing. Τοτε η προς αποδειξη σχεση ισχυει εξ' ορισμου για f κ' g χαρακτηριστικές συναρτησεις μετρησιμων υποσυνολων του X. Αυτο αμεσως επεκτεινεται σε f κ' g απλες συναρτησεις (συναρτησεις της μορφης $\sum_{i=1}^{n} a_i\mathbb{1}_{A_i}$ οπου a_i σταθερες και A_i ξενα ανα δυο μετρησιμα συνολα).

Εστω τωρα $f,g\in L^2(\mu)$. Αν $\varepsilon>0$ ειναι δοσμενο, τοτε υπαρχουν f_0, g_0 απλες συναρτησεις, ωστε $\|f-f_0\|_2<\varepsilon$ και $\|g-g_0\|_2<\varepsilon$. Τοτε χρησιμοποιοντας την ανισοτητα Cauchy-Schwartz και οτι η T διατηρει το μετρο μ

$$\left|(f\circ T^i,g)-(f,1)(1,g)\right| \leq \left|(f\circ T^i,g)-(f_0\circ T^i,g)\right| + \left|(f_0\circ T^i,g)-(f_0\circ T^i,g_0)\right|$$

$$+\left|(f_0\circ T^i,g_0)-(f_0,1)(1,g_0)\right| + \left|(f_0,1)(1,g_0)-(f,1)(1,g_0)\right| +$$

$$+\left|(f,1)(1,g_0)-(f,1)(1,g)\right| \leq$$

$$\leq \|f\circ T^i - f_0\circ T^i\|_2^2\,\|g\|_2^2 + \|f_0\circ T^i\|_2^2\,\|g-g_0\|_2^2 + \left|(f_0\circ T^i,g_0)-(f_0,1)(1,g_0)\right| +$$

$$+\left|(1,g_0)\right|\left|(f_0,1)-(f,1)\right| + \left|(f,1)\right|\left|(1,g_0)-(1,g)\right| \leq$$

$$\leq \|f-f_0\|_2^2\,\|g\|_2^2 + \|f_0\|_2^2\,\|g-g_0\|_2^2 + \left|(f_0\circ T^i,g_0)-(f_0,1)(1,g_0)\right| +$$

$$+\left|(1,g_0)\right|\|f-f_0\|_2^2 + \left|(f,1)\right|\|g-g_0\|_2^2 <$$

$$< (\|g\|_2^2 + \|f_0\|_2^2 + |(1,g_0)| + |(f_0,1)|) \varepsilon^2 + |(f_0 \circ T^i, g_0) - (f_0,1)(1,g_0)| < \varepsilon$$

αν διαλέξουμε το i αρκούντως μεγάλο, αφού $|(f_0 \circ T^i, g_0) - (f_0,1)(1,g_0)| \to 0$ μιας και f_0, g_0 είναι απλές συναρτήσεις. □

Παρατηρήσεις: (i) Οπως δείχνει η απόδειξη η πρόταση ισχύει είτε θεωρήσουμε τον χώρο $L^2(\mu)$ σαν τον χώρο των μιγαδικών συναρτήσεων με $\int |f|^2 d\mu < \infty$ και εσωτερικό γινόμενο $(f,g) = \int_X f \cdot \bar{g}\, d\mu$ είτε σαν τον χώρο των πραγματικών συναρτήσεων με $\int_X |f|^2 d\mu < \infty$ και εσωτερικό γινόμενο $(f,g) = \int_X f \cdot g\, d\mu$.

(ii) Η απόδειξη της Πρότασης 6.5 δείχνει επίσης ότι αν ο χώρος X έχει κάποια τοπολογία και οι συνεχείς συναρτήσεις $C(X)$ είναι πυκνές στον $L^2(\mu)$ τότε η Πρόταση 6.5 ισχύει και με $C(X)$ στη θέση $L^2(\mu)$.

6.6 Παράδειγμα: (Στροφές του κύκλου) Καμμία στροφή του κύκλου δεν είναι ασθενώς (και άρα ούτε ισχυρώς) mixing. Πράγματι, έστω $T_a : K \to K$, όπου $a \in K$, η στροφή $T_a(z) = a \cdot z$. Εστω $f(z) = z$. Τότε $f \in L^2(\lambda)$, όπου λ μέτρο Lebesgue και $(f,1) = \int_0^1 f(e^{2\pi i t})\,dt = \int_0^1 e^{2\pi i t} dt = 0$. Έτσι από την Πρόταση 6.5, αν ο T_a ήταν ασθενώς mixing θα έπρεπε

$$\frac{1}{n} \sum_{i=0}^{n-1} |(f \circ T_a^i, f)| \to 0.$$

Ομως για κάθε $j \in \mathbb{N}_0$

$$|(f \circ T_a^j, f)| = \left| \int_0^1 f(T_a^j(e^{2\pi i t}))\, \overline{f(e^{2\pi i t})}\, dt \right| =$$

$$= \left| \int_0^1 a^j e^{2\pi i t} e^{-2\pi i t} dt \right| = |a^j| = 1.$$

Παρατήρηση: Το παραπάνω ισχύει για οποιαδήποτε συμπαγή ομάδα: καμμία στροφή $T(x) = a x$ συμπαγούς ομάδας δεν είναι ασθενώς mixing ως προς το μέτρο Haar της ομάδας.

6.7 Παράδειγμα: (Ενδομορφισμοί του n-torus) Στην περίπτωση ενδομορφισμών του n-torus εργοδικότητα και ισχυρό (και άρα και ασθενές) mixing είναι ισοδύναμα. Συγκεκριμένα έστω $T : \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{R}^n/\mathbb{Z}^n$ ο ενδομορφισμός του n-torus που επάγεται από τον τύπο $x \mapsto Ax \pmod 1$, $x \in [0,1)^n$, όπου A πίνακας με στοιχεία στο $\mathbb{Z}$ και $\det A \neq 0$ (οπότε ο T διατηρεί το μέτρο Haar = μέτρο Lebesgue λ_n στο $[0,1)^n$ της ομάδας). Θα αποδείξουμε ότι αν ο T είναι εργοδικός τότε είναι και ισχυρα mixing.

Εστω λοιπον οτι ο T ειναι εργοδικος. Θα αποδειξουμε οτι

$$(*) \qquad (f\circ T^k, g) \to (f,1)(1,g) \qquad \forall\, f,g\in L^2.$$

Κατ' αρχην παρατηρουμε οτι αρκει να αποδειξουμε την (*) για $f\in C^\infty$ και $g\in L^\infty$. Πραγματι, αν $f,g\in L^\infty$ τοτε μπορουμε να βρουμε $\tilde f\in C^\infty$ με $\|f-\tilde f\|_2<\varepsilon$ και $\tilde g\in L^\infty$ με $\|g-\tilde g\|_2<\varepsilon$ και ετσι, οπως και στην αποδειξη της Προτασης 6.5, αν η (*) ισχυει για το ζευγαρι $\tilde f,\tilde g$, θα ισχυει και για το ζευγαρι f,g.

Εστω λοιπον $f\in C^\infty$ και $g\in L^\infty$. Τοτε $f,g\in L^2$ και αρα οι f και g εχουν αναπτυγματα Fourier στον L^2:

$$f(x)=\sum_{\alpha\in\mathbb{Z}^n}\hat f(\alpha)e^{2\pi i\alpha.x}\quad,\qquad g(x)=\sum_{\alpha\in\mathbb{Z}^n}\hat g(\alpha)e^{2\pi i\alpha.x}.$$

Τοτε

$$f\circ T^k(x)=\sum_{\alpha\in\mathbb{Z}^n}\hat f(\alpha)e^{2\pi i\alpha.(A^kx)}=\sum_{\alpha\in\mathbb{Z}^n}\hat f(\alpha)e^{2\pi i(A_t^k\alpha).x},$$

οπου A_t ο ανάστροφος του A. Η τελευταια σχεση μας δινει $\widehat{(f\circ T^k)}(A_t^k\alpha)=\hat f(\alpha)$ και αρα απο τον τυπο του Parseval παιρνουμε οτι

$$(f\circ T^k,g)=\sum_{\alpha\in\mathbb{Z}^n}\widehat{(f\circ T^k)}(\alpha)\overline{\hat g(\alpha)}=\sum_{\alpha\in\mathbb{Z}^n}\hat f(\alpha)\overline{\hat g(A_t^k\alpha)}=$$

$$=\hat f(0)\overline{\hat g(0)}+\sum_{\alpha\in\mathbb{Z}^n-\{0\}}\hat f(\alpha)\overline{\hat g(A_t^k\alpha)}.$$

Αφου $\hat f(0)\overline{\hat g(0)}=\int f\,d\lambda_n\int\bar g\,d\lambda_n=(f,1)(1,g)$, θα εχουμε τελειώσει αν αποδείξουμε οτι

$$(**) \qquad \sum_{\alpha\in\mathbb{Z}^n-\{0\}}\hat f(\alpha)\overline{\hat g(A_t^k\alpha)}\to 0 \qquad (k\to\infty).$$

Πρωτα απ' ολα παρατηρουμε οτι αν $\alpha\neq 0$ και $m\neq k$ τοτε $A_t^k\alpha\neq A_t^m\alpha$ γιατι διαφορετικα ο A θα ειχε μια ιδιοτιμη που θα ηταν ριζα της μοναδας κι' αυτο αντικρουει την εργοδικοτητα του T.

Απο αυτο αμεσως παιρνουμε οτι $|A_t^k\alpha|\to+\infty$ οταν $\alpha\neq0$ και αρα πρεπει $\hat g(A_t^k\alpha)\to 0$ καθως $k\to\infty$, αφου η σειρα $\sum|\hat g(\alpha)|^2=\|g\|_2^2<+\infty$ συγκλίνει. Ετσι

$$\lim_{k\to\infty}\hat f(\alpha)\overline{\hat g(A_t^k\alpha)}=0 \qquad \forall\,\alpha\in\mathbb{Z}^n-\{0\}$$

και η (∗∗) θα εχει αποδειχτεί (απο το M-τεστ του Weierstrass) αν δείξουμε οτι $|\hat{f}(a)\hat{g}(A_t^k a)| \leq M(a)$ για καποια $M(a)$ με $\sum_{a\in\mathbb{Z}^n-\{0\}} M(a) < +\infty$.

Ομως

$$|\hat{g}(a)| = \left|\int_{[0,1)} g(x)\, e^{2\pi i a\cdot x}\, dx\right| \leq \|g\|_\infty \qquad \forall a\in\mathbb{Z}^n$$

και επομένως

$$|\hat{f}(a)\,\hat{g}(A_t^k a)| \leq \|g\|_\infty\, |\hat{f}(a)|.$$

Αφου $f\in C^\infty$ η σειρα $\sum \|g\|_\infty |\hat{f}(a)| = \|g\|_\infty \sum |\hat{f}(a)|$ συγκλίνει και η (∗∗) αποδειχτηκε.

Παρατηρηση: Στο παραπανω παραδειγμα χρησιμοποιησαμε το οτι αν $f\in C^\infty$ τοτε $\sum |f(a)|$ συγκλίνει. Για $n=1$ αυτο μπορει να αποδειχτεί ως εξης.

Αν $f\in C^1$, ολοκληρωνοντας κατα μέρη παιρνουμε οτι (για $k\neq 1$)

$$\hat{f}_k = \int_0^1 f(x)\, e^{-2\pi i k x}\, dx = -\frac{1}{2\pi i k}\int_0^1 f(x)\, d e^{-2\pi i k x} =$$

$$= -\frac{1}{2\pi i k}\left[f(1)-f(0)\right] + \frac{1}{2\pi i k}\int_0^1 f'(x)\, e^{-2\pi i k x}\, dx =$$

$$= \frac{1}{2\pi i k}\,\widehat{(f')}_k\,,$$

αφου $f(0)=f(1)$ (η f είναι συναρτηση ορισμένη στον κυκλο K, η ισοδυναμα, στο $[0,1]$ με τα ακρα 0 και 1 ταυτισμένα). Επομένως αν $f\in C^m$ τοτε

$$\hat{f}_k = \frac{1}{(2\pi i k)^m}\,\widehat{\left(f^{(m)}\right)}_k$$

απο το οποιο παιρνουμε

$$|\hat{f}_k| = (2\pi|k|)^{-m}\left|\widehat{\left(f^{(m)}\right)}_k\right| \leq (2\pi|k|)^{-m}\|f^{(m)}\|_\infty = C_f|k|^{-m}.$$

Ετσι για $f\in C^\infty$ (στην πραγματικοτητα $f\in C^2$ φτανει) εχουμε $\sum_k |\hat{f}_k| < +\infty$.

Η αποδειξη για περισσοτερες διαστασεις μπορει να γίνει παρομοια. Συγκεκριμμένα μπορουμε να αποδείξουμε οπως και πιο πανω οτι για καποια σταθερα C_f εχουμε οτι για καθε $a=(a_1,\dots,a_n)\neq(0,\dots,0)$

$$|\hat{f}(a)| \leq \frac{C_f}{\prod_{i:a_i\neq 0} a_i^2}\,.$$

6.8 Παραδειγμα: (Μαρκοβιανα shift) Εστω $Y=\{0,1,\ldots,k-1\}$ και $X=Y^{\mathbb{Z}}$ ο χωρος των "διπλευρων" ακολουθιων $(\ldots,x_{-1},x_0,x_1,\ldots)$ με $x_i\in Y$. Εστω $\mathcal{B}$ η Borel σ-αλγεβρα του X, οπου ο X εχει την τοπολογια γινομενο προερχομενη απο την διακριτη τοπολογια στον Y και τελος εστω $T:X\to X$ η απεικονιση shift: $(\ldots,x_{-1},\dot{x}_0,x_1,\ldots)\xrightarrow{T}(\ldots,x_0,\dot{x}_1,x_2,\ldots)$.

Εστω τωρα οτι για καθε $n\in\mathbb{N}_0$ και $i_0,\ldots,i_n\in Y$ μας εχουν δοθει αριθμοι $p_n(i_0,\ldots,i_n)$ τετοιοι ωστε

(i) $p_n(i_0,\ldots,i_n)\geq 0$

(ii) $\sum_{i\in Y}p_0(i)=1$

(iii) $\sum_{i_{n+1}\in Y}p_{n+1}(i_0,\ldots,i_n,i_{n+1})=p_n(i_0,\ldots,i_n)$.

Αν ορισουμε τη συνολοσυναρτηση μ πανω στην ημι-αλγεβρα των κυλινδρων απο τη σχεση

$$\mu(\{(x_i):\ x_k=i_0,\ldots,x_{k+n}=i_n\})=p_n(i_0,\ldots,i_n)$$

για $i_0,\ldots,i_n\in Y$ και $n\in\mathbb{N}_0$, τοτε η μ επεκτεινεται, κατα μονοσημαντο τροπο, σ' ενα μετρο πιθανοτητας πανω σε Borel σ-αλγεβρα $\mathcal{B}$ (η οποια παραγεται απο τους κυλινδρους). Το μετρο μ μενει αναλλοιωτο απο τον T, δηλ.: $\mu(B)=\mu(T^{-1}B)$ $\forall B\in\mathcal{B}$. (Στην πραγματικοτητα ολα τα T-αναλλοιωτα μετρα πανω στο μετρησιμο χωρο $(X,\mathcal{B})$ παιρνονται μ' αυτο τον τροπο.)

Αν $\mathbf{p}=(p_0,\ldots,p_{k-1})$ ειναι ενα διανυσμα πιθανοτητας (δηλ.: $p_i\geq 0$, $\sum_i p_i=1$) τοτε θετοντας $p_n(i_0,\ldots,i_n)=p_{i_0}p_{i_1}\cdots p_{i_n}$ παιρνουμε το γνωστο $\mathbf{p}$-shift (Παραδειγματα 1.2 και 4.7).

Μια αλλη σημαντικη κατηγορια παραδειγματων ειναι τα Μαρκοβιανα shift. Θεωρουμε δοσμενα εναν στοχαστικο πινακα $P=(p(i,j))_{i,j\in Y}$ (δηλ.: $p(i,j)\geq 0$ και $\sum_{j\in Y}p(i,j)=1$) και ενα διανυσμα πιθανοτητας $\mathbf{p}=(p_0,\ldots,p_{k-1})$ με την ιδιοτητα $\mathbf{p}P=\mathbf{p}$ και θετουμε $p_0(i)=p_i$, $p_n(i_0,\ldots,i_n)=p_{i_0}\,p(i_0,i_1)\,p(i_1,i_2)\ldots\,p(i_{n-1},i_n)$. Ο χωρος $(X,\mathcal{B},\mu,T)$ που προκυπτει ετσι λεγεται $(\mathbf{p},P)$ Μαρκοβιανο shift.

Το (απλο) $\mathbf{p}$-shift ειναι ειδικη περιπτωση Μαρκοβιανου shift με $p(i,j)=p_j$ για ολα τα i. Τα Μαρκοβιανα shift εχουν εννοια, οπως και τα απλα shift, και στο χωρο των "μονοπλευρων" ακολουθιων $(x_0,x_1,\ldots)$ οπου $x_i\in Y$ και για να τα ξεχωριζουμε θα τα λεμε μονοπλευρα και διπλευρα shift και Μαρκοβιανα shift αντιστοιχα.

Τελος, θα υποθετουμε παντα οτι το διανυσμα $\mathbf{p}$ εχει ολα τα στοιχεια του θετικα, αφου αν $p_i=0$ για καποιο i, τοτε καθε κυλινδρος που περιεχει το i εχει μετρο 0 και αρα μπορουμε, απο πιθανοθεωρητικης σκοπιας, να περιορισουμε στη μελετη του χωρου $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu})$, που προκυπτει αν στη θεση του Y θεωρησουμε τον $\tilde{Y} = Y - \{i\}$.

Συμβολιζουμε με $p^{(n)}(i,j)$ το (i,j)-στοιχειο του πινακα P^n και παρατηρουμε οτι για οποιαδηποτε k

$$\frac{1}{p_i}\,\mu(\{(x_l):\ x_k = i,\ x_{k+n} = j\}) =$$

$$= \frac{1}{p_i} \sum_{i_1,\dots,i_{n-1}} \mu(\{(x_l):\ x_k = i,\ x_{k+1} = i_1, \dots, x_{k+n-1} = i_{n-1},\ x_{k+n} = j\}) =$$

$$= \frac{1}{p_i} \sum_{i_1,\dots,i_{n-1}} p_i\, p(i,i_1)\dots p(i_{n-1},j) = p^{(n)}(i,j).$$

Τελος, θα λεμε οτι ο πινακας P ειναι αναγωγος αν για οποιαδηποτε $i,j \in Y$ υπαρχει $n \in \mathbb{N}$ ωστε $p^{(n)}(i,j) > 0$, δηλ.: ξεκινωντας απο οποιαδηποτε θεση i στον Y μπορουμε τελικα να παμε σ' οποιαδηποτε θεση j στο Y. Ο P θα λεγεται αναγωγος και απεριοδικος αν υπαρχει ενα $n \in \mathbb{N}$ ωστε $p^{(n)}(i,j) > 0$ για ολα τα i,j, δηλ.: P^n εχει ολα του στοιχεια θετικα για καποιο n.

Προταση: Εστω $(X, \mathcal{B}, \mu, T)$ το διπλευρο $(\mathbf{p}, P)$-μαρκοβιανο shift με χωρο καταστασεων $Y = \{0, 1, \dots, k-1\}$. (Θεωρουμε οτι $p_i > 0$ $\forall i$.) Τοτε:

(A) τα εξης ειναι ισοδυναμα:

(i) το συστημα ειναι εργοδικο

(ii) ο P ειναι αναγωγος

(iii) το 1 ειναι απλη ιδιοτιμη του P.

(B) τα εξης ειναι ισοδυναμα:

(i) το συστημα ειναι ασθενως mixing

(ii) ο P ειναι αναγωγος και απεριοδικος

(iii) για καθε j εχουμε οτι $\lim_{n\to\infty} p^{(n)}(i,j) = p_j$, $\forall i \in Y$

(iv) το συστημα ειναι ισχυρα mixing.

ΑΠΟΔΕΙΞΗ: **(A)** (i) $\Rightarrow$ (ii) Εστω οτι το συστημα ειναι εργοδικο. Απο την Προταση 6.1

$$\frac{1}{N}\sum_{n=0}^{N-1} p^{(n)}(i,j) = \frac{1}{N}\sum_{n=0}^{N-1} \frac{1}{p_i}\,\mu(\{(x_l):\ x_0 = i,\ x_n = j\}) =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \frac{1}{p_i}\,\mu(\{x_0 = i\} \cap T^{-n}\{x_0 = j\}) \to \frac{1}{p_i}\,\mu(\{x_0 = i\})\mu(\{x_0 = j\})$$

$$= \frac{1}{p_i}\, p_i\, p_j = p_j.$$

Αφου $p_j > 0$ πρεπει να υπαρχη ενα $n \in \mathbb{N}$ ωστε $p^{(n)}(i,j) > 0$. Δηλαδη ο P ειναι αναγωγος.

(ii) ⇒ (i): Εστω οτι ο P ειναι αναγωγος. Θα αποδειξουμε οτι ο T ειναι εργοδικος.
Η αποδειξη θα γινει σε βηματα.

(1) Το οριο $\lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} P^n$ υπαρχει και ειναι ενας στοχαστικος πινακας Q.

Αποδειξη: Εστω $C_j = \{(x_i) : x_0 = j\}$. Τοτε απο το Θεωρημα 5.1 το οριο $f^* := \lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} 1_{C_j}\circ T^n$ υπαρχει. Αρα $N^{-1}\sum_{n=0}^{N-1} 1_{C_j}\circ T^n \cdot 1_{C_i} \to f^* 1_{C_i}$, μ-σ.π. και ολοκληρωνοντας παιρνουμε $N^{-1}\sum_{n=0}^{N-1} p^{(n)}(i,j) \to p_i^{-1}\int f^* 1_{C_i}\, d\mu =: q(i,j)$. Θετουμε $Q = (q(i,j))$ και προφανως ο Q ειναι στοχαστικος (δηλ.: $\sum_j q(i,j) = 1$) αφου καθε P^n ειναι στοχαστικος.

(2) $q(i,j) > 0 \quad \forall (i,j)$.

Αποδειξη: Για $i \in Y$ θετουμε $Y_i = \{ j \in Y : q(i,j) > 0\}$. ($Y_i \neq \emptyset$ γιατι $\sum_j q(i,j) = 1$.) Απο τον ορισμο του Q προκυπτει αμεσα οτι $QP = Q$ και αρα $q(i,l) \geq q(i,l)\,p(l,j)$ για ολα τα $l \in Y$. Επομενως $l \in Y_i$ κ' $p(l,j) > 0$ δινουν $q(i,j) > 0$, δηλαδη $j \in Y_i$. Απο αυτο ομως παιρνουμε οτι

$$(*) \qquad l \in Y_i \Rightarrow \sum_{j\in Y_i} p(l,j) = 1.$$

Αφου ομως ο P ειναι αναγωγος απο αυτο προκυπτει οτι $Y_i = Y$, γιατι πρεπει απο καθε θεση $l \in Y$ να μπορουμε να παμε τελικα σε οποιαδηποτε αλλη θεση $k \in Y$. Πιο αναλυτικα αν το $Y - Y_i$ δεν ηταν κενο τοτε για καθε $l \in Y_i$ και $k \in Y - Y_i$ θα επρεπε $p(l,k) = 0$ απο την (*). Ετσι για καθε $l \in Y_i$ και $k \in Y - Y_i$ θα ειχαμε

$$p^{(n)}(l,k) = \sum_{j\in Y} p(l,j)\, p^{(n-1)}(j,k) = \sum_{j\in Y_i} p(l,j)\, p^{(n-1)}(j,k)$$

και επαγωγικα θα παιρναμε οτι $p^{(n)}(l,k) = 0 \quad \forall n$ ο ο $l \in Y_i$ και $k \in Y - Y_i$. Αυτο ομως δεν μπορει να συμβει αφου ο P ειναι αναγωγος.

(3) Ολες οι γραμμες του Q ειναι ιδιες.

Αποδειξη: Πρωτα απ' ολα παρατηρουμε οτι $Q^2 = Q$. Αυτο γιατι $QP = Q$ προφανως και $Q^2 = Q \lim_n N^{-1}\sum_{n=0}^{N-1} P^n = \lim_N N^{-1}\sum_{n=0}^{N-1} QP^n = N^{-1}\sum_{n=0}^{N-1} Q = Q$.

Τωρα θετουμε $q_j = \max_i q(i,j)$. Αν ειχαμε $q(i,j) < q_j$ για καποιο i τοτε

$$q(l,j) = \sum_k q(l,k)\, q(k,j) \quad < \qquad \text{(απο το βημα (2))}$$

$$< q_j \sum_k q(l,k) = q_j$$

για ολα τα l και αυτο αντικρουει τον ορισμο του q_j.

(4) $q(i,j) = p_j \quad \forall (i,j)$.

Αποδειξη: Αφου $\mathbf{p}P = \mathbf{p}$ πρεπει και $\mathbf{p}Q = \lim_N N^{-1}\sum_{n=0}^{N-1} \mathbf{p}P^n = N^{-1}\sum_{n=0}^{N-1} \mathbf{p} = \mathbf{p}$.

Απο το προηγουμενο βημα, για δεδομενο j ολα τα $q(i,j)$ ειναι ιδια και εστω q_j η κοινη τιμη τους. Τοτε

$$\mathbf{p} Q = \mathbf{p} \Leftrightarrow \sum_i p_i q(i,j) = p_j \,, \quad \forall j \in Y \Leftrightarrow \sum_i p_i q_j = p_j \quad \forall j \in Y$$

$$\Leftrightarrow \quad q_j = p_j \,, \; \forall j \in Y,$$

αφου $\sum_i p_i = 1$. Δηλ.: $q_j = q(i,j) = p_j$ για ολα τα i κ' j.

(5) Ο T ειναι εργοδικος.

Αποδειξη: Απο την Προταση 6.4 αρκει να δειξουμε οτι $N^{-1}\sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) \to \mu(A)\mu(B)$ για A και B κυλινδρους. Εστω λοιπον $A = \{(x_i) : x_k = i_0, \dots, x_{k+r} = i_r\}$ και $B = \{(x_i) : x_l = j_0, \dots, x_{l+s} = j_s\}$ δυο κυλινδροι. Τοτε για $n \geqslant l+s-k$

$$\mu(T^{-n}A \cap B) = p_{j_0}\, p(j_0, j_1) \dots p(j_{s-1}, j_s)\; p^{(k+n-l-s)}(j_s, i_0)\; p(i_0, i_1) \dots p(i_{r-1}, i_r)$$

Ομως $N^{-1}\sum_{n=0}^{l+s-k} \mu(T^{-n}A \cap B) \to 0$ καθως $N \to \infty$ και απο τα βηματα (1) και (4)

$$\frac{1}{N}\sum_{n=l+s-k}^{N-1} p^{(k+n-l-s)}(j_s, i_0) \to q(j_s, i_0) = p_{i_0}$$

Επομενως

$$\frac{1}{N}\sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) \to p_{j_0}\, p(j_0, j_1) \dots p(j_{s-1}, j_s)\; p_{i_0}\, p(i_0, i_1) \dots p(i_{r-1}, i_r)\cdot$$

$$= \mu(B)\,\mu(A).$$

(ii) $\Rightarrow$ (iii) Απο τον ορισμο του Q αν $\mathbf{v} P = \mathbf{v}$ τοτε $\mathbf{v} Q = \mathbf{v}$. Δηλαδη καθε ιδιοδιανυσμα του P για την ιδιοτιμη 1 ειναι και ιδιοδιανυσμα του Q για την ιδιοτιμη 1. Οταν ομως ο P ειναι αναγωγος ξερουμε οτι $q(i,j) = p_j$ για ολα τα i, j και αρα μονο πολλαπλασια του $\mathbf{p}$ μπορει να ειναι ιδιοδιανυσματα του Q για την ιδιοτιμη 1. Αυτο ομως πρεπει να ισχυει και για τον P και ετσι η ιδιοτιμη 1 ειναι απλη.

(iii) $\Rightarrow$ (i) Απο τον ορισμο του Q εχουμε $QP = Q$ και αρα καθε γραμμη του Q ειναι ενα αριστερο ιδιοδιανυσμα του P για την ιδιοτιμη 1. Αφου το 1 ειναι απλη ιδιοτιμη του P καθε γραμμη του Q ειναι πολλαπλασιο οποιασδηποτε αλλης, εστω της πρωτης. Αφου καθε γραμμη πρεπει να αθροιζει στο 1 (ο Q ειναι στοχαστικος) και αρα ολες οι γραμμες του Q ειναι ιδιες. Αυτο ομως δινει οτι ο T ειναι εργοδικος απο την αποδειξη του (ii) $\Rightarrow$ (i).

Για το (Β) θα χρησιμοποιησουμε το εξης:

Λημμα: Εστω $\{a_n\}$ μια φραγμενη ακολουθια πραγματικων αριθμων. Τοτε

$$n^{-1}\sum_{i=0}^{n-1} |a_i| \to 0$$

αν και μονο αν υπαρχει $J\subset \mathbb{N}_0$ με $\frac{1}{n}\times$ πληθαριθμος$(J\cap\{0,1,\dots,n-1\})\to 0$ ωστε

$$\lim_{\substack{n\to\infty\\ n\notin J}} a_n = 0 .$$

Αποδειξη: Για $M\subset\mathbb{N}_0$ συμβολιζουμε με $\pi_n(M)$ τον πληθαριθμο του $M\cap\{0,1,\dots,n-1\}$
Εστω οτι $n^{-1}\sum_{i=0}^{n-1}|a_i|\to 0$. Θετουμε $J_k=\{n\in\mathbb{N}_0 : |a_n|>k^{-1}\}$. Τοτε $J_1\subset J_2\subset\dots$ και

$$\frac{1}{n}\pi_n(J_k)\le k\cdot\frac{1}{n}\sum_{i=0}^{n-1}|a_i| \to 0 \quad \text{καθως } n\to\infty \quad \forall k\in\mathbb{N}.$$

Αρα υπαρχουν ακεραιοι $0=l_1<l_2<\dots$ τετοιοι ωστε

$$\frac{1}{n}\pi_n(J_{k+1})<\frac{1}{k+1}\qquad \forall n\ge l_k .$$

Θετουμε $J=\bigcup_{k=0}^{\infty}(J_{k+1}\cap[l_k,l_{k+1}))$. Αν $n\ge l_k$ και $n\notin J$ τοτε $n\notin J_{k+1}$ και ετσι $|a_n|<(k+1)^{-1}$. Δηλαδη $a_n\to 0$ οταν $n\to\infty$ με $n\notin J$.

Μενει να αποδειξουμε οτι $n^{-1}\pi_n(J)\to 0$. Για $l_k\le n<l_{k+1}$ εχουμε οτι

$$J\cap[0,n)=(J\cap[0,l_k))\cup(J\cap[l_k,n))\subset(J_k\cap[0,l_k))\cup(J_{k+1}\cap[l_k,n))\subset$$

$$\subset (J_k\cap[0,n))\cup(J_{k+1}\cap[0,n))$$

αφου $J_1\subset J_2\subset\dots$ και ετσι

$$\frac{1}{n}\pi_n(J)\le\frac{1}{n}\pi_n(J_k)+\frac{1}{n}\pi_n(J_{k+1})<\frac{1}{k}+\frac{1}{k+1}$$

για $l_k\le n<l_{k+1}$. Δηλ.: $n^{-1}\pi_n(J)\to 0$.

Για το αντιστροφο εστω K ενα φραγμα της $\{a_n\}$. Εστω $\varepsilon>0$ και $N\in\mathbb{N}$ τετοιο ωστε $n>N$ και $\notin J\Rightarrow |a_n|<\varepsilon$. Τοτε $n\ge N+2$

$$\frac{1}{n}\sum_{i=0}^{n-1}|a_i|=\frac{1}{n}\sum_{i\in J\cap\{0,1,\dots,n-1\}}|a_i|+\frac{1}{n}\sum_{i\in\{0,1,\dots,N\}-J}|a_i|+\frac{1}{n}\sum_{i\in\{N+1,\dots,n-1\}-J}|a_i|$$

$$\le \frac{K}{n}\pi_n(J)+\frac{1}{n}\sum_{i=0}^{N}|a_i|+\varepsilon . \qquad \blacksquare$$

ΑΠΟΔΕΙΞΗ ΤΟΥ (B): (i) $\Rightarrow$ (ii): Εστω παλι $C_i = \{(x_n) : x_0 = i\}$. Απο τον ορισμο του ασθενως mixing $N^{-1}\sum_{n=0}^{N-1} |\mu(C_i \cap T^{-n}C_j) - \mu(C_i)\mu(C_j)| \to 0$. Αφου $\mu(C_k) = p_k$ και $\mu(C_i \cap T^{-n}C_j) = p_i\, p^{(n)}(i,j)$ επεται οτι $N^{-1}\sum_{n=0}^{N-1} |p^{(n)}(i,j) - p_j| \to 0$. Απο το λημμα υπαρχει ενα αραιο συνολο ακεραιων J ωστε

$$\lim_{\substack{n\to\infty \\ n \notin J}} p^{(n)}(i,j) = p_j \qquad \forall\,(i,j).$$

Αφου υποθετουμε οτι $p_j > 0$ για ολα τα j πρεπει να υπαρχει ενα $n \in \mathbb{N}$ ωστε $p^{(n)}(i,j)$ να ειναι θετικο για ολα τα i και j. Δηλαδη ο P ειναι αναγωγος και απεριοδικος.

(ii) $\Rightarrow$ (iii): Αυτο ειναι ενα βασικο αποτελεσμα της θεωριας των αλυσιδων Markov και υπαρχει σ' οποιοδηποτε βιβλιο πιθανοτητων κ/η στοχαστικων ανελιξεων.

(iii) $\Rightarrow$ (iv): Αν $A = \{(x_i) : x_m = i_0, \dots, x_{m+r} = i_r\}$ και $B = \{(x_i) : x_k = j_0, \dots, x_{k+s} = j_s\}$ ειναι δυο κυλινδροι και $n > k+s-m$ τοτε

$$\mu(T^{-n}A \cap B) = p_{j_0}\, p(j_0,j_1)\dots p(j_{s-1},j_s)\, p^{(n+m-k-s)}(j_s,i_0)\, p(i_0,i_1)\dots p(i_{r-1},i_r) \to$$

$$\to p_{j_0}\, p(j_0,j_1)\dots p(j_{s-1},j_s)\, p_{i_0}\, p(i_0,i_1)\dots p(i_{r-1},i_r) =$$

$$= \mu(A)\mu(B).$$

Απο την Προταση 6.4 ο T ειναι ισχυρα mixing.

(iv) $\Rightarrow$ (i): Προταση 6.3.

Πορισμα: Τα ιδια ισχυουν για μονοπλευρα μαρκοβιανα shift.

Πορισμα: Το μονοπλευρο και διπλευρο (απλο) $\mathbf{p}$-shift ειναι ισχυρα mixing.

ΑΠΟΔΕΙΞΗ: Το απλο $\mathbf{p}$-shift ειναι $(\mathbf{p}, P)$-μαρκοβιανο shift, με $p(i,j) = p_j \;\; \forall i$. Εχουμε οτι $P^n = P$ για ολα τα n και αρα $p^{(n)}(i,j) = p(i,j) = p_j$.

7. MIXING ΚΑΙ L^2-ΘΕΩΡΙΑ

Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας και $(f_n)_{n\in\mathbb{Z}}$ μια ακολουθια μετρησιμων μιγαδικων συναρτησεων του X. Η ακολουθια (f_n) λεγεται στασιμη αν για καθε $n\in\mathbb{Z}$, $k\in\mathbb{N}_0$ και $A_0, \dots, A_k \in \mathcal{B}(\mathbb{C})$ (Borel υποσυνολα του $\mathbb{C}$), το μετρο $\mu\{x\in X : f_n(x)\in A_0, f_{n+1}(x)\in A_1, \dots, f_{n+k}(x)\in A_k\}$ ειναι ανεξαρτητο του n. Δηλαδη αν

$$\forall n\in\mathbb{Z}, \qquad \mu\{x\in X : f_n(x)\in A_0, \dots, f_{n+k}(x)\in A_k\} = \mu\{x\in X : f_0(x)\in A_0, \dots, f_k(x)\in A_k\}$$

Αν η (f_n) ειναι στασιμη και $f_0\in L^2(\mu)$, οποτε $f_n\in L^2(\mu)$ για καθε n, οριζουμε την ακολουθια συνδιακυμανσεων της (f_n) σαν την ακολουθια των μιγαδικων αριθμων

$$\alpha_n := (f_0, f_n) = \int_X f_0 \bar{f_n}\, d\mu \qquad (n\in\mathbb{Z}).$$

Παρατηρουμε οτι απο την στασιμοτητα της (f_n)

$$\alpha_n = (f_m, f_{m+n}) \qquad \forall\, m, n\in\mathbb{Z}$$

και απο την ανισοτητα Cauchy - Schwartz

$$|\alpha_n| \leq \|f_0\|_2^2 = \alpha_0 \qquad \forall n\in\mathbb{Z}.$$

Τελος οριζουμε σαν κεντροποιημενη ακολουθια συνδιακυμανσεων της (f_n) την ακολουθια

$$\alpha_n^* = \left(f_0 - \int_X f_0\, d\mu ,\ f_n - \int_X f_n\, d\mu \right)$$

$$= (f_0, f_n) - \left| \int_X f_0\, d\mu \right|^2 .$$

7.1 Θεωρημα (Herglotz) Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας και $(f_n)_{n\in\mathbb{Z}}$ μια στασιμη ακολουθια μιγαδικων συναρτησεων του X με $f_0\in L^2(\mu)$. Τοτε υπαρχει μοναδικο μετρο πιθανοτητας ν, στον κυκλο S^1, ωστε

$$\frac{(f_0, f_n)}{\|f_0\|_2^2} = \hat{\nu}(n) := \int_{[0,1)} e^{-2\pi i n\theta}\, \nu(d\theta) \qquad \forall n\in\mathbb{Z}.$$

Το μετρο ν ονομαζεται το φασματικο μετρο της $(f_n)_{n\in\mathbb{Z}}$ (η καλυτερα της ακολουθιας συνδιακυμανσεων $(\alpha_n)_{n\in\mathbb{Z}}$).

Παρατηρήσεις: (1) Η μοναδικότητα του μέτρου ν είναι προφανής αφού ένα μέτρο πιθανότητας στον κύκλο S^1 καθορίζεται μονοσήμαντα από τους συντελεστές Fourier $\hat{\nu}(n) = \int e^{-2\pi i n\theta}\nu(d\theta)$.

(2) Αρκεί να αποδείξουμε το Θεώρημα για την περίπτωση που $\int_X f_0\, d\mu = 0$. Πράγματι αν $\int_X f_0\, d\mu \neq 0$ τότε ορίζουμε $g_n = f_n - \int_X f_0\, d\mu$. Τότε $\int_X g_n\, d\mu = \int_X g_0\, d\mu = 0$ για κάθε n, η $(g_n)_{n\in\mathbb{Z}}$ είναι προφανώς στάσιμη και $g_0 \in L^2(\mu)$ και $(g_0, g_n) = (f_0, f_n) - \left|\int_X f_0\, d\mu\right|^2$. Αν ν_* το φασματικό μέτρο της (g_n) τότε

$$\|f_0\|_2^{-2}(f_0, f_n) = \|f_0\|_2^{-2}(g_0, g_n) + \|f_0\|_2^{-2}\left|\int_X f_0\, d\mu\right|^2 =$$

$$= \frac{\|g_0\|_2^2}{\|f_0\|_2^2}\int_{[0,1)} e^{-2\pi i n\theta}\, d\nu_*(\theta) + \frac{\left|\int_X f_0\, d\mu\right|^2}{\|f_0\|_2^2}\int_{[0,1)} e^{-2\pi i n\theta}\, d\delta_{\{0\}}(\theta)$$

$$= \int_{[0,1)} e^{-2\pi i n\theta}\, d\nu(\theta)$$

όπου ν είναι το μέτρο $\nu = \frac{\|g_0\|_2^2}{\|f_0\|_2^2}\nu_* + \frac{\left|\int_X f_0\, d\mu\right|^2}{\|f_0\|_2^2}\delta_{\{0\}}$, που είναι μέτρο πιθανότητας σαν κυρτός συνδυασμός μέτρων πιθανότητας. Εδώ $\delta_{\{0\}}$ είναι το μέτρο $\delta_{\{0\}}(A) = 1$ αν και μόνο αν $0 \in A$.

Από τα παραπάνω συνάγουμε ότι αν $\int_X f_0\, d\mu \neq 0$ τότε το φασματικό μέτρο ν της (f_n) δίνει θετική μάζα στο $0 \in S^1$.

Για την απόδειξη του Θεωρήματος του Herglotz θα χρειαστούμε το επόμενο

7.2 Λήμμα: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(f_n)_{n\in\mathbb{Z}}$ μια στάσιμη ακολουθία μιγαδικών συναρτήσεων με $f_0 \in L^2(\mu)$ και $\int_X f_0\, d\mu = 0$. Έστω επίσης $r \in (0,1)$. Τότε υπάρχει, σε κάποιο χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$, μια ακολουθία μετρήσιμων και L^2 συναρτήσεων $(h_0, h_1, \ldots)$, ώστε για κάθε $n \in \mathbb{N}_0$

$$(h_n, h_{n+k})_{L^2(P)} = (f_0, f_k)\, r^{|k|} \qquad \forall k \in \mathbb{Z}\cap[-n, +\infty).$$

Η απόδειξη του λήμματος θα δοθεί μετά την απόδειξη του Θεωρήματος Herglotz.

Απόδειξη Θεωρήματος Herglotz: Ορίζουμε, για κάθε $0 < r < 1$

$$h_r(\theta) = \sum_{n=-\infty}^{\infty} \alpha_n r^{|n|} e^{2\pi i n\theta}, \qquad \theta \in [0,1)$$

όπου έχουμε θέσει $\alpha_n = (f_0, f_n)$, $n \in \mathbb{Z}$. Η h_r είναι προφανώς μια καλά ορισμένη συνεχής συνάρτηση του κύκλου S^1. Επιπλέον, αφού $\sum_{n=-\infty}^{\infty}\int_{[0,1)}|\alpha_n r^{|n|} e^{2\pi i n\theta}|\, d\theta < +\infty$,

$$\int_{[0,1)} h_r(\theta)\, d\theta = \sum_{n=-\infty}^{\infty} a_n r^{|n|} \int_{[0,1)} e^{-2\pi i n\theta} d\theta = a_0 .$$

Θα αποδείξουμε οτι για καθε $r \in (0,1)$ η σχέση

$$\nu_r(A) = \int_A \frac{h_r(\theta)}{a_0}\, d\theta \qquad \text{για} \quad A \in \mathcal{B}(S^1)$$

ορίζει ένα μέτρο πιθανοτητας στον S^1. Ηδη γνωρίζουμε οτι $\nu_r(S^1) = 1$ και προφανώς $\nu_r(\cup_i A_i) = \sum_i \nu_r(A_i)$ οποτε τα A_i ειναι ανα δυο ξένα. Αρκει λοιπον να αποδείξουμε οτι $\nu_r(A) \geq 0$ για καθε $A \in \mathcal{B}(S^1)$. Αυτο ομως θα ισχύει αν αποδείξουμε οτι για καθε συνεχή συναρτηση $g: S^1 \to \mathbb{R}$ με $g \geq 0$ ισχύει και $\int_{[0,1)} g(\theta) h_r(\theta) \frac{d\theta}{a_0} \geq 0$.

Εστω λοιπον μια συνεχης $g: S^1 \to [0,+\infty)$. Τοτε $\sqrt{g} \in L^2(d\theta)$ και εστω το αναπτυγμα Fourier της $\sqrt{g}$: $\sum_{n=-\infty}^{\infty} b_n e^{+2\pi i n\theta}$, $\theta \in [0,1)$. Τοτε εξ ορισμού

$$\Big\| \sqrt{g} - \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big\|_{L^2(d\theta)} \to 0 \qquad (N \to +\infty).$$

Τοτε

$$\int_{[0,1)} g(\theta) h_r(\theta) \frac{d\theta}{a_0} = \int_{[0,1)} \Big| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big|^2 h_r(\theta) \frac{d\theta}{a_0} +$$

$$+ \int_{[0,1)} \Big[g(\theta) - \Big| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big|^2 \Big] h_r(\theta) \frac{d\theta}{a_0}$$

και για τον τελευταιο ορο έχουμε οτι

$$\Big| \int_{[0,1)} \Big[g(\theta) - \Big| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big|^2 \Big] h_r(\theta) \frac{d\theta}{a_0} \Big| \leq$$

$$\leq \frac{\|h_r\|_\infty}{a_0} \int_{[0,1)} \Big| \sqrt{g(\theta)} - \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big| \Big| \sqrt{g(\theta)} + \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big| d\theta \leq$$

$$\leq \frac{\|h_r\|_\infty}{a_0} \Big\| \sqrt{g(\theta)} - \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big\|_{L^2(d\theta)} \Big\| \sqrt{g(\theta)} + \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big\|_{L^2(d\theta)}$$

$$\leq \frac{\|h_r\|_\infty}{a_0} \Big(\|\sqrt{g}\|_{L^2(d\theta)} + \Big\| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big\|_{L^2(d\theta)} \Big) \Big\| \sqrt{g(\theta)} - \sum_{n=-N}^{N} b_n e^{2\pi i n\theta}$$

$$\to 0 \quad , \quad \text{καθως } N \to \infty,$$

αφου ο τελευταιος παραγοντας τείνει στο 0, ενώ

$$\Big\| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \Big\|^2_{L^2(d\theta)} = \sum_{m,n=-N}^{N} b_n \bar{b}_m \int_{[0,1)} e^{2\pi i (n-m)\theta} d\theta = \sum_{n=-N}^{N} |b_n|^2 \leq \|\sqrt{g}\|^2_{L^2(d\theta)} .$$

Επομένως, για να αποδείξουμε ότι $\int_{[0,1)} g(\theta)\, h_r(\theta)\, \frac{d\theta}{a_0} \geq 0$, αρκεί να δείξουμε ότι $\int_{[0,1)} \left| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \right|^2 h_r(\theta) \frac{d\theta}{a_0} \geq 0$ για κάθε N. Έχουμε λοιπόν ότι:

$$\int_{[0,1)} \left| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \right|^2 \frac{h_r(\theta)}{a_0}\, d\theta = \sum_{k=-\infty}^{\infty} \sum_{n,m=-N}^{N} a_k r^{|k|} b_n \bar{b}_m \int_{[0,1)} e^{2\pi i (k-m+n)\theta} \frac{d\theta}{a_0} =$$

$$= \sum_{n,m=-N}^{N} \frac{a_{m-n}}{a_0} r^{|m-n|} b_n \bar{b}_m .$$

Αφού $\int_X f_0\, d\mu = 0$ από το Λήμμα 7.2 υπάρχει κάποια ακολουθία $(h_0, h_1, \dots)$, σε κάποιο χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$, τέτοια ώστε $\frac{a_{m-n}}{a_0} r^{|m-n|} = a_0^{-1} (h_{N+n}, h_{N+m})$ και άρα

$$\int_{[0,1)} \left| \sum_{n=-N}^{N} b_n e^{2\pi i n\theta} \right|^2 \frac{h_r(\theta)}{a_0}\, d\theta = \sum_{n,m=-N}^{N} b_n \bar{b}_m (h_{m+N}, h_{n+N})\, a_0^{-1} =$$

$$= \frac{1}{a_0} \int_{\Omega} \left| \sum_{n=-N}^{N} b_n h_{n+N} \right|^2 dP \geq 0$$

αφού $a_0 = \| f_0 \|_2^2 > 0$.

Αποδείξαμε λοιπόν ότι τα $\nu_r(d\theta) = \frac{h_r(\theta)}{a_0}\, d\theta$ ορίζουν για κάθε $0 < r <$ ένα μέτρο πιθανότητας στον S^1.

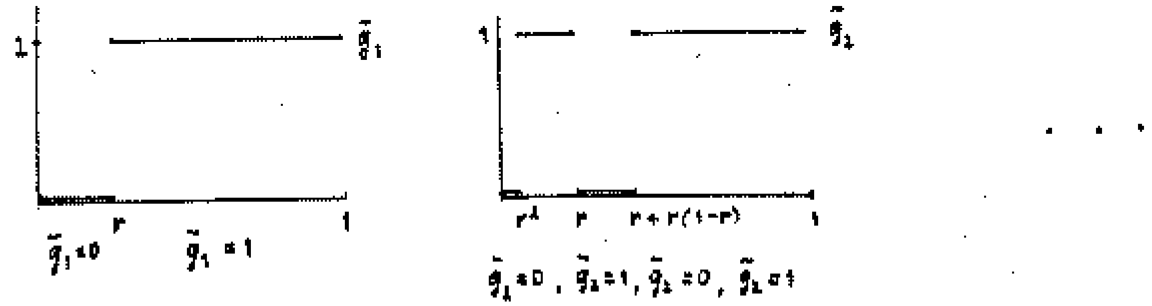
Έστω τώρα $(r_j)_{j\in\mathbb{N}}$ μια ακολουθία στον $(0,1)$ ώστε $r_j \to 1$. Από συμπάγεια, υπάρχει μια υπακολουθία $\{ \nu_{r_{j(k)}} \}_{k\in\mathbb{N}}$ η οποία συγκλίνει ασθενώς (weak*) σε κάποιο μέτρο πιθανότητας ν στον S^1. Το ν είναι το ζητούμενο μέτρο αφού

$$\frac{a_n}{a_0} = \lim_{k\to\infty} \frac{a_n}{a_0} r_{j(k)}^{|n|} = \lim_{k\to\infty} \int_{[0,1)} e^{-2\pi i n\theta}\, \nu_{r_{j(k)}}(d\theta) = \int_{[0,1)} e^{-2\pi i n\theta}\, \nu(d\theta). \blacksquare$$

Απόδειξη Λήμματος 7.2 : Ορίζουμε στον $([0,1], \mathcal{B}([0,1]), \lambda)$ (λ = Lebesgue) μια ακολουθία συναρτήσεων $\tilde{g}_n : [0,1] \to \{0,1\}$ ανεξαρτήτων μεταξύ τους ώστε $\lambda\{x : \tilde{g}_n(x) = 0\} = r = 1 - \lambda\{\tilde{g}_n(x) = 1\}$. Η ανεξαρτησία των $\tilde{g}_n$ είναι ισοδύναμη με

$$\lambda\{ x : \tilde{g}_1(x) = u_1, \dots, \tilde{g}_n(x) = u_n \} = \prod_{i=1}^{n} \lambda\{ x : \tilde{g}_i(x) = u_i \}$$

για κάθε $n \in \mathbb{N}$ και $u_i \in \{0,1\}$. Αυτό μπορεί να γίνει για παράδειγμα ως εξής:



Τωρα ορίζουμε $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \lambda) \times (X, \mathcal{A}, \mu)^{\mathbb{N}_0}$ (χωρο γινομενο μετροθεωρητικό) και για καθε $(x, \omega_0, \omega_1, \ldots) \in \Omega$ ορίζουμε

$$g_n(x, \omega_0, \omega_1, \ldots) = \tilde{g}_n(x) \qquad n \in \mathbb{N}$$

$$f_n^{(j)}(x, \omega_0, \omega_1, \ldots) = f_n(\omega_j) \qquad n \in \mathbb{Z}.$$

Τοτε επειδη το P ειναι μετρο γινομενο οι ακολουθιες $(g_n)_{n\in\mathbb{N}}$, $(f_n^{(0)})_{n\in\mathbb{Z}}$, $(f_n^{(1)})_{n\in\mathbb{Z}}$, ... ειναι ανεξαρτητες μεταξυ τους, δηλ.:

$$P\left((g_n)_{n\in\mathbb{N}} \in A,\ (f_n^{(0)})_{n\in\mathbb{Z}} \in B_0, \ldots, (f_n^{(j)})_{n\in\mathbb{Z}} \in B_j \right) =$$

$$= P\left((g_n)_{n\in\mathbb{N}} \in A \right) \quad P((f_n^{(0)}) \in B_0) \quad \ldots \quad P\left((f_n^{(j)})_{n\in\mathbb{Z}} \in B_j \right)$$

για καθε j. Επισης η συγκατανομή της $(g_n)_{n\in\mathbb{N}}$ ειναι ιδια με της $(\tilde{g}_n)_{n\in\mathbb{N}}$, δηλ.: οι συναρτησεις $g_1, g_2, \ldots$ ειναι ανεξαρτητες μεταξυ τους και $P(g_n = 1) = 1-r$ και η συγκατανομή καθε $(f_n^{(j)})_{n\in\mathbb{Z}}$ ειναι ιδια με τη συγκατανομή της $(f_n)_{n\in\mathbb{Z}}$, δηλ.

$$P\left(f_n^{(j)} \in A_0, \ldots, f_{n+k}^{(j)} \in A_k \right) = \mu\{ f_n \in A_0, \ldots, f_{n+k} \in A_k \} =$$

$$= \mu\{ f_0 \in A_0, \ldots, f_k \in A_k \}$$

για καθε j και καθε k.

Τωρα ορίζουμε $S_n = g_1 + \ldots + g_n$, $S_0 \equiv 0$ και

$$h_n(\omega) = f_n^{(S_n(\omega))}(\omega) \qquad n \in \mathbb{N}_0, \quad \omega \in \Omega.$$

Τοτε για $m \in \mathbb{N}_0$ και $n \geq 0$

$$(h_m, h_{m+n}) = \int_\Omega h_m \bar{h}_{m+n}\, dP = \int_\Omega \sum_{i=0}^{m} \sum_{j=0}^{n+m} f_m^{(i)} \overline{f_{m+n}^{(j)}}\ \mathbb{1}_{\{i\}}(S_m) \mathbb{1}_{\{j\}}(S_{n+m})\, dP$$

$$= \sum_{i=0}^{m} \int_{\{S_m = S_{n+m} = i\}} f_m^{(i)} \overline{f_{n+m}^{(i)}}\, dP \; +$$

$$+ \sum_{i=0}^{m} \sum_{j=0,\, j\neq i}^{n+m} \int_{\{S_m = i,\, S_{n+m} = j\}} f_m^{(i)} \overline{f_{m+n}^{(j)}}\, dP =$$

$$= \sum_{i=0}^{m} \int_\Omega f_m^{(i)} \overline{f_{n+m}^{(i)}}\, dP \cdot P(S_m = S_{n+m} = j) \; +$$

$$+ \sum_{i=0}^{m} \sum_{j=0,\, j\neq i}^{n+m} \int_\Omega f_m^{(i)}\, f_{n+m}^{(j)}\, dP \quad P(S_m = i, S_{n+m} = j)$$

$$= \sum_{i=0}^{m} \int_X f_m \overline{f}_{m+n}\, d\mu \;\; P(S_m = i,\, g_{m+1} = \dots = g_{m+n} = 0) +$$

$$+ \sum_{i=0}^{m} \sum_{\substack{j=0\\ j\neq i}}^{n+m} \int_\Omega f_m^{(i)}\, dP \int_\Omega \overline{f_{m+n}^{(j)}}\, dP \;\; P(S_m = i,\, S_{n+m} = j) =$$

$$= \sum_{i=0}^{m} (f_m, f_{m+n})\; P(S_m = i)\; P(g_{m+1} = 0, \dots, g_{m+n} = 0) +$$

$$+ \sum_{i=0}^{m} \sum_{\substack{j=0\\ j\neq i}}^{n+m} \int_X f_m\, d\mu \int_X \overline{f}_{n+m}\, d\mu \;\; P(S_m = i,\, S_{n+m} = j) =$$

$$= (f_m, f_{n+m})\; \lambda\{\tilde{g}_{m+1} = 0, \dots, \tilde{g}_{m+n} = 0\} \sum_{i=0}^{m} P(S_m = i) +$$

$$+ \int_X f_0\, d\mu \int_X \overline{f}_0\, d\mu \sum_{i=0}^{m} \sum_{j\neq 0}^{n+m} P(S_m = i,\, S_{n+m} = j) =$$

$$= (f_m, f_{n+m})\; r^{(n)}$$

Όμοια γίνεται υπολογισμός και για $n<0$. ∎

Εστω τώρα $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ αυτομορφισμός. Ορίζουμε $U_T : L^2(\mu) \to L^2(\mu)$ με $U_T f = f\circ T$ και λέμε ότι κάποιο $\lambda \in \mathbb{C}$ είναι ιδιοτιμή του T αν είναι ιδιοτιμή του U_T, δηλ. υπάρχει $f \in L^2(\mu)$, με $f \neq 0$ ώστε $U_T f = \lambda f$.

Παρατηρούμε ότι αν λ είναι ιδιοτιμή του T, πρέπει $|\lambda| = 1$ αφού αν $f \neq 0$ είναι η αντίστοιχη ιδιοτιμή τότε $|\lambda| \int_X |f|\, d\mu = \int_X |f\circ T|\, d\mu = \int_X |f|\, d\mu$, αφού ο T διατηρεί το μέτρο μ. Επίσης $\lambda = 1$ είναι πάντοτε ιδιοτιμή του T με αντίστοιχες ιδιοσυναρτήσεις τις σταθερές.

Αν οι σταθερές είναι οι μόνες ιδιοσυναρτήσεις του T (οπότε $\lambda = 1$ είναι η μόνη ιδιοτιμή) τότε λέμε ότι ο T έχει συνεχές φάσμα. Παρατηρούμε ότι από την Πρόταση 4.4 έπεται ότι ο T έχει συνεχές φάσμα αν και μόνο αν είναι εργοδικός και $\lambda = 1$ είναι η μόνη ιδιοτιμή του.

Παρατηρούμε ότι αν $f \in L^2(\mu)$ (η απλώς μετρήσιμη) η ακολουθία $(U_T^n f)_{n\in\mathbb{Z}} = (f\circ T^n)_{n\in\mathbb{Z}}$ είναι στάσιμη και άρα (εδώ χρειαζόμαστε $f \in L^2(\mu)$) έχει ένα μοναδικό φασματικό μέτρο ν_f. Θα χρειαστούμε το επόμενο

7.3 Λήμμα: Εστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, $T: X \to X$ αυτομορφισμός και $f \in L^2(\mu)$ όχι ταυτοτικά μηδέν. Αν ο T έχει συνεχές φάσμα και $\int_X f\, d\mu = 0$ τότε το ν_f δεν έχει άτομα, δηλ. $\nu_f(\{\theta\}) = 0$, $\forall\, \theta \in [0,1)$.

Απόδειξη: Εστω $\langle f\circ T^n : n \in \mathbb{Z} \rangle$ ο γραμμικός χώρος όλων των γραμμικών συνδυασμών (πεπερασμένων) της μορφής $\sum_{n=-N}^{M} b_n (f\circ T^n)$, όπου $b_n \in \mathbb{C}$. Εστω

$\mathcal{H}_f = \overline{\langle f\circ T^n : n\in\mathbb{Z}\rangle}$ η κλειστότητα (στον $L^2(\mu)$) του χώρου αυτού. Ορίζουμε $e_n(\theta) = e^{2\pi i n\theta}$, $\theta\in[0,1)$ και

$$\mathcal{Z}(f\circ T^n) = e_n \qquad (n\in\mathbb{Z}).$$

Τότε $(f\circ T^n, f\circ T^m)_{L^2(\mu)} = \|f\|_2^2\,(e_n, e_m)_{L^2(\nu_f)} = \|f\|_2^2\,(\mathcal{Z}(f\circ T^n), \mathcal{Z}(f\circ T^m))_{L^2(\nu_f)}$ από τον ορισμό του φασματικού μέτρου. Λόγω αυτού ο $\mathcal{Z}$ επεκτείνεται γραμμικά στον $\langle f\circ T^n : n\in\mathbb{Z}\rangle$. Πραγματι αν $\sum b_n f\circ T^n$ και $\sum c_n f\circ T^n$ είναι δυο εκφράσεις της ίδιας $h\in\langle f\circ T^n : n\in\mathbb{Z}\rangle$ τότε

$$0 = \|h-h\|^2_{L^2(\mu)} = \Big\|\sum b_n f\circ T^n - \sum c_n f\circ T^n\Big\|^2_{L^2(\mu)} =$$

$$= \sum b_n \bar{b}_m (f\circ T^n, f\circ T^m) + \sum c_n \bar{c}_m (f\circ T^n, f\circ T^m) - \sum b_n \bar{c}_m (f\circ T^n, f\circ T^m) -$$

$$- \sum \bar{b}_n c_m (f\circ T^m, f\circ T^n) =$$

$$= \|f\|_2^2\Big[\sum b_n \bar{b}_m (e_n, e_m) + \sum c_n \bar{c}_m (e_n, e_m) - \sum b_n \bar{c}_m (e_n, e_m) - \sum \bar{b}_n c_m (e_m, e_n)\Big]$$

$$= \|f\|_2^2\ \Big\|\sum b_n e_n - \sum c_n e_n\Big\|_{L^2(\nu_f)}$$

και άρα οι $\sum b_n e_n$, $\sum c_n e_n$ ορίζουν την ίδια συνάρτηση στον $L^2(\nu_f)$. Επιπλέον $\|\mathcal{Z}(h) - \mathcal{Z}(g)\|_{L^2(\nu_f)} = \|f\|_2^{-2}\,\|h-g\|_{L^2(\mu)}$ και άρα η $\mathcal{Z}$ μπορεί να επεκταθεί σε ένα γραμμικό τελεστή σ' όλο το $\mathcal{H}_f$. Ο $\mathcal{Z}$ θα έχει δε την ιδιότητα

$$\|f\|_2^2\,(\mathcal{Z}(h), \mathcal{Z}(g))_{L^2(\nu_f)} = (h,g)_{L^2(\mu)} \qquad \forall\, h, g\in\mathcal{H}_f.$$

Τώρα αν $h = \sum b_n f\circ T^n$ τότε $U_T h = h\circ T = \sum b_n f\circ T^{n+1}$ και άρα

$$\mathcal{Z}(h\circ T) = \sum b_n e_{n+1} = e_1 \cdot \sum b_n e_n = e_1\,\mathcal{Z}(h).$$

Δηλαδή η δράση του T στον $\mathcal{H}_f$ αντιστοιχεί σε πολλαπλασιασμό με την συνάρτηση $e_1(\cdot)$ στον $L^2(\nu_f)$. Τέλος ο $\mathcal{Z}$ είναι επί του $L^2(\nu_f)$ αφού οι συναρτήσεις e_n είναι πυκνές στο $C(S^1)$ και ο τελευταίος πυκνός στο $L^2(\nu_f)$ (αφού το ν_f είναι μέτρο πιθανότητας και ο S^1 μετρικός χώρος).

Έστω τώρα ότι $\nu_f(\{\theta\}) > 0$ για κάποιο θ. Ορίζουμε $g = \mathbb{1}_{\{\theta\}}\,(\nu\{\theta\})^{-1}$ και $h\in L^2(\mu)$ ώστε $\mathcal{Z}(h) = g$. Τότε

(7.1) $\quad (h, h\circ T) = \|f\|_2^2\,(\mathcal{Z}(h), \mathcal{Z}(h\circ T)) = \|f\|_2^2\,(g, e^{2\pi i\theta} g) = \|f\|_2^2\, e^{-2\pi i\theta}\,\|g\|_2^2 =$

$$= \|f\|_2^2 \; e^{-2\pi i\theta} \frac{1}{\nu_f\{\theta\}}$$

και αρα

$$|(h, h\circ T)| = \frac{\|f\|_2^2}{\nu_f\{\theta\}} = \|f\|_2^2 \, \|g\|_2^2 = \|h\|_2^2 = \|h\|_2 \, \|h\circ T\|_2$$

και η ανισοτητα Cauchy-Schwartz ισχύει σαν ισοτητα. Επεται οτι $h\circ T = \lambda h$ για καποιο $\lambda\in\mathbb{C}$ και επειδη ο T εχει συνεχές φασμα, πρεπει $\lambda = 1$ κ' $h \equiv c$ με $c\in\mathbb{C}$ σταθερά. Τοτε ομως απο την (7.2)

$$e^{2\pi i\theta} = \frac{\|f\|_2^2}{\nu_f\{\theta\}} \cdot \frac{1}{(h, h\circ T)} = \frac{\|f\|_2^2}{\nu_f\{\theta\}} \; |c|^{-2}$$

δηλ. το $e^{2\pi i\theta}$ ειναι ενας θετικος πραγματικος αριθμός. Επεται οτι $\theta = 0$.

Αποδείξαμε μεχρι τωρα οτι αν $\nu_f(\{\theta\}) > 0$ τοτε $\theta = 0$. Τωρα θα αποδείξουμε οτι $\nu(\{0\}) = 0$. Πραγματι:

$$\frac{1}{N}\sum_{n=0}^{N-1} (f, f\circ T^n) = \frac{1}{N}\sum_{n=0}^{N-1} \|f\|_2^2 \int_{[0,1)} e^{-2\pi i n\theta} d\nu_f(\theta) =$$

$$= \int_{[0,1)} \|f\|_2^2 \; \frac{1}{N}\sum_{n=0}^{N-1} e^{-2\pi i n\theta} d\nu_f(\theta).$$

Αφου $N^{-1}\sum_{n=0}^{N-1} e^{-2\pi i n 0} \to 1$ και για $\theta \neq 0$

$$\left|\frac{1}{N}\sum_{n=0}^{N-1} e^{-2\pi i n\theta}\right| = \frac{1}{N}\left|\frac{1-e^{-2\pi i N\theta}}{1-e^{-2\pi i\theta}}\right| \leq \frac{2}{N}\,|1-e^{-2\pi i\theta}| \to 0$$

επεται απο το Θεωρημα φραγμένης συγκλισης του Lebesgue οτι

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} (f, f\circ T^n) = \|f\|_2^2 \int_{[0,1)} \mathbb{1}_{\{0\}} \, d\nu_f = \|f\|_2^2 \, \nu_f(\{0\}).$$

Απο το εργοδικο θεωρημα ομως (η καλύτερα απο την Προταση 6.5)

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} (f, f\circ T^n) = (f,1)(1,f) = \left|\int_X f\,d\mu\right| = 0. \qquad \blacksquare$$

7.4 Θεώρημα: Εστω $(X, \mathcal{A}, \mu)$ χωρος πιθανοτητας και $T: X\to X$ αυτομορφισμός. Τοτε ο T ειναι ασθενως mixing αν και μονο αν εχει συνεχές φασμα.

Αποδειξη: Εστω κατ' αρχην οτι ο T ειναι ασθενως mixing. Εστω $\lambda\in\mathbb{C}$ και $f\in L^2(\mu)$ με $f\not\equiv 0$ ωστε $f\circ T = \lambda f$. Απο την Προταση 6.5

$$\frac{1}{N}\sum_{n=0}^{N-1}|(f,f\circ T^n)-(f,1)(1,f)|\to 0.$$

Αν $\lambda\neq 1$ τοτε, αφου $\int_X f\,d\mu=\int_X f\circ T\,d\mu=\lambda\int_X f\,d\mu$, πρεπει $\int_X f\,d\mu=0$.
Αρα πρεπει

$$0=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|(f,f\circ T^n)|=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\bar\lambda^{-n}(f,f)|=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\lambda^n|\,\|f\|_2^2$$

$$=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\|f\|_2^2=\|f\|_2^2$$

οποτε $f\equiv 0$. Αφου υποθεσαμε οτι $f\not\equiv 0$ πρεπει $\lambda=1$ και επιπλεον

$$0=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|(f,f\circ T^n)-(f,1)(1,f)|=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left|(f,f)-\int f\,d\mu\int\bar f\,d\mu\right|=$$

$$=\left|\left(f-\int f\,d\mu,\ f\right)\right|=\left|\left(f-\int f\,d\mu,\ f-\int f\,d\mu\right)\right|=\left\|f-\int f\,d\mu\right\|_2^2$$

αφου $\left(f-\int f\,d\mu,\ f-\int f\,d\mu\right)=\left(f-\int f\,d\mu,\ f\right)-\left(f-\int f\,d\mu,\ \int f\,d\mu\right)=$
$=\left(f-\int f\,d\mu,\ f\right)-\int\left(f-\int f\,d\mu\right)\overline{\int f\,d\mu}\,d\mu=\left(f-\int f\,d\mu,\ f\right)-\int\bar f\,d\mu\cdot\left(\int f\,d\mu-\int f\,d\mu\right)=$
$=\left(f-\int f\,d\mu,\ f\right)$.
Επεται οτι $f\equiv\int f\,d\mu$ δηλ. f σταθερη. Αποδειξαμε δηλ. οτι ο T εχει συνεχες φασμα.

Εστω τωρα ο T εχει συνεχες φασμα. Θα αποδειξουμε οτι ειναι ασθενως mixing. Αρκει να αποδειξουμε οτι

$$\frac{1}{N}\sum_{n=0}^{N-1}|(f\circ T^n,g)-(f,1)(1,g)|\to 0\qquad\forall f,g\in L^2(\mu)$$

και απο το Λημμα 7.5 παρακατω αρκει να το αποδειξουμε αυτο για $g=f$ δηλ. να αποδειξουμε οτι $N^{-1}\sum_{n=0}^{N-1}|(f\circ T^n,f)-(f,1)(1,f)|\to 0$. Επειδη

$$N^{-1}\sum_{n=0}^{N-1}|(f\circ T^n,f)-(f,1)(1,f)|=N^{-1}\sum_{n=0}^{N-1}|(f\circ T^n,f)-(f\circ T^n,1)(1,f)|=$$

$$=N^{-1}\sum_{n=0}^{N-1}\left|\left(f\circ T^n-\int f\circ T^n d\mu,\ \ f-\int f\,d\mu\right)\right|$$

(οπως παραπανω) αρκει να αποδειξουμε οτι

$$\frac{1}{N}\sum_{n=0}^{N-1}|(f\circ T^n,f)|\to 0\qquad\forall f\in L^2(\mu),\ \ \int f\,d\mu=0.$$

Τελος απο το Λημμα στη σελιδα 40 αυτων των σημειωσεων, αν $\{a_n\}$ ειναι μια φραγμενη ακολουθια πραγματικων αριθμων τοτε $N^{-1}\sum_{n=0}^{N-1}|a_n|\to 0$ αν και μονο αν $N^{-1}\sum_{n=0}^{N-1}|a_n|^2\to 0$. Ετσι αρκει να αποδειξουμε οτι

$$\frac{1}{N}\sum_{n=0}^{N-1} |(f\circ T^n, f)|^2 \to 0 \qquad \forall f\in L^2(\mu) \text{ με } \int f\, d\mu = 0.$$

Εστω λοιπον μια $f\in L^2(\mu)$ με $\int f\, d\mu = 0$. Εστω ν το φασματικο μετρο της στάσιμης L^2 ακολουθίας $(f\circ T^n)_{n\in\mathbb{Z}}$. Τοτε

$$\frac{1}{N}\sum_{n=0}^{N-1} |(f\circ T^n, f)|^2 = \frac{1}{N}\sum_{n=0}^{N-1} (f, f\circ T^n)(f\circ T^n, f) =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \int_{[0,1)} e^{-2\pi i n\theta} d\nu(\theta) \int_{[0,1)} e^{2\pi i n t}\, d\nu(t) =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \int_{[0,1)^2} e^{2\pi i n(t-\theta)}\ \nu\times\nu\,(d(\theta,t)) =$$

$$= \int_{[0,1)^2} \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i n(t-\theta)}\ \nu\times\nu\,(d(\theta,t)) \longrightarrow$$

$$\longrightarrow \int_{[0,1)^2} \mathbb{1}_{\{(x,y):x=y\}}(t,\theta)\ \nu\times\nu\,(d(\theta,t)) =$$

$$(\text{Fubini}) \quad = \int_{[0,1)} \nu(\{\theta\})\, \nu(d\theta) = 0$$

αφου το ν δεν εχει ατομα: ο T εχει συνεχές φασμα ισ' $\int f\, d\mu = 0$.
Η συγκλιση στην 4η προς 5η σειρα παραπανω επεται απο την συγκλιση

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i n\theta} \longrightarrow \mathbb{1}_{\{0\}}(\theta) \qquad \theta\in[0,1),$$

που αποδείξαμε στη σελιδα 49.

Η αποδειξη οτι ο T ειναι ασθενως mixing ειναι πληρης εκτος απο το

7.5 Λημμα: Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας, $T: X\to X$ ενδομορφισμός. Τοτε ο T ειναι ασθενως mixing αν και μονο αν

$$\frac{1}{N}\sum_{n=0}^{N-1} |(f\circ T^n, f) - (f,1)(1,f)| \to 0 \qquad \forall f\in L^2(\mu).$$

Παρατηρηση: Αναλογα αποτελέσματα που συμπληρώνουν την Προταση 6.5 ισχύουν και για εργοδικοτητα και ισχυρο mixing.

Αποδειξη: Εστω $\mathcal{H}_f$ ο μικροτερος κλειστός γραμμικος υποχωρος του $L^2(\mu)$ που περιεχει την f και τις σταθερές συναρτησεις και για τον οποιο ισχύει $U_T\mathcal{H}_f \subset \mathcal{H}_f$, δηλ. $h\in\mathcal{H}_f \Rightarrow h\circ T\in\mathcal{H}_f$. Εστω

$$\mathcal{Z}_f := \Big\{ g\in L^2(\mu) :\ N^{-1}\sum_{n=0}^{N-1} |(f\circ T^n, g) - (f,1)(1,g)| \to 0 \Big\}.$$

Τότε ο Z_f περιέχει τις σταθερές συναρτήσεις, περιέχει την f και αν $g\in Z_f$ τότε $g\circ T\in Z_f$ αφού

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n, g\circ T)-(f,1)(1,g\circ T)\right| =$$

$$=\frac{1}{N}\sum_{n=0}^{N-2}\left|(f\circ T^n, g)-(f,1)(1,g)\right| + \frac{1}{N}\left|(f,g\circ T)-(f,1)(1,g)\right|\to 0.$$

Τέλος ο Z_f είναι κλειστός και γραμμικός αφού:

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n, c_1g_1+c_2g_2)-(f,1)(1,c_1g_1+c_2g_2)\right| =$$

$$=\frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_1)\bar{c}_1+(f\circ T^n,g_2)\bar{c}_2-(f,1)(1,g_1)\bar{c}_1-(f,1)(1,g_2)\bar{c}_2\right| \le$$

$$\le \frac{|c_1|}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_1)-(f,1)(1,g_1)\right| + \frac{|c_2|}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_2)-(f,1)(1,g_2)\right|\to 0$$

που αποδεικνύει γραμμικότητα και αν $\|g_m-g\|_2\to 0$ με $g_m\in Z_f$

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g)-(f,1)(1,g)\right| \le \frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g)-(f\circ T^n,g_m)\right| +$$

$$+\frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_m)-(f,1)(1,g_m)\right| + \frac{1}{N}\sum_{n=0}^{N-1}\left|(f,1)(1,g_m)-(f,1)(1,g)\right| \le$$

$$\le \frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_m)-(f,1)(1,g_m)\right| + \|f\circ T^n\|_2\,\|g-g_m\|_2 + |(f,1)|\,\|g-g_m\|_1$$

$$\le \frac{1}{N}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_m)-(f,1)(1,g_m)\right| + 2\|f\|_2\,\|g-g_m\|_2 < \varepsilon$$

αν διαλέξουμε πρώτα ένα μεγάλο m ώστε $\|g-g_m\|_2<\varepsilon/4\|f\|_2$ και μετά ένα μεγάλο N_0 ώστε για $N\ge N_0$, $N^{-1}\sum_{n=0}^{N-1}\left|(f\circ T^n,g_m)-(f,1)(1,g_m)\right|<\varepsilon/2$.
Αυτό αποδεικνύει ότι ο Z_f είναι κλειστός.

Από τον ορισμό του $\mathcal{H}_f$ έπεται ότι $\mathcal{H}_f\subset Z_f$. Έστω τώρα $g\in\mathcal{H}_f^{\perp}$. Τότε $(1,g)=0$ αφού $1\in\mathcal{H}_f$ κ' $(f\circ T^n,g)=0$ αφού $f\in\mathcal{H}_f$ και $U\mathcal{H}_f\subset\mathcal{H}_f$. Έπεται ότι $g\in Z_f$.

Άρα έχουμε ότι $\mathcal{H}_f\subset Z_f$ και $\mathcal{H}_f^{\perp}\subset Z_f$ και άρα ο Z_f δεν μπορεί παρά να είναι όλος ο $L^2(\mu)$. ■

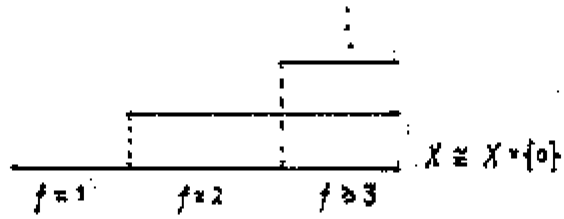
Κλείνουμε με ένα παράδειγμα ένος συστήματος που είναι ασθενώς mixing αλλά όχι ισχυρά mixing. Για το σκοπό αυτό θα χρειαστεί να μελετήσουμε πρώτα ένα άλλο παράδειγμα που είναι σημαντικό.

8. ΔΥΟ ΠΑΡΑΔΕΙΓΜΑΤΑ

8.1 Παράδειγμα: (Suspension)

Εστω $(X,\mathcal{A},\mu)$ χωρος πιθανοτητας κ' $T: X \to X$ ενας ενδομορφισμός. Εστω επίσης $f: X \to \mathbb{N}$ με $f \in L^1(\mu)$ και ορίζουμε $\mu(f) = \int_X f\, d\mu$.

Ορίζουμε εναν καινουργιο χωρο ως εξης:



$$X_f := \{ (x,n) : x \in X ,\ 0 \leq n < f(x) \}$$

$$\mathcal{A}_f := \{ A \subset X_f : \{ x \in X : (x,n) \in A \} \in \mathcal{A} ,\ \forall n \in \mathbb{N}_0 \}$$

$$\mu_f(A) := \frac{1}{\mu(f)} \sum_{n=0}^{\infty} \mu \{ x \in X : (x,n) \in A \} \qquad A \in \mathcal{A}_f .$$

Ειναι ευκολο να δει κανεις οτι ο $(X_f, \mathcal{A}_f, \mu_f)$ ειναι ενας χωρος πιθανοτητας. Ειδικα

$$\mu_f(X_f) = \frac{1}{\mu(f)} \sum_{n=0}^{\infty} \mu \{ x \in X : (x,n) \in X_f \} =$$

$$= \frac{1}{\mu(f)} \sum_{n=0}^{\infty} \mu \{ x \in X : f(x) > n \} = \frac{1}{\mu(f)} \cdot \mu(f) = 1 .$$

Ορίζουμε μια απεικονιση $T_f : X_f \to X_f$ απο την σχεση

$$X_f \ni (x,n) \longmapsto T_f(x,n) = \begin{cases} (x, n+1) & \text{αν } f(x) > n+1 \\ (Tx, 0) & \text{αν } f(x) = n+1 . \end{cases}$$

Αν για $A \in \mathcal{A}_f$ θεσουμε $A_n = \{ x \in X : (x,n) \in A \}$ τοτε

$$T_f^{-1}(A) = \bigcup_{n=1}^{\infty} (A_n \times \{n-1\}) \cup \bigcup_{n=0}^{\infty} (T^{-1}A_0 \cap \{ f(x) = n+1 \}) \times \{n\} =$$

$$= \bigcup_{n=0}^{\infty} \left[A_{n+1} \cup \left(T^{-1}A_0 \cap \{ f(x) = n+1 \} \right) \right] \times \{n\}$$

απο οπου αμεσως φαίνεται οτι η T_f ειναι μετρήσιμη $\mathcal{A}_f$ και οτι διατηρει το μ_f:

$$\mu_f(T_f^{-1}A) = \frac{1}{\mu(f)} \left(\sum_{n=1}^{\infty} \mu(A_n) + \sum_{n=0}^{\infty} \mu(T^{-1}A_0 \cap \{ f(x) = n+1 \}) \right) =$$

$$= \frac{1}{\mu(f)} \left(\sum_{n=1}^{\infty} \mu(A_n) + \mu(T^{-1}A_0) \right) = \frac{1}{\mu(f)} \left(\sum_{n=1}^{\infty} \mu(A_n) + \mu(A_0) \right) =$$

$$= \mu_f(A) .$$

Δηλαδη ο T_f ειναι ενας ενδομορφισμός του χωρου $(X_f, \mathcal{A}_f, \mu_f)$.

Αν ο αρχικος T ειναι ενας αυτομορφισμος, τοτε προφανως ο T_f ειναι 1-1 και επι και επειδη

$$T_f(A) = \bigcup_{n\geq 1}^{\infty} \left(A_{n-1} \cap \{x : f(x) > n\}\right) \times \{n\} \;\cup\; \left(\bigcup_{n=0}^{\infty} T(A_n \cap \{f(x) = n+1\})\right) \times \{0\}$$

ο T_f^{-1} ειναι μετρησιμος. Δηλαδη ο T_f ειναι και αυτος αυτομορφισμός.

Προταση: Το συστημα $(X_f, \mathcal{A}_f, \mu_f, T_f)$ ειναι εργοδικο αν και μονο αν το συστημα $(X, \mathcal{A}, \mu, T)$ ειναι εργοδικό.

Παρατηρηση: Τα αντιστοιχα αποτελεσματα για mixing, ειτε ασθενες ειτε ισχυρό, δεν ισχύουν. Για ασθενές mixing το Παραδειγμα 8.2 παρακατω ειναι ενα αντιπαραδειγμα. Για ισχυρο mixing, εστω $(X, \mathcal{A}, \mu, T)$ ενα οποιοδηποτε ισχυρα mixing συστημα (π.χ. $T(x) = 2x \pmod 1$ στον κυκλο) και $f(x) = 2$ $\forall x \in X$. Αν $X_0 = X \times \{0\}$ και $X_1 = X \times \{1\}$ η ακολουθια $\{\mu_f(T_f^{-n} X_1 \cap X_0)\}_{n\in\mathbb{N}}$ δεν συγκλινει αφου ειναι η ακολουθια $\{\frac{1}{2}, 0, \frac{1}{2}, 0, \dots\}$.

Αποδειξη Προτασης: Εστω πρωτα οτι το $(X_f, \mathcal{A}_f, T_f, \mu_f)$ ειναι εργοδικό. Εστω $A \subset X$ με $T^{-1}A = A$. Εστω οτι $\mu(A) > 0$. Οριζουμε $A_0 = A$ και $A_n = A \cap \{f(x) > n\}$ και θετουμε $\tilde{A} := \bigcup_{n\geq 0}^{\infty} A_n \times \{n\}$. Τοτε προφανως $T_f^{-1}\tilde{A} = \tilde{A}$ και αφου $\mu(A) > 0$, πρεπει $\mu_f(\tilde{A}) > 0$. Επειδη ο T_f ειναι εργοδικος επεται οτι $\mu_f(\tilde{A}) = 1$. Επεται οτι $\mu(A) = 1$.

Αντιστροφα εστω $A \subset X_f$ με $T_f^{-1} A = A$. Εστω οτι $\mu_f(A) > 0$. Οριζουμε $A_n = \{x \in X : (x,n) \in A\}$, $n \in \mathbb{N}_0$ και $B = \bigcup_{n\geq 0}^{\infty} A_n \subset X$. Απο το αναλλοιωτο του A πρεπει $Tx \in A_0$ για καθε $x \in B$. Αρα

$$\frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_{A_0} \circ T^n(x) \to 1 \qquad \forall x \in B$$

αφου το αριστερο μελος ειναι $\frac{N-1}{N}$ ή 1 για καθε $x \in B$. Απο το εργοδικο θεωρημα ομως (υποθετουμε οτι ο T ειναι εργοδικος)

$$\frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_{A_0} \circ T^n(x) \to \mu(A_0) \quad \mu\text{-σ.π.}$$

Αφου $\mu_f(A) > 0$ πρεπει $\mu(B) > 0$ και αρα $\mu(A_0) = 1 = \mu\{x : f(x) \geq 1\}$.

Εστω τωρα $n \in \mathbb{N}$. Απο το εργοδικο θεωρημα και παλι

(8.1) $$\frac{1}{N} \sum_{k=0}^{N-1} \mathbb{1}_{A_n}(T^k x) \to \mu(A_n) \quad \mu\text{-σ.π.}$$

Ομως επειδη $T_f^m A = A$ για καθε m επεται οτι για καθε $x \in B$

$$(8.2) \qquad \frac{1}{N}\sum_{k=0}^{N-1} \mathbb{1}_{A_n}\circ T^k(x) = \frac{1}{N}\sum_{k=0}^{N-1}\mathbb{1}_{\{f \geq n+1\}}\circ T^k(x) \qquad (N\in\mathbb{N}).$$

Αλλα το δεξι μελος τεινει, απο το εργοδικο θεωρημα, στο $\mu\{x\in X : f(x)>n\}$, μ-σ.π. και αφου $\mu(B)>0$ πρεπει απο τις (8.1) και (8.2)

$$\mu(A_n) = \mu\{x\in X : f(x)>n\} \qquad (n\in\mathbb{N}).$$

Επεται τωρα οτι

$$\mu_f(A) = \frac{1}{\mu(f)}\sum_{n=0}^{\infty}\mu(A_n) = \frac{1}{\mu(f)}\sum_{n=0}^{\infty}\mu\{x\in X : f(x)>n\}$$

$$= \frac{1}{\mu(f)}\,\mu(f) = 1.$$

Δηλαδη αποδειξαμε οτι αν $T_f^m A = A$ και $\mu_f(A)>0$ τοτε $\mu_f(A)=1$. Αρα ο T_f ειναι εργοδικος. ∎

8.2 Παραδειγμα: (Kakutani) (ο W. Parry το αναφερει σαν παραδειγμα Kakutani - Von Neumann.)

Εστω $X = [0,1] - \{k2^{-n} : n\in\mathbb{N}_0,\ k=0,\dots,2^n\}$, δηλ. ο X ειναι το $[0,1]$ εκτος των δυαδικων ρητων. $\mathcal{A} = \mathcal{B}(X)$ και μ = Lebesgue στον X.

Θετουμε $I_n = (1-2^{-n},\ 1-2^{-(n+1)})$, $n\in\mathbb{N}_0$ και οριζουμε

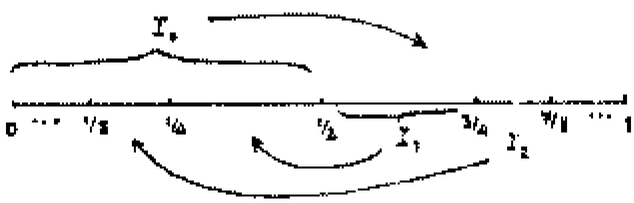
$$Tx = x - (1 - 2^{-n} - 2^{-(n+1)}) \qquad \text{οταν } x\in I_n$$

Δηλαδη:

$Tx = x + \frac{1}{2}$ $\qquad 0 < x < \frac{1}{2}$

$Tx = x - \frac{1}{4}$ $\qquad \frac{1}{2} < x < \frac{3}{4}$

$Tx = x - \frac{5}{8}$ $\qquad \frac{3}{4} < x < \frac{7}{8}$ κ.ο.κ.



Μια αλλη εκφραση του T: αν αντιστοιχισουμε στο $x \in X$ το δυαδικο αναπτυγμα του $(x_1, x_2, \dots)$, οπου $x = \sum_{i=1}^{\infty} x_i 2^{-i}$, τοτε

$$x = (x_1, x_2, \dots) \xmapsto{T} (1-x_1, 1-x_2, \dots, 1-x_i, x_{i+1}, x_{i+2}, \dots)$$

οπου $i = \min\{n\in\mathbb{N} : x_n = 0\}$. Δηλαδη $x = (1,\dots,1,0,x_{i+1},\dots) \longmapsto (0,\dots,0,1,x_{i+1},\dots)$.

Ο T ειναι προφανως ενας αυτομορφισμος του X.

Πρόταση: Το φάσμα του T (το σύνολο των ιδιοτιμών του) είναι το

$$\left\{ \exp\left(2\pi i \frac{2k-1}{2^n}\right) : n\in\mathbb{N},\ k=1,\dots,2^{n-1} \right\} \cup \{1\}.$$

Στην ιδιοτιμή 1 αντιστοιχεί η ιδιοσυνάρτηση $f_0(x)\equiv 1$ ενώ στην ιδιοτιμή $\exp(2\pi i\lambda)$, όπου $\lambda=(2k-1)2^{-n}$, αντιστοιχεί η ιδιοσυνάρτηση

$$f_\lambda(x) = \exp\left(2\pi i \frac{2k-1}{2^n}\left(x_1+2x_2+\dots+2^{n-1}x_n\right)\right) \qquad \left(x=\sum_{i=1}^{\infty} x_i 2^{-i}\right).$$

Το σύνολο των ιδιοσυναρτήσεων $\{f_\lambda\}$ του T, αποτελεί ορθοκανονική βάση του $L^2(\mu)$.

Ο T έχει δηλαδή διακριτό φάσμα: υπάρχει μια ορθοκανονική βάση του $L^2(\mu)$ που αποτελείται από ιδιοσυναρτήσεις του T.

Η απόδειξη της Πρότασης θα δοθεί στο τέλος.

Από την Πρόταση παίρνουμε αμέσως τα εξής:

(1) Ο T δεν είναι ασθενώς mixing. (Έχει "πολλές" ιδιοτιμές.)

(2) Ο T είναι εργοδικός.

(Πράγματι έστω $f\in L^2(\mu)$ με $f=f\circ T$. Αν $f=\sum a_\lambda f_\lambda$ η παράσταση της f στην βάση $\{f_\lambda\}$ – που είναι μοναδική αφού η βάση είναι ορθοκανονική – τότε $f\circ T=$ $=\sum a_\lambda f_\lambda\circ T=\sum a_\lambda e^{2\pi i\lambda} f_\lambda$ και άρα πρέπει $a_\lambda=e^{2\pi i\lambda}a_\lambda$ για όλα τα λ. Αφού $e^{2\pi i\lambda}\neq 1$ για κάθε $\lambda\neq 0$ πρέπει $a_\lambda=0$ $\forall \lambda\neq 0$ δηλ. $f=a_0=$ σταθερή μ-σ.π.)

(3) $$\| f\circ T^{2^m}-f\|_2 \longrightarrow 0 \qquad \forall f\in L^2(\mu).$$

(Πράγματι, αν $\lambda=(2k-1)2^{-n}$, τότε για $m\geq n$, $f_\lambda\circ T^{2^m}(x)=e^{2\pi i\lambda 2^m} f_\lambda(x)=$ $=\exp\left(2\pi i(2k-1)2^{m-n}\right) f_\lambda(x)=f_\lambda(x)$ για όλα τα x. Τώρα για τυχούσα $f\in L^2(\mu)$, με $f=\sum_\lambda a_\lambda f_\lambda$, διαλέγουμε, δοθέντος $\varepsilon>0$, ένα πεπερασμένο σύνολο από λ, έστω $\lambda_1,\dots,\lambda_N$, ώστε $\| f-\sum_{i=1}^N a_{\lambda_i} f_{\lambda_i}\|_2<\varepsilon$. Αν $\lambda_i=(2k_i-1)2^{-n_i}$ και θέτουμε $M=\max\{n_1,\dots,n_N\}$, τότε για κάθε $m\geq M$,

$$\| f\circ T^{2^m}-f\|_2 \leq \| f-\sum_{i=1}^{n} a_{\lambda_i} f_{\lambda_i}\|_2 + \| f\circ T^{2^m}-\sum_{i=1}^{n} a_{\lambda_i} f_{\lambda_i}\circ T^{2^m}\|_2 +$$

$$+\sum_{i=1}^{n} |a_{\lambda_i}|\, \| f_{\lambda_i}-f_{\lambda_i}\circ T^{2^m}\|_2 =$$

$$= 2\| f-\sum_{i=1}^{n} a_{\lambda_i} f_{\lambda_i}\|_2 + 0 < 2\varepsilon.)$$

Τωρα ορίζουμε μια συναρτηση $f: X \to \{1,2\}$:

$$f(x)=\begin{cases} 1 & \text{αν} \quad x\in I_{2k+1} \\ 2 & \text{αν} \quad x\in I_{2k} \end{cases}$$

για $k\in\mathbb{N}_0$. Εστω $(X_f, \mathcal{A}_f, \mu_f, T_f)$ ο αντιστοιχος χωρος suspension του Παραδειγματος 8.1.

Απο το Παραδειγμα 8.1, αφου το $(X,\mathcal{A},\mu,T)$ ειναι εργοδικό και το συστημα $(X_f,\mathcal{A}_f,\mu_f,T_f)$ ειναι εργοδικό. Θα αποδειξουμε οτι το συστημα $(X_f, \mathcal{A}_f, \mu_f, T_f)$ ειναι και ασθενώς mixing (ενω το $(X,\mathcal{A},\mu,T)$ δεν ειναι) αλλα δεν ειναι ισχυρά mixing.

Για να αποδειξουμε οτι ο T_f ειναι ασθενώς mixing, αρκει, απο το Θεωρημα 7.4, να αποδειξουμε οτι η μονη ιδιοτιμή του T_f ειναι η 1, αφου ο T_f ειναι εργοδικός. (Υπενθυμιζουμε οτι ο T ειναι αυτομορφισμός και αρα και ο T_f ειναι αυτομορφισμός.)

Εστω λοιπον $\tilde{h}\in L^2(\mu_f)$, με $\tilde{h}\neq 0$ και $\lambda\in[0,1)$ ωστε $\tilde{h}\circ T_f = e^{2\pi i\lambda}\tilde{h}$. Επειδή $\tilde{h}\circ T_f(\tilde{x}) \cdot |\tilde{h}\circ T_f(\tilde{x})|^{-1} = e^{2\pi i\lambda}\tilde{h}(\tilde{x}) \cdot |e^{2\pi i\lambda}|^{-1}|\tilde{h}(\tilde{x})|^{-1} = e^{2\pi i\lambda}\tilde{h}(\tilde{x})|\tilde{h}(\tilde{x})|^{-1}$, για καθε $\tilde{x}\in X_f$, μπορουμε να υποθεσουμε οτι $|\tilde{h}(\tilde{x})|=1 \quad \forall \tilde{x}\in X_f$.

Ορίζουμε $h: X\to\{z\in\mathbb{C} : |z|=1\}$ απο την σχέση $h(x)=\tilde{h}(x,0)$ για $x\in X$. Τοτε, εξ ορισμου της T_f

$$h\circ T(x) = \tilde{h}(Tx,0) = \tilde{h}\circ T_f^{f(x)}(x,0) = e^{2\pi i\lambda f(x)}\tilde{h}(x,0) = e^{2\pi i\lambda f(x)} h(x)$$

και ετσι, αν ως συνήθως θετουμε $S_n f(x) = f(x) + f\circ T(x) + \dots + f\circ T^{n-1}(x)$, $x\in X$,

$$h\circ T^n(x) = \exp(2\pi i\lambda S_n f(x))\, h(x) \qquad (n\in\mathbb{N},\ x\in X).$$

Λημμα: Για $n\in\mathbb{N}$, το $S_{2^n} f(x)$ παιρνει μονο δυο τιμες, s_n και s_n+1 και

$$\mu\{x\in X : S_{2^n} f(x) = s_n\} = \frac{1}{3}, \quad \mu\{x\in X : S_{2^n} f(x) = s_n+1\} = \frac{2}{3}.$$

Η αποδειξη του λημματος θα δοθει στο τέλος.

Απο το λημμα λοιπον παιρνουμε οτι (η πρωτη ισοτητα λογω του (3), σελ. 56):

$$0 = \lim_{n\to\infty} \| h\circ T^{2^n} - h\|_2^2 = \int_X |\exp(2\pi i\lambda S_{2^n} f(x)) - 1|^2 |h(x)|^2 \mu(dx) =$$

$$= \int_X |\exp(2\pi i\lambda S_{2^n} f(x)) - 1|^2 \mu(dx) =$$

$$= \frac{1}{3}|e^{2\pi i\lambda s_n} - 1|^2 + \frac{2}{3}|e^{2\pi i\lambda(s_n+1)} - 1|^2.$$

Αρα πρεπει $\exp(2\pi i\lambda(s_n+1)) \to 1$ και $\exp(2\pi i\lambda s_n) \to 1$, οποτε

$$\exp(2\pi i\lambda) = \frac{\exp(2\pi i\lambda(s_n+1))}{\exp(2\pi i\lambda s_n)} \longrightarrow 1.$$

Επομενως $e^{2\pi i\lambda}=1$ και η μονη ιδιοτιμη του T_f ειναι η 1. Οπως ηδη ειπαμε, αυτο συνεπαγεται οτι ο T_f ειναι ασθενως mixing.

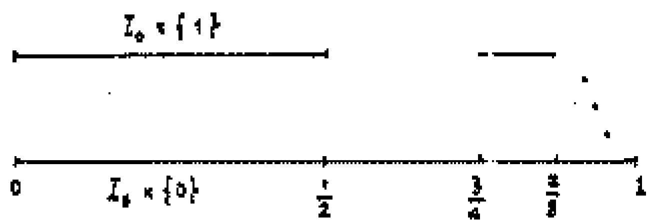
Αποδεικνυουμε τωρα οτι ο T_f δεν ειναι ισχυρα mixing. Κατ' αρχην εχουμε οτι $T_f^{f(x)}(x,0) = (Tx,0)$ για $x\in X$ οποτε

$$T_f^{S_{2^n}f(x)}(x,0) = (T^{2^n}x,0).$$

Επειδη το $S_{2^n}f(x)$ παιρνει μονο δυο τιμες

$$(T^{2^n}x,0) = \begin{cases} T_f^{s_n}(x,0), & \text{αν } S_{2^n}f(x)=s_n \\ T_f^{s_n+1}(x,0), & \text{αν } S_{2^n}f(x)=s_n+1 . \end{cases}$$

Τωρα παρατηρουμε οτι $T^2 I_0 = I_0$, οπου θυμιζουμε, $I_0=(0,\frac{1}{2})$. Αρα για καθε n πρεπει $T^{2^n}I_0 = I_0$. Δηλαδη αν $x\in I_0$ τοτε υπαρχει $y\in I_0$ με $T^{2^n}y=x$ και ετσι



$$(x,0)=(T^{2^n}y,0) = \left\{\begin{matrix} T_f^{s_n}(y,0)\\ T_f^{s_n+1}(y,0)\end{matrix}\right\} \in T_f^{s_n}(I_0\times\{0\}) \cup T_f^{s_n+1}(I_0\times\{0\}).$$

Δηλαδη

$$I_0\times\{0\} \subset T_f^{s_n}(I_0\times\{0\}) \cup T_f^{s_n+1}(I_0\times\{0\}) =$$

$$= T_f^{s_n}(I_0\times\{0\}) \cup T_f^{s_n}(I_0\times\{1\})$$

και αρα $\quad T_f^{-s_n}(I_0\times\{0\}) \subset (I_0\times\{0\})\cup(I_0\times\{1\}).$

Εστω τωρα A ενα οποιοδηποτε συνολο που δεν τεμνει τα $I_0\times\{0\}$ και $I_0\times\{1\}$ και τετοιο ωστε $\mu_f(A)>0$. (Για παραδειγμα $A=\{(x,n)\in X_f : x>\frac{1}{2}\}$.) Τοτε

$$\mu_f(T_f^{-s_n}(I_0\times\{0\})\cap A) = \mu_f(\emptyset)=0, \qquad \forall n\in\mathbb{N}$$

και αρα η ακολουθια $\mu_f(T_f^{-n}(I_0\times\{0\})\cap A)$ <u>δεν</u> μπορει να συγκλινει στο $\mu_f(I_0)\cdot\mu_f(A) = \frac{3}{10}\mu_f(A)>0$, αφου εχει απειρους ορους που ειναι 0.

Αποδείξαμε δηλαδή ότι το σύστημα $(X_f, \mathcal{A}_f, \mu_f, T_f)$ δεν είναι ισχυρά mixing.

ΑΠΟΔΕΙΞΗ ΤΗΣ ΠΡΟΤΑΣΗΣ ΣΤΗ ΣΕΛΙΔΑ 56: Κατ' αρχήν το 1 είναι ιδιοτιμή του T με αντίστοιχη ιδιοσυνάρτηση $f_0(x) = 1 \quad \forall x \in X$.

Έστω τώρα $\lambda = (2k-1)2^{-n}$ για $n \in \mathbb{N}$ και $k \in \{1, \dots, 2^{n-1}\}$. Θα αποδείξουμε ότι ο αριθμός $\exp(2\pi i\lambda)$ είναι ιδιοτιμή του T με αντίστοιχη ιδιοσυνάρτηση την

$$f_\lambda(x) = \exp\left(2\pi i \lambda(x_1 + 2x_2 + \dots + 2^{n-1}x_n)\right)$$

όπου $x = \sum_{i=1}^{\infty} x_i 2^{-i}$, δηλαδή τα x_i είναι τα (μονοσήμαντα ορισμένα) δυαδικά ψηφία του $x \in X$. Υπενθυμίζουμε ότι αν το $x \in X$ έχει δυαδικά ψηφία $x_1, x_2, \dots$ τότε το Tx έχει δυαδικά ψηφία $1-x_1, 1-x_2, \dots, 1-x_i, x_{i+1}, x_{i+2}, \dots$ όπου $i = \min\{m \in \mathbb{N} : x_m = 0\}$. Αν λοιπόν $i > n$ τότε τα n πρώτα δυαδικά ψηφία του x είναι 1 και έτσι

$$f_\lambda(x) = \exp\left(2\pi i \frac{2k-1}{2^n}(1+2+\dots+2^{n-1})\right) = \exp\left(2\pi i \frac{2k-1}{2^n}(2^n-1)\right) =$$

$$= \exp\left(-2\pi i \frac{2k-1}{2^n}\right) = \exp(-2\pi i\lambda)$$

ενώ τα n πρώτα δυαδικά ψηφία του Tx είναι 0 οπότε $f_\lambda \circ T(x) = 1$. Άρα λοιπόν $f_\lambda \circ T(x) = \exp(2\pi i\lambda) f_\lambda(x)$ σ' αυτή την περίπτωση. Αν $i \le n$ τότε τα πρώτα $i-1$ ψηφία του x είναι 1 κ' το $x_i = 0$ ενώ για το Tx τα πρώτα $i-1$ ψηφία είναι 0 κ' του $x_i = 1$. Έτσι

$$f_\lambda \circ T(x) = \exp\left(2\pi i \frac{2k-1}{2^n}(0 + 0.2 + \dots + 0\,2^{i-2} + 1.2^{i-1} + x_{i+1}2^i + \dots + x_n 2^{n-1})\right) =$$

$$= \exp\left(2\pi i \frac{2k-1}{2^n}\right)\exp\left(-2\pi i \frac{2k-1}{2^n}\right)\exp\left(2\pi i \frac{2k-1}{2^n}(2^{i-1} + 2^i x_{i+1} + \dots + 2^{n-1}x_n\right)$$

$$= \exp(2\pi i\lambda)\exp\left(2\pi i \frac{2k-1}{2^n}(2^{i-1}-1) + 2\pi i \frac{2k-1}{2^n}(2^i x_{i+1} + \dots + 2^{n-1}x_n)\right) =$$

$$= \exp(2\pi i\lambda)\exp\left(2\pi i \frac{2k-1}{2^n}(1+2+\dots+2^{i-2}) + 2\pi i \frac{2k-1}{2^n}(2^i x_{i+1} + \dots + 2^{n-1}x_n)\right) =$$

$$= \exp(2\pi i\lambda)\exp\left(2\pi i \frac{2k-1}{2^n}(1 + 2.1 + \dots + 2^{i-2}.1 + 2^{i-1}.0 + 2^i x_{i+1} + \dots + 2^{n-1}x_n\right)$$

$$= \exp(2\pi i\lambda)\; f_\lambda(x).$$

Άρα πάντα ισχύει $f_\lambda \circ T = e^{2\pi i\lambda} f_\lambda$ και έτσι το $e^{2\pi i\lambda}$ είναι ιδιοτιμή του T με αντίστοιχη ιδιοσυνάρτηση f_λ.

Έστω τώρα $\Lambda := \{0\} \cup \{ \frac{2k-1}{2^n} : n\in\mathbb{N},\ k=1,2,\dots,2^{n-1}\}$. Θα αποδείξουμε ότι το σύνολο $\{f_\lambda : \lambda\in\Lambda\}$ είναι ορθοκανονικό στον $L^2(\mu)$.

Αν $x\in X$ θα συμβολίζουμε τώρα με $d_n(x)$ τα ψηφία του δυαδικού αναπτύγματος του x, δηλαδή $x=\sum_{n=1}^{\infty} d_n(x)\cdot 2^{-n}$.

Προφανώς $|f_\lambda(x)|^2=1$ $\forall x\in X$ και $\forall \lambda\in\Lambda$ και άρα $\|f_\lambda\|_2=1$, $\forall\lambda\in\Lambda$.

Έστω τώρα $\lambda=\frac{2k-1}{2^n}$ και $\lambda'=\frac{2l-1}{2^m}$ με $\lambda\neq\lambda'$. Έστω κατ' αρχήν ότι $m\neq n$ και υποθέτουμε χωρίς βλάβη της γενικότητας ότι $m>n$. Τότε

$$\overline{f_\lambda(x)}\, f_{\lambda'}(x) = \exp\Big(-2\pi i\, \frac{2k-1}{2^n}\big(d_1(x)+\dots+2^{n-1}d_n(x)\big) + 2\pi i\, \frac{2l-1}{2^m}\big(d_1(x)+\dots+2^{m-1}d_m(x)\big)\Big)$$

$$= \exp\Big(2\pi i\, \frac{2l-1}{2}\, d_m(x)\Big) \exp\Big(2\pi i\Big[\sum_{j=1}^{n}\Big(\frac{2l-1}{2^{m-j+1}} - \frac{2k-1}{2^{n-j+1}}\Big)d_j(x) + \sum_{j=n+1}^{m-1}\frac{2l-1}{2^{m-j+1}}\, d_j(x)\Big]\Big)$$

$$= \exp\big(\pi i\,(2l-1)\, d_m(x)\big)\quad g(x)$$

όπου η $g(x)$ εξαρτάται μόνο από τα $d_1(x),\dots,d_{m-1}(x)$. Λόγω ανεξαρτησίας (πιθανοθεωρητικής) των $d_i(x)$,

$$\int_X \overline{f_\lambda(x)}\, f_{\lambda'}(x)\, d\mu(x) = \int_X \exp\big(\pi i(2l-1)\, d_m(x)\big)\cdot g(x)\, d\mu(x) =$$

$$= \int_X \exp\big(\pi i(2l-1)\, d_m(x)\big)\, d\mu(x)\cdot \int_X g(x)\, d\mu(x) =$$

$$= \frac{1}{2}\big[\exp(\pi i(2l-1)) + 1\big] \int_X g(x)\, d\mu(x) = 0$$

αφού $\exp(\pi i(2l-1)) = (e^{\pi i})^{2l-1} = (-1)^{2l-1} = -1$.

Αν $m=n$ τότε πρέπει $k\neq l$ και επίσης (λόγω αυτού) $n\geq 2$.

$$\overline{f_\lambda(x)}\, f_{\lambda'}(x) = \exp\Big(2\pi i\, \frac{2(l-k)}{2^n}\sum_{j=1}^{n} d_j(x)\, 2^{j-1}\Big) = \exp\Big(\pi i\, \frac{l-k}{2^{n-2}}\sum_{j=1}^{n} d_j(x)\, 2^{j-1}\Big).$$

Τώρα το $k-l$ γράφεται σαν $-(k-l)=2^{\nu}(2u-1)$ ($2^{\nu}\cdot$ περιττό) με $\nu<n-1$ αφού $1\leq k,l\leq 2^{n-1}$. Τότε ο αντίστοιχος παράγοντας γίνεται

$$\exp\Big(\pi i\, \frac{l-k}{2^n}\, d_{n-\nu+1}(x)\, 2^{n-\nu}\Big) = \exp\Big(\pi i\, \frac{2^{\nu}(2u-1)}{2^n}\, d_{n-\nu+1}(x)\, 2^{n-\nu}\Big) =$$

$$= \big(\exp(\pi i\, d_{n-\nu+1}(x))\big)^{2u-1}$$

Το ολοκλήρωμα αυτού του παράγοντα θα είναι $\frac{1}{2}\big(e^{\pi i(2u-1)}+1\big)=0$, πάλι, και λόγω ανεξαρτησίας των $d_i(x)$ (το ολοκλήρωμα του γινομένου = γινόμενο ολοκληρωμάτων) έπεται ότι $(f_{\lambda'}, f_\lambda) = \int_X \overline{f_\lambda}\, f_{\lambda'}\, d\mu = 0$.

Μεχρι τωρα λοιπον ξερουμε οτι το $\{f_\lambda : \lambda \in \Lambda\}$ ειναι ορθοκανονικο. Θα αποδειξουμε τωρα οτι ειναι κ' βαση του $L^2(\mu)$.

Οριζουμε $\mathcal{F}_n = \sigma(\,(j2^{-n}, (j+1)2^{-n}) : j=0,1,\ldots,2^n-1)$, η σ-αλγεβρα που παραγεται απο τα διαστηματα $(j2^{-n}, (j+1)2^{-n})$ για σταθερο n. (Απο τα διαστηματα αυτα αφαιρουμε κ' ολους τους δυαδικους ρητους αφου δεν ανηκουν στο X, αλλα για ευκολια συμβολισμου θα γραφουμε καταχρηστικα $(j2^{-n}, (j+1)2^{-n})$ και θα εννοουμε το διαστημα αυτο χωρις τους δυαδικους ρητους που περιεχει.) Μια συναρτηση f ειναι $\mathcal{F}_n$ μετρησιμη αν ειναι σταθερη πανω στα δυαδικα διαστηματα μηκους 2^{-n}. Αρα οι χαρακτηριστικες συναρτησεις $\mathbb{1}_{(j2^{-n},(j+1)2^{-n})}$ αποτελουν αλγεβρικη βαση του γραμμικου χωρου ολων των L^2 συναρτησεων f που ειναι μετρησιμες $\mathcal{F}_n$ (παραγουν το χωρο κ' ειναι προφανως γραμ. ανε.) Αρα η διασταση αυτου του χωρου ειναι 2^n.

Θετουμε $\Lambda_m = \{1\} \cup \{\lambda = \frac{2k+1}{2^n} \in \Lambda : n \le m\}$. Το Λ_n εχει 2^n στοιχεια και αρα το $\{f_\lambda : \lambda \in \Lambda_n\}$ ειναι αλγεβρικη βαση του γραμμικου χωρου των L^2 συναρτησεων που ειναι $\mathcal{F}_n$-μετρησιμες (αφου το $\{f_\lambda : \lambda \in \Lambda_n\}$ ειναι γραμμικα ανεξαρτητο συνολο λογω ορθογωνιοτητας). Αρα αν $f \in L^2(\mu)$ κ' f ειναι $\mathcal{F}_n$ μετρησιμη τοτε

$$f = \sum_{\lambda \in \Lambda_n} (f, f_\lambda)\, f_\lambda .$$

Εστω τωρα $f \in L^2(\mu)$. Η προβολη της f στο χωρο των L^2 συναρτησεων που ειναι μετρησιμες $\mathcal{F}_n$, ειναι η δεσμευμενη μεση τιμη $E(f\,|\,\mathcal{F}_n)$. Ετσι

$$E(f\,|\,\mathcal{F}_n) = \sum_{\lambda \in \Lambda_n} (\,E(f\,|\,\mathcal{F}_n)\,,\, f_\lambda)\; f_\lambda .$$

Ομως η f_λ ειναι $\mathcal{F}_n$ μετρησιμη για $\lambda \in \Lambda_n$ και ετσι

$$(\,E(f\,|\,\mathcal{F}_n), f_\lambda) = \int_X E(f\,|\,\mathcal{F}_n)\,\bar{f}_\lambda\, d\mu = \int_X E(f\cdot \bar{f}_\lambda\,|\,\mathcal{F}_n)\, d\mu = \int f \bar{f}_\lambda\, d\mu .$$

Δηλαδη

$$E(f\,|\,\mathcal{F}_n) = \sum_{\lambda \in \Lambda_n} (f, f_\lambda)\, f_\lambda \qquad \forall n \in \mathbb{N} \text{ κ' } \forall f \in L^2(\mu).$$

Ομως αν $f \in L^2(\mu)$ η $\{E(f|\mathcal{F}_n)\}_{n \in \mathbb{N}}$ ειναι μια L^2-martingale και αρα συγκλινει στον $L^2(\mu)$ στην $E(f\,|\,\sigma(\cup_{n=1}^{\infty} \mathcal{F}_n)) = E(f\,|\,\mathcal{A}) = f$ αφου η $\cup_{n=1}^{\infty} \mathcal{F}_n$ παραγει ολη τη σ-αλγεβρα των Borel συνολων του X. Αποδειξαμε δηλαδη οτι

$$f \overset{L^2}{=} \sum_{\lambda \in \Lambda} (f, f_\lambda)\, f_\lambda = \lim_{n\to\infty}\text{-}L^2 \sum_{\lambda \in \Lambda_n} (f, f_\lambda)\, f_\lambda .$$

Αποδειξαμε λοιπον οτι το $\{f_\lambda : \lambda \in \Lambda\}$ ειναι μια ορθοκανονικη βαση του $L^2(\mu)$.

Παρατήρηση: Αποδείξαμε μόνο ότι κάθε αριθμός $\exp(2\pi i\lambda)$ για $\lambda\in\Lambda$ είναι μια ιδιοτιμή του T και όχι, όπως ισχυριζόμαστε στην Πρόταση, ότι αυτές είναι όλες οι ιδιοτιμές του T. (Στο Παράδειγμα του Kakutani πουθενά χρειαζόμαστε μόνο ότι αποδείξαμε πιο πάνω.) Το ότι ο T δεν έχει άλλες ιδιοτιμές είναι συνέπεια του ότι το $\{f_\lambda : \lambda\in\Lambda\}$ είναι ορθοκανονική βάση του $L^2(\mu)$. Πράγματι, έστω c μια ιδιοτιμή του T. Δηλαδή υπάρχει $f\in L^2(\mu)$, όχι ταυτοτικά μηδέν, ώστε $f\circ T = cf$. Τότε $f\circ T(x) = \sum_{\lambda\in\Lambda}(f,f_\lambda)\, f_\lambda(T(x)) = \sum_{\lambda\in\Lambda}(f,f_\lambda)\, f_\lambda(x)\, e^{2\pi i\lambda}$ χρησιμοποιώντας την έκφραση της f στη βάση $\{f_\lambda\}$ στο σημείο $T(x)$. Όμως η $f\circ T$, σαν μέλος του $L^2(\mu)$, εκφράζεται και απ' ευθείας σ' αυτή τη βάση: $f\circ T = \sum_{\lambda\in\Lambda}(f\circ T, f_\lambda) f_\lambda = c\sum_{\lambda\in\Lambda}(f,f_\lambda) f_\lambda$. Από τις δύο αυτές εκφράσεις της $f\circ T$ παίρνουμε ότι $(f,f_\lambda)\,c = (f,f_\lambda)\exp(2\pi i\lambda)$ για κάθε $\lambda\in\Lambda$. Επειδή η f δεν είναι ταυτοτικά μηδέν πρέπει $(f,f_\lambda)\neq 0$ για τουλάχιστον ένα $\lambda\in\Lambda$, οπότε $c=\exp(2\pi i\lambda)$ για ένα τέτοιο λ. Δηλαδή ο T δεν έχει άλλες ιδιοτιμές εκτός των αριθμών $\exp(2\pi i\lambda)$, $\lambda\in\Lambda$.

ΑΠΟΔΕΙΞΗ ΤΟΥ ΛΗΜΜΑΤΟΣ ΣΤΗ ΣΕΛΙΔΑ 57: Έστω I το διάστημα $(0,2^{-n})\cap X$ του X και θεωρούμε τις εικόνες $I=T^0I, T^1I, \dots, T^{2^n-1}I$. Το I περνάει διαδοχικά από όλα τα διαστήματα $(k2^{-n},(k+1)2^{-n})\cap X$, $k=0,1,\dots,2^n-1$ και μετά από 2^n βήματα επανέρχεται στην αρχική του θέση, δηλ. $T^{2^n}I=I$. (Φυσικά το I περνάει από διαστήματα $(k2^{-n},(k+1)2^{-n})\cap X$ όχι με τη σειρά που επάγει στα διαστήματα αυτά η διάταξη των πραγματικών αριθμών.) Έπεται αμέσως ότι το ίδιο συμβαίνει αν αντί του I θεωρήσουμε οποιοδήποτε διάστημα $(k2^{-n},(k+1)2^{-n})\cap X$ στη θέση του.

Έστω τώρα $J_n=(1-2^{-n},1)\cap X$. Τότε οι εικόνες $TJ_n,\dots,T^{2^n-1}J_n$ περιέχονται εξ ολοκλήρου η κάθε μια σε ένα από τα διαστήματα $(0,\frac12),(\frac12,\frac34),\dots,(1-2^{-n+1},1-2^{-n})$ στα οποία η f είναι σταθερή. Άρα η f είναι σταθερή σε κάθε ένα από τα $TJ_n,\dots$ $\dots, T^{2^n-1}J_n$ οπότε και η $f\circ T(x)+\dots+f\circ T^{2^n-1}(x)$ είναι σταθερή για $x\in J_n$. Τώρα η f παίρνει ούτως ή άλλως δύο μόνο τιμές: $f(x)=1$ και $f(x)=2$. Μάλιστα για $x\in J_n$, $f(x)=2$ για x σ' ένα σύνολο μέτρου $2^{-n}\frac23$ και $f(x)=1$ για x σε ένα υποσύνολο του J_n μέτρου $2^{-n}\frac13$. Άρα η $S_{2^n}f(x)$ παίρνει δύο τιμές, έστω s_n και s_n+1, για $x\in J_n$ κ' $\mu(S_{2^n}f(x)=s_n,\ x\in J_n)=2^{-n}\frac13$ και $\mu(S_{2^n}f(x)=s_n+1,\ x\in J_n)=2^{-n}\frac23$. Από την παραπάνω παράγραφο όμως αν στη θέση του J_n θεωρήσουμε οποιοδήποτε διάστημα $I=(k2^{-n},(k+1)2^{-n})$ τότε οι εικόνες $I, TI,\dots,T^{2^n-1}I$ διαγράφουν τον ίδιο "κύκλο" με τις εικόνες του J_n και άρα η συνάρτηση $S_{2^n}f(x)$ παίρνει τις ίδιες δύο τιμές s_n και s_n+1 και $\mu(S_{2^n}f(x)=s_n,\ x\in I)=2^{-n}\frac13$ και $\mu(S_{2^n}f(x)=s_n+1,\ x\in I)=2^{-n}\frac23$. Άρα λοιπόν η $S_{2^n}f(x)$ παίρνει μόνο τις τιμές s_n και s_n+1 για x σ' όλο το X και $\mu(S_{2^n}f(x)=s_n)=1/3$, $\mu(S_{2^n}f(x)=s_n+1)=2/3$.

II. ΔΥΝΑΜΙΚΑ ΣΥΣΤΗΜΑΤΑ ΣΕ ΣΥΜΠΑΓΕΙΣ ΜΕΤΡΙΚΟΥΣ ΧΩΡΟΥΣ

1. Αμετάβλητα μέτρα σε συμπαγείς μετρικούς χώρους.

Εστω X ένας συμπαγής μετρικοποιήσιμος χώρος. Κάθε μέτρο Borel μ στον X είναι κανονικό, δηλαδή για κάθε σύνολο Borel $B \subset X$ και $\varepsilon > 0$ υπάρχουν ένα ανοιχτό σύνολο $V \subset X$ και ένα κλειστό σύνολο $C \subset X$ ώστε $C \subset B \subset V$ και $\mu(V \setminus C) < \varepsilon$. Συνεπώς

$$\mu(B) = \sup\{\mu(C) : C \subset B \text{ και } C \text{ κλειστό στον } X\} = \inf\{\mu(V) : B \subset V \text{ και } V \text{ ανοιχτό στον } X\}.$$

Απ' το Θεώρημα της αναπαράστασης του Riesz προκύπτει ότι το σύνολο $\mathcal{M}(X)$ των πιθανοτήτων Borel στον X βρίσκεται σε "1-1" και "επί" αντιστοιχία με το σύνολο όλων των θετικών γραμμικών μορφών $J : C(X) \to \mathbb{R}$ με $J(1) = 1$. Η αντιστοιχία αυτή ορίζεται απ' τον τύπο $J(f) = \int_X f d\mu$, όπου $f \in C(X)$, $\mu \in \mathcal{M}(X)$. Γι αυτό στην συνέχεια θα χρησιμοποιούμε τον συμβολισμό $\mu(f) = \int_X f d\mu$. Το $\mathcal{M}(X)$ εφοδιασμένο με την ασθενή τοπολογία γίνεται συμπαγής και μετρικοποιήσιμος χώρος, αφού ο $C(X)$ είναι διαχωρίσιμος.

1.1. Λήμμα Εστω $T : X \to X$ μία συνεχής και "επί" απεικόνιση. Μία πιθανότητα Borel μ είναι αμετάβλητη από την T τότε και μόνον τότε όταν

$$\mu(f \circ T) = \mu(f) \quad \text{για κάθε } f \in C(X)$$

Απόδειξη Το ευθύ είναι άμεσο απ' τον ορισμό του ολοκληρώματος. Για το αντίστροφο, λόγω της κανονικότητας του μέτρου και επειδή η T είναι επί, αρκεί να δείξουμε ότι $\mu(T^{-1}(A)) = \mu(A)$ για κάθε κλειστό σύνολο $A \subset X$. Αν το A είναι κλειστό, υπάρχει μία φθίνουσα ακολουθία συνεχών συναρτήσεων $f_n : X \to [0,1]$ με $f_n^{-1}(0) = A$, $n \in \mathbb{N}$, που συγκλίνει κατά σημείο στην χ_A. Αρα

$$\mu(T^{-1}(A)) = \int_X \chi_{T^{-1}(A)} d\mu = \int_X (\chi_A \circ T) d\mu = \lim_{n \to +\infty} \int_X (f_n \circ T) d\mu = \lim_{n \to +\infty} \mu(f_n \circ T)$$

$$= \lim_{n \to +\infty} \mu(f_n) = \int_X \chi_A d\mu = \mu(A).$$

Κάθε συνεχής και "επί" απεικόνιση $T : X \to X$ επάγει μία συνεχή απεικόνιση $T^* : \mathcal{M}(X) \to \mathcal{M}(X)$ με τύπο $T^*(\mu)(f) = \mu(f \circ T)$, $\mu \in \mathcal{M}(X)$, $f \in C(X)$.

Το μ είναι αμετάβλητο από την T τότε και μόνο τότε όταν $T^*(\mu)=\mu$.

<u>1.2. Θεώρημα</u> Για κάθε συνεχή "επί" απεικόνιση $T:X\to X$ υπάρχει τουλάχιστον ένα αμετάβλητο μέτρο πιθανότητας Borel στον X.

<u>Απόδειξη</u> Έστω $\mu_0\in\mathcal{M}(X)$. Για κάθε $n\in\mathbb{N}$ θέτουμε $\mu_n=\frac{1}{n}\sum_{k=0}^{n-1}T^{*k}(\mu_0)$
Για κάθε $f\in C(X)$ έχουμε

$$\mu_n(f)=\frac{1}{n}\sum_{k=0}^{n-1}\mu_0(f\circ T^k)$$

Προφανώς $\mu_n\in\mathcal{M}(X)$ για κάθε $n\in\mathbb{N}$. Επειδή το $\mathcal{M}(X)$ είναι συμπαγές, υπάρχει μία υπακολουθία $(\mu_{n_\lambda})_{\lambda\in\mathbb{N}}$ που συγκλίνει σε κάποιο $\mu\in\mathcal{M}(X)$, δηλαδή $\mu_{n_\lambda}(f)\to\mu(f)$ για κάθε $f\in C(X)$. Για κάθε $\lambda\in\mathbb{N}$ έχουμε

$$T^*(\mu_{n_\lambda})-\mu_{n_\lambda}=\frac{1}{n_\lambda}\left[T^{*n_\lambda}(\mu_0)-\mu_0\right]$$

και για κάθε $f\in C(X)$

$$|T^{*n_\lambda}(\mu_0)(f)-\mu_0(f)|=|\mu_0(f\circ T^{n_\lambda})-\mu_0(f)|\le|\mu_0(f\circ T^{n_\lambda})|+|\mu_0(f)|\le 2\|f\|$$

Άρα για κάθε $f\in C(X)$ έχουμε

$$|T^*(\mu)(f)-\mu(f)|\le\frac{2\|f\|}{n_\lambda}\longrightarrow 0 \quad \text{όταν} \quad \lambda\to+\infty$$

που σημαίνει ότι $T^*(\mu)=\mu$ ο.ε.δ.

Ομοίως αποδεικνύεται ότι κάθε δυναμικό σύστημα $\varphi:\mathbb{R}\times X\to X$ σε ένα συμπαγή μετρικό χώρο έχει μία τουλάχιστον αμετάβλητη πιθανότητα Borel. Η ροή φ επάγει μία ροή $\varphi^*:\mathbb{R}\times\mathcal{M}(X)\to\mathcal{M}(X)$ με τύπο

$$\varphi^*(t,\mu)(f)=\mu(f\circ\varphi_t)\qquad\text{για κάθε } f\in C(X)$$

Το $\mu\in\mathcal{M}(X)$ είναι αμετάβλητο από την φ αν είναι σταθερό σημείο της φ^*. Η απόδειξη της ύπαρξης ενός αμετάβλητου από την φ μέτρου πιθανότητας Borel γίνεται όπως στο Θεώρημα 1.2 θεωρώντας την ακολουθία

$$\mu_n=\frac{1}{n}\int_0^{n-1}\varphi^*(t,\mu_0)\,dt\ ,\ n\in\mathbb{N}.$$

Αν C είναι το κέντρο Birkhoff του φ, τότε σύμφωνα με το Θεώρημα επαναφοράς του Poincaré $\mu(C)=1$ για κάθε αμετάβλητο $\mu\in\mathcal{M}(X)$. Συνεπώς $\operatorname{supp}\mu\subset C$ για κάθε αμετάβλητο $\mu\in\mathcal{M}(X)$. Είναι επίσης φανερό ότι το $\operatorname{supp}\mu$ είναι κλειστό και αμετάβλητο υποσύνολο του X για κάθε αμετάβλητο $\mu\in\mathcal{M}(X)$.

Στην συνέχεια θα συμβολίζουμε με $\mathcal{M}_T(X)$ το σύνολο των αμετάβλητων πιθανοτήτων Borel μιας συνεχούς επί απεικόνισης $T: X\to X$ και με $\mathcal{M}_\phi(X)$ το σύνολο των αμετάβλητων πιθανοτήτων Borel ενός δυναμικού συστήματος ϕ στον X.

<u>1.5 Θεώρημα</u> Το $\mathcal{M}_T(X)$ είναι ένα (ασθενώς) συμπαγές, κυρτό σύνολο του οποίου τα ακρότατα είναι ακριβώς τα εργοδικά μέτρα της T, όταν η T είναι ομοιομορφισμός. Όμοια ισχύουν και για το $\mathcal{M}_\phi(X)$.

<u>Απόδειξη</u> Η συμπάγεια και η κυρτότητα του $\mathcal{M}_T(X)$, ως υποσυνόλου του ασθενούς $C(X)^*$, είναι προφανείς. Έστω $\mu\in\mathcal{M}_T(X)$, που δεν είναι εργοδικό. Τότε υπάρχει ένα σύνολο Borel $A\subset X$ με $T^{-1}(A)=A$ και $0<\mu(A)<1$. Θεωρούμε τα $\mu_1,\mu_2\in\mathcal{M}(X)$ με τύπους

$$\mu_1(B)=\frac{\mu(A\cap B)}{\mu(A)} \quad \text{και} \quad \mu_2(B)=\frac{\mu(B\setminus A)}{\mu(X\setminus A)}$$

για κάθε σύνολο Borel $B\subset X$. Τότε $\mu_1(A)=1\neq 0=\mu_2(A)$ και συνεπώς $\mu_1\neq\mu_2$. Επιπλέον

$$\mu_1(T^{-1}(B))=\frac{\mu(A\cap T^{-1}(B))}{\mu(A)}=\frac{\mu(T^{-1}(A)\cap T^{-1}(B))}{\mu(A)}=\frac{\mu(T^{-1}(A\cap B))}{\mu(A)}$$

$$=\frac{\mu(A\cap B)}{\mu(A)}=\mu_1(B)$$

και όμοια $\mu_2(T^{-1}(B))=\mu_2(B)$. Συνεπώς $\mu_1,\mu_2\in\mathcal{M}_T(X)$. Επιπλέον

$$\mu(B)=\mu(A)\cdot\mu_1(B)+(1-\mu(A))\,\mu_2(B)$$

Άρα το μ δεν είναι ακρότατο σημείο του $\mathcal{M}_T(X)$.

Αντίστροφα: έστω ότι το $\mu\in\mathcal{M}_T(X)$ είναι εργοδικό και $\mu=\lambda\mu_1+(1-\lambda)\mu_2$ για κάποια $\mu_1,\mu_2\in\mathcal{M}_T(X)$, $0\le\lambda\le 1$. Θα δείξουμε ότι $\mu=\mu_1=\mu_2$. Προφανώς $\mu_1\ll\mu$ και υπάρχει η παράγωγος Radon-Nikodym $\frac{d\mu_1}{d\mu}$. Επειδή το μ είναι εργοδικό, η $\frac{d\mu_1}{d\mu}$ είναι μ-σχεδόν παντού σταθερή ίση π.χ. με $\alpha\in\mathbb{R}$, γιατί είναι αμετάβλητη μ-σχεδόν παντού. Πράγματι· για κάθε σύνολο Borel $B\subset X$ έχουμε:

$$\mu_1(B)=\mu_1(T^{-1}(B))=\int_{T^{-1}(B)}\left(\frac{d\mu_1}{d\mu}\right)d\mu=\int_B\left(\frac{d\mu_1}{d\mu}\circ T^{-1}\right)d(T^*(\mu))$$

$$=\int_B\left(\frac{d\mu_1}{d\mu}\circ T^{-1}\right)d\mu$$

αφού $\mu,\mu_1\in\mathcal{M}_T(X)$. Από την μοναδικότητα της παραγώγου Radon-Nikodym

προκύπτει ότι $\frac{d\mu_1}{d\mu}\circ T = \frac{d\mu_1}{d\mu}$ μ-σχεδόν παντού.
Έχουμε τώρα

$$1 = \mu_1(X) = \int_X \left(\frac{d\mu_1}{d\mu}\right) d\mu = \int_X a\, d\mu = a$$

Άρα $\mu = \mu_1$ ο.ε.δ.

Απ το Θεώρημα Krein-Milman και το Θεώρημα 1.3 προκύπτει ότι το $\mathcal{M}_T(X)$ είναι το μικρότερο κυρτό, ασθενώς κλειστό σύνολο που περιέχει τα εργοδικά αμετάβλητα μέτρα πιθανότητας Borel. Αφού $\mathcal{M}_T(X) \neq \varphi$, προκύπτει ότι υπάρχει πάντα ένα εργοδικό αμετάβλητο μέτρο πιθανότητας Borel. Αν αυτό είναι μοναδικό, τότε το $\mathcal{M}_T(X)$ είναι μονοσύνολο.

2. Μοναδικά εργοδικά συστήματα

Έστω X ένας συμπαγής μετρικοποιήσιμος χώρος. Μία συνεχής επί απεικόνιση $T: X \to X$ λέγεται <u>μοναδικά εργοδική</u> αν το $\mathcal{M}_T(X)$ είναι μονοσύνολο. Όμοια ένα δυναμικό σύστημα $\varphi: \mathbb{R}\times X \to X$ λέγεται μοναδικά εργοδικό αν το $\mathcal{M}_\varphi(X)$ είναι μονοσύνολο.

<u>2.1. Θεώρημα</u> Για μία συνεχή επί απεικόνιση $T: X \to X$ τα ακόλουθα είναι ισοδύναμα:
(α) Η T είναι μοναδικά εργοδική.
(β) Για κάθε $f \in C(X)$ η ακολουθία συνεχών συναρτήσεων $\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k$, $n \in \mathbb{N}$ συγκλίνει ομοιόμορφα στον X σε μία σταθερά.
(γ) Για κάθε $f \in C(X)$, το όριο $\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$ υπάρχει για κάθε $x \in X$ και δεν εξαρτάται από το x.

<u>Απόδειξη</u> (α) ⇒ (β) Έστω ότι $\mathcal{M}_T(X) = \{\mu\}$. Θα δείξουμε ότι για κάθε $f \in C(X)$ ισχύει $\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k = \mu(f)$ ομοιόμορφα στον X με απαγωγή σε άτοπο. Έστω λοιπόν ότι υπάρχει $f \in C(X)$ ώστε η ακολουθία $\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k$, $n\in\mathbb{N}$ δεν συγκλίνει ομοιόμορφα στο $\mu(f)$. Δηλαδή υπάρχουν $\varepsilon > 0$ και $n_\lambda \to +\infty$, $x_\lambda \in X$ ώστε

$$\left| \frac{1}{n_\lambda}\sum_{k=0}^{n_\lambda-1} f\circ T^k(x_\lambda) - \mu(f)\right| \geq \varepsilon \quad \text{για κάθε } \lambda \in \mathbb{N}$$

Για κάθε $\lambda \in \mathbb{N}$, ο τύπος $\mu_\lambda(f) = \frac{1}{n_\lambda}\sum_{k=0}^{n_\lambda-1} f\circ T^k(x_\lambda)$, ορίζει μία θετική γραμμική μορφή $\mu_\lambda: C(X) \to \mathbb{R}$ με $\mu_\lambda(1) = 1$ και συνεπώς ένα στοιχείο του $\mathcal{M}(X)$.

Αφού το $\mathcal{M}(X)$ είναι συμπαγές, μπορούμε να υποθέσουμε ότι η ακολουθία $(\mu_\lambda)_{\lambda\in\omega}$ συγκλίνει σε κάποιο $\mu_0 \in \mathcal{M}(X)$. Λόγω των προηγουμένων έχουμε $|\mu_0(f) - \mu(f)| \geq \varepsilon$ και συνεπώς $\mu_0 \neq \mu$. Αρκεί να δείξουμε τώρα ότι $\mu_0 \in \mathcal{M}_T(X)$

Για κάθε $g \in C(X)$ έχουμε:

$$\mu_0(g\circ T) = \lim_{\lambda\to+\infty} \mu_\lambda(g\circ T) = \lim_{\lambda\to+\infty} \frac{1}{n_\lambda}\sum_{k=1}^{n_\lambda} g\circ T^k(x_\lambda) \quad \text{και}$$

$$\frac{1}{n_\lambda}\sum_{k=1}^{n_\lambda} g\circ T^k(x) = \mu_\lambda(g) + \frac{1}{n_\lambda}\left[g\circ T^{n_\lambda}(x_\lambda) - g(x_\lambda)\right] \quad \text{ενώ}$$

$$\frac{1}{n_\lambda}\left|g\circ T^{n_\lambda}(x_\lambda) - g(x_\lambda)\right| \leq \frac{2\|g\|}{n_\lambda} \to 0 \quad \text{όταν } \lambda\to+\infty$$

Άρα $\mu_0(g\circ T) = \lim_{\lambda\to+\infty}\mu_\lambda(g) = \mu_0$.

Το (β) $\Rightarrow$ (γ) είναι τετριμένο

(γ) $\Rightarrow$ (α) Λόγω της υπόθεσης ορίζεται η $\mu : C(X) \to \mathbb{R}$ με τύπο

$$\mu(f) = \lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k \in \mathbb{R}\,, \quad f\in C(X)$$

που είναι προφανώς μία θετική γραμμική μορφή με $\mu(1)=1$ και ορίζει ένα στοιχείο του $\mathcal{M}_T(X)$. Θα δείξουμε ότι $\mathcal{M}_T(X) = \{\mu\}$. Έστω $\nu \in \mathcal{M}_T(X)$

Για κάθε $f\in C(X)$ και $n\in\mathbb{N}$ έχουμε:

$$\nu(f) = \frac{1}{n}\sum_{k=0}^{n-1}\nu(f\circ T^k) = \nu\left(\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right)$$

Αφού $\left|\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x)\right| \leq \|f\|$ για κάθε $x\in X$, προκύπτει από το Θεώρημα της κυριαρχημένης σύγκλισης ότι

$$\nu(f) = \nu\left(\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right) = \nu(\mu(f)) = \mu(f) \quad \text{οεδ.}$$

Με όμοιο τρόπο, αντικαθιστώντας τα αθροίσματα με ολοκληρώματα Riemann και χρησιμοποιώντας το Θεώρημα του Fubini αποδεικνύεται το ακόλουθο:

<u>2.2. Θεώρημα.</u> Για ένα δυναμικό σύστημα $\varphi : \mathbb{R}\times X \to X$ τα επόμενα είναι ισοδύναμα:

(α) Το σύστημα φ είναι μοναδικά εργοδικό

(β) Για κάθε $f\in C(X)$ υπάρχει μία σταθερά $\mu(f)$ ώστε $\lim_{t\to\pm\infty}\frac{1}{t}\int_0^t f(\varphi_s(x))ds = \mu(f)$ ομοιόμορφα στον X.

(γ) Για κάθε $f\in C(X)$ και $x\in X$ το όριο $\lim_{t\to\pm\infty}\frac{1}{t}\int_0^t f(\varphi_s(x))\,ds$ υπάρχει και δεν εξαρτάται από το x.

Η δυναμική συμπεριφορά των μοναδικά εργοδικών δυναμικών συστημάτων περιγράφεται από την ακόλουθη πρόταση.

<u>2.3. Πρόταση</u> Αν $\phi: \mathbb{R}\times X \to X$ είναι ένα μοναδικά εργοδικό δυναμικό σύστημα στον συμπαγή μετρικοποιήσιμο χώρο X με αμετάβλητη πιθανότητα Borel μ, τότε το $\operatorname{supp}\mu$ είναι ελάχιστο σύνολο και $\operatorname{supp}\mu \subset L^+(x)\cap L^-(x)$ για κάθε $x\in X$. Επιπλέον το περιορισμένο δυναμικό σύστημα στο $\operatorname{supp}\mu$ είναι μοναδικά εργοδικό.

<u>Απόδειξη</u> Αφού το $\operatorname{supp}\mu$ είναι κλειστό και αμετάβλητο, αρκεί να δείξουμε τον δεύτερο ισχυρισμό. Έστω $x_0 \in \operatorname{supp}\mu$ και V μία ανοικτή περιοχή του x_0 στον X. Τότε $\mu(V)>0$ και αφού $\mu(V)=\sup\{\mu(F) : F\subset V \text{ και } F \text{ κλειστό στον } X\}$, υπάρχει $F\subset V$ κλειστό στον X με $\mu(F)>0$. Υπάρχει μία συνεχής συνάρτηση $f: X\to[0,1]$ με $f^{-1}(1)=F$ και $f^{-1}(0)=X\setminus V$. Για κάθε $x\in X$ έχουμε τώρα απ' το Θεώρημα 2.2,

$$\lim_{t\to\pm\infty} \frac{1}{t}\int_0^t f(\phi_s(x))\,ds = \mu(f) \geq \int_F f\,d\mu = \mu(F) > 0$$

Συνεπώς υπάρχουν $t_1<0<t_2$ με $\phi_{t_1}(x), \phi_{t_2}(x) \in V$. Αυτό δείχνει ότι $\operatorname{supp}\mu \subset L^+(x)\cap L^-(x)$ για κάθε $x\in X$. Ο τελευταίος ισχυρισμός είναι προφανής.

Στην συνέχεια θα δούμε μερικά παραδείγματα μοναδικά εργοδικών συστημάτων.

<u>2.4. Πρόταση</u> Έστω $T: X\to X$ μία συνεχής επί απεικόνιση ώστε:

(i) η οικογένεια $\{T^k : k\in\mathbb{Z}^+\}$ είναι ισοσυνεχής (ως προς κάποια συμβιβαστή μετρική).

(ii) υπάρχει $x_0\in X$ ώστε $\overline{\{T^k(x_0) : k\in\mathbb{Z}^+\}}=X$.

Τότε η T είναι μοναδικά εργοδική.

<u>Απόδειξη</u> Για κάθε $f\in C(X)$, η οικογένεια συνεχών συναρτήσεων $f\circ T^k: X\to\mathbb{R}$, $k\in\mathbb{Z}^+$, είναι επίσης ισοσυνεχής και το ίδιο ισχύει και για την

$$f_n = \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k \quad , n\in\mathbb{N}$$

που είναι ομοιόμορφα φραγμένη από το $\|f\|$. Από το Θεώρημα του Ascoli υπάρχει μία ακολουθία $n_\lambda\to+\infty$ ώστε η $(f_{n_\lambda})_{\lambda\in\mathbb{N}}$ να συγκλίνει ομοιόμορφα σε κάποια $g\in C(X)$. Τότε έχουμε για κάθε $x\in X$,

$$g(T(x)) = \lim_{\lambda\to+\infty} \frac{1}{n_\lambda}\sum_{k=1}^{n_\lambda} f\circ T^k(x) = g(x)$$

Ειδικά για $x=x_0$ προκύπτει ότι $g(T^k(x_0)) = g(x_0)$ για κάθε $k\in\mathbb{Z}^+$ και λόγω της υπόθεσης (ii) πρέπει αναγκαστικά να είναι σταθερή ίση με $g(x_0)$ σ' όλο του X. Για κάθε $\mu\in\mathcal{M}_T(X)$ έχουμε τώρα

$$\mu(f) = \lim_{\lambda\to+\infty} \frac{1}{n_\lambda}\sum_{k=0}^{n_\lambda-1}\mu(f\circ T^k) = \lim_{\lambda\to+\infty}\mu(f_n) = \mu(g) = g(x_0).$$

Αυτό δείχνει ότι το $\mathcal{M}_T(X)$ είναι μονοσύνολο.

Σημειώνουμε ότι αν η $T: X\to X$ είναι ομοιομορφισμός, τότε η υπόθεση για την ισοσυνέχεια της οικογένειας $\{T^k : k\in\mathbb{Z}\}$ ως προς κάποια συμβιβαστή μετρική του X είναι ισοδύναμη με το ότι η T είναι ισομετρία ως προς κάποια συμβιβαστή μετρική του X.

Όμοια έχουμε:

2.5. Πρόταση Κάθε ισοσυνεχές δυναμικό σύστημα $\varphi: \mathbb{R}\times X\to X$ με τουλάχιστον μία πυκνή τροχιά είναι μοναδικά εργοδικό.

Έστω τώρα G μία $1^η$ αριθμήσιμη, συμπαγής τοπολογική ομάδα (Hausdorff). Τότε, η G είναι μετρικοποιήσιμη και η τοπολογία της ορίζεται από μία μετρική d που είναι αμετάβλητη από τις αριστερές (και δεξιές) μεταφορές. Δηλαδή $d(gx,gy)=d(x,y)$ για κάθε $g\in G$, $x,y\in G$.

Παραδείγματος χάριν, αν $G=S^1$, τότε η d είναι η ευκλείδεια απόσταση από το $\mathbb{C}$, αφού $|gx-gy| = |g|\,|x-y| = |x-y|$ για κάθε $g\in S^1$. Γενικότερα, αν $G=\mathbb{T}^n$ είναι ο n-torus και $d(x,y)=\sum_{i=1}^n |x_i-y_i|$, όπου $x=(x_1,\dots,x_n)$ και $y=(y_1,\dots,y_n)$, τότε για κάθε $g=(g_1,\dots,g_n)\in\mathbb{T}^n$ έχουμε:

$$d(gx,gy) = \sum_{i=1}^n |g_ix_i-g_iy_i| = \sum_{i=1}^n |g_i|\,|x_i-y_i| = \sum_{i=1}^n |x_i-y_i| = d(x,y).$$

Έτσι για κάθε $g\in G$, αν $T_g: G\to G$ είναι η αριστερή μεταφορά $T_g(x)=gx$, τότε το $\{T_g^k : k\in\mathbb{Z}^+\}$ είναι ισοσυνεχής οικογένεια ομοιομορφισμών της G, που αφήνουν αμετάβλητο το μέτρο Haar της G.

2.6. Πρόταση Έστω G μία $1^η$ αριθμήσιμη συμπαγής τοπολογική ομάδα Hausdorff και $g\in G$. Η μεταφορά $T_g: G\to G$ είναι μοναδικά εργοδική τότε και μόνον τότε αν έχει μία πυκνή τροχιά στην G.

Απόδειξη Αν η T_g είναι μοναδικά εργοδική, τότε αφού είναι ομοιομορφισμός, από την Πρόταση 2.3 το $\operatorname{supp}\mu$ είναι ελάχιστο σύνολο, όπου το μ είναι το

μέτρο Haar στην G. Αφού όμως $G = supp\,\mu$, κάθε τροχιά είναι πυκνή στην G.
Το αντίστροφο είναι άμεσο από την Πρόταση 2.4.

<u>2.7. Παρατήρηση</u> Αν G είναι μία συμπαγής τοπολογική ομάδα Hausdorff για την οποία υπάρχουν $g \in G, x_0 \in G$ ώστε το $\{g^n x_0 : n \in \mathbb{Z}\}$ να είναι πυκνό στην G, τότε η G είναι και αναγκαία αβελιανή. Πράγματι· επειδή η απεικόνιση $h: G \to G$ με $h(x) = x x_0^{-1}$ είναι ομοιομορφισμός, η αβελιανή υποομάδα $\{g^n : n \in \mathbb{Z}\}$ είναι πυκνή στην G. Έστω τώρα $x, y \in G$. Υπάρχουν ακολουθίες $(g^{n_k})_{k\in\mathbb{N}}$, $(g^{m_k})_{k\in\mathbb{N}}$ με $x = \lim_{k\to+\infty} g^{n_k}$ και $y = \lim_{k\to+\infty} g^{m_k}$. Συνεπώς

$$xy = \lim_{k\to+\infty} g^{n_k} \cdot \lim_{k\to+\infty} g^{m_k} = \lim_{k\to+\infty} (g^{n_k}\cdot g^{m_k}) = \lim_{k\to+\infty} (g^{m_k}\cdot g^{n_k}) = \lim_{k\to+\infty} g^{m_k} \cdot \lim_{k\to+\infty} g^{n_k} = yx.$$

<u>2.8. Θεώρημα (Kronecker)</u> Αν οι πραγματικοί αριθμοί $1, a_1, \ldots, a_n$ είναι γραμμικώς ανεξάρτητοι πάνω στο σώμα $\mathbb{Q}$, τότε κάθε τροχιά της μεταφοράς

$$T(e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_n}) = (e^{2\pi i(\theta_1+a_1)}, \ldots, e^{2\pi i(\theta_n+a_n)})$$

του n-torus T^n είναι πυκνή στο T^n. Συνεπώς η T είναι μοναδικά εργοδική.

<u>Απόδειξη</u> Βλ. Κεφάλαιο XXIII του βιβλίου των G.H. Hardy και E.M. Wright, An introduction to the Theory of Numbers, (Fifth Edition), Oxford 1979.

Ειδικά για τον μοναδιαίο κύκλο $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ ισχύει το ακόλουθο:

<u>2.9. Θεώρημα (Denjoy - van Kampen - Fürstenberg)</u> Κάθε ομοιομορφισμός $h: S^1 \to S^1$ χωρίς περιοδικά σημεία είναι μοναδικά εργοδικός.

Αντίστοιχο αποτέλεσμα για τον n-torus T^n, $n \geq 2$ δεν ισχύει.

<u>2.10. Θεώρημα (Fürstenberg)</u> Υπάρχει μία αναλυτική αμφιδιαφόριση $h: T^2 \to T^2$ που διατηρεί το μέτρο Haar, κάθε τροχιά της είναι πυκνή στον T^2 και είναι (Haar) σχεδόν παντού συζυγής με την μεταφορά $h_\gamma(z_1, z_2) = (\gamma z_1, z_2)$ για κάποιο $\gamma \in S^1$. Αφού η h_γ δεν είναι εργοδική, προκύπτει ότι ούτε η h είναι. Συνεπώς η h δεν είναι μοναδικά εργοδική.

Η απόδειξη των δύο τελευταίων Θεωρημάτων βρίσκεται στην εργασία του H. Fürstenberg, Strict ergodicity and transformation of the torus, Amer. J. Math. 83 (1961), 573-601.

<u>2.11. Παρατήρηση</u> Οι συνεχείς επί ενδομορφισμοί συμπαγών τοπολογικών ομάδων δεν είναι μοναδικά εργοδικοί, αφού εκτός από το μέτρο Haar διατηρούν και το σημειακό μέτρο του ουδετέρου στοιχείου.

3. Η εργοδική ανάλυση των αμετάβλητων μέτρων

Έστω X ένας συμπαγής μετρικοποιήσιμος χώρος και $T: X \to X$ μια συνεχής επί απεικόνιση. Ένα σύνολο Borel $E \subset X$ λέγεται μηδενικής πιθανότητας αν $\mu(E)=0$ για κάθε $\mu \in M_T(X)$. Αν $\mu(E)>0$ για κάποιο $\mu \in M_T(X)$, το E λέγεται θετικής πιθανότητας. Το E λέγεται μέγιστης πιθανότητας αν $\mu(E)=1$ για κάθε $\mu \in M_T(X)$.

3.1 Ορισμός. Ένα σημείο $x \in X$ λέγεται ημικανονικό (quasi-regular) αν για κάθε $f \in C(X)$ το όριο

$$\tilde{f}(x) = \lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x)$$

υπάρχει στο $\mathbb{R}$.

3.2. Θεώρημα. Το σύνολο Q όλων των ημικανονικών σημείων είναι αμετάβλητο, Borel και μέγιστης πιθανότητας.

Απόδειξη. Η αμεταβλητότητα του Q είναι άμεση από την εκτίμηση

$$\left| \frac{1}{n} \sum_{k=1}^{n} f \circ T^k(x) - \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) \right| = \frac{1}{n} \left| f \circ T^k(x) - f(x) \right| \le \frac{2\|f\|}{n}$$

Έστω $\{f_n : n \in \mathbb{N}\}$ ένα αριθμήσιμο πυκνό υποσύνολο του $C(X)$, που υπάρχει γιατί ο X είναι μετρικοποιήσιμος. Θεωρούμε τα σύνολα

$$E_n = \left\{ x \in X : \text{ το όριο } \lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_n \circ T^k(x) \text{ δεν υπάρχει στο } \mathbb{R} \right\}, \quad n \in \mathbb{N}$$

Για κάθε $n, m, \lambda \in \mathbb{N}$ το σύνολο

$$E_{n,m,\lambda} \left\{ x \in X : \left| \frac{1}{n_1} \sum_{k=0}^{n_1-1} f_n \circ T^k(x) - \frac{1}{n_2} \sum_{k=0}^{n_2-1} f_n \circ T^k(x) \right| \le \frac{1}{m} \text{ για κάθε } n_1, n_2 \ge \lambda \right\}$$

είναι κλειστό και

$$X \setminus E_n = \bigcap_{m=1}^{\infty} \bigcup_{\lambda=1}^{\infty} E_{n,m,\lambda}$$

Άρα τα E_n, $n \in \mathbb{N}$ είναι σύνολα Borel. Είναι προφανές ότι αν $E = \bigcup_{n=1}^{\infty} E_n$, τότε $Q \subset X \setminus E$. Από το εργοδικό θεώρημα του Birkhoff επίσης έχουμε $\mu(E_n)=0$ για κάθε $n \in \mathbb{N}$ και συνεπώς $\mu(X \setminus E) = 1$. Αρκεί λοιπόν να δείξουμε ότι $X \setminus E \subset Q$. Έστω $x \in X \setminus E$, $f \in C(X)$ και $\varepsilon > 0$. Υπάρχει $n \in \mathbb{N}$ με $\|f - f_n\| < \frac{\varepsilon}{3}$. Επειδή $x \in X \setminus E_n$ υπάρχει $\lambda \in \mathbb{N}$ ώστε

$$\left| \frac{1}{n_1} \sum_{k=0}^{n_1-1} f_n \circ T^k(x) - \frac{1}{n_2} \sum_{k=0}^{n_2-1} f_n \circ T^k(x) \right| < \frac{\varepsilon}{3}$$

για κάθε $n_1, n_2 \ge \lambda$. Συνεπώς, για κάθε $n_1, n_2 \ge \lambda$ έχουμε:

$$\left|\frac{1}{n_1}\sum_{k=0}^{n_1-1} f\circ T^k(x) - \frac{1}{n_2}\sum_{k=0}^{n_2-1} f\circ T^k(x)\right| \leq$$

$$\left|\frac{1}{n_1}\sum_{k=0}^{n_1-1} f\circ T^k(x) - \frac{1}{n_1}\sum_{k=0}^{n_1-1} f_n\circ T^k(x)\right| + \left|\frac{1}{n_1}\sum_{k=0}^{n_1-1} f_n\circ T^k(x) - \frac{1}{n_2}\sum_{k=0}^{n_2-1} f_n\circ T^k(x)\right| + \left|\frac{1}{n_2}\sum_{k=0}^{n_2-1} f_n\circ T^k(x) - \frac{1}{n_2}\sum_{k=0}^{n_2-1} f\circ T^k(x)\right|$$

$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$.

Αυτό δείχνει ότι το όριο $\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x)$ υπάρχει στο $\mathbb{R}$. ο.ε.δ.

Είναι προφανές ότι για κάθε $x\in Q$ η απεικόνιση $\mu_x : C(X)\to\mathbb{R}$ με τύπο $\mu_x(f) = \tilde f(x)$ είναι μία θετική γραμμική μορφή με $\mu_x(1)=1$, και $\mu_x(f\circ T) = \mu_x(f)$. Συνεπώς $\mu_x \in \mathcal{M}_T(X)$ για κάθε $x\in Q$. Στη συνέχεια θα μελετήσουμε την σχέση των μ_x, $x\in Q$ μεταξύ τους καθώς και με τα υπόλοιπα στοιχεία του $\mathcal{M}_T(X)$, αν υπάρχουν. Κατ'αρχήν είναι προφανές ότι $\mu_x = \mu_y$, αν $y = T^k(x)$ για κάποιο $k\in\mathbb{Z}^+$.

<u>3.3. Λήμμα</u> Για κάθε $f\in C(X)$ και $\mu\in\mathcal{M}_T(X)$ ισχύει:

$$\int_X f\,d\mu = \int_Q\left(\int_X f\,d\mu_x\right)d\mu$$

<u>Απόδειξη</u> Η συνάρτηση $Q\ni x\longrightarrow \mu_x(f)\in\mathbb{R}$ είναι μετρήσιμη ως κατά σημείο όριο συνεχών συναρτήσεων. Για κάθε $n\in\mathbb{N}$ έχουμε:

$$\int_X f\,d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\int_X f\,d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\int_X (f\circ T^k)\,d\mu = \int_X\left(\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right)d\mu$$

Απ' το Θεώρημα της κυριαρχημένης σύγκλισης και το Θεώρημα 3.2 έχουμε

$$\int_Q\left(\int_X f\,d\mu_x\right)d\mu = \lim_{n\to+\infty}\int_Q\left(\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right)d\mu = \int_X f\,d\mu \quad \text{ο.ε.δ.}$$

Έστω $\mu\in\mathcal{M}_T(X)$. Απ' το εργοδικό Θεώρημα του Birkhoff, για κάθε φραγμένη μετρήσιμη συνάρτηση $f: Q\to\mathbb{R}$ το όριο $\tilde f(x) = \lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x)$ υπάρχει μ-σχεδόν παντού. Θεωρούμε τον διανυσματικό χώρο

$$E(\mu) = \left\{ f: Q\to\mathbb{R} \text{ φραγμένη, μετρήσιμη ώστε } \int_X f\,d\mu_x = \tilde f(x)\ \mu\text{-σχεδόν παντού}\right\}$$

Σύμφωνα με το Θεώρημα 3.2, $C(X)\in E(\mu)$. Για την επέκταση της ισχύος του Λήμματος 3.3 στις φραγμένες μετρήσιμες συναρτήσεις θα χρειαστούμε μία σειρά από λήμματα.

<u>3.4. Λήμμα</u> Αν $(f_n)_{n\in\mathbb{N}}$ είναι μία ομοιόμορφα φραγμένη ακολουθία στον $E(\mu)$ και $f_n \to f$ κατά σημείο, τότε $f \in E(\mu)$.

<u>Απόδειξη</u> Από το Θεώρημα της κυριαρχημένης σύγκλισης έχουμε

$$\int_X f\,d\mu_x = \lim_{n\to+\infty}\int_X f_n\,d\mu_x = \lim_{n\to+\infty}\tilde{f}_n(x)$$

και από το εργοδικό Θεώρημα του Birkhoff

$$\int_X|\tilde{f}-\tilde{f}_n|\,d\mu \le \int_X \widetilde{|f-f_n|}\,d\mu = \int_X|f-f_n|\,d\mu \longrightarrow 0 \quad , \text{όταν } n\to+\infty.$$

Άρα $\tilde{f}_n \to \tilde{f}$ στον $L^1(\mu)$ και συνεπώς υπάρχει μία υπακολουθία $(f_{n_\lambda})_{\lambda\in\mathbb{N}}$ ώστε $\tilde{f}_{n_\lambda} \to \tilde{f}$ μ-σχεδόν παντού. Προκύπτει ότι

$$\int_X f\,d\mu_x = \lim_{\lambda\to+\infty}\int_X f_{n_\lambda}\,d\mu_x = \lim_{\lambda\to+\infty}\tilde{f}_{n_\lambda}(x) = \tilde{f}(x) \quad \mu\text{-σχεδόν παντού}$$

<u>3.5. Λήμμα</u> Για κάθε κλειστό σύνολο $A\subset X$ ισχύει $\chi_A \in E(\mu)$.

<u>Απόδειξη</u> Υπάρχει μία ακολουθία συνεχών συναρτήσεων $f_n: X\to[0,1]$, $n\in\mathbb{N}$, με $f_n \to \chi_A$ κατά σημείο. Το συμπέρασμα είναι λοιπόν άμεσο από το Λήμμα 3.4.

<u>3.6. Λήμμα.</u> Για κάθε σύνολο Borel $A\subset X$ ισχύει $\chi_A \in E(\mu)$.

<u>Απόδειξη</u> Υπάρχει μία ακολουθία κλειστών συνόλων $A_1\subset A_2\subset\cdots\subset A$ με $\mu(A\setminus\bigcup_{n=1}^{\infty}A_n)=0$, από την κανονικότητα του μέτρου μ. Συνεπώς $\chi_{A_n}\to\chi_A$ μ-σχεδόν παντού και οι χ_{A_n}, $n\in\mathbb{N}$ κυριαρχούνται από την χ_A. Αν το εργοδικό Θεώρημα του Birkhoff έχουμε:

$$\int_X|\tilde{\chi}_{A_n}-\tilde{\chi}_A|\,d\mu \le \int_X\widetilde{|\chi_{A_n}-\chi_A|}\,d\mu = \int_X|\chi_{A_n}-\chi_A|\,d\mu \to 0, \quad \text{όταν } n\to+\infty$$

Άρα $\tilde{\chi}_{A_n}\to\tilde{\chi}_A$ στον $L^1(\mu)$ και συνεπώς υπάρχει μία υπακολουθία $(\chi_{A_{n_\lambda}})_{\lambda\in\mathbb{N}}$ ώστε $\tilde{\chi}_{A_{n_\lambda}}\to\tilde{\chi}_A$ μ-σχεδόν παντού. Από το Θεώρημα της κυριαρχημένης σύγκλισης έχουμε τώρα

$$\int_X\chi_A\,d\mu_x \ge \limsup_{n\to+\infty}\int_X\chi_{A_n}\,d\mu_x = \limsup_{n\to+\infty}\tilde{\chi}_{A_n}(x) \ge \lim_{\lambda\to+\infty}\tilde{\chi}_{A_{n_\lambda}}(x) = \tilde{\chi}_A(x)$$

μ-σχεδόν παντού. Ομοίως $\int_X\chi_{X\setminus A}\,d\mu_x \ge \tilde{\chi}_{X\setminus A}(x)$ μ-σχεδόν παντού. Έτσι

$$\int_X \chi_A \, d\mu_x = 1 - \int_X \chi_{X\setminus A} \, d\mu_x \leq 1 - \tilde{\chi}_{X\setminus A}(x) = \tilde{\chi}_A(x) \leq \int_X \tilde{\chi}_A \, d\mu_x$$

μ-σχεδόν παντού ο.ε.δ.

3.7. Πρόταση Έστω $\mu \in \mathcal{M}_T(X)$. Για κάθε φραγμένη μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$ ισχύει

$$\int_X f \, d\mu_x = \bar{f}(x) \quad \mu\text{-σχεδόν παντού στον } X$$

Απόδειξη Αρκεί να δείξουμε ότι ο $E(\mu)$ ταυτίζεται με τον χώρο όλων των φραγμένων μετρήσιμων πραγματικών συναρτήσεων. Κάθε θετική μετρήσιμη φραγμένη συνάρτηση $f : X \to \mathbb{R}$ είναι το κατά σημείο όριο μιας ακολουθίας γραμμικών συνδυασμών χαρακτηριστικών συναρτήσεων συνόλων Borel. Συνεπώς $f \in E(\mu)$ από τα Λήμματα 3.4 και 3.6. Τέλος κάθε μετρήσιμη φραγμένη συνάρτηση γράφεται διαφορά δύο θετικών μετρήσιμων φραγμένων συναρτήσεων.

3.8. Πόρισμα Για κάθε $\mu \in \mathcal{M}_T(X)$ και μετρήσιμη φραγμένη συνάρτηση $f : X \to \mathbb{R}$ ισχύει

$$\int_Q \left(\int_X f \, d\mu_x \right) d\mu = \int_X f \, d\mu$$

3.9. Πρόταση: Έστω $\mu \in \mathcal{M}_T(X)$. Για κάθε $f \in C(X)$ ισχύει

$$\int_X |\tilde{f} - \tilde{f}(x)|^2 \, d\mu_x = 0 ,$$

μ-σχεδόν για κάθε $x \in Q$.

Απόδειξη Επειδή $\mu_x \in \mathcal{M}_T(X)$ έχουμε

$$\int_X \tilde{f} \, d\mu_x = \lim_{n\to+\infty} \int_X \left(\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \right) d\mu_x = \lim_{n\to+\infty} \int_X f \, d\mu_x = \tilde{f}(x). \quad \text{Άρα}$$

$$\int_X |\tilde{f} - \tilde{f}(x)|^2 \, d\mu_x = \tilde{f}(x)^2 - 2 \tilde{f}(x) \int_X \tilde{f} \, d\mu_x + \int_X \tilde{f}^2 \, d\mu_x = \int_X \tilde{f}^2 \, d\mu_x - \tilde{f}(x)^2$$

Ολοκληρώνοντας ως προς μ παίρνουμε

$$\int_Q \left(\int_X |\tilde{f} - \tilde{f}(x)|^2 \, d\mu_x \right) d\mu = \int_Q \left(\int_X \tilde{f}^2 \, d\mu_x \right) d\mu - \int_Q \tilde{f}^2 \, d\mu$$

Επειδή η $\tilde{f} : Q \to \mathbb{R}$ είναι φραγμένη και μετρήσιμη, το ίδιο είναι και η $\tilde{f}^2$.

Συνεπώς απ' το πόρισμα 3.8 έχουμε

$$\int_Q \left(\int_X |\tilde f - \tilde f(x)|^2 \, d\mu_x \right) d\mu = 0$$

και άρα $\int_X |\tilde f - \tilde f(x)|^2 \, d\mu_x = 0$ μ-σχεδόν για κάθε $x \in Q$.

<u>3.10. Θεώρημα</u> Το σύνολο

$$V = \{x \in Q : \int_X |\tilde f - \tilde f(x)|^2 \, d\mu_x = 0 \text{ για κάθε } f \in C(X)\}$$

$$= \{x \in Q : \mu_x = \mu_y, \ \mu_x\text{-σχεδόν για κάθε } y \in Q\}$$

είναι αμετάβλητο, Borel και μέγιστης πιθανότητας.

<u>Απόδειξη</u> Η αμεταβλητότητα του V είναι προφανής. Έστω $\{f_n : n \in \mathbb{N}\}$ ένα αριθμήσιμο πυκνό υποσύνολο του $C(X)$. Για κάθε $n \in \mathbb{N}$ το σύνολο

$$E_n = \{x \in Q : \int_X |\tilde f_n - \tilde f_n(x)|^2 \, d\mu_x > 0\}$$

είναι Borel και μηδενικής πιθανότητας, σύμφωνα με την πρόταση 3.9. Συνεπώς το σύνολο $E = \bigcup_{n=1}^{\infty} E_n$ είναι Borel και μηδενικής πιθανότητας, ενώ προφανώς $V \subset Q \setminus E$. Αρκεί να δείξουμε ότι $Q \setminus E \subset V$. Έστω $x \in Q \setminus E$, $f \in C(X)$ και $\varepsilon > 0$. Υπάρχει $n \in \mathbb{N}$ με $\|f - f_n\| < \varepsilon$. Αφού $x \in Q \setminus E_n$,

$$\int_X |\tilde f_n - \tilde f_n(x)|^2 \, d\mu_x = 0$$

ενώ

$$|\tilde f - \tilde f(x)|^2 \le |\tilde f - \tilde f_n|^2 + |\tilde f_n - \tilde f_n(x)|^2 + |\tilde f_n(x) - \tilde f(x)|^2$$

Προκύπτει ότι

$$\int_X |\tilde f - \tilde f(x)|^2 \, d\mu_x < 2\varepsilon^2$$

Άρα $\int_X |\tilde f - \tilde f(x)|^2 \, d\mu_x = 0$, ο.ε.δ.

<u>3.11. Θεώρημα</u> Για κάθε $x \in V$, το μ_x είναι εργοδικό.

<u>Απόδειξη</u> Έστω $x \in V$. Τότε για κάθε μετρήσιμη φραγμένη συνάρτηση $f : X \to \mathbb{R}$ ισχύει $\tilde f(x) = \int_X f \, d\mu_x = \int_X f \, d\mu_y = \tilde f(y)$, μ_x-σχεδόν για κάθε $y \in Q$.

Έστω τώρα $A \subset X$ ένα σύνολο Borel με $A = T^{-1}(A)$. Έχουμε

$$\int_X \chi_A \tilde f(x) \, d\mu_x = \int_X \chi_A \tilde f \, d\mu_x, \text{ συνεπώς } \tilde f(x) \mu_x(A) = \int_X \chi_A \tilde f \, d\mu_x$$

Αφού το A είναι αμετάβλητο από την T, η χ_A είναι αμετάβλητη, δηλαδή $\chi_A \circ T = \chi_A$. Έτσι έχουμε:

$$\tilde{f}(x)\,\mu_x(A) = \int_X \chi_A \tilde{f}\, d\mu_x = \int_X \chi_A \Big(\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\Big)\, d\mu_x$$

$$= \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \int_X (\chi_A\circ T^k)\cdot(f\circ T^k)\, d\mu_x$$

$$= \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \int_X \chi_A f\, d\mu_x = \int_A f\, d\mu_x$$

Ειδικά για $f = \chi_A$ παίρνουμε:

$$\tilde{\chi}_A(x)\cdot\mu_x(A) = \int_A \chi_A\, d\mu_x$$

και συνεπώς

$$\mu_x(A)\cdot\mu_x(A) = \mu_x(A),$$

δηλαδή $\mu_x(A) = 0$ ή 1 ο.ε.δ.

<u>3.12. Πρόταση.</u> Το σύνολο $D = \{x \in Q : x \in \operatorname{supp}\mu_x\}$ είναι αμετάβλητο, Borel και μέγιστης πιθανότητας.

<u>Απόδειξη</u> Η αμεταβλητότητα του D είναι άμεση από την αμεταβλητότητα του μ_x και την συνέχεια της T. Για κάθε $m \in \mathbb{N}$ υπάρχουν $x_{1m}, \dots, x_{k_m m} \in X$ ώστε $X = S(x_{1m}, \frac{1}{m}) \cup \dots \cup S(x_{k_m m}, \frac{1}{m})$, όπου στον X θεωρούμε μία συμβιβαστή μετρική. Υπάρχουν συνεχείς συναρτήσεις $f_{n,m} : X \to [0,1]$ με

$$S(x_{nm}, \tfrac{1}{m}) = f_{n,m}^{-1}(1) \quad \text{και} \quad X \setminus S(x_{nm}, \tfrac{2}{m}) = f_{n,m}^{-1}(0)$$

$1 \le n \le k_m$, $m \in \mathbb{N}$.

Η συνάρτηση $\tilde{f}_{n,m} : Q \to [0,1]$ είναι ως γνωστόν μετρήσιμη και αμετάβλητη. Συνεπώς τα σύνολα $E_{n,m} = \{x \in Q : \tilde{f}_{n,m}(x) = 0\}$ είναι Borel και αμετάβλητα.

Για κάθε $\mu \in \mathcal{M}_T(X)$ έχουμε

$$0 = \int_{E_{n,m}} \tilde{f}_{n,m}\, d\mu = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \int_{E_{n,m}} (f_{n,m}\circ T^k)\, d\mu = \int_{E_{n,m}} f_{n,m}\, d\mu \ge \int_{E_{n,m}\cap S_{n,m}} f_{n,m}\, d\mu$$

$$= \mu(E_{n,m} \cap S_{n,m})$$

όπου έχουμε θέσει για απλότητα $S_{n,m} = S(x_{nm}, \frac{1}{m})$, $1 \le n \le k_m$, $m \in \mathbb{N}$.

Συνεπώς το σύνολο Borel

$$E = Q \setminus \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{k_m} E_{n,m} \cap S_{n,m}$$

είναι μέγιστης πιθανότητας. Θα δείξουμε ότι $D = E$.

Εστω $x\in D$. Αν $x\in S_{n,m}$ για κάποια $1\le n\le k_m$, $m\in\mathbb{N}$, υπάρχει $\varepsilon>0$ ώστε $S(x,\varepsilon)\subset S_{n,m}$ και $\mu_x(S(x,\varepsilon))>0$. Συνεπώς

$$0<\mu_x(S(x,\varepsilon))\le \mu_x(S_{n,m}) < \int_X f_{n,m}\,d\mu_x = \tilde{f}(x),$$

που σημαίνει ότι $x\in Q\setminus E_{n,m}\cap S_{n,m}$. Αυτό δείχνει ότι $D\subset E$.
Εστω τώρα ότι $x\notin D$, δηλαδή υπάρχει $\varepsilon>0$ ώστε $\mu_x(S(x,\varepsilon))=0$. Εστω $m\in\mathbb{N}$ με $\frac{2}{m}<\frac{\varepsilon}{2}$. Υπάρχει $1\le n\le k_m$ με $x\in S_{n,m}$. Τότε

$$x\in S_{n,m}\subset S\left(x_{nm},\tfrac{2}{m}\right)\subset S(x,\varepsilon).$$

Άρα

$$0=\mu_x(S(x,\varepsilon))=\int_X \chi_{S(x,\varepsilon)}\,d\mu_x \ge \int_X f_{n,m}\,d\mu_x=\tilde{f}_{n,m}(x)$$

που σημαίνει ότι $x\in E_{n,m}\cap S_{n,m}$. ο.ε.δ.

Από τα μέχρι τώρα αποτελέσματα προκύπτει ότι το σύνολο $R=U\cap D$ είναι Borel, μετρήσιμο και μέγιστης πιθανότητας. Τα σημεία του R λέγονται <u>κανονικά</u> σημεία. Αν $x\in R$, τότε το μ_x είναι εργοδικό και $x\in\operatorname{supp}\mu_x$.

<u>3.13. Θεώρημα.</u> Εστω $\mu\in\mathcal{M}_T(X)$. Τότε κάθε $f\in L^1(\mu)$ είναι μ_x-ολοκληρώσιμη μ-σχεδόν για όλα τα $x\in R$ και

$$\int_X\left(\int_X f\,d\mu_x\right)d\mu=\int_X f\,d\mu$$

<u>Απόδειξη.</u> Το συμπέρασμα κατ' αρχήν ισχύει όταν η f είναι φραγμένη και μετρήσιμη, απ' το πόρισμα 3.8. Αν τώρα η $f\in L^1(\mu)$ είναι μη-αρνητική, είναι κατά σημείο όριο μιας αύξουσας ακολουθίας $(f_n)_{n\in\mathbb{N}}$ φραγμένων μετρήσιμων συναρτήσεων. Άρα

$$\int_X f\,d\mu_x=\lim_{n\to+\infty}\int_X f_n\,d\mu_x=\lim_{n\to+\infty}\tilde{f}_n(x)\le f(x)$$

μ-σχεδόν για κάθε $x\in X$. Η ακολουθία $(\tilde{f}_n)_{n\in\mathbb{N}}$ είναι επίσης αύξουσα και άνω φραγμένη από την f. Άρα

$$\int_X\left(\int_X f\,d\mu_x\right)=\int_X\left(\lim_{n\to+\infty}\tilde{f}_n(x)\right)d\mu=\lim_{n\to+\infty}\int_X\tilde{f}_n\,d\mu=\lim_{n\to+\infty}\int_X f_n\,d\mu=\int_X f\,d\mu$$

Αν η $f\in L^1(\mu)$ δεν είναι μη-αρνητική, τότε μπορούμε να την γράψουμε διαφορά δύο μη-αρνητικών. Το θεώρημα αποδείχθηκε.

<u>3.14. Πόρισμα.</u> Για κάθε σύνολο Borel $A\subset X$ και $\mu\in\mathcal{M}_T(X)$ ισχύει

$$\mu(A)=\int_X \mu_x(A)\,d\mu$$

3.15. Πόρισμα. Ένα σύνολο Borel $E\subset X$ είναι μεγίστης πιθανότητας τότε και μόνον τότε όταν $\mu(E)=1$ για κάθε εργοδικό $\mu\in\mathcal{M}_T(X)$.

Απόδειξη. Αν $\mu(E)=1$ για κάθε εργοδικό $\mu\in\mathcal{M}_T(X)$, τότε $\mu_x(E)=1$ για κάθε $x\in R$ και από το πόρισμα 3.14 έχουμε $\mu(E)=\int_R \mu_x(E)\,d\mu=1$ για κάθε $\mu\in\mathcal{M}_T(X)$.

3.16. Πόρισμα. Για κάθε εργοδικό $\mu\in\mathcal{M}_T(X)$ υπάρχει ένα αμετάβλητο σύνολο Borel $E\subset R$ με $\mu(E)=1$ ώστε $\mu=\mu_x$ για κάθε $x\in E$.

Απόδειξη. Έστω $F=\operatorname{supp}\mu$. Το F είναι κλειστό αμετάβλητο και $\mu(F)=1$. Έστω $\{f_n : n\in\mathbb{N}\}$ ένα αριθμήσιμο πυκνό υποσύνολο του $C(X)$. Για κάθε $n\in\mathbb{N}$, η $\tilde f_n : R\to\mathbb{R}$ είναι αμετάβλητη. Αφού το μ είναι εργοδικό, η $\tilde f_n$ είναι μ-σχεδόν παντού σταθερή. Δηλαδή, υπάρχει ένα αμετάβλητο σύνολο Borel $E_n\subset F\cap R$ με $\mu(E_n)=1$ ώστε η $\tilde f_n|E_n$ να είναι σταθερή. Το $E=\bigcap_{n=1}^{\infty}E_n$ είναι αμετάβλητο σύνολο Borel και $\mu(E)=1$. Έστω τώρα $f\in C(X)$ και $\varepsilon>0$. Υπάρχει $n\in\mathbb{N}$ με $\|f-f_n\|<\frac{\varepsilon}{2}$. Για κάθε $x,y\in E$ έχουμε

$$|\tilde f(x)-\tilde f(y)|\le|\tilde f(x)-\tilde f_n(x)|+|\tilde f_n(y)-\tilde f(y)|<\varepsilon .$$

Αυτό δείχνει ότι η $\tilde f|E$ είναι σταθερή για κάθε $f\in C(X)$. Αφού το μ είναι εργοδικό για κάθε $x\in E$ και $f\in C(X)$ έχουμε

$$\int_X f\,d\mu_x=\tilde f(x)=\int_X f\,d\mu \qquad \text{ο.ε.δ.}$$

3.17. Παράδειγμα. Έστω $T:[0,1]\to[0,1]$ με τύπο $T(x)=\frac{1}{2}(x+x^2)$. Για κάθε $0\le x<1$ έχουμε $\lim_{n\to\infty}T^n(x)=0$ ενώ $T(0)=0, T(1)=1$. Συνεπώς για κάθε $f\in C[0,1]$ έχουμε

$$0\le \tilde f(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(T^k(x))=f(0)\ ,\quad 0\le x<1$$

και προφανώς $\tilde f(1)=f(1)$. Άρα $Q=[0,1]$ και $\mu_x=\delta_0$, για $0\le x<1$, $\mu_1=\delta_1$. Έτσι $V=[0,1]$ και $R=\{0,1\}$. Από το πόρισμα 3.16 προκύπτει ότι τα δ_0,δ_1 είναι τα μοναδικά εργοδικά στοιχεία του $\mathcal{M}_T([0,1])$. Συνεπώς το $\mathcal{M}_T([0,1])$ είναι το ευθύγραμμο τμήμα στον $C[0,1]$ με άκρα τα δ_0,δ_1.

III. ΠΙΘΑΝΟΘΕΩΡΗΤΙΚΗ ΕΝΤΡΟΠΙΑ

0. ΔΕΣΜΕΥΜΕΝΗ ΠΙΘΑΝΟΤΗΤΑ κ' ΜΕΣΗ ΤΙΜΗ

0.1 ΘΕΩΡΗΜΑ - ΟΡΙΣΜΟΣ: Εστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $f\in L^1(X,\mathcal{A},\mu)$. Εστω $\mathcal{G}$ μια σ-υποάλγεβρα της $\mathcal{A}$. Τοτε υπάρχει μια συνάρτηση $E(f|\mathcal{G})$ με τις εξής ιδιότητες: (i) $E(f|\mathcal{G})$ είναι $\mathcal{G}$-μετρήσιμη.

(ii) $E(f|\mathcal{G})\in L^1(\mu)$.

(iii) $\int_G E(f|\mathcal{G})\,d\mu = \int_G f\,d\mu \quad \forall G\in\mathcal{G}$.

Η $E(f|\mathcal{G})$ είναι η δεσμευμένη μέση τιμή της f δεδομένης της σ-αλγεβρας $\mathcal{G}$ και είναι μοναδικά ορισμένη, με την εξής έννοια: αν g είναι μια οποιαδήποτε άλλη συνάρτηση που ικανοποιεί τις (i)-(iii) τότε $g=E(f|\mathcal{G})$, μ-σ.π.

Απόδειξη: Εστω κατ' αρχην οτι η f παίρνει τιμές στο $[0,+\infty]$ και $f\not\equiv 0$. Τότε η σχέση $\mu_f(G)=\int_G f\,d\mu$, $G\in\mathcal{G}$, ορίζει ένα μέτρο πάνω στην $\mathcal{G}$, το οποίο είναι απολύτως συνεχές ως προς το $\mu|_{\mathcal{G}}$ (το μέτρο μ περιορισμένο πάνω στην $\mathcal{G}$ μόνο). Απο το θεώρημα Radon-Nikodym υπάρχει μια συνάρτηση, την οποία συμβολίζουμε με $E(f|\mathcal{G})$, με $E(f|\mathcal{G})\geq 0$ και τέτοια ώστε να ισχύουν οι (i)-(iii).

Τώρα αν $f\equiv 0$ ορίζουμε $E(f|\mathcal{G})\equiv 0$. Αν η f παίρνει τιμές στο $[-\infty,+\infty]$ τοτε γράφουμε $f=f^+-f^-$ και ορίζουμε $E(f|\mathcal{G})=E(f^+|\mathcal{G})-E(f^-|\mathcal{G})$, ενώ αν η f παίρνει τιμές στο $\mathbb{C}$ γράφουμε $f=f_1+if_2$ και ορίζουμε $E(f|\mathcal{G}) = E(f_1|\mathcal{G})+iE(f_2|\mathcal{G})$.

Τέλος εστω μια συνάρτηση g που ικανοποιεί τις (i)-(iii). Τοτε τα σύνολα $A=\{x\in X : E(f|\mathcal{G})(x)<g(x)\}$ και $B=\{x\in X : E(f|\mathcal{G})(x)>g(x)\}$ ανήκουν στην $\mathcal{G}$. Αρα

$$\int_X |g-E(f|\mathcal{G})|\,d\mu = \int_A (g-E(f|\mathcal{G}))\,d\mu + \int_B (E(f|\mathcal{G})-g)\,d\mu =$$

(λόγω της (iii)) $= \int_A f\,d\mu - \int_A f\,d\mu + \int_B f\,d\mu - \int_B f\,d\mu = 0$

και αρα $g=E(f|\mathcal{G})$, μ-σ.π. ■

ΠΑΡΑΤΗΡΗΣΕΙΣ: (i) Από την ιδιότητα (iii) του θεωρήματος $\int_X E(f|\mathcal{G})\,d\mu=\int_X f\,d\mu$.

(ii) Αν πάρουμε για $\mathcal{G}$ την τετριμμένη σ-άλγεβρα $\mathcal{G}=\{\emptyset, X\}$ τότε $E(f|\mathcal{G})=\int_X f\,d\mu$.

(iii) Η παραπάνω απόδειξη δείχνει επίσης ότι αν $f\geq 0$ τότε $E(f|\mathcal{G})\geq 0$ αφού $E(f|\mathcal{G})=\frac{d\mu_f}{d\mu|_{\mathcal{G}}}$

ΣΥΜΒΟΛΙΣΜΟΣ: Αν $A\in\mathcal{A}$ και $f=\mathbb{1}_A$ τότε γράφουμε $\mu(A|\mathcal{G})=E(\mathbb{1}_A|\mathcal{G})$ και η $\mu(A|\mathcal{G})$ είναι η δεσμευμένη πιθανότητα του A δεδομένης της $\mathcal{G}$.

0.2 ΠΡΟΤΑΣΗ: Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας, $\{B_1,\dots,B_n\}$ μια πεπερασμένη διαμέριση του X, με $B_1,\dots,B_n\in\mathcal{A}$ και θέτουμε $\mathcal{G}=\sigma(B_1,\dots,B_n)$, η σ-άλγεβρα που παράγουν τα $B_1,\dots,B_n$. Τότε για κάθε $A\in\mathcal{A}$, η δεσμευμένη πιθανότητα $\mu(A|\mathcal{G})$ είναι σταθερή πάνω σε κάθε B_i με $\mu(B_i)>0$ και η τιμή της είναι

$$\mu(A|\mathcal{G})(x)=\frac{\mu(A\cap B_i)}{\mu(B_i)}\quad\text{για } \mu\text{-σχεδόν κάθε } x\in B_i.$$

Απόδειξη: Απλώς επαληθεύουμε ότι η συνάρτηση $g(x)=\frac{\mu(A\cap B_i)}{\mu(B_i)}$, για $x\in B_i$ με $\mu(B_i)>0$, επαληθεύει τα (i)–(iii) του Θεωρήματος 0.1. ■

Η δεσμευμένη μέση τιμή ικανοποιεί τις συνήθεις ιδιότητες μέσων τιμών. Μεταξύ άλλων $E(af+bg)=aE(f|\mathcal{G})+bE(g|\mathcal{G})$ (για σταθερές a,b) μ-σ.π., δηλαδή ο τελεστής $E(\cdot|\mathcal{G}):L^1(X,\mathcal{A},\mu)\to L^1(X,\mathcal{G},\mu)$ είναι γραμμικός. Για την απόδειξη απλώς επαληθεύουμε ότι το δεξί μέλος ικανοποιεί τις (i)–(iii) του ορισμού: αν f και g είναι L^1 τότε και η $aE(f|\mathcal{G})+bE(f|\mathcal{G})$ είναι L^1 (αφού κάθε μια από τις $E(f|\mathcal{G})$ και $E(g|\mathcal{G})$ είναι) και επίσης η $aE(f|\mathcal{G})+bE(g|\mathcal{G})$ είναι $\mathcal{G}$-μετρήσιμη σαν άθροισμα $\mathcal{G}$-μετρήσιμων συναρτήσεων. Τέλος

$$\int_G\left[aE(f|\mathcal{G})+bE(g|\mathcal{G})\right]d\mu=a\int_G E(f|\mathcal{G})\,d\mu+b\int_G E(g|\mathcal{G})\,d\mu=$$
$$=a\int_G f\,d\mu+b\int_G g\,d\mu=\int_G(af+bg)\,d\mu$$

για κάθε $G\in\mathcal{G}$.

Επιπλέον ισχύουν και οι εξής ιδιότητες:

(1) Αν η f είναι $\mathcal{G}$-μετρήσιμη $E(fg|\mathcal{G})=fE(g|\mathcal{G})$ (υποθέτουμε ότι $f,g\in L^2$)

(2) Αν $\mathcal{G}_1\subseteq\mathcal{G}_2$ τότε $E(E(f|\mathcal{G}_2)|\mathcal{G}_1)=E(f|\mathcal{G}_1)$.

(Η $E(E(f|\mathcal{G}_1)|\mathcal{G}_2)=E(f|\mathcal{G}_1)$ είναι προφανής για $\mathcal{G}_1\subseteq\mathcal{G}_2$ λόγω της ιδιότητας (

(3) Αν η $\varphi:\mathbb{R}\to\mathbb{R}$ είναι κυρτή τότε $E(\varphi\circ f|\mathcal{G})\leq\varphi(E(f|\mathcal{G}))$ (για $f:X\to\mathbb{R}$)

Αυτή είναι η ανισότητα Jensen για δεσμευμένες μέσες τιμές.

Η απόδειξη της (1) έχει ως εξής. Κατ' αρχήν η $f E(g|\mathcal{G})$ είναι L^1 και $\mathcal{G}$-μετρήσιμη. Άρα αρκεί να αποδείξουμε την ιδιότητα (iii) του Θεωρήματος-Ορισμού. Αν $f = 1_G$ για κάποιο $G \in \mathcal{G}$ τότε για κάθε $F \in \mathcal{G}$

$$\int_F f E(g|\mathcal{G}) d\mu = \int_{F \cap G} E(g|\mathcal{G})\, d\mu = \int_{F\cap G} g\, d\mu = \int_F fg\, d\mu .$$

Άρα η (iii) ισχύει όταν $f = 1_G$ για $G \in \mathcal{G}$. Από γραμμικότητα παίρνουμε την (iii) για f πεπερασμένο γραμμικό συνδυασμό χαρακτηριστικών συναρτήσεων $\sum_i c_i 1_{G_i}$, με κάθε $G_i \in \mathcal{G}$ και προχωράμε κατά τα συνήθη.

Για την (2) έχουμε τα εξής. Κατ' αρχήν η $E(f|\mathcal{G}_1)$ είναι εξ' ορισμού $\mathcal{G}_1$-μετρήσιμη και L^1. Τώρα αν $G \in \mathcal{G}_1$ τότε αναγκαστικά και $G \in \mathcal{G}_2$ και έτσι

$$\int_G E(f|\mathcal{G}_2)\, d\mu = \int_X 1_G E(f|\mathcal{G}_2)\, d\mu = \int_X E(f 1_G|\mathcal{G}_2)\, d\mu =$$

$$= \int_X 1_G f\, d\mu = \int_G f\, d\mu = \int_G E(f|\mathcal{G}_1)\, d\mu .$$

Άρα $E(E(f|\mathcal{G}_2)|\mathcal{G}_1) = E(f|\mathcal{G}_1)$, μ-σ.π.

Αποδείξεις της (3) υπάρχουν σε βιβλία Θεωρίας Πιθανοτήτων. Σημειώνουμε επίσης χωρίς απόδειξη ότι τα θεωρήματα μονότονης και κυριαρχημένης σύγκλισης ισχύουν και για δεσμευμένες μέσες τιμές. Τέλος έχουμε την εξής χρήσιμη πρόταση.

0.3 ΠΡΟΤΑΣΗ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, $\mathcal{G}$ μια σ-υποάλγεβρα της $\mathcal{A}$ και $P: L^2(X, \mathcal{A}, \mu) \to L^2(X, \mathcal{G}, \mu)$ η ορθογώνια προβολή του $L^2(X, \mathcal{A}, \mu)$ επί του $L^2(X, \mathcal{G}, \mu)$. Τότε ο περιορισμός του γραμμικού τελεστή $E(\cdot|\mathcal{G})$ στο $L^2(X, \mathcal{A}, \mu)$ ταυτίζεται με τον P, δηλαδή $E(f|\mathcal{G}) = P(f)$ $\forall f \in L^2(X, \mathcal{A}, \mu)$.

Απόδειξη: Κατ' αρχήν αν $f \in L^2(X, \mathcal{A}, \mu)$ τότε $E(f|\mathcal{G}) \in L^2(X, \mathcal{A}, \mu)$ και αφού η $E(f|\mathcal{G})$ είναι και $\mathcal{G}$-μετρήσιμη έχουμε ότι $E(f|\mathcal{G}) \in L^2(X, \mathcal{G}, \mu)$. Τώρα αφού $P(f) \in L^2(X, \mathcal{G}, \mu)$ η $P(f)$ είναι αναγκαστικά και L^1 και φυσικά $\mathcal{G}$-μετρήσιμη. Επίσης

$$\int_G P(f)\, d\mu = \int_X 1_G P(f)\, d\mu = (P(f), 1_G) = (f, P(1_G)) = (f, 1_G) = \int_G f\, d\mu$$

που αποδεικνύει το (iii) του Θεωρήματος 0.1. Άρα $P(f) = E(f|\mathcal{G})$.

1. ΕΝΤΡΟΠΙΑ ΔΙΑΜΕΡΙΣΗΣ

Εστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας. Με τον όρο πεπερασμένη μετρήσιμη διαμέριση (πμδ) του X θα εννοούμε μια πεπερασμένη διαμέριση $\{A_1,\dots,A_k\}$ του X, όπου $A_i\in\mathcal{A}$ για κάθε i. Τέτοιες διαμερίσεις θα τις συμβολίζουμε παρακάτω με μικρά ελληνικά γράμματα.

Αν $\mathcal{C}$ είναι μια οποιαδήποτε οικογένεια υποσυνόλων του X τότε $\sigma(\mathcal{C})$ θα συμβολίζει την σ-άλγεβρα που παράγει η $\mathcal{C}$, δηλαδή την μικρότερη σ-άλγεβρα που περιέχει την οικογένεια $\mathcal{C}$.

Αν α είναι μια πμδ του X τότε η $\sigma(\alpha)$ είναι μια πεπερασμένη σ-υποάλγεβρα της $\mathcal{A}$. Αντίστροφα, έστω $\mathcal{C}$ μια οποιαδήποτε πεπερασμένη υποάλγεβρα της $\mathcal{A}$ (όχι αναγκαστικά σ-άλγεβρα). Αν $\mathcal{C}=\{C_1,\dots,C_n\}$ τότε η οικογένεια $\{A_1\cap\dots\cap A_n : A_i = C_i \text{ ή } A_i = X\setminus C_i\}$ αποτελεί μια π.μ.δ. του X. Επομένως, υπάρχει μια αμφιμονοσήμαντη αντιστοιχία μεταξύ πεπερασμένων μετρήσιμων διαμερίσεων του X και πεπερασμένων υποαλγεβρών της $\mathcal{A}$.

Εστω α και β πμδ του X. Το join των α και β ορίζεται σαν

$$\alpha\vee\beta=\{A\cap B : A\in\alpha,\ B\in\beta\}$$

και είναι επίσης μια πμδ του X. Τον συμβολισμό "$\vee$" θα τον χρησιμοποιούμε και ως εξής. Αν $\mathcal{C}$ και $\mathcal{D}$ είναι σ-υποάλγεβρες της $\mathcal{A}$ (όχι κατ' ανάγκη πεπερασμένες) τότε $\mathcal{C}\vee\mathcal{D}=\sigma(\mathcal{C},\mathcal{D})$ θα συμβολίζει την μικρότερη σ-άλγεβρα που περιέχει τις $\mathcal{C}$ και $\mathcal{D}$. Αν α είναι μια πμδ του X και $\mathcal{C}$ μια σ-υποάλγεβρα της $\mathcal{A}$ τότε πάλι $\alpha\vee\mathcal{C}=\sigma(\alpha,\mathcal{C})$ θα είναι η μικρότερη σ-άλγεβρα που περιέχει τις α και $\mathcal{C}$.

Εστω α και β πμδ του X. Τότε $\alpha\le\beta$ σημαίνει ότι κάθε μέλος της α είναι ένωση στοιχείων της β (δηλαδή η β είναι λεπτότερη της α).

Τέλος αν $\mathcal{C}$ και $\mathcal{D}$ είναι σ-υποάλγεβρες της $\mathcal{A}$ τότε ο συμβολισμός $\mathcal{C}\overset{\circ}{\subseteq}\mathcal{D}$ θα σημαίνει ότι για κάθε $C\in\mathcal{C}$ υπάρχει $D\in\mathcal{D}$ τέτοιο ώστε $\mu(C\triangle D)=0$. $\mathcal{C}\overset{\circ}{=}\mathcal{D}$ θα σημαίνει ότι $\mathcal{C}\overset{\circ}{\subseteq}\mathcal{D}$ και $\mathcal{D}\overset{\circ}{\subseteq}\mathcal{C}$. Για π.μ.δ. α και β, $\alpha\overset{\circ}{=}\beta$ θα σημαίνει $\sigma(\alpha)\overset{\circ}{=}\sigma(\beta)$.

Εστω α μια π.μ.δ. του X και $\mathcal{F}$ μια σ-υποάλγεβρα της $\mathcal{A}$. Ορίζουμε την δεσμευμένη πληροφορία της διαμέρισης α δεδομένης της σ-άλγεβρας $\mathcal{F}$ σαν

$$I(\alpha|\mathcal{F}) = -\sum_{A\in\alpha} \mathbb{1}_A \log \mu(A|\mathcal{F}).$$

Η $I(\alpha|\mathcal{F})$ δεν είναι κατ' ανάγκην $\mathcal{F}$-μετρήσιμη συνάρτηση. Στην ειδική περίπτωση που $\mathcal{F}=\sigma(\beta)$, όπου β είναι μια π.μ.δ. του X γράφουμε $I(\alpha|\beta)$ αντί $I(\alpha|\sigma($ και έχουμε από την Πρόταση 0.2

$$I(\alpha|\beta) = -\sum_{\substack{A\in\alpha\\ B\in\beta,\,\mu(B)>0}} \mathbb{1}_{A\cap B} \log \frac{\mu(A\cap B)}{\mu(B)}$$

Τέλος η (αδέσμευτη) πληροφορία της διαμέρισης α ορίζεται σαν

$$I(\alpha) = -\sum_{A\in\alpha} \mu(A) \log \mu(A).$$

Προφανώς, $I(\alpha) = I(\alpha|\{\emptyset, X\})$, όπου $\{\emptyset, X\}$ η τετριμμένη σ-άλγεβρα.

Η εντροπία μιας διαμέρισης είναι η μέση τιμή της πληροφορίας:

$$\begin{aligned} H(\alpha|\mathcal{F}) &= \int_X I(\alpha|\mathcal{F})\,d\mu \\ &= -\sum_{A\in\alpha} \int_A \log \mu(A|\mathcal{F})\,d\mu \\ &= -\sum_{A\in\alpha} \int_X \mu(A|\mathcal{F}) \log \mu(A|\mathcal{F})\,d\mu . \end{aligned}$$

Αν $\mathcal{F}=\sigma(\beta)$, όπου β π.μ.δ. του X, τότε γράφουμε πάλι $H(\alpha|\beta)$ αντί $H(\alpha|\sigma(\beta))$.

$$H(\alpha|\beta) = \int_X I(\alpha|\beta)\,d\mu = \sum_{\substack{A\in\alpha\\ B\in\beta,\,\mu(B)>0}} \mu(A\cap B) \log \frac{\mu(A\cap B)}{\mu(B)}.$$

Τέλος

$$H(\alpha) = \int_X I(\alpha)\,d\mu = -\sum_{A\in\alpha} \mu(A) \log \mu(A)$$

και προφανώς πάλι $H(\alpha) = H(\alpha|\{\emptyset, X\})$.

ΠΑΡΑΤΗΡΗΣΕΙΣ : (i) Σε όλο αυτό το κεφάλαιο $0.\log 0$ θα είναι 0.

(ii) Από τον ορισμό της εντροπίας παίρνουμε άμεσα τα εξής.

$H(\alpha|\mathcal{F}) \geq 0$. Άρα και $H(\alpha) \geq 0$, $H(\alpha|\beta) \geq 0$.

Αν $\alpha \overset{\circ}{=} \beta$ τότε $H(\alpha|\mathcal{F}) = H(\beta|\mathcal{F})$. Άρα και $H(\alpha|\gamma) = H(\beta|\gamma)$ και $H(\alpha) = H$

Αν $\mathcal{F} \overset{\circ}{=} \mathcal{G}$ τότε $H(\alpha|\mathcal{G}) = H(\alpha|\mathcal{F})$.

1.1 ΠΡΟΤΑΣΗ (ΙΔΙΟΤΗΤΕΣ ΔΕΣΜΕΥΜΕΝΗΣ ΕΝΤΡΟΠΙΑΣ)

Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και α, β, ξ, η π.μ.δ. του X. Τότε:

(i) $H(\alpha \vee \beta \mid \xi) = H(\alpha \mid \xi) + H(\beta \mid \alpha \vee \xi)$.

(ii) $\alpha \leq \beta \Rightarrow H(\alpha \mid \xi) \leq H(\beta \mid \xi)$.

(iii) $\xi \leq \eta \Rightarrow H(\alpha \mid \xi) \geq H(\alpha \mid \eta)$

(iv) $H(\alpha \vee \beta \mid \xi) \leq H(\alpha \mid \xi) + H(\beta \mid \xi)$.

(v) $H(\alpha \mid \xi) = H(\alpha)$ ανν α κ' ξ είναι ανεξάρτητες

(δηλ.: $\mu(A \cap B) = \mu(A)\,\mu(B) \quad \forall\, A \in \alpha,\ B \in \xi$).

(vi) $H(\alpha \mid \xi) = 0$ ανν $\sigma(\alpha) \overset{\cdot}{\subseteq} \sigma(\xi)$ (δηλ. $\forall A \in \alpha,\ B \in \xi,\ \mu(A \cap B) \in \{0, \mu(B)\}$).

(vii) Αν $T: X \to X$ είναι ένας ενδομορφισμός τότε $H(T^{-1}\alpha \mid T^{-1}\xi) = H(\alpha \mid \xi)$.

Απόδειξη: Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι κάθε στοιχείο των διαμερίσεων α, β, ξ, η έχει θετικό μέτρο. Αν, για παράδειγμα $\alpha = \{A_1, \dots, A_n\}$ και $\mu(A_1) > 0, \dots, \mu(A_k) > 0$, $\mu(A_{k+1}) = \dots = \mu(A_n) = 0$ ορίζουμε την π.μ.δ. $\tilde{\alpha} = \{A_1, \dots, A_{k-1}, A_k \cup A_{k+1} \cup \dots \cup A_n\}$ του X και επειδή $\alpha \doteq \tilde{\alpha}$ μπορούμε να χρησιμοποιήσουμε την προηγούμενη Παρατήρηση.

(i)
$$H(\alpha \vee \beta \mid \xi) = - \sum_{A \in \alpha,\, B \in \beta,\, C \in \xi} \mu(A \cap B \cap C) \log \frac{\mu(A \cap B \cap C)}{\mu(C)} .$$

Όμως $\dfrac{\mu(A \cap B \cap C)}{\mu(C)} \quad \dfrac{\mu(A \cap B \cap C)}{\mu(A \cap C)} \ \dfrac{\mu(A \cap C)}{\mu(C)}$ εκτός και αν $\mu(A \cap C) = 0$, οπότε όμως και το αριστερό μέλος είναι 0. Άρα

$$H(\alpha \vee \beta \mid \xi) = - \sum_{\substack{A \in \alpha,\, B \in \beta,\, C \in \xi \\ \mu(A \cap C) > 0}} \mu(A \cap B \cap C) \log \frac{\mu(A \cap C)}{\mu(C)} -$$

$$- \sum_{\substack{A \in \alpha\ B \in \beta,\, C \in \xi \\ \mu(A \cap C) > 0}} \mu(A \cap B \cap C) \log \frac{\mu(A \cap B \cap C)}{\mu(A \cap C)} =$$

$$= - \sum_{A \in \alpha,\, C \in \xi} \mu(A \cap C) \log \frac{\mu(A \cap C)}{\mu(C)} -$$

$$- \sum_{\substack{B \in \beta,\ D \in \alpha \vee \xi \\ \mu(D) > 0}} \mu(B \cap D) \log \frac{\mu(B \cap D)}{\mu(D)} =$$

$$= H(\alpha \mid \xi) + H(\beta \mid \alpha \vee \xi).$$

(ii) $\alpha \leq \beta \Rightarrow H(\beta|\xi) = H(\alpha\vee\beta|\xi) \overset{(i)}{=} H(\alpha|\xi) + H(\beta|\alpha\vee\xi) \geq H(\alpha|\xi)$

αφού από την Παρατήρηση παραπάνω $H(\beta|\alpha\vee\xi) \geq 0$.

(iii) Η συνάρτηση $\varphi : [0,1] \to [-e^{-1}, 0]$ με τύπο $\varphi(x) = x\log x$ $\quad (\varphi(0)=0)$

είναι γνησίως κυρτή. Άρα για οποιαδήποτε $A\in\alpha$, $C\in\xi$

$$\varphi\left(\sum_{D\in\eta}\frac{\mu(D\cap C)}{\mu(C)}\frac{\mu(A\cap D)}{\mu(D)}\right) \leq \sum_{D\in\eta}\frac{\mu(D\cap C)}{\mu(C)}\,\varphi\left(\frac{\mu(A\cap D)}{\mu(D)}\right).$$

Όμως αφού $\xi\leq\eta$ τα $C\in\xi$ είναι ακριβώς ένωση στοιχείων της η και άρα

$$\sum_{D\in\eta}\frac{\mu(D\cap C)}{\mu(C)}\frac{\mu(A\cap D)}{\mu(D)} = \frac{\mu(A\cap C)}{\mu(C)}$$

Άρα

$$\frac{\mu(A\cap C)}{\mu(C)}\log\frac{\mu(A\cap C)}{\mu(C)} = \varphi\left(\frac{\mu(A\cap C)}{\mu(C)}\right) =$$

$$= \varphi\left(\sum_{D\in\eta}\frac{\mu(D\cap C)}{\mu(C)}\frac{\mu(A\cap D)}{\mu(D)}\right) \leq \sum_{D\in\eta}\frac{\mu(D\cap C)}{\mu(C)}\frac{\mu(A\cap D)}{\mu(D)}\log\frac{\mu(A\cap D)}{\mu(D)}$$

Πολλαπλασιάζοντας και τα δύο μέλη με $\mu(C)$ παίρνουμε

$$\mu(A\cap C)\log\frac{\mu(A\cap C)}{\mu(C)} \leq \sum_{D\in\eta}\mu(D\cap C)\frac{\mu(A\cap D)}{\mu(D)}\log\frac{\mu(A\cap D)}{\mu(D)}$$

και αθροίζοντας ως προς $C\in\xi$

$$\sum_{C\in\xi}\mu(A\cap C)\log\frac{\mu(A\cap C)}{\mu(C)} \leq \sum_{C\in\xi}\sum_{D\in\eta}\mu(D\cap C)\frac{\mu(A\cap D)}{\mu(D)}\log\frac{\mu(A\cap D)}{\mu(D)}$$

$$= \sum_{D\in\eta}\left(\sum_{C\in\xi}\mu(D\cap C)\right)\frac{\mu(A\cap D)}{\mu(D)}\log\frac{\mu(A\cap D)}{\mu(D)}$$

$$= \sum_{D\in\eta}\mu(A\cap D)\log\frac{\mu(A\cap D)}{\mu(D)}.$$

Τέλος αθροίζοντας ως προς $A\in\alpha$ παίρνουμε ότι

$$-H(\alpha|\xi) \leq -H(\alpha|\eta).$$

(iv) Από τις (i) και (iii), επειδή $\xi \leq \alpha \vee \xi$,

$$H(\alpha \vee \beta \mid \xi) = H(\alpha \mid \xi) + H(\beta \mid \alpha \vee \xi) \leq H(\alpha \mid \xi) + H(\beta \mid \xi).$$

(v) Έστω ότι οι α και ξ είναι ανεξάρτητες. Τότε $\mu(A \cap B) = \mu(A)\mu(B)$ $\forall A \in \alpha, B \in \xi$.

Επομένως

$$H(\alpha \mid \xi) = -\sum_{A \in \alpha} \sum_{B \in \xi} \mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)} = -\sum_{A \in \alpha} \mu(A) \log \mu(A) = H(\alpha).$$

Αντίστροφα έστω ότι

$$(*) \qquad -\sum_{A \in \alpha} \sum_{B \in \xi} \mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)} = -\sum_{A \in \alpha} \mu(A) \log \mu(A).$$

Η συνάρτηση $\phi(x) = x \log x$ είναι γνήσια κυρτή. Άρα για κάθε (σταθερό) $A \in \alpha$

$$(**) \qquad -\sum_{B \in \xi} \mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)} = -\sum_{B \in \xi} \mu(B) \phi\left(\frac{\mu(A \cap B)}{\mu(B)} \right) \leq$$

$$\leq -\phi\left(\sum_{B \in \xi} \frac{\mu(A \cap B)}{\mu(B)} \mu(B) \right) = -\phi(\mu(A)) = -\mu(A) \log \mu(A),$$

με ισότητα αν και μόνο αν η τιμή του $\mu(A \mid B) := \dfrac{\mu(A \cap B)}{\mu(B)}$ είναι ανεξάρτητη του B. Στην περίπτωση αυτή η κοινή τιμή των $\mu(A \mid B)$ πρέπει αναγκαστικά να είναι η $\mu(A)$ (γιατί: αν a είναι η κοινή τιμή των $\mu(A \mid B)$ τότε $\mu(A \cap B) = a\mu(B)$ για κάθε $B \in \xi$ και αθροίζοντας ως προς $B \in \xi$ βλέπουμε ότι $a = \mu$). Τώρα από την $(*)$ βλέπουμε ότι η $(**)$ πρέπει να ισχύει με ισότητα για κάθε $A \in \alpha$. Άρα πρέπει $\mu(A \cap B) = \mu(A)\mu(B)$ $\forall A \in \alpha, B \in \xi$. Επί. η α πρέπει να είναι ανεξάρτητη της ξ.

(vi) $\sigma(\alpha) \overset{\circ}{\subseteq} \sigma(\xi)$ σημαίνει ότι $\forall A \in \sigma(\alpha)$ $\exists B \in \sigma(\xi)$ τ.ω. $\mu(A \triangle B) = 0$. Επομένως, αν $A \in \alpha$ και $B \in \xi$ τότε ή $\mu(A \cap B) = \mu(B)$ ή $\mu(A \cap B) = 0$. Σε κάθε περίπτωση

$$H(\alpha \mid \xi) = -\sum_{A \in \alpha} \sum_{B \in \xi} \mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)} =$$

$$= -\sum_{A \in \alpha} \left(\sum_{\substack{B \in \xi \\ \mu(A \cap B) = \mu(B)}} \mu(B) \log 1 + \sum_{\substack{B \in \xi \\ \mu(A \cap B) = 0}} 0 \cdot \log \frac{\mu(A \cap B)}{\mu(B)} \right) = 0.$$

Αντίστροφα, αν $H(\alpha|\xi)=0$ τότε πρέπει $\mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(B)}=0$, $\forall A\in\alpha, B\in\xi$, που φυσικά είναι ότι ή $\mu(A\cap B)=0$ ή $\mu(A\cap B)=\mu(B)$ για όλα τα $A\in\alpha$, $B\in\xi$.

vii) Έστω $T: X\to X$ μια $\mathcal{A}$ μετρήσιμη απεικόνιση που διατηρεί το μέτρο μ (δηλαδή ο T είναι ένας ενδομορφισμός). Τότε

$$H(T^{-1}\alpha\,|\,T^{-1}\xi) = -\sum_{A\in\alpha}\sum_{B\in\xi}\mu(T^{-1}A\cap T^{-1}B)\log\frac{\mu(T^{-1}A\cap T^{-1}B)}{\mu(T^{-1}B)}$$

$$= -\sum_{A\in\alpha}\sum_{B\in\xi}\mu(T^{-1}(A\cap B))\log\frac{\mu(T^{-1}(A\cap B))}{\mu(T^{-1}B)}$$

$$= -\sum_{A\in\alpha}\sum_{B\in\xi}\mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(B)} = H(\alpha|\xi). \blacksquare$$

Από τις ιδιότητες της δεσμευμένης εντροπίας παίρνουμε αμέσως τις εξής ιδιότητες για την αδέσμευτη εντροπία.

1.2 ΠΡΟΤΑΣΗ (ΙΔΙΟΤΗΤΕΣ ΑΔΕΣΜΕΥΤΗΣ ΕΝΤΡΟΠΙΑΣ)

Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και α,β π.μ.δ. Τότε

(i) $H(\alpha\vee\beta) = H(\alpha) + H(\beta|\alpha)$.

(ii) $\alpha\le\beta \Rightarrow H(\alpha)\le H(\beta)$.

(iii) $H(\alpha)\ge H(\alpha|\beta)$

(iv) $H(\alpha\vee\beta)\le H(\alpha)+H(\beta)$

(v) Αν $T: X\to X$ είναι ένας ενδομορφισμός τότε $H(T^{-1}\alpha)=H(\alpha)$.

Η απόδειξη αυτών των ιδιοτήτων είναι άμεση συνέπεια των αντίστοιχων ιδιοτήτων για δεσμευμένη εντροπία και του ότι $H(\alpha\,|\,\{\emptyset,X\}) = H(\alpha)$.

Στην Πρόταση 1.1 είδαμε ιδιότητες της δεσμευμένης εντροπίας δεδομένης μιας πεπερασμένης σ-άλγεβρας $\mathcal{F}$. Οι ιδιότητες αυτές ισχύουν και για δεσμευμένη εντροπία δεδομένης μια αυθαίρετης σ-άλγεβρας $\mathcal{F}$.

1.3 ΠΡΟΤΑΣΗ (ΙΔΙΟΤΗΤΕΣ ΔΕΣΜΕΥΜΕΝΗΣ ΕΝΤΡΟΠΙΑΣ)

Εστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, α και β π.μ.δ. του X και $\mathcal{F}, \mathcal{G}$ σ-υποάλγεβρες της $\mathcal{F}$. Τότε

(i) $H(\alpha \vee \beta \mid \mathcal{F}) = H(\alpha \mid \mathcal{F}) + H(\beta \mid \alpha \vee \mathcal{F})$.

(ii) $\alpha \leq \beta \Rightarrow H(\alpha \mid \mathcal{F}) \leq H(\beta \mid \mathcal{F})$.

(iii) $\mathcal{F} \subseteq \mathcal{G} \Rightarrow H(\alpha \mid \mathcal{F}) \geq H(\alpha \mid \mathcal{G})$.

Αρα και $H(\alpha) \geq H(\alpha \mid \mathcal{F})$ (αφού $H(\alpha) = H(\alpha \mid \{\emptyset, X\})$).

(iv) $H(\alpha \vee \beta \mid \mathcal{F}) \leq H(\alpha \mid \mathcal{F}) + H(\beta \mid \mathcal{F})$

(v) $H(\alpha \mid \mathcal{F}) = H(\alpha)$ ανν η α είναι ανεξάρτητη της $\mathcal{F}$.

(vi) $H(\alpha \mid \mathcal{F}) = 0$ ανν $\sigma(\alpha) \stackrel{\cdot}{\subseteq} \mathcal{F}$ (δηλ. $\forall A \in \alpha\ \exists F \in \mathcal{F}$ με $\mu(A \triangle F) = 0$).

Η απόδειξη μπορεί να δοθεί απ' ευθείας με χρήση ιδιοτήτων της δεσμευμένης μέσης τιμής. Ειδικά στην περίπτωση που ο χώρος πιθανότητας $(X, \mathcal{A}, \mu)$ έχει αριθμήσιμη βάση τότε οι ιδιότητες (i)-(iv) παραπάνω συνάγονται και σαν όρισμα των αντίστοιχων ιδιοτήτων της Πρότασης 1.1 και της Πρότασης 1.4 παρακάτω. (Βλέπε και Παρατηρήσεις μετά το τέλος της απόδειξης της Πρότασης 1.4.)

Εμείς εδώ θα αποδείξουμε μόνο τα (v) και (vi) της Πρότασης 1.3.

Απόδειξη της (v): Εστω ότι η α είναι ανεξάρτητη της $\mathcal{F}$. Τότε για κάθε A: $\mu(A \mid \mathcal{F}) = \mu(A)$, μ-σ.π. και άρα

$$H(\alpha \mid \mathcal{F}) = -\sum_{A \in \alpha} \int \mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F})\, d\mu = -\sum_{A \in \alpha} \mu(A) \log \mu(A) = H(\alpha).$$

Αντίστροφα, έστω ότι $H(\alpha) = H(\alpha \mid \mathcal{F})$. Εστω $B \in \mathcal{F}$. Τότε η $\sigma(B) = \{X, \emptyset, B, B^c\}$ περιέχεται στην $\mathcal{F}$ και άρα από την ιδιότητα (iii) (της οποίας η απόδειξη είναι ανεξάρτητη της παρούσας ιδιότητας) $H(\alpha) = H(\alpha \mid \{\emptyset, X\}) \geq H(\alpha \mid \sigma(B)) \geq H(\alpha \mid \mathcal{F}) = H(\alpha)$ από την υπόθεσή μας. Αρα $H(\alpha) = H(\alpha \mid \sigma(B))$. Η $\sigma(B)$ όμως παράγεται από την π.μ.δ. $\{B, B^c\}$ και άρα από την Πρόταση 1.1 (v) η α πρέπει να είναι ανεξάρτητη της $\{B, B^c\}$, δηλαδή $\mu(A \cap B) = \mu(A)\mu(B)$ $\forall A \in \alpha$. Αφού το B ήταν ένα αυθαίρετο στοιχείο της $\mathcal{F}$ έπεται ότι $\mu(A \cap B) = \mu(A)\mu(B)$ για κάθε $A \in \alpha$ και για κάθε $B \in \mathcal{F}$. Αυτό σημαίνει ακριβώς ότι η α είναι ανεξάρτητη της $\mathcal{F}$.

Απόδειξη της (vi): Αν $\sigma(\alpha) \subseteq \mathcal{F}$ τότε για κάθε στοιχείο A της α υπάρχει $B \in \mathcal{F}$ με $\mu(A \triangle B) = 0$. Τότε όμως $\mu(A \mid \mathcal{F}) \in \{0,1\}$ για κάθε $A \in \alpha$. (Πράγματι, έστω $B \in \mathcal{F}$ με $\mu(A \triangle B) = 0$. Τότε $\mu(A \mid \mathcal{F}) = \mathbb{1}_B$, μ-σ.π., γιατί $\int_F \mathbb{1}_B \, d\mu =$ $= \mu(B \cap F) = \mu(A \cap F) = \int_F \mu(A \mid \mathcal{F}) d\mu$, $\forall F \in \mathcal{F}$.) Άρα $\mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F}) = 0$ για όλα τα $A \in \alpha$ και έτσι

$$H(\alpha \mid \mathcal{F}) = -\sum_{A \in \alpha} \int \mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F}) \, d\mu = 0.$$

Αντίστροφα, αν ισχύει η παραπάνω σχέση, τότε επειδή $-\mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F}) \geq 0$ για όλα τα $A \in \alpha$, πρέπει $-\mu(A \mid \mathcal{F}) \log \mu(A \mid \mathcal{F}) = 0$ για κάθε $A \in \alpha$. Αυτό όμως σημαίνει ότι $\mu(A \mid \mathcal{F}) \in \{0,1\}$ δηλαδή $\mu(A \mid \mathcal{F}) = \mathbb{1}_B$, μ-σ.π., για κάποιο $B \in \mathcal{F}$. Τότε

$$\mu(A \cap B^c) = \int_{B^c} \mathbb{1}_A \, d\mu = \int_{B^c} \mu(A \mid \mathcal{F}) \, d\mu = \int_{B^c} \mathbb{1}_B \, d\mu = 0$$

και

$$\mu(A^c \cap B) = \int_B \mathbb{1}_{A^c} \, d\mu = \int_B \mu(A^c \mid \mathcal{F}) \, d\mu = \int_B [1 - \mu(A \mid \mathcal{F})] \, d\mu =$$

$$= \int_B (1 - \mathbb{1}_B) \, d\mu = \int_B \mathbb{1}_{B^c} \, d\mu = 0.$$

Άρα $\mu(A \triangle B) = 0$. Αποδείξαμε δηλαδή ότι $\forall A \in \alpha \; \exists B \in \mathcal{F}$ με $\mu(A \triangle B) = 0$. ∎

Αποδεικνύουμε τώρα μια άλλη χρήσιμη ιδιότητα της εντροπίας.

1.4 ΠΡΟΤΑΣΗ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, α μια π.μ.δ. του X και $\{\mathcal{F}_n\}$ μια ακολουθία σ-υποαλγεβρών της $\mathcal{A}$. Τότε:

(i) Αν $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots$ τότε $H(\alpha \mid \mathcal{F}_n) \to H(\alpha \mid \bigvee_{n=1}^{\infty} \mathcal{F}_n)$.

(ii) Αν $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \ldots$ τότε $H(\alpha \mid \mathcal{F}_n) \to H(\alpha \mid \bigcap_{n=1}^{\infty} \mathcal{F}_n)$.

Η πρόταση αυτή είναι άμεση συνέπεια του θεωρήματος σύγκλισης για martingales του Doob. Εμείς θα δώσουμε μια απόδειξη που στηρίζεται στο επόμενο λήμμα.

1.5 ΛΗΜΜΑ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $\{\mathcal{F}_n\}$ μια μονότονη ακολουθία σ-υποαλγεβρών της $\mathcal{A}$. Αν $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots$ θέτουμε $\mathcal{F} = \bigvee_{n=1}^{\infty} \mathcal{F}_n$ ενώ αν $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \ldots$ θέτουμε $\mathcal{F} = \bigcap_{n=1}^{\infty} \mathcal{F}_n$. Τότε για κάθε $f \in L^2(X, \mathcal{A}, \mu)$

$$\| E(f|\mathcal{F}_n) - E(f|\mathcal{F})\|_2 \to 0 .$$

Απόδειξη: Έστω πρώτα ότι $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$. Έστω $B \in \mathcal{F}$. Για κάθε $k \in \mathbb{N}$ διαλέγουμε ένα $B_k' \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$ με $\mu(B \triangle B_k') < k^{-1}$. Αυτό μπορεί να γίνει γιατί η $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ είναι μια άλγεβρα που παράγει την σ-άλγεβρα $\mathcal{F}$. Έστω ότι $B_k' \in \mathcal{F}_{n_k}$, όπου n_k ο μικρότερος ακέραιος ώστε $B_k' \in \mathcal{F}_{n_k}$. Τότε θέτω $B_1 = \dots = B_{n_1 - 1} = B'$, όπου B' ένα οποιοδήποτε στοιχείο της $\mathcal{F}_1$, $B_{n_1} = B_1'$, $B_{n_1+1} = \dots = B_{n_2-1} = B_{n_1}'$ κ.ο.κ. Παίρνουμε έτσι μια ακολουθία $\{B_n\}$ με $B_n \in \mathcal{F}_n$ και $\mu(B \triangle B_n) \to 0$. Τώρα το $\mu(B|\mathcal{F}_n) = E(1_B|\mathcal{F}_n)$ είναι το στοιχείο εκείνο του $L^2(X, \mathcal{F}_n, \mu)$ που ελαχιστοποιεί την απόσταση από το $1_B \in L^2(X, \mathcal{F}, \mu)$ (Πρόταση 0.3). Άρα

$$\| \mu(B|\mathcal{F}_n) - 1_B \|_2^2 \leq \| 1_{B_n} - 1_B \|_2^2 = \mu(B \triangle B_n) \to 0 .$$

Επειδή γραμμικοί συνδυασμοί χαρακτηριστικών συναρτήσεων είναι πυκνοί στον $L^2(X, \mathcal{F}, \mu)$ έπεται ότι $\| E(f|\mathcal{F}_n) - f \|_2 \to 0$, $\forall f \in L^2(X, \mathcal{F}, \mu)$. Τέλος αν $f \in L^2(X, \mathcal{A}, \mu)$

$$\| E(f|\mathcal{F}_n) - E(f|\mathcal{F})\|_2 = \| E(E(f|\mathcal{F})|\mathcal{F}_n) - E(f|\mathcal{F})\|_2 \to 0 .$$

Αν $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots$ θέτουμε $f_n = E(f|\mathcal{F}_n)$. Τότε

$$\int f_{n+k}(f_n - f_{n+k})\,d\mu = \int E(f|\mathcal{F}_{n+k}) E(f|\mathcal{F}_n)\,d\mu - \int E(f|\mathcal{F}_{n+k}) E(f|\mathcal{F}_{n+k})\,d\mu =$$

$$= \int E(E(f|\mathcal{F}_{n+k})f \,|\, \mathcal{F}_n)\,d\mu - \int E(E(f|\mathcal{F}_{n+k})f \,|\, \mathcal{F}_{n+k})\,d\mu =$$

$$= \int f E(f|\mathcal{F}_{n+k})\,d\mu - \int f E(f|\mathcal{F}_{n+k})\,d\mu = 0 .$$

Άρα

$$\| f_n \|_2^2 = \| f_{n+k} - f_{n+k} + f_n \|_2^2 = \| f_{n+k} - f_n \|_2^2 + \| f_{n+k} \|_2^2 .$$

Επομένως η ακολουθία $\{\|f_n\|_2\}$ είναι μια φθίνουσα ακολουθία αριθμών και άρα συγκλίνει σε έναν αριθμό $\alpha \in [0, +\infty)$. Όμως τότε $\| f_{n+k} - f_n \|_2^2 = \|f_n\|_2^2 - \|f_{n+k}\|_2^2 \to \alpha - \alpha = 0$, καθώς $n, k \to 0$. Δηλαδή η $\{f_n\}$ είναι Cauchy στον L^2 και άρα $f_n \xrightarrow{L^2} g$ για κάποια συνάρτηση $g \in L^2(X, \mathcal{A}, \mu)$. Από το θεώρημα Cauchy-Schwartz, $\| f_n 1_A - g 1_A \|_1 \leq \| f_n - g \|_2 \| 1_A \|_2 \to 0$ και άρα $\int_A f_n\,d\mu \to \int_A g\,d\mu$ για κάθε $A \in \mathcal{A}$. Αφού για $F \in \mathcal{F}$ όμως $\int_F f_n\,d\mu = \int_F f\,d\mu$ $\forall n \in \mathbb{N}$, πρέπει $g = E(f|\mathcal{F})$ (αφού η g είναι και $\mathcal{F}$-μετρήσιμη και L^1). ∎

Η απόδειξη που δώσαμε για την περίπτωση $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots$ δουλεύει με κατάλληλες τροποποιήσεις και για την περίπτωση $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$.

Απόδειξη Πρότασης 1.4: Απο το λήμμα $\mu(A|\mathcal{F}_n) \to \mu(A|\mathcal{F})$ στον L^2 και αρα και κατά μέτρο, δηλαδή

$$\mu\{x\in X : |\mu(A|\mathcal{F}_n)(x) - \mu(A|\mathcal{F})(x)| > \epsilon\} \to 0 \quad \forall \epsilon>0$$

για όλα τα $A\in\alpha$. Επομένως

$$(*) \qquad -\sum_{A\in\alpha} \mu(A|\mathcal{F}_n)\log\mu(A|\mathcal{F}_n) \longrightarrow -\sum_{A\in\alpha} \mu(A|\mathcal{F})\log\mu(A|\mathcal{F})$$

κατά μέτρο. Αφού η συνάρτηση $x \mapsto x\log x$ είναι φραγμένη στο $[0,1]$ και οι δύο πλευρές της $(*)$ φράσσονται από το φράγμα της $x\mapsto x\log x$, $x\in[0,1]$ επί τον πληθάριθμο της διαμέρισης α. Η κατά μέτρο σύγκλιση τώρα δίνει σύγκλιση στον L^1, δηλαδή

$$H(\alpha|\mathcal{F}_n) = -\int \sum_{A\in\alpha} \mu(A|\mathcal{F}_n)\log\mu(A|\mathcal{F}_n)\,d\mu \to$$
$$\to -\int \sum_{A\in\alpha} \mu(A|\mathcal{F})\log\mu(A|\mathcal{F})\,d\mu = H(\alpha|\mathcal{F}). \quad \blacksquare$$

ΠΑΡΑΤΗΡΗΣΕΙΣ: (i) Απο το θεώρημα σύγκλισης των martingales, του Doob, η σύγκλιση στο Λήμμα 1.5 είναι και σχεδόν παντού. Συγκεκριμένα το θεώρημα του Doob δίνει οτι

$$E(f|\mathcal{F}_n) \to E(f|\bigvee_{n=1}^{\infty}\mathcal{F}), \quad \mu\text{-σ.π. κ' στον } L^1$$

αν $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$ ενώ αν $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots$

$$E(f|\mathcal{F}_n) \to E(f|\bigcap_{n=1}^{\infty}\mathcal{F}_n), \quad \mu\text{-σ.π. κ' στον } L^1$$

για κάθε $f\in L^1(X,\mathcal{A},\mu)$.

(ii) Εστω οτι ο $(X,\mathcal{A},\mu)$ είναι ενας χωρος πιθανότητας που έχει αριθμήσιμη βάση, δηλαδή υπάρχει ακολουθία $\{A_n\}$ στοιχείων της $\mathcal{A}$ τέτοια ώστε: για καθε $\epsilon>0$ και για κάθε $A\in\mathcal{A}$ υπάρχει $n\in\mathbb{N}$ με $\mu(A\triangle A_n)<\epsilon$. (Αυτό είναι ισοδύναμο με το να πούμε οτι ο $L^2(X,\mathcal{A},\mu)$ είναι διαχωρίσιμος. Αν ο X είναι μετρικός χώρος και $\mathcal{A}$ είναι τα Borel υποσύνολα του X τότε ο $(X,\mathcal{A},\mu)$ έχει αριθμήσιμη βάση.) Εστω $\mathcal{F}$ μια σ-υποάλγεβρα της $\mathcal{A}$. Τότε μπορούμε να βρούμε πεπερασμένες σ-άλγεβρες $\mathcal{F}_n$, τέτοιες ώστε $\bigvee_{n=1}^{\infty}\mathcal{F}_n \doteq \mathcal{F}$. Οι πεπερασμένες σ-άλγεβρες $\mathcal{F}_n$ παράγονται από πμδ του X (βλ. αρχή αυτής της ενότητας) και άρα ισχύουν οι ιδιότητες της Πρότασης 1.1 για τις $H(\alpha|\mathcal{F}_n)$. Χρησιμοποιώντας την Πρόταση 1.4 παίρνουμε τις ιδιότητες της Πρότασης 2.3 για την $H(\alpha|\mathcal{F})$.

2. ΕΝΤΡΟΠΙΑ ΕΝΔΟΜΟΡΦΙΣΜΟΥ

2.1 ΛΗΜΜΑ: Έστω $\{a_n\}$ μια ακολουθία πραγματικών αριθμών με $a_{n+m} \leq a_n + a_m$ για όλα τα n και m. Τότε η $\{\frac{a_n}{n}\}$ έχει όριο και $\lim\limits_{n\to\infty} \frac{a_n}{n} = \inf\limits_{n} \frac{a_n}{n}$.
(Το όριο $\lim\limits_{n\to\infty}\frac{a_n}{n}$ μπορεί να είναι $-\infty$ αλλά αν η $\{a_n\}$ είναι κάτω φραγμένη τότε το όριο είναι ≥ 0.)

Απόδειξη: Σταθεροποιούμε ένα $m\in\mathbb{N}$ και για $n\in\mathbb{N}$ γράφουμε $n = mk+i$ με $0\leq i<m$. Τότε

$$\frac{a_n}{n} = \frac{a_{mk+i}}{mk+i} \leq \frac{a_{mk}}{mk+i} + \frac{a_i}{mk+i} \leq \frac{k\,a_m}{mk+i} + \frac{a_i}{mk+i}.$$

Αφήνοντας το $n\to\infty$, πρέπει και $k\to\infty$ και παίρνουμε ότι $\limsup\limits_{n\to\infty}\frac{a_n}{n} \leq \frac{a_m}{m}$.
Αφού το m ήταν αυθαίρετα επιλεγμένο έπεται ότι $\limsup\limits_{n\to\infty}\frac{a_n}{n} \leq \inf\limits_{n}\frac{a_m}{m}$ και έτσι πρέπει αναγκαστικά $\frac{a_n}{n} \to \inf \frac{a_m}{m}$ καθώς $n\to\infty$. (Προφανώς αν $a_n \geq M$ $\forall n\in\mathbb{N}$, τότε $\lim\limits_{n\to\infty}\frac{a_n}{n} \geq 0$.) ∎

2.2 ΘΕΩΡΗΜΑ – ΟΡΙΣΜΟΣ: Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $T: X\to X$ ένας ενδομορφισμός. Τότε για οποιαδήποτε π.μ.δ. α του X το όριο

$$h(T,\alpha) := \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big)$$

υπάρχει και καλείται εντροπία του T ως προς την διαμέριση α.

Απόδειξη: Έστω $a_n = H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$. Τότε

$$a_{n+m} = H\Big(\bigvee_{i=0}^{n+m-1} T^{-i}\alpha\Big) \leq H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) + H\Big(\bigvee_{i=n}^{n+m-1} T^{-i}\alpha\Big) =$$
$$= a_n + H\Big(T^{-n}\Big(\bigvee_{i=0}^{m-1} T^{-i}\alpha\Big)\Big) = a_n + H\Big(\bigvee_{i=0}^{m-1} T^{-i}\alpha\Big) = a_n + a_m.$$

Από το Λήμμα 2.1 το όριο υπάρχει. ∎

ΠΑΡΑΤΗΡΗΣΗ: $h(T,\alpha)\geq 0$.

2.3 ΠΡΟΤΑΣΗ: Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $T: X\to X$ ένας ενδομορφισμός. Τότε για οποιαδήποτε π.μ.δ. α του X η ακολουθία $n^{-1}H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$ φθίνει πρός το όριό της $h(T,\alpha)$.

Απόδειξη: Θα αποδείξουμε ότι η ακολουθία $n^{-1}H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$ είναι φθίνουσα

Από την ιδιότητα $H(\alpha \vee \beta) = H(\alpha) + H(\beta \mid \alpha)$ για π.μ.δ. α και β του X έχουμε επαγωγικά ότι

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) = H(\alpha) + \sum_{j=1}^{n-1} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{j} T^{-i}\alpha\right).$$

Πράγματι, για $n=1$ η σχέση αυτή είναι προφανής ($\sum_1^{1-1} = 0$). Έστω ότι η σχέση ισχύει για $n=p$. Τότε

$$H\left(\bigvee_{i=0}^{p} T^{-i}\alpha\right) = H\left(\alpha \vee \bigvee_{i=1}^{p} T^{-i}\alpha\right) = H\left(\bigvee_{i=1}^{p} T^{-i}\alpha\right) + H\left(\alpha \,\Big|\, \bigvee_{i=1}^{p} T^{-i}\alpha\right) =$$

$$= H\left(T^{-1}\bigvee_{i=0}^{p-1} T^{-i}\alpha\right) + H\left(\alpha \,\Big|\, \bigvee_{i=1}^{p} T^{-i}\alpha\right) =$$

$$= H\left(\bigvee_{i=0}^{p-1} T^{-i}\alpha\right) + H\left(\alpha \,\Big|\, \bigvee_{i=1}^{p} T^{-i}\alpha\right) =$$

(από υπόθεση επαγωγής) $$= H(\alpha) + \sum_{j=1}^{p-1} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{j} T^{-i}\alpha\right) + H\left(\alpha \,\Big|\, \bigvee_{i=1}^{p} T^{-i}\alpha\right).$$

Από αυτή τη σχέση τώρα παίρνουμε ότι

$$n\, H\left(\bigvee_{i=0}^{n} T^{-i}\alpha\right) = n\left[H\left(\bigvee_{i=1}^{n} T^{-i}\alpha\right) + H\left(\alpha \,\Big|\, \bigvee_{i=1}^{n} T^{-i}\alpha\right)\right] =$$

$$= n\, H\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) + n\, H\left(\alpha \,\Big|\, \bigvee_{i=1}^{n} T^{-i}\alpha\right) \leq$$

$$\leq n\, H\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) + \sum_{j=1}^{n-1} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{j} T^{-i}\alpha\right) + H(\alpha) =$$

$$= (n+1)\, H\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right). \quad \blacksquare$$

2.4 ΠΟΡΙΣΜΑ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ενδομορφισμός. Τότε για κάθε π.μ.δ. α του X,

$$h(T, \alpha) = \lim_{n\to\infty} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{n} T^{-i}\alpha\right) = H\left(\alpha \,\Big|\, \bigvee_{i=1}^{\infty} T^{-i}\alpha\right).$$

Απόδειξη: Αφού $\bigvee_{i=1}^{n} T^{-i}\alpha \leq \bigvee_{i=1}^{n+1} T^{-i}\alpha$ έπεται ότι $\sigma(\bigvee_{i=1}^{n} T^{-i}\alpha) \subseteq \sigma(\bigvee_{i=1}^{n+1} T^{-i}\alpha)$ και από την Πρόταση 1.4 το όριο $\lim_{n\to\infty} H(\alpha \mid \bigvee_{i=1}^{n} T^{-i}\alpha)$ υπάρχει και ισούται με $H(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha)$.

Τώρα από την απόδειξη της Πρότασης 2.3

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) = H(\alpha) + \sum_{j=1}^{n-1} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{j} T^{-i}\alpha\right).$$

Διαιρώντας με n και τα δύο μέλη και παίρνοντας όρια παίρνουμε ότι

$$h(T, \alpha) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n-1} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{j} T^{-i}\alpha\right) = \lim_{n\to\infty} H\left(\alpha \,\Big|\, \bigvee_{i=1}^{n} T^{-i}\alpha\right).$$

(Αν, για πραγματικούς αριθμούς, $a_n \to a$ τότε $n^{-1}\sum_{i=1}^{n} a_i \to a$.) $\blacksquare$

2.5 ΟΡΙΣΜΟΣ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Η (μετροθεωρητική) εντροπία της T (ως προς το μέτρο μ) ορίζεται σαν

$$h(T) = \sup_{\alpha} h(T,\alpha)$$

όπου το supremum το παίρνουμε ως πρός όλες τις π.μ.δ. α του X.

ΠΑΡΑΤΗΡΗΣΕΙΣ: (i) $h(T) \in [0,+\infty]$.

(ii) $h(id_X)=0$ όπου id_X η ταυτοτική απεικόνιση του X. Απο την άλλη αν $h(T)=0$ τότε για κάθε π.μ.δ. α του X πρέπει $h(T,\alpha)=0$. Δηλαδή, για οποιαδήποτε π.μ.δ. α του X η $\bigvee_{i=0}^{n-1} T^{-i}\alpha$ δεν "αλλάζει πολύ" με το n.

2.6 ΠΡΟΤΑΣΗ (ΙΔΙΟΤΗΤΕΣ ΤΗΣ $h(T,\alpha)$)

Εστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας, $T: X\to X$ ένας ενδομορφισμός και α,β π.μ.δ.

(i) $h(T,\alpha) \le H(\alpha)$.

(ii) $h(T,\alpha\vee\beta) \le h(T,\alpha) + h(T,\beta)$.

(iii) $\alpha\le\beta \Rightarrow h(T,\alpha)\le h(T,\beta)$.

(iv) $h(T,\alpha) \le h(T,\beta) + H(\alpha|\beta)$.

(v) $h(T,T^{-1}\alpha) = h(T,\alpha)$.

(vi) $h(T,\alpha) = h(T,\bigvee_{i=0}^{n-1} T^{-i}\alpha) \quad n\ge 1$

Αν ο T είναι αυτομορφισμός, $h(T,\alpha) = h(T,\bigvee_{i=-m}^{n} T^{-i}\alpha)$, $m,n\ge 1$.

Απόδειξη:

(i) $h(T,\alpha) = \lim_{n\to\infty} n^{-1}H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \le \lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} H(T^{-i}\alpha) = \lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} H(\alpha) = H(\alpha)$.

(ii) $H(\bigvee_{i=0}^{n-1} T^{-i}(\alpha\vee\beta)) = H(\bigvee_{i=0}^{n-1} T^{-i}\alpha \vee \bigvee_{i=0}^{n-1} T^{-i}\beta) \le H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) + H(\bigvee_{i=0}^{n-1} T^{-i}\beta)$.

Διαιρώντας με n και παίρνοντας όρια παίρνουμε ότι $h(T,\alpha\vee\beta)\le h(T,\alpha)+h(T,\beta)$.

(iii) $\alpha\le\beta \Rightarrow \bigvee_{i=0}^{n-1} T^{-i}\alpha \le \bigvee_{i=0}^{n-1} T^{-i}\beta \Rightarrow H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \le H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \Rightarrow h(T,\alpha)\le h(T,\beta)$.

(iv) Εχουμε διαδοχικά:

$$H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \le H(\bigvee_{i=0}^{n-1} T^{-i}\alpha \vee \bigvee_{i=0}^{n-1} T^{-i}\beta) \quad (\text{απο την (iii)})$$

(από Πρόταση 1.1(ii)) $= H(\bigvee_{i=0}^{n-1} T^{-i}\beta) + H(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid \bigvee_{i=0}^{n-1} T^{-i}\beta) \le$

(από Πρόταση 1.1(iv)) $\le H(\bigvee_{i=0}^{n-1} T^{-i}\beta) + \sum_{i=0}^{n-1} H(T^{-i}\alpha \mid \bigvee_{i=0}^{n-1} T^{-i}\beta) \le$

(από Πρόταση 1.1(iii)) $\le H(\bigvee_{i=0}^{n-1} T^{-i}\beta) + \sum_{i=0}^{n-1} H(T^{-i}\alpha \mid T^{-i}\beta) =$

(από Πρόταση 1.1 (viii)) $= H(\bigvee_{i=0}^{n-1} T^{-i}\beta) + n\,H(\alpha|\beta)$.

Διαιρώντας με n και παίρνοντας όρια έπεται ότι $h(T,\alpha)\le h(T,\beta) + H(\alpha|\beta)$.

v) $H(\bigvee_{i=1}^{n} T^{-i}\alpha) = H(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$. Άρα

$$h(T, T^{-1}\alpha) = \lim_{n\to\infty} n^{-1} H(\bigvee_{i=0}^{n-1} T^{-i}(T^{-1}\alpha)) = \lim_{n\to\infty} n^{-1} H(\bigvee_{i=1}^{n} T^{-i}\alpha) = h(T,\alpha).$$

vi) $$h(T, \bigvee_{i=0}^{n} T^{-i}\alpha) = \lim_{k\to\infty} k^{-1} H(\bigvee_{i=0}^{k-1} T^{-i}(\bigvee_{i=0}^{n} T^{-i}\alpha)) =$$

$$= \lim_{k\to\infty} k^{-1} H(\bigvee_{i=0}^{n+k-1} T^{-i}\alpha) = h(T,\alpha).$$

vii) $$h(T, \bigvee_{i=-m}^{n} T^{-i}\alpha) = h(T, \bigvee_{i=0}^{n+m+1} T^{-i}\alpha) = h(T,\alpha).$$ ■

2.7 ΘΕΩΡΗΜΑ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός.

(i) $h(T^k) = k\, h(T)$, $k \geq 0$.

(ii) Αν ο T είναι αυτομορφισμός τότε $h(T^k) = |k|\, h(T)$ για όλα τα $k \in \mathbb{Z}$.

Απόδειξη:

(i) Κατ' αρχήν

$$H(\bigvee_{j=0}^{n-1} T^{-kj}(\bigvee_{i=0}^{k-1} T^{-i}\alpha)) = H(\bigvee_{i=0}^{kn-1} T^{-i}\alpha).$$

Διαιρώντας με n και παίρνοντας όρια παίρνουμε ότι

$$h(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha) = k\, h(T,\alpha).$$

Επομένως

$$k\,h(T) = k \sup_\alpha h(T,\alpha) = \sup_\alpha h(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha) \leq \sup_\alpha h(T^k, \alpha) = h(T^k).$$

Αφού επίσης

$$h(T^k, \alpha) \leq h(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha) = k\, h(T,\alpha)$$

για κάθε π.μ.δ. α του X, παίρνουμε ότι $h(T^k) \leq k\,h(T)$ και άρα τελικά $h(T^k) = k\, h(T)$.

(ii) Αρκεί να δείξουμε ότι $h(T^{-1}) = h(T)$. Για το σκοπό αυτό φθάνει να αποδείξουμε ότι $h(T^{-1},\alpha) = h(T,\alpha)$ για οποιαδήποτε π.μ.δ. α του X. Έχουμε

$$h(T^{-1},\alpha) = \lim_{n\to\infty} n^{-1} H(\bigvee_{i=0}^{n-1} T^{i}\alpha) = \lim_{n\to\infty} n^{-1} H(T^{n-1} \bigvee_{i=0}^{n-1} T^{-i}\alpha) =$$

$$(\text{από Πρόταση 1.2(v)}) = \lim_{n\to\infty} n^{-1} H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) = h(T,\alpha).$$ ■

2.8 ΠΡΟΤΑΣΗ: Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, $T: X \to X$ ενδομορφισμός κ' α μια π.μ.δ. του X. Τότε $h(T,\alpha) = 0$ ανν $\sigma(\alpha) \overset{\circ}{\subseteq} \bigvee_{i=1}^{\infty} T^{-i}\alpha$.

Απόδειξη: $h(T,\alpha) = H(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha)$. Άρα $h(T,\alpha) = 0$ ανν $H(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha) = 0$ ανν $\alpha \overset{\circ}{\subseteq} \bigvee_{i=1}^{\infty} T^{-i}\alpha$ από την Πρόταση 1.1 (vi). ■

Μπορούμε τώρα να αποδείξουμε ότι ένας γνήσια μη αντιστρέψιμος μετασχηματισμός πρέπει να έχει θετική εντροπία.

2.9. ΠΟΡΙΣΜΑ: Εστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Αν $h(T)=0$ τότε $T^{-1}\mathcal{A} \doteq \mathcal{A}$.

Απόδειξη: Απο τη μετρησιμότητα του T έχουμε ότι $T^{-1}\mathcal{A} \subseteq \mathcal{A}$. Εστω τώρα $A \in \mathcal{A}$. Θέτουμε $\alpha = \{A, A^c\}$. Τότε $\sigma(\alpha) = \{\emptyset, X, A, A^c\}$. Αφού $h(T)=0$ πρέπει $h(T,\alpha)=0$ και έτσι από την Πρόταση 2.8 πρέπει $\sigma(\alpha) \dot{\subseteq} \bigvee_{i=1}^{\infty} T^{-i}\alpha$. Ομως $\alpha \subseteq \mathcal{A}$ και άρα $\bigvee_{i=1}^{\infty} T^{-i}\alpha \subseteq T^{-1}\mathcal{A}$. Επομένως $\sigma(\alpha) \dot{\subseteq} T^{-1}\mathcal{A}$ και συγκεκριμένα $A \dot{\in} T^{-1}\mathcal{A}$. Αφού το A ήταν αυθαίρετα επιλεγμένο έπεται ότι $\mathcal{A} \dot{\subseteq} T^{-1}\mathcal{A}$. ∎

ΠΑΡΑΤΗΡΗΣΗ: Αν ο X είναι ένας διαχωρίσιμος πλήρης μετρικός χώρος και $\mathcal{A} = \mathcal{B}(X)$ η σ-άλγεβρα των Borel υποσυνόλων του X, ή αν ο χώρος $(X, \mathcal{A}, \mu)$ είναι ένας χώρος Lebesgue, τότε η σχέση $\mathcal{A} \doteq T^{-1}\mathcal{A}$ δίνει ότι ο T είναι "σχεδόν αντιστρέψιμος", δηλαδή υπάρχει $N \in \mathcal{A}$, με $\mu(N)=0$, τέτοιο ώστε $T(X \setminus N) = X \setminus N$ και ο $T|_{X \setminus N}$ είναι ένας αυτομορφισμός του $X \setminus N$.

Το Πόρισμα 2.9 ισχύει σε μια πιο γενική μορφή:

2.10 ΠΟΡΙΣΜΑ: Εστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Εστω $\mathcal{F}$ μια σ-υποάλγεβρα της $\mathcal{A}$ τέτοια ώστε $T^{-1}\mathcal{F} \subseteq \mathcal{F}$. Αν $h(T)=0$ τότε $T^{-1}\mathcal{F} \doteq \mathcal{F}$.

Απόδειξη: Θεωρούμε τον ενδομορφισμό $\tilde{T}: (X, \mathcal{F}, \mu) \to (X, \mathcal{F}, \mu)$. Αφού $h(T)=0$ πρέπει $h(T,\alpha)=0$ για κάθε π.μ.δ. α του X με $\alpha \subseteq \mathcal{A}$. Τότε προφανώς $h(T,\beta)=0$ για κάθε πεπερασμένη διαμέριση β του X με κάθε στοιχείο της β στην σ-άλγεβρα $\mathcal{F}$. Αρα $h(\tilde{T})=0$. Εφαρμόζοντας τώρα το Πόρισμα 2.9 στον $\tilde{T}: (X, \mathcal{F}, \mu) \to (X, \mathcal{F}, \mu)$ παίρνουμε ότι $T^{-1}\mathcal{F} \doteq \mathcal{F}$. ∎

3. ΓΕΝΝΗΤΟΡΕΣ - ΥΠΟΛΟΓΙΣΜΟΙ

Εστω $(X,\mathcal{A},\mu)$ ένας χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Μια π.μ.δ. α του X είναι ένας μονόπλευρος γεννήτορας του συστήματος $(X,\mathcal{A},\mu,T)$ αν $\bigvee_{i=0}^{\infty} T^{-i}\alpha \stackrel{\circ}{=} \mathcal{A}$. Αν ο T είναι αυτομορφισμός τότε η π.μ.δ. α του X είναι ένας αμφίπλευρος γεννήτορας του συστήματος αν $\bigvee_{i=-\infty}^{\infty} T^{-i}\alpha \stackrel{\circ}{=} \mathcal{A}$.

Το επόμενο Θεώρημα είναι το βασικότερο εργαλείο υπολογισμού εντροπίας.

3.1 ΘΕΩΡΗΜΑ (Kolmogorov-Sinai)

Εστω $(X,\mathcal{A},\mu)$ ένας χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Αν α είναι ένας γεννήτορας του συστήματος τότε $h(T)=h(T,\alpha)$.
Αν ο T είναι αυτομορφισμός και α είναι ένας αμφίπλευρος γεννήτορας τότε πάλι $h(T)=h(T,\alpha)$.

Απόδειξη: Εξ ορισμού $h(T)=\sup_{\beta} h(T,\beta)$. Αρκεί λοιπόν να αποδείξουμε ότι $h(T,\alpha) \geq h(T,\beta)$ για κάθε π.μ.δ. β του X. Εστω λοιπόν β μια π.μ.δ. του X. Απο την Πρόταση 2.6 (iv)

$$h(T,\beta) \leq h(T,\bigvee_{i=0}^{n-1} T^{-i}\alpha) + H(\beta \mid \bigvee_{i=0}^{n-1} T^{-i}\alpha) =$$

(από Πρόταση 2.6(vi)) $= h(T,\alpha) + H(\beta \mid \bigvee_{i=0}^{n-1} T^{-i}\alpha)$.

Τώρα $\sigma(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \subseteq \sigma(\bigvee_{i=0}^{n} T^{-i}\alpha)$ και άρα από την Πρόταση 1.4

$$H(\beta \mid \bigvee_{i=0}^{n-1} T^{-i}\alpha) \to H(\beta \mid \bigvee_{i=0}^{\infty} T^{-i}\alpha).$$

Αφού η α είναι ένας γεννήτορας, $\bigvee_{i=0}^{\infty} T^{-i}\alpha \stackrel{\circ}{=} \mathcal{A}$ και άρα $\sigma(\beta) \stackrel{\circ}{\subseteq} \bigvee_{i=0}^{\infty} T^{-i}\alpha$. Τότε όμως $H(\beta \mid \bigvee_{i=0}^{\infty} T^{-i}\alpha)=0$ απο την Πρόταση 1.3(vi) και έτσι $h(T,\beta) \leq h(T,\alpha)$.

Η περίπτωση που ο T είναι αυτομορφισμός αποδεικνύεται παρόμοια. ■

ΠΑΡΑΤΗΡΗΣΗ: Αν ένας αυτομορφισμός T έχει έναν μονόπλευρο γεννήτορα, δηλαδή αν ήδη $\bigvee_{i=0}^{\infty} T^{-i}\alpha \stackrel{\circ}{=} \mathcal{A}$, τότε $h(T)=0$. Πράγματι, ο α είναι και αμφίπλευρος γεννήτορας και άρα $h(T)=h(T,\alpha)$. Τώρα προφανώς $\sigma(\alpha) \subseteq \mathcal{A} = T^{-1}\mathcal{A} \stackrel{\circ}{=} \bigvee_{i=1}^{\infty} T^{-i}\alpha$. Αρα απο το Πόρισμα 2.4 $h(T,\alpha) = H(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha) = 0$ από την Πρόταση 1.3(vi) (αφού $\sigma(\alpha) \stackrel{\circ}{\subseteq} \bigvee_{i=1}^{\infty} T^{-i}\alpha$).

3.2 ΠΑΡΑΔΕΙΓΜΑ: (ΣΤΡΟΦΕΣ ΤΟΥ ΚΥΚΛΟΥ)

Έστω $X = \{z \in \mathbb{C} : |z| = 1\}$ ο μοναδιαίος κύκλος, $\mathcal{A} = \mathcal{B}(X)$ τα Borel υποσύνολα του X και $\mu = \lambda$, το μέτρο Lebesgue. Αν $T : X \to X$ είναι μια (οποιαδήποτε) στροφή $T(z) = az$, όπου $a \in X$, τότε $h(T) = 0$.

Απόδειξη: Έστω ότι $a = e^{2\pi i\theta}$. Αν $\theta \in [0,1) \cap \mathbb{Q}$, έστω $\theta = \frac{m}{n}$, με $m \in \mathbb{Z}$ και $n \in \mathbb{N}$, τότε T^n είναι η ταυτοτική του X και άρα $h(T^n) = 0$. Από το Θεώρημα 2.7,

$h(T) = \frac{1}{n} h(T^n) = 0$.

Αν $\theta \in [0,1) \cap \mathbb{Q}^c$, τότε η τροχιά $\{T^n(-1) : n \in \mathbb{N}_0\}$ του $-1 = e^{2\pi i \frac{1}{2}}$ είναι πυκνό σύνολο στον X. Έστω α η π.μ.δ. του X που αποτελείται από τα δύο ημικύκλια $[0, e^{\pi i})$ και $[e^{\pi i}, 1)$. Τότε η α είναι ένας γεννήτορας του συστήματος και μάλιστα μονόπλευρος, ακριβώς επειδή η τροχιά του $e^{\pi i} = -1$ είναι πυκνή.

Τώρα η $\bigvee_{i=0}^{k-1} T^{-i}\alpha$ έχει $2(k+1)$ στοιχεία. Επίσης,

$$\max\left\{-\sum_{i=1}^{l} p_i \log p_i : p_i \in [0,1],\ \sum_{i=1}^{l} p_i = 1\right\} = \log l .$$

Άρα

$$\frac{1}{k} H\Big(\bigvee_{i=0}^{k-1} T^{-i}\alpha\Big) = \frac{1}{k} \sum_{A \in \bigvee_{i=0}^{k-1} T^{-i}\alpha} \mu(A)\log\mu(A) \le \frac{\log(2k+2)}{k} \to 0 .$$

Έπεται ότι $h(T) = h(T, \alpha) = 0$.

(Ότι $h(T) = 0$ έπεται και από την προηγούμενη παρατήρηση αφού ο T είναι αντιστρέψιμος και η α είναι ένας μονόπλευρος γεννήτορας.) ∎

3.3 ΠΑΡΑΔΕΙΓΜΑ: (ΑΠΛΟ SHIFT)

(i) Έστω $N \in \mathbb{N}$ και $X = \{0, 1, \dots, N-1\}^{\mathbb{N}_0}$, $\mathcal{A} = \mathcal{B}(X)$ τα Borel υποσύνολα του X (όπου θεωρούμε ότι ο X έχει την τοπολογία γινόμενο που προέρχεται από την διακριτή τοπολογία στο $\{0, 1, \dots, N-1\}$) και μ ένα οποιοδήποτε μέτρο πιθανότητας της μορφής $\mu = \mu_p = \mu_p \times \cdots$ (μέτρο γινόμενο) όπου $\mu_p(\{i\}) = p_i$, $0 \le i < N$ για κάποιο διάνυσμα πιθανότητας $p = (p_0, \dots, p_{N-1})$. Έστω τέλος T το shift, δηλ.: $T(\{x_0, x_1, \dots\}) = \{x_1, x_2, \dots\}$. Τότε $h(T) = -\sum_{i=0}^{N-1} p_i \log p_i$.

Πράγματι η π.μ.δ. α που αποτελείται από τους κυλίνδρους $\{(x_i)_{i=0}^{\infty} : x_0 = i\}$, $0 \le i < N$, αποτελεί έναν (μονόπλευρο φυσικά) γεννήτορα του συστήματος, αφού ένα στοιχείο της $\bigvee_{i=0}^{n} T^{-i}\alpha$ είναι ένας κύλινδρος της μορφής

$$A_0 \cap T^{-1}A_1 \cap \dots \cap T^{-n}A_n = \{(x_i) : x_0 = j_0, x_1 = j_1, \dots, x_n = j_n\}$$

για κάποια $0 \le j_0 < N$, $0 \le j_1 < N$, ..., $0 \le j_n < N$. Τώρα για κάθε n,

$$\frac{1}{n+1} H\left(\bigvee_{i=0}^{n} T^{-i}\alpha\right) =$$

$$= -\frac{1}{n+1} \sum_{i_0,\dots,i_n} \mu\{(x_i) : x_0 = i_0, \dots, x_n = i_n\} \log \mu\{(x_i) : x_0 = i_0, \dots, x_n = i_n\} =$$

$$= -\frac{1}{n+1} \sum_{i_0,\dots,i_n} p_{i_0} \cdots p_{i_n} \log p_{i_0} \cdots p_{i_n} =$$

$$= -\frac{1}{n+1} \sum_{i_0 \dots i_n} p_{i_0} \cdots p_{i_n} (\log p_{i_0} + \dots + \log p_{i_n}) =$$

$$= -\frac{1}{n+1} \left(\sum_{i_0,\dots,i_n} p_{i_0} \cdots p_{i_n} \log p_{i_0} + \dots + \sum_{i_0,\dots,i_n} p_{i_0} \cdots p_{i_n} \log p_{i_n} \right) =$$

$$= -\frac{1}{n+1} \left(\sum_{i_0=0}^{N-1} p_{i_0} \log p_{i_0} + \dots + \sum_{i_n=0}^{N-1} p_{i_n} \log p_{i_n} \right) = -\sum_{i=0}^{N-1} p_i \log p_i .$$

Άρα $h(T) = h(T, \alpha) = -\sum_{i=0}^{N-1} p_i \log p_i$.

Ο ίδιος υπολογισμός δείχνει $h(T) = -\sum_{i=0}^{N-1} p_i \log p_i$ και για το αμφίπλευρο shift, δηλαδή όταν $X = \{0,1,\dots,N-1\}^{\mathbb{Z}}$. ∎

3.4 ΠΑΡΑΔΕΙΓΜΑ: (MARKOV SHIFT)

i) Έστω X, $\mathcal{A}$ και T όπως στο Παράδειγμα 3.3 (i) και έστω $p = (p_0, \dots, p_N)$ ένα διάνυσμα πιθανότητας και $P = (p(i,j))_{0 \le i,j < N}$ ένας στοχαστικός πίνακας (δηλ. $p(i,j) \ge 0$ για όλα τα i,j και $\sum_j p(i,j) = 1$ για κάθε i) τέτοια ώστε $pP = p$. Έστω τέλος μ το (αναλλοίωτο ως προς T) μέτρο Markov που αντιστοιχεί στα p και P. Έστω α πάλι η π.μ.δ. του X που αποτελείται από τους κυλίνδρους $\{(x_i) : x_0 = i\}$, $0 \le i < N$. Τότε η α είναι ένας γεννήτορας για το σύστημα. Θα αποδείξουμε ότι

$$h(T) = h(T, \alpha) = -\sum_{i,j} p_i \, p(i,j) \log p(i,j).$$

$$H\Big(\bigvee_{i=0}^{n} T^{-i}\alpha\Big) = -\sum_{i_0=0}^{N-1}\cdots\sum_{i_n=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-1},i_n)\big[\log p_{i_0} + \log p(i_0,i_1) + \cdots + \log p(i_{n-1},i_n)\big] =$$

$$= -\sum_{i_0=0}^{N-1}\cdots\sum_{i_n=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-1},i_n)\big[\log p_{i_0} + \log p(i_0,i_1) + \cdots + \log p(i_{n-2},i_{n-1})\big] -$$

$$-\sum_{i_0=0}^{N-1}\cdots\sum_{i_n=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-1},i_n)\log p(i_{n-1},i_n) =$$

$$= -\sum_{i_0=0}^{N-1}\cdots\sum_{i_{n-1}=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-2},i_{n-1})\big[\log p_{i_0} + \log p(i_0,i_1) + \cdots + \log p(i_{n-2},i_{n-1})\big]\cdot\sum_{i_n=0}^{N-1} p(i_{n-1},i_n) -$$

$$-\sum_{i_0=0}^{N-1}\cdots\sum_{i_{n-1}=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-2},i_{n-1})\sum_{j=0}^{N-1} p(i_{n-1},j)\log p(i_{n-1},j) =$$

$\left(\sum_{i_n=0}^{N-1} p(i_{n-1},i_n)=1\right)$

$$= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) - \sum_{i_0=0}^{N-1}\cdots\sum_{i_{n-1}=0}^{N-1} p_{i_0}\, p(i_0,i_1)\cdots p(i_{n-2},i_{n-1})\sum_{j=0}^{N-1} p(i_{n-1},j)\log p(i_{n-1},j) =$$

(αφού $pP=p$)

$$= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) - \sum_{i_1=0}^{N-1}\cdots\sum_{i_{n-1}=0}^{N-1} p_{i_1}\, p(i_1,i_2)\cdots p(i_{n-2},i_{n-1})\sum_{j=0}^{N-1} p(i_{n-1},j)\log p(i_{n-1},j)$$

$$= \cdots = H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) - \sum_{i_{n-1}=0}^{N-1} p_{i_{n-1}}\, p(i_{n-1},j)\log p(i_{n-1},j)$$

Δηλαδή για κάθε n,

$$H\Big(\bigvee_{i=0}^{n} T^{-i}\alpha\Big) = H\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) - \sum_{i,j} p_i\, p(i,j)\log p(i,j).$$

Επομένως

$$H\Big(\bigvee_{i=0}^{n} T^{-i}\alpha\Big) = H(\alpha) - n\sum_{i,j} p_i\, p(i,j)\log p(i,j)$$

που δίνει ότι

$$h(T) = h(T,\alpha) = \lim_{n\to\infty}\frac{1}{n+1} H\Big(\bigvee_{i=0}^{n} T^{-i}\alpha\Big) = -\sum_{i,j} p_i\, p(i,j)\log p(i,j).$$

(ii) Με τον ίδιο υπολογισμό παίρνουμε ότι $h(T) = -\sum_{i,j} p_i\, p(i,j)\log p(i,j)$ για το αμφίπλευρο Markov shift, όταν δηλαδή $X = \{0,1,\ldots,N-1\}^{\mathbb{Z}}$. ■

Κλείνουμε αυτήν την ενότητα με ένα παράδειγμα ενδομορφισμού με $h_\mu(T)=\infty$.

3.5 ΠΑΡΑΔΕΙΓΜΑ: Έστω $X = \prod_{i=0}^{\infty}[0,1)$ με $\mathcal{A} = \mathcal{B}([0,1))\times\mathcal{B}([0,1))\times\cdots$ και $\mu = \lambda\times\lambda\times$

όπου λ το μέτρο Lebesgue στο $[0,1]$. $T: X \to X$ είναι πάλι το shift. Τότε $h(T)=+\infty$.

Για κάθε n ορίζουμε την διαμέριση α_n του X που αποτελείται από τα σύνολα

$$A_i^{(n)} = \{ (x_i) : \frac{i-1}{n} \le x_0 < \frac{i}{n} \} , \quad 1 \le i \le n .$$

Τότε $\mu(A_i^{(n)}) = n^{-1}$ για κάθε $1 \le i \le n$, αφού το μ είναι μέτρο γινόμενο. Επίσης ένα στοιχείο $A_{i_0}^{(n)} \cap T^{-1}A_{i_1}^{(n)} \cap \dots \cap T^{-k+1}A_{i_{k-1}}^{(n)}$ της $\bigvee_{j=0}^{k-1} T^{-j}\alpha_n$ έχει μέτρο

$$\mu (A_{i_0}^{(n)} \cap T^{-1} A_{i_1}^{(n)} \cap \dots \cap T^{-k+1} A_{i_{k-1}}^{(n)}) = n^{-k}$$

και έτσι

$$h(T,\alpha_n) = \lim_{k\to\infty} -\frac{1}{k} \sum_{i_0=1}^{n} \dots \sum_{i_{k-1}=1}^{n} n^{-k} \log n^{-k} = \log n .$$

Όμως $h(T) = \sup_\alpha h(T,\alpha) \ge h(T,\alpha_n) = \log n$ για κάθε $n \in \mathbb{N}$ και άρα $h(T)=+\infty$. ■

4. ΑΛΛΕΣ ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΗΣ $h(T)$

4.1 ΛΗΜΜΑ: Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $r \in \mathbb{N}$. Τότε για κάθε $\epsilon>0$ υπάρχει $\delta>0$ τέτοιο ώστε: αν $\alpha=\{A_1,\dots,A_r\}$ και $\beta=\{B_1,\dots,B_r\}$ είναι δυο π.μ.δ. με r στοιχεία η καθε μία, τέτοιες ώστε $\sum_{i=1}^{r} \mu(A_i \triangle B_i) < \delta$ τότε $H(\alpha|\beta)+H(\beta|\alpha)<\epsilon$.

Απόδειξη: Θέτουμε $\gamma := \{ A_i \cap B_j : i \ne j \} \cup \{ \bigcup_{i=1}^{r} A_i \cap B_i \}$. Τότε η γ είναι μια π.μ.δ. του X και $\alpha\vee\beta = \alpha\vee\gamma$. Άρα $H(\beta|\alpha) + H(\alpha) = H(\alpha\vee\beta) = H(\alpha\vee\gamma) \le H(\alpha)+H(\gamma)$. Επομένως $H(\beta|\alpha) \le H(\gamma)$. Ομοια, επειδή και $\alpha\vee\beta=\beta\vee\gamma$, παίρνουμε $H(\alpha|\beta) \le H(\gamma)$. Άρα αρκεί να αποδείξουμε ότι αν επιλέξουμε κατάλληλα το δ τότε $H(\gamma)<\epsilon$.

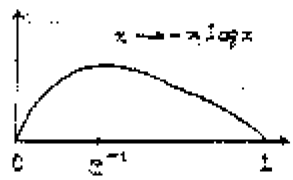
Παρατηρούμε ότι για $i \ne j$

$$\mu(A_i \cap B_j) \le \mu \Big(\bigcup_{i=1}^{r} (A_i \triangle B_i) \Big) \le \sum_{i=1}^{r} \mu(A_i \triangle B_i) < \delta .$$

Έτσι παίρνουμε και ότι

$$\mu \Big(\bigcup_{i=1}^{r} (A_i \cap B_i) \Big) = 1 - \sum_{i\ne j} \mu(A_i \cap B_j) > 1 - r(r-1)\delta$$

(αφού η γ είναι διαμέριση). Επειδή η συνάρτηση $x \mapsto -x\log x$ τείνει στο 0 όταν $x \to 0^+$ ή $x \to 1^-$ μπορούμε λοιπόν να διαλέξουμε το δ αρκούντως μικρό για να έχουμε $x \in [0,\delta) \cup (r(r-1)\delta, 1] \Rightarrow$ $\Rightarrow -x\log x < \frac{\epsilon}{r(r-1)+1}$. Τότε $-\mu(C)\log\mu(C) < \frac{\epsilon}{r(r-1)+1}$, για κάθε $C \in \gamma$ και άρα $H(\gamma)<\epsilon$. ■



4.2 ΠΡΟΤΑΣΗ : Εστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Εστω $\mathcal{A}_0 \subseteq \mathcal{A}$ μια άλγεβρα τέτοια ώστε $\sigma(\mathcal{A}_0) \overset{\circ}{=} \mathcal{A}$. Τότε για κάθε π.μ.δ. α του X και κάθε $\epsilon > 0$ υπάρχει μια π.μ.δ. α_ϵ του X με $\alpha_\epsilon \subseteq \mathcal{A}_0$ (δηλ. κάθε στοιχείο της α_ϵ ανήκει στην άλγεβρα $\mathcal{A}_0$) τέτοια ώστε $H(\alpha|\alpha_\epsilon) + H(\alpha_\epsilon|\alpha) < \epsilon$.

Απόδειξη : Αφού η $\mathcal{A}_0$ είναι άλγεβρα και $\sigma(\mathcal{A}_0) \overset{\circ}{=} \mathcal{A}$, για κάθε $A \in \mathcal{A}$ υπάρχει $B \in \mathcal{A}_0$ τέτοιο ώστε $\mu(A \triangle B) \le \delta$, για οποιοδήποτε δοσμένο $\delta > 0$. Εστω λοιπόν ότι $\alpha = \{A_1, \dots, A_r\}$ και $B_1, \dots, B_r \in \mathcal{A}_0$ τέτοια ώστε $\mu(A_i \triangle B_i) < \delta$ για κάθε $1 \le i \le r$, όπου το δ θα επιλεγεί κατάλληλα παρακάτω. Θέτουμε $N = \bigcap_{i \ne j} B_i \cap B_j$ και $C_i = B_i - N$ για όλα τα $1 \le i < r$ και $C_r = X - \bigcup_{i=1}^{r-1} C_i$. Τότε η $\alpha_\epsilon = \{C_1, \dots, C_r\}$ είναι προφανώς μια π.μ.δ. του X με $\alpha_\epsilon \subseteq \mathcal{A}_0$.

Τώρα $B_i \cap B_j \subseteq (A_i \triangle B_i) \cup (A_j \cap B_j)$ και άρα $\mu(B_i \cap B_j) < 2\delta$. Επομένως, $\mu(N) < 2r(r-1)\delta$. Αφού $C_i \triangle A_i \subseteq (B_i \triangle A_i) \cap N$, για $1 \le i < r$, έπεται ότι

$$\mu(C_i \triangle A_i) \le \delta + 2r(r-1)\delta = \delta[2r(r-1)+1], \quad 1 \le i < r.$$

Αφού επίσης $C_r \triangle A_r \subseteq \bigcup_{i=1}^{r-1} (C_i \triangle A_i)$, έχουμε ότι $\mu(C_r \triangle A_r) < (r-1)[2r(r-1)+1]\delta$. Έτσι

$$\sum_{i=1}^{r} \mu(A_i \triangle B_i) < 2(r-1)[2r(r-1)+1]\delta.$$

Τώρα μπορούμε να διαλέξουμε το δ κατάλληλα από το προηγούμενο λήμμα ώστε να έχουμε $H(\alpha|\alpha_\epsilon) + H(\alpha_\epsilon|\alpha) < \epsilon$. ■

4.3 ΠΡΟΤΑΣΗ : Εστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Αν $\mathcal{A}_0$ είναι μια άλγεβρα που παράγει την $\mathcal{A}$, δηλ. $\sigma(\mathcal{A}_0) \overset{\circ}{=} \mathcal{A}$

$$h(T) = \sup_{\alpha \subseteq \mathcal{A}_0} h(T, \alpha)$$

Απόδειξη : Προφανώς $\sup_{\alpha \subseteq \mathcal{A}_0} h(T,\alpha) \le \sup_{\alpha \subseteq \mathcal{A}} h(T,\alpha) = h(T)$.
Εστω τώρα $\epsilon > 0$ αυθαίρετο και $\alpha \subseteq \mathcal{A}$ μια π.μ.δ. του X, τέτοια ώστε $h(T,\alpha) \ge h(T) - \epsilon/2$. Από την προηγούμενη Πρόταση υπάρχει $\alpha_0 \subseteq \mathcal{A}_0$, π.μ.δ. του X ώστε $H(\alpha|\alpha_0) + H(\alpha_0|\alpha) \le \epsilon/2$. Τότε όμως (Πρόταση 2.6 (iv))

$$h(T,\alpha) \le h(T,\alpha_0) + H(\alpha|\alpha_0) \le h(T,\alpha_0) + \frac{\epsilon}{2}.$$

Επομένως, $h(T) - \epsilon \le h(T,\alpha_0) \le \sup_{\beta \subseteq \mathcal{A}_0} h(T,\beta)$. ■

Το επόμενο θεώρημα είναι μια εναλλακτική λύση όταν δεν υπάρχει γεννήτορας.

4.4 ΘΕΩΡΗΜΑ: Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός. Έστω $\alpha_1 \leq \alpha_2 \leq \cdots$ μια αύξουσα ακολουθία π.μ.δ. του X με $\bigvee_{n=1}^{\infty} \alpha_n \doteq \mathcal{A}$. Τότε

$$h(T) = \lim_{n \to \infty} h(T, \alpha_n).$$

Απόδειξη: Θέτουμε $\mathcal{A}_0 = \bigcup_{n=1}^{\infty} \mathcal{A}_n$ όπου $\mathcal{A}_n$ είναι η άλγεβρα που παράγει η α_n. Τότε η $\mathcal{A}_0$ είναι άλγεβρα και $\sigma(\mathcal{A}_0) \doteq \mathcal{A}$. Από την Πρόταση 4.3 λοιπόν

$$h(T) = \sup_{\alpha \subset \mathcal{A}_0} h(T, \alpha).$$

Τώρα αν α είναι μια οποιαδήποτε π.μ.δ. του X με $\alpha \subset \mathcal{A}_0$ τότε $\alpha \subset \mathcal{A}_n$ για κάποιο n και άρα $\alpha \leq \alpha_n$ για αυτό το n. Από την Πρόταση 2.6 (iii)

$$h(T, \alpha) \leq h(T, \alpha_n) \leq \lim_{m \to \infty} h(T, \alpha_m).$$

Αφού αυτό ισχύει για κάθε π.μ.δ. α του X με $\alpha \subset \mathcal{A}_0$ έπεται ότι

$$h(T) \leq \lim_{n \to \infty} h(T, \alpha_n).$$

Εξ' ορισμού όμως $h(T) \geq h(T, \alpha_n)$ για κάθε n και άρα $h(T) = \lim_{n \to \infty} h(T, \alpha_n)$. ■

Χρησιμοποιώντας την Πρόταση 4.3 μπορούμε επίσης να υπολογίσουμε την εντροπία ενός ευθέως γινομένου.

4.5 ΠΡΟΤΑΣΗ: Έστω $(X_i, \mathcal{A}_i, \mu_i)$ και T_i ένας ενδομορφισμός του $(X_i, \mathcal{A}_i, \mu_i)$ για $i = 1, 2$. Έστω $X = X_1 \times X_2$, $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ και $\mu = \mu_1 \times \mu_2$ το (ευθύ) γινόμενο των δύο χώρων και $T: X \to X$ ο ενδομορφισμός $T(x_1, x_2) = (T_1(x_1), T_2(x_2))$. Τότε $h(T) = h(T_1) + h(T_2)$.

Απόδειξη: Έστω $\mathcal{R}$ η οικογένεια όλων των πεπερασμένων ενώσεων μετρήσιμων ορθογωνίων. (Ένα μετρήσιμο ορθογώνιο είναι ένα σύνολο της μορφής $A_1 \times A_2$ όπου $A_i \in \mathcal{A}_i$). Τότε η $\mathcal{R}$ είναι άλγεβρα και εξ' ορισμού της σ-άλγεβρας-γινόμενο $\sigma(\mathcal{R}) \doteq \mathcal{A}_1 \times \mathcal{A}_2 = \mathcal{A}$. Άρα από την Πρόταση 4.3

$$h(T) = \sup_{\alpha \subset \mathcal{R}} h(T, \alpha).$$

Τώρα όταν α είναι μια π.μ.δ. του X με $\alpha \subset \mathcal{R}$ τότε υπάρχουν π.μ.δ.

α_1 του X_1 και α_2 του X_2 (με $\alpha_i \subset \mathcal{A}_i$ ουσικά) τέτοιες ώστε $\alpha \leq \alpha_1 \times \alpha_2$.

(Αν $\alpha = \{C_1, \dots, C_m\}$ όπου $C_i = \bigcup_{j=1}^{n_i} A_j^{(i)} \times B_j^{(i)}$ τότε $\alpha \leq \alpha_1 \times \alpha_2$ όπου $\alpha_1 = \{A_j^{(i)} : j = 1, \dots, n_i, \; i = 1, \dots, m\}$ και $\alpha_2 = \{B_j^{(i)} : j = 1, \dots, n_i, \; i = 1, \dots, m\}$.)

Επομένως,

$$h(T, \alpha) \leq h(T, \alpha_1 \times \alpha_2)$$

και άρα

$$h(T) \leq \sup_{\substack{\alpha_1 \subset \mathcal{A}_1 \\ \alpha_2 \subset \mathcal{A}_2}} h(T, \alpha_1 \times \alpha_2).$$

Τώρα αν α_1 και α_2 είναι π.μ.δ. των X_1 και X_2 αντίστοιχα η $\alpha_1 \times \alpha_2$ είναι προφανώς μία π.μ.δ. του X και άρα $h(T) \geq \sup_{\substack{\alpha_1 \subset \mathcal{A}_1 \\ \alpha_2 \subset \mathcal{A}_2}} h(T, \alpha_1 \times \alpha_2)$ και έτσι έχουμε τελικά ότι

$$h(T) = \sup_{\substack{\alpha_1 \subset \mathcal{A}_1 \\ \alpha_2 \subset \mathcal{A}_2}} h(T, \alpha_1 \times \alpha_2).$$

Έστω τώρα ότι έχουμε π.μ.δ. α_1 και α_2 των X_1 και X_2 αντίστοιχα. Τότε

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}(\alpha_1 \times \alpha_2)\right) = H\left(\left(\bigvee_{i=0}^{n-1} T_1^{-i}\alpha_1\right) \times \left(\bigvee_{i=0}^{n-1} T_2^{-i}\alpha_2\right)\right) =$$

$$= - \sum_{\substack{A_1 \in \bigvee_{i=0}^{n-1} T_1^{-i}\alpha_1 \\ A_2 \in \bigvee_{i=0}^{n-1} T_2^{-i}\alpha_2}} \mu(A_1 \times A_2) \log \mu(A_1 \times A_2) =$$

$$= - \sum_{A_1, A_2} \mu_1(A_1)\mu_2(A_2) \log \mu_1(A_1) - \sum_{A_1, A_2} \mu_1(A_1)\mu_2(A_2) \log \mu_2(A_2) =$$

$$= - \sum_{A_1} \mu_1(A_1) \log \mu(A_1) - \sum_{A_2} \mu_2(A_2) \log \mu_2(A_2) =$$

$$= H\left(\bigvee_{i=0}^{n-1} T_1^{-i}\alpha_1\right) + H\left(\bigvee_{i=0}^{n-1} T_2^{-i}\alpha_2\right).$$

Άρα $h(T, \alpha_1 \times \alpha_2) = h(T_1, \alpha_1) + h(T_2, \alpha_2)$, για οποιεσδήποτε δύο π.μ.δ. α_1, α_2 των X_1 και X_2 αντίστοιχα. Αφού και $h(T) = \sup h(T, \alpha_1 \times \alpha_2)$ το ζητούμενο έπεται. ∎

4.6 ΠΑΡΑΔΕΙΓΜΑ : (ΣΤΡΟΦΕΣ ΤΟΥ n-TORUS)

Έστω $X = S^1 \times \dots \times S^1$ ο n-torus με μέτρο Lebesgue και $T : X \to X$ ο ενδομορφισμός $T(z_1, \dots, z_n) = (a_1 z_1, \dots, a_n z_n)$, όπου $a_i \in \mathbb{C}$ με $|a_i| = 1$, $1 \le i \le n$. Ορίζουμε την στροφή $T_i : S^1 \to S^1$, $T_i(z) = a_i z$ του 1-torus. Ήδη γνωρίζουμε ότι $h(T_i) = 0$ (Παράδειγμα 3.2). Από την Πρόταση 4.5 έπεται ότι

$$h(T) = h(T_1) + \dots + h(T_n) = 0.$$

Δηλαδή κάθε στροφή του n-torus (εργοδική ή μή) έχει εντροπία 0. ∎

Το παραπάνω αποτέλεσμα γενικεύεται. Συγκεκριμένα έστω $X = G$, $\mathcal{A} = \mathcal{B}(G)$ και μ το μέτρο Haar της G, όπου G είναι μιά μετρικοποιήσιμη, συμπαγής αβελιανή ομάδα. Έστω $a \in G$ και $T(x) = ax$, για $x \in G$. Τότε $h(T) = 0$.

Μια απόδειξη του παραπάνω αποτελέσματος, που στηρίζεται στα θεωρία χαρακτήρων και στο Παράδειγμα 4.6, υπάρχει στο βιβλίο του Walters, An Introduction to Ergodic Theory, Springer Verlag, 1982.

5. Το Θεώρημα Shannon - McMillan - Breiman

5.1 Θεώρημα: Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός που είναι εργοδικός. Τότε για κάθε π.μ.δ. α του X

$$\frac{1}{n} I\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) \to h(T, \alpha) \quad \mu\text{-σ.π. και στο } L^1.$$

Παρατηρήσεις (i) Αν η α είναι ένας γεννήτορας τότε $\frac{1}{n} I(\vee_{i=0}^{n-1} T^{-i}\alpha) \to h(T)$ μ-σ.π., από το παραπάνω θεώρημα και το Θεώρημα Sinai - Kolmogorov.

(ii) Έστω $x \in X$ και $A_n(x)$ το στοιχείο εκείνο της διαμέρισης $\vee_{i=0}^{n-1} T^{-i}\alpha$ που περιέχει το x. Τότε $I(\vee_{i=0}^{n-1} T^{-i}\alpha)(x) = -\log \mu(A_n(x))$. Από το παραπάνω θεώρημα, αν $h(T, \alpha) > 0$ τότε $\mu(A_n(x)) \longrightarrow 0$ εκθετικά με "ταχύτητα" $e^{-nh(T,\alpha)}$

Απόδειξη:

(1) Αν β και γ είναι δυο π.μ.δ. του X τότε $I(\beta \vee \gamma) = I(\gamma | \beta) + I(\beta)$.

Η απόδειξη αυτής της σχέσης είναι όπως και για την αντίστοιχη ιδιότητα για εντροπίες (Πρόταση 1.1 (i)):

$$I(\beta \vee \gamma) = -\sum_{B \in \beta, C \in \gamma} 1\!\!1_{B \cap C} \log \mu(B \cap C) = -\sum_{B \in \beta, C \in \gamma} 1\!\!1_{B \cap C} \log \frac{\mu(B \cap C)}{\mu(B)} -$$

$$- \sum_{B \in \beta, C \in \gamma} 1\!\!1_{B \cap C} \log \mu(B) = I(\gamma | \beta) - \sum_{B \in \beta} 1\!\!1_B \log \mu(B) =$$

$$= I(\gamma | \beta) + I(\beta).$$

Από την (1) παίρνουμε άμεσα ότι

$$(2) \qquad I\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) = I\left(\alpha \,\Big|\, \bigvee_{i=1}^{n-1} T^{-i}\alpha\right) + I\left(\bigvee_{i=1}^{n-1} T^{-i}\alpha\right) =$$

$$= I\left(\alpha \,\Big|\, \bigvee_{i=1}^{n-1} T^{-i}\alpha\right) + I\left(T^{-1}\alpha \,\Big|\, \bigvee_{i=2}^{n-1} T^{-i}\alpha\right) + \dots + I(T^{-n+1}\alpha).$$

Τώρα για κάθε $k = 0, 1, \dots, n-1$ ισχύει

(3) $$I\Big(T^{-k}\alpha \,\Big|\, \bigvee_{i=k+1}^{n-1} T^{-i}\alpha\Big) = I\Big(\alpha \,\Big|\, \bigvee_{i=1}^{n-k-1} T^{-i}\alpha\Big)\circ T^{k}$$

(όπου με το $\bigvee_{i=j+1}^{j} T^{-i}\alpha$ εννοούμε την τετριμμένη σ-άλγεβρα $\{\emptyset, X\}$ για ευκολία συμβολισμού). Η απόδειξη της (3) έχει ως εξής:

$$I\Big(T^{-k}\alpha \,\Big|\, \bigvee_{i=k+1}^{n-1} T^{-i}\alpha\Big) = -\sum_{\substack{A\in T^{-k}\alpha \\ B\in \bigvee_{i=k+1}^{n-1} T^{-i}\alpha}} \mathbb{1}_{A\cap B}\log\frac{\mu(A\cap B)}{\mu(B)} =$$

$$-\sum_{\substack{A\in\alpha \\ B_1,\dots,B_{n-k-1}\in\alpha}} \mathbb{1}_{T^{-k}A\cap T^{-k-1}B_1\cap\dots\cap T^{-n+1}B_{n-k-1}} \log\frac{\mu(T^{-k}A\cap T^{-k-1}B_1\cap\dots\cap T^{-n+1}B_{n-k-1})}{\mu(T^{-k-1}B_1\cap\dots\cap T^{-n+1}B_{n-k-1})}$$

$$= -\sum_{\substack{A\in\alpha \\ B_1,\dots,B_{n-k-1}\in\alpha}} \mathbb{1}_{T^{-k}(A\cap T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1})} \log\frac{\mu(T^{-k}(A\cap T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1}))}{\mu(T^{-k}(T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1}))}$$

$$= -\sum_{\substack{A\in\alpha \\ B_1,\dots,B_{n-k-1}\in\alpha}} \mathbb{1}_{A\cap T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1}}\circ T^{k} \log\frac{\mu(A\cap T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1})}{T^{-1}B_1\cap\dots\cap T^{-n+k+1}B_{n-k-1})}$$

$$= -\sum_{\substack{A\in\alpha \\ B\in \bigvee_{i=1}^{n-k-1} T^{-i}\alpha}} \mathbb{1}_{A\cap B}\circ T^{k}\log\frac{\mu(A\cap B)}{\mu(B)} =$$

$$= \Bigg(-\sum_{\substack{A\in\alpha \\ B\in \bigvee_{i=1}^{n-k-1} T^{-i}\alpha}} \mathbb{1}_{A\cap B}\log\frac{\mu(A\cap B)}{\mu(B)}\Bigg)\circ T^{k} =$$

$$= I\Big(\alpha \,\Big|\, \bigvee_{i=1}^{n-k-1} T^{-i}\alpha\Big)\circ T^{k}$$

Από τις (2) και (3) λοιπόν παίρνουμε ότι

$$I\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) = \sum_{k=0}^{n-1} I\Big(\alpha \,\Big|\, \bigvee_{i=1}^{n-k-1} T^{-i}\alpha\Big)\circ T^{k}$$

και αν θέσουμε $f_0 = I(\alpha)$, $f_1 = I(\alpha \mid T^{-1}\alpha)$, ..., $f_n = I(\alpha \mid \bigvee_{i=1}^{n} T^{-i}\alpha)$, τότε

(4) $$I\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha\Big) = \sum_{k=0}^{n-1} f_{n-k-1}\circ T^{k}.$$

Θέτουμε $f = I(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha)$. Από το εργοδικό θεώρημα,

$$\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k \longrightarrow \int f\,d\mu = H(\alpha \mid \bigvee_{i=1}^{\infty} T^{-i}\alpha) = h(T,\alpha)$$

καθώς $n\to\infty$, μ-σ.π., αφού $f\in L^1(X,\mathcal{A},\mu)$. Αρκεί λοιπόν να αποδείξουμε ότι

$$(5)\qquad \frac{1}{n}\sum_{k=0}^{n-1} |f-f_{n-k-1}|\circ T^k \longrightarrow 0 \qquad \mu\text{-σ.π.}$$

και θα έχουμε αποδείξει το θεώρημα. Για τον σκοπό αυτό θα χρειαστούμε το εξής:

Maximal inequality: Έστω $f^* = \sup_{n\in\mathbb{N}_0} f_n$. Τότε $\int f^* d\mu < +\infty$.

Απόδειξη. Ορίζουμε, για δοσμένο $\lambda>0$,

$$\tau(x) = \inf\{\, n\in\mathbb{N}_0 : f_n(x) > \lambda \,\}$$

όπου $\inf\emptyset = +\infty$. Τότε $\{x\in X : f^*(x) > \lambda\} = \{x\in X : \tau(x) < +\infty\} = \bigcup_{n=0}^{\infty}\{x\in X : \tau(x)=n\}$

Άρα για οποιοδήποτε $A\in\mathcal{A}$:

$$\mu(A\cap\{f^*>\lambda\}) = \sum_{n=0}^{\infty}\mu(A\cap\{\tau=n\}).$$

Τώρα αν $A\in\alpha$, τότε $f_n(x) = -\log\mu(A\mid \bigvee_{i=1}^{n} T^{-i}\alpha)$ για $x\in A$ και έτσι αν

αν $x\in A\cap\{x'\in X : f_n(x')>\lambda\}$ τότε $\mu(A\mid \bigvee_{i=1}^{n} T^{-i}\alpha)(x) < e^{-\lambda}$. Αφού $f_n(x)>\lambda$

όταν $\tau(x)=n$, παίρνουμε ότι (από την (iii) του ορισμού της δεσμευμένης μέσης τιμής)

$$\mu(A\cap\{\tau=n\}) = \int_{\{\tau=n\}}\mu(A\mid \bigvee_{i=1}^{n} T^{-i}\alpha)\,d\mu < e^{-\lambda}\mu(\{\tau=n\}).$$

Επομένως για οποιοδήποτε $A\in\alpha$ πρέπει

$$\mu(A\cap\{f^*>\lambda\}) \leq e^{-\lambda}\mu(\{\tau<+\infty\}) \leq e^{-\lambda}.$$

Αθροίζοντας ως προς $A\in\alpha$ παίρνουμε ότι $\mu(\{f^*>\lambda\}) \leq |\alpha|\, e^{-\lambda}$ όπου $|\alpha|$ συμβολίζει

τον πληθάριθμο της α και έτσι

$$\int f^* d\mu \leq |\alpha|\int_0^{\infty} e^{-\lambda}\,d\lambda = |\alpha| < +\infty.$$

(Για να αποφύγουμε να κάνουμε χρήση του $|\alpha|<+\infty$ μπορούμε να αντικαταστήσουμε

την τελευταία γραμμή με το

$$\begin{aligned}\int f^* d\mu &= \int_0^{\infty}\mu(\{f^*>\lambda\})\,d\lambda = \sum_{A\in\alpha}\int_0^{\infty}\mu(A\cap\{f^*>\lambda\})\,d\lambda \leq \\ &\leq \sum_{A\in\alpha}\int_0^{\infty}\min\{e^{-\lambda},\mu(A)\}\,d\lambda = \sum_{A\in\alpha}\Big[\int_0^{-\log\mu(A)}\mu(A)\,d\lambda + \int_{-\log\mu(A)}^{\infty} e^{-\lambda}\,d\lambda\Big] \\ &= \sum_{A\in\alpha}\big[-\mu(A)\log\mu(A) + \mu(A)\big] = H(\alpha)+1 < +\infty.\ )\end{aligned}$$

Μένει λοιπόν τώρα να αποδείξουμε την (5). Ορίζουμε $F_n = \sup_{m \geq n} |f_m - f|$, για κάθε $n \in \mathbb{N}_0$. Από το Θεώρημα του Doob για την σύγκλιση martingales (το οποίο δεν μπορέσαμε να αποφύγουμε τελικά, - βλ. και Παρατήρηση (i) στην σελίδα 13 αυτών των σημειώσεων)

$$\lim_{n\to\infty} f_n = f \quad \mu\text{-σ.π.}$$

και άρα $\lim_{n\to\infty} F_n = 0$, μ-σ.π. Τώρα $0 \leq F_n \leq F_0$ για κάθε n και επιπλέον $F_0 \leq f^* + f$. Αφού $f^* \in L^1(X,\mathcal{A},\mu)$ και $f \in L^1(X,\mathcal{A},\mu)$ ($f \geq 0$ και $\int f\, d\mu = h(T,\alpha) < +\infty$) έπεται ότι $F_0 \in L^1(X,\mathcal{A},\mu)$. Άρα $F_n \to 0$ στον L^1.

Έστω $\epsilon > 0$. Επιλέγουμε $m \in \mathbb{N}$ αρκούντως μεγάλο ώστε $\int F_m\, d\mu < \epsilon$. Τότε

(6) $$\frac{1}{n}\sum_{i=0}^{n-1} |f_{n-i-1} - f| \circ T^i = \frac{1}{n}\sum_{i=0}^{n-m-1} |f_{n-i-1} - f| \circ T^i + \frac{1}{n}\sum_{i=n-m}^{n-1} |f_{n-i-1} - f| \circ T^i$$

και $n^{-1}\sum_{i=0}^{n-m-1} |f_{n-i-1} - f| \circ T^i \leq n^{-1}\sum_{i=0}^{n-m-1} F_m \circ T^i$, για όλα τα $n \geq m+1$. Από το εργοδικό θεώρημα όμως $n^{-1}\sum_{i=0}^{n-m-1} F_m \circ T^i \to \int F_m\, d\mu$ και από την επιλογή του m παίρνουμε ότι

(7) $$\overline{\lim_{n\to\infty}}\ \frac{1}{n}\sum_{i=0}^{n-m-1} |f_{n-i-1} - f| \circ T^i < \epsilon.$$

Από την άλλη, για $n \geq m$,

(8) $$\frac{1}{n}\sum_{i=n-m}^{n-1} |f_{n-i-1} - f| \circ T^i = \frac{1}{n}\sum_{k=0}^{m-1} |f_k - f| \circ T^{n-k-1} \leq$$

$$\leq \frac{1}{n}\sum_{k=0}^{m-1} F_0 \circ T^{n-k-1} = \frac{1}{n}\sum_{k=1}^{m} F_0 \circ T^{n-k}.$$

Όμως αφού η T διατηρεί το μ,

$$\sum_{n=k}^{\infty} \mu(\{F_0 \circ T^{n-k} \geq \epsilon n\}) = \sum_{n=k}^{\infty} \mu(\{F_0 \geq \epsilon n\}) < +\infty$$

αφού $F_0 \in L^1(X,\mathcal{A},\mu)$ και άρα από το Λήμμα Borel-Cantelli $\overline{\lim}_{n\to\infty} F_0 \circ T^{n-k} < 2\epsilon$ για οποιοδήποτε k. Άρα (m είναι σταθερό) $\overline{\lim}_{n\to\infty} n^{-1}\sum_{i=0}^{m-1} F_0 \circ T^{n-k} < 2\epsilon$ και έτσι από την (8)

(9) $$\overline{\lim_{n\to\infty}}\ \frac{1}{n}\sum_{i=n-m}^{n-1} |f_{n-i-1} - f| \circ T^i < 2\epsilon.$$

Συνδυάζοντας τις (6), (7) και (9) παίρνουμε την (5). ∎

Σαν εφαρμογή του Θεωρήματος Shannon – McMillan – Breiman θα δώσουμε μια απόδειξη του τύπου του Abramov για την εντροπία του επαγόμενου μετασχηματισμού.

Εστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και $T : X \to X$ ένας αυτομορφισμός, δηλαδή ο T είναι ένας μετρήσιμος, αντιστρέψιμος (με T^{-1} μετρήσιμο) μετασχηματισμός που διατηρεί το μέτρο μ. (Η υπόθεση ότι ο T είναι αυτομορφισμός και όχι απλά ένας ενδομορφισμός, γίνεται μόνο για να απλουστεύσουμε λίγο τα πράγματα.) Εστω τώρα Y ένα $\mathcal{A}$-μετρήσιμο υποσύνολο του X με $\mu(Y) > 0$. Ορίζουμε

$$\tau_Y(x) = \inf\{n \in \mathbb{N} : T^n(x) \in Y\}$$

για $x \in \bigcup_{n=1}^{\infty} T^{-n} Y$ και $\tau_Y(x) = +\infty$ αλλού.

Υποθέτουμε τώρα επίσης ότι ο T είναι εργοδικός (και αυτή η υπόθεση γίνεται μόνο για να απλουστεύσουμε τα πράγματα). Τότε $\tau_Y(x) < +\infty$ για μ-σχεδόν κάθε x και ο "επαγόμενος μετασχηματισμός" $T_Y : Y \to Y$ που δίνεται από την σχέση

$$T_Y(x) = T^{\tau_Y(x)}(x), \quad x \in Y$$

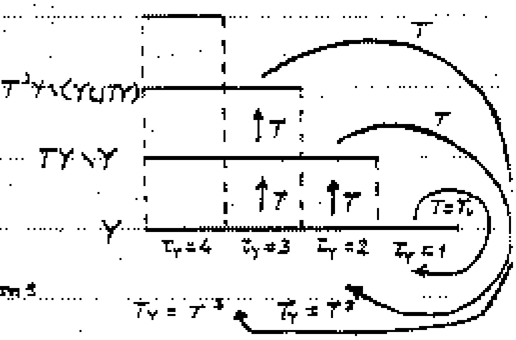
είναι καλά ορισμένος μ-σχεδόν παντού. Εστω $\mathcal{A}_Y$ η σ-άλγεβρα που αποτελείται από όλα τα σύνολα της μορφής $A \cap Y$, όπου $A \in \mathcal{A}$ και μ_Y το μέτρο

$$\mu_Y(A) = \frac{\mu(A)}{\mu(Y)}, \quad A \in \mathcal{A}_Y.$$

Τότε ο T_Y είναι ένας εργοδικός ενδομορφισμός του $(Y, \mathcal{A}_Y, \mu_Y)$, δηλαδή ο T_Y διατηρεί το μέτρο μ_Y και είναι εργοδικός. (Μια απόδειξη δίνεται στο τέλος αυτής της ενότητας.)

ΘΕΩΡΗΜΑ (ABRAMOV): $\quad h_{\mu_Y}(T_Y) = \left(\int_Y \tau_Y \, d\mu_Y\right) \cdot h_\mu(T)$

Απόδειξη: Θα υποθέσουμε εδώ ότι η συνάρτηση τ_Y είναι φραγμένη (πάνω στο Y). Εστω N το φράγμα της τ_Y και α_Y μια π.μ.δ. του Y που είναι λεπτότερη της διαμέρισης $\xi := \{\{x \in Y : \tau_Y(x) = n\} : n = 1, \ldots, N\}$. Η α_Y επάγει μια διαμέριση α του X : η α θα αποτελείται από όλα τα σύνολα της α_Y, όλα τα σύνολα TA για εκείνα τα $A \in \alpha_Y$ για τα οποία $\tau_Y(x) > 1 \ \forall x \in A$, όλα τα σύνολα $T^2 A$ για εκείνα τα $A \in \alpha_Y$ για τα οποία $\tau_Y(x) > 2 \ \forall x \in A$ κ.ο.κ. Είναι εύκολο να ελέγξει κανείς ότι η α πράγματι αποτελεί μια π.μ.δ. του X (εκτός ίσως από ένα υπόλοιπο μέτρου μηδέν). Το ότι η α είναι πεπερασμένη είναι συνέπεια της υπόθεσης ότι η συνάρτηση τ_Y είναι φραγμένη.



Έστω τώρα $A_n(x)$ το στοιχείο εκείνο της διαμέρισης $\bigvee_{i=0}^{n} T_Y^{-i}\alpha_Y$ που περιέχει το σημείο $x\in Y$. Εφαρμόζοντας το Θεώρημα Shannon-McMillan-Breiman (στον χώρο $(Y, \mathcal{A}_Y, \mu_Y, T_Y)$ για την διαμέριση α_Y) παίρνουμε ότι

$$\frac{1}{n}\log \mu_Y(A_n(x)) \longrightarrow h_{\mu_Y}(T_Y, \alpha_Y)$$

για μ_Y-σχεδόν κάθε $x\in Y$. Επειδή $\frac{1}{n}\log\mu(Y)\to 0$ έχουμε τελικά ότι

$$(*)\qquad \frac{1}{n}\log \mu(A_n(x)) \longrightarrow h_{\mu_Y}(T_Y, \alpha_Y) \qquad \mu_Y\text{-σ. κάθε } x\in Y.$$

Θέτουμε $r_n(x) = \tau_Y(x) + \tau_Y\circ T_Y(x) + \dots + \tau_Y\circ T_Y^{n}(x)$. Από το εργοδικό θεώρημα

$$(**)\qquad \frac{1}{n}\, r_n(x) \longrightarrow \int_Y \tau_Y\, d\mu_Y \qquad \mu_Y\text{-σ. κάθε } x\in Y.$$

Επιπλέον, επειδή η συνάρτηση τ_Y είναι σταθερή πάνω σε κάθε στοιχείο της διαμέρισης α_Y, η συνάρτηση r_n είναι σταθερή πάνω σε κάθε στοιχείο της $\bigvee_{i=0}^{n} T_Y^{-i}\alpha_Y$.

Έστω τώρα $A_0\cap T_Y^{-1}A_1\cap\dots\cap T_Y^{-n}A_n$ ένα στοιχείο της διαμέρισης $\bigvee_{i=0}^{n} T_Y^{-i}\alpha_Y$. Έστω k_j η τιμή της συνάρτησης τ_Y πάνω στο σύνολο A_j (υπενθυμίζουμε ότι έχουμε υποθέσει ότι $\alpha_Y \geq \xi$). Τότε η συνάρτηση r_n παίρνει την τιμή $k_0+k_1+\dots+k_n$ πάνω στο σύνολο $A_0\cap T_Y^{-1}A_1\cap\dots\cap T_Y^{-n}A_n$. Από την κατασκευή της α τα σύνολα $TA_0, T^2A_0, \dots, T^{k_0-1}A_0$ είναι στοιχεία της διαμέρισης α και προφανώς $A_0 = A_0\cap T^{-1}(TA_0)\cap\dots\cap T^{-k_0+1}(T^{k_0-1}A_0)$. Όμοια τα σύνολα $TA_1, \dots, T^{k_1-1}A_1$ είναι στοιχεία της α και $A_1 = A_1\cap T^{-1}(TA_1)\cap\dots\cap T^{-k_1+1}(T^{k_1-1}A_1)$ οπότε και $T^{-k_0}A_1 = T^{-k_0}A_1\cap T^{-(k_0+1)}(TA_1)\cap\dots\cap T^{-(k_0+k_1-1)}(T^{k_1-1}A_1)$. Επειδή όμως $A_0\cap T_Y^{-1}A_1 = A_0\cap T^{-k_0}A_1$ έπεται ότι $A_0\cap T_Y^{-1}A_1 = A_0\cap T^{-1}(TA_0)\cap\dots\cap T^{-k_0+1}(T^{k_0-1}A_0)\cap$ $\cap\, T^{-k_0}A_1\cap T^{-(k_0+1)}(TA_1)\cap\dots\cap T^{-(k_0+k_1-1)}(T^{k_1-1}A_1)$. Με όμοιο τρόπο τελικά έχουμε

$$A_0\cap T_Y^{-1}A_1\cap\dots\cap T_Y^{-n}A_n = A_0\cap T^{-1}(TA_0)\cap\dots\cap T^{-k_0+1}(T^{k_0-1}A_0)\cap$$
$$\cap\, T^{-k_0}A_1\cap T^{-k_0-1}(TA_1)\cap\dots\cap T^{-(k_0+k_1-1)}(T^{k_1-1}A_1)\cap$$
$$\vdots$$
$$\cap\, T^{-(k_0+\dots+k_{n-1})}A_n\cap T^{-(k_0+\dots+k_{n-1}+1)}(TA_n)\cap\dots\cap T^{-(k_0+\dots+k_n)}(T^{k_n-1}A$$

και το τελευταίο σύνολο ανήκει στην διαμέριση $\bigvee_{i=0}^{k_1+\dots+k_n} T^{-i}\alpha$ του X. Δηλαδή για $x\in Y$, το στοιχείο $A_n(x)$ της διαμέρισης $\bigvee_{i=0}^{n} T_Y^{-i}\alpha_Y$ του Y είναι και στοιχείο της διαμέρισης $\bigvee_{i=0}^{r} T^{-i}\alpha$ του X, όπου $r=r_n(x)$. Από το Θεώρημα Shannon-McMillan-Breiman λοιπόν, εφαρμοσμένο στον χώρο $(X,\mathcal{A},\mu,T)$ αυτή τη φορά

(***) $$\frac{1}{r_n(x)}\log\mu(A_n(x)) \longrightarrow h_\mu(T,\alpha)$$

για μ-σχεδόν κάθε $x\in Y$ και άρα και για μ_Y σχεδόν κάθε $x\in Y$. Διαλέγοντας ένα $x\in Y$ για το οποίο να ισχύουν οι (*) (**) και (***) παίρνουμε τελικά ότι

$$h_{\mu_Y}(T_Y,\alpha_Y) = h_\mu(T,\alpha)\int_Y r_Y\,d\mu_Y$$

Αφού η σχέση αυτή ισχύει κάθε $\alpha_Y \in \xi$ και αφού για τυχούσα π.μ.δ. β του Y έχουμε ότι $\alpha_Y := \beta\vee\xi \geq \xi$, έπεται ότι

$$h_{\mu_Y}(T_Y,\beta)\le h_{\mu_Y}(T_Y,\alpha_Y) = h_\mu(T,\alpha)\int_Y r_Y\,d\mu_Y \le h_\mu(T)\int_Y r_Y\,d\mu_Y$$

για οποιαδήποτε π.μ.δ. β του Y. Άρα

(****) $$h_{\mu_Y}(T_Y)\le h_\mu(T)\int_Y r_Y\,d\mu_Y .$$

Για την αντίστροφη ανισότητα ξεκινάμε με μια π.μ.δ. β του X. Τότε υπάρχει μια διαμέριση α_Y του Y τέτοια ώστε αν α είναι η διαμέριση του X που προκύπτει από την α_Y όπως στο πρώτο μέρος αυτής της απόδειξης (βλέπε τέλος σελίδας 32) τότε $\alpha\geq\beta$. Η α_Y μπορεί να κατασκευαστεί ως εξής: ονομάζουμε ξ_i την κοινή εκλέπτυνση των εξής δύο διαμερίσεων του $T^iY\setminus(Y\cup TY\cup\dots\cup T^{i-1}Y)$: της διαμέρισης β_i που επάγει η β στο σύνολο $T^iY\setminus(Y\cup TY\cup\dots\cup T^{i-1}Y)$ (δηλ. μη κενές τομές στοιχείων της β με το σύνολο αυτό) και της διαμέρισης $\{T^i(\{r_Y=n\}) : i<n\le N\}$. Κατόπιν θέτουμε $\xi_0\vee T^{-1}\xi_1\vee\dots\vee T^{-(N-1)}\xi_{N-1} = \alpha_Y$. Όπως και στο πρώτο μέρος $h_{\mu_Y}(T_Y,\alpha_Y)=h_\mu(T,\alpha)\int_Y r_Y\,d\mu_Y$. Τώρα αφού $\beta\le\alpha$ έπεται ότι $h_\mu(T,\beta)\le h_\mu(T,\alpha)$ και άρα

$$h_\mu(T,\beta)\cdot\int_Y r_Y\,d\mu_Y \le h_{\mu_Y}(T_Y)$$

για οποιαδήποτε π.μ.δ β του X. Από αυτή παίρνουμε

$$h_\mu(T)\int_Y r_Y\,d\mu_Y\le h_{\mu_Y}(T_Y)$$

που μαζί με την (****) αποδεικνύουν το θεώρημα. ∎

ΠΑΡΑΤΗΡΗΣΗ: Ο λόγος που κάναμε την υπόθεση τ_Y φραγμένη στην παραπάνω απόδειξη ήταν για να είναι όλες οι διαμερίσεις μας πεπερασμένες. Από τα λίγα που θα αναφέρουμε στην επόμενη ενότητα για (άπειρες) αριθμήσιμες διαμερίσεις θα πρέπει να γίνει φανερό ότι η απόδειξη αυτή δουλεύει και χωρίς την υπόθεση τ_Y φραγμένη.

Θα αποδείξουμε τώρα τον ισχυρισμό της σελίδας 32 ότι δηλαδή ο T_Y είναι ένας εργοδικός ενδομορφισμός του $(Y, \mathcal{A}_Y, \mu_Y)$.

ΛΗΜΜΑ: Ο επαγόμενος μετασχηματισμός T_Y διατηρεί το μέτρο μ_Y.

Απόδειξη: Έστω $A \in \mathcal{A}_Y$. Θέτουμε $A_n = \{x \in A : \tau_Y \circ T_Y^{-1}(x) = n\}$ για $n \in \mathbb{N}$. (Ο T_Y είναι προφανώς αντιστρέψιμος εφόσον ο T είναι αντιστρέψιμος.) Παρατηρούμε δε ότι $T_Y^{-1} A_n = T^{-n} A_n$. Επιπλέον τα A_n είναι προφανώς ξένα ανά δύο και $\bigcup_{n=1}^{\infty} A_n = A$. Επίσης τα $T_Y^{-1} A_n$ είναι ξένα ανά δύο και $\bigcup_{n=1}^{\infty} T_Y^{-1} A_n = T_Y^{-1} A$. Τέλος $T_Y^{-1} A_n \subseteq Y$ και άρα $\mu_Y(T_Y^{-1} A_n) = \mu(T_Y^{-1} A_n) / \mu(Y)$. Έτσι

$$\mu_Y(T_Y^{-1} A) = \sum_{n=1}^{\infty} \mu_Y(T_Y^{-1} A_n) = \frac{1}{\mu(Y)} \sum_{n=1}^{\infty} \mu(T_Y^{-1} A_n) =$$

$$= \frac{1}{\mu(Y)} \sum_{n=1}^{\infty} \mu(T^{-n} A_n) = \frac{1}{\mu(Y)} \sum_{n=1}^{\infty} \mu(A_n) = \frac{\mu(A)}{\mu(Y)} = \mu_Y(A). \quad \blacksquare$$

ΛΗΜΜΑ: Αν ο T είναι εργοδικός τότε και ο επαγόμενος μετασχηματισμός T_Y είναι.

Απόδειξη: Έστω $A \in \mathcal{A}_Y$ τέτοιο ώστε $T_Y^{-1} A = A$. Έστω ότι $\mu_Y(A) > 0$. Τότε $\mu(A) > 0$. Τώρα αφού $A \subseteq Y$ προφανώς $\mathbb{1}_A \leq \mathbb{1}_Y$ σ' όλον του X και άρα $n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_A \circ T^i \leq n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_Y \circ T^i$. Επειδή όμως $T_Y A = A$ για μ-σχεδόν κάθε $x \in A$ έχουμε $n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_A \circ T^i = n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_Y \circ T^i$. Από το εργοδικό θεώρημα για μ-σχεδόν κάθε $x \in X$, $n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_A \circ T^i \to \mu(A)$ και $n^{-1} \sum_{i=0}^{n-1} \mathbb{1}_Y \circ T^i \to \mu(Y)$. Αφού $\mu(A) > 0$ πρέπει $\mu(A) = \mu(Y)$ ή ισοδύναμα $\mu_Y(A) = 1$. $\blacksquare$

6. ΑΡΙΘΜΗΣΙΜΕΣ ΔΙΑΜΕΡΙΣΕΙΣ

Η εντροπία $H(\xi)$ και η δεσμευμένη εντροπία $H(\xi|\mathcal{F})$ ορίζεται και για αυθαίρετες αριθμήσιμες διαμερίσεις $\xi \subseteq \mathcal{A}$. Η Πρόταση 1.1, η Πρόταση 1.2 και η Πρόταση 1.3 ισχύουν και για (άπειρες) αριθμήσιμες διαμερίσεις. Επίσης ισχύει το Θεώρημα 2.2 που ορίζει την εντροπία $h(T,\xi)$ με ξ μια αυθαίρετη αριθμήσιμη διαμέριση του X (από μετρήσιμα σύνολα). Για αυθαίρετες αριθμήσιμες διαμερίσεις ξ η εντροπία $H(\xi)$ μπορεί να πάρει και την τιμή $+\infty$. Στην περίπτωση αυτή φυσικά $h(T,\xi)=+\infty$. Επίσης ισχύουν η Πρόταση 1.4 και το Πόρισμα 2.4 για αριθμήσιμες διαμερίσεις ξ με $H(\xi)<+\infty$. Τέλος, ισχύει η Πρόταση 2.6 που αφορά ιδιότητες της $h(T,\xi)$.

Η εντροπία ενός ενδομορφισμού έχει οριστεί σαν $h(T)=\sup h(T,\alpha)$ όπου το supremum είναι ως προς όλες τις πεπερασμένες διαμερίσεις $\alpha \subseteq \mathcal{A}$. Ισχύει όμως και το εξής:

6.1 ΠΡΟΤΑΣΗ: Εστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $T: X \to X$ ένας ενδομορφισμός.

$$h(T)=\sup\{h(T,\xi):\ \xi \subseteq \mathcal{A} \text{ αριθμήσιμη διαμέριση με } H(\xi)<+\infty\}.$$

Απόδειξη: Εστω $h'(T)$ το δεξί μέλος. Τότε προφανώς $h'(T) \geq h(T)$. Για την αντίστροφη ανισότητα, έστω ξ μια αριθμήσιμη διαμέριση με $H(\xi)<+\infty$. Θέτουμε $\xi_n=\{A_1,A_2,\dots,A_n,\ \bigcup_{i>n}A_i\}$ για $n=1,2,\dots$ όπου $A_1,A_2,\dots$ είναι μια αρίθμηση των στοιχείων της ξ. Τότε προφανώς $\xi_n \leq \xi$ και άρα και $\bigvee_{i=1}^{\infty}T^{-i}\xi_n \leq \bigvee_{i=1}^{\infty}T^{-i}\xi$.

Επομένως

$$h(T,\xi_n)=H(\xi_n \mid \bigvee_{i=1}^{\infty}T^{-i}\xi_n) \geq H(\xi_n \mid \bigvee_{i=1}^{\infty}T^{-i}\xi)$$

και άρα

$$\lim_{n\to\infty} h(T,\xi_n) \geq \lim_{n\to\infty} H(\xi_n \mid \bigvee_{i=1}^{\infty}T^{-i}\xi)$$

(τα όρια προφανώς υπάρχουν αφού $\xi_1 \leq \xi_2 \leq \dots$) Εστω $\mathcal{F}$ η σ-άλγεβρα $\bigvee_{i=1}^{\infty}T^{-i}\xi$. Τότε $I(\xi_n|\mathcal{F}) \to I(\xi|\mathcal{F})$, μ-σ.π. Πράγματι, αν $x \in A_k$ τότε $I(\xi_n|\mathcal{F})(x)=-\log\mu(A_k|\mathcal{F})(x)=I(\xi|\mathcal{F})(x)$ για όλα τα $n \geq k$. Επίσης, για $k>n$, έχουμε $I(\xi_n|\mathcal{F})(x)=-\log\mu(\bigcup_{i>n}A_i|\mathcal{F})(x) \leq -\log\mu(A_k|\mathcal{F})(x)=I(\xi|\mathcal{F})(x)$. Άρα για

$I(\xi_n | \mathcal{F}) \le I(\xi | \mathcal{F})$ και από το θεώρημα κυριαρχημένης σύγκλισης, αφού $\int I(\xi|\mathcal{F}) d\mu = H(\xi|\mathcal{F}) \le H(\xi) < +\infty$, $H(\xi_n | \mathcal{F}) \to H(\xi | \mathcal{F})$. (Εναλλακτικά βλέπει κανείς ότι $I(\xi_n | \mathcal{F}) \le I(\xi_{n+1} | \mathcal{F})$ και χρησιμοποιεί το θεώρημα μονότονης σύγκλισης.) Επομένως $\lim_{n\to\infty} h(T, \xi_n) \ge \lim_{n\to\infty} H(\xi_n | V_{i=1}^{\infty} T^{-i}\xi) = H(\xi | V_{i=1}^{\infty} \xi)$ και λόγω μονοτονίας $h(T, \xi_n) \nearrow h(T, \xi)$.

Από την προηγούμενη παράγραφο, αν η ξ είναι μια αριθμήσιμη διαμέριση με $H(\xi) < +\infty$, τότε η $h(T, \xi)$ προσεγγίζεται από $h(T, \xi_n)$ για ξ_n πεπερασμένες διαμερίσεις και έτσι τελικά $h(T) = h'(T)$. ∎

Μια αριθμήσιμη διαμέριση ξ του X είναι ένας γεννήτορας για τον αυτομορφισμό T αν $V_{n=-\infty}^{\infty} T^{-n}\xi \doteq \mathcal{A}$. Ισχύει και πάλι το Θεώρημα Kolmogorov-Sinai: Αν T είναι ένας αυτομορφισμός και ξ ένας γεννήτορας με $H(\xi) < +\infty$ τότε $h(T) = h(T, \xi)$.

Το βασικό θεώρημα για την ύπαρξη γεννητόρων ξ με $H(\xi) < \infty$ είναι το εξής:

ΘΕΩΡΗΜΑ (ROHLIN): Εστω $(X, \mathcal{A}, \mu)$ ένας χώρος Lebesgue (ορισμός στην επόμενη ενότητα) μη ισόμορφος (μετροθεωρητικά) μ' ένα πεπερασμένο σύνολο. Εστω T ένας αυτομορφισμός του $(X, \mathcal{A}, \mu)$. Τότε ο T έχει έναν γεννήτορα ξ με $H(\xi) < +\infty$ αν κ' μόνο αν $h(T) < +\infty$ και ο T είναι μη περιοδικός (δηλ.: $\mu\{x \in X ;\ T^n(x) = x$ για κάποιο $n \in \mathbb{Z} \setminus \{0\}\} = 0$).

Τέλος αναφέρουμε το επόμενο σημαντικό θεώρημα.

ΘΕΩΡΗΜΑ (KRIEGER): Εστω T ένας αυτομορφισμός ενός χώρου Lebesgue $(X, \mathcal{A}, \mu)$. Αν $h(T) < +\infty$ τότε ο T έχει έναν πεπερασμένο γεννήτορα α. Μπορούμε δε να διαλέξουμε τον α έτσι ώστε το πλήθος των στοιχείων του $|\alpha|$ να είναι στο διάστημα $[e^{h(T)}, e^{h(T)} + 1]$.

Το Θεώρημα Shannon-McMillan-Breiman ισχύει κ' για αριθμήσιμες διαμερίσεις ξ με $H(\xi) < +\infty$. Η απόδειξη είναι όπως και για πεπερασμένες διαμερίσεις.

ΠΑΡΑΡΤΗΜΑ

Δύο χώροι πιθανότητας $(X_1, \mathcal{A}_1, \mu_1)$ και $(X_2, \mathcal{A}_2, \mu_2)$ λέγονται ισόμορφοι αν υπάρχουν σύνολα $M_i \in \mathcal{A}_i$ με $\mu_i(M_i) = 1$ και ένας αντιστρέψιμος μετασχηματισμός $\phi : M_1 \to M_2$ που διατηρεί τα μέτρα μ_1 και μ_2 δηλ. $\mu_1(\phi^{-1}A) = \mu_2(A)$ για κάθε $A \in \mathcal{A}_2$, με $A \subseteq M_2$. (Οι χώροι M_i θεωρούνται εφοδιασμένοι με τις σ-άλγεβρες $\{A \cap M_i : A \in \mathcal{A}_i\}$.)

ΘΕΩΡΗΜΑ: Έστω X ένας πλήρης διαχωρίσιμος μετρικός χώρος, $\mathcal{B}(X)$ η Borel σ-άλγεβρα του X και μ ένα μέτρο πιθανότητας ορισμένο στην $\mathcal{B}(X)$. Αν $\mu(\{x\}) = 0$ για κάθε $x \in X$ τότε ο χώρος $(X, \mathcal{B}(X), \mu)$ είναι ισόμορφος με τον $([0,1], \mathcal{B}([0,1]), \lambda)$ όπου λ το μέτρο Lebesgue και η πλήρωση $(X, \overline{\mathcal{B}(X)}, \mu)$ του χώρου $(X, \mathcal{B}(X), \mu)$ είναι ισόμορφη με τον $([0,1], \mathcal{L}([0,1]), \lambda)$, όπου $\mathcal{L}([0,1]) = \overline{\mathcal{B}([0,1])}$ είναι τα Lebesgue-μετρήσιμα υποσύνολα του $[0,1]$ (πλήρωση του $\mathcal{B}([0,1])$). Διαφορετικά, υπάρχει ένα αριθμήσιμο πλήθος σημείων (πεπερασμένο ή άπειρο) $\{x_i\}$ με $\mu(\{x_i\}) > 0$ και ο χώρος $(X, \mathcal{B}(X), \mu)$ είναι ισόμορφος μ' έναν χώρο που αποτελείται από σημεία $\{y_i\}$ με μέτρα $\mu(\{x_i\})$ μαζί με τον $([0,s], \mathcal{B}([0,s]), \lambda)$, όπου $s = 1 - \sum_i \mu(\{x_i\})$. Και στην περίπτωση αυτή ισχύει ανάλογο αποτέλεσμα για την πλήρωση του χώρου $(X, \mathcal{B}(X), \mu)$.

Μια απόδειξη του Θεωρήματος αυτού υπάρχει στο βιβλίο του Royden (Θεώρημα 9, σελ. 327).

Πολλές φορές στην εργοδική θεωρία θεωρούμε χώρους Lebesgue: Ένας χώρος πιθανότητας $(X, \mathcal{A}, \mu)$ είναι ένας χώρος Lebesgue αν είναι ισόμορφος με έναν χώρο πιθανότητας που είναι ξένη ένωση ενός αριθμήσιμου χώρου $\{y_i\}$ (πεπερασμένου ή απείρου, πιθανώς κενού) με μέτρο του $y_i = p_i > 0$ και του χώρου $([0,s], \mathcal{L}([0,s]), \lambda)$, όπου πάλι $s = 1 - \sum_i p_i$ και $\mathcal{L}([0,s])$ τα Lebesgue μετρήσιμα υποσύνολα του $[0,s]$.

Από το προηγούμενο θεώρημα, αν X είναι διαχωρίσιμος πλήρης μετρικός χώρος και $\mathcal{A}$ η πλήρωση του $\mathcal{B}(X)$, ως προς κάποιο μέτρο πιθανότητας μ στην $\mathcal{B}(X)$, τότε ο $(X, \mathcal{A}, \mu)$ είναι ένας χώρος Lebesgue.

Έστω τώρα $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας. Ορίζουμε μια σχέση ισοδυναμίας $\sim$ στην $\mathcal{A}$ λέγοντας ότι $A \sim B$ αν $\mu(A \triangle B) = 0$. Έστω $\tilde{\mathcal{A}}$ η οικογένεια των κλάσεων ισοδυναμίας. Τότε η $\tilde{\mathcal{A}}$ είναι μια σ-άλγεβρα Boole με πράξεις το συμπλήρωμα, την ένωση και την τομή επαγώμενες από την $\mathcal{A}$. Το μέτρο μ επάγει ένα μέτρο $\tilde{\mu}$ επί της $\tilde{\mathcal{A}}$ μέσω της σχέσης $\tilde{\mu}(\tilde{A}) = \mu(A)$ για οποιονδήποτε αντιπρόσωπο A της κλάσης $\tilde{A}$. Το ζεύγος $(\tilde{\mathcal{A}}, \tilde{\mu})$ λέγεται measure algebra.

Δύο χώροι πιθανότητας $(X_1, \mathcal{A}_1, \mu_1)$ και $(X_2, \mathcal{A}_2, \mu_2)$ λέγονται συζυγείς αν οι αντίστοιχες measure algebras $(\tilde{\mathcal{A}}_1, \tilde{\mu}_1)$ και $(\tilde{\mathcal{A}}_2, \tilde{\mu}_2)$ είναι ισόμορφες, δηλαδή υπάρχει μια 1-1 και επί απεικόνιση $\Phi : \tilde{\mathcal{A}}_2 \to \tilde{\mathcal{A}}_1$ πού διατηρεί συμπληρώματα, αριθμήσιμες ενώσεις και τομές και που είναι τέτοια ώστε $\tilde{\mu}_1(\Phi \tilde{A}) = \tilde{\mu}_2(\tilde{A})$ για κάθε $\tilde{A} \in \tilde{\mathcal{A}}_2$.

Δύο ισόμορφοι χώροι $(X_1, \mathcal{A}_1, \mu_1)$ και $(X_2, \mathcal{A}_2, \mu_2)$ είναι και συζυγείς: αν $\phi : M_1 \to M_2$ είναι ο (μετροθεωρητικός) ισομορφισμός τότε η $\Phi(\tilde{A}) = \widetilde{(\phi^{-1}(A \cap M_2)}$ ορίζει έναν ισομορφισμό για measure algebras. Για τα αντίστροφα έχουμε το εξής:

ΘΕΩΡΗΜΑ: Έστω, για $i = 1, 2$, X_i ένας διαχωρίσιμος πλήρης μετρικός χώρος, $\mathcal{B}(X_i)$ η Borel σ-άλγεβρα του X_i και μ_i ένα μέτρο πιθανότητας επί της $\mathcal{B}(X_i)$. Έστω $\Phi : \tilde{\mathcal{A}}_2 \to \tilde{\mathcal{A}}_1$ ένας ισομορφισμός για measure algebras. Τότε υπάρχουν $M_1 \in \mathcal{B}(X_1)$ και $M_2 \in \mathcal{B}(X_2)$, με $\mu_i(M_i) = 1$, $i = 1, 2$ και μια αντιστρέψιμη απεικόνιση $\phi : M_1 \to M_2$ που διατηρεί τα μέτρα μ_1, μ_2 τέτοια ώστε $\Phi(\tilde{A}) = \widetilde{(\phi^{-1}(A \cap M_2))}$. Επιπλέον, αν ψ είναι ένας άλλος ισομορφισμός των $(X_1, \mathcal{B}(X_1), \mu_1)$ και $(X_2, \mathcal{B}(X_2), \mu_2)$ που επάγει τον Φ τότε $\mu_1(\{x \in X_1 : \phi(x) \neq \psi(x)\}) = 0$.

Μια απόδειξη αυτού του θεωρήματος υπάρχει στην σελίδα 329 (Θεώρημα 12) του βιβλίου του Royden. Επίσης ισχύει το ανάλογο αποτέλεσμα για χώρους Lebesgue.

Μπορούμε τώρα να αποδείξουμε την Παρατήρηση της σελίδας 18.

ΠΡΟΤΑΣΗ: Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος Lebesgue ή ένας χώρος πιθανότητας όπου X είναι διαχωρίσιμος πλήρης μετρικός χώρος και $\mathcal{A} = \mathcal{B}(X)$. Έστω T ένας ενδομορφισμός του $(X, \mathcal{A}, \mu)$. Τότε ο T είναι σχεδόν αντιστρέψιμος αν κ' μόνο αν $T^{-1}\mathcal{A} \doteq \mathcal{A}$.

Απόδειξη: Έστω $(\tilde{\mathcal{A}}, \mu)$ η measure algebra που επάγει η $\mathcal{A}$ και το μ. Έστω $\tilde{T} : \tilde{\mathcal{A}} \to \tilde{\mathcal{A}}$ ορισμένη από την $\tilde{T}(\tilde{A}) = \widetilde{(T^{-1}A)}$. Τότε η $\tilde{T}$ είναι 1-1.

Αν ο T είναι σχεδόν αντιστρέψιμος τότε επάγει πάνω στην measure algebra την ίδια απεικόνιση που επάγει και ο αντιστρέψιμος μετασχηματισμός. Έτσι $\tilde{T}\tilde{\mathcal{A}} = \tilde{\mathcal{A}}$.

(Αυτό ισχύει χωρίς καμμία υπόθεση για την μορφή του χώρου $(X, \mathcal{A}, \mu)$ φυσικά.)

Αν τώρα $\tilde{T}\tilde{\mathcal{A}} = \tilde{\mathcal{A}}$ τότε η απεικόνιση $\tilde{T} : \tilde{\mathcal{A}} \to \tilde{\mathcal{A}}$ είναι 1-1 και επί και άρα πρέπει να επάγεται από έναν αντιστρέψιμο μετασχηματισμό ορισμένο σ' ένα υποσύνολο του X μέτρου 1, από το προηγούμενο θεώρημα. Άρα ο T είναι ίσος σχεδόν παντού μ' έναν αντιστρέψιμο μετασχηματισμό. ∎

IV. ΤΟΠΟΛΟΓΙΚΗ ΕΝΤΡΟΠΙΑ

1. Εντροπία των ανοιχτών καλυμμάτων

Εστω X ένας συμπαγής χώρος Hausdorff. Αν α, β είναι δύο ανοιχτά καλύμματα του X, τότε το $\alpha\vee\beta=\{A\cap B : A\in\alpha, B\in\beta\}$ είναι ανοιχτό κάλυμμα του X. Αν το β είναι εκλέπτυνση στου α γράφουμε $\beta\succ\alpha$ ή $\alpha\prec\beta$. Η σχέση $\succ$ είναι ως γνωστόν κατεύθυνση στο σύνολο όλων των ανοιχτών καλυμμάτων του X. Αν $T : X\to X$ είναι μια συνεχής απεικόνιση, τότε το $T^{-1}(\alpha)=\{T^{-1}(A) : A\in\alpha\}$ είναι επίσης ανοιχτό κάλυμμα του X. Προφανώς

$$T^{-1}(\alpha\vee\beta)=T^{-1}(\alpha)\vee T^{-1}(\beta) \quad \text{και}$$

$$T^{-1}(\alpha)\prec T^{-1}(\beta), \text{ όταν } \alpha\prec\beta$$

Επειδή ο X είναι συμπαγής, κάθε ανοιχτό κάλυμμα του X έχει πεπερασμένο υποκάλυμμα. Συμβολίζουμε με $N(\alpha)$ το ελάχιστο πλήθος στοιχείων του ανοιχτού καλύμματος α που απαιτούνται για να καλύψουν τον X. Δηλαδή,

$$N(\alpha)=\min\{|\beta| : \text{το } \beta \text{ είναι πεπερασμένο υποκάλυμμα του } \alpha\}$$

1.1 Λήμμα. Εστωσαν α, β δύο ανοιχτά καλύμματα του X. Τότε:

(i) $N(\alpha)\geq 1$ και $N(\alpha)=1$ τότε και μόνον τότε όταν $X\in\alpha$.

(ii) Αν $\alpha\prec\beta$, τότε $N(\alpha)\leq N(\beta)$

(iii) $N(\alpha\vee\beta)\leq N(\alpha)\cdot N(\beta)$

(iv) Αν $T: X\to X$ είναι μία συνεχής απεικόνιση, τότε $N(T^{-1}(\alpha))\leq N(\alpha)$. Αν η T είναι "επί", τότε $N(T^{-1}(\alpha))=N(\alpha)$.

Απόδειξη. Το (i) είναι τετριμμένο.

(ii) Εστω $\{B_1,\dots,B_{N(\beta)}\}$ ένα ελάχιστο πεπερασμένο υποκάλυμμα του β. Τότε για κάθε $1\leq k\leq N(\beta)$, υπάρχει $A_k\in\alpha$ ώστε $B_k\subset A_k$. Συνεπώς, $X=A_1\cup\dots\cup A_{N(\beta)}$ και άρα $N(\alpha)\leq N(\beta)$.

(iii) Αν $\{A_1,\dots,A_{N(\alpha)}\}$ και $\{B_1,\dots,B_{N(\beta)}\}$ είναι ελάχιστα πεπερασμένα υποκαλύμματα των α και β αντίστοιχα, τότε το $\{A_i\cap B_j : 1\leq i\leq N(\alpha), 1\leq j\leq N(\beta)\}$ είναι πεπερασμένο υποκάλυμμα του $\alpha\vee\beta$. Άρα $N(\alpha\vee\beta)\leq N(\alpha)N(\beta)$.

(iv) Αν $\{A_1,\dots,A_{N(\alpha)}\}$ είναι ένα ελάχιστο πεπερασμένο υποκάλυμμα του α, τότε το $\{T^{-1}(A_1),\dots,T^{-1}(A_{N(\alpha)})\}$ είναι πεπερασμένο υποκάλυμμα του $T^{-1}(\alpha)$ και συνεπώς $N(T^{-1}(\alpha)) \leq N(\alpha)$. Αν η T είναι "επί" και $\{T^{-1}(A_1),\dots,T^{-1}(A_{N(T^{-1}(\alpha))})\}$ είναι ένα ελάχιστο πεπερασμένο υποκάλυμμα του $T^{-1}(\alpha)$, τότε το πεπερασμένο σύνολο $\{A_1,\dots,A_{N(T^{-1}(\alpha))}\} \subset \alpha$ επίσης καλύπτει τον X. Άρα $N(T^{-1}(\alpha)) = N(\alpha)$.

__1.2. Λήμμα.__ Έστω $(a_n)_{n\in\mathbb{N}}$ μια ακολουθία μη-αρνητικών πραγματικών αριθμών ώστε $a_{n+m} \leq a_n + a_m$ για κάθε $n,m \in \mathbb{N}$. Τότε το $\lim_{n\to+\infty} \frac{a_n}{n}$ υπάρχει και ισούται με $a = \inf\{\frac{a_n}{n} : n\in\mathbb{N}\}$.

__Απόδειξη__ Θα αποδείξουμε το συμπέρασμα με απαγωγή στο άτοπο. Έστω ότι υπάρχει $\varepsilon>0$ και μια υπακολουθία $(a_{n_k})_{k\in\mathbb{N}}$ ώστε $a+\varepsilon \leq \frac{a_{n_k}}{n_k}$ για κάθε $k\in\mathbb{N}$. Υπάρχει $m\in\mathbb{N}$ ώστε

$$a \leq \frac{a_m}{m} < a+\varepsilon \leq \frac{a_{n_k}}{n_k} \quad , k\in\mathbb{N}.$$

Για κάθε $k\in\mathbb{N}$ με $n_k \geq m$, υπάρχουν $q_k\in\mathbb{N}$ και $0\leq \lambda_k < m$ ώστε $n_k = q_k m + \lambda_k$. Συνεπώς,

$$\frac{a_{n_k}}{n_k} = \frac{a_{q_k m+\lambda_k}}{q_k m+\lambda_k} \leq \frac{a_{q_k m} + a_{\lambda_k}}{q_k m+\lambda_k} \leq \frac{q_k a_m + a_{\lambda_k}}{q_k m+\lambda_k} \leq \frac{q_k a_m}{q_k m} + \frac{a_{\lambda_k}}{q_k m}.$$

Άρα $0 < a+\varepsilon - \frac{a_m}{m} \leq \frac{a_{n_k}}{n_k} - \frac{a_m}{m} \leq \frac{1}{q_k m}\cdot \max\{a_1,\dots,a_{m-1}\}$

Προφανώς $\lim_{k\to+\infty} q_k = +\infty$ και συνεπώς παίρνοντας το όριο στην προηγούμενη ανισότητα έχουμε $0 < a+\varepsilon - \frac{a_m}{m} \leq 0$, αντίφαση.

__1.3. Θεώρημα.__ Αν α είναι ένα ανοιχτό κάλυμμα του X και $T: X\to X$ είναι μια συνεχής απεικόνιση, τότε το

$$h(T,\alpha) = \lim_{n\to+\infty} \frac{1}{n} \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha))$$

υπάρχει στο $\mathbb{R}$ και λέγεται εντροπία της T ως προς το ανοιχτό κάλυμμα α.

__Απόδειξη__ Θέτουμε $a_n = \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha)) \geq 0$. Από το Λήμμα 1.2 αρκεί να δείξουμε ότι $a_{n+m} \leq a_n + a_m$ για κάθε $n,m\in\mathbb{N}$. Έχουμε

$$\begin{aligned} a_{n+m} &= \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n-m+1}(\alpha)) \\ &\leq \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha)) + \log N(T^{-n}(\alpha) \vee \dots \vee T^{-n-m+1}(\alpha)) \\ &= a_n + \log N(T^{-n}(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-m+1}(\alpha))) \\ &\leq a_n + \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-m+1}(\alpha)) \\ &= a_n + a_m. \quad \text{ο.ε.δ.} \end{aligned}$$

<u>1.4. Λήμμα.</u> Έστω $T : X \to X$ μια συνεχής απεικόνιση και α, β δυο ανοιχτά καλύματα. Τότε ισχύουν τα ακόλουθα:

(i) $h(T,\alpha) \geq 0$.

(ii) Αν $\alpha \prec \beta$, τότε $h(T,\alpha) \leq h(T,\beta)$.

(iii) $h(T,\alpha) \leq \log N(\alpha)$.

<u>Απόδειξη</u> Το (i) είναι τετριμμένο.

(ii) Αν $\alpha \prec \beta$, τότε $\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-n+1}(\alpha) \prec \beta \vee T^{-1}(\beta) \vee \cdots \vee T^{-n+1}(\beta)$ για κάθε $n \in \mathbb{N}$ και το συμπέρασμα είναι άμεσο από το Λήμμα 1.1 (ii).

(iii) Από το Λήμμα 1.1 (iii) έχουμε για κάθε $n \in \mathbb{N}$

$$N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-n+1}(\alpha)) \leq \prod_{k=0}^{n-1} N(T^{-k}(\alpha)) \leq (N(\alpha))^n$$

και το συμπέρασμα είναι άμεσο.

Έτσι για κάθε συνεχή απεικόνιση $T : X \to X$ το

$$h(T) = \sup\{h(T,\alpha) : \alpha \text{ ανοιχτό κάλυμμα του } X\}$$
$$= \sup\{h(T,\alpha) : \alpha \text{ πεπερασμένο ανοιχτό κάλυμμα του } X\}$$

υπάρχει και είναι ένας μη-αρνητικός αριθμός. Ο $h(T)$ λέγεται <u>τοπολογική εντροπία</u> της T. Προφανώς δυο τοπολογικά συζυγείς απεικονίσεις έχουν την ίδια τοπολογική εντροπία. Επίσης $h(id_X) = 0$

<u>1.5. Λήμμα.</u> Αν ο $T : X \to X$ είναι ομοιομορφισμός τότε $h(T) = h(T^{-1})$.

<u>Απόδειξη.</u> Από το Λήμμα 1.1 (iv), για κάθε ανοιχτό κάλυμμα α του X έχουμε $N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-n+1}(\alpha)) = N(T^{n-1}(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-n+1}(\alpha)) =$ $N(\alpha \vee \cdots \vee (T^{-1})^{-n+1}(\alpha))$ για κάθε $n \in \mathbb{N}$ και το συμπέρασμα είναι άμεσο.

<u>1.6. Θεώρημα.</u> Αν $T : X \to X$ είναι ένας ομοιομορφισμός τότε

$$h(T^k) = |k| \cdot h(T) \quad \text{για κάθε } k \in \mathbb{Z}.$$

<u>Απόδειξη</u> Έστω $k \in \mathbb{N}$. Για κάθε ανοιχτό κάλυμμα α του X, αν θέσουμε $\beta = \alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-k+1}(\alpha)$, έχουμε:

$$h(T^k) \geq h(T^k, \beta) = \lim_{n \to +\infty} \frac{1}{n} \log N(\beta \vee T^{-k}(\beta) \vee \cdots \vee T^{-nk+k}(\beta))$$

$$= \lim_{n \to +\infty} \frac{1}{n} \log N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-nk+1}(\alpha))$$

$$= k \cdot \lim_{n \to +\infty} \frac{1}{nk} \log N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-nk+1}(\alpha)) = k\, h(T,\alpha)$$

Άρα $h(T^k) \geq k\, h(T)$.

Από το Λήμμα 1.1 (ii) έχουμε από την άλλη μεριά

$$N(\alpha \vee T^{-k}(\alpha) \vee \cdots \vee T^{-nk+k}(\alpha)) \leq N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-nk+1}(\alpha))$$

και συνεπώς

$$h(T^k,\alpha) = \lim_{n\to+\infty} \frac{1}{n} \log N(\alpha \vee T^{-k}(\alpha) \vee \cdots \vee T^{-nk+k}(\alpha))$$

$$\leq k \cdot \lim_{n\to+\infty} \frac{1}{nk} \log N(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-nk+1}(\alpha))$$

$$= k h(T,\alpha) \leq k h(T).$$

Άρα $h(T^k) \leq k h(T)$. Δείξαμε λοιπόν ότι $h(T^k) = k h(T)$ για κάθε $k \in \mathbb{Z}^+$. Από το Λήμμα 1.5 για κάθε $k \in \mathbb{N}$ έχουμε τώρα

$$h(T^{-k}) = h((T^{-1})^k) = k\, h(T^{-1}) = |-k|\, h(T).$$

Το Θεώρημα αποδείχθει.

2. Τοπολογική εντροπία και μετρικές.

Εστω X ένας συμπαγής μετρικοποιήσιμος χώρος και d μία συμβιβαστή μετρική του X, δηλαδή η d παράγει την τοπολογία του X. Εστω $T: X \to X$ μια συνεχής απεικόνιση. Εστω $n \in \mathbb{N}$ και $\varepsilon > 0$. Ενα σύνολο $F \subset X$ λέμε ότι <u>(n,ε)-παράγει</u> τον X (ως προς T) αν για κάθε $x \in X$ υπάρχει $y \in F$ ώστε $d(T^k(x), T^k(y)) \leq \varepsilon$ για κάθε $0 \leq k \leq n-1$.

<u>2.1. Λήμμα.</u> Εστω $T: X \to X$ μια συνεχής απεικόνιση. Για κάθε $n \in \mathbb{N}$ και $\varepsilon > 0$ υπάρχει ένα πεπερασμένο σύνολο $F \subset X$ που (n,ε)-παράγει τον X με $|F| \leq m^n$, όπου το $m \in \mathbb{N}$ εξαρτάται μόνο από το $\varepsilon > 0$.

<u>Απόδειξη.</u> Επειδή ο X είναι συμπαγής, έχει ένα πεπερασμένο ανοιχτό κάλυμμα $\{V_1, \ldots, V_m\}$ με $\mathrm{diam}(V_i) \leq \varepsilon$, $1 \leq i \leq m$. Σε κάθε μη-κενό από τα σύνολα

$$V_{i_0} \cap T^{-1}(V_{i_1}) \cap \cdots \cap T^{-n+1}(V_{i_{n-1}}), \quad 1 \leq i_0, \ldots, i_{n-1} \leq m$$

θεωρούμε ένα στοιχείο $x_{i_0 \cdots i_{n-1}}$. Το πεπερασμένο σύνολο

$$F = \{ x_{i_0 \cdots i_{n-1}} : 1 \leq i_0, \ldots, i_{n-1} \leq m \}$$

έχει ισχύ το πολύ m^n. Εστω τώρα $x \in X$. Για κάθε $0 \leq k \leq n-1$ υπάρχει $1 \leq i_k \leq m$ ώστε $T^k(x) \in V_{i_k}$. Συνεπώς

$$x \in V_{i_0} \cap T^{-1}(V_{i_1}) \cap \cdots \cap T^{-n+1}(V_{i_{n-1}})$$

από όπου προκύπτει ότι

$$d(T^k(x), T^k(x_{i_0\cdots i_{n-1}})) \le diam(V_{i_k}) \le \varepsilon \quad , \quad 0\le k\le n-1.$$

Για κάθε συνεχή απεικόνιση $T: X\longrightarrow X$, $n\in\mathbb{N}$ και $\varepsilon>0$ θέτουμε:

$r_T(n,\varepsilon) = \inf\{|F| : \text{το } F \ (n,\varepsilon)\text{-παράγει τον } X\}$ και

$$\overline{r}_T(\varepsilon) = \limsup_{n\to+\infty} \frac{1}{n}\log r_T(n,\varepsilon)$$

Ένα σύνολο $E\subset X$ λέγεται <u>(n,ε)-διακριτό</u>, όπου $n\in\mathbb{N}$, $\varepsilon>0$, αν για κάθε $x,y\in E$, $x\neq y$ ισχύει $\max\{d(T^k(x),T^k(y)) : 0\le k\le n-1\}>\varepsilon$. Αν το $E\subset X$ είναι ένα (n,ε)-διακριτό σύνολο και το $F\subset X$ $(n,\frac{\varepsilon}{2})$-παράγει τον X, τότε υπάρχει μία συνάρτηση $\varphi: E\to F$ με $d(T^k(x), T^k(\varphi(x)))\le\frac{\varepsilon}{2}$ για κάθε $0\le k\le n-1$. Η φ είναι "1-1" γιατί αν $\varphi(x)=\varphi(y)$, όπου $x,y\in E$, τότε

$$d(T^k(x),T^k(y)) \le d(T^k(x),T^k(\varphi(x)))+d(T^k(\varphi(y)),T^k(y)) \le \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

για κάθε $0\le k\le n-1$, και συνεπώς $x=y$. Άρα $|E|\le|F|$. Αυτό δείχνει ότι $|E|\le r_T(n,\frac{\varepsilon}{2})$ και κατά συνέπεια

$s_T(n,\varepsilon)=\sup\{|E| : \text{το } E\subset X \text{ είναι } (n,\varepsilon)\text{-διακριτό}\}\le r_T(n,\frac{\varepsilon}{2})$

Επιπλέον $r_T(n,\varepsilon)\le s_T(n,\varepsilon)$. Πραγματικά υπάρχει ένα (n,ε)-διακριτό σύνολο $E\subset X$ με $|E|=s_T(n,\varepsilon)$. Το E (n,ε)-παράγει τον X γιατί αν αυτό δεν συμβαίνει, τότε υπάρχει $x\in X\setminus E$ ώστε

$$\max\{d(T^k(x),T^k(y)) : 0\le k\le n-1\}>\varepsilon$$

για κάθε $y\in E$. Συνεπώς το $E\cup\{x\}$ είναι (n,ε)-διακριτό, που είναι αντίφαση. Συμπερασματικά λοιπόν έχουμε:

$$r_T(n,\varepsilon)\le s_T(n,\varepsilon)\le r_T(n,\tfrac{\varepsilon}{2}) \text{ για κάθε } n\in\mathbb{N} \text{ και } \varepsilon>0.$$

Θέτουμε $\overline{s}_T(\varepsilon)=\limsup_{n\to+\infty}\frac{1}{n}\log s_T(n,\varepsilon)$. Είναι προφανές από τους ορισμούς ότι για $0<\varepsilon_1<\varepsilon_2$ έχουμε $\overline{r}_T(\varepsilon_1)\ge\overline{r}_T(\varepsilon_2)$ και $\overline{s}_T(\varepsilon_1)\ge\overline{s}_T(\varepsilon_2)$. Συνεπώς το

$$h_d(T)=\lim_{\varepsilon\to0}\overline{r}_T(\varepsilon)=\lim_{\varepsilon\to0}\overline{s}_T(\varepsilon)$$

υπάρχει στο $[0,+\infty]$.

<u>2.2. Θεώρημα.</u> Έστω X ένας συμπαγής μετρικοποιήσιμος χώρος και d μία συμβιβαστή μετρική στον X. Τότε $h(T)=h_d(T)$ για κάθε συνεχή απεικόνιση $T: X\to X$.

Απόδειξη. Θα δείξουμε πρώτα ότι $h(T) \leq h_d(T)$, δείχνοντας ότι $h(T,\alpha) \leq h_d(T)$ για κάθε πεπερασμένο ανοιχτό κάλυμμα $\alpha = \{A_1, \dots, A_p\}$ του X. Έστω $\delta > 0$ ένας αριθμός Lebesgue του α και $n \in \mathbb{N}$. Θεωρούμε ένα σύνολο $F \subset X$ που $(n, \frac{\delta}{2})$-παράγει του X με $|F| = r_T(n, \frac{\delta}{2})$. Για κάθε $z \in F$ και $0 \leq k \leq n-1$, υπάρχει $1 \leq i_k(z) \leq p$ ώστε $S(T^k(z), \delta) \subset A_{i_k(z)}$. Θέτουμε

$$C(z) = A_{i_0}(z) \cap T^{-1}(A_{i_1(z)}) \cap \dots \cap T^{-n+1}(A_{i_{n-1}(z)}) \in \alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha).$$

Τότε $X = \bigcup_{z \in F} C(z)$ γιατί για κάθε $x \in X$ υπάρχει $z \in F$ ώστε

$$\max\{d(T^k(x), T^k(z)) : 0 \leq k \leq n-1\} \leq \frac{\delta}{2} < \delta$$

και συνεπώς $x \in T^{-k}(S(T^k(z), \delta)) \subset T^{-k}(A_{i_k(z)})$, $0 \leq k \leq n-1$. Αυτό σημαίνει ότι το $\{C(z) : z \in F\}$ είναι ένα πεπερασμένο υποκάλυμμα του $\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha)$. Προκύπτει λοιπόν ότι

$$N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha)) \leq |F| = r_T(n, \frac{\delta}{2}) \quad \text{για κάθε } n \in \mathbb{N}$$

και κατά συνέπεια

$$h(T,\alpha) = \lim_{n \to +\infty} \frac{1}{n} \log N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha))$$

$$\leq \limsup_{n \to +\infty} \frac{1}{n} \log r_T(n, \frac{\delta}{2}) = \bar{r}_T(\frac{\delta}{2}) \leq h_d(T)$$

Θα δείξουμε τώρα ότι $h(T) \geq h_d(T)$. Έστω $\varepsilon > 0$. Υπάρχει ένα πεπερασμένο ανοιχτό κάλυμμα $\alpha = \{V_1, \dots, V_q\}$ του X με $\operatorname{diam}(V_k) \leq \varepsilon$, $1 \leq k \leq q$. Έστω $n \in \mathbb{N}$ και $E \subset X$ ένα (n,ε)-διακριτό σύνολο με $|E| = s_T(n,\varepsilon)$. Αν $x, y \in E$ και $x, y \in V_{i_0} \cap T^{-1}(V_{i_1}) \cap \dots \cap T^{-n+1}(V_{i_{n-1}})$ για κάποια $1 \leq i_0, \dots, i_{n-1} \leq q$, τότε

$$\max\{d(T^k(x), T^k(y)) : 0 \leq k \leq n-1\} \leq \max\{\operatorname{diam}(V_{i_k}) : 1 \leq k \leq n-1\} \leq \varepsilon$$

που σημαίνει $x = y$. Αυτό δείχνει ότι

$$N(\alpha \vee T^{-1}(\alpha) \vee \dots \vee T^{-n+1}(\alpha)) \geq |E| = s_T(n,\varepsilon) \quad \text{και}$$

$$h(T,\alpha) \geq \limsup_{n \to +\infty} \frac{1}{n} \log s_T(n,\varepsilon) = \bar{s}_T(\varepsilon) \quad \text{για κάθε } \varepsilon > 0.$$

Άρα $h(T,\alpha) \geq \lim_{\varepsilon \to 0} \bar{s}_T(\varepsilon) = \sup\{\bar{s}_T(\varepsilon) : \varepsilon > 0\} = h_d(T)$ ο.ε.δ.

Ιδιαίτερα μέσω προκύπτει από το Θεώρημα 2.2 ότι το $h_d(T)$ είναι ανεξάρτητο από την συμβιβαστή μετρική d του X.

<u>2.3. Θεώρημα.</u> Έστω $\varphi: \mathbb{R}\times X \longrightarrow X$ ένα δυναμικό σύστημα στον συμπαγή μετρικοποιήσιμο χώρο X. Τότε $h(\varphi_t) = |t|\, h(\varphi_1)$ για κάθε $t\in\mathbb{R}$.

<u>Απόδειξη</u> Λόγω του Θεωρήματος 1.6 αρκεί ν' αποδείξουμε ότι

$$h(\varphi_t) \leq \frac{t}{s}\, h(\varphi_s) \quad \text{για κάθε } t,s>0$$

Θεωρούμε μία συμβιβαστή μετρική d στον X. Λόγω της συμπάγειας, για κάθε $\varepsilon>0$ υπάρχει $\delta>0$ ώστε

$$d(\varphi_r(x), \varphi_r(y)) \leq \varepsilon \quad \text{για κάθε } x,y\in X \text{ με } d(x,y)\leq\delta \text{ και } 0\leq r\leq s.$$

Έστω $n\in\mathbb{N}$ και $F\subset X$ ένα πεπερασμένο σύνολο που $\left(\left[\frac{nt}{s}\right]+1, \delta\right)$-παράγει τον X ως προς φ_s, τότε το F (n,ε)-παράγει τον X ως προς φ_t. Έτσι έχουμε:

$$r_{\varphi_t}(n,\varepsilon) \leq r_{\varphi_s}\left(\left[\frac{nt}{s}\right]+1, \delta\right) \quad \text{για κάθε } n\in\mathbb{N}.$$

Απ' αυτό προκύπτει ότι

$$\overline{r}_{\varphi_t}(\varepsilon) \leq \overline{r}_{\varphi_s}(\delta)\cdot \limsup_{n\to+\infty} \frac{1}{n}\left(\left[\frac{nt}{s}\right]+1\right) = \left(\frac{t}{s}\right)\overline{r}_{\varphi_s}(\delta) \leq \left(\frac{t}{s}\right) h(\varphi_s)$$

για κάθε $\varepsilon>0$ ο.ε.δ.

Στην συνέχεια θα παρουσιάσουμε μερικά κλασικά παραδείγματα εκτίμησης της τοπολογικής εντροπίας.

<u>2.4. Πρόταση</u> Έστω $X\subset\mathbb{R}^m$ ένα συμπαγές σύνολο και $T: X\longrightarrow X$ μία απεικόνιση Lipschitz με σταθερά $\alpha>0$, δηλαδή

$$d(T(x), T(y)) \leq \alpha\, d(x,y)$$

για κάθε $x,y\in X$, όπου d είναι η ευκλείδεια μετρική. Τότε

$$h(T) \leq \max\{0, m\log\alpha\}$$

<u>Απόδειξη</u> Αν $\alpha\leq 1$, τότε η T δεν μεγαλώνει τις αποστάσεις και συνεπώς κάθε σύνολο $F\subset X$ που $(1,\varepsilon)$-παράγει τον X, $\varepsilon>0$, και (n,ε)-παράγει τον X για κάθε $n\in\mathbb{N}$. Άρα

$$\overline{r}_T(\varepsilon) = \limsup_{n\to+\infty} \frac{1}{n}\log r_T(1,\varepsilon) = 0 \quad \text{για κάθε } \varepsilon>0$$

και συνεπώς $h(T)=0$. Υποθέτουμε λοιπόν στην συνέχεια ότι $\alpha>1$.
Υπάρχει $b>0$ ώστε $X\subset[-b,b]^m$, αφού το X είναι συμπαγές. Για κάθε $0<\delta<b$, το σύνολο

$$F(\delta) = \{(r_1\delta,\dots,r_m\delta) : r_1,\dots,r_m\in\mathbb{Z} \text{ και } |r_i\delta|<2b,\ 1\leq i\leq m\}$$

έχει ισχύ $|F(\delta)| \leq \left(\frac{5b}{\delta}\right)^m$, γιατί οι ακέραιοι που περιέχονται στο διάστημα $\left[-\frac{2b}{\delta}, \frac{2b}{\delta}\right]$ είναι $2\left[\frac{2b}{\delta}\right]+1 \leq \frac{5b}{\delta}$. Οι κύβοι $\prod_{i=1}^{m}[r_i\delta, r_i\delta+\delta]$, όπου $(r_1\delta,\dots,r_m\delta)\in F(\delta)$ καλύπτουν το X και το πλήθος τους είναι $< \left(\frac{5b}{\delta}\right)^m$. Για κάθε $(r_1\delta,\dots,r_m\delta)\in F(\delta)$ ώστε $|r_i\delta|\leq b$ και $X\cap\prod_{i=1}^{m}(r_i\delta, r_i\delta+\delta)\neq\phi$ επιλέγουμε ένα σημείο στο $X\cap\prod_{i=1}^{m}(r_i\delta, r_i\delta+\delta)$. Έστω F_0 το σύνολο που προκύπτει μ' αυτόν τον τρόπο και $G(\delta)=F_0\cup(X\cap F(\delta))$. Τότε $|G(\delta)|\leq 2\left(\frac{5b}{\delta}\right)^m$ και για κάθε $x\in X$ υπάρχει ένα $y\in G(\delta)$ με $d(x,y)\leq\sqrt{m}\,\delta$. Συνεπώς

$$d(T^k(x),T^k(y))\leq a^n\delta\sqrt{m}\ ,\quad 0\leq k\leq n-1\ ,\quad n\in\mathbb{N}$$

που σημαίνει ότι το $G(\delta)$ $(n, a^n\delta\sqrt{m})$-παράγει το X ως προς T. Αλλιώς, για κάθε $\varepsilon>0$ το $G\left(\frac{\varepsilon}{a^n\sqrt{m}}\right)$ (n,ε)-παράγει το X ως προς T και

$$\left|G\left(\frac{\varepsilon}{a^n\sqrt{m}}\right)\right|\leq 2\left(\frac{5ba^n\sqrt{m}}{\varepsilon}\right)^m = 2a^{nm}\left(\frac{5\sqrt{m}\,b}{\varepsilon}\right)^m$$

Άρα $\bar{r}_T(\varepsilon)=\limsup_{n\to+\infty}\frac{1}{n}\left[\log\left(2\left(\frac{5\sqrt{m}\,b}{\varepsilon}\right)^m\right)+mn\log a\right]=\log a$, για κάθε $\varepsilon>0$.

Συνεπώς $h(T)=\lim_{\varepsilon\to 0}\bar{r}_T(\varepsilon)\leq m\log a$.

<u>2.5. Παράδειγμα</u>. Αν $T_\alpha: S^1\to S^1$ είναι μία στροφή του S^1 κατά γωνία $2\pi a$, $a\in\mathbb{R}$, τότε $h(T_a)=0$.

<u>2.6. Θεώρημα</u> Έστω M μία συμπαγής πολλαπλότητα Riemann και $T: M\to M$ μία C^1 απεικόνιση. Τότε

$$h(T)\leq\max\{0, m\log a\}$$

όπου $a=\sup\{\|dT(x)\|_x : x\in M\}$, όπου $(\|\ \|_x)_{x\in M}$ είναι η norm που ορίζει η μετρική Riemann.

<u>Απόδειξη</u> Επειδή η M είναι συμπαγής και η T C^1 απεικόνιση, $a<+\infty$. Επίσης $d(T(x),T(y))\leq a\,d(x,y)$ για κάθε $x,y\in M$, όπου d είναι η γεωδεσιακή απόσταση Riemann. Έτσι αν $a\leq 1$, τότε $h(T)=0$, όπως στην αρχή της απόδειξης της πρότασης 2.4. Υποθέτουμε στη συνέχεια ότι $a>1$. Κάθε $x\in M$ έχει μία ανοιχτή περιοχή V_x για την οποία υπάρχει μία C^∞ αμφιδιαφόριση $\varphi_x:\mathbb{R}^m\to V_x$ με $\varphi_x(0)=x$. Επειδή η απεικόνιση

$$\mathbb{R}^m\ni z\longrightarrow\|d\varphi_x(z)\|_{\varphi_x(z)}\in\mathbb{R}$$

είναι συνεχής, υπάρχει $A_x>0$ ώστε $\|d\varphi_x(z)\|_{\varphi_x(z)}\leq A_x$ για κάθε $\|z\|\leq 3$.

Έτσι για κάθε $z,w \in \overline{S(0,3)}$ έχουμε

$$d(\varphi_k(z),\varphi_k(w)) \le \int_0^1 \|d\varphi_k(z+t(w-z))(w-z)\|\,dt \le \int_0^1 \|d\varphi_k(z+t(w-z))\|\cdot\|w-z\|\,dt$$

$$\le A_k\|w-z\| .$$

Επειδή η M είναι συμπαγής προκύπτει από τα προηγούμενα ότι υπάρχουν C^∞ εμφυτεύσεις $\varphi_1,\dots,\varphi_r : \mathbb{R}^m \to M$ ώστε $M=\bigcup_{i=1}^{r} \varphi_i(S(0,1))$ και $A>0$ ώστε $d(\varphi_i(z),\varphi_i(w)) \le A\|z-w\|$ για κάθε $z,w\in\overline{S(0,3)}$, $1\le i\le r$.
Για κάθε $0<\delta\le 1$ θέτουμε

$$E(\delta)=\{(r_1\delta,\dots,r_m\delta) : r_i\in\mathbb{Z} \text{ και } |r_i\delta|<2 \;,\; 1\le i\le m\}$$

Τότε $|E(\delta)|\le\left(\frac{5}{\delta}\right)^m$ και για κάθε $z\in S(0,1)$ υπάρχει $w\in E(\delta)$ με $\|z-w\|\le\sqrt{m}\,\delta$. Το σύνολο $F(\delta)=\bigcup_{i=1}^{r}\varphi_i(E(\delta))$ $(n,a^nA\sqrt{m}\,\delta)$-παράγει την M για κάθε $n\in\mathbb{N}$ και $|F(\delta)|\le\left(\frac{5}{\delta}\right)^m\cdot r$. Για κάθε $\varepsilon>0$ το $F\left(\frac{\varepsilon}{a^nA\sqrt{m}}\right)$ (n,ε)-παράγει την M, $n\in\mathbb{N}$, και συνεπώς

$$r_T(n,\varepsilon)\le\left(\frac{5a^nA\sqrt{m}}{\varepsilon}\right)r = a^{nm}\cdot r\left(\frac{5\sqrt{m}\,A}{\varepsilon}\right)^m$$

Άρα $\bar{r}_T(\varepsilon)\le m\log a$ για κάθε $\varepsilon>0$ και συνεπώς $h(T)\le m\log a$.

<u>2.7. Θεώρημα</u> Κάθε ομοιομορφισμός $T: S^1\to S^1$ έχει $h(T)=0$.

<u>Απόδειξη</u> Στο S^1 θεωρούμε την μετρική

$$d(x,y)=\frac{1}{2\pi}(\text{μήκος του μικρού τόξου με άκρα } x,y)$$

οπότε $\operatorname{diam} S^1=1$. Έστω $\delta_0>0$ ώστε $d(T^{-1}(x),T^{-1}(y))\le\frac{1}{4}$, όταν $d(x,y)\le\delta_0$. Θεωρούμε ένα οποιοδήποτε $0<\delta<\delta_0$. Προφανώς $r_T(1,\delta)\le\left[\frac{1}{\delta}\right]+1$, γιατί αν προσπαθήσουμε να διαμερίσουμε το S^1, ξεκινώντας π.χ. από το σημείο 1, σε διαδοχικά τόξα μήκους $2\pi\delta$ με τελευταίο τόξο μήκους $\le 2\pi\delta$, το πεπερασμένο σύνολο των άκρων των τόξων έχει $\left[\frac{1}{\delta}\right]+1$ σημεία και προφανώς $(1,\delta)$-παράγει τον T. Θα δείξουμε ότι

$$r_T(n,\delta)\le n\left(\left[\frac{1}{\delta}\right]+1\right)$$

για κάθε $n\in\mathbb{N}$ με επαγωγή. Έστω λοιπόν ότι το $F\subset X$ $(n-1,\delta)$-παράγει τον T και $|F|\le(n-1)\left(\left[\frac{1}{\delta}\right]+1\right)$. Επειδή ο T είναι ομοιομορφισμός $|T^{n-1}(F)|=|F|$. Στο $T^{n-1}(F)$ μπορούμε να προσθέσουμε ακόμα ένα σύνολο A με $\left[\frac{1}{\delta}\right]+1$ σημεία ώστε τα τόξα

μεταξύ των διαδοχικών σημείων του $T^{n-1}(F) \cup A$ να έχουν μήκος $\leq 2\pi\delta$.
Θέτουμε $F' = F \cup T^{-n+1}(A)$. Τότε

$$|F'| \leq |F| + |T^{-n+1}(A)| = |F| + |A| \leq n\left(\left[\frac{1}{\delta}\right]+1\right)$$

Θα δείξουμε ότι το F' (n,δ)-παράγει τον T. Έστω $x \in S'$. Από την επαγωγική υπόθεση υπάρχει $y \in F$ ώστε $d(T^k(x), T^k(y)) \leq \delta$, $0 \leq k \leq n-2$. Αν $d(T^{n-1}(x), T^{n-1}(y)) \leq \delta$, τότε έχουμε τελειώσει. Έστω λοιπόν ότι $d(T^{n-1}(x), T^{n-1}(y)) > \delta$. Ένα από τα δύο τόξα, π.χ. το I, με άκρα $T^{n-1}(x), T^{n-1}(y)$ απεικονίζεται από τον T^{-1} στο μικρό τόξο J με άκρα $T^{n-2}(x), T^{n-2}(y)$ που έχει μήκος $\leq 2\pi\delta$. Αφού $d(T^{n-1}(x), T^{n-1}(y)) > \delta$, υπάρχει $\alpha \in A$ ώστε $d(T^{n-1}(x), \alpha) \leq \delta$ και $\alpha \in I$. Έτσι $\alpha = T^{n-1}(z)$ για κάποιο $z \in T^{-n+1}(A) \subset F'$, $T^{n-1}(z) \in I$ και $d(T^{n-1}(x), T^{n-1}(z)) \leq \delta$. Συνεπώς $T^{n-2}(z) \in J$ και $d(T^{n-2}(x), T^{n-2}(z)) \leq \delta$. Από την άλλη μεριά $\operatorname{diam} T^{-1}(J) \leq \frac{1}{4}$. Άρα το $T^{-1}(J)$ είναι το μικρό τόξο με άκρα $T^{n-3}(x), T^{n-3}(y)$. Άρα $\operatorname{diam} T^{-1}(J) \leq \delta$, από την υπόθεση της επαγωγής. Αφού $T^{n-3}(z) \in T^{-1}(J)$ έχουμε $d(T^{n-3}(x), T^{n-3}(z)) \leq \delta$.
Επαναλαμβάνοντας αυτή την διαδικασία έχουμε τελικά ότι
$d(T^k(x), T^k(z)) \leq \delta$ για κάθε $0 \leq k \leq n-1$ και η επαγωγή είναι τέλεια.
Προκύπτει τώρα ότι

$$0 \leq \bar{r}_T(\delta) \leq \limsup_{n\to+\infty} \frac{1}{n} \log\left(n\left(\left[\frac{1}{\delta}\right]+1\right)\right) = \limsup_{n\to+\infty} \frac{\log n}{n} = 0$$

Άρα $h(T)=0$.

<u>2.8. Πόρισμα</u> Κάθε ομοιομορφισμός $T:[0,1]\to[0,1]$ έχει $h(T)=0$.

<u>Απόδειξη</u> Αν $T:[0,1]$ είναι ένας ομοιομορφισμός, τότε ο T είναι αύξων, οπότε $T(0)=0$ και $T(1)=1$ ή φθίνων, οπότε $T(0)=1$ και $T(1)=0$. Σε κάθε περίπτωση ο $T^2:[0,1]\to[0,1]$ είναι αύξων ομοιομορφισμός και συνεπώς επάγει έναν ομοιομορφισμό $R: S'\to S'$. Συνεπώς $h(R)=0$ από το Θεώρημα 2.7. Άρα $h(T) = \frac{1}{2}h(T^2) = 0$.

Μέχρι τώρα δεν έχουμε δώσει συγκεκριμένα παραδείγματα συνεχών απεικονίσεων με μη-μηδενική τοπολογική εντροπία. Το βασικότερο παράδειγμα είναι το shift που θα παρουσιάσει στην επόμενη παράγραφο στο γενικότερο πλαίσιο των expansive ομοιομορφισμών.

Θα κλείσουμε την παράγραφο αυτή περιγράφοντας την σχέση της τοπολογικής εντροπίας με τον τοπολογικό βαθμό. Υπενθυμίζουμε σύντομα την έννοια του βαθμού μιας απεικόνισης. Εστω M μιά προσανατολίσιμη, συμπαγής, συνεκτική m-πολλαπλότητα και $f: M \to M$ μια συνεχής απεικόνιση. Τότε $H_m(M;\mathbb{Z}) \cong \mathbb{Z}$ και η f επάγει εναν ομομορφισμό ομάδων $f_*: H_m(M;\mathbb{Z}) \to H_m(M;\mathbb{Z})$. Ο ακέραιος αριθμός $\deg f = f_*(1)$ λέγεται βαθμός της f και περιγράφει πόσες φορές η f "τυλίγει" την M γύρω από τον εαυτό της, λαμβανομένου υπόψιν και του προσανατολισμού. Αν η f είναι C^2 και το $y \in M$ μία κανονική τιμή της f, δηλαδή η $Df(x): T_xM \to T_yM$ είναι "επί" για κάθε $x \in f^{-1}(y)$, τότε

$$\deg f = \sum_{i=1}^{k} \varepsilon(x_i)$$

όπου $f^{-1}(y) = \{x_1, \dots, x_k\}$ και

$$\varepsilon(x) = \begin{cases} +1, & \text{αν η } Df(x) \text{ διατηρεί τον προσανατολισμό} \\ -1, & \text{αν η } Df(x) \text{ αναστρέφει τον προσανατολισμό} \end{cases}$$

Για παράδειγμα, αν $M = S^1$ και $f(z) = z^k$, τότε $\deg f = k$, $k \in \mathbb{Z}$.

<u>2.9. Πρόταση.</u> Εστω M μια προσανατολίσιμη, συμπαγής, συνεκτική πολλαπλότητα. Αν η $f: M \to M$ είναι μια C^2 submersion επί της M,

$$h(f) \geq \log|\deg f|.$$

<u>Απόδειξη.</u> Θεωρούμε μία μετρική Riemann στην M. Επειδή η f είναι submersion, είναι τουλάχιστον $|\deg f|$-fold απεικόνιση κάλυψης (covering map). Έτσι, λόγω της συμπάγειας της M, υπάρχει $\varepsilon > 0$ ώστε $d(z_1, z_2) > \varepsilon$ για κάθε $z_1, z_2 \in f^{-1}(y)$ και $y \in M$. Εστω $E \neq \emptyset$ ενα πεπερασμένο $(1,\varepsilon)$-διακριτό σύνολο, δηλαδή $d(x,y) > \varepsilon$ για κάθε $x, y \in E$. Αν $x, y \in f^{-1}(E)$, τότε $f(x), f(y) \in E$ και συνεπώς $d(f(x), f(y)) > \varepsilon$. Αρα το $f^{-1}(E)$ είναι $(2,\varepsilon)$-διακριτό. Επαγωγικά, το $f^{-n}(E)$ είναι $(n+1,\varepsilon)$-διακριτό για κάθε $n \in \mathbb{N}$. Όμως

$$|f^{-n}(E)| \geq |\deg f|^n \cdot |E|.$$

Προκύπτει ότι $s_f(n,\varepsilon) \geq |\deg f|^{n-1} \cdot |E|$ για κάθε $n \in \mathbb{N}$ και συνεπώς

$$\bar{s}_f(\varepsilon) = \limsup_{n\to+\infty} \frac{1}{n} \log s_f(n,\varepsilon) \geq \lim_{n\to+\infty} \frac{1}{n} \log(|\deg f|^{n-1} \cdot |E|) = \log|\deg f|$$

από όπου το συμπέρασμα.

<u>2.10. Ορισμός</u> Έστω M μια συμπαγής πολλαπλότητα. Μια C^1 απεικόνιση $f: M \to M$ επί του M λέγεται expanding αν υπάρχει μία μετρική Riemann στην M και $\lambda > 1$ ώστε $\|Df^n(x)v\| \geq \lambda^n \|v\|$ για κάθε $v \in T_xM$, $x \in M$ και $n \in \mathbb{N}$.

Για παράδειγμα η $f: S^1 \to S^1$ με $f(z) = z^k$, $k \geq 2$, είναι expanding αφού με την συνηθισμένη μετρική Riemann έχουμε $\|Df^n(z)v\| = k^n \|v\|$ για κάθε $v \in T_zS^1$ και $z \in S^1$, $n \in \mathbb{N}$.

Αν η M είναι προσανατολίσιμη συμπαγής (και συνεκτική) πολλαπλότητα και η $f: M \to M$ expanding, τότε η f είναι $|\deg f|$-απεικόνιση καλύψης που διατηρεί ή αναστρέφει τον προσανατολισμό.

<u>2.11. Θεώρημα</u>. Έστω M μία προσανατολίσιμη συμπαγής συνεκτική πολλαπλότητα. Αν η $f: M \to M$ είναι expanding, τότε $h(f) = \log|\deg f|$.

<u>Απόδειξη</u> Υπάρχει $\varepsilon_0 > 0$ ώστε για κάθε $x \in M$ η ανοιχτή μπάλα $S(x,\varepsilon)$ να είναι evenly covered ισχυρά γεωδαισιακά κυρτή περιοχή κανονικών συντεταγμένων στο x, για κάθε $0 < \varepsilon \leq \varepsilon_0$. Έστω $F \neq \emptyset$ ένα πεπερασμένο σύνολο που $(1,\varepsilon)$-παράγει την M ως προς f, δηλαδή το $\{S(x,\varepsilon) : x \in F\}$ είναι ανοιχτό κάλυμμα της f. Για κάθε $n \in \mathbb{N}$, το ίδιο ισχύει και για το $\{f^{-n}(S(x,\varepsilon)) : x \in F\}$. Έστω $y \in M$. Υπάρχει $x \in F$ ώστε $y \in f^{-n}(S(x,\varepsilon))$. Αν το $z \in f^{-n}(x)$ ανήκει στην ίδια συνεκτική συνιστώσα του $f^{-n}(S(x,\varepsilon))$ που περιέχει και το y, τότε

$$\lambda^i d(f^{n-i}(z), f^{n-i}(y)) \leq d(f^n(z), f^n(y)) < \varepsilon$$

για κάθε $0 \leq i \leq n$. Αυτό δείχνει ότι το $f^{-n}(F)$ $(n+1,\varepsilon)$-παράγει την M ως προς f, αφού $\lambda > 1$. Όμως

$$|f^{-n}(F)| = |\deg f|^n \cdot |F| \quad \text{για κάθε } n \in \mathbb{N}.$$

Συνεπώς $r_f(n,\varepsilon) \leq |\deg f|^{n-1} \cdot |F|$ για κάθε $n \in \mathbb{N}$. Έτσι έχουμε:

$$\bar{r}_f(\varepsilon) = \limsup_{n \to +\infty} \frac{1}{n} \log r_f(n,\varepsilon) \leq \limsup_{n \to +\infty} \frac{1}{n} \log\left(|\deg f|^{n-1} \cdot |F|\right) = \log|\deg f|$$

για κάθε $0 < \varepsilon \leq \varepsilon_0$. Άρα $h(f) \leq \log|\deg f|$ και σε συνδυασμό με την Πρόταση 2.9 έχουμε το συμπέρασμα.

3. Expansive ομοιομορφισμοί

Έστω (X,d) ένας συμπαγής μετρικός χώρος. Ένας ομοιομορφισμός $T: X \to X$ λέγεται *expansive* αν υπάρχει $\delta>0$ ώστε

$$\sup\{d(T^n(x),T^n(y)) : n\in\mathbb{Z}\} > \delta$$

για κάθε $x,y\in X$ με $x\neq y$.

Θα δείξουμε παρακάτω ότι η expansiveness είναι τοπολογική ιδιότητα, δηλαδή ανεξάρτητη από την μετρική d (Η σταθερά δ εξαρτάται όμως από την μετρική). Γι' αυτό θα χρειαστούμε την έννοια του γεννήτορα που συνδέεται άμεσα με την expansiveness. Ένα πεπερασμένο ανοιχτό κάλυμμα α του X λέγεται *γεννήτορας* ενός ομοιομορφισμού $T: X\to X$ αν για κάθε ακολουθία $(A_n)_{n\in\mathbb{Z}}$ στοιχείων του α το σύνολο

$$\bigcap_{n\in\mathbb{Z}} T^{-n}(\bar{A}_n)$$

είναι το πολύ μονοσύνολο.

3.1. Λήμμα. Ο ομοιομορφισμός $T: X\to X$ έχει γεννήτορα τότε και μόνο τότε όταν υπάρχει ένα πεπερασμένο ανοιχτό κάλυμμα α του X ώστε για για κάθε ακολουθία $(A_n)_{n\in\mathbb{Z}}$ στοιχείων του α το $\bigcap_{n\in\mathbb{Z}} T^{-n}(A_n)$ είναι το πολύ μονοσύνολο.

Απόδειξη. Το ευθύ είναι τετριμμένο. Για το αντίστροφο, έστω $\alpha=\{A_1,\dots,A_k\}$ ένα πεπερασμένο ανοιχτό κάλυμμα του X που ικανοποιεί την υπόθεση. Έστω $\eta>0$ ένας αριθμός Lebesgue του α. Υπάρχει ένα πεπερασμένο ανοιχτό κάλυμμα $\beta=\{B_1,\dots,B_\lambda\}$ του X ώστε $\operatorname{diam}\bar{B}_i<\eta$, $1\le i\le\lambda$. Έτσι αν $(B_{i_n})_{n\in\mathbb{Z}}$ είναι μία ακολουθία στοιχείων του β, υπάρχει μία ακολουθία $(A_{j_n})_{n\in\mathbb{Z}}$ στοιχείων του α ώστε $\bar{B}_{i_n}\subset A_{j_n}$, $n\in\mathbb{Z}$. Αφού

$$\bigcap_{n\in\mathbb{Z}} T^{-n}(\bar{B}_{i_n}) \subset \bigcap_{n\in\mathbb{Z}} T^{-n}(A_{j_n})$$

το $\bigcap_{n\in\mathbb{Z}} T^{-n}(\bar{B}_{i_n})$ είναι το πολύ μονοσύνολο.

3.2. Πρόταση. Ένας ομοιομορφισμός $T: X\to X$ είναι expansive τότε και μόνο τότε όταν έχει γεννήτορα.

Απόδειξη. Έστω ότι ο T είναι expansive με σταθερά $\delta>0$. Θεωρούμε ένα πεπερασμένο ανοικτό κάλυμμα α του X από ανοικτές μπάλλες ακτίνας $\delta/2$. Έστω $(A_n)_{n\in\mathbb{Z}}$ μία ακολουθία στοιχείων του α. Αν $x,y\in\bigcap_{n\in\mathbb{Z}} T^{-n}(\bar{A}_n)$, τότε $d(T^n(x),T^n(y))\le\delta$ για κάθε $n\in\mathbb{Z}$ και επομένως $x=y$. Άρα το α είναι γεννήτορας του T. Αντίστροφα, έστω ότι ο T έχει έναν γεννήτορα α. Έστω $\delta>0$ ένας αριθμός Lebesgue του α. Αν $x,y\in X$ και $d(T^n(x),T^n(y))\le\delta/2$ για κάθε $n\in\mathbb{Z}$, τότε υπάρχουν $A_n\in\alpha$, $n\in\mathbb{Z}$ ώστε $T^n(x),T^n(y)\in A$. Συνεπώς $x,y\in\bigcap_{n\in\mathbb{Z}} T^{-n}(A_n)$ που είναι το πολύ μονοσύνολο. Άρα $x=y$, που δείχνει ότι ο T είναι expansive.

<u>3.3. Πόρισμα</u> Η expansiveness είναι τοπολογική ιδιότητα.

<u>3.4. Πόρισμα.</u> Ο ομοιομορφισμός $T:X\to X$ είναι expansive τότε και μόνον τότε όταν ο T^k είναι expansive για κάποιο $k\in\mathbb{Z}\setminus\{0\}$.

<u>Απόδειξη</u> Αν το α είναι ένας γεννήτορας του T, τότε το $\alpha\vee T^{-1}(\alpha)\vee\cdots\vee T^{-k+1}(\alpha)$ είναι ένας γεννήτορας του T^k, ενώ κάθε γεννήτορας του T^k είναι και του T.

<u>3.5. Παράδειγμα.</u> Έστω $k\in\mathbb{N}$, $X=\prod_{n\in\mathbb{Z}}\{0,1,\ldots,k-1\}$ και $T:X\to X$ το αριστερό shift στα k-σύμβολα. Αν $x=(x_n)_{n\in\mathbb{Z}}$, $y=(y_n)_{n\in\mathbb{Z}}\in X$ και $x\ne y$, υπάρχει $n_0\in\mathbb{Z}$ ώστε $x_{n_0}\ne y_{n_0}$. Συνεπώς

$$d(T^{n_0}(x),T^{n_0}(y))=\sum_{n\in\mathbb{Z}}\frac{|x_{n+n_0}-y_{n+n_0}|}{2^{|n|}}\ge|x_{n_0}-y_{n_0}|\ge1$$

Άρα το shift T είναι expansive ομοιομορφισμός. Ένας γεννήτορας του T είναι το ανοικτό κάλυμμα $\alpha=\{A_0,A_1,\ldots,A_{k-1}\}$ του X, όπου

$$A_i=\{x=(x_n)_{n\in\mathbb{Z}}\in X:\ x_0=i\}\ ,\ 0\le i\le k-1 .$$

Πράγματι $x\in\bigcap_{n\in\mathbb{Z}}T^{-n}(A_{i_n})$ τότε και μόνον τότε όταν $T^n(x)\in A_{i_n}$ για κάθε $n\in\mathbb{Z}$ δηλαδή $x=(x_{i_n})_{n\in\mathbb{Z}}$. Παρατηρούμε ότι το α αποτελείται από ξένα μεταξύ τους ανοικτά-κλειστά σύνολα. Συνεπώς $N(\alpha\vee T^{-1}(\alpha)\vee\cdots\vee T^{-n+1}(\alpha))=k^n$, $n\in\mathbb{N}$.

<u>3.6. Πρόταση</u> Έστω $T:X\to X$ ένας expansive ομοιομορφισμός και α ένας γεννήτορας του T. Τότε για κάθε $\varepsilon>0$ υπάρχει $n\in\mathbb{N}$ ώστε $\operatorname{diam}A<\varepsilon$ για κάθε $A\in\bigvee_{k=-n}^{n}T^{-k}(\alpha)$.

<u>Απόδειξη</u> Προχωρούμε με απαγωγή στο άτοπο. Έστω ότι υπάρχει $\varepsilon>0$ ώστε για κάθε $n\in\mathbb{N}$ να υπάρχει ένα $A_n\in\bigvee_{k=-n}^{n}T^{-k}(\alpha)$ με $\operatorname{diam}A_n\ge\varepsilon$. Υπάρχουν λοιπόν $x_n,y_n\in A_n$ με $d(x_n,y_n)>\frac{\varepsilon}{2}$, $n\in\mathbb{N}$. Υπάρχουν επίσης

$A_{-n,n},\dots,A_{n,n}\in\alpha$ ώστε $A_n=\bigcap_{k=-n}^{n}T^{-k}(A_{k,n})$. Επειδή ο X είναι συμπαγής μπορούμε να υποθέσουμε ότι οι ακολουθίες $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ συγκλίνουν σε σημεία $x,y\in X$ αντίστοιχα. Προφανώς $x\neq y$. Επειδή το α είναι πεπερασμένο, άπειρο πλήθος από τα σύνολα $A_{0,n}$, $n\in\mathbb{N}$ ταυτίζονται μ' ένα σύνολο $B_0\in\alpha$. Τότε $x,y\in\overline{B}_0$. Όμοια για κάθε $k\in\mathbb{Z}$ άπειρο πλήθος από τα σύνολα $A_{k,n}$, $n\geq|k|$, ταυτίζονται μ' ένα σύνολο $B_k\in\alpha$. Τότε $x,y\in T^{-k}(\overline{B}_k)$. Προκύπτει λοιπόν ότι

$$x,y\in\bigcap_{k\in\mathbb{Z}}T^{-k}(\overline{B}_k)$$

και $x\neq y$, που αντιφάσκει με το γεγονός ότι το α είναι γεννήτορας του T.

Με άλλα λόγια αν το α είναι γεννήτορας του T, τότε η οικογένεια $\bigcup\{\bigvee_{k=-n}^{n}T^{-k}(\alpha) : n\in\mathbb{N}\}$ είναι βάση της τοπολογίας του X.

<u>3.7. Θεώρημα</u>. Έστω $T:X\to X$ ένας expansive ομοιομορφισμός και α ένας γεννήτορας του T. Τότε $h(T)=h(T,\alpha)$.

<u>Απόδειξη</u> Έστω β ένα ανοιχτό κάλυμμα του X και $\eta>0$ ένας αριθμός Lebesgue του β. Από την Πρόταση 3.6 υπάρχει $n\in\mathbb{N}$ ώστε $\operatorname{diam}A<\eta$ για κάθε $A\in\bigvee_{k=-n}^{n}T^{-k}(\alpha)$. Συνεπώς $\beta<\bigvee_{k=-n}^{n}T^{-k}(\alpha)$ και έχουμε:

$$h(T,\beta)\leq h\Big(T,\bigvee_{k=-n}^{n}T^{-k}(\alpha)\Big)=\lim_{m\to+\infty}\frac{1}{m}\log N\Big(\bigvee_{i=0}^{m-1}T^{-i}\Big(\bigvee_{k=-n}^{n}T^{-k}(\alpha)\Big)\Big)$$

$$=\lim_{m\to+\infty}\frac{1}{m}\log N\Big(\bigvee_{k=-n}^{n+m-1}T^{-k}(\alpha)\Big)$$

$$\leq\lim_{m\to+\infty}\frac{1}{m}\log N\Big(\bigvee_{k=0}^{n+m-1}T^{-k}(\alpha)\Big)$$

$$=\lim_{m\to+\infty}\frac{n+m-1}{m}\cdot\frac{1}{n+m-1}\log N\Big(\bigvee_{k=0}^{n+m-1}T^{-k}(\alpha)\Big)$$

$$=h(T,\alpha)$$

Αυτό δείχνει ότι $h(T,\beta)\leq h(T,\alpha)$ για κάθε ανοιχτό κάλυμμα β του X. Άρα $h(T)=h(T,\alpha)$.

<u>3.8. Πόρισμα</u> Κάθε expansive ομοιομορφισμός έχει πεπερασμένη εντροπία.

<u>3.9. Παράδειγμα</u> Έστω $k\in\mathbb{N}$, $X=\prod_{n\in\mathbb{Z}}\{0,1,\dots,k-1\}$ και $T:X\to X$ το shift.

Χρησιμοποιώντας το Θεώρημα 3.7 θα υπολογίσουμε το $h(T)$. Θεωρήσουμε τον γεννήτορα $\alpha=\{A_0,\dots,A_{k-1}\}$ του T, όπου $A_i=\{x=(x_n)_{n\in\mathbb{Z}} : x_0=i\}$, $0\le i<k$, οπότε σύμφωνα με το Θεώρημα 3.7

$$h(T)=h(T,\alpha)=\lim_{n\to+\infty}\frac{1}{n}\log N(\alpha\vee T^{-1}(\alpha)\vee\cdots\vee T^{-n+1}(\alpha))=\lim_{n\to+\infty}\frac{1}{n}\log k^n=\log k$$

3.10. Παράδειγμα. Το παράδειγμα 3.9 μπορεί να γενικευτεί. Χρησιμοποιώντας τον συμβολισμό του παραδείγματος 3.9, έστω $\Lambda\subset X$ ένα T-αμετάβλητο συμπαγές σύνολο. Ο ομοιομορφισμός $T|\Lambda:\Lambda\to\Lambda$ λέγεται subshift. Αν $\alpha=\{A_0,\dots,A_{k-1}\}$ είναι ο φυσικός γεννήτορας του T, τότε ο $\alpha'=\{\Lambda\cap A_0,\dots,\Lambda\cap A_{k-1}\}$ είναι γεννήτορας του $T|\Lambda$ και συνεπώς $h(T|\Lambda)=h(T|\Lambda,\alpha')$. Ο $N(\bigvee_{i=0}^{n-1}(T|\Lambda)^{-i}(\alpha'))$ είναι ακριβώς ο αριθμός $\theta_n(\Lambda)$ των n-άδων $(i_0,\dots,i_{n-1})$ για τις οποίες

$$\{(x_n)_{n\in\mathbb{Z}}\in\Lambda : x_0=i_0,\dots,x_{n-1}=i_{n-1}\}\neq\phi$$

και $h(T|\Lambda)=\lim_{n\to+\infty}\frac{1}{n}\log\theta_n(\Lambda)$.

Μια ιδιαίτερα σημαντική κατηγορία subshifts είναι τα subshifts πεπερασμένου τύπου. Έστω $A=(a_{ij})\in\mathbb{R}^{k\times k}$ με $a_{ij}=0$ ή 1, $0\le i,j\le k-1$. Το υποσύνολο του X,

$$X_A=\{x=(x_n)_{n\in\mathbb{Z}} : a_{x_n x_{n+1}}=1 \text{ για κάθε } n\in\mathbb{Z}\}$$

είναι συμπαγές και T-αμετάβλητο. Η $T_A=T|X_A : X_A\to X_A$ λέγεται <u>subshift πεπερασμένου τύπου</u>. Έστω ότι $X_A\neq\phi$. Τότε,

$$\{(x_n)_{n\in\mathbb{Z}}\in X_A : x_0=i_0,\dots,x_{n-1}=i_{n-1}\}\neq\phi$$

τότε και μόνον τότε όταν $a_{i_0i_1}a_{i_1i_2}\cdots a_{i_{n-2}i_{n-1}}=1$. Άρα

$$\theta_n(X_A)=\sum_{i_0,\dots,i_{n-1}=0}^{k-1}1=\sum_{i_0,\dots,i_{n-1}=0}^{k-1}a_{i_0i_1}\cdots a_{i_{n-2}i_{n-1}}=\sum_{i_0,i_{n-1}=0}^{k-1}a^{n-1}_{i_0i_{n-1}}=\|A^{n-1}\|$$

όπου $A^{n-1}=(a^{n-1}_{ij})$. Άρα

$$h(T_A)=\lim_{n\to+\infty}\frac{1}{n}\log\|A^{n-1}\|=\lim_{n\to+\infty}\log(\|A^{n-1}\|^{1/n})=\log r$$

όπου r είναι η φασματική ακτίνα του A.

Η εντροπία ενός θερμοδυναμικού ομοιομορφισμού $T:X\to X$ συνδέεται άμεσα με το πλήθος των περιοδικών σημείων του. Θέτουμε $p_n(T)=|\{x\in X: T^n(x)=x\}|$.

<u>3.11. Πρόταση</u>. Εστω $T:X\to X$ ενας expansive ομοιομορφισμος και α ένας γεννήτορας του. Τότε $p_n(T)\leq |\alpha|^n$.

<u>Απόδειξη</u> Εστω $x\in X$ με $T^n(x)=x$. Υπαρχουν $A_0,\dots,A_{n-1}\in\alpha$ ώστε $x\in A_0\cap T(A_1)\cap\dots\cap T^{n-1}(A_{n-1})$. Συνεπώς

$$x=T^{kn}(x)\in T^{kn}(A_0\cap T(A_1)\cap\dots\cap T^{n-1}(A_{n-1})) \text{ για καθε } k\in\mathbb{Z}.$$

Αφου το α είναι γεννήτορας του T προκύπτει οτι

$$\{x\}=\bigcap_{k\in\mathbb{Z}} T^{kn}(A_0\cap T(A_1)\cap\dots\cap T^{n-1}(A_{n-1})).$$

Ετσι για καθε n-αδα $(A_0,\dots,A_{n-1})$ στοιχείων του α παίρνουμε με τον προηγουμενο φοπο το πολύ ένα περιοδικό σημείο περιοδου n. Αρα $p_n(T)\leq|\alpha|^n$.

<u>3.12. Θεώρημα.</u> Αν ο T είναι expansive ομοιομορφισμος τότε

$$h(T)\geq \limsup_{n\to+\infty}\frac{1}{n}\log p_n(T)$$

<u>Απόδειξη</u>. Εστω α ένας γεννήτορας του T. Για καθε $n\in\mathbb{N}$ εχουμε:

$$p_n(T)\leq N(\alpha\vee T^{-1}(\alpha)\vee\dots\vee T^{-n+1}(\alpha))$$

γιατι αν τα $x,y\in X$ είναι περιοδικά σημεία περιοδου n και $A_0,\dots,A_{n-1}\in\alpha$ με $x,y\in\bigcap_{k=0}^{n-1}T^{-k}(A_k)$, τότε

$$x,y\in\bigcap_{k\in\mathbb{Z}}T^{-kn}(A_0\cap T^{-1}(A_1)\cap\dots\cap T^{-n+1}(A_{n-1}))$$

που είναι το πολύ μονοσύνολο. Ετσι εχουμε

$$\limsup_{n\to+\infty}\frac{1}{n}\log p_n(T)\leq h(T,\alpha)=h(T).$$

<u>4. Τοπολογικη και πιθανοθεωρητικη εντροπια.</u>

Εστω X ένας συμπαγης μετρικοποιησιμος χώρος και d μια συμβιβαστή μετρική στον X. Συμβολίζουμε με $\mathcal{M}(X)$ το σύνολο των μέτρων πιθανοτητας Borel στον X. Εστω $T:X\to X$ μια συνεχης απεικονιση και $\mathcal{M}_T(X)$ το σύνολο των T-αμετάβλητων μετρων πιθανοτητας Borel. Στην παράγραφο αυτή θα δείξουμε οτι $h(T)=\sup\{h_\mu(T):\mu\in\mathcal{M}_T(X)\}$. Με $h_\mu(T)$

συμβολίζουμε την εντροπία του T ως προς το μέτρο $\mu \in \mathcal{M}_T(X)$.

4.1. Λήμμα. Έστω $\mu \in \mathcal{M}(X)$. Τότε

(α) Για κάθε $x \in X$ και $\varepsilon > 0$ υπάρχει $0 < \delta < \varepsilon$ ώστε $\mu(\partial S(x,\delta)) = 0$.

(β) Για κάθε $\varepsilon > 0$ υπάρχει μια διαμέριση $\mathcal{P} = \{A_1, \dots, A_k\}$ του $(X, \mathcal{B}(X))$, ώστε $\operatorname{diam}(A_i) < \varepsilon$ και $\mu(\partial A_i) = 0$, $1 \le i \le k$, όπου $\mathcal{B}(X)$ είναι τα σύνολα Borel του X.

Απόδειξη. Το (α) είναι προφανές. Για το (β), από το (α) υπάρχει ένα πεπερασμένο ανοικτό κάλυμμα $\{B_1, \dots, B_k\}$ του X από ανοικτές μπάλλες ακτίνας $< \frac{\varepsilon}{2}$ με $\mu(\partial B_i) = 0$, $1 \le i \le k$. Θέτουμε $A_1 = \overline{B}_1$ και $A_i = \overline{B}_i \setminus (\overline{B}_1 \cup \dots \cup \overline{B}_{i-1})$, $1 < i \le k$. Τότε η $\mathcal{P} = \{A_1, \dots, A_k\}$ είναι διαμέριση του $(X, \mathcal{B}(X))$, γιατί $A_i \cap A_j = \emptyset$ για $i \ne j$, με $\operatorname{diam}(A_i) < \varepsilon$. Δεδομένου ότι $\partial A_i \subset \bigcup_{i=1}^{k} \partial B_i$ έχουμε $\mu(\partial A_i) = 0$, $1 \le i \le k$.

Για κάθε $\mu \in \mathcal{M}(X)$ και διαμέριση $\mathcal{P}$ του $(X, \mathcal{B}(X))$ συμβολίζουμε με $H_\mu(\mathcal{P})$ την εντροπία της $\mathcal{P}$ ως προς την πιθανότητα μ.

4.2. Λήμμα. Αν $\mu_1, \dots, \mu_n \in \mathcal{M}(X)$ και $a_1, \dots, a_n \ge 0$ με $a_1 + \dots + a_n = 1$, τότε $H_{a_1\mu_1 + \dots + a_n\mu_n}(\mathcal{P}) \ge a_1 H_{\mu_1}(\mathcal{P}) + \dots + a_n H_{\mu_n}(\mathcal{P})$ για κάθε διαμέριση $\mathcal{P}$ του $(X, \mathcal{B}(X))$.

Απόδειξη. Το συμπέρασμα είναι άμεσο απ' τους ορισμούς και την κυρτότητα της συνεχούς συνάρτησης $\psi : [0, +\infty) \to \mathbb{R}$ με τύπο

$$\psi(t) = \begin{cases} t \log t, & \text{όταν } t > 0 \\ 0, & \text{όταν } t = 0 \end{cases}$$

4.3. Θεώρημα. Έστω $T : X \to X$ μια συνεχής απεικόνιση. Αν $\mu_1, \dots, \mu_n \in \mathcal{M}_T(X)$ και $a_1, \dots, a_n \ge 0$ με $a_1 + \dots + a_n = 1$, τότε

$$h_{a_1\mu_1 + \dots + a_n\mu_n}(T) = a_1 h_{\mu_1}(T) + \dots + a_n h_{\mu_n}(T).$$

Απόδειξη. Αρκεί να δείξουμε ότι $h_{a\mu_1 + (1-a)\mu_2}(T) = a h_{\mu_1}(T) + (1-a) h_{\mu_2}(T)$ για κάθε $\mu_1, \mu_2 \in \mathcal{M}_T(X)$ και $0 < a < 1$.

Θεωρούμε την συνεχή κυρτή συνάρτηση $\psi(t) = t\log t$, $t > 0$, $\psi(0) = 0$.

Για κάθε σύνολο Borel $B\subset X$ με $\mu_1(B)\mu_2(B)>0$, έχουμε

$$0 \geq \psi(\alpha\mu_1(B)+(1-\alpha)\mu_2(B)) - \alpha\psi(\mu_1(B)) - (1-\alpha)\psi(\mu_2(B))$$
$$=(\alpha\mu_1(B)+(1-\alpha)\mu_2(B))\log(\alpha\mu_1(B)+(1-\alpha)\mu_2(B))$$
$$-\alpha\mu_1(B)\log\mu_1(B) - (1-\alpha)\mu_2(B)\log\mu_2(B)$$
$$=\alpha\mu_1(B)\left[\log(\alpha\mu_1(B)+(1-\alpha)\mu_2(B)) - \log(\alpha\mu_1(B))\right]$$
$$+(1-\alpha)\mu_2(B)\left[\log(\alpha\mu_1(B)+(1-\alpha)\mu_2(B)) - \log((1-\alpha)\mu_2(B))\right]$$
$$+\alpha\mu_1(B)\left[\log(\alpha\mu_1(B)) - \log\mu_1(B)\right]$$
$$+(1-\alpha)\mu_2(B)\left[\log((1-\alpha)\mu_2(B)) - \log\mu_2(B)\right]$$
$$\geq 0+0+\alpha\mu_1(B)\log\alpha + (1-\alpha)\mu_2(B)\log(1-\alpha)$$

Έτσι για κάθε διαμέριση $\mathcal{P}$ του $(X,\mathcal{B}(X))$ έχουμε:

$$0\leq H_{\alpha\mu_1+(1-\alpha)\mu_2}(\mathcal{P}) - \alpha H_{\mu_1}(\mathcal{P}) - (1-\alpha)H_{\mu_2}(\mathcal{P})$$
$$\leq -(\alpha\log\alpha + (1-\alpha)\log(1-\alpha)) < (1-\alpha)+\alpha = 1$$

αφού τα μ_1,μ_2 είναι μέτρα πιθανότητας.

Συνεπώς για κάθε διαμέριση $\mathcal{P}$ του $(X,\mathcal{B}(X))$ έχουμε

$$h_{\alpha\mu_1+(1-\alpha)\mu_2}(T,\mathcal{P}) = \alpha h_{\mu_1}(T,\mathcal{P}) + (1-\alpha)h_{\mu_2}(T,\mathcal{P})$$

και άρα $h_{\alpha\mu_1+(1-\alpha)\mu_2}(T)\leq \alpha h_{\mu_1}(T)+(1-\alpha)h_{\mu_2}(T)$.

Για την αντίθετη ανισότητα θεωρούμε $\varepsilon>0$. Επιλέγουμε διαμερίσεις $\mathcal{P}_1,\mathcal{P}_2$ του $(X,\mathcal{B}(X))$ ώστε

$$h_{\mu_i}(T,\mathcal{P}_i) > \begin{cases} h_{\mu_i}(T)-\varepsilon, & \text{αν } h_{\mu_i}(T)<+\infty \\ \frac{1}{\varepsilon}, & \text{αν } h_{\mu_i}(T)=+\infty \end{cases}, \quad i=1,2 .$$

Θέτοντας $\mathcal{P}=\mathcal{P}_1\vee\mathcal{P}_2$ έχουμε από την προηγούμενη ισότητα

$$h_{\alpha\mu_1+(1-\alpha)\mu_2}(T,\mathcal{P}) > \begin{cases} \alpha h_{\mu_1}(T)+(1-\alpha)h_{\mu_2}(T)-\varepsilon, & \text{αν } h_{\mu_1}(T)<+\infty \text{ και } h_{\mu_2}(T)<+\infty \\ \frac{1}{\varepsilon}, & \text{αν } h_{\mu_1}(T)=+\infty \text{ ή } h_{\mu_2}(T)=+\infty \end{cases}$$

Αυτό δείχνει ότι $h_{\alpha\mu_1+(1-\alpha)\mu_2}(T)\geq \alpha h_{\mu_1}(T)+(1-\alpha)h_{\mu_2}(T)$ ο.ε.δ.

<u>4.4 Πρόταση</u> Για κάθε συνεχή απεικόνιση $T:X\longrightarrow X$ και $\mu\in\mathcal{M}_T(X)$ ισχύει $h_\mu(T)\leq h(T)$.

<u>Απόδειξη.</u> Έστω $\mathcal{P}=\{A_1,\dots,A_k\}$ μία διαμέριση του $(X,\mathcal{B}(X))$.
Έστω $0<\varepsilon<\frac{1}{k\log k}$. Επειδή το μ είναι κανονικό, υπάρχουν συμπαγή

σύνολα $C_i \subset A_i$ με $\mu(A_i \setminus C_i) < \varepsilon$, $1 \le i \le k$. Θέτουμε $C_0 = X \setminus \bigcup_{i=1}^k C_i$.
Τότε $C_0 \subset (\bigcup_{i=1}^k A_i \setminus \bigcup_{i=1}^k C_i) \cup (X \setminus \bigcup_{i=1}^k A_i) \subset (\bigcup_{i=1}^k (A_i \setminus C_i)) \cup (X \setminus \bigcup_{i=1}^k A_i)$
και συνεπώς $\mu(C_0) \le \mu(\bigcup_{i=1}^k (A_i \setminus C_i)) = \sum_{i=1}^k \mu(A_i \setminus C_i) < k\varepsilon$. Επίσης
αν $Q = \{C_0, \dots, C_k\}$, τότε

$$H_\mu(\mathcal{P} | Q) = - \sum_{i=0}^k \sum_{j=1}^k \mu(C_i)\, \psi\left(\frac{\mu(C_i \cap A_j)}{\mu(C_i)} \right)$$

όπου όπως πάντα $\psi(t) = t\log t$, $t > 0$, και $\psi(0) = 0$. Αφού για $i \neq 0$
έχουμε $\frac{\mu(C_i \cap A_j)}{\mu(C_i)} = 0$ ή 1 αν $i \neq j$ ή $i = j$ αντίστοιχα, προκύπτει ότι

$$H_\mu(\mathcal{P}|Q) = -\mu(C_0) \sum_{j=1}^k \psi\left(\frac{\mu(C_0 \cap A_j)}{\mu(C_0)} \right) = -k\mu(C_0) \sum_{j=1}^k \frac{1}{k} \psi\left(\frac{\mu(C_0 \cap A_j)}{\mu(C_0)} \right)$$

$$\le k\mu(C_0) \left[-\psi\left(\sum_{j=1}^k \frac{1}{k} \cdot \frac{\mu(C_0 \cap A_j)}{\mu(C_0)} \right) \right] = k\mu(C_0)\left(-\psi\left(\frac{1}{k}\right)\right)$$

$$= k\mu(C_0) \cdot \frac{1}{k} \log k < k\varepsilon \log k < 1$$

Επειδή $C_0 \cup C_i = X \setminus \bigcup_{j \neq i} C_j$, $0 < i \le k$, το $\alpha = \{C_0 \cup C_1, \dots, C_0 \cup C_k\}$
είναι ανοιχτό κάλυμμα του X. Από την κυρτότητα της ψ για κάθε
$n \in \mathbb{N}$ έχουμε

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(Q) \right) \le \log \left| \bigvee_{i=0}^{n-1} T^{-i}(Q) \right|$$

Κάθε στοιχείο του $\bigvee_{i=0}^{n-1} T^{-i}(\alpha)$ είναι η ένωση 2^n στοιχείων της
$\bigvee_{i=0}^{n-1} T^{-i}(Q)$. Έτσι 2^n μη-κενά στοιχεία της $\bigvee_{i=0}^{n-1} T^{-i}(Q)$ δίνουν το
ίδιο μη-κενό στοιχείο του $\bigvee_{i=0}^{n-1} T^{-i}(\alpha)$. Άρα $\left|\bigvee_{i=0}^{n-1} T^{-i}(Q)\right| \le 2^n N\left(\bigvee_{i=0}^{n-1} T^{-i}(\alpha)\right)$
και συνεπώς

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(Q) \right) \le \log\left(2^n N\left(\bigvee_{i=0}^{n-1} T^{-i}(\alpha) \right) \right).$$

Άρα $h_\mu(T, \mathcal{P}) \le h_\mu(T, Q) + H_\mu(\mathcal{P}|Q) \le h(T, \alpha) + \log 2 + 1$
$\le h(T) + \log 2 + 1$

Αυτό δείχνει ότι $h_\mu(T) \le h(T) + \log 2 + 1$ για κάθε συνεχή απεικόνιση
$T: X \to X$. Εφαρμόζοντας το αποτέλεσμα αυτό για την T^n, $n \in \mathbb{N}$, έχουμε
$n h_\mu(T) = h_\mu(T^n) \le h(T^n) + \log 2 + 1 = n h(T) + \log 2 + 1$
για κάθε $n \in \mathbb{N}$. Άρα $h_\mu(T) \le h(T)$ ο.ε.δ.

4.5. Πρόταση. Έστω $T:X\to X$ μια συνεχής απεικόνιση. Τότε για κάθε $\varepsilon>0$ υπάρχει $\mu\in\mathcal{M}_T(X)$ ώστε $\bar{s}_T(\varepsilon)\le h_\mu(T)$.

Απόδειξη. Έστω $n\in\mathbb{N}$ και E_n ένα (n,ε)-διακριτό σύνολο με $|E_n|=s_T(n,\varepsilon)$. Θέτουμε

$$\sigma_n=\frac{1}{s_T(n,\varepsilon)}\sum_{x\in E_n}\delta_x\in\mathcal{M}(X),$$

δηλαδή το σ_n είναι το ατομικό μέτρο ομοιόμορφα κατανεμημένο στα στοιχεία του E_n. Θέτουμε επίσης $\mu_n=\frac{1}{n}\sum_{i=0}^{n-1}\sigma_n\circ T^{-i}\in\mathcal{M}(X)$, $n\in\mathbb{N}$. Επειδή ο $\mathcal{M}(X)$ είναι ασθενώς συμπαγής, υπάρχει μια ακολουθία $n_\lambda\to+\infty$ ώστε η $(\mu_{n_\lambda})_{\lambda\in\mathbb{N}}$ να συγκλίνει ασθενώς σε κάποιο $\mu\in\mathcal{M}(X)$ και $\lim_{\lambda\to+\infty}\frac{1}{n_\lambda}\log s_T(n_\lambda,\varepsilon)=\bar{s}_T(\varepsilon)$. Όπως είναι γνωστόν $\mu\in\mathcal{M}_T(X)$. Θα δείξουμε ότι $\bar{s}_T(\varepsilon)\le h_\mu(T)$.

Από το Λήμμα 4.1 υπάρχει μια διαμέριση $\mathcal{P}=\{A_1,\dots,A_k\}$ του $(X,\mathcal{B}(X))$ με $\operatorname{diam}(A_i)<\varepsilon$ και $\mu(\partial A_i)=0$, $1\le i\le k$. Επειδή το E_n είναι (n,ε)-διακριτό, κάθε στοιχείο της διαμέρισης $\bigvee_{i=0}^{n-1}T^{-i}(\mathcal{P})$ περιέχει το πολύ ενα στοιχείο του E_n. Αφού λοιπόν $|E_n|=s_T(n,\varepsilon)$, $s_T(n,\varepsilon)$ στοιχεία της $\bigvee_{i=0}^{n-1}T^{-i}(\mathcal{P})$ έχουν σ_n-μέτρο $\frac{1}{s_T(n,\varepsilon)}$, ενώ τα άλλα μηδέν. Απ' αυτά προκύπτει ότι

$$H_{\sigma_n}\Big(\bigvee_{i=0}^{n-1}T^{-i}(\mathcal{P})\Big)=-\sum_{i=1}^{s_T(n,\varepsilon)}\frac{1}{s_T(n,\varepsilon)}\log\frac{1}{s_T(n,\varepsilon)}=\log s_T(n,\varepsilon)$$

Έστω τώρα $1<q<n$. Για κάθε $0\le j\le q-1$, θέτουμε $a(j)=\left[\frac{n-j}{q}\right]$. Τότε ισχύουν τα ακόλουθα:

(α) $a(0)\ge a(1)\ge\dots\ge a(q-1)$.

(β) Για κάθε $0\le j\le q-1$ ισχύει

$$\{0,1,\dots,n-1\}=\{j+rq+i:\ 0\le r\le a(j)-1 \text{ και } 0\le i\le q-1\}\cup S$$

όπου $S=\{0,1,\dots,j-1,\ j+a(j)q,\ j+a(j)q+1,\dots,n-1\}$ και συνεπώς $|S|=n-(j+a(j)q-j)\le n-(n-q+j)=q+j<2q$, γιατί $j+a(j)q\ge j+\left[\frac{n-j}{q}-1\right]q$ = ο μέγιστος ακέραιος που είναι μικρότερος ή ίσος με $j+\frac{n-j}{q}-1=n-q$.

(γ) Οι αριθμοί $j+rq$, $0\le j\le q-1$, $0\le r\le a(j)-1$ είναι όλοι μικρότεροι από $n-q$, γιατί $j+rq\le j+(a(j)-1)q\le j+\left[\frac{n-j}{q}-1\right]q=n-q$, και διαφορετικοί μεταξύ τους, αφού $0 \le j \le q-1$.

Από το (β) τώρα έχουμε
$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \bigvee_{r=0}^{a(j)-1} T^{-(j+rq)}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) \vee \bigvee_{\ell \in S} T^{-\ell}(\mathcal{P}) \quad \text{και} \quad |S| < 2q.$$

Άρα
$$\log s_T(n,\varepsilon) = H_{\sigma_n}\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\Big)$$
$$\le \sum_{r=0}^{a(j)-1} H_{\sigma_n}\Big(T^{-(rq+j)}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big)\Big) + \sum_{\ell \in S} H_{\sigma_n}\big(T^{-\ell}(\mathcal{P})\big)$$
$$\le \sum_{r=0}^{a(j)-1} H_{\sigma_n \circ T^{-(j+rq)}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + 2q\log k.$$

Προσθέτοντας τώρα κατά μέλη έχουμε
$$q\log s_T(n,\varepsilon) = \sum_{j=0}^{q-1} \log s_T(n,\varepsilon) \le \sum_{r=0}^{a(j)-1}\sum_{j=0}^{q-1} H_{\sigma_n \circ T^{-(j+rq)}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + 2q^2\log k$$
$$\le \sum_{m=0}^{n-1} H_{\sigma_n \circ T^{-m}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + 2q^2\log k$$
λόγω του (γ).

Διαιρώντας τώρα με n και χρησιμοποιώντας το Λήμμα 4.2 έχουμε
$$\frac{q}{n}\log s_T(n,\varepsilon) \le \sum_{m=0}^{n-1} \frac{1}{n} H_{\sigma_n \circ T^{-m}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + \frac{2q^2}{n}\log k$$
$$\le H_{\mu_n}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + \frac{2q^2}{n}\log k$$

Αν τώρα $A_{i_0}, \dots, A_{i_{q-1}} \in \mathcal{P}$, τότε $\partial\Big(\bigcap_{m=0}^{q-1} T^{-m}(A_{i_m})\Big) \subset \bigcup_{m=0}^{q-1} T^{-m}(\partial A_{i_m})$ και συνεπώς $\mu\Big(\partial \bigcap_{m=0}^{q-1} T^{-m}(A_{i_m})\Big) = 0$. Δηλαδή $\mu(\partial B) = 0$ για κάθε $B \in \bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})$. Συνεπώς $\lim_{\lambda\to+\infty} \mu_{n_\lambda}(B) = \mu(B)$ για κάθε $B \in \bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})$.

Απ' αυτό προκύπτει ότι
$$\lim_{\lambda\to+\infty} H_{\mu_{n_\lambda}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) = H_\mu\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big).$$

Έτσι για $n = n_\lambda$ στην τελευταία ανισότητα και διαιρώντας με q έχουμε
$$\frac{1}{n_\lambda}\log s_T(n,\varepsilon) \le \frac{1}{q} H_{\mu_{n_\lambda}}\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) + \frac{2q}{n_\lambda}\log k$$

για κάθε $\lambda \in \mathbb{N}$ και $1 < q < n_\lambda$, οπότε παίρνοντας όρια για $\lambda \to +\infty$,
$$\overline{s}_T(\varepsilon) \le \frac{1}{q} H_\mu\Big(\bigvee_{i=0}^{q-1} T^{-i}(\mathcal{P})\Big) \quad \text{για κάθε } q > 1$$

Αρα $\underline{\Sigma}_T(\varepsilon) \leq h_\mu(T,\mathcal{P}) \leq h_\mu(T)$ ο.ε.δ.

Από τις Προτάσεις 4.4 και 4.5 προκύπτει το βασικό συμπέρασμα της παρούσας παραγράφου.

<u>4.6. Θεώρημα</u>. Για κάθε συνεχή απεικόνιση $T: X \to X$ ισχύει

$$h(T) = \sup\{h_\mu(T) : \mu \in \mathcal{M}_T(X)\}.$$