Εντροπία, υπερσυσταλτότητα και οι ανισότητες του **Talagrand**

Διπλωματική Εργασία

Μηνάς Πάφης

Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
Αθήνα – 2020

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Σε αυτή την εργασία συνοψίζουμε μελέτες των τελευταίων δεκαετιών δίνοντας έμφαση σε έναν ενιαίο τρόπο αντιμετώπισης (που βασίζεται στην εντροπία, την λογαριθμική ανισότητα Sobolev και την υπερσυσταλτότητα) τεσσάρων διάσημων ανισοτήτων του Talagrand. Αυτή η ενιαία αντιμετώπιση εμφανίστηκε σε μια πρόσφατη εργασία του Ledoux (βλ. [54]). Μελετάμε την ανισότητα απόστασης από την κυρτή θήκη, την L^1-L^2-ανισότητα για την διασπορά, την ανισότητα μεταφοράς με τετραγωνικό κόστος και την ανισότητα για το supremum εμπειρικών διαδικασιών. Παρουσιάζουμε επίσης τις αρχικές αποδείξεις που έδωσε ο Talagrand και συζητάμε, ενδεικτικά, κάποιες από τις εφαρμογές τους σε διάφορες περιοχές των Μαθηματικών.

1.1 Οι τέσσερις ανισότητες του Talagrand

Σε αυτή την παράγραφο δίνουμε απλά το περιεχόμενο των τεσσάρων ανισοτήτων του Talagrand, όπως εμφανίστηκαν στην πρώτη τους έκδοση. Όλες οι ανισότητες δημοσιεύτηκαν κατά την περίοδο 1994-96. Οι τρεις πρώτες μάλιστα δημοσιεύτηκαν σε σχετικά μικρά άρθρα.

Η ανισότητα απόστασης από την κυρτή θήκη. Έστω $(\Omega_i, \mathcal{A}_i, \mu_i)$, $1 \leqslant i \leqslant n$ χώροι πιθανότητας. Θεωρούμε τον χώρο γινόμενο $(\Omega, \mathcal{A}, \mathbb{P}) = (\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \mu_i)$. Δοθέντων $A \subseteq \Omega$ και $x \in \Omega$, ορίζουμε

$$U_A(x) = \{(s_i)_{1\leqslant i\leqslant n} \in \{0,1\}^n \mid \text{ υπάρχει } y \in A \text{ ώστε αν } s_i = 0 \text{ τότε } x_i = y_i \text{ για κάθε } 1 \leqslant i \leqslant n\}.$$

Συμβολίζουμε με $V_A(x)$ την κυρτή θήκη $\mathrm{conv}(U_A(x))$ την οποία θεωρούμε ως υποσύνολο του $\mathbb{R}^n$. Τότε, $0 \in V_A(x)$ αν και μόνο αν $x \in A$. Θεωρούμε τη συνάρτηση

$$d_A(x) = \mathrm{dist}(0, V_A(x)).$$

Τότε, για κάθε $A \in \mathcal{A}$ ισχύει ότι

$$\int_\Omega e^{\frac{1}{4}d_A^2} d\mathbb{P} \leqslant \frac{1}{\mathbb{P}(A)}. \tag{1.1.1}$$

Αυτή η ανισότητα έχει πολυάριθμες εφαρμογές σε διάφορες περιοχές της θεωρίας πιθανοτήτων, της στατιστικής και της βελτιστοποίησης, αλλά και στη συνδυαστική και στα διακριτά μαθηματικά. Ο Talagrand απέδειξε αυτή την ανισότητα στο [91], γενικεύοντας από πολλές απόψεις προηγούμενο δικό του αποτέλεσμα για τον διακριτό κύβο (βλ. [90]). Ενδιάμεσα, οι Johnson και Schechtman είχαν δώσει στο [46] μια απλή γενίκευση του αποτελέσματος της εργασίας [90], και αυτή η γενίκευση στάθηκε το κίνητρο για την γενική διατύπωση της ανισότητας στο [91]. Εκτός από την (1.1.1), ο Talagrand απέδειξε και άλλες ανισότητες αυτού του τύπου, στις οποίες υπεισέρχονται αποστάσεις διαφορετικές από την d_A. Παρουσιάζουμε όλα αυτά τα αποτελέσματα στο Κεφάλαιο 3. Η νέα απόδειξη που παρουσιάζουμε βασίζεται στη μέθοδο της εντροπίας και οφείλεται στους Boucheron, Lugosi και Massart (βλ. [21] και [22]).

Το πρώτο, στη σειρά αυτών των άρθρων, αποτέλεσμα του Talagrand προήλθε από προβλήματα της περιοχής των πιθανοτήτων σε χώρους Banach. Μια εφαρμογή που είχε στο νου του ήταν μια ισχυρή εκδοχή της κλασικής ανισότητας Kahane-Khintchine (βλ. [55] και [47]). Περιγράφουμε αυτή την εφαρμογή στο Κεφάλαιο 3.

Το πλήθος των αναφορών στην εργασία [91] (δείτε για παράδειγμα τα βιβλία και άρθρα επισκόπησης [94], [89], [58], [53], [67], [32], [98], [99], [22], [3], [100] και άλλα) δείχνει το εύρος των εφαρμογών αυτής της ανισότητας.

Η L^1–L^2-ανισότητα διασποράς. Έστω $\{0,1\}^n$ ο διακριτός κύβος και $p \in (0,1)$. Εφοδιάζουμε τον διακριτό κύβο με το μέτρο γινόμενο πιθανότητας μ_p το οποίο δίνει βάρος p στο 1 και $1-p$ στο 0. Για $1 \leqslant r \leqslant \infty$ συμβολίζουμε με $\|\cdot\|_r$ τη συνήθη νόρμα του $L^r(\mu_p)$. Για κάθε $1 \leqslant i \leqslant n$ και για κάθε $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ ορίζουμε $U_i(x)$ το σημείο του $\{0,1\}^n$ που παίρνουμε αν στην i-οστή συντεταγμένη του x αντικαταστήσουμε το x_i με $1-x_i$ και αφήσουμε τις υπόλοιπες συντεταγμένες του x αμετάβλητες. Αν $f : \{0,1\}^n \to \mathbb{R}$ είναι μια συνάρτηση, θέτουμε

$$\Delta_i f(x) = \begin{cases} (1-p)(f(x) - f(U_i(x))) & , \text{ αν } x_i = 1 \\ p(f(x) - f(U_i(x))) & , \text{ αν } x_i = 0. \end{cases}$$

Τότε, υπάρχει σταθερά $K > 0$ ώστε για κάθε συνάρτηση $f : \{0,1\}^n \to \mathbb{R}$ με $\int_{\{0,1\}^n} f \, d\mu_p = 0$ να ισχύει ότι

$$\|f\|_2^2 \leqslant K \log\Big(\frac{2}{p(1-p)}\Big) \sum_{i=1}^n \frac{\|\Delta_i f\|_2^2}{\log\Big(e\frac{\|\Delta_i f\|_2}{\|\Delta_i f\|_1}\Big)}.$$

Η L^1-L^2-ανισότητα διασποράς στον διακριτό κύβο προσέφερε μια εναλλακτική προσέγγιση σε ένα θεώρημα των Kahn-Kalai-Linial στην Boolean ανάλυση και αποτέλεσε κεντρικό εργαλείο της θεωρητικής πληροφορικής. Μοιάζει αυτή τη στιγμή να δίνει το μόνο γενικό επιχείρημα που εξασφαλίζει την ύπαρξη της λεγόμενης υπερ-συγκέντρωσης που παίζει κρίσιμο ρόλο σε πολλά μοντέλα της σύγχρονης θεωρίας πιθανοτήτων (διήθηση, τυχαίους πίνακες, spin glasses). Τα βιβλία και άρθρα επισκόπησης [12], [23], [38], [6], [87] προσφέρουν πολλές πληροφορίες για τις σχετικές εξελίξεις σε αυτές τις περιοχές.

Στην πρώτη παράγραφο του Κεφαλαίου 4 περιγράφουμε αναλυτικά την προέλευση της L^1-L^2 ανισότητας. Αφετηρία του Talagrand ήταν ένα αποτέλεσμα του Russo [81] σχετικά με φαινόμενα κατωφλίου για μονότονα υποσύνολα του διακριτού κύβου που έχουν μικρή εξάρτηση από κάθε συντεταγμένη. Ταυτόχρονα, χρησιμοποιώντας την έδωσε εναλλακτική απόδειξη του θεωρήματος

Kahn-Kalai-Linial από το [48]. Η αρχική απόδειξη του Talagrand στο [93] χρησιμοποιεί ανάλυση Fourier (την ανισότητα Bonami-Beckner) και δεν αναφέρεται ρητά στην υπερσυσταλτότητα. Τροποποιεί και αξιοποιεί κάποιες από τις ιδέες που είχαν χρησιμοποιηθεί στο [48]. Η δεύτερη απόδειξη που παρουσιάζουμε βασίζεται πλήρως στην υπερσυσταλτότητα.

Η L^1-L^2 ανισότητα ισχύει και στον χώρο του Gauss και είναι, κάτω από προϋποθέσεις, ισχυρότερη από την ανισότητα Poincaré. Παρουσιάζουμε μια εφαρμογή αυτής της οπτικής στην γεωμετρική συναρτησιακή ανάλυση: στο Κεφάλαιο 4 περιγράφουμε ένα πρόσφατο αποτέλεσμα των Βαλέττα και Παούρη, το οποίο δίνει την καλύτερη γνωστή «εξάρτηση από το ε» στο κλασικό θεώρημα του Dvoretzky.

Μπορούμε να δούμε την L^1-L^2 ανισότητα ως μια δυϊκή μορφή της λογαριθμικής ανισότητας Sobolev, με την έννοια ότι η τελευταία εξασφαλίζει ότι αν η κλίση μιας συνάρτησης f είναι στον L^2 τότε η f είναι στον χώρο Orlicz $L^2\log L$, ενώ η ανισότητα του Talagrand εξασφαλίζει ότι αν η κλίση της f είναι στον $L^2(\log L)^{-1}$ τότε η f είναι στον L^2. Αυτή η πτυχή της ανισότητας τονίζεται στο [17]. Εναλλακτικές μορφές της L^1-L^2 ανισότητας, οι οποίες προκύπτουν άμεσα από την λογαριθμική ανισότητα Sobolev, παρουσιάζονται στα [37] και [79].

Ανισότητα μεταφοράς με τετραγωνικό κόστος. Έστω $(X,\mathcal{A})$, $(Y,\mathcal{B})$ δύο μετρήσιμοι χώροι. Θεωρούμε τον χώρο γινόμενο $(X\times Y,\mathcal{A}\otimes\mathcal{B})$ και ένα μέτρο $\pi:\mathcal{A}\otimes\mathcal{B}\to[0,\infty]$ στον $(X\times Y,\mathcal{A}\otimes\mathcal{B})$. Τότε:

(α) Το $\pi_X:\mathcal{A}\to[0,\infty]$ με $\pi_X(A)=\pi(A\times Y)$ είναι μέτρο στον μετρήσιμο χώρο $(X,\mathcal{A})$ και θα το λέμε *πρώτο περιθώριο μέτρο* του π.

(β) Το $\pi_Y:\mathcal{B}\to[0,\infty]$ με $\pi_Y(B)=\pi(X\times B)$ είναι μέτρο στον μετρήσιμο χώρο $(Y,\mathcal{B})$ και θα το λέμε *δεύτερο περιθώριο μέτρο* του π.

Έστω τώρα μ,ν δύο Borel μέτρα πιθανότητας στον $\mathbb{R}^n$. Θεωρούμε την $w:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ με $w(x,y)=\sum_{i=1}^n(x_i-y_i)^2$. Ορίζουμε

$$T_w(\mu,\nu)=\inf\int_{\mathbb{R}^n\times\mathbb{R}^n}w(x,y)\,d\pi(x,y)$$

όπου το infimum είναι πάνω από όλα τα Borel μέτρα πιθανότητας π στον $\mathbb{R}^n\times\mathbb{R}^n$ για τα οποία έχουμε $\pi_X=\mu$ και $\pi_Y=\nu$. Έστω επίσης $d\gamma(x)=\frac{1}{(2\pi)^{n/2}}e^{-\|x\|_2^2/2}dx$ το τυπικό μέτρο Gauss στα Borel υποσύνολα του $\mathbb{R}^n$.

Τότε, για κάθε μέτρο πιθανότητας $\mu\ll\gamma$ (απολύτως συνεχές ως προς το γ_n) με $f=\frac{d\mu}{d\gamma}$ ισχύει ότι

$$T_w(\mu,\gamma)\leqslant 2\int_{\mathbb{R}^n}f\,\log f\,d\gamma=2\int_{\mathbb{R}^n}\log f\,d\mu.$$

Η ανισότητα μεταφοράς με τετραγωνικό κόστος αποτέλεσε θεμέλιο λίθο στη σχέση μεταξύ των μερικών διαφορικών εξισώσεων, της θεωρίας πιθανοτήτων και της γεωμετρίας, όπως τονίστηκε για πρώτη φορά από τους Otto και Villani, και σε συνδυασμό με την λογαριθμική ανισότητα Sobolev οδήγησε σε κάτω φράγματα για την καμπυλότητα και στην ανάλυση μετρικών χώρων μέτρου που αναπτύχθηκε από τους Lott, Villani, Sturm, Ambrosio, Gigli, Savaré.

Η αρχική απόδειξη της ανισότητας στο [95] χρησιμοποιεί μεταφορά του μέτρου στη μία διάσταση, και το πέρασμα στις μεγαλύτερες διαστάσεις γίνεται με ένα επιχείρημα tensorization. Μάλιστα, ο

Talagrand αποδεικνύει ισχυρότερα αντίστοιχα αποτελέσματα για γινόμενα του εκθετικού μέτρου, και χρησιμοποιεί την περίπτωση του μέτρου Gauss ως ένα απλούστερο παράδειγμα για να παρουσιάσει τις ιδέες του. Περιγράφουμε το επιχείρημα στο Κεφάλαιο 5. Αφετηρία του γι' αυτά τα αποτελέσματα ήταν η σύνδεση μεταξύ ανισοτήτων της θεωρίας της μεταφοράς του μέτρου και του φαινομένου της συγκέντρωσης του μέτρου, η οποία είχε παρατηρηθεί από την Marton στα [59] και [60].

Η παρατήρηση ότι η ανισότητα του Talagrand για τη μεταφορά με τετραγωνικό κόστος είναι συνέπεια της λογαριθμικής ανισότητας Sobolev, οφείλεται στους Otto και Villani, οι οποίοι στο [72] απέδειξαν αυτή τη συνεπαγωγή σε ένα πολύ γενικό πλαίσιο. Η προσέγγισή τους βασίζεται σε παλιότερη δουλειά του Otto, και έπαιξε πολύ σημαντικό ρόλο στη μελέτη συναρτησιακών ανισοτήτων σε μετρικούς χώρους πιθανότητας (βλ. [101], [102], [9]). Η δεύτερη απόδειξη που παρουσιάζουμε στο Κεφάλαιο 5 για την ανισότητα του Talagrand, προέρχεται από τα [15], [16]. Ένα ενδιαφέρον ερώτημα είναι αν μπορεί κανείς να αποδείξει και τις υπόλοιπες τρεις ανισότητες αυτής της εργασίας με μεθόδους μεταφοράς του μέτρου, και το ερώτημα αυτό μελετάται ως ένα βαθμό στο [22, Κεφάλαιο 8].

Ανισότητα για το supremum εμπειρικών κατανομών. Έστω $(\Omega, \mathcal{A})$ ένας μετρήσιμος χώρος και $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές με τιμές στον Ω. Έστω $\mathcal{F}$ μια αριθμήσιμη οικογένεια μετρήσιμων συναρτήσεων ορισμένων στο Ω. Θεωρούμε την τυχαία μεταβλητή

$$Z = \sup_{f \in \mathcal{F}} \sum_{i=1}^{n} f(X_i)$$

και θέτουμε

$$U = \sup_{f \in \mathcal{F}} \|f\|_\infty \quad \text{και} \quad V = \mathbb{E}\Big(\sup_{f \in \mathcal{F}} \sum_{i=1}^{n} f(X_i)^2 \Big).$$

Τότε, υπάρχει σταθερά $K > 0$ ώστε για κάθε $t > 0$ να ισχύει ότι

$$\mathbb{P}\Big(|Z - \mathbb{E}(Z)| \geqslant t \Big) \leqslant K \exp\Big[- \frac{t}{KU} \log\Big(1 + \frac{tU}{V}\Big)\Big].$$

Η ανισότητα για το supremum εμπειρικών διαδικασιών αποτελεί σημαντικό και απαραίτητο εργαλείο για την απειροδιάστατη στατιστική. Τα βιβλία [63], [64], [22], [39] και [103] δίνουν διάφορα παραδείγματα εφαρμογών στην σύγχρονη στατιστική.

Στο Κεφάλαιο 6 περιγράφουμε την αρχική απόδειξη που έδωσε ο Talagrand στο [96], η οποία είναι εξαιρετικά πολύπλοκη. Στο [74] δόθηκε μια απλούστερη απόδειξη, η οποία συνδυάζει την ανισότητα απόστασης από την κυρτή θήκη με ένα επιχείρημα συμμετρικοποίησης. Η δεύτερη απόδειξη που παρουσιάζουμε βασίζεται στη μέθοδο της εντροπίας και τη λογαριθμική ανισότητα Sobolev, εμφανίστηκε για πρώτη φορά στο [51], και πήρε ακριβή μορφή στο [62] και σε διάφορες μεταγενέστερες εργασίες. Στο άρθρο [55] του Massart γίνεται πολύ προσεκτική ανάλυση των αριθμητικών σταθερών που υπεισέρχονται στην ανισότητα και οι οποίες είναι αρκετά σημαντικές για τις εφαρμογές. Στο βιβλίο [22] παρουσιάζονται διάφορες άλλες ανισότητες για εμπειρικές ανελίξεις και δίνονται αναφορές στις δουλειές που διαδοχικά βελτίωσαν τις τιμές των σταθερών σε όλες αυτές τις ανισότητες.

1.2 Λίγα λόγια για τη σημασία και την προέλευση των ανισοτήτων του Talagrand

Τα μαθηματικά επιτεύγματα του M. Talagrand επηρέασαν σημαντικά τις επιστημονικές εξελίξεις των τελευταίων δεκαετιών. Το έργο του περιλαμβάνει αποτελέσματα σε διάφορους τομείς όπως η θεωρία μέτρου, η συναρτησιακή ανάλυση, η γεωμετρία χώρων Banach, οι στοχαστικές ανελίξεις, η Boolean ανάλυση, οι ανισότητες συγκέντρωσης, οι ισοπεριμετρικές ανισότητες κλπ.

Οι ανισότητες του Talagrand, εκτός ίσως από την L^1-L^2-ανισότητα διασποράς, είναι εμπνευσμένες από το φαινόμενο της συγκέντρωσης του μέτρου και τις εφαρμογές του στη γεωμετρία χώρων Banach και τις πιθανότητες σε χώρους Banach. «Η ιδέα της συγκέντρωσης του μέτρου, που επινοήθηκε από τον V. Milman, είναι αναμφίβολα μία από τις πιο σπουδαίες ιδέες όλων των εποχών στην Ανάλυση», αναφέρει χαρακτηριστικά ο Talagrand.

Κύριο χαρακτηριστικό των ανισοτήτων του Talagrand αποτελεί το γεγονός ότι είναι ελεύθερες διάστασης, δηλαδή οι σταθερές δεν εξαρτώνται από το μέγεθος του δείγματος – υπάρχουν μάλιστα και επεκτάσεις τους σε απειροδιάστατα συστήματα.

Μάλιστα, είναι εμπνευσμένες (αλλά ταυτόχρονα παρέχουν ισχυρές προεκτάσεις της) από την ακόλουθη ανισότητα συγκέντρωσης για το μέτρο Gauss: Έστω $F : \mathbb{R}^n \to \mathbb{R}$ Lipschitz συνεχής συνάρτηση με σταθερά Lipschitz

$$\|F\|_{\mathrm{Lip}} = \sup\Big\{\frac{|F(x)-F(y)|}{\|x-y\|_2} : x,y\in\mathbb{R}^n, x\neq y\Big\}.$$

(Η $\|\cdot\|_{\mathrm{Lip}}$ αποτελεί μάλιστα ημινόρμα.) Τότε,

$$\gamma\Big(F \geqslant \int_{\mathbb{R}^n} F\,d\gamma + t\Big) \leqslant \exp\Big(-\frac{t^2}{2\|F\|_{\mathrm{Lip}}^2}\Big)$$

για κάθε $t\geqslant 0$, όπου

$$d\gamma(x) = \frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2} dx$$

είναι το τυπικό μέτρο Gauss για τα σύνολα Borel του $\mathbb{R}^n$.

ΚΕΦΑΛΑΙΟ 2

Εντροπία, λογαριθμική ανισότητα Sobolev, υπερσυσταλτότητα

Στα παρακάτω ο $(\Omega, \mathcal{A}, \mu)$ είναι ένας χώρος πιθανότητας και, για κάθε $1 \leqslant p \leqslant \infty$, η $\|\cdot\|_p$ είναι η συνήθης νόρμα του $L^p(\mu)$.

2.1 Εντροπία

Ορισμός 2.1.1 (εντροπία)**.** Έστω $f : \Omega \to \mathbb{R}$ μετρήσιμη συνάρτηση με $f \geqslant 0$ και $f \in L^1(\mu)$. Ορίζουμε την *εντροπία* $\mathrm{Ent}_\mu(f)$ *της* f *ως προς* μ ως εξής:

$$\mathrm{Ent}_\mu(f) = \int_\Omega f \log f \, d\mu - \int_\Omega f \, d\mu \cdot \log\Big(\int_\Omega f \, d\mu\Big).$$

Λήμμα 2.1.2. *Η* $\varphi : (0, \infty) \to \mathbb{R}$, $\varphi(x) = x \log x$ *είναι κυρτή, κάτω φραγμένη και επεκτείνεται συνεχώς στο* 0 *αν θέσουμε* $\varphi(0) = 0$. *Θα γράφουμε* $0 \cdot \log 0 = 0$.

Απόδειξη. Το γεγονός ότι $\lim_{x \to 0^+} \varphi(x) = 0$ είναι άμεσο.

Παρατηρούμε ότι $\varphi'(x) = \log x + 1$ και $\varphi''(x) = \frac{1}{x} > 0$ για κάθε $x > 0$. Οπότε, η φ είναι κυρτή και $\min(\varphi) = \varphi(e^{-1}) = -\frac{1}{e}$. Άρα, η φ είναι και κάτω φραγμένη. □

Παρατήρηση 2.1.3. Θέτουμε $g = f \log f$. Τότε,

$$g^- = -f \log f \cdot \mathbf{1}_{\{f<1\}} \leqslant e^{-1}$$

από το Λήμμα 2.1.2. Άρα,

$$\int_\Omega g^- d\mu \leqslant e^{-1} < \infty.$$

Επίσης, προφανώς,

$$\int_\Omega f \, d\mu \cdot \log\Big(\int_\Omega f \, d\mu\Big) = \varphi\Big(\int_\Omega f \, d\mu\Big) \in \mathbb{R}.$$

Άρα, η $\mathrm{Ent}_\mu(f)$ ορίζεται καλά και παίρνει τιμές στο $\mathbb{R} \cup \{\infty\}$.

Πρόταση 2.1.4 (ιδιότητες της εντροπίας). *Έστω $f : \Omega \to \mathbb{R}$ μετρήσιμη συνάρτηση με $f \geqslant 0$ και $f \in L^1(\mu)$.*

(i) $\mathrm{Ent}_\mu(f) \in [0, \infty]$.

(ii) *Η απεικόνιση $f \mapsto \mathrm{Ent}_\mu(f)$ είναι ομογενής τάξης 1.*

Απόδειξη. (i) Αφού η φ είναι κυρτή και $f \in L^1(\mu)$, από την ανισότητα Jensen έχουμε ότι

$$\int_\Omega \varphi(f)\, d\mu \geqslant \varphi\Big(\int_\Omega f\, d\mu\Big)$$

ή ισοδύναμα

$$\int_\Omega f \log f\, d\mu \geqslant \int_\Omega f\, d\mu \cdot \log\Big(\int_\Omega f\, d\mu\Big).$$

Συνεπώς, $\mathrm{Ent}_\mu(f) \geqslant 0$.

(ii) Για κάθε $\lambda > 0$ έχουμε

$$\begin{aligned}
\mathrm{Ent}_\mu(\lambda f) &= \int_\Omega (\lambda f) \log(\lambda f)\, d\mu - \int_\Omega \lambda f\, d\mu \cdot \log\Big(\int_\Omega \lambda f\, d\mu\Big) \\
&= \lambda \int_\Omega f \log f\, d\mu + \lambda \log \lambda \int_\Omega f\, d\mu - \lambda \int_\Omega f\, d\mu \cdot \log \lambda - \lambda \int_\Omega f\, d\mu \cdot \log\Big(\int_\Omega f\, d\mu\Big) \\
&= \lambda\, \mathrm{Ent}_\mu(f).
\end{aligned}$$

Άρα, η εντροπία είναι ομογενής τάξης 1. □

Παρατήρηση 2.1.5. Έστω ότι για κάποιο μέτρο πιθανότητας ν στον $(\Omega, \mathcal{A})$ ισχύει ότι $d\nu = f\, d\mu$, δηλαδή $f = \frac{d\nu}{d\mu}$. Τότε,

$$\mathrm{Ent}_\mu(f) = \int_\Omega \log f\, d\nu.$$

Η ποσότητα αυτή ονομάζεται *σχετική εντροπία* του ν ως προς το μ και θα συμβολίζεται με $H(\nu \mid \mu)$.

Μια πολύ χρήσιμη περιγραφή της εντροπίας δίνεται από το επόμενο λήμμα.

Λήμμα 2.1.6. *Αν $f : \Omega \to [0, \infty)$ είναι μετρήσιμη συνάρτηση με $f \in L^1(\mu)$ τότε ισχύει ότι*

$$\mathrm{Ent}_\mu(f) = \sup\Big\{ \int fg\, d\mu \mid g \text{ φραγμένη μετρήσιμη με } \int e^g\, d\mu \leqslant 1 \Big\}.$$

Απόδειξη. Η σχέση που θέλουμε να δείξουμε είναι ομογενής, οπότε μπορούμε να υποθέσουμε χωρίς περιορισμό της γενικότητας ότι $\int_\Omega f\, d\mu = 1$.

Ισχυρισμός: Για κάθε $u \geqslant 0$ και για κάθε $v \in \mathbb{R}$ ισχύει ότι

$$uv \leqslant u \log u - u + e^v.$$

Πράγματι, έστω $v \in \mathbb{R}$. Θεωρούμε την $h : [0, \infty) \to \mathbb{R}$, $h(u) = u \log u - u - uv + e^v$.

Η h είναι παραγωγίσιμη στο $[0, \infty)$ με $h'(u) = \log u - v$. Άρα, $h(u) \geqslant h(e^v)$ για κάθε $u \geqslant 0$. Όμως,

$$h(e^v) = v\, e^v - e^v - v\, e^v + e^v = 0.$$

Οπότε, $h(u) \geqslant 0$ για κάθε $u \geqslant 0$ και δείξαμε τον ισχυρισμό.

Από τον ισχυρισμό έπεται ότι αν η $g : \Omega \to \mathbb{R}$ είναι μετρήσιμη και φραγμένη με $\int_\Omega e^g d\mu \leqslant 1$ τότε

$$\begin{aligned}\int_\Omega fg\,d\mu &\leqslant \int_\Omega f\log f\,d\mu - \int_\Omega f\,d\mu + \int_\Omega e^g\,d\mu\\ &= \int_\Omega f\log f\,d\mu - 1 + \int_\Omega e^g\,d\mu \leqslant \int_\Omega f\log f\,d\mu\\ &= \int_\Omega f\log f\,d\mu - \int_\Omega f\,d\mu\,\cdot\,\log\Big(\int_\Omega f\,d\mu\Big)\\ &= \mathrm{Ent}_\mu(f).\end{aligned}$$

Παίρνοντας το supremum βλέπουμε ότι

$$\sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη με } \int e^g\ d\mu \leqslant 1\Big\} \leqslant \mathrm{Ent}_\mu(f).$$

Για την αντίστροφη ανισότητα θεωρούμε τις $f_n = \min\left\{\max\left\{f, \frac{1}{n}\right\}, n\right\}$. Παρατηρούμε ότι

$$f_n(x) = \begin{cases} \max\left\{f(x), \frac{1}{n}\right\} & ,\text{ αν } f(x) < 1\\ 1 & ,\text{ αν } f(x) = 1\\ \min\{f(x), n\} & ,\text{ αν } f(x) > 1.\end{cases}$$

Θα δείξουμε ότι $f_n \longrightarrow f$ κατά σημείο. Πράγματι, έστω $x \in \Omega$. Αν $f(x) = 0$ τότε $f_n(x) = \frac{1}{n} \to 0 = f(x)$. Αν $f(x) > 0$ τότε υπάρχει $n_0 \in \mathbb{N}$ ώστε για κάθε $n \geqslant n_0$ να έχουμε $\frac{1}{n} < f(x) < n$ και τότε για κάθε $n \geqslant n_0$ έχουμε $f_n(x) = f(x) \to f(x)$.

Παρατηρούμε επίσης ότι $\frac{1}{n} \leqslant f_n \leqslant n$ για κάθε $n \in \mathbb{N}$. Οπότε, $\int_\Omega f_n\,d\mu > 0$ για κάθε $n \in \mathbb{N}$, και από το θεώρημα κυριαρχημένης σύγκλισης έχουμε

$$\int_\Omega f_n\,d\mu \longrightarrow \int_\Omega f\,d\mu.$$

Ορίζονται λοιπόν καλά οι $g_n := \log\left(\frac{f_n}{\int_\Omega f_n\,d\mu}\right)$ για $n \in \mathbb{N}$, είναι μετρήσιμες, με $\int_\Omega e^{g_n} d\mu = 1$, και προφανώς είναι φραγμένες λόγω των φραγμάτων που ικανοποιούν οι f_n και συνεπώς των φραγμάτων των ολοκληρωμάτων τους. Οπότε,

$$\int_\Omega fg_n\,d\mu \leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη με } \int e^g\ d\mu \leqslant 1\Big\}.$$

Έχουμε ότι

$$\int_\Omega fg_n\,d\mu = \int_\Omega f\log f_n\,d\mu - \int_\Omega f\,d\mu\cdot\log\Big(\int_\Omega f_n\,d\mu\Big).$$

Τώρα,

$$\int_\Omega f\,d\mu\cdot\log\Big(\int_\Omega f_n\,d\mu\Big) \longrightarrow \int_\Omega f\,d\mu\cdot\log\Big(\int_\Omega f\,d\mu\Big) = 1\cdot\log 1 = 0.$$

Επίσης,

$$\begin{aligned}\int_\Omega f\log f_n\,d\mu &= \int_{\{f>0\}} f\log f_n\,d\mu\\ &= \int_{\{0<f<1\}} f\log\Big(\max\Big\{f, \frac{1}{n}\Big\}\Big)\,d\mu + \int_{\{f>1\}} f\log(\min\{f, n\})\,d\mu.\end{aligned}$$

Τώρα, από το θεώρημα μονότονης σύγκλισης έχουμε ότι

$$\int_{\{f>1\}} f\log(\min\{f,n\})\,d\mu \longrightarrow \int_{\{f>1\}} f\log f\,d\mu,$$

και, πάλι από το θεώρημα μονότονης σύγκλισης, έχουμε ότι

$$\begin{aligned}\int_{\{0<f<1\}} f\log\Big(\max\Big\{f,\frac{1}{n}\Big\}\Big)\,d\mu &= -\int_{\{0<f<1\}} f\log\Big(\max\Big\{f,\frac{1}{n}\Big\}^{-1}\Big)\,d\mu\\ &\longrightarrow -\int_{\{0<f<1\}} f\log\frac{1}{f}\,d\mu = \int_{\{0<f<1\}} f\log f\,d\mu.\end{aligned}$$

Από όλα τα παραπάνω συμπεραίνουμε ότι

$$\int_\Omega fg_n\,d\mu \longrightarrow \int_{\{f>0\}} f\log f\,d\mu = \int_\Omega f\log f\,d\mu = \mathrm{Ent}_\mu(f).$$

Οπότε,

$$\mathrm{Ent}_\mu(f) \leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη με } \int e^g\,d\mu \leqslant 1\Big\}$$

και η απόδειξη είναι πλήρης. □

Παρατήρηση 2.1.7. Στο πρώτο σκέλος της απόδειξης, η υπόθεση ότι η g είναι φραγμένη δεν χρειάστηκε. Αρκούσε να υποθέσουμε ότι η fg είναι ολοκληρώσιμη. Επίσης, στο δεύτερο σκέλος, η προσέγγιση έγινε από ακολουθία $(g_n)_{n\in\mathbb{N}}$ φραγμένων μετρήσιμων συναρτήσεων με $\int e^{g_n}d\mu = 1$ για κάθε $n\in\mathbb{N}$. Συνεπώς, έχουμε δείξει το εξής:

$$\begin{aligned}&\sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη}, \int e^g\,d\mu = 1\Big\}\\ &\quad\leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ μετρήσιμη}, fg \text{ ολοκληρώσιμη}, \int e^g\,d\mu = 1\Big\}\\ &\quad\leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ μετρήσιμη}, fg \text{ ολοκληρώσιμη}, \int e^g\,d\mu \leqslant 1\Big\}\\ &\quad\leqslant \mathrm{Ent}_\mu(f)\\ &\quad\leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη}, \int e^g\,d\mu \leqslant 1\Big\}\\ &\quad\leqslant \sup\Big\{\int fg\,d\mu \mid g \text{ μετρήσιμη}, fg \text{ ολοκληρώσιμη}, \int e^g\,d\mu \leqslant 1\Big\}.\end{aligned}$$

Άρα τελικά έχουμε την ισότητα

$$\begin{aligned}\mathrm{Ent}_\mu(f) &= \sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη}, \int e^g\,d\mu = 1\Big\}\\ &= \sup\Big\{\int fg\,d\mu \mid g \text{ μετρήσιμη}, fg \text{ ολοκληρώσιμη}, \int e^g\,d\mu \leqslant 1\Big\}\\ &= \sup\Big\{\int fg\,d\mu \mid g \text{ φραγμένη μετρήσιμη}, \int e^g\,d\mu \leqslant 1\Big\}.\end{aligned}$$

Πρόταση 2.1.8. *Αν $f : \Omega \to \mathbb{R}$ είναι μετρήσιμη συνάρτηση με $\int_\Omega f\,d\mu = 1$ και η g είναι μετρήσιμη συνάρτηση ώστε η fg να είναι ολοκληρώσιμη, τότε*

$$\int_\Omega fg\,d\mu \leqslant \mathrm{Ent}_\mu(f) + \log\Big(\int_\Omega e^g d\mu\Big).$$

Απόδειξη. Έχουμε ότι

$$\begin{aligned}\int_\Omega fg\,d\mu - \int_\Omega f\log f\,d\mu &= \int_\Omega (f\log(e^g) - f\log f)\,d\mu \\ &= \int_\Omega f\log\Big(\frac{e^g}{f}\Big)\,d\mu,\end{aligned}$$

με τη σύμβαση ότι $0 \cdot \log\Big(\frac{e^{g(x)}}{0}\Big) = 0$. Τώρα, αφού $\int_\Omega f\,d\mu = 1$, η f είναι σχεδόν παντού πεπερασμένη, και αφού η λογαριθμική συνάρτηση είναι κοίλη, από την ανισότητα Jensen έχουμε ότι

$$\int_\Omega f\log\Big(\frac{e^g}{f}\Big)\,d\mu \leqslant \log\Big(\int_\Omega e^g d\mu\Big).$$

Άρα,

$$\int_\Omega fg\,d\mu - \mathrm{Ent}_\mu(f) \leqslant \log\Big(\int_\Omega e^g d\mu\Big),$$

το οποίο είναι το ζητούμενο. □

Λήμμα 2.1.9. *Ισχύει ότι*

$$\mathrm{Ent}_\mu(f) = \inf_{c>0}\Big\{\int_\Omega [f(\log f - \log c) - (f - c)]\,d\mu\Big\}.$$

Απόδειξη. Για $c > 0$ έχουμε ότι

$$\begin{aligned}\int_\Omega [f(\log f - \log c) - (f - c)]\,d\mu &= \int_\Omega f\log f\,d\mu + c - (\log c + 1)\int_\Omega f\,d\mu \\ &= \int_\Omega f\log f\,d\mu + h(c),\end{aligned}$$

όπου $h(c) := c - (\log c + 1)\int_\Omega f\,d\mu$. Τώρα, $h'(c) = 1 - \frac{1}{c}\int_\Omega f\,d\mu$, άρα

$$\min_{c>0} h(c) = h\Big(\int_\Omega f\,d\mu\Big).$$

Έπεται ότι

$$\begin{aligned}\inf_{c>0}\Big\{\int_\Omega [f(\log f - \log c) - (f - c)]\,d\mu\Big\} &= \int_\Omega f\log f\,d\mu + h\Big(\int_\Omega f\,d\mu\Big) \\ &= \int_\Omega f\log f\,d\mu - \int_\Omega f\,d\mu \cdot \log\Big(\int_\Omega f\,d\mu\Big) = \mathrm{Ent}_\mu(f).\end{aligned}$$

□

2.1.1 Εντροπία σε χώρους γινόμενα

Έστω $(\Omega_i, \mathcal{A}_i, \mu_i)$, $1 \leqslant i \leqslant n$ χώροι πιθανότητας. Θεωρούμε τον χώρο γινόμενο $(X, \mathcal{A}, P) = (\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \mu_i)$ εφοδιασμένο με το μέτρο γινόμενο,

Αν $f : X \to \mathbb{R}$ είναι συνάρτηση ορισμένη στον X και $1 \leqslant i \leqslant n$, σταθεροποιούμε $x_1, \ldots, x_{i-1}$, $x_{i+1}, \ldots, x_n$ και ορίζουμε $f_i : \Omega_i \to \mathbb{R}$ ως εξής:

$$f_i(x_i) = f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

Παρατήρηση 2.1.10. (α) Αν η f είναι μετρήσιμη τότε για να είναι η f_i μετρήσιμη αρκεί να ισχύει ότι $\{x_j\} \in \mathcal{A}_j$ για κάθε $j \in [n] \setminus \{i\}$, αφού

$$f_i^{-1}((-\infty, \beta]) = f^{-1}((-\infty, \beta] \cap (\{x_1\} \times \cdots \times \{x_{i-1}\} \times \Omega_i \times \{x_{i+1}\} \times \cdots \times \{x_n\})).$$

Αυτό θα υποθέτουμε ότι ισχύει παρακάτω χωρίς βλάβη της γενικότητας.

(β) Αν $f \in L^q(P)$ για κάποιον $1 \leqslant q \leqslant \infty$ τότε $f_i \in L^q(\mu_i)$ για κάθε $i = 1, \ldots, n$. Πράγματι, αν $1 \leqslant q < \infty$ και $f \in L^q(\mu)$ έχουμε ότι

$$+\infty > \int_X |f|^q \, d\mu = \int_{\prod\limits_{j\neq i} \Omega_j} \int_{\Omega_i} |f(x_1, \ldots, x_n)|^q d\mu_i(x_i) d \otimes_{j\neq i} m_j(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Από το θεώρημα Tonelli έχουμε ότι

$$\int_{\Omega_i} |f(x_1, \ldots, x_n)|^q \, d\mu_i(x_i) = \int_{\Omega_i} |f_i(x_i)|^q d\mu_i(x_i) < \infty \tag{2.1.1}$$

για $\otimes_{j\neq i}\mu_j$-σχεδόν κάθε $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \prod_{j\neq i} \Omega_j$. Οπότε, επιλέγοντας μια τέτοια διατεταγμένη $(n-1)$-άδα έχουμε το ζητούμενο. Μάλιστα μπορούμε, χωρίς περιορισμό της γενικότητας, να υποθέσουμε ότι η (2.1.1) ισχύει πάντα, άρα έχουμε το ζητούμενο για οποιαδήποτε επιλογή των $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.

Για $q = +\infty$ μπορούμε να υποθέσουμε ότι η f είναι φραγμένη στο X και το ζητούμενο έπεται άμεσα.

Πρόταση 2.1.11. *Έστω $f : X \to [0, \infty)$ με $f \in L^1(P)$. Τότε,*

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^n \int \mathrm{Ent}_{\mu_i}(f_i) \, dP,$$

όπου στο δεξιό μέλος η ολοκλήρωση ως προς dP για κάθε $1 \leqslant i \leqslant n$ γίνεται πάνω στις συντεταγμένες $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

Απόδειξη. Έστω $g : X \to \mathbb{R}$ μετρήσιμη και φραγμένη, με $\int e^g \, dP \leqslant 1$. Ορίζουμε

$$g^1(x_1, \ldots, x_n) = \log\left(\frac{e^{g(x_1,\ldots,x_n)}}{\int_X e^{g(y_1,x_2,\ldots,x_n)} d\mu_1(y_1)}\right)$$

και, για κάθε $i = 2, \ldots, n$,

$$g^i(x_1, \ldots, x_n) = \log\left(\frac{\int e^{g(y_1,\ldots,y_{i-1},x_i,\ldots,x_n)} \, d\mu_1(y_1) \ldots d\mu_{i-1}(y_{i-1})}{\int e^{g(y_1,\ldots,y_i,x_{i+1},\ldots,x_n)} \, d\mu_1(y_1) \ldots d\mu_i(y_i)}\right).$$

Παρατηρούμε ότι

$$\sum_{i=1}^{n} g^i = \log\left(\frac{e^g}{\int e^g\, dP}\right) = g - \log\Big(\int_X e^g dP\Big) \geqslant g - \log 1 = g.$$

Άρα, $g \leqslant \sum_{i=1}^n g^i$. Επίσης, εύκολα βλέπουμε ότι

$$\int_{\Omega_i} e^{(g^i)_i} d\mu_i = 1$$

και ότι οι $(g^i)_i$ είναι φραγμένες αφού η g είναι φραγμένη. Τώρα,

$$\begin{aligned}\int_X fg\, dP &\leqslant \sum_{i=1}^{n}\int_X fg^i\, dP = \sum_{i=1}^{n}\int_X \left(\int_{\Omega_i} f_i (g^i)_i\, d\mu_i\right) dP\\ &\leqslant \sum_{i=1}^{n}\int_X \mathrm{Ent}_{\mu_i}(f_i)\, dP,\end{aligned}$$

και συμπεραίνουμε ότι

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^{n}\int_X \mathrm{Ent}_{\mu_i}(f_i)\, dP$$

από το Λήμμα 2.1.6. □

Λήμμα 2.1.12. *Έστω $(\Omega, \mathcal{A}, \mu)$ χώρος πιθανότητας και $f \in L^1(\mu)$ με $f \geqslant 0$. Τότε,*

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2}\int_\Omega\int_\Omega [f(x) - f(y)]\,[\log f(x) - \log f(y)]\, d\mu(x)\, d\mu(y).$$

Απόδειξη. Παρατηρούμε ότι

$$\begin{aligned}&\frac{1}{2}\int_\Omega\int_\Omega [f(x) - f(y)]\,[\log f(x) - \log f(y)]\, d\mu(x)\, d\mu(y)\\ &\qquad = \frac{1}{2}\Big[2\int_\Omega f\log f\, d\mu - 2\int_\Omega f\, d\mu \cdot \int_\Omega \log f\, d\mu\Big]\\ &\qquad = \int_\Omega f\log f\, d\mu - \int_\Omega f\, d\mu \cdot \int_\Omega \log f\, d\mu\\ &\qquad \geqslant \int_\Omega f\log f\, d\mu - \int_\Omega f\, d\mu \cdot \log\Big(\int_\Omega f\, d\mu\Big)\\ &\qquad = \mathrm{Ent}_\mu(f),\end{aligned}$$

από την ανισότητα Jenssen, αφού η λογαριθμική συνάρτηση είναι κοίλη. □

Πρόταση 2.1.13. *Έστω $(X, \mathcal{A}, P)$ ο χώρος γινόμενο που ορίσαμε παραπάνω, και $f : X \to [0, \infty)$ με $f \in L^1(P)$. Τότε:*

(i) *Ισχύει ότι*

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^{n}\int_X \iint_{\{f_i(x_i)\geqslant f_i(y_i)\}} [f_i(x_i) - f_i(y_i)]\,[\log f_i(x_i) - \log f_i(y_i)]\, d\mu_i(x_i)\, d\mu_i(y_i)\, dP.$$

(ii) *Ισχύει ότι*

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^{n} \int_X \inf_{c_i>0} \int_{\Omega_i} [f_i(\log f_i - \log c_i) - (f_i - c_i)]\, d\mu_i\, dP.$$

Απόδειξη. (i) Από την Πρόταση 2.1.11 έχουμε ότι

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^{n} \int_X \mathrm{Ent}_{\mu_i}(f_i)\, dP.$$

Το ζητούμενο έπεται άμεσα με εφαρμογή του Λήμματος 2.1.12 παίρνοντας υπόψιν και τη συμμετρία των συναρτήσεων εντός των ολοκληρωμάτων.

(ii) Έπεται άμεσα με χρήση της Πρότασης 2.1.11 και του Λήμματος 2.1.9. □

Ορισμός 2.1.14 (διασπορά)**.** Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $f \in L^2(\mu)$. Τότε ορίζουμε την *διασπορά* της f

$$\mathrm{Var}_\mu(f) = \int_X \Big(f - \int_X f\, d\mu)^2\, d\mu = \int_X f^2\, d\mu - \Big(\int_X f\, d\mu\Big)^2.$$

Η διασπορά έχει τις ακόλουθες βασικές ιδιότητες:

(i) $\mathrm{Var}_\mu(f) \in [0, +\infty]$.

(ii) $\mathrm{Var}_\mu(f) < \infty$ αν και μόνο αν $f \in L^2(\mu)$.

(iii) $\mathrm{Var}_\mu(f) = 0$ αν και μόνο αν $f \equiv$ σταθερά.

(iv) $\mathrm{Var}_\mu(af + b) = a^2 \mathrm{Var}_\mu(f)$ για κάθε $a, b \in \mathbb{R}$.

Θεώρημα 2.1.15 (ανισότητα Efron-Stein)**.** *Έστω $(X, \mathcal{A}, P)$ ο χώρος γινόμενο που ορίσαμε παραπάνω και έστω $f \in L^2(P)$. Τότε,*

$$\mathrm{Var}_P(f) \leqslant \sum_{i=1}^{n} \int_X \mathrm{Var}_{\mu_i}(f_i)\, dP.$$

Απόδειξη. Μπορούμε να υποθέσουμε ότι η f είναι φραγμένη και ότι $\int_X f\, dP = 0$ και $\int_X f^2\, dP = 1$, οπότε $\mathrm{Var}_P(f) = 1$. Παρατηρούμε επίσης ότι

$$\int_X \mathbb{E}_{\mu_i}(f_i)\, dP = \int_X f\, dP = 0 \quad \text{και} \quad \int_X \mathbb{E}_{\mu_i}(f_i^2)\, dP = \int_X f^2\, dP = 1 \tag{2.1.2}$$

για κάθε $i = 1, \ldots, n$. Θεωρούμε την συνάρτηση $g = (1 + \varepsilon f)^2$ και παρατηρούμε ότι

$$\begin{aligned}
\frac{1}{\varepsilon^2}\mathrm{Ent}_P(g) &= \frac{1}{\varepsilon^2}\Big(2\int_X (1+\varepsilon f)^2 \log(1+\varepsilon f)\, dP - \int_X (1+\varepsilon f)^2\, dP \log\Big(\int_X (1+\varepsilon f)^2\, dP\Big)\Big) \\
&= \frac{1}{\varepsilon^2}\Big(2\int_X (\varepsilon f + \frac{3}{2}\varepsilon^2 f^2 + O(\varepsilon^3))\, dP - (1+\varepsilon^2)\log(1+\varepsilon^2)\Big) \\
&= \frac{1}{\varepsilon^2}\Big(3\varepsilon^2 - \varepsilon^2 + O(\varepsilon^3)\Big) = 2 + O(\varepsilon),
\end{aligned}$$

δηλαδή
$$\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon^2}\mathrm{Ent}_P(g) = 2\mathrm{Var}_P(f).$$
Κάνοντας παρόμοιο υπολογισμό για τις $\mathrm{Ent}_{\mu_i}(g_i) = \mathrm{Ent}_{\mu_i}((1+\varepsilon f_i)^2)$, $i = 1, \dots, n$, ολοκληρώνοντας στον X ως προς P, προσθέτοντας και παίρνοντας υπ' όψιν μας τις (2.1.2), βλέπουμε ότι
$$\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon^2}\sum_{i=1}^{n}\int_X \mathrm{Ent}_{\mu_i}(g_i)\,dP = 2\sum_{i=1}^{n}\int_X \mathrm{Var}_{\mu_i}(f_i)\,dP.$$
Όμως, από την Πρόταση 2.1.11 έχουμε
$$\mathrm{Ent}_P(g) \leqslant \sum_{i=1}^{n}\int \mathrm{Ent}_{\mu_i}(g_i)\,dP,$$
συνεπώς, $\mathrm{Var}_P(f) \leqslant \sum_{i=1}^{n}\int_X \mathrm{Var}_{\mu_i}(f_i)\,dP$. □

2.2 Ημιομάδες Markov

Ορισμός 2.2.1. Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας. Μια οικογένεια $(P_t)_{t\geqslant 0}$ γραμμικών τελεστών που ορίζονται στην κλάση $\mathcal{L}_{b,\mu}(X) = \{f : X \to \mathbb{R} \mid f \text{ φραγμένη και μετρήσιμη}\}$ λέγεται *ημιομάδα Markov* αν ικανοποιεί τα εξής:

(i) Για κάθε $f \in \mathcal{L}_{b,\mu}(X)$ και για κάθε $t \geqslant 0$ ισχύει ότι $P_t(f) \in \mathcal{L}_{b,\mu}(X)$.

(ii) Ο P_0 είναι ο ταυτοτικός τελεστής.

(iii) $P_{t+s} = P_t \circ P_s$ για κάθε $t, s \geqslant 0$.

(iv) $P_t(\mathbf{1}) = \mathbf{1}$ για κάθε $t \geqslant 0$, όπου $\mathbf{1}$ είναι η σταθερή συνάρτηση με τιμή 1.

(v) Αν $f \in \mathcal{L}_{b,\mu}(X)$ με $f \geqslant 0$ τότε για κάθε $t \geqslant 0$ ισχύει ότι $P_t(f) \geqslant 0$.

Ορισμός 2.2.2. Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και $(P_t)_{t\geqslant 0}$ μια ημιομάδα Markov. Το μέτρο πιθανότητας μ λέγεται *αναλλοίωτο* για την $(P_t)_{t\geqslant 0}$ αν
$$\int_X P_t f\,d\mu = \int_X f\,d\mu$$
για κάθε $f \in \mathcal{L}_{b,\mu}(X)$.

Λήμμα 2.2.3. *Αν η $\varphi : I \to \mathbb{R}$ είναι κυρτή συνάρτηση, όπου I ανοικτό διάστημα στο $\mathbb{R}$, τότε υπάρχει οικογένεια $(f_a)_{a\in I}$ γραμμικών συναρτήσεων ώστε $\varphi = \sup\limits_{a\in I} f_a$.*

Απόδειξη. Έστω $a \in I$. Αφού η φ είναι κυρτή συνάρτηση, έχει ευθεία στήριξης στο $(a, \varphi(a))$, δηλαδή υπάρχει ευθεία $(\varepsilon_a) : y = \lambda_a x + b_a$ που διέρχεται από το $(a, \varphi(a))$ και $\varphi(x) \geqslant \lambda_a x + b_a$ για κάθε $x \in I$. Ορίζουμε $f_a : I \to \mathbb{R}$ με $f_a(x) = \lambda_a x + b_a$ και θεωρούμε την οικογένεια γραμμικών συναρτήσεων $(f_a)_{a\in I}$.

Έστω $x \in I$. Για κάθε $a \in I$ έχουμε ότι $\varphi(x) \geqslant f_a(x)$ και $\varphi(x) = f_x(x)$. Συνεπώς, $\varphi(x) = \sup\limits_{a\in I} f_a(x)$. □

Πρόταση 2.2.4. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t \geqslant 0}$ ημιομάδα Markov. Αν η $\varphi : \mathbb{R} \to \mathbb{R}$ είναι κυρτή συνάρτηση τότε για κάθε $f \in \mathcal{L}_{b,\mu}(X)$ ισχύει ότι $P_t(\varphi(f)) \geqslant \varphi(P_t f)$ για κάθε $t \geqslant 0$.*

Απόδειξη. Αφού η φ είναι κυρτή, υπάρχει οικογένεια γραμμικών συναρτήσεων $(\psi_a)_{a \in A}$ ώστε $\varphi = \sup\limits_{a \in A} \psi_a$.

Έστω τώρα $x \in X$. Τότε, αφού κάθε ψ_a είναι γραμμική, για κάθε $a \in A$ έχουμε

$$\psi_a(P_t f(x)) = P_t(\psi_a(f))(x) \leqslant P_t\Big(\sup_{a \in A} \psi_a(f)\Big)(x) = P_t(\varphi(f))(x).$$

Συνεπώς, $\varphi(P_t f(x)) = \sup_{a \in A} \psi_a(P_t f(x)) \leqslant P_t(\varphi(f))(x)$. □

Πόρισμα 2.2.5. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t \geqslant 0}$ ημιομάδα Markov. Τότε, για κάθε $1 \leqslant p < \infty$ και για κάθε $f \in \mathcal{L}_{b,\mu}(X)$ ισχύει ότι*

$$|P_t f|^p \leqslant P_t(|f|^p).$$

Πρόταση 2.2.6. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t \geqslant 0}$ ημιομάδα Markov για την οποία το μ είναι αναλλοίωτο. Τότε, για κάθε $f \in \mathcal{L}_{b,\mu}(X)$ και για κάθε $1 \leqslant p \leqslant \infty$ ισχύει ότι $\|P_t f\|_p \leqslant \|f\|_p$.*

Απόδειξη. Έστω ότι το μ είναι αναλλοίωτο για την ημιομάδα Markov $(P_t)_{t \geqslant 0}$ και έστω $f \in \mathcal{L}_{b,\mu}(X)$ και $1 \leqslant p < \infty$. Τότε,

$$\|P_t f\|_p^p = \int_X |P_t f|^p \, d\mu \leqslant \int_X P_t(|f|^p) \, d\mu = \int_X |f|^p d\mu = \|f\|_p^p.$$

Συνεπώς,

$$\|P_t f\|_p \leqslant \|f\|_p.$$

Επίσης, αφού ο $(X, \mathcal{A}, \mu)$ είναι χώρος πιθανότητας και από τις $f, P_t f \in \mathcal{L}_{b,\mu}(X)$ έπεται ότι $f, P_t f \in L^p(\mu)$ για κάθε $1 \leqslant p < \infty$, παίρνοντας όριο για $p \to \infty$ στην προηγούμενη ανισότητα έχουμε ότι $\|P_t f\|_\infty \leqslant \|f\|_\infty$. □

Πρόταση 2.2.7. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t \geqslant 0}$ ημιομάδα Markov για την οποία το μ είναι αναλλοίωτο. Τότε, για κάθε $1 \leqslant p \leqslant \infty$ η $(P_t)_{t \geqslant 0}$ μπορεί να επεκταθεί στον $L^p(\mu)$ ώστε για κάθε $f \in L^p(\mu)$ να ισχύει ότι $\|P_t f\|_p \leqslant \|f\|_p$.*

Απόδειξη. Αν $f \in L^p(\mu)$ τότε υπάρχει ακολουθία $(s_n)_{n \in \mathbb{N}}$ από απλές μετρήσιμες συναρτήσεις ώστε $\|s_n - f\|_p \longrightarrow 0$. Προφανώς για κάθε $n \in \mathbb{N}$ ισχύει ότι $s_n \in \mathcal{L}_{b,\mu}(X)$ και

$$\|P_t(s_n) - P_t(s_m)\|_p = \|P_t(s_n - s_m)\|_p \leqslant \|s_n - s_m\|_p$$

για κάθε $m, n \in \mathbb{N}$. Η $(s_n)_{n \in \mathbb{N}}$ είναι βασική ως συγκλίνουσα, άρα η $(P_t(s_n))_{n \in \mathbb{N}}$ είναι βασική. Αφού ο $L^p(\mu)$ είναι χώρος Banach, υπάρχει $g \in L^p(\mu)$ ώστε $\|P_t(s_n) - g\|_p \to 0$.

Ορίζουμε $P_t(f) = g$. Αν $(r_n)_{n \in \mathbb{N}}$ είναι μια άλλη ακολουθία απλών μετρήσιμων συναρτήσεων με $\|r_n - f\|_p \to 0$ τότε $\|P_t(r_n) - P_t(s_n)\|_p \leqslant \|r_n - s_n\|_p$ άρα

$$\lim_{n \to \infty} P_t(r_n) = g = P_t(f)$$

στον $L^p(\mu)$. Συνεπώς, η $P_t(f)$ ορίζεται καλά. Τέλος, $\|P_t(s_n)\|_p \leqslant \|s_n\|_p$ για κάθε $n \in \mathbb{N}$, άρα $\|P_t f\|_p \leqslant \|f\|_p$. □

Παρατήρηση 2.2.8. Στον ορισμό της ημιομάδας Markov $(P_t)_{t\geqslant 0}$ με την ιδιότητα ότι το μ είναι αναλλοίωτο ως προς αυτήν (που θα παραλείπεται στο εξής) απαιτούμε επιπλέον και μια ιδιότητα συνέχειας: Για κάθε $f \in L^2(\mu)$ απαιτούμε να ισχύει $\lim\limits_{t\to 0} \|P_t f - f\|_2 = 0$.

Πόρισμα 2.2.9. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t\geqslant 0}$ ημιομάδα Markov. Τότε, για κάθε $f \in L^2(\mu)$ η συνάρτηση $t \mapsto \mathrm{Var}_\mu(P_t f)$ είναι φθίνουσα.*

Απόδειξη. Έστω $t > s \geqslant 0$. Τότε,

$$\begin{aligned}\mathrm{Var}_\mu(P_t f) &= \Big\|P_t(f) - \int_X P_t(f)\,d\mu\Big\|_2^2 = \Big\|P_t(f) - \int_X f\,d\mu\Big\|_2^2 = \Big\|P_t\Big(f - \int_X f\,d\mu\Big)\Big\|_2^2 \\ &= \Big\|P_{t-s}\Big(P_s\Big(f - \int_X f\,d\mu\Big)\Big)\Big\|_2^2 \leqslant \Big\|P_s\Big(f - \int_X f\,d\mu\Big)\Big\|_2^2 = \mathrm{Var}_\mu(P_s f).\end{aligned}$$

□

Ορισμός 2.2.10. Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t\geqslant 0}$ μια ημιομάδα Markov. Ορίζουμε $D(L)$ το σύνολο των $f \in L^2(\mu)$ για τις οποίες υπάρχει το $\lim\limits_{t\to 0^+} \frac{P_t f - f}{t}$ στον $L^2(\mu)$, και τον *γεννήτορα* $L : D(f) \to L^2(\mu)$ της $(P_t)_{t\geqslant 0}$,

$$L(f) = \lim_{t\to 0^+} \frac{P_t f - f}{t}.$$

Η κλάση $D(L)$ είναι το *πεδίο* του L.

Παρατήρηση 2.2.11. (α) Αποδεικνύεται ότι το σύνολο $D(L)$ είναι πυκνό στον $L^2(\mu)$ (θεώρημα Hille-Yosida).

(β) Αν $f \in D(L)$ τότε $P_t f \in D(L)$ για κάθε $t \geqslant 0$, και $L(P_t f) = P_t(L(f))$, δηλαδή οι L, P_t αντιμετατίθενται. Πράγματι,

$$\Big\|\frac{P_h(P_t f) - P_t f}{h} - P_t(L(f))\Big\|_2 = \Big\|P_t\Big(\frac{P_h f - f}{h} - L(f)\Big)\Big\|_2 \leqslant \Big\|\frac{P_h f - f}{h} - L(f)\Big\|_2 \to 0$$

καθώς $h \to 0^+$.

(γ) Ο $L : D(L) \to L^2(\mu)$ είναι γραμμικός τελεστής.

Πρόταση 2.2.12. *Αν $f \in D(L)$ τότε $\frac{d}{dt}(P_t f) = P_t(L(f)) = LP_t(f)$.*

Απόδειξη. Έστω $f \in D(L)$ και $L(f) = \lim\limits_{t\to 0^+} \frac{P_t f - f}{t}$. Τότε,

$$\frac{d}{dt}(P_t f) = \lim_{h\to 0} \frac{P_{t+h} f - P_t f}{h} = \lim_{h\to 0} P_t\Big(\frac{P_h f - f}{h}\Big).$$

Όμως,

$$\Big\|P_t\Big(\frac{P_h f - f}{h}\Big) - P_t(L(f))\Big\|_2 \leqslant \Big\|\frac{P_h f - f}{h} - L(f)\Big\|_2 \to 0$$

καθώς $h \to 0$. Οπότε,

$$\frac{d}{dt}(P_t f) = \lim_{h\to 0} P_t\Big(\frac{P_h f - f}{h}\Big) = P_t(L(f)).$$

Η άλλη ισότητα έπεται από την Παρατήρηση 2.2.11 (β). □

Παρατήρηση 2.2.13. (α) Η σχέση μεταξύ του γεννήτορα L και της ημιομάδας $(P_t)_{t\geqslant 0}$ εκφράζεται από την $P_t = e^{tL}$. Αυτή η έκφραση γίνεται ακριβής όταν $D(L) = L^2(\mu)$, και μπορούμε να γράψουμε

$$e^{tL} = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k = P_t, \quad t \geqslant 0.$$

(β) Η ημιομάδα Markov $(P_t)_{t\geqslant 0}$ που έχει γεννήτορα τον L με πεδίο το $D(L)$ είναι μοναδική.

Ορισμός 2.2.14. Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t\geqslant 0}$ μια ημιομάδα Markov. Η $(P_t)_{t\geqslant 0}$ λέγεται *συμμετρική* ή *αντιστρέψιμη* ως προς μ αν για κάθε $f, g \in L^2(\mu)$ και κάθε $t \geqslant 0$ ισχύει ότι

$$\int_X f \cdot P_t g \, d\mu = \int_X P_t f \cdot g \, d\mu.$$

Συμβολισμός. Αν $f, g \in L^2(\mu)$ τότε θέτουμε

$$\langle f, g \rangle_\mu = \int_X f \cdot g \, d\mu.$$

Οπότε, με αυτό τον συμβολισμό, η $(P_t)_{t\geqslant 0}$ είναι συμμετρική αν και μόνο αν

$$\langle f, P_t g \rangle_\mu = \langle P_t f, g \rangle_\mu$$

για κάθε $f, g \in L^2(\mu)$. Ισοδύναμα, αν ο P_t είναι αυτοσυζυγής και, επίσης ισοδύναμα, αν ο γεννήτορας L είναι αυτοσυζυγής, αφού $P_t = e^{tL}$ στην περίπτωση που $D(L) = L^2(\mu)$.

Ορισμός 2.2.15. Έστω $(P_t)_{t\geqslant 0}$ μια ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Τότε, η *ενέργεια* ή *μορφή Dirichlet* της ημιομάδας $(P_t)_{t\geqslant 0}$ ορίζεται από την

$$\mathcal{E}(f, g) = -\langle f, Lg \rangle_\mu, \qquad f, g \in L^2(\mu).$$

Συμβολισμός. Έστω (X, d) μετρικός χώρος που είναι τοπικά συμπαγής. Συμβολίζουμε με $C_0(X)$ τον χώρο όλων των συνεχών συναρτήσεων $f : X \to \mathbb{R}$ με την εξής ιδιότητα: υπάρχει $x_0 \in X$ ώστε για κάθε $\varepsilon > 0$ να υπάρχει $M_\varepsilon > 0$ τέτοιο ώστε αν $x \in X$ και $d(x, x_0) > M_\varepsilon$ τότε $|f(x)| < \varepsilon$.

Ορισμός 2.2.16. Έστω (X, d) μετρικός χώρος που είναι τοπικά συμπαγής και $L : D_0(L) \to C_0(X)$ γραμμικός τελεστής, όπου $D_0(L) \subseteq C_0(X)$. Θα λέμε ότι ο L ικανοποιεί την *αρχή του θετικού μεγίστου* αν για κάθε $f \in D_0(L)$ που έχει μέγιστο στο $x_0 \in X$ με $f(x_0) \geqslant 0$ ισχύει ότι $Lf(x_0) \leqslant 0$.

Ορισμός 2.2.17. Έστω (X, d) τοπικά συμπαγής μετρικός χώρος και $L : D_0(L) \to C_0(X)$ γραμμικός τελεστής, όπου $D_0(L) \subseteq C_0(X)$. Τότε, ο L θα λέγεται *γεννήτορας Markov* αν ικανοποιεί τα εξής:

(i) $\overline{D_0(L)} = C_0(X)$, δηλαδή το πεδίο του L είναι πυκνό στον $C_0(X)$.

(ii) Ο L ικανοποιεί την αρχή θετικού μεγίστου.

(iii) Υπάρχει $\lambda_0 > 0$ ώστε ο $\lambda_0 I - L$ να είναι επί του $C_0(X)$.

Θεώρημα 2.2.18. *Έστω (X,d) τοπικά συμπαγής μετρικός χώρος και $L : D_0(L) \to C0(X)$ γραμμικός τελεστής, όπου $D_0(L) \subseteq C_0(X)$. Τότε, τα παρακάτω είναι ισοδύναμα:*

(i) *Ο L είναι γεννήτορας Markov.*

(ii) *Η $(P_t)_{t\geqslant 0} = (e^{tL})_{t\geqslant 0} = \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} L^k\right)_{t\geqslant 0}$ είναι ημιομάδα Markov στον $C_0(X)$ με γεννήτορα τον L και πεδίο του L τον $D_0(L)$.*

Παρατήρηση 2.2.19. (α) Η απόδειξη του θεωρήματος έπεται από την θεωρία Hille-Yosida και παραλείπεται.

(β) Λέγοντας ότι η $(P_t)_{t\geqslant 0}$ είναι ημιομάδα Markov στον $C_0(X)$ εννοούμε ότι κάθε $P_t : C_0(X) \to C_0(X)$ είναι γραμμικός τελεστής, και ικανοποιούνται τα εξής:

(i) $P_0 = I$.

(ii) $P_s \circ P_t = P_{s+t}$ για κάθε $s,t \geqslant 0$.

(iii) Αν $f \in C_0(X)$ και $f \geqslant 0$ τότε $P_t f \geqslant 0$ για κάθε $t \geqslant 0$.

iv) $P_t\mathbf{1} = \mathbf{1}$ για κάθε $t \geqslant 0$.

(γ) Λέγοντας ότι ο $D_0(L)$ είναι το πεδίο του γεννήτορα L εννοούμε ότι ο $D_0(L)$ αποτελείται από όλες τις $f \in C_0(X)$ για τις οποίες υπάρχει το $\lim\limits_{t\to 0^+} \frac{P_t f - f}{t}$ στον $C_0(X)$.

2.3 Ανισότητα Poincaré και λογαριθμική ανισότητα Sobolev

Ορισμός 2.3.1. Έστω $(X,\mathcal{A},\mu)$ χώρος πιθανότητας και $(P_t)_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και πεδίο του L το $D(L) = L^2(\mu)$.

(α) Λέμε ότι το μ ικανοποιεί μια *ανισότητα Poincaré με σταθερά* $\beta > 0$ ως προς $(P_t)_{t\geqslant 0}$ αν για κάθε $f \in L^2(\mu)$ ισχύει ότι
$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{\beta}\mathcal{E}(f,f).$$

(β) Λέμε ότι το μ ικανοποιεί μια *λογαριθμική ανισότητα Sobolev με σταθερά* $\beta > 0$ ως προς $(P_t)_{t\geqslant 0}$ αν για κάθε $f \in L^2(\mu)$ ισχύει ότι
$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{1}{\beta}\mathcal{E}(f,f).$$

Στη συνέχεια αποδεικνύουμε την ανισότητα Poincaré και τη λογαριθμική ανισότητα Sobolev στον διακριτό κύβο $X := \{-1,1\}^n$.

Ορισμός 2.3.2. Η $d_n : \{-1,1\}^n \times \{-1,1\}^n \to [0,\infty)$ με
$$d_n(x,y) = \frac{1}{2}\|x-y\|_1 = \frac{1}{2}\sum_{i=1}^{n} |x_i - y_i|$$
λέγεται *μετρική Hamming* και είναι μετρική επί του διακριτού κύβου $X = \{-1,1\}^n$.

Παρατήρηση 2.3.3. Ο $X = \{-1, 1\}^n$, εφοδιασμένος με τη μετρική Hamming και το μέτρο πιθανότητας γινόμενο μ_p^n, $0 < p < 1$, που ορίσαμε στην εισαγωγή, γίνεται μετρικός χώρος πιθανότητας. Μάλιστα, είναι συμπαγής ως πεπερασμένο σύνολο, και ειδικότερα είναι τοπικά συμπαγής.

Πρόταση 2.3.4. *Κάθε $f : X \to \mathbb{R}$ είναι συνεχής και φραγμένη. Ειδικότερα, $C(X) = C_0(X) = L^q(\mu_p^n) = \mathcal{L}_{b,\mu}(X)$ για κάθε $1 \leqslant q \leqslant \infty$.*

Απόδειξη. Προφανές, αφού ο (X, d_n) είναι πεπερασμένος μετρικός χώρος. □

Πρόταση 2.3.5. *Έστω $(f_t)_{t\geqslant 0}, f, g : X \to \mathbb{R}$. Αν $\lim\limits_{t\to 0^+} \frac{f_t(x)-f(x)}{t} = g(x)$ για κάθε $x \in X$ τότε*

$$\lim_{t\to 0^+} \left\| \frac{f_t - f}{t} - g \right\|_2 = 0.$$

Απόδειξη. Έστω $X = \{x_1, \ldots, x_{2^n}\}$ μια αρίθμηση του διακριτού κύβου. Τότε έχουμε ότι

$$\mu_p^n(\{x_i\}) \leqslant (\max\{p, 1-p\})^n = c$$

για κάθε $i = 1, \ldots, 2^n$. Τώρα λοιπόν έχουμε ότι, για $t > 0$,

$$\begin{aligned}\left\| \frac{f_t - f}{t} - g \right\|_2^2 &= \int_X \left| \frac{f_t(x) - f(x)}{t} - g(x) \right|^2 d\mu_p^n(x) = \sum_{i=1}^{2^n} \left| \frac{f_t(x_i) - f(x_i)}{t} - g(x_i) \right|^2 \mu_p^n(\{x_i\}) \\ &\leqslant c \sum_{i=1}^{2^n} \left| \frac{f_t(x_i) - f(x_i)}{t} - g(x_i) \right|^2 \longrightarrow c \cdot 0 = 0\end{aligned}$$

όταν $t \to 0^+$, και έχουμε το ζητούμενο. □

Παρατήρηση 2.3.6. Από τα παραπάνω συμπεραίνουμε ότι αν $L : C_0(X) \to C_0(X)$ είναι γεννήτορας Markov, τότε την ημιομάδα Markov $(P_t)_{t\geqslant 0}$ που παράγει μπορούμε να την δούμε ως $P_t : L^2(\mu_p^n) \to L^2(\mu_p^n)$ που έχει γεννήτορα τον L και πεδίο του γεννήτορα $D(L)$ την κλάση όλων των $f \in L^2(\mu_p^n)$ για τις οποίες υπάρχει το $\lim\limits_{t\to 0^+} \frac{P_t f - f}{t}$ στον $L^2(\mu_p^n)$. Όμως, $C_0(X) = D_0(L)$, όπου $D_0(L)$ είναι η κλάση όλων των $f \in C_0(X)$ για τις οποίες $\lim\limits_{t\to 0^+} \frac{P_t f - f}{t} = g(x)$ για κάθε $x \in X$, και συνεπώς $D_0(L) = D(L)$, δηλαδή $D(L) = L^2(\mu_p^n)$.

Ορίζουμε τώρα $L_i : C_0(X) \to C_0(X)$ με

$$L_i(f) = \int_{\{-1,1\}} f_i(x_i)\, d\mu_p(x_i) - f$$

και $L = \sum_{i=1}^n L_i$, δηλαδή

$$L(f) = \sum_{i=1}^n \int_{\{-1,1\}} f_i(x_i)\, d\mu_p(x_i) - nf.$$

Πρόταση 2.3.7. *Ο $L : C_0(X) \to C_0(X)$ που ορίσαμε προηγουμένως είναι γεννήτορας Markov και συνεπώς η $(P_t)_{t\geqslant 0}$ με $P_t = e^{tL} = \sum_{k=0}^\infty \frac{t^k}{k!} L^k$ είναι ημιομάδα Markov.*

Απόδειξη. (i) Το πεδίο $D_0(L)$ του L είναι προφανώς πυκνό στον $C_0(X)$, αφού $D_0(L) = C_0(X)$.

(ii) Έστω $f \in C_0(X)$ και $x_0 \in X$ με $\max(f) = f(x_0) \geqslant 0$. Έστω ότι $x_0 = (y_1, \dots, y_n)$. Τότε,

$$\begin{aligned} Lf(x_0) &= \sum_{i=1}^{n} \int_{\{-1,1\}} f(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n)\, d\mu_p(x_i) - nf(x_0) \\ &\leqslant \sum_{i=1}^{n} \int_{\{-1,1\}} f(x_0)\, d\mu_p(x_i) - nf(x_0) = \sum_{i=1}^{n} f(x_0) - nf(x_0) = 0. \end{aligned}$$

(iii) Θέλουμε να δείξουμε ότι υπάρχει $\lambda > 0$ ώστε ο $\lambda I - L$ να είναι επί του $C_0(X)$. Ισοδύναμα, θέλουμε να δείξουμε ότι υπάρχει $\lambda > 0$ ώστε για κάθε $g : X \to \mathbb{R}$ υπάρχει $f : X \to \mathbb{R}$ ώστε

$$\begin{aligned} (\lambda - n) f(y_1, \dots, y_n) + \sum_{i=1}^{n} (q \cdot f(y_1, \dots, y_{i-1}, -1, y_{i+1}, \dots, y_n) \\ + p \cdot f(y_1, \dots, y_{i-1}, 1, y_{i+1}, \dots, y_n)) = g(y_1, \dots, y_n) \end{aligned}$$

για κάθε $(y_1, \dots, y_n) \in X$, όπου $q = 1 - p$. Η τελευταία σχέση είναι ισοδύναμη με ένα (εν γένει) μη ομογενές γραμμικό σύστημα 2^n εξισώσεων και 2^n αγνώστων, με πίνακα συντελεστών ισοδύναμο με τον

$$\begin{pmatrix} \lambda + a_{11} & a_{12} & \dots & a_{1,2^n} \\ a_{21} & \lambda + a_{22} & \dots & a_{2,2^n} \\ \vdots & \vdots & & \vdots \\ a_{2^n,1} & a_{2^n,2} & \dots & \lambda + a_{2^n,2^n} \end{pmatrix}$$

για κάποιους $a_{ij} \in \mathbb{R}$, $1 \leqslant i, j \leqslant 2^n$. Η ορίζουσα του τελευταίου πίνακα είναι πολυώνυμο βαθμού 2^n ως προς λ, και συνεπώς για $\lambda \to +\infty$ η ορίζουσα του πίνακα τείνει στο $+\infty$. Οπότε, υπάρχει $\lambda_0 > 0$ ώστε η ορίζουσα του πίνακα να είναι μη μηδενική και συνεπώς το σύστημα έχει πάντα λύση. $\square$

Πρόταση 2.3.8. *Ο $L : L^2(\mu_p^n) \to L^2(\mu_p^n)$ που ορίσαμε προηγουμένως είναι αυτοσυζυγής.*

Απόδειξη. Έστω $f, g \in L^2(\mu_p^n)$. Θα δείξουμε ότι $\langle f, Lg \rangle_{\mu_p^n} = \langle Lf, g \rangle_{\mu_p^n}$, δηλαδή

$$\int_X f \Big(\sum_{i=1}^{n} \int_{\{-1,1\}} g_i \, d\mu_p - ng \Big) d\mu_p^n = \int_X g \Big(\sum_{i=1}^{n} \int_{\{-1,1\}} f_i \, d\mu_p - nf \Big) d\mu_p^n,$$

ή, ισοδύναμα,

$$\int_X f \sum_{i=1}^{n} \int_{\{-1,1\}} g_i \, d\mu_p \, d\mu_p^n = \int_X g \sum_{i=1}^{n} \int_{\{-1,1\}} f_i \, d\mu_p \, d\mu_p^n.$$

Τώρα,

$$
\begin{aligned}
&\int_X f\Big(\sum_{i=1}^n \int_{\{-1,1\}} g_i\Big)\, d\mu_p\, d\mu_p^n \\
&= \int_X f(y_1,\dots,y_n)\sum_{i=1}^n (q\cdot g(y_1,\dots,-1,\dots,y_n)+p\cdot g(y_1,\dots,1,\dots,y_n))\, d\mu_p^n(y_1,\dots,\widehat{y}_i,\dots,y_n) \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}}\int_{\{-1,1\}} f(y_1,\dots,y_n)(q\cdot g(y_1,\dots,-1,\dots,y_n)+p\cdot g(y_1,\dots,1,\dots,y_n)) \\
&\qquad\qquad d\mu_p(y_i)\, d\mu_p^{n-1}(y_1,\dots,\widehat{y}_i,\dots,y_n) \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} (q\cdot f(y_1,\dots,-1,\dots,y_n)+p\cdot f(y_1,\dots,1,\dots,y_n)) \\
&\qquad\quad \times (q\cdot g(y_1,\dots,-1,\dots,y_n)+p\cdot g(y_1,\dots,1,\dots,y_n))\, d\mu_p^{n-1}(y_1,\dots,\widehat{y}_i,\dots,y_n) \\
&= \int_X g\Big(\sum_{i=1}^n \int_{\{-1,1\}} f_i\Big)\, d\mu_p\, d\mu_p^n,
\end{aligned}
$$

κάνοντας την ίδια διαδικασία αφού καταλήξαμε σε συμμετρική μορφή ως προς f και g. □

Πόρισμα 2.3.9. *Το μ_p^n είναι αναλλοίωτο στην $(P_t)_{t\geqslant 0}$ που έχει γεννήτορα τον L με πεδίο τον $L^2(\mu_p^n)$, και επίσης η $(P_t)_{t\geqslant 0}$ είναι συμμετρική.*

Απόδειξη. Το γεγονός ότι η $(P_t)_{t\geqslant 0}$ είναι συμμετρική έπεται από το ότι έχει αυτοσυζυγή γεννήτορα. Για το ότι το μ_p^n είναι αναλλοίωτο στην $(P_t)_{t\geqslant 0}$ έχουμε ότι για κάθε $f\in L^2(\mu_p^n)$,

$$\langle P_t f, \mathbf{1}\rangle_{\mu_p^n} = \langle f, P_t\mathbf{1}\rangle_{\mu_p^n}$$

άρα

$$\int_X P_t f\cdot \mathbf{1}\, d\mu_p^n = \int_X f\cdot P_t\mathbf{1}\, d\mu_p^n,$$

ή, ισοδύναμα,

$$\int_X P_t f\, d\mu_p^n = \int_X f\, d\mu_p^n.$$

□

Λήμμα 2.3.10. *Για κάθε $f\in L^2(\mu_p^n)$ ισχύει ότι*

$$\mathcal{E}(f,f) = \sum_{i=1}^n \mathrm{Var}_{\mu_p}(f_i)\, d\mu_p^n.$$

Απόδειξη. Έχουμε

$$
\begin{aligned}
\mathcal{E}(f,f) &= -\langle f, Lf\rangle_{\mu_p^n} = -\int_X f\Big(\sum_{i=1}^n \int_{\{-1,1\}} f_i\, d\mu_p - nf\Big)\, d\mu_p^n \\
&= n\int_X f^2\, d\mu_p^n - \sum_{i=1}^n \int_X f \int_{\{-1,1\}} f_i\, d\mu_p\, d\mu_p^n \\
&= \sum_{i=1}^n \int_X \Big(f^2 - f\int_{\{-1,1\}} f_i\, d\mu_p\Big)\, d\mu_p^n \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} \Big[\int_{\{-1,1\}} f^2(y_1,\dots,y_n) - f(y_1,\dots,y_n) \\
&\qquad \times \int_{\{-1,1\}} f(y_1,\dots,x_i,\dots,y_n)\, d\mu_p(x_i)\Big]\, d\mu_p(y_i)\, d\mu_p^{n-1}(y_1,\dots,\widehat{y}_i,\dots,y_n) \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} \int_{\{-1,1\}} f_i^2\, d\mu_p - \Big(\int_{\{-1,1\}} f_i\, d\mu_p\Big)^2\, d\mu_p^{n-1} \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} \mathrm{Var}_{\mu_p}(f_i)\, d\mu_p^{n-1} = \sum_{i=1}^n \int_X \mathrm{Var}_{\mu_p}(f_i)\, d\mu_p^n.
\end{aligned}
$$

□

Πρόταση 2.3.11 (ανισότητα Poincaré στον διακριτό κύβο)**.** *Το μ_p^n ικανοποιεί ανισότητα Poincaré με σταθερά $\beta = 1$ ως προς την ημιομάδα Markov $(P_t)_{t\geqslant 0} = (e^{-Lt})_{t\geqslant 0}$ που έχουμε ορίσει.*

Απόδειξη. Από την ανισότητα Efron-Stein και από το Λήμμα 2.3.10 έχουμε ότι, για κάθε $f \in L^2(\mu_p^n)$,

$$\mathrm{Var}_{\mu_p^n}(f) \leqslant \sum_{i=1}^n \mathrm{Var}_{\mu_p}(f_i)\, d\mu_p^n = \mathcal{E}(f,f),$$

και έπεται το ζητούμενο. □

Θεώρημα 2.3.12 (λογαριθμική ανισότητα Sobolev στον χώρο δύο σημείων)**.** *Ο $(Y, \mathbb{P}(Y), \mu_p)$, όπου $Y = \{-1,1\}$ και μ_p το μέτρο Bernoulli, ικανοποιεί λογαριθμική ανισότητα Sobolev ως προς την $(P_t)_{t\geqslant 0}$ με σταθερά $r = \frac{p-q}{\log p - \log q}$ αν $p \neq \frac{1}{2}$ και $r = \frac{1}{2}$ αν $p = \frac{1}{2}$.*

Απόδειξη. Αρχικά, από την Παρατήρηση 2.1.7 και την Πρόταση 2.3.4 έχουμε ότι αν $f : Y \to \mathbb{R}$ τότε

$$
\begin{aligned}
\mathrm{Ent}_{\mu_p}(f^2) &= \sup\Big\{\int_Y f^2 g\, d\mu_p \mid g : Y \to \mathbb{R}, \int_Y e^g\, d\mu_p = 1\Big\} \\
&= \sup\Big\{\langle f^2, g\rangle_{\mu_p} \mid g : Y \to \mathbb{R}, \int_Y e^g\, d\mu_p = 1\Big\}.
\end{aligned}
$$

Για $g \neq 0$ με $\int_Y e^g\, d\mu_p = 1$ ορίζουμε

$$a(g) = \inf\Big\{\frac{\mathcal{E}(f,f)}{\langle f^2, g\rangle_{\mu_p}} : \mathrm{Ent}_{\mu_p}(f^2) > 0, \langle f^2, g\rangle_{\mu_p} > 0\Big\}$$

που είναι καλά ορισμένο και ικανοποιεί την $a(g) \geqslant 0$. Ορίζουμε επίσης

$$\kappa = \inf\Big\{a(g) : g \neq 0, \int_Y e^g \, d\mu_p = 1\Big\} > 0.$$

Θα δείξουμε ότι

$$\kappa \operatorname{Ent}_{\mu_p}(f^2) \leqslant \mathcal{E}(f,f) \tag{2.3.1}$$

για κάθε $f : Y \to \mathbb{R}$. Αρχικά από το Λήμμα 2.3.10, για $n = 1$, έχουμε $\mathcal{E}(f,f) = \operatorname{Var}_{\mu_p}(f) \geqslant 0$ για κάθε $f \in L^2(\mu_p)$. Οπότε αν $f \equiv 0$ ή γενικότερα $\operatorname{Ent}_{\mu_p}(f^2) = 0$ τότε η (2.3.1) ισχύει τετριμμένα.

Έστω τώρα ότι $\operatorname{Ent}_{\mu_p}(f^2) > 0$. Θεωρούμε ακολουθία $(g_n)_{n\in\mathbb{N}}$ με $\int_Y e^{g_n} \, d\mu_p = 1$ για κάθε $n \in \mathbb{N}$ και $\langle f^2, g_n\rangle_{\mu_p} \to \operatorname{Ent}_{\mu_p}(f^2) > 0$. Υποθέτουμε μάλιστα χωρίς περιορισμό της γενικότητας ότι $\langle f^2, g_n\rangle_{\mu_p} > 0$ για κάθε $n \in \mathbb{N}$, οπότε $g_n \neq 0$ για κάθε $n \in \mathbb{N}$. Συνεπώς,

$$\mathcal{E}(f,f) \geqslant \langle f^2, g_n\rangle_{\mu_p} a(g_n) \geqslant \kappa \langle f^2, g_n\rangle_{\mu_p}$$

για κάθε $n \in \mathbb{N}$, και όταν $n \to \infty$ έχουμε

$$\kappa \operatorname{Ent}_{\mu_p}(f^2) \leqslant \mathcal{E}(f,f),$$

και αυτό για κάθε $f \in L^2(\mu_p)$.

Θα υπολογίσουμε τώρα το $a(g)$ για τυχούσα $g \in L^2(\mu_p)$, $g \neq 0$, με $\int_Y e^g \, d\mu_p = 1$. Έχουμε $pe^{g(1)} + qe^{g(-1)} = 1$ και $g(1)g(-1) < 0$ αφού $p + q = 1$. Θέτουμε $a = g(1)$ και $b = g(-1)$. Αν $f \in L^2(\mu_p)$ τότε

$$\begin{aligned}\operatorname{Var}_{\mu_p}(f) &= \int_Y f^2 \, d\mu_p - \Big(\int_Y f \, d\mu_p\Big)^2 = p \cdot f^2(1) - q \cdot f^2(-1) - (p \cdot f(1) + q \cdot f(-1))^2 \\ &= pq(f(1) - f(-1))^2.\end{aligned}$$

Οπότε θα υπολογίσουμε το

$$A = \inf\Big\{pq \frac{(f(1) - f(-1))^2}{pf^2(1)a + qf^2(-1)b} : f \in L^2(\mu_p), pf^2(1)a + qf^2(-1)b > 0\Big\},$$

όπου $A \geqslant 0$. Έστω λοιπόν $f \in L^2(\mu_p)$ με $pf^2(1)a + qf^2(-1)b \neq 0$.

(α) Αν $f(-1) = 0$, τότε

$$pq \frac{(f(1) - f(-1))^2}{pf^2(1)a + qf^2(-1)b} = \frac{q}{a}.$$

(β) Αν $f(1) = 0$, τότε

$$pq \frac{(f(1) - f(-1))^2}{pf^2(1)a + qf^2(-1)b} = \frac{p}{b}.$$

(γ) Αν $f(-1)f(1) \neq 0$, τότε

$$pq \frac{(f(1) - f(-1))^2}{pf^2(1)a + qf^2(-1)a} = \frac{pq\left(\frac{f(1)}{f(-1)} - 1\right)^2}{p\left(\frac{f(1)}{f(-1)}\right)^2 a + b}.$$

Παρατηρούμε ότι αν $\tilde{f} : L^2(\mu_p) \to \mathbb{R}$ είναι η συνάρτηση με $\tilde{f}(-1) = 1$ και

$$\tilde{f}(1) = \begin{cases} \frac{f(1)}{f(-1)} & , \text{ αν } f(1)f(-1) > 0 \\ -\frac{f(1)}{f(-1)} & , \text{ αν } f(1)f(-1) < 0, \end{cases}$$

έχουμε ότι

$$\frac{pq(\tilde{f}(1) - \tilde{f}(-1))^2}{p\tilde{f}^2(1)a + q\tilde{f}^2(-1)b} = \frac{pq\left(\frac{f(1)}{f(-1)} - 1\right)^2}{p\left(\frac{f(1)}{f(-1)}\right)^2 a + qb} \leqslant \frac{pq\left(\frac{f(1)}{f(-1)} - 1\right)^2}{p\left(\frac{f(1)}{f(-1)}\right)^2 a + qb}.$$

Συνεπώς, αφού η $\tilde{f}$ μας δίνει μικρότερη τιμή στην ζητούμενη ποσότητα, μπορούμε να υποθέσουμε χωρίς περιορισμό της γενικότητας ότι $f > 0$, $f(1) = x$ για κάποιο $x > 0$ και $f(-1) = 1$. Άρα, θα υπολογίσουμε το

$$\inf_{x>0} \left\{ pq\frac{(x-1)^2}{px^2 a + qb} : px^2 a + qb > 0 \right\}.$$

Διακρίνοντας τις περιπτώσεις $a > 0$ και $a < 0$ καταλήγουμε στο ότι το inf επιτυγχάνεται για $x = \frac{qb}{pa}$ και παίρνει την τιμή $\frac{p}{b} + \frac{q}{a}$. Συνεπώς, για να υπολογίσουμε το $a(g)$ πρέπει να συγκρίνουμε τα $\frac{p}{b} + \frac{q}{a}, \frac{p}{b}, \frac{q}{a}$.

Αν, για παράδειγμα, $b < 0$ τότε το $\frac{p}{b} < 0$ απορρίπτεται αφού $A \geqslant 0$. Όμως, $1 = pe^a + qe^b > e^{ap+bq}$, αφού η e^x είναι γνησίως κυρτή και $ab < 0$ άρα $a \neq b$. Συνεπώς, $ap + bq < 0$, δηλαδή $\frac{p}{b} + \frac{q}{a} > 0$, και μάλιστα $\frac{p}{b} + \frac{q}{a} < \frac{q}{a}$. Όμοια αν $a < 0$.

Καταλήγουμε λοιπόν στο ότι $a(g) \geqslant \frac{p}{b} + \frac{q}{a} = A$. Τότε,

$$\begin{aligned} \kappa &= \inf\left\{ a(g) : g \neq 0, \int_Y e^g \, d\mu_p = 1 \right\} \geqslant \inf\left\{ \frac{p}{b} + \frac{q}{a} : pe^a + qe^b = 1 \right\} \\ &= \inf\left\{ \frac{p}{\log s} + \frac{q}{\log t} : pt + qs = 1 \right\}, \end{aligned}$$

θέτοντας $t = e^a$ και $s = e^b$.

Ορίζουμε λοιπόν

$$h(t) = \frac{p}{\log s} + \frac{q}{\log t},$$

όπου $pt + qs = 1$ για $t \in (0,1) \cup (1, 1/p)$, και τότε

$$\kappa \geqslant \inf\{h(t) : t \in (0,1) \cup (1, 1/p)\}.$$

Υπολογίζοντας τα όρια, η h επεκτείνεται σε συνεχή συνάρτηση $\tilde{h} : [0, 1/p] \to \mathbb{R}$ με

$$\tilde{h}(0) = -\frac{p}{\log q}, \quad \tilde{h}(1) = \frac{1}{2}, \quad \tilde{h}(1/p) = -\frac{q}{\log p}.$$

Η $\tilde{h}$ έχει συνεπώς ολικό ελάχιστο στο $[0, 1/p]$. Έχουμε ότι

$$\tilde{h}'(t) = \frac{p^2}{qs(\log s)^2} - \frac{q}{t(\log t)^2}, \quad t \in (0,1) \cup (1, 1/p).$$

Παρατηρούμε ότι $\lim_{t \to 0^+} \tilde{h}'(t) = -\infty$ και $\lim_{t \to 0^-} \tilde{h}'(t) = +\infty$. Οπότε η $\tilde{h}$ δεν παρουσιάζει ολικό ελάχιστο στα άκρα της. Συνεπώς, θα το παρουσιάζει είτε στο 1 είτε σε σημείο στο οποίο μηδενίζεται η παράγωγός της. Παρατηρούμε ακόμη ότι ορίζεται η $h'(1) = \frac{1-2p}{12(-1+p)}$.

Τώρα, για $t \neq 1$, η επίλυση της $h'(t) = 0$, $pt + qs = 1$ είναι ισοδύναμη με τις

$$p\sqrt{t}\log t = -q\sqrt{s}\log s \quad \text{και} \quad pt + qs = 1,$$

λαμβάνοντας υπ' όψιν ότι $\log s \cdot \log t = ab < 0$.

Τώρα, $p = \frac{1-s}{t-s}$ και $q = \frac{1-t}{s-t}$, αφού $t \neq s$ λόγω της $t \neq 1$. Συνεπώς, αναγόμαστε στην επίλυση της

$$\frac{\sqrt{t}\log t}{1-t} = \frac{\sqrt{s}\log s}{1-s}.$$

Η συνάρτηση $\varphi : (0,1) \cup (1,\infty) \to \mathbb{R}$, $\varphi(t) = \frac{\sqrt{t}\log t}{1-t}$ ικανοποιεί τις

$$\lim_{t\to 0}\varphi(t) = \lim_{t\to+\infty}\varphi(t) = 0, \quad \lim_{t\to 1}\varphi(t) = -1, \quad \varphi(t) = \varphi(1/t)$$

για κάθε $t \in (0,1)\cup(1,\infty)$. Επίσης, η φ είναι αύξουσα στο $(1,\infty)$ και φθίνουσα στο $(0,1)$. Οπότε, αν $\varphi(t) = \varphi(s)$ τότε είτε $t = s = 1$, που οδηγεί σε άτοπο, είτε $t = \frac{1}{s} = \frac{q}{p}$.

(α) Αν $p = \frac{1}{2}$ τότε βλέπουμε ότι η μοναδική ρίζα της $\tilde{h}'$ είναι στο 1 και συνεπώς

$$\min \tilde{h} = \tilde{h}(1) = \frac{1}{2}.$$

(β) Αν $p \neq \frac{1}{2}$ τότε η μοναδική ρίζα της $\tilde{h}'$ είναι στο $\frac{q}{p}$ και

$$\min \tilde{h} = \tilde{h}\Big(\frac{q}{p}\Big) = \frac{q-p}{\log q - \log p}.$$

Οπότε, $\inf h(t) = \frac{p-q}{\log p - \log q}$ αν $p \neq \frac{1}{2}$ και $\inf h(t) = \frac{1}{2}$ αν $p = \frac{1}{2}$.

Τελικά έχουμε το εξής: αφού $0 < r \leqslant \kappa$,

$$r\mathrm{Ent}_{\mu_p}(f^2) \leqslant \kappa\mathrm{Ent}_{\mu_p}(f^2) \leqslant \mathcal{E}(f,f),$$

δηλαδή

$$\mathrm{Ent}_{\mu_p}(f^2) \leqslant \frac{1}{r}\, pq\,(f(1) - f(-1))^2,$$

όπου $r = \frac{p-q}{\log p - \log q}$ αν $p \neq \frac{1}{2}$ και $r = \frac{1}{2}$ αν $p = \frac{1}{2}$. □

Πόρισμα 2.3.13 (λογαριθμική ανισότητα Sobolev στον διακριτό κύβο)**.** *Ο $(X, \mathbb{P}(X), \mu_p^n)$, όπου $X = \{-1,1\}^n$ και μ_p^n το μέτρο γινόμενο που αντιστοιχεί στο μ_p, ικανοποιεί λογαριθμική ανισότητα Sobolev με σταθερά $r = \frac{p-q}{\log p - \log q}$ αν $p \neq \frac{1}{2}$ και $r = \frac{1}{2}$ αν $p = \frac{1}{2}$.*

Απόδειξη. Από την Πρόταση 2.1.11 έχουμε

$$\mathrm{Ent}_{\mu_p^n}(f^2) \leqslant \sum_{i=1}^n \int \mathrm{Ent}_{\mu_p}(f_i^2)\, d\mu_p^n \leqslant \frac{1}{r}\sum_{i=1}^n \mathrm{Var}_{\mu_p}(f_i^2)\, d\mu_p^n = \frac{1}{r}\mathcal{E}(f,f),$$

εφαρμόζοντας το Θεώρημα 2.3.12 και το Λήμμα 2.3.10. □

2.4 Υπερσυσταλτότητα στον διακριτό κύβο

Λήμμα 2.4.1. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $1 \leqslant p < q \leqslant \infty$. Τότε, $L^q(\mu) \subseteq L^p(\mu)$ και $\|f\|_p \leqslant \|f\|_q$ για κάθε $f : X \to \mathbb{R}$ μετρήσιμη.*

Απόδειξη. Αν $\|f\|_q = +\infty$ τότε δεν έχουμε κάτι να δείξουμε. Έστω τώρα $f \in L^q(\mu)$.

(i) Αν $q = +\infty$ τότε $\int |f|^p \, d\mu \leqslant \int \|f\|_\infty^p \, d\mu = \|f\|_\infty^p$, άρα $f \in L^p(\mu)$ και $\|f\|_p \leqslant \|f\|_\infty$.

(ii) Αν $q < \infty$ τότε από την ανισότητα Hölder

$$\int_X |f|^p \, d\mu \leqslant \Big(\int_X (|f|^p)^{\frac{q}{p}} \, d\mu \Big)^{\frac{p}{q}} \Big(\int_X \mathbf{1}^r \, d\mu \Big)^{\frac{1}{r}} = \Big(\int_X |f|^q \, d\mu \Big)^{\frac{p}{q}} = \|f\|_q^p < \infty,$$

όπου r είναι ο συζυγής εκθέτης του q/p. Άρα, $f \in L^p(\mu)$ και $\|f\|_p \leqslant \|f\|_q$. □

Πόρισμα 2.4.2. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και $(P_t)_{t \geqslant 0}$ ημιομάδα Markov ως προς την οποία το μ είναι αναλλοίωτο. Τότε, για κάθε $1 \leqslant p < q \leqslant \infty$ και για κάθε $f \in L^p(\mu)$ ισχύει ότι $\|P_t f\|_p \leqslant \|f\|_q$.*

Απόδειξη. Έστω $f \in L^p(\mu)$. Τότε, από την Πρόταση 2.2.7 και το παραπάνω λήμμα, $\|P_t f\|_p \leqslant \|f\|_p \leqslant \|f\|_q$. □

Στη συνέχεια θα εξετάσουμε στη συγκεκριμένη περίπτωση του διακριτού κύβου, όπου δηλαδή $(X, \mathcal{A}, \mu) = (\{-1,1\}^n, \mathcal{P}(X), \mu_p^n)$, υπό ποιες προϋποθέσεις ισχύει μια «αντίστροφη» ανισότητα της μορφής $\|P_t f\|_q \leqslant \|f\|_p$, όταν $1 \leqslant p < q \leqslant \infty$ και $t \geqslant 0$, ως προς την ημιομάδα Markov $(P_t)_{t \geqslant 0}$ που είχαμε ορίσει μέσω της $Lf = \sum_{i=1}^n \int_{\{-1,1\}^n} f_i d\mu_p^n - nf$.

Λήμμα 2.4.3. *Έστω $L : L^q(\mu_p^n) \to L^q(\mu_{p_n})$, $Lf = \sum_{i=1}^n \int_{\{-1,1\}^n} f_i d\mu_p^n - nf$, ο γεννήτορας Markov που έχουμε ορίσει και $(P_t)_{t \geqslant 0}$ η ημιομάδα Markov που παράγει. Τότε, για κάθε $f, g \in L^q(\mu_p^n)$ ισχύει ότι*

$$\mathcal{E}(f,g) = \frac{1}{2} \sum_{i=1}^n \int_X \int_{\{-1,1\}} \int_{\{-1,1\}} (f_i(x_i) - f_i(y_i))(g_i(x_i) - g_i(y_i)) \, d\mu_p(x_i) \, d\mu_p(y_i) \, d\mu_p^n(x).$$

Απόδειξη. Αρχικά,

$$\begin{aligned}
\mathcal{E}(f,g) &= \int_X f(-Lg)\,d\mu_p^n = -\int_X f\Big(\sum_{i=1}^n \int_{\{-1,1\}} g_i\,d\mu_p - ng\Big)\,d\mu_p^n \\
&= \int_X \Big(nfg - \sum_{i=1}^n f\int_{\{-1,1\}} g_i\,d\mu_p\Big)\,d\mu_p^n \\
&= \sum_{i=1}^n \int_X \Big(fg - f\int_{\{-1,1\}} g_i\,d\mu_p\Big)\,d\mu_p^n \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} \Big(\int_{\{-1,1\}} f_i(y_i)g_i(y_i) - f_i(y_i)\int_{\{-1,1\}} g_i(x_i)\,d\mu_p(x_i)\Big)\,d\mu_p^{n-1} \\
&= \sum_{i=1}^n \int_{\{-1,1\}^{n-1}} \Big(\int_{\{-1,1\}} f_i(y_i)g_i(y_i)\,d\mu_p(y_i) \\
&\qquad - \int_{\{-1,1\}} f_i(x_i)\,d\mu_p(x_i)\cdot\int_{\{-1,1\}} g_i(x_i)\,d\mu_p(x_i)\Big)\,d\mu_p^{n-1} \\
&= \sum_{i=1}^n \int_X \Big(\int_{\{-1,1\}} f_i(x_i)g_i(x_i)\,d\mu_p(x_i) \\
&\qquad - \int_{\{-1,1\}} f_i(x_i)\,d\mu_p(x_i)\int_{\{-1,1\}} g_i(x_i)\,d\mu_p(x_i)\Big)\,d\mu_p^n
\end{aligned}$$

με το συμβολισμό που έχουμε εισαγάγει.

Από την άλλη πλευρά,

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^n \int_X\int_{\{-1,1\}}\int_{\{-1,1\}} (f_i(x_i)-f_i(y_i))(g_i(x_i)-g_i(y_i))\,d\mu_p(x)\,d\mu_p(y)\,d\mu_p^n(x) \\
&= \frac{1}{2}\Big[\sum_{i=1}^n \int_X \Big(2\int_{\{-1,1\}} f_i(x_i)g_i(x_i)\,d\mu_p(x_i) \\
&\qquad - 2\int_{\{-1,1\}} f_i(y_i)\,d\mu_p(y_i)\int_{\{-1,1\}} g_i(y_i)d\mu_p(y_i)\Big)\,d\mu_p^n\Big],
\end{aligned}$$

και έχουμε αυτό που θέλουμε. □

Λήμμα 2.4.4. *Για κάθε $x > y \geqslant 0$ και για κάθε $q \geqslant 1$ ισχύει ότι*

$$\frac{4(q-1)}{q^2}(x^{q/2}-y^{q/2})^2 \leqslant (x^{q-1}-y^{q-1})(x-y).$$

Απόδειξη. Για $q = 1$ ισχύει τετριμμένα. Έστω ότι $q > 1$. Αν $y = 0$ και $x > 0$ παίρνουμε ότι $\frac{4(q-1)}{q^2}x^q \leqslant x^q$ και ότι η $4(q-1) \leqslant q^2$ είναι ισοδύναμη με την $(q-2)^2 \geqslant 0$ που ισχύει.

Έστω λοιπόν ότι $0 < y < x$. Τότε, χργσιμοποιώντας και την ανισότητα Cauchy-Schwarz, έχουμε

$$\begin{aligned}
\Big[\frac{x^{q/2}-y^{q/2}}{x-y}\Big]^2 &= \Big[\frac{q}{2(x-y)}\int_y^x s^{\frac{q}{2}-1}ds\Big]^2 \leqslant \frac{q^2}{4(x-y)^2}\int_y^x s^{q-2}ds\int_y^x \mathbf{1}\,ds \\
&= \frac{q^2}{4(x-y)}\int_y^x s^{q-2}ds = \frac{q^2}{4(x-y)}\,\frac{x^{q-1}-y^{q-1}}{q-1}.
\end{aligned}$$

Δηλαδή έχουμε δείξει ότι

$$\Big[\frac{x^{q/2}-y^{q/2}}{x-y}\Big]^2 \leqslant \frac{q^2}{4(x-y)}\,\frac{x^{q-1}-y^{q-1}}{q-1},$$

που είναι ισοδύναμη με την

$$\frac{4(q-1)}{q^2}(x^{q/2}-y^{q/2})^2 \leqslant (x^{q-1}-y^{q-1})(x-y).$$

$\square$

Λήμμα 2.4.5. *Έστω $f : X \to \mathbb{R}$ συνάρτηση. Έστω επίσης $(P_t)_{t\geqslant 0}$ ημιομάδα Markov με το μ_p^n αναλλοίωτο ως προς αυτήν, με γεννήτορα L και $D(L) = L^2(\mu)$. Τότε, για κάθε $x \in X$ η $\psi_x : [0,\infty) \to \mathbb{R}$, $\psi_x(t) = P_t f(x)$ είναι συνεχής.*

Απόδειξη. Έστω $f : X \to \mathbb{R}$ συνάρτηση. Είναι φανερό ότι $f \in L^2(\mu_p^n)$. Αφού $D(L) = L^2(\mu_p^n)$ έχουμε $\lim\limits_{h\to 0^+} \frac{P_hf-f}{f} = Lf$ στον $L^2(\mu_p^n)$, άρα $\left\|\frac{P_hf-f}{h}\right\|_2 \to \|Lf\|_2$ καθώς $h \to 0^+$, και ειδικότερα,

$$\lim_{h\to 0^+}\|P_hf-f\|_2 = 0.$$

Έστω $\varepsilon > 0$. Τότε υπάρχει $\delta > 0$ ώστε αν $0 \leqslant h < \delta$ τότε $\|P_hf-f\|_2 < \varepsilon$. Έστω τώρα $t \geqslant 0$ και $s \in (t-\delta, t+\delta) \cap [0,+\infty)$. Αν $s \in (t, t+\delta)$ τότε

$$\|P_tf-P_sf\|_2 = \|P_t(P_{s-t}(f)-f)\|_2 \leqslant \|P_{s-t}(f)-f\|_2 < \varepsilon,$$

αφού $0 \leqslant s-t < \delta$. Όμοια ελέγχεται και η άλλη περίπτωση και καταλήγουμε στο ότι $\|P_tf-P_sf\|_2 < \varepsilon$ για κάθε $s \in (t-\delta, t+\delta) \cap [0,\infty)$.

Έστω τώρα $\{x_1,\ldots,x_{2^n}\}$ μια αρίθμηση του X και $x_j \in X$. Από τα παραπάνω, υπάρχει $\delta_j > 0$ ώστε $\|P_tf-P_sf\|_2 < \varepsilon\sqrt{\mu_p^n(\{x_j\})}$ για κάθε $s \in (t-\delta_j, t+\delta_j) \cap [0,\infty)$, αφού $\mu_p^n(\{x_j\}) > 0$. Όμως,

$$\|P_tf-P_sf\|_2 = \Big(\sum_{i=1}^{2^n}|P_tf(x_j)-P_sf(x_j)|^2\mu_p^n(\{x_j\})\Big)^{1/2}.$$

Άρα, για κάθε $i = 1,\ldots,2^n$,

$$\begin{aligned}|P_tf(x_i)-P_sf(x_i)| &= \frac{1}{\sqrt{\mu_p^n(\{x_i\})}}\Big(|P_tf(x_i)-P_sf(x_i)|^2\mu_p^n(\{x_i\})\Big)^{1/2}\\ &\leqslant \frac{1}{\sqrt{\mu_p^n(\{x_i\})}}\|P_tf-P_sf\|_2 < \varepsilon\end{aligned}$$

για κάθε $s \in (t-\delta_i, t+\delta_i) \cap [0,\infty)$.

Οπότε, πράγματι η $\psi_x : [0,\infty) \to \mathbb{R}$ είναι συνεχής για κάθε $x \in X$. $\square$

Λήμμα 2.4.6. *Έστω $f : X \to \mathbb{R}$ συνάρτηση. Έστω επίσης $(P_t)_{t\geqslant 0}$ ημιομάδα Markov με το μ_p^n αναλλοίωτο ως προς αυτήν, με γεννήτορα L και $D(L) = L^2(\mu)$. Τότε, για κάθε $t \geqslant 0$ και για κάθε $[a,b] \subseteq [0,\infty)$ με $t \in [a,b]$ η $\Phi : [a,b] \times X \to \mathbb{R}$, $\Phi(s,x) = P_sf(x)$ είναι φραγμένη.*

Απόδειξη. Έστω $t \geqslant 0$. Θεωρούμε διάστημα $[a,b] \subseteq [0,\infty)$ με $t \in [a,b]$. Τότε, για κάθε $x \in X$, η $\psi_x : [a,b] \to \mathbb{R}$, $\psi_x(s) = P_s f(x)$ είναι συνεχής από το προηγούμενο λήμμα. Συνεπώς, υπάρχει $M_x > 0$ ώστε $|\psi_x(s)| \leqslant M_x$ για κάθε $s \in [a,b]$.

Θέτουμε $M = \max\{M_x : x \in X\}$. Τότε έχουμε ότι

$$|\Phi(s,x)| = |\psi_x(s)| \leqslant M_x \leqslant M$$

για κάθε $(s,x) \in [a,b] \times X$. □

Θεώρημα 2.4.7 (ανισότητα υπερσυσταλτότητας στον διακριτό κύβο)**.** *Έστω* $L : L^2(\mu_p^n) \to L^2(\mu_{p^n})$, $Lf = \sum_{i=1}^n \int_{\{-1,1\}^n} f_i d\mu_p^n - nf$, *ο γεννήτορας Markov που έχουμε ορίσει και* $(P_t)_{t\geqslant 0}$ *η ημιομάδα Markov που παράγει. Τότε, για κάθε* $1 < p < q < \infty$ *και* $t > 0$ *με* $e^{4rt} \geqslant \frac{q-1}{p-1}$ *ισχύει ότι* $\|P_t f\|_q \leqslant \|f\|_p$, *όπου*

$$r = \begin{cases} \frac{p-(1-p)}{\log p - \log(1-p)} & , \text{ αν } p \neq \frac{1}{2} \\ \frac{1}{2} & , \text{ αν } p = \frac{1}{2}. \end{cases}$$

Απόδειξη. Θα το δείξουμε για τις $f : X \to [0,\infty)$ και τότε στη γενική περίπτωση, χρησιμοποιώντας το Πόρισμα 2.2.5, θα έχουμε

$$\|P_t f\|_q = \Big(\int_X |P_t f|^q \, d\mu_p^n\Big)^{1/q} \leqslant \Big(\int_X P_t(|f|)^q \, d\mu_p^n\Big)^{1/q} = \|P_t(|f|)\|_q \leqslant \|\,|f|\,\|_p = \|f\|_p.$$

Ορίζουμε $q(t) = (p-1)e^{4rt} + 1$, $t \geqslant 0$. Έχουμε ότι

$$\begin{aligned} \frac{d}{dt}(P_t f)^{q(t)} &= q'(t)(P_t f)^{q(t)} \log P_t f + q(t)(P_t f)^{q(t)-1} L P_t f \\ &= q'(t)(P_t f)^{q(t)} \log P_t f + q(t)(P_t f)^{q(t)-1} P_t(Lf). \end{aligned}$$

Με βάση το προηγούμενο λήμμα, για κάθε $t_0 \geqslant 0$ και κάθε διάστημα $[a,b] \subseteq [0,\infty)$ με $t_0 \in [a,b]$ υπάρχει $M > 0$ ώστε $\left|\frac{d}{dt}(P_t f(x)^{q(t)})\right| \leqslant M$ για κάθε $x \in X$ και κάθε $t \in [a,b]$. Συνεπώς, από το γενικευμένο θεώρημα κυριαρχημένης σύγκλισης,

$$\frac{d}{dt}\Big(\int_X (P_t f)^{q(t)} d\mu_p^n\Big) = q'(t)\int_X (P_t f)^{q(t)} \log P_t f \, d\mu_p^n + q(t)\int_X (P_t f)^{q(t)-1} L P_t f \, d\mu_p^n. \tag{2.4.1}$$

Οπότε,

$$\begin{aligned} &\|P_t f\|_{q(t)}^{q(t)-1} \frac{d}{dt}\|P_t f\|_{q(t)} \\ &= \Big(\int_X (P_t f)^{q(t)} d\mu_p^n\Big)^{\frac{q(t)-1}{q(t)}} \frac{d}{dt}\Big(\Big[\int_X (P_t f)^{q(t)} d\mu_p^n\Big]^{\frac{1}{q(t)}}\Big) \\ &= -\frac{q'(t)}{q^2(t)}\int_X (P_t f)^{q(t)} d\mu_p^n \log \int_X (P_t f)^{q(t)} d\mu_p^n + \frac{1}{q(t)}\frac{d}{dt}\Big(\int_X (P_t f)^{q(t)} d\mu_p^n\Big) \\ &= -\frac{q'(t)}{q^2(t)}\int_X (P_t f)^{q(t)} d\mu_p^n \log \int_X (P_t f)^{q(t)} d\mu_p^n + \frac{q'(t)}{q^2(t)}\int_X (P_t f)^{q(t)} d\mu_p^n \log \int_X (P_t f)^{q(t)} d\mu_p^n \\ &\qquad + \int_X (P_t f)^{q(t)-1} P_t(Lf) \, d\mu_p^n \\ &= \frac{q'(t)}{q^2(t)}\mathrm{Ent}_{\mu_p^n}((P_t f)^{q(t)}) + \int_X (P_t f)^{q(t)-1} P_t f \, d\mu_p^n. \end{aligned}$$

Γράφοντας για συντομία $q = q(t)$ και $q' = q'(t)$, έχουμε ότι

$$\begin{aligned}(2.4.2)\qquad q^2\|P_tf\|_q^{q-1}\cdot\frac{d}{dt}\|P_tf\|_q &= q'\mathrm{Ent}_{\mu_p^n}((P_tf)^q) + q^2\int_X (P_tf)^{q-1}LP_tf\,d\mu_p^n\\ &= q'\mathrm{Ent}_{\mu_p^n}((P_tf)^q) - q^2\mathcal{E}((P_tf)^{q-1},P_tf).\end{aligned}$$

Ισχυρισμός 2.4.8. Ισχύει ότι

$$\mathcal{E}((P_tf)^{q-1},P_tf) \geqslant \frac{4(q-1)}{q^2}\mathcal{E}((P_tf)^{\frac{q}{2}},(P_tf)^{\frac{q}{2}}).$$

Πράγματι, από το Λήμμα 2.4.3 και το Λήμμα 2.4.4,

$$\begin{aligned}\mathcal{E}((P_tf)^{q-1},P_tf) &= \frac{1}{2}\sum_{i=1}^n\int_X\int_{\{-1,1\}}\int_{\{-1,1\}}(((P_tf)_i(x_i))^{q-1}-((P_tf)_i(y_i))^{q-1})\\ &\qquad\times((P_tf)_i(x_i)-(P_tf)_i(y_i))\,d\mu_p(x_i)\,d\mu_p(y_i)\,d\mu_p^n\\ &\geqslant \frac{1}{2}\sum_{i=1}^n\int_X\int_{\{-1,1\}}\int_{\{-1,1\}}\frac{4(q-1)}{q^2}\big[(((P_tf)_i(x_i))^{\frac{q}{2}}-((P_tf)_i(y_i))^{\frac{q}{2}}\big]^2\\ &\qquad d\mu_p(x_i)\,d\mu_p(y_i)\,d\mu_p^n\\ &= \frac{4(q-1)}{q^2}\mathcal{E}((P_tf)^{\frac{q}{2}},(P_tf)^{\frac{q}{2}}).\end{aligned}$$

Οπότε, από την (2.4.2),

$$\begin{aligned}q^2\|P_tf\|_q^{q-1}\frac{d}{dt}\|P_tf\|_q &\leqslant q'\mathrm{Ent}_{\mu_p^n}((P_tf)^q) - 4(q-1)\mathcal{E}((P_tf)^{\frac{q}{2}},(P_tf)^{\frac{q}{2}})\\ &\leqslant 4r(p-1)e^{4rt}\mathrm{Ent}_{\mu_p^n}((P_tf)^q) - 4r(q-1)\mathrm{Ent}_{\mu_p^n}((P_tf)^q),\end{aligned}$$

εφαρμόζοντας τη λογαριθμική ανισότητα Sobolev για τον διακριτό κύβο για την $(P_tf)^{q/2}$. Άρα,

$$q^2\|P_tf\|_q^{q-1}\frac{d}{dt}\|P_tf\|_q \leqslant 4r\mathrm{Ent}_{\mu_p^n}((P_tf)^q)((p-1)e^{4rt}-(q-1)) = 0.$$

Δηλαδή, για κάθε $t \geqslant 0$ έχουμε

$$\frac{d}{dt}\|P_tf\|_{q(t)} \leqslant 0.$$

Δείξαμε ότι η $t \mapsto \|P_tf\|_{q(t)}$ είναι φθίνουσα, οπότε αν $q > p$ και $t > 0$ ώστε $e^{4rt} \geqslant \frac{q-1}{p-1}$, δηλαδή αν $q(t) \geqslant q$, τότε έχουμε ότι

$$\|f\|_p = \|P_0f\|_{q(0)} \geqslant \|P_tf\|_{q(t)} \geqslant \|P_tf\|_q$$

από το Λήμμα 2.4.1. □

Πόρισμα 2.4.9. *Για κάθε* $1 < p < q < \infty$ *και για κάθε* $t > 0$ *με* $e^{2t} \geqslant \frac{q-1}{p-1}$ *ισχύει ότι*

$$\Big(\frac{1}{2}|a+e^{-t}b|^q + \frac{1}{2}|a-e^{-t}b|^q\Big)^{1/q} \leqslant \Big(\frac{1}{2}|a+b|^p + \frac{1}{2}|a-b|^p\Big)^{1/p}$$

για κάθε $a, b \in \mathbb{R}$.

Απόδειξη. Για $n = 1$, στον $(X, \mathcal{P}(\{-1,1\}), \mu_{1/2})$ έχουμε $r = 1/2$. Οπότε, αν $a, b \in \mathbb{R}$ θεωρούμε την $f(x) = a + bx$, $x \in X$ και από το θεώρημα έχουμε ότι για κάθε $1 < p < q < \infty$ και για κάθε $t > 0$ με $e^{2t} \geqslant \frac{q-1}{p-1}$ ισχύει ότι $\|P_t f\|_q \leqslant \|f\|_p$.

Όμως,

$$\|f\|_p = \Big(\int_X |a+bx|^p d\mu_{1/2}(x)\Big)^{1/p} = \Big(\frac{1}{2}|a+b|^p + \frac{1}{2}|a-b|^p\Big)^{1/p}$$

που είναι το δεξιό μέλος της προς απόδειξη ανισότητας. Έχουμε επίσης

$$Lf(x) = \int_X f\, d\mu_{1/2} - f(x) = \frac{a+b}{2} + \frac{a-b}{2} - (a+bx) = -bx.$$

Έστω ότι $L^k f(x) = (-1)^k bx$. Τότε,

$$L^{k+1} f(x) = L(L^k f)(x) = L((-1)^k bx) = \frac{(-1)^k b - (-1)^k b}{2} - (-1)^k bx = (-1)^{k+1} bx.$$

Οπότε, από επαγωγή,

$$L^k f(x) = \begin{cases} a + bx & , \text{αν } k = 0 \\ (-1)^k bx & , \text{αν } k \geqslant 1. \end{cases}$$

Άρα,

$$P_t f(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k f(x) = a + \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} bx = a + e^{-t} bx.$$

Έπεται ότι

$$\|P_t f\|_q = \Big(\frac{1}{2}|a + e^{-t} b|^q + \frac{1}{2}|a - e^{-t} b|^q\Big)^{1/q},$$

που είναι το αριστερό μέλος της ανισότητας που θέλουμε να δείξουμε. □

2.5 Τυπικό μέτρο του Gauss

Σημειώνουμε ότι στα παρακάτω, όταν είναι σαφές ότι δουλεύουμε στον $\mathbb{R}^n$, για κάθε $x \in \mathbb{R}^n$ θα γράφουμε $|x| = \big(\sum_{i=1}^n |x_i|^2\big)^{1/2}$, την Ευκλείδεια απόσταση δηλαδή του x από την αρχή των αξόνων.

Ορισμός 2.5.1. Το $\gamma_n : \mathcal{B}(\mathbb{R}^n) \to [0,1]$ με

$$\gamma_n(B) = \frac{1}{(2\pi)^{n/2}} \int_B e^{-|x|^2/2} d\lambda_n(x),$$

ή ισοδύναμα, $d\gamma_n(x) = (2\pi)^{-n/2} e^{-|x|^2/2} d\lambda_n(x)$, όπου λ_n είναι το μέτρο Lebesgue στον $\mathbb{R}^n$, θα λέγεται *τυπικό μέτρο του Gauss στον* $\mathbb{R}^n$.

Παρατήρηση 2.5.2. Όταν είναι σαφές ότι είμαστε στον $\mathbb{R}^n$ θα συμβολίζουμε απλά με γ και λ το τυπικό μέτρο του Gauss και το μέτρο Lebesgue στον $\mathbb{R}^n$ αντίστοιχα.

Πρόταση 2.5.3. *Το τυπικό μέτρο του Gauss στον $\mathbb{R}^n$ είναι Borel μέτρο πιθανότητας στον $\mathbb{R}^n$ με $\gamma_n \ll \lambda_n$: το γ_n είναι απολύτως συνεχές ως προς το λ_n.*

Απόδειξη. Το ότι το γ_n είναι μέτρο Borel στον $\mathbb{R}^n$ με $\gamma_n \ll \lambda_n$ είναι σαφές από τον τρόπο ορισμού του. Τώρα,

$$\begin{aligned}\gamma_n(\mathbb{R}^n) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-|x|^2/2} d\lambda_n(x) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \frac{1}{(2\pi)^{n/2}} e^{-\sum_{i=1}^n |x_i|^2} d\lambda_1(x_1) \cdots d\lambda_n(x_n) \\ &= \prod_{i=1}^n \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2} d\lambda_i(x_i) = \prod_{i=1}^n 1 = 1.\end{aligned}$$

Άρα, έχουμε το ζητούμενο. □

Ορισμός 2.5.4. Έστω $(X, \mathcal{A}, \mu)$ χώρος μέτρου, $(Y, \mathcal{B})$ μετρήσιμος χώρος και $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ μετρήσιμη συνάρτηση. Ορίζουμε την *εικόνα του* μ *μέσω της* f ή *push-forward του* μ *ως προς* f ως $f * \mu = \mu \circ f^{-1} : \mathcal{B} \to [0, \infty)$ με τύπο

$$f * \mu(B) = \mu \circ f^{-1}(B) = \mu(f^{-1}(B)).$$

Πρόταση 2.5.5. *Έστω $(X, \mathcal{A}, \mu)$ χώρος μέτρου, $(Y, \mathcal{B})$ μετρήσιμος χώρος και $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ μετρήσιμη συνάρτηση. Τότε:*

(α) *Το $f * \mu$ είναι μέτρο στον $(Y, \mathcal{B})$ με $f * \mu(Y) = \mu(X)$.*

(β) *Αν η $g : (Y, \mathcal{B}) \to \mathbb{R}$ είναι μετρήσιμη και είτε $g \geqslant 0$ είτε $g \in L^1(f * \mu)$, τότε*

$$\int_Y g \, d(\mu \circ f^{-1}) = \int_X g \circ f \, d\mu.$$

Απόδειξη. (α) Προφανές.

(β) Αν $g = \mathbf{1}_B$ για κάποιο $B \in \mathcal{B}$, έχουμε ότι

$$\int_Y \mathbf{1}_B d(\mu \circ f^{-1}) = \mu \circ f^{-1}(B) = \mu(f^{-1}(B)) = \int_X \mathbf{1}_{f^{-1}(B)} \, d\mu = \int_X \mathbf{1}_B \circ f \, d\mu.$$

Αν η g είναι απλή μετρήσιμη, δηλαδή $g = \sum_{i=1}^n \beta_i \mathbf{1}_{B_i}$ για κάποια $n \in \mathbb{N}$, $\beta_i \in \mathbb{R}$, $B_i \in \mathcal{B}$, $1 \leqslant i \leqslant n$, τότε το ζητούμενο έπεται από τη γραμμικότητα του ολοκληρώματος.

Αν $g \geqslant 0$, τότε προσεγγίζοντας την g από αύξουσα ακολουθία μη αρνητικών μετρήσιμων συναρτήσεων και εφαρμόζοντας το θεώρημα μονότονης σύγκλισης, από τα παραπάνω προκύπτει πάλι το ζητούμενο.

Αν $g \in L^1(f * \mu)$, εφαρμόζουμε το αποτέλεσμα για τις g^+, g^-. □

Πρόταση 2.5.6. *Αν $m, n \in \mathbb{N}$ τότε $\gamma_{n+m} = \gamma_n \otimes \gamma_m$.*

Απόδειξη. Αρκεί να δείξουμε ότι $\gamma_{n+m}(A \times B) = \gamma_n(A)\gamma_m(B)$ για κάθε $A \in \mathcal{B}(\mathbb{R}^n)$ και $B \in \mathcal{B}(\mathbb{R}^m)$.

Γράφουμε

$$\begin{aligned}
\gamma_{n+m}(A\times B) &= \frac{1}{(2\pi)^{\frac{n+m}{2}}}\int_{A\times B} e^{-\sum_{i=1}^{n+m}x_i^2/2}d\lambda_{n+m}(x)\\
&= \frac{1}{(2\pi)^{\frac{n+m}{2}}}\int_A\int_B e^{-\sum_{i=1}^{n+m}x_i^2/2}d\lambda_n(x_1,\dots,x_n)\,d\lambda_m(x_{n+1},\dots,x_{n+m})\\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_A e^{-\sum_{i=1}^{n}x_i^2/2}d\lambda_n(x_1,\dots,x_n)\\
&\qquad\times\frac{1}{(2\pi)^{\frac{m}{2}}}\int_B e^{-\sum_{i=n+1}^{n+m}x_i^2/2}d\lambda_m(x_{n+1},\dots,x_{n+m})\\
&= \gamma_n(A)\gamma_m(B).
\end{aligned}$$

□

Πρόταση 2.5.7. *Έστω $\vartheta\in\mathbb{R}$ και $\varphi_\vartheta:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ η συνάρτηση*

$$\varphi_\vartheta(x,y)=\cos\vartheta\cdot x+\sin\vartheta\cdot y.$$

Τότε, $(\gamma_n\otimes\gamma_n)\circ\varphi_\vartheta^{-1}=\gamma_n$.

Απόδειξη. Θεωρούμε τον γραμμικό μετασχηματισμό $L:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n$ που έχει ως πίνακα τον

$$A=\begin{bmatrix}\cos\vartheta\cdot I_n & \sin\vartheta\cdot I_n\\ -\sin\vartheta\cdot I_n & \cos\vartheta\cdot I_n\end{bmatrix}.$$

Παρατηρούμε ότι ο A είναι ορθογώνιος και ότι

$$L(x,y)=(\cos\vartheta\cdot x+\sin\vartheta\cdot y,\cos\vartheta\cdot y-\sin\vartheta\cdot x)=(\varphi_\vartheta(x,y),\cos\vartheta\cdot y-\sin\vartheta\cdot x).$$

Τότε,

$$\begin{aligned}
(\gamma_n\otimes\gamma_n)\circ\varphi_\vartheta^{-1}(B) &= \gamma_n\otimes\gamma_n(\varphi_\vartheta^{-1}(B))=\gamma_n\otimes\gamma_n(L^{-1}(B\times\mathbb{R}^n))\\
&= \frac{1}{(2\pi)^{\frac{2n}{2}}}\int_{L^{-1}(B\times\mathbb{R}^n)} e^{-|(x,y)|^2/2}d\lambda_{2n}(x,y)\\
&= \frac{1}{(2\pi)^{\frac{2n}{2}}}\int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-|(x,y)|^2/2}\mathbf{1}_{B\times\mathbb{R}^n}\circ L\,d\lambda_{2n}(x,y)\\
&= \frac{1}{(2\pi)^{\frac{2n}{2}}}\int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-|L^{-1}(x,y)|^2/2}\mathbf{1}_{B\times\mathbb{R}^n}\,d\lambda_{2n}(x,y)\\
&= \frac{1}{(2\pi)^{\frac{2n}{2}}}\int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-|(x,y)|^2/2}\mathbf{1}_{B\times\mathbb{R}^n}\,d\lambda_{2n}(x,y)\\
&= \frac{1}{(2\pi)^{\frac{2n}{2}}}\int_{B\times\mathbb{R}^n} e^{-|(x,y)|^2/2}\,d\lambda_{2n}(x,y)\\
&= \gamma_{2n}(B\times\mathbb{R}^n)=\gamma_n(B)\gamma_n(\mathbb{R}^n)=\gamma_n(B),
\end{aligned}$$

όπου πριν χρησιμοποιήσαμε τον τύπο αλλαγής μεταβλητής, λαμβάνοντας υπ' όψιν ότι ο L είναι ορθογώνιος και συνεπώς $|\det L|=1$, καθώς και το ότι $|L(x,y)|=|L^{-1}(x,y)|=|(x,y)|$ για κάθε $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$.

Συνεπώς, δείξαμε ότι για κάθε $B\in\mathcal{B}(\mathbb{R}^n)$ έχουμε $(\gamma_n\otimes\gamma_n)\circ\varphi_\vartheta^{-1}(B)=\gamma_n(B)$, και επομένως έχουμε το ζητούμενο. □

2.6 Ημιομάδα Ornstein-Uhlenbeck

Ορισμός 2.6.1. Η *ημιομάδα Ornstein-Uhlenbeck* $(T_t)_{t\geqslant 0}$ στον $\mathbb{R}^n$ ορίζεται ως εξής: Για $t \geqslant 0$, ο $T_t : L^1(\gamma_n) \to L^1(\gamma_n)$ είναι ο τελεστής

$$T_t f(x) = \int_{\mathbb{R}^n} f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)\, d\gamma_n(y), \qquad x \in \mathbb{R}^n.$$

Παρατήρηση 2.6.2. Η $(T_t)_{t\geqslant 0}$ είναι καλά ορισμένη και $\|T_t f\|_1 \leqslant \|f\|_1$ για κάθε $f \in L^1(\gamma_n)$.

Πράγματι, έστω $t \geqslant 0$. Παρατηρούμε ότι $(e^{-t})^2 + \big(\sqrt{1-e^{-2t}}\big)^2 = 1$. Οπότε, υπάρχει $\vartheta \in \mathbb{R}$ ώστε $e^{-t} = \cos\vartheta$ και $\sqrt{1-e^{-2t}} = \sin\vartheta$. Συνεπώς, αν $f \in L^1(\gamma_n)$, χρησιμοποιώντας την $\gamma_n = (\gamma_n \otimes \gamma_n) \circ \varphi_\vartheta^{-1}$ που δείξαμε πριν, έχουμε

$$\begin{aligned} +\infty > \int_{\mathbb{R}^n} |f(z)|\, d\gamma_n(z) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(\cos\vartheta\cdot x + \sin\vartheta \cdot y)|\, d\gamma_n(y) d\gamma_n(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)|\, d\gamma_n(y) d\gamma_n(x), \end{aligned} \tag{2.6.1}$$

απ' όπου έπεται ότι

$$\int_{\mathbb{R}^n} |f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)|\, d\gamma_n(y) < +\infty$$

για γ_n-σχεδόν κάθε $x \in \mathbb{R}^n$ και μάλιστα, χωρίς περιορισμό της γενικότητας, για κάθε $x \in \mathbb{R}^n$. Επομένως,

$$T_t f(x) = \int_{\mathbb{R}^n} f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)\, d\gamma_n(y) \in \mathbb{R}$$

για κάθε $x \in \mathbb{R}^n$. Τέλος, από την (2.6.1) παίρνουμε ότι

$$\int_{\mathbb{R}^n} |T_t f(x)|\, d\gamma_n(x) \leqslant \int_{\mathbb{R}^n} |f(z)|\, d\gamma_n(z) < +\infty,$$

και άρα $T_t f \in L^1(\gamma_n)$ και $\|T_t f\|_1 \leqslant \|f\|_1$.

Πρόταση 2.6.3. *Για κάθε $f \in L^1(\gamma_n)$ και $t \geqslant 0$ ισχύει ότι*

$$\int_{\mathbb{R}^n} T_t f\, d\gamma_n = \int_{\mathbb{R}^n} f\, d\gamma_n,$$

δηλαδή το μέτρο Gauss είναι αναλλοίωτο ως προς την $(T_t)_{t\geqslant 0}$.

Απόδειξη. Ακριβώς όπως στην παρατήρηση έχουμε ότι

$$\int_{\mathbb{R}^n} f(z)\, d\gamma_n(z) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)\, d\gamma_n(y) d\gamma_n(x) = \int_{\mathbb{R}^n} T_t f(x)\, d\gamma_n(x).$$

□

Πρόταση 2.6.4. *Για $p \geqslant 1$, αν $f \in L^p(\gamma_n)$ τότε $T_t f \in L^p(\gamma_n)$ και $\|T_t f\|_p \leqslant \|f\|_p$ για κάθε $t \geqslant 0$.*

Απόδειξη. Για $p=1$ έχουμε ήδη αποδείξει τον ισχυρισμό. Έστω λοιπόν $p>1$ και $t\geqslant 0$. Τότε, αν $f\in L^p(\gamma_n)$ έχουμε ότι

$$+\infty > \int_{\mathbb{R}^n} |f(z)|^p\,d\gamma_n(z) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)|^p\,d\gamma_n(y)d\gamma_n(x). \tag{2.6.2}$$

Επομένως, η

$$x\mapsto \int_{\mathbb{R}^n} |f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)|^p\,d\gamma_n(y)$$

είναι ολοκληρώσιμη. Τώρα, από την ανισότητα Hölder και την (2.6.2) παίρνουμε ότι

$$\begin{aligned}\|T_tf\|_p^p &\leqslant \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n}|f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)|\,d\gamma_n(x)\Big)^p d\gamma_n(y)\\ &\leqslant \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n}|f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)|^p\,d\gamma_n(x)\Big)d\gamma_n(y)\\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)|^p\,d\gamma_n(y)d\gamma_n(x)\\ &= \|f\|_p^p<+\infty.\end{aligned}$$

Έπεται ότι $T_tf\in L^p(\gamma_n)$ και $\|T_tf\|_p\leqslant \|f\|_p$. □

Πρόταση 2.6.5. *Για κάθε $t,s\geqslant 0$ ισχύει ότι $T_{t+s}=T_t\circ T_s$. Επίσης, $T_0=I$.*

Απόδειξη. Έστω $f\in L^1(\mu)$ και $t,s>0$. Τότε,

$$\begin{aligned}&T_t(T_sf)(x)=\int_{\mathbb{R}^n} T_sf\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y)\\ &=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\big(e^{-s}\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)+\sqrt{1-e^{-2s}}z\big)\,d\gamma_n(z)\,d\gamma_n(y)\\ &=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\big(e^{-(s+t)}x+e^{-s}\sqrt{1-e^{-2t}}y+\sqrt{1-e^{-2s}}z\big)\,d\gamma_n(z)\,d\gamma_n(y)\\ &=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\Big(e^{-(s+t)}x+\sqrt{1-e^{-2(t+s)}}\Big(\frac{e^{-s}\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2(t+s)}}}y+\frac{\sqrt{1-e^{-2s}}}{\sqrt{1-e^{-2(t+s)}}}z\Big)\Big)\,d\gamma_n(z)\,d\gamma_n(y).\end{aligned}$$

Παρατηρούμε ότι

$$\Big(\frac{e^{-s}\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2(t+s)}}}\Big)^2+\Big(\frac{\sqrt{1-e^{-2s}}}{\sqrt{1-e^{-2(t+s)}}}\Big)^2=1.$$

Οπότε, όπως πριν, παίρνουμε ότι

$$T_t(T_sf)(x)=\int_{\mathbb{R}^n} f\big(e^{-(s+t)}x+\sqrt{1-e^{-2(s+t)}}w\big)\,d\gamma_n(w)=T_{s+t}f(x).$$

Αν $t=0$ ή $s=0$ τότε το αποτέλεσμα είναι άμεσο, αφού

$$T_0f(x)=\int_{\mathbb{R}^n} f(x)\,d\gamma_n(y)=f(x)$$

για κάθε $x\in\mathbb{R}^n$. Ο υπολογισμός αυτός δείχνει επίσης ότι $T_0=I$. □

Πρόταση 2.6.6. *Έστω $f,g\in L^1(\gamma_n)$. Τότε ισχύουν τα παρακάτω:*

(i) *Αν η f υποτεθεί φραγμένη και συνεχής, τότε*

$$\lim_{t\to+\infty} T_t f(x) = \int_{\mathbb{R}^n} f\,d\gamma_n, \qquad x\in\mathbb{R}^n.$$

(ii) *Αν $f\geqslant g$ τότε $T_t f\geqslant T_t g$ για κάθε $t\geqslant 0$.*

(iii) *Για κάθε $t\geqslant 0$ και $\lambda,\mu\in\mathbb{R}$ ισχύει ότι $T_t(\lambda f+\mu g)=\lambda T_t f+\mu T_t g$.*

(iv) *$T_t\mathbf{1}=\mathbf{1}$ για κάθε $t\geqslant 0$.*

Απόδειξη. Για την (i) αρχικά έχουμε ότι

$$T_t f(x) = \int_{\mathbb{R}^n} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y).$$

Τώρα, $\lim_{t\to\infty} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big) = f(y)$ λόγω συνέχειας της f, και αφού η f είναι φραγμένη συνάρτηση, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης παίρνουμε ότι

$$\lim_{t\to\infty} T_t f(x) = \int_{\mathbb{R}^n} \lim_{t\to\infty} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y) = \int_{\mathbb{R}^n} f(y)\,d\gamma_n(y).$$

Οι (ii) και (iii) ελέγχονται άμεσα από τον ορισμό. Τέλος, για την (iv) παρατηρούμε ότι

$$T_t\mathbf{1}(x) = \int_{\mathbb{R}^n} \mathbf{1}\,d\gamma_n(y) = \gamma_n(\mathbb{R}^n) = 1$$

για κάθε $x\in\mathbb{R}^n$. □

Πρόταση 2.6.7. *Έστω $f,g\in L^2(\gamma_n)$. Τότε ισχύουν τα παρακάτω:*

(i) *$(T_t(fg))^2\leqslant T_t(f^2)T_t(g^2)$ για κάθε $t\geqslant 0$.*

(ii) *Για κάθε $t\geqslant 0$ ισχύει ότι*
$$\int_{\mathbb{R}^n} (T_t f)g\,d\gamma_n = \int_{\mathbb{R}^n} (T_t g)f\,d\gamma_n.$$

Απόδειξη. (i) Για κάθε $t\geqslant 0$ έχουμε ότι

$$\begin{aligned}(T_t(fg)(x))^2 &= \Big(\int_{\mathbb{R}^n} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)g\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y)\Big)^2\\ &\leqslant \Big(\int_{\mathbb{R}^n} f^2\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y)\Big)\Big(\int_{\mathbb{R}^n} g^2\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\,d\gamma_n(y)\Big)\\ &= T_t(f^2)T_t(g^2)\end{aligned}$$

εφαρμόζοντας την ανισότητα Cauchy-Schwarz.

(ii) Έστω $t\geqslant 0$. Γράφουμε

$$\begin{aligned}\int_{\mathbb{R}^n} (T_t f)g\,d\gamma_n &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)g(x)\,d\gamma_n(y)d\gamma_n(x)\\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)g(x)e^{-\frac{|(x,y)|^2}{2}}\,d\lambda(y)d\lambda(x).\end{aligned} \tag{2.6.3}$$

Θεωρούμε $\vartheta \in \mathbb{R}$ ώστε $\cos\vartheta = e^{-t}$, $\sin\vartheta = \sqrt{1-e^{-2t}}$, και τον γραμμικό μετασχηματισμό $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ με πίνακα

$$A = \begin{bmatrix} \cos\vartheta \cdot I_n & \sin\vartheta \cdot I_n \\ -\sin\vartheta \cdot I_n & \cos\vartheta \cdot I_n \end{bmatrix}.$$

Παρατηρούμε ότι ο A είναι ορθογώνιος με

$$A^{-1} = A^t = \begin{bmatrix} \cos\vartheta \cdot I_n & -\sin\vartheta \cdot I_n \\ \sin\vartheta \cdot I_n & \cos\vartheta \cdot I_n \end{bmatrix}.$$

και $L(x,y) = (\cos\vartheta \cdot x + \sin\vartheta \cdot y, -\sin\vartheta \cdot x + \cos\vartheta \cdot y)$ και $L^{-1}(x,y) = (\cos\vartheta \cdot x - \sin\vartheta \cdot y, \sin\vartheta \cdot x + \cos\vartheta \cdot y)$.

Συνεπώς, κάνοντας αλλαγή μεταβλητής, αφού $|\det L| = 1$, παίρνουμε από την (2.6.3) ότι

$$\begin{aligned} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (T_t f) g \, d\gamma_n &= \frac{1}{(2\pi)^{\frac{2n}{2}}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x) g(\cos\vartheta \cdot x - \sin\vartheta \cdot y) e^{-\frac{|L^{-1}(x,y)|^2}{2}} d\lambda(x) d\lambda(y) \\ &= \frac{1}{(2\pi)^{\frac{2n}{2}}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x) g(\cos\vartheta \cdot x + \sin\vartheta \cdot y) e^{-\frac{|(x,y)|^2}{2}} d\lambda(x) d\lambda(y) \\ &= \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} g\big(e^{-t}x + \sqrt{1-e^{-2t}} y\big) \, d\gamma_n(y) d\gamma_n(x) \\ &= \int_{\mathbb{R}^n} f(x) (T_t g)(x) \, d\gamma_n(x), \end{aligned}$$

όπου στη μεσαία ισότητα χρησιμοποιήσαμε την ορθογωνιότητα του L και την συμμετρία του μέτρου Lebesgue, ότι δηλαδή

$$\int_{\mathbb{R}^n} h(-x) \, d\lambda(x) = \int_{\mathbb{R}^n} h(x) \, d\lambda(x)$$

για κάθε $h \in L^1(\lambda)$. $\square$

Ορισμός 2.6.8. Αν $g_1, \ldots, g_n \in L^1(\gamma)$, ορίζουμε

$$T_t(g_1, \ldots, g_n) := (T_t(g_1), \ldots, T_t(g_n)), \quad t \geqslant 0.$$

Επίσης, συμβολίζουμε με $C_b^k(\mathbb{R}^n)$ την κλάση όλων των συναρτήσεων $f : \mathbb{R}^n \to \mathbb{R}$ που έχουν φραγμένες και συνεχείς μερικές παραγώγους έως και k-στης τάξης.

Πρόταση 2.6.9. *Έστω $f \in C_b^1(\mathbb{R}^n)$. Τότε, για κάθε $t \geqslant 0$ ισχύει ότι*

$$\nabla T_t(f) = e^{-t} T_t(\nabla f).$$

Απόδειξη. Έχουμε ότι

$$\begin{aligned}\frac{\partial T_t f}{\partial x_i}(x) &= \lim_{h\to 0}\frac{T_t f(x+he_i)-T_t f(x)}{h}\\ &= \lim_{h\to 0}\int_{\mathbb{R}^n}\frac{f(e^{-t}(x+he_i)+\sqrt{1-e^{-2t}}y)-f(e^{-t}x+\sqrt{1-e^{-2t}}y)}{h}\,d\gamma_n(y)\\ &= \int_{\mathbb{R}^n}\lim_{h\to 0}\frac{f(e^{-t}(x+he_i)+\sqrt{1-e^{-2t}}y)-f(e^{-t}x+\sqrt{1-e^{-2t}}y)}{h}\,d\gamma_n(y)\\ &= \int_{\mathbb{R}^n} e^{-t}\frac{\partial f}{\partial x_i}(e^{-t}x+\sqrt{1-e^{-2t}}y)\,d\gamma_n(y)\\ &= e^{-t}T_t\Big(\frac{\partial f}{\partial x_i}\Big)(x),\end{aligned}$$

λόγω του γενικευμένου θεωρήματος κυριαρχημένης σύγκλισης, αφού η ∇f είναι φραγμένη, σε χώρο πιθανότητας. Συνεπώς,

$$\nabla(T_t f) = e^{-t}\Big(T_t\Big(\frac{\partial f}{\partial x_1}\Big),\dots,T_t\Big(\frac{\partial f}{\partial x_n}\Big)\Big) = e^{-t}T_t(\nabla f).$$

□

Ορισμός 2.6.10. Στον $C_b^2(\mathbb{R}^n)$ ορίζουμε τον γεννήτορα L της ημιομάδας Ornstein-Uhlenbeck ως εξής: Αν $f\in C_b^2(\mathbb{R}^n)$ και $x\in\mathbb{R}^n$, θέτουμε

$$Lf(x) = \lim_{t\to 0^+}\frac{T_t f(x)-f(x)}{t}.$$

Πρόταση 2.6.11. *Ο L είναι καλά ορισμένος και ικανοποιεί τις εξής σχέσεις:*

(i) $(Lf)(x) = \Delta f(x) - \langle x,\nabla f(x)\rangle$, *για κάθε* $x\in\mathbb{R}^n$.

(ii) $\frac{d}{dt}(T_t f) = L(T_t f) = T_t(Lf)$.

Απόδειξη. Έστω $f\in C_b^k(\mathbb{R}^n)$ και $x\in\mathbb{R}^n$.
(i) Για $t>0$ και $h\neq 0$ έχουμε ότι

$$\frac{T_{t+h}f(x)-T_t f(x)}{h} = \int_{\mathbb{R}^n}\frac{f(e^{-(t+h)}x+\sqrt{1-e^{-2(t+h)}}y)-f(e^{-t}x+\sqrt{1-e^{-2t}}y)}{h}\,d\gamma_n(y)$$

και

$$\begin{aligned}&\lim_{h\to 0}\frac{f(e^{-(t+h)}x+\sqrt{1-e^{-2(t+h)}}y)-f(e^{-t}x+\sqrt{1-e^{-2t}}y)}{h}\\ &\quad= -\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y),e^{-t}x\rangle + \Big\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y),\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\Big\rangle = g(t,y).\end{aligned}$$

Τώρα, θεωρώντας $0<t_1<t<t_2<+\infty$, αφού η ∇f είναι φραγμένη, υπάρχουν $M_1,M_2>0$ ώστε $|g(t,y)|\leqslant M_1+M_2\sum_{i=1}^n|y_i|$ για κάθε $t\in[t_1,t_2]$ και για κάθε $y\in\mathbb{R}^n$, και

$$\int_{\mathbb{R}^n}\Big(M_1+M_2\sum_{i=1}^n|y_i|\Big)\,d\gamma_n(y)<+\infty.$$

Οπότε, από το θεώρημα κυριαρχημένης σύγκλισης,

$$\begin{aligned}\lim_{h\to 0}\frac{T_{t+h}f(x)-T_tf(x)}{h} &= -\int_{\mathbb{R}^n}\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y), e^{-t}x\rangle\,d\gamma_n(y)\\ &\quad+\int_{\mathbb{R}^n}\Big\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y,\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\Big\rangle\,d\gamma_n(y).\end{aligned}$$

Επομένως, για να δείξουμε ότι ο L είναι καλά ορισμένος, αρκεί να δείξουμε ότι υπάρχει το όριο της τελευταίας ποσότητας όταν $t\to 0^+$. Προφανώς,

$$\lim_{t\to 0^+}\int_{\mathbb{R}^n}-\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y), e^{-t}x\rangle\,d\gamma_n(y) = -\int_{\mathbb{R}^n}\langle\nabla f(x),x\rangle\,d\gamma_n(y) = -\langle\nabla f(x),x\rangle,$$

αφού $f\in L^2(\gamma_n)$ και η ∇f είναι φραγμένη. Για τον δεύτερο όρο του δεξιού μέλους έχουμε ότι

$$\begin{aligned}&\int_{\mathbb{R}^n}\Big\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y,\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\Big\rangle\,d\gamma_n(y)\\ &=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\Big\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y,\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y\Big\rangle e^{-|y|^2/2}\,dy\\ &=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\Big\langle\nabla f(e^{-t}x+\sqrt{1-e^{-2t}}y\frac{e^{-2t}}{\sqrt{1-e^{-2t}}},ye^{-|y|^2/2}\Big\rangle\,dy\\ &=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\langle\nabla K_t(y),\nabla h(y)\rangle\,dy,\end{aligned}\tag{2.6.4}$$

όπου

$$K_t(y)=f(e^{-t}x+\sqrt{1-e^{-2t}}y)\frac{e^{-2t}}{1-e^{-2t}}\quad\text{και}\quad h(y)=\frac{1}{(2\pi)^{n/2}}e^{-|y|^2/2}.$$

Έχουμε ότι

$$\int_{\mathbb{R}^n}\langle\nabla K_t(y),\nabla h(y)\rangle\,dy=\int_{\mathbb{R}^n}\Delta K_t(y)e^{-|y|^2/2}dy.$$

Τώρα,

$$\Delta K_t(y)=\sum_{i=1}^n\frac{\partial^2}{\partial y_i^2}f(e^{-t}x+\sqrt{1-e^{-2t}}y)\longrightarrow\sum_{i=1}^n\frac{\partial^2 f}{\partial y_i^2}(x)=\Delta f(x)$$

όταν $t\to 0^+$. Οπότε, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης παίρνουμε τελικά ότι

$$\begin{aligned}\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\Delta K_t(y)e^{-|y|^2/2}dy &\longrightarrow\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\Delta f(x)e^{-|y|^2/2}dy\\ &=\Delta f(x)\int_{\mathbb{R}^n}\mathbf{1}\,d\gamma_n(y)=\Delta f(x).\end{aligned}$$

Άρα, πράγματι ο L είναι καλά ορισμένος και

$$Lf(x)=\Delta f(x)-\langle x,\nabla f(x)\rangle.$$

(ii) Για $t>0$ έχουμε

$$\frac{d}{dt}T_tf(x)=\lim_{h\to 0}\frac{T_{t+h}f(x)-T_tf(x)}{h}=\lim_{h\to 0}T_t\Big(\frac{T_hf(x)-f(x)}{h}\Big)=T_t(Lf(x)),$$

αφού $f \in C_b^2(\mathbb{R}^n)$. Υπολογίζοντας το ίδιο όριο με διαφορετικό τρόπο, παίρνουμε ότι

$$\frac{d}{dt}T_t f(x) = \lim_{h \to 0} \frac{T_h(T_t f(x)) - T_t f(x)}{h} = L(T_t f(x)),$$

και άρα έχουμε τη ζητούμενη ιδιότητα. □

Παρατήρηση 2.6.12. Αν $f, g \in C_b^2(\mathbb{R}^n)$ τότε η μορφή Dirichlet της ημιομάδας Ornstein-Uhlenbeck ορίζεται ως

$$\mathcal{E}(f,g) = -\int_{\mathbb{R}^n} f \cdot Lg \, d\gamma_n = \int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle \, d\gamma_n,$$

από τον τύπο του Green.

Πρόταση 2.6.13. *Αν $g \in C_b^2(\mathbb{R}^n)$ τότε*

$$\int_{\mathbb{R}^n} Lg \, d\gamma_n = 0.$$

Απόδειξη. Παρατηρούμε ότι

$$\int_{\mathbb{R}^n} Lg \, d\gamma_n = \int_{\mathbb{R}^n} \mathbf{1} \cdot Lg \, d\gamma_n = -\int_{\mathbb{R}^n} \langle \nabla \mathbf{1}, \nabla g \rangle, d\gamma_n = 0.$$

□

2.7 Λογαριθμική ανισότητα Sobolev στον χώρο του Gauss

Ορισμός 2.7.1. Έστω μ ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$. Θα λέμε ότι το μ ικανοποιεί *λογαριθμική ανισότητα Sobolev με σταθερά* $\beta > 0$ αν για κάθε $f \in L^2(\mu)$ ισχύει ότι

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{1}{\beta} \int_{\mathbb{R}^n} |\nabla f|^2 \, d\mu.$$

Λήμμα 2.7.2. *Έστω $\varphi \in C_b^2(\gamma_n)$ με $\varphi \geqslant \varepsilon$ για κάποιο $\varepsilon > 0$. Τότε,*

$$\int_{\mathbb{R}^n} \frac{d}{dt}(T_t\varphi \cdot \log T_t\varphi) \, d\gamma_n = -\int_{\mathbb{R}^n} \frac{|\nabla(T_t\varphi)|^2}{T_t\varphi} \, d\gamma_n.$$

Απόδειξη. Αρχικά, αφού $\varphi \geqslant \varepsilon$ έχουμε ότι $T_t\varphi \geqslant \varepsilon$. Τώρα,

$$\frac{d}{dt}(T_t\varphi \log(T_t\varphi)) = LT_t\varphi \cdot \log(T_t\varphi) + LT_t\varphi.$$

Οπότε,

$$\begin{aligned}
\int_{\mathbb{R}^n} \frac{d}{dt}(T_t\varphi \cdot \log T_t\varphi) \, d\gamma_n &= \int_{\mathbb{R}^n} LT_t\varphi \cdot \log(T_t\varphi) \, d\gamma_n + \int_{\mathbb{R}^n} LT_t\varphi \, d\gamma_n \\
&= -\int_{\mathbb{R}^n} \langle \nabla T_t\varphi, \nabla \log T_t\varphi \rangle \, d\gamma_n = -\int_{\mathbb{R}^n} \left\langle \nabla T_t\varphi, \frac{\nabla T_t\varphi}{T_t\varphi} \right\rangle d\gamma_n \\
&= -\int_{\mathbb{R}^n} \frac{|\nabla(T_t\varphi)|^2}{T_t\varphi} \, d\gamma_n,
\end{aligned}$$

όπου χρησιμοποιήσαμε και την $\int_{\mathbb{R}^n} LT_t\varphi \, d\gamma_n = 0$ (βλ. Πρόταση 2.6.13). □

Λήμμα 2.7.3. *Έστω* $\varphi \in C_b^2(\gamma_n)$ *με* $\varphi \geqslant 0$ *και* $\int_{\mathbb{R}^n} \varphi \, d\gamma_n = 1$. *Τότε,*

$$\mathrm{Ent}_{\gamma_n}(\varphi) \leqslant \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla\varphi|^2}{\varphi} \, d\gamma_n.$$

Απόδειξη. Μπορούμε να υποθέσουμε ότι $\varphi \geqslant \varepsilon$ για κάποιο $\varepsilon > 0$. Έχουμε ότι

$$\begin{aligned}\varphi \log \varphi &= \varphi \log \varphi - \int_{\mathbb{R}^n} \varphi \, d\gamma_n \cdot \log \Big(\int_{\mathbb{R}^n} \varphi \, d\gamma_n \Big) \\ &= T_0\varphi \cdot \log T_0\varphi - \lim_{t\to+\infty} (T_t\varphi \cdot \log T_t\varphi) = - \Big[T_t\varphi \cdot \log T_t \varphi \Big]_0^{+\infty} \\ &= - \int_0^\infty \frac{d}{dt} (T_t\varphi \cdot \log T_t\varphi) \, d\gamma_n.\end{aligned}$$

Χρησιμοποιώντας και το Λήμμα 2.7.2, γράφουμε

$$\begin{aligned}\mathrm{Ent}_{\gamma_n}(\varphi) &= \int_{\mathbb{R}^n} \varphi \log \varphi \, d\gamma_n = - \int_{\mathbb{R}^n} \int_0^{+\infty} \frac{d}{dt} (T_t\varphi \cdot \log T_t\varphi) \, dt \, d\gamma_n \\ &= - \int_0^{+\infty} \frac{d}{dt} \Big(\int_{\mathbb{R}^n} T_t\varphi \cdot \log T_t\varphi \, d\gamma_n \Big) \, dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^n} \frac{|\nabla T_t\varphi|^2}{T_t\varphi} \, d\gamma_n \\ &= \int_0^{+\infty} \int_{\mathbb{R}^n} \frac{|e^{-t} T_t(\nabla\varphi)|^2}{T_t\varphi} \, d\gamma_n \, dt \\ &= \int_0^{+\infty} e^{-2t} \int_{\mathbb{R}^n} \frac{|T_t(\nabla\varphi)|^2}{T_t\varphi} \, d\gamma_n \, dt.\end{aligned} \tag{2.7.1}$$

Η εναλλαγή παραγώγου και ολοκληρώματος στην τρίτη ισότητα επιτρέπεται καθώς όλες οι ποσότητες είναι είτε φραγμένες είτε απόλυτα ολοκληρώσιμες. Τώρα,

$$\begin{aligned}|T_t(\nabla\varphi)(x)|^2 &= \sum_{i=1}^n \Big(\int_{\mathbb{R}^n} \frac{\partial\varphi}{\partial x_i} (e^{-t}x + \sqrt{1-e^{-2t}}y) \frac{e^{-|y|^2/2}}{(2\pi)^{n/2}} \, dy \Big)^2 \\ &\leqslant \Big(\int_{\mathbb{R}^n} \Big(\sum_{i=1}^n \Big(\frac{\partial\varphi}{\partial x_i} (e^{-t}x + \sqrt{1-e^{-2t}}y) \Big)^2 \Big(\frac{e^{-|y|^2/2}}{(2\pi)^{n/2}} \Big)^2 \Big)^{1/2} \, dy \Big) \\ &= T_t(|\nabla\varphi|)(x)^2,\end{aligned}$$

δηλαδή,

$$|T_t(\nabla\varphi)|^2 \leqslant T_t(|\nabla\varphi|)^2. \tag{2.7.2}$$

Όμως,

$$\frac{T_t(|\nabla\varphi|)^2}{T_t\varphi} = \frac{T_t\left(\frac{|\nabla\varphi|\sqrt{\varphi}}{\sqrt{\varphi}}\right)^2}{T_t\varphi} \leqslant \frac{T_t\left(\frac{|\nabla\varphi|^2}{\varphi}\right) T_t\varphi}{T_t\varphi} = T_t\Big(\frac{|\nabla\varphi|^2}{\varphi}\Big). \tag{2.7.3}$$

Οπότε, από τις (2.7.1), (2.7.2) και (2.7.3) έχουμε ότι

$$\begin{aligned}\mathrm{Ent}_{\gamma_n}(\varphi) &\leqslant \int_0^{+\infty} e^{-2t} \int_{\mathbb{R}^n} T_t\Big(\frac{|\nabla\varphi|^2}{\varphi}\Big) \, d\gamma_n \, dt = \int_0^{+\infty} e^{-2t} \int_{\mathbb{R}^n} \frac{|\nabla\varphi|^2}{\varphi} \, d\gamma_n \, dt \\ &= \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla\varphi|^2}{\varphi} \, d\gamma_n,\end{aligned}$$

λόγω του αναλλοίωτου του γ_n ως προς την $(T_t)_{t\geqslant 0}$. □

Θεώρημα 2.7.4 (λογαριθμική ανισότητα Sobolev)**.** *Για κάθε $f \in C_b^2(\gamma_n)$ ισχύει η ανισότητα*

$$\int_{\mathbb{R}^n} f^2 \log |f| \, d\gamma_n - \frac{1}{2}\int_{\mathbb{R}^n} f^2 \, d\gamma_n \cdot \log \left(\int_{\mathbb{R}^n} f^2 \, d\gamma_n \right) \leqslant \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n,$$

δηλαδή

$$\mathrm{Ent}_{\gamma_n}(f) \leqslant \int |\nabla f|^2 d\gamma_n.$$

Απόδειξη. Θα υποθέσουμε ότι $f \in C_b^\infty(\mathbb{R}^n)$ και $f \geqslant c > 0$. Θέτουμε $\varphi = f^2$, οπότε $\nabla f = \frac{\nabla \varphi}{2\sqrt{\varphi}}$ και η ανισότητα παίρνει τη μορφή

$$\int_{\mathbb{R}^n} \varphi \log \varphi \, d\gamma_n - \int_{\mathbb{R}^n} \varphi \, d\gamma_n \cdot \log \left(\int_{\mathbb{R}^n} \varphi \, d\gamma_n \right) \leqslant \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla \varphi|^2}{\varphi} \, d\gamma_n,$$

η οποία αποδείχθηκε στο Λήμμα 2.7.3. □

ΚΕΦΑΛΑΙΟ 3

Η ανισότητα της απόστασης από την κυρτή θήκη

3.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζουμε δύο αποδείξεις της ανισότητας απόστασης από την κυρτή θήκη. Έστω $(\Omega_i, \mathcal{A}_i, \mu_i)$, $1 \leqslant i \leqslant n$ χώροι πιθανότητας. Θεωρούμε τον χώρο γινόμενο $(\Omega, \mathcal{A}, \mathbb{P}) = (\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \mu_i)$. Δοθέντων $A \subseteq \Omega$ και $x \in \Omega$, ορίζουμε

$$U_A(x) = \{(s_i)_{1 \leqslant i \leqslant n} \in \{0,1\}^n \mid \text{ υπάρχει } y \in A \text{ ώστε αν } s_i = 0 \text{ τότε } x_i = y_i \text{ για κάθε } 1 \leqslant i \leqslant n\}.$$

Συμβολίζουμε με $V_A(x)$ την κυρτή θήκη $\mathrm{conv}(U_A(x))$ την οποία θεωρούμε ως υποσύνολο του $\mathbb{R}^n$. Τότε, $0 \in V_A(x)$ αν και μόνο αν $x \in A$. Θεωρούμε τη συνάρτηση

$$d_A(x) = \mathrm{dist}(0, V_A(x)).$$

Η ανισότητα του Talagrand ισχυρίζεται ότι για κάθε $A \in \mathcal{A}$ ισχύει ότι

$$\int_\Omega e^{\frac{1}{4}d_A^2} d\mathbb{P} \leqslant \frac{1}{\mathbb{P}(A)}. \tag{3.1.1}$$

Παρουσιάζουμε αρχικά μια απόδειξη που βασίζεται στη μέθοδο της εντροπίας και οφείλεται στους Boucheron, Lugosi και Massart. Στη συνέχεια παρουσιάζουμε την αρχική απόδειξη του Talagrand ξεκινώντας από προγενέστερα αποτελέσματα που οδήγησαν σε αυτήν. Ο Talagrand κατέληξε στην διατύπωση της ανισότητας απόστασης από την κυρτή θήκη, ξεκινώντας από ένα αποτέλεσμα που απέδειξε για τον διακριτό κύβο. Ενδιάμεσα, οι Johnson και Schechtman γενίκευσαν αυτό το αποτέλεσμα του Talagrand, και αυτή η γενίκευση στάθηκε το κίνητρο για την γενική διατύπωση της ανισότητας. Εκτός από την (3.1.1), ο Talagrand απέδειξε και άλλες ανισότητες αυτού του τύπου, στις οποίες υπεισέρχονται αποστάσεις διαφορετικές από την d_A. Παρουσιάζουμε όλα αυτά τα αποτελέσματα, και στο τέλος του κεφαλαίου περιγράφουμε κάποιες ενδεικτικές εφαρμογές.

3.2 Απόδειξη μέσω εντροπίας

Αρχικά αναφέρουμε το πλαίσιο στο οποίο θα εργαστούμε. Θεωρούμε τους χώρους πιθανότητας $(\Omega_i, \mathcal{A}_i, \mu_i)$, $1 \leqslant i \leqslant n$ και τον χώρο γινόμενο $(X, \mathcal{A}, P) = (\times_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \mu_i)$.

Στα παρακάτω θα υποθέσουμε ότι όλα τα σύνολα και οι συναρτήσεις που χρησιμοποιούνται είναι μετρήσιμα. Αυτό είναι εφικτό αν κάνουμε την υπόθεση ότι οι Ω_i είναι Πολωνικοί, τα μ_i μέτρα Borel και τα σύνολα στα οποία δουλεύουμε είναι συμπαγή (θα εννοείται στη συνέχεια χωρίς περαιτέρω αναφορά). Στη γενική περίπτωση μπορούμε να οδηγηθούμε μέσω επιχειρημάτων προσέγγισης.

Ορισμός 3.2.1. Η *απόσταση Hamming* στον X με βάρος $a = (a_1, \ldots, a_n) \in [0, \infty)^n \setminus \{\mathbf{0}_n\}$ ορίζεται ως $d_a : X \times X \to [0, \infty)$,

$$d_a(x, y) = \sum_{i=1}^n a_i \mathbf{1}_{\{x_i \neq y_i\}}.$$

Παρατήρηση 3.2.2. Η d_a είναι μετρική στον X για κάθε $a \in [0, \infty)^n \setminus \{\mathbf{0}_n\}$.

Ορισμός 3.2.3. Για $A \in \mathcal{A}$, $A \neq \emptyset$, $x \in A$ και $a \in [0, \infty)^n$, $a \neq \mathbf{0}_n$, ορίζουμε τα εξής:

(i) $U_A(x) = \{(s_i)_{1 \leqslant i \leqslant n} \in \{0, 1\}^n :$ υπάρχει $y \in A$ ώστε αν $s_i = 0$ για κάποιο $1 \leqslant i \leqslant n$ τότε $x_i = y_i\}$.

(ii) $V_A(x) = \mathrm{conv}(U_A(x))$.

(iii) $f_c(A, x) = d_A(x) = \mathrm{dist}(0, V_A(x))$, ως προς την Ευκλείδεια απόσταση του $\mathbb{R}^n$.

(iv) $d_a(x, A) = \mathrm{dist}_{d_a}(x, A)$.

(v) $F_A(x) = \sup\limits_{|a|=1} d_a(x, A)$.

Παρατήρηση 3.2.4. (α) Ισχύει ότι

$$\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\} \subseteq U_A(x) \quad \text{και} \quad d_A(x) = \mathrm{dist}(0, \mathrm{conv}\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\}).$$

Πράγματι, αν $y \in A$ και $s = (\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n}$, τότε αν $s_i = 0$ έχουμε $x_i = y_i$. Οπότε, $s \in U_A(x)$.

Από τη σχέση υποσυνόλου παίρνουμε ότι

$$\mathrm{conv}\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\} \subseteq V_A(x),$$

άρα

$$d_A(x) = \mathrm{dist}(0, V_A(x)) \leqslant \mathrm{dist}(0, \mathrm{conv}\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\}).$$

Για την άλλη ανισότητα, έστω $\sum_{j=1}^k \lambda_j s^{(j)} \in V_A(x)$, όπου $\lambda_1, \ldots, \lambda_k \geqslant 0$ με $\sum_{j=1}^k \lambda_j = 1$ και $s^{(j)} \in U_A(x)$ για κάθε $1 \leqslant j \leqslant k$. Για κάθε $1 \leqslant j \leqslant k$ υπάρχει $y^{(j)} \in A$ ώστε αν $s_i^{(j)} = 0$ να έχουμε $y_i^{(j)} = x_i$, δηλαδή $\mathbf{1}_{\{y_i^{(j)} \neq x_i\}} = 0$. Οπότε, $\mathbf{1}_{\{y_i^{(j)} \neq x_i\}} \leqslant s_i^{(j)}$, και αυτό για κάθε $1 \leqslant i \leqslant n$ και $1 \leqslant j \leqslant k$. Συνεπώς,

$$\Big| \sum_{j=1}^k \lambda_j (\mathbf{1}_{\{y_i \neq x_i\}}) \Big| \leqslant \Big| \sum_{j=1}^k \lambda_j s^{(j)} \Big|,$$

αφού όλες οι συντεταγμένες είναι μη αρνητικές. Έπεται ότι

$$\mathrm{dist}(0, \mathrm{conv}\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\}) \leqslant \Big| \sum_{j=1}^{k} \lambda_j s^{(j)} \Big|,$$

και αυτό για το τυχόν σημείο του $V_A(x)$. Παίρνοντας infimum έχουμε το ζητούμενο.

(β) $V_A(x) \subseteq [0,1]^n \subseteq \sqrt{n} B_2^n = \{x \in \mathbb{R}^n : |x| \leqslant \sqrt{n}\}$.

(γ) $F_A(x) \leqslant \sqrt{n}$, αφού $d_a(x,y) \leqslant \sum_{i=1}^n a_i \leqslant \sqrt{n}$ για κάθε $x, y \in X$ και κάθε $a \in [0,\infty)^n$ με $|a| = 1$.

Λήμμα 3.2.5. *Έστω $a \in [0,\infty)^n$ με $|a| = 1$. Τότε,*

$$\inf_{y \in V_A(x)} \langle a, y \rangle = \inf_{s \in U_A(x)} \langle a, s \rangle = d_a(x, A).$$

Απόδειξη. Ισχύει προφανώς ότι

$$\inf_{y \in V_A(x)} \langle a, y \rangle \leqslant \inf_{s \in U_A(x)} \langle a, s \rangle.$$

Αν τώρα $y \in V_A(x)$, τότε $y = \sum_{i=1}^k \lambda_i s_i$ για κάποια $k \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_k \geqslant 0$ με $\lambda_1 + \cdots + \lambda_k = 1$ και $s_1, \ldots, s_k \in U_A(x)$. Οπότε,

$$\langle a, y \rangle = \sum_{i=1}^k \lambda_i \langle a, s_i \rangle \geqslant \sum_{i=1}^k \lambda_i \inf_{s \in U_A(x)} \langle a, s \rangle = \inf_{s \in U_A(x)} \langle a, s \rangle,$$

και άρα έπεται άμεσα η πρώτη ισότητα.

Για την απόδειξη της άλλης ισότητας, αρχικά παρατηρούμε ότι

$$\inf_{s \in U_A(x)} \langle a, s \rangle = \inf\{\langle a, s \rangle \mid s = (\mathbf{1}_{\{x_i \neq y_i\}})_{1 \leqslant i \leqslant n}, y \in A\}.$$

Πράγματι, αφού $\{(\mathbf{1}_{\{y_i \neq x_i\}})_{1 \leqslant i \leqslant n} \mid y \in A\} \subseteq U_A(x)$, έχουμε ότι

$$\inf_{s \in U_A(x)} \langle a, s \rangle \leqslant \inf\{\langle a, s \rangle \mid s = (\mathbf{1}_{\{x_i \neq y_i\}})_{1 \leqslant i \leqslant n}, y \in A\}.$$

Αντίστροφα, αν $\tilde{s} \in U_A(x)$, υπάρχει $\tilde{y} \in A$ τέτοιο ώστε $\mathbf{1}_{\{x_i \neq \tilde{y}_i\}} \leqslant \tilde{s}_i$ για κάθε $1 \leqslant i \leqslant n$. Οπότε,

$$\inf\{\langle a, s \rangle \mid s = (\mathbf{1}_{\{x_i \neq y_i\}})_{1 \leqslant i \leqslant n}\} \leqslant \sum_{i=1}^n a_i \mathbf{1}_{\{x_i \neq \tilde{y}_i\}} \leqslant \sum_{i=1}^n a_i \tilde{s}_i = \langle a, \tilde{s} \rangle,$$

για το τυχόν $\tilde{s} \in U_A(x)$. Παίρνοντας infimum έχουμε το ζητούμενο.

Άρα, αν $s = \big(\mathbf{1}_{\{x_i \neq y_i\}}\big)_{1 \leqslant i \leqslant n}$ για κάποιο $y \in A$, τότε

$$\langle a, s \rangle = \sum_{i=1}^n a_i s_i = \sum_{i=1}^n a_i \cdot \mathbf{1}_{\{x_i \neq y_i\}} = d_a(x, y).$$

Αντίστροφα, αν $y \in A$ τότε θεωρώντας το $s = \big(\mathbf{1}_{\{x_1 \neq y_1\}}, \ldots, \mathbf{1}_{\{x_n \neq y_n\}}\big)$ παίρνουμε ότι $\langle a, s \rangle = d_a(x,y)$, και προφανώς $s \in U_A(x)$. Οπότε,

$$\{\langle a, s \rangle : s \in U_A(x)\} = \{d_a(x, y) : y \in A\},$$

και επομένως ισχύει η δεύτερη ισότητα. □

Πρόταση 3.2.6. *Για κάθε $A \in \mathcal{A}$ ισχύει ότι $F_A = d_A$.*

Απόδειξη. Έστω ότι $d_A(x) < r$ για κάποιο $r > 0$. Τότε υπάρχει $z \in V_A(x)$ ώστε $|z| < r$. Έστω τώρα $a \in [0, \infty)^n$ με $|a| = 1$. Τότε έχουμε ότι

$$\inf_{y \in V_A(x)} \langle a, y \rangle \leqslant \langle a, z \rangle \leqslant |a| \cdot |z| < r.$$

Από το Λήμμα 3.2.5 έχουμε $\inf\limits_{y \in V_A(x)} \langle a, y \rangle = d_a(x, A)$, οπότε παίρνοντας supremum βλέπουμε ότι $F_A(x) \leqslant r$. Συνεπώς, για $r \to d_A(x)^+$ παίρνουμε ότι $F_A(x) \leqslant d_A(x)$.

Για την αντίστροφη ανισότητα υποθέτουμε ότι $d_A(x) > 0$ διότι αλλιώς δεν έχουμε κάτι να δείξουμε. Έστω $0 < \delta < 1$. Τότε υπάρχει $z \in V_A(x)$ με

$$0 < d_A(x)^2 \leqslant |z|^2 < d_A(x)^2 + \delta.$$

Από την κυρτότητα του $V_A(x)$ έχουμε ότι για κάθε $\vartheta \in (0, 1)$ και κάθε $y \in V_A(x)$ ισχύει ότι $\vartheta y + (1 - \vartheta) z \in V_A(x)$. Οπότε,

$$|z + \vartheta(y - z)|^2 = |\vartheta y + (1 - \vartheta) z|^2 \geqslant d_A(x)^2 > |z|^2 - \delta,$$

άρα

$$2\vartheta \langle y - z, z \rangle + \vartheta^2 |y - z|^2 > -\delta. \tag{3.2.1}$$

Θέτουμε λοιπόν $a = \frac{z}{|z|}$ και έχουμε ότι

$$\langle a, y \rangle = \frac{\langle z, y \rangle}{|z|} = \frac{\langle y - z, z \rangle}{|z|} + |z| > |z| - \frac{\vartheta |y - z|^2}{2|z|} - \frac{\delta}{2\vartheta |z|} \geqslant |z| - \frac{2\vartheta n}{|z|} - \frac{\delta}{2\vartheta |z|},$$

χρησιμοποιώντας την (3.2.1) και την Παρατήρηση 3.2.4 (β).

Οπότε, χρησιμοποιώντας πάλι το Λήμμα 3.2.5, έχουμε ότι

$$F_A(x) \geqslant d_a(x, A) = \inf_{y \in V_A(x)} \langle a, y \rangle > |z| - \frac{2\vartheta n}{|z|} - \frac{\delta}{2\vartheta |z|} \geqslant d_A(x) - \frac{2\vartheta n}{d_A(x)} - \frac{\delta}{2\vartheta d_A(x)}.$$

Παίρνοντας διαδοχικά $\delta \to 0^+$ και μετά $\vartheta \to 0^+$ έχουμε ότι $F_A(x) \geqslant d_A(x)$. □

Σημείωση 3.2.7. Για να ελαφρύνουμε το συμβολισμό θα γράφουμε F αντί για F_A όταν δεν υπάρχει σύγχυση ως προς το σύνολο στο οποίο αναφερόμαστε.

Πρόταση 3.2.8. *Για κάθε $\emptyset \neq A \in \mathcal{A}$ ισχύει ότι*

$$\mathrm{Ent}_P(e^{\lambda F^2}) \leqslant 4\lambda^2 \int_X F^2 e^{\lambda F^2} dP$$

για κάθε $\lambda \geqslant 0$.

Απόδειξη. Έστω $x = (x_1, \ldots, x_n) \in X$ και $\varepsilon > 0$. Τότε υπάρχει $a \in [0, \infty)^n$ με $|a| = 1$ ώστε $F(x) \leqslant d_a(x, A) + \varepsilon$.

Έστω τώρα $1 \leqslant i \leqslant n$ και $y_i \in \Omega_i$. Θέτουμε $y = (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)$. Τότε,

$$F_i(x_i) - F_i(y_i) = F(x) - F(y) \leqslant d_a(x, A) + \varepsilon - d_a(y, A) \leqslant d_a(x, y) + \varepsilon,$$

αφού η $d_a(\cdot, A) : X \to [0, \infty)$ είναι Lipschitz με σταθερά 1. Οπότε,

$$F_i(x_i) - F_i(y_i) \leqslant d_a(x, y) + \varepsilon = a_i \cdot \mathbf{1}_{\{x_i \neq y_i\}} + \varepsilon \leqslant a_i(x) + \varepsilon.$$

Ισχυρισμός. Αν $F_i(x_i) \geqslant F_i(y_i)$ και $\lambda \geqslant 0$, τότε

$$(\lambda F_i(x_i)^2 - \lambda F_i(y_i)^2)(e^{\lambda F_i(x_i)^2} - e^{\lambda F_i(y_i)^2}) \leqslant \lambda^2 (F_i(x_i)^2 - F_i(y_i)^2)^2 e^{\lambda F_i(x_i)^2}. \tag{3.2.2}$$

Πράγματι, αν $\lambda = 0$ ή $F_i(x_i) = F_i(y_i)$ τότε ο ισχυρισμός ισχύει τετριμμένα.

Έστω τώρα ότι $\lambda > 0$ και $F_i(x_i) > F_i(y_i) \geqslant 0$. Θεωρούμε την $g(y) = e^{\lambda y}$. Τότε υπάρχει $\xi \in [F_i(y_i)^2, F_i(x_i)^2]$ ώστε

$$\frac{g(F_i(x_i)^2) - g(F_i(y_i)^2)}{F_i(x_i)^2 - F_i(y_i)^2} = \lambda e^{\lambda \xi} \leqslant \lambda e^{\lambda F_i(x_i)^2},$$

και από αυτό έπεται το ζητούμενο.

Τώρα,

$$\begin{aligned} \lambda^2 (F_i(x_i)^2 - F_i(y_i)^2)^2 e^{\lambda F_i(x_i)^2} &= \lambda^2 (F_i(x_i) + F_i(y_i))^2 (F_i(x_i) - F_i(y_i))^2 e^{\lambda F_i(x_i)^2} \\ &\leqslant 4\lambda^2 F_i(x_i)^2 (a_i(x) + \varepsilon)^2 e^{\lambda F_i(x_i)^2}. \end{aligned} \tag{3.2.3}$$

Εφαρμόζοντας τώρα την Πρόταση 2.1.13 (i) για την $e^{\lambda F^2}$, σε συνδυασμό με τις (3.2.2) και (3.2.3) παίρνουμε ότι

$$\begin{aligned} \mathrm{Ent}_P(e^{\lambda F^2}) &\leqslant 4\lambda^2 \sum_{i=1}^n \int_X (a_i + \varepsilon)^2 F(x)^2 e^{\lambda F(x)^2} \, dP(x) \\ &\leqslant 4\lambda^2 \int_X (1 + 2\varepsilon\sqrt{n} + n\varepsilon^2) F^2 e^{\lambda F^2} \, dP, \end{aligned}$$

λαμβάνοντας υπ' όψιν ότι $|a|^2 = \sum_{i=1}^n a_i^2 = 1$ και ότι $\|a\|_1 \leqslant \sqrt{n}\,|a|$. Τέλος, αφού η F είναι φραγμένη, για $\varepsilon \to 0^+$, από το θεώρημα κυριαρχημένης σύγκλισης παίρνουμε το ζητούμενο. □

Πόρισμα 3.2.9. *Έστω $A \in \mathcal{A}$, $A \neq \emptyset$. Τότε,*

$$\mathrm{Var}_P(F^2) \leqslant 8 \int_X F^2 \, dP.$$

Απόδειξη. Για $\lambda > 0$ έχουμε ότι

$$\frac{\mathrm{Ent}_P(e^{\lambda F^2})}{\lambda^2} \leqslant 4 \int_X F^2 e^{\lambda F^2} \, dP.$$

Παίρνοντας όρια για $\lambda \to 0^+$ και εφαρμόζοντας δύο φορές τον κανόνα de L' Hospital στο αριστερό μέλος, παίρνουμε ότι

$$\frac{\mathrm{Var}_P(F^2)}{2} \leqslant 4 \int_X F^2 \, dP,$$

και έχουμε το ζητούμενο. □

Πρόταση 3.2.10. *Έστω $A \in \mathcal{A}$, $A \neq \emptyset$. Τότε,*

$$\int_X e^{\lambda F_A^2} dP \leqslant e^{\frac{\lambda M}{1-4\lambda}}$$

για κάθε $0 \leqslant \lambda < \frac{1}{4}$, όπου $M = \int_X F_A^2 \, dP$.

Απόδειξη. Θεωρούμε τη συνάρτηση

$$\Lambda(\lambda) = \int_X e^{\lambda F_A^2} \, dP, \quad \lambda \in \mathbb{R}$$

που είναι καλά ορισμένη, αφού η F_A είναι φραγμένη, και παραγωγίσιμη, με

$$\Lambda'(\lambda) = \int_X F_A^2 e^{\lambda F_A^2} \, dP.$$

Παρατηρούμε ότι, για $\lambda > 0$,

$$\begin{aligned} \lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda) &= \int_X \lambda F_A^2 e^{\lambda F_A^2} \, dP - \int_X e^{\lambda F_A^2} \, dP \cdot \log\Big(\int_X e^{\lambda F_A^2} \, dP\Big) \\ &= \mathrm{Ent}_P(e^{\lambda F_A^2}) \leqslant 4\lambda^2 \int_X F_A^2 e^{\lambda F_A^2} \, dP \\ &= 4\lambda^2 \Lambda'(\lambda), \end{aligned}$$

χρησιμοποιώντας την Πρόταση 3.2.8. Έχουμε λοιπόν ότι

$$\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda) \leqslant 4\lambda^2\Lambda'(\lambda) \tag{3.2.4}$$

για κάθε $\lambda > 0$.

Παρατηρούμε ότι $\Lambda(\lambda) \geqslant 1$ για κάθε $\lambda > 0$ και $\Lambda(\lambda) = 1$ αν και μόνο αν $F_A \equiv 0$.

(α) Αν $P(A) = 1$, τότε $F_A \equiv 0$ με πιθανότητα 1 και η ανισότητα ισχύει τετριμμένα.

(β) Αν $P(A) < 1$ τότε $P(F_A > 0) > 0$ αφού $F_A(x) > 0$ για κάθε $x \in A^c$. Σε αυτή την περίπτωση, $\Lambda(\lambda) > 1$ για κάθε $\lambda > 0$, και θεωρούμε την $K(\lambda) = \log\Lambda(\lambda)$.

Παρατηρούμε ότι

$$\lambda(1-4\lambda)K'(\lambda) = \lambda(1-4\lambda)\frac{\Lambda'(\lambda)}{\Lambda(\lambda)} = \frac{\lambda\Lambda'(\lambda) - 4\lambda^2\Lambda'(\lambda)}{\Lambda(\lambda)} \leqslant \log\Lambda(\lambda) = K(\lambda),$$

χρησιμοποιώντας την (3.2.4). Οπότε, έχουμε ότι $\lambda(1-4\lambda)K'(\lambda) \leqslant K(\lambda)$, άρα

$$\frac{K'(\lambda)}{K(\lambda)} \leqslant \frac{1}{\lambda(1-4\lambda)}$$

για κάθε $\lambda \in \big(0, \frac{1}{4}\big)$. Ισοδύναμα,

$$(\log K)'(\lambda) \leqslant \frac{1}{\lambda(1-4\lambda)}$$

για κάθε $\lambda \in \big(0, \frac{1}{4}\big)$. Συνεπώς, αν $0 < \delta \leqslant \lambda < \frac{1}{4}$, έχουμε ότι

$$\log K(\lambda) - \log K(\delta) = \int_\delta^\lambda (\log K)'(u) \, du \leqslant \int_\delta^\lambda \frac{1}{u(1-4u)} \, du = \log\Big(\frac{\lambda}{1-4\lambda}\,\frac{1-4\delta}{\delta}\Big),$$

δηλαδή,

$$\log K(\lambda) - \log K(\delta) \leqslant \log\Big(\frac{\lambda}{1-4\lambda}\,\frac{1-4\delta}{\delta}\Big). \tag{3.2.5}$$

Παρατηρούμε ότι

$$\lim_{\delta\to 0^+}\frac{K(\delta)}{\delta} = \lim_{\delta\to 0^+} K'(\delta) = \lim_{\delta\to 0^+}\frac{\int_X F_A^2 e^{\delta F_A^2}\,dP}{\int_X e^{\delta F_A^2}\,dP} = \int_X F_A^2\,dP = M > 0,$$

οπότε

$$\lim_{\delta\to 0^+}\frac{K(\delta)}{M\delta} = 1.$$

Γράφουμε

$$\log K(\lambda) - \log K(\delta) = \log K(\lambda) - \log\frac{K(\delta)}{M\delta} - \log M - \log\delta.$$

Τότε, η (3.2.5) μας δίνει

$$\log K(\lambda) - \log\frac{K(\delta)}{M\delta} - \log M \leqslant \log\Big(\frac{\lambda}{1-4\lambda}\cdot(1-4\delta)\Big),$$

κα για $\delta \to 0^+$,

$$\log K(\lambda) - \log M \leqslant \log\Big(\frac{\lambda}{1-4\lambda}\Big),$$

ή ισοδύναμα $K(\lambda) \leqslant \frac{M\lambda}{1-\lambda}$, που μας δίνει ότι

$$\int_X e^{\lambda F_A^2} dP \leqslant e^{\frac{\lambda M}{1-4\lambda}}$$

για κάθε $0 < \lambda < \frac{1}{4}$. Μάλιστα η ανισότητα προφανώς ισχύει και για $\lambda = 0$.

□

Πόρισμα 3.2.11. *Έστω $A \in \mathcal{A}$, $A \neq \emptyset$. Τότε,*

$$\int_X e^{\frac{1}{14}F_A^2} dP \leqslant e^{\frac{4}{5P(A)}}.$$

Απόδειξη. Αν $P(A) = 1$, τότε $F_A = 0$ με πιθανότητα 1 και η ανισότητα ισχύει.

Αν $P(A) < 1$, από την Πρόταση 3.2.10 με $\lambda = \frac{1}{14}$ έχουμε ότι

$$\int_X e^{\frac{1}{14}F_A^2} dP \leqslant e^{\frac{M}{10}},$$

όπου $M = \int_X F_A^2\,dP$. Τώρα, από το Πόρισμα 3.2.9,

$$8M \geqslant \mathrm{Var}_P(F_A^2) = \int_X (F_A^2 - M)^2\,dP \geqslant \int_A (F_A^2 - M)^2\,dP = \int_A M^2\,dP = M^2 P(A),$$

αφού $F_A \cdot \mathbf{1}_A \equiv 0$. Συνεπώς, $M^2 P(A) \leqslant 8M$ άρα

$$M \leqslant \frac{8}{P(A)}.$$

Οπότε τελικά,

$$\int_X e^{\frac{1}{14}F_A^2} dP \leqslant e^{\frac{4}{5P(A)}}.$$

□

Παρατήρηση 3.2.12. Αυτά τα πρώτα συμπεράσματα είναι ασθενέστερα από την ανισότητα απόστασης από την κυρτή θήκη του Talagrand, και ως προς την αριθμητική σταθερά και ως προς την εξάρτηση από το A, ειδικά για ενδεχόμενα A με μικρή πιθανότητα πραγματοποίησης.

Για να φτάσουμε σε καλύτερα συμπεράσματα είναι απαραίτητο να αναπτύξουμε στη συνέχεια την προηγούμενη ανάλυση στις αρνητικές τιμές του λ.

Λήμμα 3.2.13. *Έστω $A \in \mathcal{A}$, $A \neq \emptyset$. Γράφουμε $F := F_A$. Θεωρούμε $x = (x_1, \ldots, x_n) \in \Omega$ και $y_i \in \Omega$, $1 \leqslant i \leqslant n$. Αν*

$$F_i(y_i) = F(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n) = F(y),$$

τότε

$$F_i(x_i)^2 - F_i(y_i)^2 = F(x)^2 - F(y)^2 \leqslant 1.$$

Απόδειξη. Υπενθυμίζουμε πρώτα ότι

$$\begin{aligned} F(x)^2 = d_A(x)^2 &= \inf\{|w|^2 : w \in \mathrm{conv}(U_A(x))\} \\ &= \inf\{|w|^2 : w \in \mathrm{conv}((\mathbf{1}_{\{x_j \neq z_j\}})_{1 \leqslant j \leqslant n} : z \in A)\} = \inf \Big| \sum \vartheta_z s(z) \Big|^2, \end{aligned}$$

όπου το τελευταίο infimum είναι πάνω από όλα τα πεπερασμένα αθροίσματα $\sum \vartheta_z s(z)$ με $\vartheta_z \geqslant 0$, $\sum \vartheta_z = 1$ και $s(z) = (\mathbf{1}_{\{x_j \neq z_j\}})_{1 \leqslant j \leqslant n}$, $z \in A$.

Έστω $\delta > 0$. Τότε μπορούμε να βρούμε πεπερασμένο υποσύνολο $J \subseteq A$ και $\vartheta_z \geqslant 0$, $z \in J$ με $\sum_{z \in J} \vartheta_z = 1$ ώστε αν $\hat{s}(z) = (\mathbf{1}_{\{y_i \neq z_i\}})_{1 \leqslant i \leqslant n}$, όπου $z = (z_1, \ldots, z_n)$, να ισχύει

$$\Big| \sum_{z \in J} \vartheta_z \hat{s}(z) \Big|^2 \leqslant d_A(y)^2 + \delta.$$

Για κάθε $z \in J$ θεωρούμε το $s(z) = (\mathbf{1}_{\{x_i \neq z_i\}})_{1 \leqslant i \leqslant n}$. Τότε,

$$d_A(x)^2 - d_A(y)^2 - \delta \leqslant \Big| \sum_{z \in J} \vartheta_z s(z) \Big|^2 - \Big| \sum_{z \in J} \vartheta_z \hat{s}(z) \Big|^2.$$

Τα $s(z), \hat{s}(z)$ διαφέρουν το πολύ στην i-συντεταγμένη, αφού $s_i(z) = \mathbf{1}_{\{x_i \neq z_i\}}$ και $\hat{s}_i(z) = \mathbf{1}_{\{y_i \neq z_i\}}$. Οπότε,

$$\begin{aligned} d_A(x)^2 - d_A(y)^2 - \delta &\leqslant \Big(\sum_{z \in J} \vartheta_z \mathbf{1}_{\{x_i \neq z_i\}} \Big)^2 - \Big(\sum_{z \in J} \vartheta_z \mathbf{1}_{\{y_i \neq z_i\}} \Big)^2 \\ &\leqslant \Big(\sum_{z \in J} \vartheta_z \cdot 1 \Big)^2 - \Big(\sum_{z \in J} \vartheta_z \cdot 0 \Big)^2 = 1. \end{aligned}$$

Αφού το $\delta > 0$ ήταν τυχόν, έπεται ότι $d_A(x)^2 - d_A(y)^2 \leqslant 1$. □

Πρόταση 3.2.14. *Για κάθε $\emptyset \neq A \in \mathcal{A}$ θέτουμε $F := F_A$. Τότε, ισχύει ότι*

$$\mathrm{Ent}_P(e^{-\lambda F^2}) \leqslant 4\lambda^2 e^{\lambda} \int_X F^2 e^{-\lambda F^2} dP$$

για κάθε $\lambda \geqslant 0$.

Απόδειξη. Έστω $x = (x_1, \ldots, x_n) \in X$ και $\varepsilon > 0$. Τότε, υπάρχει $a \in [0, \infty)^n$ με $|a| = 1$ ώστε

$$F(x) \leqslant d_a(x, A) + \varepsilon.$$

Έστω τώρα $1 \leqslant i \leqslant n$ και $y_i \in \Omega_i$. Θέτουμε $y = (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)$. Τότε,

$$\begin{aligned} F_i(x_i) - F_i(y_i) = F(x) - F(y) &\leqslant d_a(x, A) + \varepsilon - d_a(y, A) \leqslant d_a(x, y) + \varepsilon = a_i \cdot \mathbf{1}_{\{x_i \neq y_i\}} + \varepsilon \\ &\leqslant a_i + \varepsilon, \end{aligned}$$

δηλαδή

$$F_i(x_i) - F_i(y_i) \leqslant a_i + \varepsilon. \tag{3.2.6}$$

Ισχυρισμός. Αν $F_i(x_i) \geqslant F_i(y_i)$ και $\lambda \geqslant 0$ τότε ισχύει ότι

$$[-\lambda F_i(x_i)^2 + \lambda F_i(y_i)^2]\,[e^{-\lambda F_i(x_i)^2} - e^{-\lambda F_i(y_i)^2}] \leqslant \lambda^2 [F_i(x_i)^2 - F_i(y_i)^2]^2 e^{-\lambda F_i(y_i)^2}.$$

Πράγματι, υποθέτουμε ότι $F_i(x_i) > F_i(y_i)$ και ότι $\lambda > 0$ διότι αλλιώς ο ισχυρισμός προφανώς ισχύει ως ισότητα. Θεωρούμε την $g(y) = e^{-\lambda y}$. Τότε, από το θεώρημα μέσης τιμής υπάρχει $\xi \in (F_i(y_i)^2, F_i(x_i)^2)$ ώστε

$$\frac{e^{-\lambda F_i(x_i)^2} - e^{-\lambda F_i(y_i)^2}}{F_i(x_i)^2 - F_i(y_i)^2} = -\lambda e^{-\lambda \xi} \geqslant -\lambda e^{-\lambda F_i(y_i)^2}.$$

Πολλαπλασιάζοντας με $(-\lambda)$ την τελευταία ανισότητα παίρνουμε εύκολα το ζητούμενο.

Οπότε τώρα,

$$\begin{aligned} [-\lambda F_i(x_i)^2 + \lambda F_i(y_i)^2]\,[e^{-\lambda F_i(x_i)^2} - e^{-\lambda F_i(y_i)^2}] &\leqslant \lambda^2 [F_i(x_i)^2 - F_i(y_i)^2]^2 e^{-\lambda F_i(y_i)^2} \\ &\leqslant \lambda^2 [F_i(x_i)^2 - F_i(y_i)^2]^2 e^{-\lambda (F_i(x_i)^2 - 1)} \\ &\leqslant 4\lambda^2 e^{\lambda} (a_i(x) + \varepsilon)^2 F_i(x_i)^2 e^{-\lambda F_i(x_i)^2}, \end{aligned}$$

όπου χρησιμοποιήσαμε το Λήμμα 3.2.13, την (3.2.6) και το ότι $F_i(y_i) \leqslant F_i(x_i)$.

Τώρα, όπως ακριβώς στην περίπτωση όπου $\lambda > 0$ (δείτε την απόδειξη της Πρότασης 3.2.8) καταλήγουμε στο ότι

$$\mathrm{Ent}_P(e^{-\lambda F^2}) \leqslant 4\lambda^2 e^{\lambda} \int_X F^2 e^{-\lambda F^2} dP$$

εφαρμόζοντας την Πρόταση 2.1.13 και παίρνοντας όριο καθώς $\varepsilon \to 0^+$. □

Πρόταση 3.2.15. *Για κάθε $\emptyset \neq A \in \mathcal{A}$ ισχύει ότι*

$$\log\Big(\int_X e^{-\lambda F_A^2} dP\Big) \leqslant -\frac{\lambda M}{1 + 8\lambda}$$

για κάθε $0 \leqslant \lambda \leqslant \frac{1}{2}$, όπου $M = \int_X F_A^2 \, dP$.

Απόδειξη. Αν $P(A) = 1$ τότε $F_A = 0$ με πιθανότητα 1 και $M = 0$. Οπότε, η ανισότητα ισχύει ως ισότητα.

Έστω ότι $P(A) < 1$. Τότε, $P(F_A > 0) > 0$. Θεωρούμε την

$$\Lambda(\lambda) = \int_X e^{-\lambda F_A^2} dP, \quad \lambda \in \mathbb{R}.$$

Έχουμε

$$\Lambda'(\lambda) = -\int_X F_A^2 e^{-\lambda F_A^2} dP.$$

Παρατηρούμε ότι, από την Πρόταση 3.2.14,

$$\begin{aligned}\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda) &= -\int_X \lambda F_A^2 e^{-\lambda F_A^2} dP - \int_X e^{-\lambda F_A^2} dP \cdot \log\Big(\int_X e^{-\lambda F_A^2} dP\Big) \\ &= \mathrm{Ent}_P(e^{-\lambda F_A^2}) \leqslant 4\lambda^2 e^{\lambda} \int_X F_A^2 e^{-\lambda F_A^2} dP.\end{aligned}$$

Οπότε, για $0 \leqslant \lambda \leqslant 1/2$ παίρνουμε ότι

$$\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda) \leqslant 8\lambda^2 \int_X F_A^2 e^{-\lambda F_A^2} dP = -8\lambda^2\Lambda'(\lambda).$$

Θεωρούμε την $K(\lambda) = -\log\Lambda(\lambda) \geqslant 0$, $\lambda \in [0, 1/2]$, καθώς έχουμε ότι $0 < \Lambda(\lambda) < 1$ για κάθε $\lambda \in (0, 1/2]$ και $K(0) = 0$.

Τότε,

$$\lambda(1+8\lambda)K'(\lambda) = -\frac{\lambda\Lambda'(\lambda) + 8\lambda^2\Lambda'(\lambda)}{\Lambda(\lambda)} \geqslant -\log\Lambda(\lambda) = K(\lambda),$$

δηλαδή

$$\lambda(1+8\lambda)K'(\lambda) \geqslant K(\lambda),$$

ή ισοδύναμα,

$$\frac{K'(\lambda)}{K(\lambda)} \geqslant \frac{1}{\lambda(1+8\lambda)}$$

για κάθε $\lambda \in (0, 1/2]$, δηλαδή

$$(\log K)'(\lambda) \geqslant \frac{1}{\lambda(1+8\lambda)}$$

για κάθε $\lambda \in (0, 1/2]$. Τώρα, αν $0 < \delta \leqslant \lambda \leqslant 1/2$ έχουμε ότι

$$\begin{aligned}\log K(\lambda) - \log K(\delta) &= \int_\delta^\lambda (\log K)'(u)\, du \\ &\geqslant \int_\delta^\lambda \frac{1}{u(1+8u)}\, du \\ &= \int_\delta^\lambda \Big(\frac{1}{u} - \frac{8}{1+8u}\Big)\, du \\ &= \log\Big(\frac{\lambda(1+8\delta)}{\delta(1+8\lambda)}\Big) = \log\Big(\frac{\lambda}{1+8\lambda}\Big) - \log\delta + \log(1+8\delta).\end{aligned}$$

Παρατηρούμε ότι

$$\lim_{\delta\to 0^+}\frac{K(\delta)}{\delta}=\lim_{\delta\to 0^+}K'(\delta)=\lim_{\delta\to 0^+}\frac{\int_X F_A^2 e^{-\delta F_A^2}dP}{\int_X e^{-\delta F_A^2}dP}=\int_X F_A^2\,dP=M.$$

Όπότε,

$$\log(\lambda(1+8\lambda)^{-1})+\log(1+8\delta)\leqslant \log K(\lambda)-\log\frac{K(\delta)}{M\delta}-\log M,$$

και καθώς $\delta\to 0^+$ έχουμε

$$\log K(\lambda)-\log M\geqslant \log(\lambda(1+8\lambda)^{-1})$$

ή ισοδύναμα

$$K(\lambda)\geqslant\frac{M\lambda}{1+8\lambda}.$$

Έπεται ότι

$$\log\Big(\int_X e^{-\lambda F_A^2}dP\Big)\leqslant -\frac{M\lambda}{1+8\lambda}$$

για κάθε $\lambda\in[0,1/2]$. □

Θεώρημα 3.2.16 (ανισότητα απόστασης από την κυρτή θήκη)**.** *Για κάθε $\emptyset\neq A\in\mathcal{A}$ ισχύει ότι*

$$\int_X e^{\frac{1}{14}F_A^2}dP\leqslant\frac{1}{P(A)}.$$

Απόδειξη. Για $\lambda=\frac{1}{2}$ η Πρόταση 3.2.15 μας δίνει ότι

$$-\frac{M}{10}\geqslant\log\Big(\int_X e^{-\frac{F_A^2}{2}}dP\Big)\geqslant\log\Big(\int_A e^{-\frac{F_A^2}{2}}dP\Big)=\log\Big(\int_A \mathbf{1}\,dP\Big)=\log P(A)$$

αφού $F_A\cdot\mathbf{1}_A\equiv 0$, άρα

$$\frac{M}{10}\leqslant\log\Big(\frac{1}{P(A)}\Big).$$

Εφαρμόζοντας την Πρόταση 3.2.10 με $\lambda=\frac{1}{14}$ έχουμε ότι

$$\int_X e^{\frac{1}{14}F_A^2}dP\leqslant e^{\frac{M}{10}}\leqslant\frac{1}{P(A)}.$$

□

Παρατήρηση 3.2.17. Δείξαμε επομένως την ανισότητα απόστασης από την κυρτή θήκη του Talagrand, αλλά με διαφορετική σταθερά. Δεν είναι ξεκάθαρο ωστόσο ούτε αν η σταθερά $\frac{1}{4}$ μπορεί να επιτευχθεί με αυτή τη μέθοδο ούτε και αν αυτή είναι η βέλτιστη σταθερά.

3.3 Η αρχική απόδειξη του Talagrand

3.3.1 Ένα ισοπεριμετρικό θεώρημα για τον κύβο

Θεωρούμε το σύνολο $E_2^n = \{-1,1\}^n$, το οποίο ταυτίζουμε με το σύνολο των κορυφών του κύβου $Q_n = [-1,1]^n$ στον $\mathbb{R}^n$. Στο E_2^n ορίζουμε το ομοιόμορφο μέτρο πιθανότητας μ_n που δίνει μάζα 2^{-n} σε κάθε σημείο. Για κάθε μη κενό $A \subseteq E_2^n$ θέτουμε

$$\varphi_A(x) = \inf\{|x-y| : y \in \mathrm{conv}(A)\}.$$

Το βασικό θεώρημα αυτής της παραγράφου είναι το εξής.

Θεώρημα 3.3.1. *Για κάθε $A \subseteq E_2^n$, $A \neq \emptyset$,*

$$\int_{E_2^n} \exp(\varphi_A^2/8)\, d\mu_n \leqslant \frac{1}{\mu_n(A)}.$$

Απόδειξη. Με επαγωγή ως προς το πλήθος των σημείων του A. Αν $\mathrm{card}(A) = 1$ δηλαδή $A = \{y\}$, τότε

$$\varphi_A(x) = \inf\{|x-z| : z \in \mathrm{conv}(A) = \{y\}\} = |x-y|.$$

Άρα,

$$\int_{E_2^n} \exp(\varphi_A^2/8)\, d\mu_n = \int_{E_2^n} e^{|x-y|^2/8}\, d\mu_n = \frac{1}{2^n} \sum_{x \in E_2^n} e^{|x-y|^2/8}.$$

Κάθε $x \in E_2^n$ διαφέρει από το y σε i θέσεις, $i = 0, 1, \ldots, n$. Το πλήθος των $x \in E_2^n$ που διαφέρουν σε i θέσεις από το y είναι ίσο με $\binom{n}{i}$. Παρατηρούμε ότι $|x-y|^2 = 4i$ όταν το x διαφέρει απο το y σε i θέσεις. Άρα,

$$\begin{aligned}\int_{E_2^n} e^{|x-y|^2/8}\, d\mu_n &= \frac{1}{2^n} \sum_{i=0}^n \binom{n}{i} e^{i/2} = \frac{1}{2^n}\Big(1+e^{1/2}\Big)^n = \Big(\frac{1+e^{1/2}}{2}\Big)^n \\ &\leqslant 2^n = \frac{1}{\mu_n(A)},\end{aligned}$$

αφού $e^{1/2} \leqslant e \leqslant 3$.

Έστω τώρα ότι $\mathrm{card}(A) \geqslant 2$. Εξετάζουμε πρώτα την περίπτωση $n = 1$. Αναγκαστικά έχουμε $A = E_1$, επομένως $\varphi_A(x) = 0$ για κάθε $x \in E_1$. Άρα,

$$\int_{E_2^1} \exp(\varphi_A^2/8)\, d\mu_1 = \int_{E_2^1} e^0\, d\mu_1 = 1 = 1/\mu_1(A).$$

Για το επαγωγικό βήμα θεωρούμε $A \subseteq E_{n+1}$ με $\mathrm{card}(A) \geqslant 2$. Μπορούμε να υποθέσουμε ότι

$$A = (A_1 \times \{1\}) \cup (A_{-1} \times \{-1\})$$

όπου $A_1, A_{-1} \neq \emptyset$. Πράγματι, αν για παράδειγμα $A_{-1} = \emptyset$ τότε από το λήμμα που ακολουθεί και την επαγωγική υπόθεση έχουμε

$$\begin{aligned}\int_{E_2^{n+1}} e^{\varphi_A^2/8}\, d\mu_{n+1} &\leqslant \int_{E_2^n} e^{\varphi_{A_1}^2/8}\, d\mu_n \leqslant \frac{1}{\mu_n(A_1)} = \frac{2^n}{\mathrm{card}(A_1)} = \frac{2^n}{\mathrm{card}(A)} \\ &\leqslant \frac{2^{n+1}}{\mathrm{card}(A)} = \frac{1}{\mu_{n+1}(A)}.\end{aligned}$$

Μπορούμε επίσης, χωρίς περιορισμό της γενικότητας, να υποθέσουμε ότι $\mathrm{card}A_{-1} \leqslant \mathrm{card}A_1$.

Λήμμα 3.3.2. *Για κάθε* $x \in E_2^n$ *ισχύει ότι*

$$\varphi_A((x,1)) \leqslant \varphi_{A_1}(x).$$

Απόδειξη. Αρκεί να δείξουμε ότι

$$\{|x-y| : y \in \mathrm{conv}(A_1)\} \subseteq \{|(x,1)-z| : z \in \mathrm{conv}(A)\}.$$

Έστω $y \in \mathrm{conv}A_1$. Τότε, $y = \sum_{i=1}^m t_i x_i$ όπου $t_i \geqslant 0$ με $\sum_{i=1}^m t_i = 1$ και $x_i \in A_1$. Τότε όμως $(x_i, 1) \in A$ και

$$\sum_{i=1}^m t_i(x_i,1) = \Big(\sum_{i=1}^m t_i x_i, \sum_{i=1}^m t_i\Big) = (y,1),$$

δηλαδή $(y,1) \in \mathrm{conv}A$. Αφού $|x-y| = |(x,1)-(y,1)|$ και

$$|(x,1)-(y,1)| \in \{|(x,1)-z| : z \in \mathrm{conv}(A)\},$$

έχουμε το ζητούμενο. □

Λήμμα 3.3.3. *Για κάθε* $x \in E_2^n$ *και κάθε* $0 \leqslant a \leqslant 1$,

$$\varphi_A^2((x,-1)) \leqslant 4a^2 + a\varphi_{A_1}^2(x) + (1-a)\varphi_{A_{-1}}^2(x).$$

Απόδειξη. Έστω $z_i \in \mathrm{conv}A_i$ $(i = 1, -1)$. Τότε, όπως προηγουμένως, $(z_i, i) \in \mathrm{conv}A$. Το $\mathrm{conv}A$ είναι κυρτό, άρα

$$z := a(z_1,1) + (1-a)(z_{-1},-1) = (az_1 + (1-a)z_{-1}, 2a-1) \in \mathrm{conv}A.$$

Έχουμε

$$\begin{aligned}|(x,-1)-z|^2 &= |(x - az_1 - (1-a)z_{-1}, -2a)|^2 = |(x - az_1 - (1-a)z_{-1}, 0)|^2 + |(0,-2a)|^2\\ &\leqslant (a|x-z_1| + (1-a)|x-z_{-1}|)^2 + 4a^2 \leqslant a|x-z_1|^2 + (1-a)|x-z_{-1}|^2 + 4a^2,\end{aligned}$$

όπου στην τελευταία ανισότητα χρησιμοποιήθηκε η κυρτότητα της $x \mapsto x^2$. Αφού τα $z_i \in \mathrm{conv}A_i$ ήταν τυχόντα, έπεται ότι

$$\varphi_A^2(x,-1) \leqslant a\varphi_{A_1}^2(x) + (1-a)\varphi_{A_{-1}}^2(x) + 4a^2.$$

□

Χρησιμοποιώντας τα δύο λήμματα, γράφουμε

$$\begin{aligned}\int_{E_2^{n+1}} e^{\varphi_A^2/8}\, d\mu_{n+1} &= \frac{1}{2^{n+1}} \sum_{x\in E_2^{n+1}} e^{\varphi_A^2(x)/8}\\ &= \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\varphi_A^2((x,1))/8} + \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\varphi_A^2((x,-1))/8}\\ &\leqslant \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\varphi_{A_1}^2(x)/8} + \frac{1}{2^{n+1}} e^{a^2/2} \sum_{x\in E_2^n} e^{a\varphi_{A_1}^2(x)/8 + (1-a)\varphi_{A_{-1}}^2(x)/8}\\ &\leqslant \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\varphi_{A_1}^2(x)/8} + \frac{1}{2^{n+1}} e^{a^2/2} \Big(\sum_{x\in E_2^n} e^{\varphi_{A_1}^2(x)/8}\Big)^a \Big(\sum_{x\in E_2^n} e^{\varphi_{A_{-1}}^2(x)/8}\Big)^{1-a}\\ &= \frac{1}{2}\int_{E_2^n} e^{\varphi_{A_1}^2/8}\, d\mu_n + \frac{1}{2} e^{a^2/2} \Big(\int_{E_2^n} e^{\varphi_{A_1}^2/8}\, d\mu_n\Big)^a \Big(\int_{E_2^n} e^{\varphi_{A_{-1}}^2/8}\, d\mu_n\Big)^{1-a}.\end{aligned}$$

Θέτουμε

$$u_1 = \int_{E_2^n} e^{\varphi_{A_1}^2/8}\, d\mu_n \;\;,\;\; v_1 = \frac{1}{\mu_n(A_1)}$$

και

$$u_{-1} = \int_{E_2^n} e^{\varphi_{A_{-1}}^2/8}\, d\mu_n \;\;,\;\; v_{-1} = \frac{1}{\mu_n(A_{-1})}.$$

Από την επαγωγική υπόθεση έχουμε $u_1 \leqslant v_1$ και $u_{-1} \leqslant v_{-1}$. (επίσης, η $\text{card}A_{-1} \leqslant \text{card}A_1$ γράφεται $v_1 \leqslant v_{-1}$). Άρα η προηγούμενη ανισότητα παίρνει τη μορφή

$$\begin{aligned}\int_{E_2^{n+1}} e^{\varphi_A^2/8}\, d\mu_{n+1} &\leqslant \frac{1}{2}u_1 + \frac{1}{2}e^{a^2/2}(u_1)^a(u_{-1})^{1-a} \leqslant \frac{1}{2}u_1 + \frac{1}{2}e^{a^2/2}(v_1)^a(v_{-1})^{1-a}\\ &\leqslant \frac{v_1}{2}(1 + e^{a^2/2}(v_1/v_{-1})^{a-1}).\end{aligned}$$

Η τελευταία ποσότητα γίνεται ελάχιστη για $a = -\log(v_1/v_{-1})$. Η τιμή $-\log(v_1/v_{-1})$ είναι περίπου ίση με $1 - v_1/v_{-1}$. Επιλέγουμε $a_0 = 1 - v_1/v_{-1}$. Αφού $v_1 \leqslant v_{-1}$ έχουμε $0 \leqslant a_0 \leqslant 1$, οπότε μπορούμε να γράψουμε

$$\int_{E_2^{n+1}} e^{\varphi_A^2/8}\, d\mu_{n+1} \leqslant \frac{v_1}{2}(1 + e^{a_0^2/2}(1-a_0)^{a_0-1}).$$

Λήμμα 3.3.4. *Για κάθε $0 \leqslant a \leqslant 1$ έχουμε*

$$1 + e^{a^2/2}(1-a)^{a-1} \leqslant \frac{4}{2-a}.$$

Απόδειξη. Απλές πράξεις δείχνουν ότι η ανισότητα που ζητάμε είναι ισοδύναμη με την

$$g(a) = \log(2+a) - \log(2-a) - a^2/2 - (a-1)\log(1-a) \geqslant 0.$$

Παραγωγίζοντας βλέπουμε ότι $g'' \geqslant 0$ και $g'(0) = 0$. Άρα η g είναι αύξουσα στο $[0,1]$. Αφού $g(0) = 0$, έπεται το ζητούμενο. □

Χρησιμοποιώντας το Λήμμα 3.3.4 έχουμε

$$\begin{aligned}\int_{E_2^{n+1}} e^{\varphi_A^2/8}\, d\mu_{n+1} &\leqslant \frac{v_1}{2}\frac{4}{2-a_0} = \frac{2v_1}{1 + v_1/v_{-1}} = \frac{2}{1/v_1 + 1/v_{-1}} = \frac{2}{\mu_n(A_1) + \mu_n(A_{-1})}\\ &= \frac{1}{\mu_{n+1}(A)}.\end{aligned}$$

Η τελευταία ισότητα είναι φανερή αφού $\mu_{n+1}(A_i \times \{i\}) = \mu_n(A_i)/2$, $i = \pm 1$. Έτσι ολοκληρώνονται το επαγωγικό βήμα και η απόδειξη του Θεωρήματος 3.3.1. □

Πόρισμα 3.3.5. *Για κάθε $t > 0$ έχουμε ότι*

$$\mu_n(\varphi_A \geqslant t) \leqslant \frac{1}{\mu_n(A)}e^{-t^2/8}.$$

Απόδειξη. Έχουμε ότι

$$\begin{aligned}e^{t^2/8}\mu_n(\varphi_A \geqslant t) &= e^{t^2/8}\int_{\{\varphi_A \geqslant t\}} d\mu_n = \int_{\{\varphi_A \geqslant t\}} e^{t^2/8} d\mu_n \leqslant \int_{\{\varphi_A \geqslant t\}} e^{\varphi_A^2/8}\, d\mu_n\\ &\leqslant \int_{E_2^n} e^{\varphi_A^2/8}\, d\mu_n \leqslant \frac{1}{\mu_n(A)},\end{aligned}$$

χρησιμοποιώντας το Θεώρημα 3.3.1. □

3.3.2 Επέκταση σε γινόμενα πεπερασμένων υποσυνόλων χώρων με νόρμα

Έστω $(X_i, \|\cdot\|_i)$, $1 \leqslant i \leqslant n$ χώροι με νόρμα. Θεωρούμε τώρα τον χώρο γινόμενο $X^{(n)} = X_1 \times \cdots \times X_n$ εφοδιασμένο με τη νόρμα $\|\cdot\| = \|\cdot\|_{X^{(n)}} : X^{(n)} \to [0,\infty)$,

$$\|x\| = \|(x_1,\dots,x_n)\| = \Big(\sum_{i=1}^n \|x_i\|_i^2\Big)^{1/2}.$$

Για κάθε $1 \leqslant i \leqslant n$ θεωρούμε ένα πεπερασμένο $\Omega_i \subseteq X_i$ με διάμετρο $\mathrm{diam}(\Omega_i) \leqslant 1$. Έστω P_i μέτρο πιθανότητας στο Ω_i. θέτουμε

$$\Omega = \Omega_1 \times \Omega_2 \ldots \times \Omega_n$$

και

$$P = P^n = P_1 \otimes P_2 \otimes \ldots \otimes P_n$$

το μέτρο γινόμενο στο Ω. Για κάθε $A \subseteq \Omega$ ορίζουμε

$$\varphi_A(t) = \mathrm{dist}(t, \mathrm{conv}(A)),$$

την απόσταση του t από την κυρτή θήκη $\mathrm{conv}(A)$ του A στον $X^{(n)}$.

Θεώρημα 3.3.6. *Για κάθε $A \subseteq \Omega$, $A \neq \emptyset$,*

$$\int_\Omega e^{\varphi_A^2}/4\, dP \leqslant \frac{1}{P(A)}.$$

Απόδειξη. Με επαγωγή ως προς n.

$n=1$: Γράφουμε

$$\begin{aligned}\int_\Omega e^{\varphi_A^2/4}\, dP &= \int_\Omega e^{\varphi_A^2(t)/4}\, dP(t) = \int_A e^{\varphi_A^2(t)/4}\, dP(t) + \int_{\Omega\setminus A} e^{\varphi_A^2(t)/4}\, dP(t) \\ &= P(A) + \int_{\Omega\setminus A} e^{\varphi_A^2(t)/4}\, dP(t).\end{aligned}$$

Η τελευταία ισότητα ισχύει αφού $\varphi_A(t) = 0$ για κάθε $t \in A$. Επίσης, για κάθε $t \in \Omega \setminus A$ ισχύει $\varphi_A(t) \leqslant 1$. Πράγματι,

$$\varphi_A(x) = \inf\{\|x-y\| : y \in \mathrm{conv}(A)\} \leqslant \inf\{\|x-y\| : y \in A\} \leqslant \mathrm{diam}(\Omega_1) \leqslant 1.$$

Άρα,

$$\begin{aligned}\int_\Omega e^{\varphi_A^2/4}\, dP = P(A) + \int_{\Omega\setminus A} e^{\varphi_A^2(t)/4} dP(t) &\leqslant P(A) + \int_{\Omega\setminus A} e^{1/4} dP(t) \\ = P(A) + (1-P(A))e^{1/4} &\leqslant \frac{1}{P(A)}.\end{aligned}$$

Η τελευταία ανισότητα ισχύει γιατί

$$\max_{r\in[0,1]} \big\{r(r + (1-r)e^{1/4})\big\} = 1.$$

Επαγωγικό βήμα: Έστω $A \subseteq \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{n+1}$. Για κάθε $\omega \in \Omega_{n+1}$ θέτουμε

$$A_\omega = \{t \in \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n : (t, \omega) \in A\}.$$

Έστω $v \in \Omega_{n+1}$ τέτοιο ώστε

$$P^n(A_v) = \max_{\omega \in \Omega_{n+1}} P^n(A_\omega).$$

Λήμμα 3.3.7. *Για κάθε $t \in \Omega_1 \times \ldots \times \Omega_n$,*

$$\varphi_A^2(t, v) \leqslant \varphi_{A_v}^2(t).$$

Απόδειξη. Αρκεί να δείξουμε ότι

$$\{\|t - y\|_{X^{(n)}} : y \in \mathrm{conv}(A_v)\} \subseteq \{\|(t, v) - \omega\|_{X^{(n+1)}} : \omega \in \mathrm{conv}(A)\}.$$

Όμως, αν $y \in \mathrm{conv}(A_v)$ τότε $(y, v) \in \mathrm{conv}(A)$ και $\|t - y\|_{X^{(n)}} = \|(t, v) - (y, v)\|_{X^{(n+1)}}$. □

Λήμμα 3.3.8. *Για κάθε $t \in \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ ισχύει*

$$\varphi_A^2(t, \omega) \leqslant \inf_{a \in [0,1]} (a\varphi_{A_\omega}^2(t) + (1 - a)\varphi_{A_v}^2(t) + (1 - a)^2).$$

Απόδειξη. Αν $z_\omega \in \mathrm{conv}(A_\omega)$ και $z_v \in \mathrm{conv}(A_v)$ τότε $(z_\omega, \omega) \in \mathrm{conv}(A)$ και $(z_v, v) \in \mathrm{conv}(A)$. Αφού το $\mathrm{conv}(A)$ είναι κυρτό σύνολο, έχουμε

$$z_a := (1 - a)(z_v, v) + a(z_\omega, \omega) = (az_\omega + (1 - a)z_v, a\omega + (1 - a)v) \in \mathrm{conv}(A).$$

Χρησιμοποιώντας και την $\mathrm{diam}(\Omega_i) \leqslant 1$ έχουμε ότι

$$\begin{aligned}
\|(t, \omega) - z_a\|_{X^{(n+1)}}^2 &= \|(t - az_\omega - (1 - a)z_v, (1 - a)(\omega - v)\|_{X^{(n+1)}}^2 \\
&= \|t - az_\omega - (1 - a)z_v\|_{X^{(n)}}^2 + \|(1 - a)(\omega - v)\|_{n+1}^2 \\
&= \|t - az_\omega - (1 - a)z_v\|_{X^{(n)}}^2 + (1 - a)^2\|\omega - v\|_{n+1}^2 \\
&= \|at + (1 - a)t - az_\omega - (1 - a)z_v\|_{X^{(n)}}^2 + (1 - a)^2\|\omega - v\|_{n+1}^2 \\
&\leqslant \|a(t - z_\omega) + (1 - a)(t - z_v)\|_{X^{(n)}}^2 + (1 - a)^2.
\end{aligned}$$

Από την τριγωνική ανισότητα και την κυρτότητα της συνάρτησης $x \mapsto x^2$ έχουμε ότι

$$\begin{aligned}
\|a(t - z_\omega) + (1 - a)(t - z_v)\|_{X^{(n)}}^2 &\leqslant (a\|t - z_\omega\|_{X^{(n)}} + (1 - a)\|t - z_v\|_{X^{(n)}})^2 \\
&\leqslant a\|t - z_\omega\|_{X^{(n)}}^2 + (1 - a)\|t - z_v\|_{X^{(n)}}^2.
\end{aligned}$$

Άρα,

$$\varphi_A^2(t, \omega) \leqslant a\|t - z_\omega\|_{X^{(n)}}^2 + (1 - a)\|t - z_v\|_{X^{(n)}}^2 + (1 - a)^2.$$

Επειδή τα z_ω, z_v είναι τυχόντα στοιχεία των A_ω και A_v, έπεται ότι

$$\varphi_A^2(t, \omega) \leqslant a\varphi_{A_\omega}^2(t) + (1 - a)\varphi_{A_v}^2(t) + (1 - a)^2$$

για κάθε $a \in [0, 1]$. □

Χρησιμοποιώντας τα παραπάνω λήμματα και την ανισότητα Hölder γράφουμε

$$
\begin{aligned}
\int_\Omega e^{\frac{\varphi_A^2}{4}}\,dP &= \int_{\Omega_1\times\Omega_2\times\ldots\times\Omega_n\times\{v\}} e^{\frac{\varphi_A^2(s)}{4}}dP(s) + \sum_{\omega\neq v}\int_{\Omega_1\times\ldots\times\Omega_n\times\{\omega\}} e^{\frac{\varphi_A^2(s)}{4}}dP(s)\\
&= P_{n+1}(\{v\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_A^2(t,v)}{4}}dP^n(t) + \sum_{\omega\neq v}P_{n+1}(\{\omega\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_A^2(t,\omega)}{4}}dP^n(t)\\
&\leqslant P_{n+1}(\{v\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_v}^2}{4}}\,dP^n\\
&\qquad+\sum_{\omega\neq v}P_{n+1}(\{\omega\})\int_{\Omega_1\times\Omega_2\times\ldots\times\Omega_n} e^{\frac14(a\varphi_{A_\omega}^2(t)+(1-a)\varphi_{A_v}^2(t)+(1-a)^2)}dP^n(t)\\
&\leqslant P_{n+1}(\{v\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_v}^2}{4}}\,dP^n + \sum_{\omega\neq v}P_{n+1}(\{\omega\})e^{\frac{(1-a)^2}{4}}\\
&\qquad\times\Big(\int_{\Omega_1\times\Omega_2\times\ldots\times\Omega_n} e^{\frac{\varphi_{A_\omega}^2(t)}{4}}dP^n(t)\Big)^a\Big(\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_v}^2(t)}{4}}dP^n(t)\Big)^{1-a}\\
&= P_{n+1}(\{v\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_v}^2}{4}}\,dP^n + \sum_{\omega\neq v}P_{n+1}(\{\omega\})e^{\frac{(1-a)^2}{4}}\\
&\qquad\times\Big(\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_\omega}^2}{4}}\,dP^n\Big)^a\Big(\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\frac{\varphi_{A_v}^2}{4}}\,dP^n\Big)^{1-a}
\end{aligned}
$$

για κάθε $a\in[0,1]$. Άρα,

$$
\begin{aligned}
\int_\Omega e^{\varphi_A^2/4}\,dP &\leqslant P_{n+1}(\{v\})\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\varphi_{A_v}^2/4}\,dP^n + \sum_{\omega\neq v}P_{n+1}(\{\omega\})\\
&\qquad\times\inf_{0\leqslant a\leqslant 1}\Big(\Big(\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\varphi_{A_\omega}^2/4}\,dP^n\Big)^a\Big(\int_{\Omega_1\times\Omega_2\ldots\times\Omega_n} e^{\varphi_{A_v}^2/4}\,dP^n\Big)^{1-a}e^{\frac{(1-a)^2}{4}}\Big).
\end{aligned}
$$

Χρησιμοποιώντας και την επαγωγική υπόθεση, παίρνουμε

$$
\begin{aligned}
\int_\Omega e^{\varphi_A^2/4}\,dP &\leqslant P_{n+1}(\{v\})\frac{1}{P^n(A_v)} + \sum_{\omega\neq v}P_{n+1}(\{\omega\})\inf_{0\leqslant a\leqslant 1}\Big(\Big(\frac{1}{P^n(A_\omega)}\Big)^a\Big(\frac{1}{P^n(A_v)}\Big)^{1-a}e^{\frac{(1-a)^2}{4}}\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + \sum_{\omega\neq v}P_{n+1}(\{\omega\})\inf_{0\leqslant a\leqslant 1}\Big(\Big(\frac{P^n(A_v)}{P^n(A_\omega)}\Big)^a e^{\frac{(1-a)^2}{4}}\Big)\Big).
\end{aligned}
$$

Το ελάχιστο της $\frac{e^{(1-a)^2/4}}{\lambda^a}$ όπου $0\leqslant\lambda\leqslant 1$ λαμβάνεται στο σημείο

$$
a(\lambda)=\begin{cases} 1+2\log\lambda & ,\ \text{αν}\ \ 2\log\lambda>-1\\ 0 & ,\ \text{αλλιώς,}\end{cases}
$$

και η αντίστοιχη τιμή είναι

$$
g(\lambda)=\frac{e^{(1-a(\lambda))^2/4}}{\lambda^{a(\lambda)}}=\begin{cases} e^{-\log\lambda-(\log\lambda)^2} & ,\ \text{αν}\ \ 2\log\lambda>-1\\ e^{1/4} & ,\ \text{αλλιώς.}\end{cases}
$$

Είχαμε καταλήξει στην

$$
\int_\Omega e^{\varphi_A^2/4}\,dP \leqslant \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + \sum_{\omega\neq v}P_{n+1}(\{\omega\})\inf_{0\leqslant a\leqslant 1}\Big(\Big(\frac{P^n(A_v)}{P^n(A_\omega)}\Big)^a e^{\frac{(1-a)^2}{4}}\Big)\Big).
$$

Αφού $\lambda_\omega = \frac{P^n(A_\omega)}{P^n(A_v)} \leqslant 1$, η ανισότητα γράφεται στη μορφή

$$\int_\Omega e^{\varphi_A^2/4}\,dP \leqslant \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + \sum_{\omega\neq v} P_{n+1}(\{\omega\})g(\lambda_\omega)\Big).$$

Ισχυρισμός. Για κάθε $0 \leqslant \lambda \leqslant 1$ ισχύει $g(\lambda) \leqslant 2-\lambda$ (η απόδειξη του ισχυρισμού είναι τετριμμένη).

Από τον ισχυρισμό έχουμε:

$$\begin{aligned}
\int_\Omega e^{\varphi_A^2/4}\,dP &\leqslant \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + \sum_{\omega\neq v} P_{n+1}(\{\omega\})(2-\lambda_\omega)\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + 2(1-P_{n+1}(\{v\})) - \sum_{\omega\neq v}\frac{P_{n+1}(\{\omega\})P^n(A_\omega)}{P^n(A_v)}\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + 2(1-P_{n+1}(\{v\})) - \sum_{\omega\neq v}\frac{P^{n+1}(A_\omega\times\{\omega\})}{P^n(A_v)}\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + 2(1-P_{n+1}(\{v\})) - \frac{P(A)-P^{n+1}(A_v\times\{v\})}{P^n(A_v)}\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + (1-P_{n+1}(\{v\}))\times\Big(2-\Big(\frac{P(A)-P^{n+1}(A_v\times\{v\})}{P^n(A_v)(1-P_{n+1}(v))}\Big)\Big)\Big)\\
&= \frac{1}{P^n(A_v)}\Big(P_{n+1}(\{v\}) + (1-P_{n+1}(\{v\}))\times\Big(2-\Big(\frac{P(A)-P^n(A_v)P_{n+1}(\{v\})}{(1-P_{n+1}(\{v\}))P^n(A_v)}\Big)\Big)\Big).
\end{aligned}$$

Θέτουμε

$$q = P_{n+1}(\{v\}) \text{ και } t = \frac{P(A)-P^n(A_v)P_{n+1}(\{v\})}{(1-P_{n+1}(\{v\}))P^n(A_v)}.$$

Τότε,

$$\int_\Omega e^{\varphi_A^2/4}\,dP \leqslant \frac{q+(1-q)(2-t)}{P^n(A_v)}.$$

Παρατηρούμε ότι $0 \leqslant q, t \leqslant 1$. Πράγματι,

$$\begin{aligned}
t &= \frac{P(A)-P^n(A_v)P_{n+1}(\{v\})}{(1-P_{n+1}(\{v\}))P^n(A_v)} = \frac{\sum_{\omega\neq v}P^{n+1}(A_\omega\times\{\omega\})}{(1-P_{n+1}(\{v\}))P^n(A_v)}\\
&= \frac{\sum_{\omega\neq v}P^n(A_\omega)P_{n+1}(\{\omega\})}{(1-P_{n+1}(\{v\}))P^n(A_v)} \leqslant \frac{\sum_{\omega\neq v}P_{n+1}(\{\omega\})}{1-P_{n+1}(\{v\})}\\
&= 1.
\end{aligned}$$

Από τον ορισμό του t έπεται ότι

$$\frac{1}{P(A)} = \frac{1}{P^n(A_v)((1-q)t+q)}.$$

Αν λοιπόν δείξουμε ότι για κάθε $q, t \in [0,1]$ ισχύει $q+(1-q)(2-t) \leqslant \frac{1}{q+(1-q)t}$ θα πάρουμε την

$$\int_\Omega e^{\varphi_A^2/4}\,dP \leqslant \frac{1}{P(A)}.$$

Όμως,

$$\begin{aligned}(q+(1-q)(2-t))(q+(1-q)t) &= q^2+2(1-q)q+(1-q)^2t(2-t)\\ &\leqslant q^2+2(1-q)q+(1-q)^2=(q+(1-q))^2=1,\end{aligned}$$

και αυτό ολοκληρώνει την απόδειξη του θεωρήματος. □

3.3.3 Γινόμενα χώρων πιθανότητας – έλεγχος με ένα σημείο

Έστω $(\Omega,\mathcal{A},\mu)$ χώρος πιθανότητας χωρίς άτομα (υπενθυμίζουμε ότι αν $(X,\mathcal{A},\mu)$ είναι ένας χώρος μέτρου τότε το $A\in\mathcal{A}$ λέγεται *άτομο* αν $\mu(A)>0$ και για κάθε $B\in\mathcal{A}$ με $B\subseteq A$ και $\mu(B)<\mu(A)$ ισχύει ότι $\mu(B)=0$).

Θεωρούμε το μέτρο γινόμενο $P:=P^N=\mu\otimes\cdots\otimes\mu$ (N φορές) στον $X=\Omega\times\cdots\times\Omega$. Ένα σημείο $x\in X$ γράφεται στη μορφή $x=(x_1,\dots,x_N)$. Η απόσταση στον X ορίζεται από την

$$d(x,y)=\mathrm{card}\{1\leqslant i\leqslant N:\ x_i\neq y_i\}.$$

Αν $A\subseteq X$, ορίζουμε

$$f(A,x)=\min\{d(y,x):\ y\in A\}.$$

Σκοπός μας είναι να δείξουμε το εξής θεώρημα «ελέγχου με ένα σημείο».

Θεώρημα 3.3.9. *Για κάθε $t>0$ έχουμε ότι*

$$\int e^{tf(A,x)}dP^N(x)\leqslant\frac{1}{P^N(A)}\Big(\frac{1}{2}+\frac{e^t+e^{-t}}{4}\Big)^N\leqslant\frac{1}{P^N(A)}e^{t^2N/4}.$$

Πόρισμα 3.3.10. *Ειδικότερα, για κάθε $k\leqslant N$ ισχύει ότι*

$$P^N(\{x:f(A,x)\geqslant k\})\leqslant\frac{1}{P^N(A)}e^{-k^2/N}.$$

Η απόδειξη θα γίνει με επαγωγή ως προς N και θα βασιστεί στο εξής λήμμα.

Λήμμα 3.3.11. *Έστω $g:\Omega\to[0,1]$ μετρήσιμη. Τότε,*

$$\int_\Omega\min\Big\{e^t,\frac{1}{g(\omega)}\Big\}\,d\mu(\omega)\int_\Omega g(\omega)d\mu(\omega)\leqslant a(t),$$

όπου $a(t)=\frac{1}{2}+\frac{e^t+e^{-t}}{4}$.

Απόδειξη. Έστω ότι έχει γίνει η απόδειξη για μετρήσιμες συναρτήσεις $h:\Omega\to[e^{-t},1]$. Έστω ότι $0\leqslant g\leqslant 1$. Ορίζουμε $h=\max\{e^{-t},g\}$. Τότε η h ικανοποιεί την

$$\int_\Omega\min\Big\{e^t,\frac{1}{\max\{g,e^{-t}\}}\Big\}\,d\mu\int_\Omega\max\{g,e^{-t}\}d\mu\leqslant a(t).$$

Διακρίνοντας τις περιπτώσεις $g(\omega)>e^{-t}$ και $g(\omega)\leqslant e^{-t}$, εύκολα βλέπουμε ότι

$$\min\big\{e^t,\frac{1}{\max\{g,e^{-t}\}}\big\}=\min\big\{e^t,\frac{1}{g}\big\}.$$

Άρα,

$$\begin{aligned}\int_\Omega \min\{e^t, 1/g\}\,d\mu \cdot \int_\Omega g\,d\mu &= \int_\Omega \min\Big\{e^t, \frac{1}{\max\{g, e^{-t}\}}\Big\}\,d\mu \cdot \int_\Omega g\,d\mu \\ &\leqslant \int_\Omega \min\Big\{e^t, \frac{1}{\max\{g, e^{-t}\}}\Big\}\,d\mu \cdot \int_\Omega \max\{g, e^{-t}\}\,d\mu \\ &\leqslant a(t).\end{aligned}$$

Αρκεί λοιπόν να δείξουμε το ζητούμενο για $g : e^{-t} \leqslant g \leqslant 1$.

Γράφουμε $C = \{g : \Omega \to [e^{-t}, 1]\}$ και για κάθε $b \in [e^{-t}, 1]$ θεωρούμε το $C_b = \{g \in C : \int_\Omega g\,d\mu = b\}$. Η συνάρτηση $F : C_b \to \mathbb{R}$ με

$$F(g) = \int_\Omega \frac{1}{g} d\mu$$

είναι κυρτή στο C_b αφού η $x \to \frac{1}{x}$ είναι κυρτή στο $\mathbb{R}^+$. Θεωρούμε το C_b σαν υποσύνολο κάποιου $(L^p)^* = L^q$, $p > 1$.

Ισχυρισμός. Το C_b είναι w^*-κλειστό υποσύνολο της B_{L^q}.

Πράγματι, αφού ο L^q είναι αυτοπαθής χώρος έχουμε ότι οι w και w^* τοπολογίες του L^q ταυτίζονται. Όμως, το C_b είναι κυρτό και θα δείξουμε ότι είναι και $\|\cdot\|_q$-κλειστό. Τότε, από το θεώρημα Mazur έπεται ο ισχυρισμός. Έστω $(g_n)_{n\in\mathbb{N}}$ στο C_b με $\|g_n - g\|_q \to 0$. Τότε,

$$\Big| \int_\Omega g_n\,d\mu - \int_\Omega g\,d\mu \Big| \leqslant \int_\Omega |g_n - g|\,d\mu \leqslant \|g_n - g\|_q$$

από την ανισότητα Hölder. Άρα, $b = \int_\Omega g_n\,d\mu \to \int_\Omega g\,d\mu$, και συνεπώς $\int_\Omega g\,d\mu = b$. Θα δείξουμε ακόμη ότι $g(\Omega) \subseteq [e^{-t}, 1]$. Έστω ότι δεν ισχύει. Τότε, χωρίς περιορισμό της γενικότητας, έστω ότι $\mu(\{g > 1\}) > 0$. Άρα, υπάρχει $\varepsilon > 0$ ώστε $\mu(\{g \geqslant 1 + \varepsilon\}) > 0$. Συνεπώς,

$$0 < \varepsilon \cdot \mu(\{g \geqslant 1 + \varepsilon\}) \leqslant \int_\Omega |g - g_n|\,d\mu \leqslant \|g_n - g\|_q \to 0,$$

το οποίο είναι άτοπο. Οπότε, $g \in C_b$.

Από το θεώρημα του Alaoglu έπεται ότι το C_b είναι w^*-συμπαγές, και από το θεώρημα Krein-Milman ισχύει ότι

$$C_b = \overline{\mathrm{conv}\{x_i : x_i \in \mathrm{ext}(C_b)\}}^{w^*}.$$

Λήμμα 3.3.12. *Τα ακραία σημεία του C_b είναι της μορφής*

$$\chi_A e^{-t} + \chi_{\Omega\setminus A}, \quad A \in \mathcal{A}.$$

Απόδειξη. Με απαγωγή σε άτοπο. Έστω $w \in C_b$ ακραίο σημείο του C_b η οποία δεν είναι της μορφής $\chi_A e^{-t} + \chi_{\Omega\setminus A}$. Εαν $\Gamma = \{\omega : e^{-t} < w(\omega) < 1\}$, τότε $\mu(\Gamma) > 0$. Θέτουμε

$$\Gamma_\delta = \{\omega : e^{-t} + \delta < w(\omega) < 1 - \delta\}.$$

Τότε, $\lim_{\delta\to 0^+} \mu(\Gamma_\delta) = \mu(\Gamma)$. Άρα, υπάρχει $\delta > 0$ τέτοιος ώστε $\mu(\Gamma_\delta) > 0$. Αφού ο χώρος δεν έχει άτομα, υπάρχουν Δ_1, Δ_2 τέτοια ώστε $\Delta_1 \cup \Delta_2 = \Gamma_\delta$, $\Delta_1 \cap \Delta_2 = \emptyset$ και $\mu(\Delta_1) > 0, \mu(\Delta_2) > 0$.

Ορίζουμε

$$w_1(\omega) = \begin{cases} w(\omega) + \delta\frac{\mu(\Delta_2)}{\mu(\Delta_1)+\mu(\Delta_2)} & , \text{αν } \omega \in \Delta_1 \\ w(\omega) - \delta\frac{\mu(\Delta_1)}{\mu(\Delta_1)+\mu(\Delta_2)} & , \text{αν } \omega \in \Delta_2 \\ w(\omega) & , \text{αν } \omega \notin \Gamma_\delta \end{cases}$$

και

$$w_2(\omega) = \begin{cases} w(\omega) - \delta\frac{\mu(\Delta_2)}{\mu(\Delta_1)+\mu(\Delta_2)} & , \text{αν } \omega \in \Delta_1 \\ w(\omega) + \delta\frac{\mu(\Delta_1)}{\mu(\Delta_1)+\mu(\Delta_2)} & , \text{αν } \omega \in \Delta_2 \\ w(\omega) & , \text{αν } \omega \notin \Gamma_\delta \end{cases}$$

Τότε, $w = \frac{w_1+w_2}{2}$ και

$$\begin{aligned} \int_\Omega w_1 d\mu &= \int_{\Delta_1} w(\omega) + \delta\frac{\mu(\Delta_2)}{\mu(\Delta_1)+\mu(\Delta_2)} \\ &\quad + \int_{\Delta_2} w(\omega) - \delta\frac{\mu(\Delta_1)}{\mu(\Delta_1)+\mu(\Delta_2)} + \int_{\Omega\setminus\Gamma_\delta} w d\mu \\ &= \int_\Omega w d\mu = b. \end{aligned}$$

Όμοια, $\int_\Omega w_2 d\mu = b$. Επίσης για τις w_1, w_2 ισχύει $e^{-t} \leqslant w_1 \leqslant 1$ και $e^{-t} \leqslant w_2 \leqslant 1$. Άρα $w_1, w_2 \in C_b$. Επομένως η w δεν είναι ακραίο σημείο του C_b. □

Η F είναι w^*-συνεχής, και αφού είναι κυρτή τότε το $A = \{x \in C_b : F(x) = \max(F\,|_{C_b})\}$ είναι κλειστό και κυρτό υποσύνολο του συμπαγούς συνόλου C_b. Άρα, $A \cap \mathrm{ext}(C_b) \neq \emptyset$ και η $F|_{C_b}$ παίρνει μέγιστο σε ακραίο σημείο του C_b. Το ίδιο ισχύει για την

$$H(g) = \int_\Omega \min\Big\{e^t, \frac{1}{g(\omega)}\Big\} d\mu(\omega) \int_\Omega g(\omega) d\mu(\omega), \;\; g \in C_b.$$

Άρα, για κάθε $b \in [e^{-t}, 1]$ αρκεί να ελέγξουμε όλες τις $u = \chi_A e^{-t} + \chi_{\Omega\setminus A}$ που ανήκουν στο C_b. Ισοδύναμα, αρκεί να ελέγξουμε όλες τις $u = \chi_A e^{-t} + \chi_{\Omega\setminus A}$. Έστω u μία τέτοια συνάρτηση και έστω $\mu(A) = p$. Τότε, έχουμε :

$$\begin{aligned} \int_\Omega \min\Big\{e^t, \frac{1}{u(\omega)}\Big\} d\mu(\omega) \int_\Omega u(\omega) d\mu(\omega) &= \Big(\int_\Omega \frac{1}{u} d\mu\Big)\Big(\int_\Omega u d\mu\Big) \\ &= (pe^t + (1-p))(pe^{-t} + (1-p)). \end{aligned}$$

Εαν θεωρήσουμε την συνάρτηση $z(p) = (pe^t + (1-p))(pe^{-t} + (1-p))$, $p \in [0,1]$, τότε παρατηρούμε ότι $z(1-p) = z(p)$. Επίσης $z''(p) = 2(e^t - 1)(e^{-t} - 1) \leqslant 0$, δηλαδή η z είναι κοίλη. Άρα,

$$z(1/2) \geqslant \frac{z(p) + z(1-p)}{2} = \frac{z(p) + z(p)}{2} = z(p).$$

Άρα, το μέγιστο της z είναι

$$z(1/2) = \Big(\frac{1}{2} + \frac{1}{2}e^t\Big)\Big(\frac{1}{2} + \frac{1}{2}e^{-t}\Big) = \frac{1}{2} + \frac{e^{-t} + e^t}{4} = a(t).$$

Αυτό αποδεικνύει το Λήμμα 3.3.11. □

Απόδειξη του Θεωρήματος 3.3.9. Με επαγωγή ως προς N. Έστω ότι $N = 1$ και έστω $A \in \mathcal{A}$. Από το Λήμμα 3.3.11 για την $g = \chi_A$, έχουμε

$$\int_\Omega \min\{e^t, 1/\chi_A\} d\mu(\omega) \int_\Omega \chi_A d\mu(\omega) \leqslant a(t)$$

άρα

$$\int_A \min\Big\{e^t, \frac{1}{\chi_A}\Big\} d\mu(\omega) + \int_{\Omega\setminus A} \min\Big\{e^t, \frac{1}{\chi_A}\Big\} d\mu(\omega) \leqslant \frac{a(t)}{\mu(A)}.$$

Όμως αν $\omega \in A$ τότε $\min\Big\{e^t, \frac{1}{\chi_A}\Big\} = 1 = e^{tf(A,\omega)}$ αφού $f(A,\omega) = 0$, και αν $\omega \notin A$ τότε $\min\Big\{e^t, \frac{1}{\chi_A}\Big\} = e^t = e^{tf(A,\omega)}$ αφού $f(A,\omega) = 1$. Άρα,

$$\int_\Omega e^{tf(A,\omega)} d\mu(\omega) \leqslant \frac{a(t)}{\mu(A)}.$$

Επαγωγικό βήμα: Υποθέτουμε οτι το θεώρημα ισχύει για το N και θα το δείξουμε για το $N+1$. Θεωρούμε $A \subseteq \Omega^{N+1}$ και θέτουμε

$$A(\omega) = \{x \in \Omega^N : (x,\omega) \in A\}, \ \ \omega \in \Omega$$

και

$$\Gamma = \{x \in \Omega^N : \exists \omega \in \Omega \,\mu\varepsilon\, (x,\omega) \in A\}.$$

Παρατηρούμε ότι $A(\omega) \subseteq \Gamma$ για κάθε $\omega \in \Omega$.

Ισχυρισμός 1. $f(A,(x,\omega)) \leqslant f(A(\omega),x)$.

Απόδειξη. Έχουμε

$$f(A,(x,\omega)) = \min\{d((x,\omega),z) : z \in A\}$$

και

$$f(A(\omega),x) = \min\{d(x,y) : y \in A(\omega)\}$$

όπου $d(x,y) = \mathrm{card}\{i \leqslant N : x_i \neq y_i\}$. Αρκεί να δείξουμε ότι

$$\{d(x,y) : y \in A(\omega)\} \subseteq \{d((x,\omega),z) : z \in A\}.$$

Όμως, αν $y \in A(\omega)$ τότε $(y,\omega) \in A$ και $d((y,\omega),(x,\omega)) = d(x,y)$. □

Ισχυρισμός 2. $f(A,(x,\omega)) \leqslant f(\Gamma,x) + 1$.

Απόδειξη. Έστω $y \in \Gamma$ με $d(x,y) = f(\Gamma,x)$. Τότε, υπάρχει ω' τέτοιο ώστε $(y,\omega') \in A$. Άρα,

$$\begin{aligned} d((y,\omega'),(x,\omega)) &\leqslant d((y,\omega'),(x,\omega')) + d((x,\omega),(x,\omega')) = d(y,x) + d(\omega',\omega) \\ &\leqslant d(x,y) + 1. \end{aligned}$$

Άρα $d((y,\omega'),(x,\omega)) \leqslant f(\Gamma,x) + 1$. Επομένως, $f(A,(x,\omega)) \leqslant f(\Gamma,x) + 1$. □

Χρησιμοποιώντας την επαγωγική υπόθεση και τους δύο ισχυρισμούς, παίρνουμε:

$$\int_{\Omega^N} e^{tf(A,(x,\omega))} dP^N(x) \leqslant \int_{\Omega^N} e^{tf(A(\omega),x)} dP^N(x) \leqslant \frac{a^N(t)}{P^N(A(\omega))}$$

και

$$\int_{\Omega^N} e^{tf(A,(x,\omega))} dP^N(x) \leqslant e^t \int_{\Omega^N} e^{tf(\Gamma,x)} dP^N(x) \leqslant e^t \frac{a^N(t)}{P^N(\Gamma)}.$$

Άρα,

$$\int_{\Omega^N} e^{tf(A,(x,\omega))} dP^N(x) \leqslant \min\Big\{e^t \frac{a^N(t)}{P^N(\Gamma)}, \frac{a^N(t)}{P^N(A(\omega))}\Big\}.$$

Αν ολοκληρώσουμε ως προς ω την προηγούμενη σχέση παίρνουμε

$$\int_\Omega e^{tf(A,(x,\omega))} dP^N(\omega) d\mu(\omega) \leqslant a^N(t) \int_\Omega \min\Big\{\frac{e^t}{P^N(\Gamma)}, \frac{1}{P^N(A(\omega))}\Big\} d\mu(\omega).$$

Από το Λήμμα 3.3.11 για την $g(\omega) = P^N(A(\omega))/P^N(\Gamma)$ έπεται ότι

$$\int_\Omega \min\Big\{\frac{e^t}{P^N(\Gamma)}, \frac{1}{P^N(A(\omega))}\Big\} d\mu(\omega) \leqslant \frac{a(t)}{\int_\Omega P^N(A(\omega)) d\mu(\omega)}.$$

Από το θεώρημα Fubini,

$$P(A) = \int_\Omega P^N(A(\omega)) d\mu(\omega).$$

Άρα, τελικά

$$\int_{\Omega^{N+1}} e^{tf(A,(x,\omega))} \leqslant \frac{a^{N+1}(t)}{P(A)}.$$

Αυτό ολοκληρώνει την απόδειξη του θεωρήματος. Για να πάρουμε την δεύτερη ανισότητα, γράφουμε

$$\frac{1}{2} + \frac{e^t + e^{-t}}{4} = 1 + \sum_{n=1}^{\infty} \frac{t^{2n}}{2(2n)!}$$

και παρατηρώντας ότι $2(2n)! \geqslant 4^n n!$, παίρνουμε

$$a(t) = 1 + \sum_{n=1}^{\infty} \frac{t^{2n}}{2(2n)!} \leqslant 1 + \sum_{n=1}^{\infty} \frac{t^{2n}}{4^n\, n!} = 1 + \sum_{n=1}^{\infty} \frac{(t^2/4)^n}{n!} = \sum_{n=0}^{\infty} \frac{(t^2/4)^n}{n!} = e^{t^2/4}.$$

Επομένως,

$$\int e^{tf(A,x)} dP(x) \leqslant \frac{1}{P(A)} \left(\frac{1}{2} + \frac{e^t + e^{-t}}{4}\right)^N \leqslant \frac{1}{P(A)} e^{t^2 N/4}.$$

□

Απόδειξη του Πορίσματος 3.3.10. Γράφουμε

$$e^{tk} P^N(\{x : f(A,x) \geqslant k\}) = \int_{\{x:f(A,x)\geqslant k\}} e^{tk} \leqslant \int_\Omega e^{tf(A,x)} dP^N(x) \leqslant \frac{1}{P^N(A)} e^{t^2 N/4}.$$

Για $t = \frac{2k}{N}$ έχουμε

$$P^N(\{x : f(A,x) \geqslant k\}) \leqslant \frac{1}{P^N(A)} e^{-k^2/N}.$$

□

Παρατήρηση 3.3.13. Τα αποτελέσματα αυτής της παραγράφου γενικεύονται στην περίπτωση $\Omega = \Omega_1 \times \cdots \times \Omega_N$ και $P = \mu_1 \times \cdots \times \mu_N$, όπου $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i \leqslant N$ χώροι πιθανότητας με άτομα. Η απόδειξη είναι ακριβώς η ίδια.

3.3.4 Γινόμενα χώρων πιθανότητας – έλεγχος με q σημεία

Έστω $(\Omega, \mathcal{A}, \mu)$ χώρος πιθανότητας. Θεωρούμε το μέτρο γινόμενο $P := P^N = \mu \otimes \cdots \otimes \mu$ (N φορές) στον $\Omega^N = \Omega \times \cdots \times \Omega$. Ένα σημείο $x \in X$ γράφεται στη μορφή $x = (x_1, \ldots, x_N)$. Αν $q \geqslant 2$ και $A_1, A_2, \ldots, A_q$ μετρήσιμα υποσύνολα του Ω^N, για κάθε $x \in \Omega^N$ θέτουμε

$$f(A_1, A_2, \ldots, A_q, x) = \min\{\mathrm{card}\{i \leqslant N : x_i \notin \{y_i^1, \ldots, y_i^q\}\} : y^1 \in A_1, \ldots, y^q \in A_q\}.$$

Σκοπός μας είναι να δείξουμε το εξής θεώρημα «ελέγχου με q σημεία».

Θεώρημα 3.3.14. *Έστω $q \geqslant 2$ και $A_1, \ldots, A_q \subseteq \Omega^N$ μετρήσιμα. Τότε,*

$$\int_{\Omega^N} q^{f(A_1,\ldots,A_q,x)} dP(x) \leqslant \frac{1}{\prod_{i \leqslant q} P(A_i)}.$$

Ειδικότερα, για κάθε μετρήσιμο $A \subseteq \Omega^N$ και κάθε $k > 0$,

$$P(\{f(A, \ldots, A, x) \geqslant k\}) \leqslant \frac{1}{q^k P(A)^q}.$$

Λήμμα 3.3.15. *Έστω $q \geqslant 2$ και $g : \Omega \to \mathbb{R}$ μετρήσιμη, τέτοια ώστε $1/q \leqslant g \leqslant 1$. Τότε,*

$$\Big(\int_\Omega \frac{1}{g} d\mu\Big)\Big(\int_\Omega g d\mu\Big)^q \leqslant 1.$$

Απόδειξη. Δείχνουμε πρώτα ότι

$$\int_\Omega \frac{1}{g} d\mu + q \int_\Omega g d\mu \leqslant q + 1. \tag{3.3.1}$$

Πράγματι, εύκολα ελέγχουμε ότι αν $q^{-1} \leqslant x \leqslant 1$ τότε

$$\frac{1}{x} + qx \leqslant q + 1$$

άρα, αφού $1/q \leqslant g \leqslant 1$ έχουμε $1/g(\omega) + qg(\omega) \leqslant q + 1$ για κάθε $\omega \in \Omega$, και ολοκληρώνοντας παίρνουμε την (3.3.1).

Παρατηρούμε ότι αν $a, b > 0$ με $a + qb \leqslant 1$ τότε

$$\begin{aligned} a + qb \leqslant q + 1 &\Rightarrow a - 1 + q(b - 1) \leqslant 0 \Rightarrow \log a + q \log b \leqslant 0 \\ &\Rightarrow e^{\log a + q \log b} \leqslant e^0 \Rightarrow ab^q \leqslant 1. \end{aligned}$$

Άρα, αν θέσουμε $a = \int_\Omega 1/g d\mu$ και $b = \int_\Omega g d\mu$ έχουμε το ζητούμενο. □

Πόρισμα 3.3.16. *Έστω $g_i : \Omega \to [0, 1]$ μετρήσιμες, $1 \leqslant i \leqslant q$. Τότε,*

$$\Big(\int_\Omega \min\Big\{q, \min_{i \leqslant q} \frac{1}{g_i}\Big\} d\mu\Big) \cdot \prod_{i \leqslant q} \int_\Omega g_i d\mu \leqslant 1.$$

Απόδειξη. Θέτουμε $g = \big(\min\{q, \min_{i\leqslant q} 1/g_i\}\big)^{-1}$. Χρησιμοποιώντας την $0 \leqslant g_i \leqslant 1$ ελέγχουμε ότι

$$\frac{1}{q} \leqslant g \leqslant 1.$$

Από το Λήμμα 3.3.15 έπεται ότι

$$\int_\Omega \min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i}\Big\}\,d\mu \cdot \Big(\int_\Omega g d\mu\Big)^q \leqslant 1.$$

Όμως,

$$\prod_{i\leqslant q}\int_\Omega g_i d\mu \leqslant \Big(\int_\Omega g d\mu\Big)^q.$$

Επομένως,

$$\Big(\int_\Omega \min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i}\Big\}\,d\mu\Big)\cdot\prod_{i\leqslant q}\int_\Omega g_i d\mu \leqslant 1.$$

□

Απόδειξη του Θεωρήματος 3.3.14. Με επαγωγή ως προς N. Για $N = 1$, θεωρούμε τις συναρτήσεις

$$g_i(\omega) = \begin{cases} 1 & , \text{ αν } \omega \in A_i \\ \frac{1}{n} & , \text{ αλλιώς} \end{cases}$$

όπου $n > q$, $n \in \mathbb{N}$. Από το Πόρισμα 3.3.16 έπεται ότι

$$\int_\Omega \min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i}\Big\}\,d\mu \leqslant \frac{1}{\prod_{i\leqslant q}\Big(\mu(A_i) + \frac{(1-\mu(A_i))}{n}\Big)}$$

δηλαδή

$$\int_{A_1\cup A_2\cup\ldots\cup A_q}\min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i}\Big\}\,d\mu + \int_{\Omega\setminus A_1\cup A_2\ldots\cup A_n}\min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i}\Big\}\,d\mu$$
$$\leqslant \frac{1}{\prod_{i\leqslant q}\Big(\mu(A_i) + \frac{(1-\mu(A_i))}{n}\Big)}.$$

Αν $x \in A_1 \cup A_2 \ldots \cup A_q$, τότε υπάρχει $i_0 \leqslant q$ με $x \in A_{i_0}$. Αφού $x \in A_{i_0}$, έχουμε

$$\min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i(x)}\Big\} = \min\{q, 1\} = 1 = q^0 = q^{f(A_1,A_2,\ldots,A_q,x)}.$$

Αν $x \notin A_1, A_2, \ldots, A_q$, τότε

$$\min\Big\{q, \min_{i\leqslant q}\frac{1}{g_i(x)}\Big\} = \min\{q, n\} = q$$

και

$$f(A_1,\ldots,A_q,x) = \min(\mathrm{card}\{i \leqslant 1, x_i \notin \{y_i^1,\ldots y_i^q\}\} : y_i^1 \in A_1,\ldots,y_i^q \in A_q) = 1.$$

Έπεται ότι

$$\int_{A_1\cup\ldots\cup A_q} q^{f(A_1,A_2,\ldots,A_q,x)}d\mu + \int_{\Omega\setminus A_1\cup\ldots\cup A_q} q^{f(A_1,\ldots A_q,x)}d\mu \leqslant \frac{1}{\prod_{i\leqslant q}\left(\mu(A_i)+\frac{(1-\mu(A_i))}{n}\right)}.$$

Αφήνοντας το $n\to\infty$ παίρνουμε

$$\int_\Omega q^{f(A_1,\ldots,A_q,x)}d\mu \leqslant \frac{1}{\prod_{i\leqslant q}\mu(A_i)}.$$

Επαγωγικό βήμα: Υποθέτουμε οτι το θεώρημα ισχύει για το N και θα το δείξουμε για το $N+1$. Έστω

$$A_1,\ldots A_q \subseteq \Omega^{N+1}.$$

Ορίζουμε

$$B_i = \{x\in\Omega^N : \exists\omega\in\Omega, (x,\omega)\in A_i\}$$

και

$$A_i(\omega) = \{x\in\Omega^N : (x,\omega)\in A_i\}$$

Ισχυρισμός 1. $f(A_1,\ldots,A_q,(x,\omega)) \leqslant 1+f(B_1,B_2,\ldots,B_q,x)$.

Απόδειξη. Έστω $y^1\in B_1,\ldots,y^q\in B_q$. Τότε, υπάρχουν $\omega^1,\ldots,\omega^q\in\Omega$ τ.ω. $(y^1,\omega^1)\in A_1,\ldots(y^q,\omega^q)\in A_q$. Θέτουμε $z^i=(y^i,\omega^i)$ και $p=(x,\omega)$. Παρατηρούμε ότι

$$\begin{aligned} f(A_1,A_2,\ldots,A_q,(x,\omega)) &\leqslant \mathrm{card}\{k\leqslant N+1 : p_k\notin\{z_k^1,\ldots,z_k^q\}\} \\ &= \sum_{k=1}^{N+1}(1-\chi_{\{z_k^1,\ldots,z_k^q\}}(p_k)) \\ &= \sum_{k=1}^{N}(1-\chi_{\{z_k^1,\ldots,z_k^q\}}(x_k)) + (1-\chi_{\{\omega^1,\ldots,\omega^q\}}(\omega)) \\ &\leqslant \mathrm{card}\{i\leqslant N : x_i\notin\{y_i^1,\ldots,y_i^q\}\}+1. \end{aligned}$$

Αφού τα $y^1,\ldots,y^q$ ήταν τυχόντα σημεία των $B_1,B_2,\ldots,B_q$, έχουμε το ζητούμενο. □

Ισχυρισμός 2. Έστω $j\leqslant q$. Τότε,

$$f(A_1,\ldots,A_q,(x,\omega)) \leqslant f(C_1,\ldots,C_q,x),$$

όπου $C_i=B_i$ για $i\neq j$ και $C_j=A_j(\omega)$.

Απόδειξη. Θέτουμε $z=(x,\omega)$. Αρκεί να δείξουμε οτι

$$\begin{aligned} &\{\mathrm{card}\{k\leqslant N : x_k\notin\{y_k^1,\ldots,y_k^q\} : y^i\in A_i \quad i\neq j, \quad y^j\in A_j(\omega)\} \\ &\qquad\subseteq \{\mathrm{card}\{k\leqslant N+1 : z_k\notin\{y_k^1,\ldots,y_k^q\}\} : y^i\in A_i\} \end{aligned}$$

Έστω οτι $y^i\in A_i$ για $i\neq j$ και $y^j\in A_j(\omega)$ (δηλαδή $(y^j,\omega)\in A_j$). Θέτουμε $b^i=y^i$ και $b^j=y^j$.Τότε,

$$\mathrm{card}\{k\leqslant N+1 : z_k\notin\{b_k^1,\ldots,b_k^q\}\} = \mathrm{card}\{k\leqslant N : x_k\notin\{y_k^1,\ldots,y_k^q\}\}$$

και αυτό δείχνει το ζητούμενο. □

Συνεχίζουμε την απόδειξη ως εξής: ορίζουμε

$$g_i(\omega) = \frac{P(A_i(\omega))}{P(B_i)}, \qquad \omega \in \Omega.$$

Έστω $\omega \in \Omega$ και i_0 τέτοιος ώστε $\min \frac{1}{g_i(\omega)} = \frac{1}{g_{i_0}(\omega)}$. Τότε από τους δύο ισχυρισμούς έχουμε

$$f(A_1, \ldots, A_q, (x, \omega)) \leqslant 1 + f(B_1, \ldots, B_q, x)$$

και

$$f(A_1, \ldots, A_q, (x, \omega)) \leqslant f(C_1, \ldots, C_q, x)$$

όπου $C_i = B_i$ για $i \neq i_0$ και $C_{i_0} = A_{i_0}(\omega)$. Άρα,

$$\begin{aligned}\int_{\Omega^N} q^{f(A_1,\ldots,A_q,(x,\omega))} dP(x) &\leqslant \int_{\Omega^N} q^{1+f(B_1,\ldots,B_q,x)} dP(x)\\ &\leqslant q \int_{\Omega^N} q^{f(B_1,\ldots,B_q,x)} dP(x) \leqslant \frac{q}{\prod_{i\leqslant q} P(B_i)}.\end{aligned}$$

Επίσης

$$\begin{aligned}\int_{\Omega^N} q^{f(A_1,\ldots,A_q,(x,\omega))} dP(x) &\leqslant \int_{\Omega^N} q^{f(C_1,\ldots,C_q,x)} dP(x)\\ &\leqslant \frac{1}{\prod_{i\leqslant q} P(C_i)} = \frac{1}{P(A_{i_0}(\omega)) \prod_{i\neq i_0,\, i\leqslant q} P(B_i)}.\end{aligned}$$

Άρα,

$$\begin{aligned}\int_{\Omega^N} q^{f(A_1,\ldots,A_q,(x,\omega))} dP(x) &\leqslant \min\Big\{ \frac{q}{\prod_{i\leqslant q} P(B_i)}, \frac{1}{P(A_{i_0}(\omega)) \prod_{\{i\neq i_0,\, i\leqslant q\}} P(B_i)} \Big\}\\ &= \frac{1}{\prod_{i\leqslant q} P(B_i)} \min\Big\{ q, \frac{1}{\frac{P(A_{i_0}(\omega))}{P(B_{i_0})}} \Big\}\\ &= \frac{1}{\prod_{i\leqslant q} P(B_i)} \min\Big\{ q, \min_{i\leqslant q} \frac{1}{g_i(\omega)} \Big\}\end{aligned}$$

για κάθε $\omega \in \Omega$. Από το Πόρισμα 3.3.16 και το θεώρημα του Fubini παίρνουμε

$$\begin{aligned}\int_{\Omega}\int_{\Omega^N} q^{f(A_1,\ldots,A_q,(x,\omega))} dP(x) d\mu(\omega) &\leqslant \frac{1}{\prod_{i\leqslant q} P(B_i)} \int_{\Omega} \min\Big\{ q, \min_{i\leqslant q} \frac{1}{g_i(\omega)} \Big\} d\mu(\omega)\\ &\leqslant \frac{1}{\prod_{i\leqslant q} P(B_i)} \quad \frac{1}{\prod_{i\leqslant q} \int_{\Omega} \frac{P(A_i(\omega))}{P(B_i)} d\mu(\omega)}\\ &= \frac{1}{\prod_{i\leqslant q} \int_{\Omega} P(A_i(\omega)) d\mu(\omega)}\\ &= \frac{1}{\prod_{i\leqslant q} P(A_i)}.\end{aligned}$$

Έχοντας αποδείξει τον πρώτο ισχυρισμό του θεωρήματος παρατηρούμε ότι, από την ανισότητα Markov, αν $A \subseteq \Omega^N$ μετρήσιμο και $k > 0$, έχουμε ότι

$$P(\{f(A, \ldots, A, x) \geqslant k\}) \leqslant \frac{1}{q^k} \int_{\Omega^N} q^{f(A,\ldots,A,x)} dP(x) \leqslant \frac{1}{q^k P(A)^q}.$$

□

3.3.5 Γινόμενα χώρων πιθανότητας – κυρτή θήκη

Έστω $(\Omega, \mathcal{A}, \mu)$ χώρος πιθανότητας. Θεωρούμε το μέτρο γινόμενο $P := P^N = \mu \otimes \cdots \otimes \mu$ (N φορές) στον $\Omega^N = \Omega \times \cdots \times \Omega$. Ένα σημείο $x \in \Omega^N$ γράφεται στη μορφή $x = (x_1, \ldots, x_N)$. Αν $A \subseteq \Omega^N$ και $x \in \Omega^N$, έχουμε ορίσει

$$U_A(x) = \{(s_i)_{i \leqslant N} \in \{0,1\}^N : \exists y \in A, s_i = 0 \Rightarrow x_i = y_i\}.$$

Έστω $V_A(x)$ η κυρτή θήκη του $U_A(x)$. Συμβολίζουμε με $f_C(A, x)$ την Ευκλείδεια απόσταση του 0 άπο το $V_A(x)$.

Λήμμα 3.3.17. *Ισχύει η ισοδυναμία:* $0 \in V_A(x) \Leftrightarrow x \in A$.

Απόδειξη. Εαν $0 \in V_A(x)$, τότε αφού $0 \in \mathrm{ext}([0,1]^N)$ πρέπει να ισχύει $0 \in U_A(x)$. Όμως τότε, $x \in A$. Αντίστροφα, αν $x \in A$ τότε $0 \in U_A(x)$, άρα $0 \in V_A(x)$. □

Λήμμα 3.3.18 (θεώρημα Καραθεοδωρή)**.** *Έστω* $B \subseteq \mathbb{R}^n$ *και* $x \in \mathrm{conv}(B)$. *Τότε, το* x *μπορεί να γραφτεί σαν κυρτός συνδυασμός το πολύ* $n+1$ *στοιχείων του* B.

Απόδειξη. Έστω $x \in \mathrm{conv}(B)$. Τότε $x = \sum_{i=1}^k \lambda_i x_i$ για κάποια $k \in \mathbb{N}$, $x_i \in B$, $\lambda_i > 0$, $1 \leqslant i \leqslant k$ με $\lambda_1 + \cdots + \lambda_k = 1$.

Αν $k \leqslant n+1$ δεν έχουμε κάτι να δείξουμε. Έστω λοιπόν ότι $k > n+1$. Τότε τα $x_2 - x_1, \ldots, x_k - x_1$ είναι $k-1 > n$ το πλήθος και συνεπώς είναι γραμμικώς εξαρτημένα. Επομένως, υπάρχουν $\mu_2, \ldots, \mu_k \in \mathbb{R}$, όχι όλα μηδέν, ώστε $\sum_{i=2}^k \mu_i(x_i - x_1) = 0$. Ορίζουμε $\mu_1 = -\sum_{i=2}^k \mu_i$. Τότε $\sum_{i=1}^k \mu_i x_i = 0$ και $\sum_{i=1}^k \mu_i = 0$, και υπάρχει $1 \leqslant i_1 \leqslant k$ ώστε $\mu_{i_1} > 0$ αφού τα μ_i δεν είναι όλα ίσα με 0.

Τώρα, αν $a \in \mathbb{R}$ έχουμε ότι

$$x = \sum_{i=1}^k \lambda_i x_i = \sum_{i=1}^k \lambda_i x_i - a \sum_{i=1}^k \mu_i x_i = \sum_{i=1}^k (\lambda_i - a\mu_i) x_i.$$

Θέτουμε

$$a = \min\left\{\frac{\lambda_i}{\mu_i} : \mu_i > 0\right\} = \frac{\lambda_{i_0}}{\mu_{i_0}}$$

για κάποιο $1 \leqslant i_0 \leqslant k$. Τότε, $a > 0$ και $\lambda_i - a\mu_i \geqslant 0$ για κάθε $1 \leqslant i \leqslant k$. Επιπλέον,

$$\lambda_{i_0} - a\mu_{i_0} = 0.$$

Οπότε, $x = \sum_{i=1}^k (\lambda_i - a\mu_i) x_i$, όπου $\lambda_i - a\mu_i \geqslant 0$ για κάθε $1 \leqslant i \leqslant k$, $\lambda_{i_0} - a\mu_{i_0} = 0$ και $\sum_{i=1}^k (\lambda_i - a\mu_i) = 1 - a \cdot 0 = 1$.

Άρα, το x γράφεται σαν κυρτός συνδυασμός $k-1$ στοιχείων του B. Με επαγωγή έπεται το ζητούμενο. □

Λήμμα 3.3.19. *Για κάθε* $x \in \Omega^N$ *τα* $U_A(x), V_A(x)$ *είναι συμπαγή.*

Απόδειξη. Το σύνολο $U_A(x)$ είναι πεπερασμένο σύνολο, άρα συμπαγές. Θεωρούμε το $B = U_A(x) \times \cdots \times U_A(x) \times D$ (το $U_A(x)$ παίρνεται $N+1$ φορές) όπου

$$D = \left\{(\lambda_1, \ldots, \lambda_{N+1}) \in \mathbb{R}^{N+1} : \lambda_i \geqslant 0 \text{ για κάθε } 1 \leqslant i \leqslant N+1 \text{ και } \sum_{i=1}^{N+1} \lambda_i = 1\right\}.$$

Τότε, το B είναι συμπαγές υποσύνολο του $\mathbb{R}^{(N+1)^2}$. Θεωρούμε την $h : \mathbb{R}^{(N+1)^2} \to \mathbb{R}^N$ με

$$h(x_1, \ldots, x_{N+1}, \mu_1, \ldots, \mu_{N+1}) = \sum_{i=1}^{N+1} \mu_i x_i,$$

που είναι συνεχής. Από το Λήμμα 3.3.18 έπεται ότι $V_A(x) = h(B)$ και συνεπώς το $V_A(x)$ είναι συμπαγές ως εικόνα συμπαγούς συνόλου μέσω συνεχούς συνάρτησης. □

Θεωρούμε σαν «t-επέκταση» του A το σύνολο

$$A_t = \{x \in \Omega^N : f_C(A, x) \leqslant t\}.$$

Σκοπός μας είναι να αποδείξουμε το εξής θεώρημα:

Θεώρημα 3.3.20. *Για κάθε μετρήσιμο $A \subseteq \Omega^N$ έχουμε*

$$\int_{\Omega^N} e^{f_C^2(A,x)/4} dP(x) \leqslant \frac{1}{P(A)}.$$

Ειδικότερα,

$$P(A_t) \geqslant 1 - \frac{1}{P(A)} e^{-t^2/4}.$$

Για να εξηγήσουμε τον τρόπο με τον οποίο χρησιμοποιείται το Θεώρημα 3.3.20, χρειαζόμαστε το εξής λήμμα.

Λήμμα 3.3.21. *Αν $x \in A_t$, τότε για κάθε $(a_i)_{i \leqslant N} \in \mathbb{R}^N$ υπάρχει $y \in A$ με την ιδιότητα*

$$\sum_{\{i \leqslant N : x_i \neq y_i\}} a_i \leqslant t \sqrt{\sum_{i \leqslant N} a_i^2}.$$

Απόδειξη. Έστω $(a_i)_{i \leqslant N} \in \mathbb{R}^N$. Ορίζουμε $\bar{a} : \mathbb{R}^N \to \mathbb{R}$ με $\bar{a}(y) = \sum_{i \leqslant N} a_i y_i$. Το $\bar{a}$ είναι φραγμένο γραμμικό συναρτησοειδές αφού, από την ανισότητα Cauchy-Schwarz,

$$|\bar{a}(y)| \leqslant \Big(\sum_{i=1}^{N} a_i^2 \Big)^{1/2} \Big(\sum_{i=1}^{N} y_i^2 \Big)^{1/2},$$

οπότε $\|\bar{a}\| \leqslant |a|$. Όμως, $\bar{a}(a) = |a| \cdot |a|$, άρα $\|\bar{a}\| = |a|$.

Το $V_A(x)$ είναι συμπαγές, επομένως υπάρχει $y_0 \in V_A(x)$ τέτοιο ώστε $|y_0| = f_C(A, x)$. Άρα,

$$\min\{\bar{a}(y) : y \in V_A(x)\} \leqslant \bar{a}(y_0) \leqslant |\bar{a}| \cdot |y_0| = |\bar{a}| f_C(A, x) \leqslant t \|\bar{a}\|.$$

Ισχυρισμός. $\min\{\bar{a}(y) : y \in V_A(x)\} = \min\{\bar{a}(y) : y \in U_A(x)\}$.

Πράγματι, έστω ότι $\bar{a}(y_0) = \min\{\bar{a}(y) : y \in V_A(x)\}$ για κάποιο $y_0 \in V_A(x)$. Τότε, $y_0 = \sum_{i=1}^{k} \lambda_i y_i$, όπου $\lambda_i \geqslant 0$, $y_i \in U_A(x)$ για κάθε $1 \leqslant i \leqslant k$ και $\sum_{i=1}^{k} \lambda_i = 1$. Αν $\bar{a}(y_i) > \bar{a}(y_0)$ για κάθε $1 \leqslant i \leqslant k$, τότε καταλήγουμε στην $\bar{a}(y_0) > \bar{a}(y_0)$, άτοπο. Άρα, υπάρχει $1 \leqslant i_0 \leqslant k$ ώστε $\bar{a}(y_{i_0}) = \bar{a}(y_0)$.

Επομένως, υπάρχει $s \in U_A(x)$ τέτοιο ώστε $\bar{a}(s) \leqslant t\|\bar{a}\|$. Έστω λοιπόν $y \in A$ ώστε $s_i = 0 \implies x_i = y_i$ για κάθε $1 \leqslant i \leqslant N$. Έχουμε ότι

$$\sum_{\{i\leqslant N: y_i \neq x_i\}} a_i = \sum_{i\leqslant N} a_i \cdot \mathbf{1}_{\{x_i \neq y_i\}} = \sum_{i\leqslant N} a_i s_i = \bar{a}(s) \leqslant t\sqrt{\sum_{i\leqslant N} a_i^2}.$$

□

Χρησιμοποιώντας το Λήμμα 3.3.21 μπορούμε να συγκρίνουμε το Θεώρημα 3.3.9 με το Θεώρημα 3.3.20. Το Θεώρημα 3.3.9 μας δίνει την ανισότητα

$$P^N(\{x : f(A,x) \geqslant k\}) \leqslant \frac{1}{P^N(A)} e^{-k^2/N}$$

όπου

$$f(A,x) = \min\{d(y,x) : \ y \in A\}$$

και

$$d(x,y) = \mathrm{card}\{1 \leqslant i \leqslant N : \ x_i \neq y_i\}$$

($X = \Omega^N$ και A μετρήσιμο υποσύνολο του X με θετικό μέτρο). Αν στο Λήμμα 3.3.21 πάρουμε $a_i = 1$ και $t = k/\sqrt{N}$, βλέπουμε οτι

$$f_C(A,x) \leqslant \frac{k}{\sqrt{N}} \implies \exists y \in A : d(x,y) \leqslant k.$$

Δηλαδή,

$$\{x \in X : f(A,x) \geqslant k\} \subseteq \{x \in X : f_C(A,x) \geqslant k/\sqrt{N}\}.$$

Εφαρμόζοντας το Θεώρημα 3.3.20 βλέπουμε οτι

$$P^N(\{x : f(A,x) \geqslant k\}) \leqslant P^N(x : f_C(A,x) \geqslant k/\sqrt{N}\}) \leqslant \frac{1}{P^N(A)} \exp(-k^2/4N),$$

δηλαδή το Θεώρημα 3.3.9 με ελαφρώς χειρότερη αριθμητική σταθερά. Το μεγάλο όμως πλεονέκτημα του Θεωρήματος 3.3.20 είναι οτι μας αφήνει το περιθώριο να επιλέγουμε τα (a_i), κάτι που έχει σημασία για κάποιες εφαρμογές.

Απόδειξη του Θεωρήματος 3.3.20. Η απόδειξη της πρώτης ανισότητας γίνεται με επαγωγή ως προς N. Για $N = 1$ έχουμε: αν $x \in A$ τότε $0 \in V_A(x)$, άρα $f_C(A,x) = d(0, V_A(x)) = \min_{y \in V_A(x)} |y| = 0$. Επίσης, αν $x \notin A$ τότε $U_A(x) = \{1\}$ άρα $V_A(x) = \{1\}$, επομένως $f_C(A,x) = \min_{y \in V_A(x)} |y| = 1$. Αυτό σημαίνει ότι

$$\begin{aligned}\int e^{f_C^2(A,x)/4} dP(x) &= \int_A e^{f_C^2(A,x)/4} dP(x) + \int_{\Omega \setminus A} e^{f_C^2(A,x)/4} dP(x) \\ &= P(A)e^0 + e^{1/4}(1 - P(A)) \leqslant \frac{1}{P(A)}.\end{aligned}$$

(έχουμε ήδη χρησιμοποιήσει την τελευταία ανισότητα).

Επαγωγικό βήμα: Υποθέτουμε ότι το θεώρημα ισχύει για το N και θα το δείξουμε για το $N+1$. Έστω $A \subseteq \Omega^{N+1}$ μετρήσιμο. Για κάθε $\omega \in \Omega$ θέτουμε

$$A(\omega) = \{x \in \Omega^N : (x, \omega) \in A\}$$

και

$$B = \{x \in \Omega^N : \exists \omega \in \Omega, (x, \omega) \in A\}.$$

Ισχυρισμός 1. Εαν $s \in U_{A(\omega)}(x)$, τότε $(s, 0) \in U_A((x, \omega)))$.

Απόδειξη. Έστω $z = (x, \omega)$. Αφού $s \in U_{A(\omega)}(x)$, υπάρχει $y \in A(\omega)$ τέτοιο ώστε $s_i = 0 \Rightarrow x_i = y_i$. Επειδή $y \in A(\omega)$, έχουμε $v = (y, \omega) \in A$. Το $\mu = (s, 0) \in U_A(x, \omega)$ αφού για κάθε $i \leqslant N$

$$\mu_i = s_i = 0 \Rightarrow z_i = v_i,$$

ενώ για $i = N+1$ έχουμε $z_{N+1} = v_{N+1} = \omega$. □

Ισχυρισμός 2. Εαν $t \in U_B(x)$, τότε $(t, 1) \in U_A((x, \omega))$ για κάθε $\omega \in \Omega$.

Απόδειξη. Έστω $\omega \in \Omega$. Αφού $t \in U_B(x)$ τότε υπάρχει $y \in B$ τέτοιο ώστε, για κάθε $1 \leqslant i \leqslant n$, $(t_i = 0 \Rightarrow y_i = x_i)$ Αφού $y \in B$, υπάρχει $\tilde{\omega} \in \Omega$ ώστε $(y, \tilde{\omega}) \in A$. Τότε, $(t, 1) \in U_A(x, \omega)$: από τον ορισμό αρκεί να ελέγξουμε οτι για κάθε $i \leqslant N$

$$t_i = 0 \Rightarrow x_i = y_i,$$

το οποίο ισχύει. □

Ισχυρισμός 3. Για κάθε $\lambda \in [0, 1]$,

$$f_C^2(A, (x, \omega)) \leqslant (1-\lambda)^2 + \lambda f_C^2(A(\omega), x) + (1-\lambda) f_C^2(B, x).$$

Απόδειξη. Έστω $z = (x, \omega)$. Αν $s \in V_{A(\omega)}(x)$ και $t \in V_B(x)$, παρατηρούμε οτι

$$s \in V_{A(\omega)}(x) \Rightarrow (s, 0) \in V_A(x, \omega)$$

και

$$t \in V_B(x) \Rightarrow (t, 1) \in V_A(x, \omega).$$

Αφού το $V_A(x, \omega)$ είναι κυρτό,

$$\lambda(s, 0) + (1-\lambda)(t, 1) = (\lambda s + (1-\lambda)t, 1-\lambda) \in V_A((x, \omega)).$$

Από τριγωνική ανισότητα και από το γεγονος οτι η x^2 είναι κυρτή, έχουμε

$$\begin{aligned} |(\lambda s + (1-\lambda)t, 1-\lambda)|^2 &= |(\lambda s + (1-\lambda)t, 0)|^2 + |(0, 1-\lambda)|^2 \\ &= |\lambda s + (1-\lambda)t|^2 + (1-\lambda)^2 \\ &\leqslant (\lambda|s| + (1-\lambda)|t|)^2 + (1-\lambda)^2 \\ &\leqslant \lambda|s|^2 + (1-\lambda)|t|^2 + (1-\lambda)^2. \end{aligned}$$

Άρα
$$f_C^2(A,z) \leqslant (1-\lambda)^2 + \lambda|s|^2 + (1-\lambda)|t|^2.$$
Αφού τα t, s είναι τυχόντα στοιχεία των $V_B(x)$ και $V_{A(\omega)}(x)$, έχουμε
$$f_C^2(A,z) \leqslant \lambda f_C^2(A(\omega),x) + (1-\lambda) f_C^2(B,x) + (1-\lambda)^2.$$

□

Από την προηγούμενη ανισότητα, από την ανισότητα Hölder και από την επαγωγική υπόθεση, έχουμε
$$\begin{aligned}\int_{\Omega^N} e^{f_C^2(A,(x,\omega))/4} dP(x) &\leqslant e^{(1-\lambda)^2/4} \int_{\Omega^N} e^{\lambda f_C^2(A(\omega),x)/4} e^{(1-\lambda) f_C^2(B,x)/4} dP(x)\\ &\leqslant e^{(1-\lambda)^2/4} \Big(\int_{\Omega^N} e^{f_C^2(A(\omega),x)/4} dP(x)\Big)^{\lambda} \Big(\int_{\Omega^N} e^{f_C^2(B,x)/4} dP(x)\Big)^{1-\lambda}\\ &\leqslant e^{(1-\lambda)^2/4} \Big(\frac{1}{P(A(\omega))}\Big)^{\lambda} \Big(\frac{1}{P(B)}\Big)^{1-\lambda}\\ &= \frac{1}{P(B)} e^{(1-\lambda)^2/4} \Big(\frac{P(A(\omega))}{P(B)}\Big)^{-\lambda}\end{aligned}$$
για κάθε $\lambda \in [0,1]$. Άρα,
$$\int_{\Omega^N} e^{f_C^2(A,(x,\omega))/4} dP(x) \leqslant \inf_{\lambda\in[0,1]} \frac{1}{P(B)} e^{(1-\lambda)^2/4} \Big(\frac{P(A(\omega))}{P(B)}\Big)^{-\lambda}.$$
Έχουμε δείξει οτι για κάθε $0 \leqslant r \leqslant 1$ ισχύει η
$$\inf_{0\leqslant\lambda\leqslant 1} r^{-\lambda} e^{(1-\lambda)^2/4} \leqslant 2 - r.$$
Αν πάρουμε
$$r = \frac{P(A(\omega))}{P(B)} \leqslant 1,$$
βλέπουμε οτι
$$\int_{\Omega^N} e^{f_C^2(A,(x,\omega))/4} dP(x) \leqslant \frac{1}{P(B)} \Big(2 - \frac{P(A(\omega))}{P(B)}\Big).$$
Από το θεώρημα Fubini έπεται ότι
$$\begin{aligned}\int_{\Omega}\int_{\Omega^N} e^{f_C^2(A,(x,\omega))/4} dP(x) d\mu(\omega) &\leqslant \frac{1}{P(B)} \int_{\Omega} \Big(2 - \frac{P(A(\omega))}{P(B)}\Big) d\mu(\omega)\\ &= \frac{1}{P(B)} \Big(2 - \frac{1}{P(B)} \int_{\Omega} P(A(\omega)) d\mu(\omega)\Big)\\ &\leqslant \frac{1}{P(B)} \cdot \frac{P(B)}{(P\times\mu)(A)}\\ &= \frac{1}{(P\times\mu)(A)}.\end{aligned}$$

Η τελευταία ανισότητα ισχύει λόγω της $x(2-x) \leqslant 1$, $x \in [0,2]$. Αυτό ολοκληρώνει την απόδειξη του πρώτου ισχυρισμού του Θεωρήματος 3.3.20. Τώρα, για κάθε $t > 0$,

$$e^{t^2/4} P(\Omega^N \setminus A_t) \leqslant \int_{\Omega^N \setminus A_t} e^{f_C^2(A,x)/4} dP(x) \leqslant \int_{\Omega^N} e^{f_C^2(A,x)/4} dP(x) \leqslant \frac{1}{P(A)}.$$

Άρα,

$$P(A_t) \geqslant 1 - \frac{1}{P(A)} e^{-t^2/4}.$$

□

3.4 Εφαρμογές

Κλείνουμε αυτό το κεφάλαιο με κάποιες ενδεικτικές εφαρμογές της ανισότητας απόστασης από την κυρτή θήκη. Για την πρώτη, που αποτέλεσε και το κίνητρο για τον Talagrand, δίνουμε αρχικά τον ορισμό και μερικά βασικά αποτελέσματα για την διάμεσο μιας πραγματικής τυχαίας μεταβλητής.

3.4.1 Διάμεσος

Ορισμός 3.4.1. Έστω $(\Omega, \mathcal{A}, P)$ χώρος πιθανότητας και $X : \Omega \to \mathbb{R}$ μια πραγματική τυχαία μεταβλητή. Ένας αριθμός $m \in \mathbb{R}$ θα λέγεται *μέσος Lévy* ή *διάμεσος* της X αν $P(X \leqslant m) \geqslant \frac{1}{2}$ και $P(X \geqslant m) \geqslant \frac{1}{2}$.

Λήμμα 3.4.2. *Κάθε πραγματική τυχαία μεταβλητή έχει διάμεσο.*

Απόδειξη. Θεωρούμε τον $m = \sup\{a \in \mathbb{R} : P(X \geqslant a) \geqslant 1/2\}$ και θα δείξουμε ότι ο m είναι διάμεσος για την X. Αν $a > m$ τότε $P(X \geqslant a) < \frac{1}{2}$, άρα $P(X > m) \leqslant \frac{1}{2}$ παίρνοντας όριο καθώς $a \to m^+$. Οπότε, $P(X \leqslant m) \geqslant \frac{1}{2}$.

Αν $a < m$, τότε υπάρχει $a < s < m$ ώστε $P(X \geqslant s) \geqslant \frac{1}{2}$, άρα $P(X \geqslant a) \geqslant P(X \geqslant s) \geqslant \frac{1}{2}$. Παίρνοντας όριο καθώς $a \to m^-$ έχουμε ότι $P(X \geqslant m) \geqslant \frac{1}{2}$. □

Πρόταση 3.4.3. *Έστω $(\Omega_i, \mathcal{A}_i, \mu_i)$ χώροι πιθανότητας, $1 \leqslant i \leqslant n$ και έστω $(X, \mathcal{A}, P) = (\times_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \mu_i)$ ο χώρος γινόμενο. Έστω $F : X \to \mathbb{R}$ μετρήσιμη συνάρτηση τέτοια ώστε για κάθε $x \in X$ υπάρχει $a = a(x) \in [0, +\infty)^n$ με $|a| = 1$ ώστε για κάθε $y \in X$ να ισχύει*

$$F(x) \leqslant F(y) + d_a(x,y).$$

Αν M είναι μια διάμεσος της F τότε για κάθε $r \geqslant 0$ ισχύει ότι

$$P(|F - M| \geqslant r) \leqslant 4e^{-\frac{r^2}{4}}.$$

Απόδειξη. Θεωρούμε το σύνολο $A = \{F \leqslant M\}$. Στην Πρόταση 3.2.6 έχουμε δείξει ότι $d_A(x) = \sup_{|a|=1} d_a(x, A)$, όπου $d_A(x) = \mathrm{dist}(0, V_A(x))$.

Έστω $y \in A$. Τότε, για κάθε $x \in X$ υπάρχει $a(x) \in [0, \infty)^n$ με $|a(x)| = 1$ ώστε

$$F(x) \leqslant F(y) + d_a(x,y) \leqslant M + d_a(x,A) \leqslant M + d_A(x).$$

Οπότε αν $x \in \{F \geqslant M+r\}$ τότε $d_A(x) \geqslant r$, δηλαδή $\{F \geqslant M+r\} \subseteq \{d_A \geqslant r\}$. Συνεπώς, από την ανισότητα της απόστασης από την κυρτή θήκη έχουμε ότι

$$P(F \geqslant M+r) \leqslant P(d_A \geqslant r) \leqslant \frac{1}{P(A)} e^{-\frac{r^2}{4}}.$$

Άρα,

$$\frac{1}{2} P(F \geqslant M+r) \leqslant P(A)P(F \geqslant M+r) \leqslant e^{-\frac{r^2}{4}},$$

δηλαδή

$$P(F \geqslant M+r) \leqslant 2e^{-\frac{r^2}{4}}. \tag{3.4.1}$$

Τώρα, θεωρώντας το $\tilde{A} = \{F \leqslant M-r\}$, εντελώς όμοια παίρνουμε ότι

$$P(F \leqslant M-r)P(F \geqslant M) = P(\tilde{A})P(F \geqslant M-r+r) \leqslant e^{-\frac{r^2}{4}},$$

δηλαδή

$$P(F \leqslant M-r) \leqslant 2e^{-\frac{r^2}{4}}. \tag{3.4.2}$$

Οι (3.4.1) και (3.4.2) μας δίνουν το ζητούμενο. □

Πόρισμα 3.4.4. *Έστω $(X, \mathcal{A}, P)$ ένας χώρος πιθανότητας γινόμενο με φορέα το $[0,1]^n$. Τότε, για κάθε συνάρτηση Lipschitz $F : X \to \mathbb{R}$ με $\|F\|_{\mathrm{Lip}} \leqslant 1$ και για κάθε $r \geqslant 0$ έχουμε ότι*

$$P(|F-M| \geqslant r) \leqslant 4e^{-\frac{r^2}{4n}},$$

όπου M είναι μια διάμεσος της F ως προς P.

Απόδειξη. Χωρίς περιορισμό της γενικότητας μπορούμε να δουλέψουμε στο $[0,1]^n$. Έστω $x, y \in [0,1]^n$. Τότε,

$$F(x) - F(y) \leqslant |x-y| \leqslant \sum_{i=1}^{n} |x_i - y_i| \leqslant \sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq y_i\}}.$$

Συνεπώς,

$$\frac{F(x)-F(y)}{\sqrt{n}} \leqslant \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \mathbf{1}_{\{x_i \neq y_i\}} = d_a(x,y),$$

όπου $a = \frac{1}{\sqrt{n}}(1, \ldots, 1)$ και $|a| = 1$. Θεωρούμε λοιπόν την $\tilde{F} = \frac{F}{\sqrt{n}}$, η οποία ικανοποιεί τις υποθέσεις της Πρότασης 3.4.3. Το ζητούμενο έπεται από την Πρόταση 3.4.3 παρατηρώντας ότι ο M είναι διάμεσος της F αν και μόνο αν ο $\frac{M}{\sqrt{n}}$ είναι διαμεσος της $\tilde{F}$. □

Λήμμα 3.4.5. *Αν $(X, \mathcal{A}, P)$ είναι ένας χώρος πιθανότητας και $F : X \to \mathbb{R}$ είναι μετρήσιμη συνάρτηση με*

$$P(|F-m| \geqslant r) \leqslant a e^{-\kappa r^2}$$

για κάθε $r \geqslant 0$ και κάποιον $m \in \mathbb{R}$, τότε

$$\Big| \int_X F \, dP - m \Big| \leqslant \frac{a\sqrt{\pi}}{2\sqrt{\kappa}}.$$

Απόδειξη. Έχουμε ότι

$$\Big|\int_X F\,dP - m\Big| = \Big|\int_X (F-m)\,dP\Big| \leqslant \int_X |F-m|\,dP$$
$$= \int_0^\infty P(|F-m| \geqslant r)\,dr \leqslant \int_0^\infty ae^{-\kappa r^2}\,dr = \frac{a\sqrt{\pi}}{2\sqrt{\kappa}}.$$

□

Πρόταση 3.4.6. *Έστω $(X,\mathcal{A},P)$ χώρος πιθανότητας γινόμενο και $F : X \to \mathbb{R}$.*

(α) *Αν η F ικανοποιεί τις συνθήκες της Πρότασης 3.4.3, τότε για κάθε $r > 8\sqrt{\pi}$ έχουμε ότι*

$$P\Big(\Big|F - \int_X F\,dP\Big| \geqslant r\Big) \leqslant 4e^{-\frac{r^2}{16}}.$$

(β) *Αν ο $(X,\mathcal{A},P)$ και η F ικανοποιούν τις υποθέσεις του Πορίσματος 3.4.4, τότε για $r > 8\sqrt{\pi}$ ισχύει ότι*

$$P\Big(\Big|F - \int_X F\,dP\Big| \geqslant r\Big) \leqslant 4e^{-\frac{r^2}{16n}}.$$

Απόδειξη. (α) Έστω M μια διάμεσος της F. Η Πρόταση 3.4.3 και το Λήμμα 3.4.5 μας δίνουν ότι

$$\Big|\int_X F\,dP - M\Big| \leqslant 4\sqrt{\pi}.$$

Οπότε, για $r > 8\sqrt{\pi}$ έχουμε ότι

$$P\Big(\Big|F - \int_X F\,dP\Big| \geqslant r\Big) \leqslant P(|F-M| \geqslant r/2) + P\Big(\Big|M - \int_X F\,dP\Big| \geqslant \frac{r}{2}\Big)$$
$$\leqslant 4e^{-\frac{r^2}{16}}.$$

(β) Έστω M μια διάμεσος της F. Από το Πόρισμα 3.4.4 και το Λήμμα 3.4.5 παίρνουμε ότι

$$\Big|\int_X F\,dP - M\Big| \leqslant 4\sqrt{\pi n}.$$

Οπότε, όπως πριν, έπεται το ζητούμενο. □

3.4.2 Ανισότητα Kahane-Khintchine

Οι συναρτήσεις Rademacher $r_i : E_2^n \to \mathbb{R}$, $i = 1,\dots,n$, ορίζονται ως εξής:

$$r_i(\epsilon) = \epsilon_i, \tag{3.4.3}$$

όπου $\epsilon = (\epsilon_1,\dots,\epsilon_n)$. Η $\{r_i\}_{i=1}^n$ είναι ορθοκανονική ακολουθία στον $L_2(E_2^n,\mu_n)$, όπου μ_n είναι το ομοιόμορφο μέτρο πιθανότητας στον E_2^n. Έπεται ότι, για κάθε ακολουθία $\{a_i\} \in \ell_2^n$,

$$\int_{E_2^n} \Big|\sum_i a_i r_i(\epsilon)\Big|^2 d\mu_n(\epsilon) = \sum_{i=1}^n a_i^2. \tag{3.4.4}$$

Θεώρημα 3.4.7 (Khintchine). *Υπάρχουν σταθερές $A_p, B_p > 0$ ($p > 0$) με την εξής ιδιότητα: για κάθε $n \in \mathbb{N}$ και για κάθε $a = (a_1, \ldots, a_n) \in \ell_2^n$,*

$$A_p \left(\sum_{i=1}^n a_i^2\right)^{1/2} \leqslant \left(\int_{E_2^n} \left|\sum_{i=1}^n a_i r_i(\epsilon)\right|^p d\mu_n(\epsilon)\right)^{1/p} \leqslant B_p \left(\sum_{i=1}^n a_i^2\right)^{1/2}. \tag{3.4.5}$$

Παρατήρηση 3.4.8. Ισοδύναμα, η ανισότητα του Khintchine γράφεται στη μορφή

$$A_p \Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_2} \leqslant \Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_p} \leqslant B_p \Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_2}. \tag{3.4.6}$$

Αν A_p^*, B_p^* είναι οι βέλτιστες σταθερές για τις οποίες ισχύει το Θεώρημα 3.4.7, από την ανισότητα του Hölder είναι φανερό ότι $A_p^* = 1$ αν $p \geqslant 2$ και $B_p^* = 1$ αν $0 < p \leqslant 2$. Οι ακριβείς τιμές των A_p^*, B_p^* έχουν υπολογιστεί από τους Szarek (A_1^*) και Haagerup (για κάθε p).

Η ανισότητα Kahane-Khintchine γενικεύει την ανισότητα του Khintchine.

Θεώρημα 3.4.9. *Υπάρχει σταθερά K ώστε για κάθε χώρο με νόρμα X, για κάθε $n \in \mathbb{N}$, για κάθε $x_1, \ldots, x_n \in X$ και για κάθε $p \geqslant 1$,*

$$\left(\int_{E_2^n} \|\sum_{i\leqslant n} \epsilon_i x_i\|^p d\mu_n\right)^{1/p} \leqslant 2\int_{E_2^n} \|\sum_{i\leqslant n} \epsilon_i x_i\| \, d\mu_n + K\sigma\sqrt{p}, \tag{3.4.7}$$

όπου

$$\sigma^2 = \sup\big\{\sum_{i\leqslant n} |x^*(x_i)|^2 : x^* \in X^*, \|x^*\| \leqslant 1\big\}. \tag{3.4.8}$$

Πρόταση 3.4.10. *Η ανισότητα Kahane-Khintchine γενικεύει την ανισότητα του Khintchine.*

Απόδειξη. Για $(X, \|\cdot\|) = (\mathbb{R}, |\cdot|)$ παίρνουμε ότι

$$\Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_p} \leqslant 2\Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_1} + K\sigma\sqrt{p}$$

για κάθε $p \geqslant 1$ και για κάθε $a_1, \ldots, a_n \in \mathbb{R}$. Όμως,

$$\Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_1} \leqslant \Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_2}$$

και

$$\sigma \leqslant \Big(\sum_{i=1}^n a_i^2\Big)^{1/2} = \Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_2}.$$

Οπότε,

$$\Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_p} \leqslant (2 + K\sqrt{p})\Big\|\sum_{i=1}^n a_i r_i\Big\|_{L_2},$$

και $B_p = 2 + K\sqrt{p}$.

Η άλλη ανισότητα προκύπτει άμεσα στην περίπτωση που $p \geqslant 2$, με $A_p = 1$, ενώ αν $1 \leqslant p < 2$, θεωρώντας την $f = \sum_{i=1}^n a_i r_i$ έχουμε ότι

$$\begin{aligned}\|f\|_{L_2}^2 &= \int_{E_2^n} |f|^2 \, d\mu_n(\epsilon) = \int_{E_2^n} |f|^{\frac{p}{2}} |f|^{2-\frac{p}{2}} \, d\mu_n(\epsilon) \\ &\leqslant \Big(\int_{E_2^n} |f|^p \, d\mu_n(\epsilon)\Big)^{1/2} \Big(\int_{E_2^n} |f|^{4-p} \, d\mu_n(\epsilon)\Big)^{1/2} \\ &= \|f\|_{L_p}^{\frac{p}{2}} \|f\|_{L_{4-p}}^{\frac{4-p}{2}}\end{aligned}$$

από την ανισότητα Cauchy-Schwarz. Όμως $4-p>2$, άρα $\|f\|_{L_{4-p}} \leqslant B_{4-p}\|f\|_{L_2}$. Συνεπώς,

$$\|f\|_{L_2}^2 \leqslant \|f\|_{L_p}^{\frac{p}{2}} B_{4-p}^{\frac{4-p}{2}} \|f\|_{L_2}^{\frac{4-p}{2}},$$

και τελικά

$$\|f\|_{L_p} \geqslant \Big(\frac{1}{B_{4-p}}\Big)^{\frac{4-p}{p}} \|f\|_{L_2}.$$

□

Η απόδειξη του Θεωρήματος 3.4.9 θα βασιστεί στο θεώρημα του Talagrand για τον διακριτό κύβο:

Για κάθε $A \subseteq E_2^n$,

$$\int_{E_2^n} \exp(\varphi_A^2/8) d\mu_n \leqslant \frac{1}{\mu_n(A)}$$

όπου

$$\varphi_A(x) = \inf\{\|x-y\|_2 : y \in \mathrm{conv}(A)\}.$$

Συνέπεια αυτού του θεωρήματος είναι η συγκέντρωση των κυρτών Lipschitz συναρτήσεων γύρω από τον μέσο Lévy τους.

Θεώρημα 3.4.11. *Θεωρούμε μια κυρτή Lipschitz συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ με σταθερά Lipschitz σ. Έστω M ένας μέσος Lévy της f στον E_2^n. Τότε, για κάθε $t > 0$ έχουμε*

$$\mu_n(\{|f-M| \geqslant t\}) \leqslant 4e^{-t^2/8\sigma^2}. \tag{3.4.9}$$

Απόδειξη. Για τον M ισχύουν οι $\mu_n(\{f \geqslant M\}) \geqslant 1/2$ και $\mu_n(\{f \leqslant M\}) \geqslant 1/2$.

Θέτουμε $A = \{f \leqslant M\}$. Αφού η f είναι κυρτή, για κάθε $y \in \mathrm{conv}(A)$ έχουμε $f(y) \leqslant M$. Αν λοιπόν $f(x) \geqslant M+t$ για κάποιο $x \in E_2^n$, τότε $f(x) \geqslant M+t \geqslant f(y)+t$ για κάθε $y \in \mathrm{conv}(A)$. Άρα, $\sigma\|x-y\|_2 \geqslant |f(x)-f(y)| \geqslant t$. Αυτό σημαίνει ότι

$$\varphi_A(x) \geqslant t/\sigma. \tag{3.4.10}$$

Από το Πόρισμα 3.3.5 και από την $\mu_n(A) \geqslant 1/2$ έχουμε

$$\begin{aligned}\mu_n(\{f \geqslant M+t\}) &\leqslant \mu_n(\{\varphi_A \geqslant t/\sigma\}) \leqslant \frac{1}{\mu_n(A)} e^{-t^2/8\sigma^2} \\ &\leqslant 2e^{-t^2/8\sigma^2}.\end{aligned} \tag{3.4.11}$$

Έστω $t > 0$ και $B = \{f \leqslant M - t\}$. Αν $u < t$, όπως πριν ελέγχουμε ότι

$$f(x) \geqslant M - t + u \implies \varphi_B(x) \geqslant u/\sigma, \tag{3.4.12}$$

απ' όπου παίρνουμε

$$\begin{aligned} \mu_n(\{f(x) \geqslant M\}) &\leqslant \mu_n(\{f(x) \geqslant M - t + u\}) \leqslant \mu_n(\{\varphi_B \geqslant u/\sigma\}) \\ &\leqslant \frac{1}{\mu_n(B)} e^{-u^2/8\sigma^2}. \end{aligned} \tag{3.4.13}$$

Όμως $1/2 \leqslant \mu_n(\{f(x) \geqslant M\})$, άρα $\mu_n(B) \leqslant 2e^{-u^2/8\sigma^2}$. Αφήνοντας το u να τείνει στο t παίρνουμε

$$\mu_n(B) \leqslant 2e^{-t^2/8\sigma^2}. \tag{3.4.14}$$

Συνδυάζοντας τα παραπάνω βλέπουμε ότι

$$\begin{aligned} \mu_n(\{|f - M| > t\}) &= \mu_n(\{f \geqslant M + t\}) + \mu_n(\{f \leqslant M - t\}) \\ &\leqslant 2e^{-t^2/8\sigma^2} + 2e^{-t^2/8\sigma^2} \\ &= 4e^{-t^2/8\sigma^2}. \end{aligned} \tag{3.4.15}$$

□

Πόρισμα 3.4.12. *Έστω X χώρος με νόρμα και $(x_i)_{i \leqslant n}$ ακολουθία διανυσμάτων στον X. Θέτουμε*

$$\sigma^2 = \sup\Big\{ \sum_{i \leqslant n} |x^*(x_i)|^2 : x^* \in X^*, \|x^*\| \leqslant 1 \Big\}. \tag{3.4.16}$$

Αν M είναι μέσος Lévy της $\| \sum_{i \leqslant n} \epsilon_i x_i \|$ στον E_2^n τότε, για κάθε $t \geqslant 0$,

$$\mu_n\left(\{ | \, \| \sum_{i \leqslant n} \epsilon_i x_i \| - M \, | \geqslant t \} \right) \leqslant 4e^{-t^2/8\sigma^2}. \tag{3.4.17}$$

Απόδειξη. Θεωρούμε την $f(u) = \| \sum_{i \leqslant n} u_i x_i \|$, $u \in \mathbb{R}^n$. Χρησιμοποιώντας τις ιδιότητες της νόρμας ελέγχουμε εύκολα ότι η f είναι κυρτή συνάρτηση. Έστω $x^* \in X^*$ με $\|x^*\| \leqslant 1$ και $u, v \in \mathbb{R}^n$. Από την ανισότητα Cauchy-Schwarz έχουμε

$$\begin{aligned} \left| x^*\Big(\sum_{i \leqslant n} u_i x_i - \sum_{i \leqslant n} v_i x_i \Big) \right| &= \left| \sum_{i \leqslant n} u_i x^*(x_i) - \sum_{i \leqslant n} v_i x^*(x_i) \right| \\ &= \left| \sum_{i \leqslant n} (u_i - v_i) x^*(x_i) \right| \\ &\leqslant \left(\sum_{i \leqslant n} |x^*(x_i)|^2 \right)^{1/2} \left(\sum_{i \leqslant n} (u_i - v_i)^2 \right)^{1/2} \\ &\leqslant \sigma \|u - v\|_2. \end{aligned} \tag{3.4.18}$$

Από το Θεώρημα Hahn-Banach συμπεραίνουμε ότι

$$|f(u) - f(v)| \leqslant \Big\| \sum_{i\leqslant n} u_i x_i - \sum_{i\leqslant n} v_i x_i \Big\| \leqslant \sigma \|u - v\|_2, \tag{3.4.19}$$

επομένως η f είναι Lipschitz συνεχής με σταθερά σ. Εφαρμόζοντας το Θεώρημα 3.4.11 για την f έχουμε το ζητούμενο. □

Μπορούμε τώρα να δώσουμε μία απόδειξη της ανισότητας Khintchine-Kahane με βέλτιστη εξάρτηση από το p.

Απόδειξη του θεωρήματος 3.4.9. Θεωρούμε την $f : E_2^n \to \mathbb{R}^+$ με

$$f(\epsilon_1, \ldots, \epsilon_n) = \Big| \, \| \sum_{i\leqslant n} \epsilon_i x_i \| - M \, \Big|. \tag{3.4.20}$$

Από το Πόρισμα 3.4.12, κάνοντας την αλλαγή μεταβλητής $x = t^2/8\sigma^2$ έχουμε

$$\begin{aligned} \int_{E_2^n} \Big| \, \| \sum_{i\leqslant n} \epsilon_i x_i \| - M \, \Big|^p d\mu_n(\epsilon) &= p \int_0^\infty t^{p-1} \mu_n(\epsilon : | \| \sum_{i\leqslant n} \epsilon_i x_i \| - M | \geqslant t) dt \\ &\leqslant 4 \int_0^\infty p t^{p-1} e^{-t^2/8\sigma^2} dt \\ &= 2^{p+1} p (\sqrt{2}\sigma)^p \int_0^\infty e^{-x} x^{p/2-1} dx \leqslant (K\sigma\sqrt{p})^p. \end{aligned}$$

Άρα,

$$\left(\int_{E_2^n} | \, \| \sum_{i\leqslant n} \epsilon_i x_i \| - M |^p d\mu_n(\epsilon) \right)^{1/p} \leqslant K\sigma\sqrt{p}. \tag{3.4.21}$$

Από την τριγωνική ανισότητα,

$$\left(\int_{E_2^n} \| \sum_{i\leqslant n} \epsilon_i x_i \|^p d\mu_n(\epsilon) \right)^{1/p} \leqslant M + K_1 \sigma p^{1/2} \tag{3.4.22}$$

για κάθε $p \geqslant 1$. Τέλος, παρατηρούμε ότι $M \leqslant 2 \int_{E_2^n} \| \sum_{i\leqslant n} \epsilon_i x_i \| \, d\mu_n$ από την ανισότητα του Markov. □

3.4.3 Χρωματικός αριθμός τυχαίων γραφημάτων

Ορισμός 3.4.13. Για λόγους απλότητας θα ονομάζουμε *γράφημα με κορυφές τα στοιχεία του* $V = \{1, \ldots, n\}$ ένα υποσύνολο του $E_0 = \{(i,j) \in V \times V \mid i < j\}$. Αν $(i,j) \in G$, τότε θα λέμε ότι τα i και j *συνδέονται με μια ακμή.*

Ένα $I \subseteq V$ θα λέγεται *ανεξάρτητο* αν για κάθε $i, j \in I$ με $i \neq j$ ισχύει ότι $(i,j) \notin G$, δηλαδή αν κάθε δύο διακεκριμένα στοιχεία του I δεν συνδέονται με ακμή. Ο *χρωματικός αριθμός* $\chi(G, A)$ ενός $A \subseteq V$ είναι ο μικρότερος αριθμός ανεξάρτητων συνόλων που καλύπτουν το A.

Παρατήρηση 3.4.14. Με άλλα λόγια, ο χρωματικός αριθμός $\chi(G,A)$ είναι ο ελάχιστος αριθμός χρωμάτων με τόν οποίο μπορούμε να χρωματίσουμε τα στοιχεία του A έτσι ώστε κάθε δύο στοιχεία που έχουν το ίδιο χρώμα να μην συνδέονται με ακμή.

Θέτουμε τώρα

$$\chi(G,m) = \inf\{\chi(G,A) \mid \mathrm{card}(A) = m\}.$$

Ορισμός 3.4.15. Δοθέντος $0 < p < 1$, ορίζουμε το *τυχαίο γράφημα* $G = G(n,p)$ τοποθετώντας στο G την ακμή $(i,j) \in E_0$ με πιθανότητα p, ανεξάρτητα από τις υπόλοιπες ακμές.

Θεώρημα 3.4.16. *Έστω $k \in \mathbb{N}$ και $t > 0$. Τότε, υπάρχει $a \in \mathbb{Z}$ ώστε*

$$\begin{aligned} &P(\chi(G(n,p),m) \in [a-k,a]) \\ &\qquad \geqslant 1 - 2e^{-\frac{t^2}{8}} - P(\sup\{\chi(G(n,p),F) \mid F \subseteq V,\ \mathrm{card}(F) \leqslant t\sqrt{m}\} > k). \end{aligned}$$

Απόδειξη. Θέτουμε

$$b = P(\sup\{\chi(G(n,p),F) \mid F \subseteq V,\ \mathrm{card}(F) \leqslant t\sqrt{m}\} > k)$$

και a τον μέγιστο ακέραιο για τον οποίο

$$P(\chi(G(n,p),m) \geqslant a) \geqslant e^{-\frac{t^2}{8}} + b \quad \text{και} \quad P(\chi(G(n,p),m) > a) < e^{-\frac{t^2}{8}} + b. \tag{3.4.23}$$

Για να εφαρμόσουμε την ανισότητα απόστασης από την κυρτή θήκη πρέπει να αναπαραστήσουμε τον παραπάνω χώρο πιθανότητας σαν χώρο γινόμενο.

Για $2 \leqslant j \leqslant n$ θέτουμε $\Omega_j = \{0,1\}^{j-1}$ και $\Omega' = \prod_{2\leqslant j\leqslant n} \Omega_j$. Γράφουμε το τυχόν στοιχείο $\omega \in \Omega'$ ως $\omega = (\omega_j)_{2\leqslant j\leqslant n}$, όπου $\omega_j = (\omega_{i,j})_{1\leqslant i\leqslant j-1} \in \Omega_j$.

Στο ω αντιστοιχούμε το γράφημα $G(\omega)$ έτσι ώστε για $i < j$ να έχουμε $(i,j) \in G(\omega)$ αν και μόνο αν $\omega_{i,j} = 1$. Συνεπώς, το $G(n,p)$ έχει την ίδια κατανομή με το $G(\omega)$ για το σύνηθες μέτρο γινόμενο στο Ω', που δίνει βάρος p στο 1 και $1-p$ στο 0.

Έστω λοιπόν τώρα

$$A = \big\{\omega \in \Omega' \mid \chi(G(\omega),m) \geqslant a \ \text{ και } \ \sup\{\chi(G(\omega),F) \mid F \subseteq V,\ \mathrm{card}(F) \leqslant t\sqrt{m}\} \leqslant k\big\}.$$

Τότε, από την (3.4.23) και τον ορισμό του β, έχουμε ότι

$$P(A) \geqslant e^{-\frac{t^2}{8}}.$$

Θέτουμε τώρα

$$B = \Big\{\omega \in \Omega' \mid \text{ για κάθε } (a_j)_{2\leqslant j\leqslant n} \text{ υπάρχει } \omega' \in A \text{ ώστε } \sum_{j=2}^{n} a_j \mathbf{1}_{\{\omega_j \neq \omega'_j\}} \leqslant t\sqrt{\sum_{j=2}^{n} a_j^2}\Big\}.$$

Τότε, από το Λήμμα 3.3.21 έχουμε ότι αν $A_t = \{\omega \in \Omega' \mid f_C(A,x) \leqslant t\}$ ισχύει ότι $A_t \subseteq B$. Οπότε,

$$P(B) \geqslant P(A_t) \geqslant 1 - \frac{1}{P(A)} e^{-\frac{t^2}{4}} \geqslant 1 - e^{-\frac{t^2}{8}}$$

από το Θεώρημα 3.3.20.

Ισχυρισμός. Αν $\omega \in B$, τότε $\chi(G(\omega), m) \geqslant a - k$.

Πράγματι, έστω $\omega \in B$. Θέτουμε $r = \chi(G(\omega), m)$. Έστω $F \subseteq V$ με $\mathrm{card}(F) = m$, ώστε $\chi(G(\omega), F) = r$. Θέτουμε

$$a_j = \begin{cases} 1 & , \text{ αν } j \in F \text{ για κάθε } 2 \leqslant j \leqslant n, \\ 0 & , \text{ αλλιώς,} \end{cases}$$

και από τον ορισμό του B υπάρχει $\omega' \in A$ ώστε αν $J = \{j \in F : \omega_j \neq \omega'_j\}$ τότε $\mathrm{card}(J) \leqslant t\sqrt{m}$.

Όμως, προφανώς ισχύει ότι

$$\chi(G(\omega'), F \setminus J) = \chi(G(\omega), F \setminus J) \leqslant r.$$

Συνεπώς, αφού $\omega' \in A$ έχουμε ότι

$$\begin{aligned} a &\leqslant \chi(G(\omega'), m) \leqslant \chi(G(\omega'), F) \leqslant \chi(G(\omega'), F \setminus J) + \chi(G(\omega'), J) \\ &\leqslant r + \chi(G(\omega'), J) \leqslant r + \sup\{\chi(G(\omega'), K) \mid \mathrm{card}(K) \leqslant t\sqrt{m}\} \leqslant r + k, \end{aligned}$$

δηλαδή $r \geqslant a - k$, άρα $\chi(G(\omega), m) \geqslant a - k$. Οπότε έχουμε ότι

$$B \subseteq \{\omega \in \Omega' \mid \chi(G(\omega), m) \geqslant a - k\}.$$

Συνεπώς,

$$\begin{aligned} 1 - e^{-\frac{t^2}{8}} &\leqslant P(B) \leqslant P(\chi(G(\omega), m) \geqslant a - k\} \\ &= P(\chi(G(\omega), m) \in [a - k, a]) + P(\chi(G(\omega), m) > a) \\ &\leqslant P(\chi(G(\omega), m) \in [a - k, a]) + e^{-\frac{t^2}{8}} + b, \end{aligned}$$

άρα

$$P(\chi(G(\omega), m) \in [a - k, a]) \geqslant 1 - 2e^{-\frac{t^2}{8}} - b,$$

και δείξαμε αυτό που θέλαμε. □

Ορισμός 3.4.17. Έστω $r \in \mathbb{N}$ και $\ell = (i, j) \in E_0$. Θέτουμε $N(G, e)$ το πλήθος των ανεξάρτητων υποσυνόλων του V πληθικότητας r που περιέχουν τα i, j, δηλαδή

$$N(G, \ell) = \mathrm{card}\{F \subseteq V \mid F \text{ ανεξάρτητο}, \ \mathrm{card}(F) = r, \ i, j \in F\}.$$

Θεώρημα 3.4.18. *Έστω $u > 0$ ώστε*

$$P\Big(u\sqrt{\sum_{\ell \in E_0} N(G(n, p), \ell)^2} \leqslant \sum_{\ell \in E_0} N(G(m.p), \ell)\Big) > \frac{1}{2}.$$

Τότε,

$$P(\text{ το } G(n, p) \text{ δεν περιέχει ανεξάρτητο σύνολο πληθικότητας } r) \leqslant 2\exp\Big(-\frac{u^2}{r^2(r-1)^2}\Big).$$

Απόδειξη. Θεωρούμε το $\Omega = \{0, 1\}$ εφοδιασμενο με το μέτρο πιθανότητας που δίνει βάρος p στο 1 και $1 - p$ στο 0. Έστω P το αντίστοιχο μέτρο γινόμενο στον Ω^{E_0}.

Για $x \in (x_\ell)_{\ell \in E_0} \in \Omega^{E_0}$ ορίζουμε $G(x)$ από τη σχέση

$$\ell = (i, j) \in G(x) \quad \Longleftrightarrow \quad x_\ell = 1.$$

Τότε το $G(x)$ έχει την ίδια κατανομή με το $G(n, p)$. Θεωρούμε το

$$A = \{y \in \Omega^{E_0} \mid \text{ το } G(n, p) \text{ δεν περιέχει ανεξάρτητο σύνολο πληθικότητας } r\}.$$

Θέτουμε $t_0 = 2\sqrt{\log\left(\frac{2}{P(A)}\right)}$. Τότε,

$$P(A_{t_0}) \geqslant 1 - \frac{1}{P(A)} e^{-\frac{t_0^2}{4}} \geqslant \frac{1}{2}$$

από το Θεώρημα 3.3.20. Οπότε, προφανώς,

$$A_{t_0} \cap \Big\{x \in \Omega^{E_0} \mid u\sqrt{\sum_{\ell \in E_0} N(G(x), \ell)^2} \leqslant \sum_{\ell \in E_0} N(G(x), \ell)\Big\} \neq \emptyset.$$

Συνεπώς, λόγω του Λήμματος 3.3.21, υπάρχει $x \in \Omega^{E_0}$ με

$$a\sqrt{\sum_{\ell \in E_0} N(G(x), \ell)^2} \leqslant \sum_{\ell \in E_0} N(G(x), \ell) \tag{3.4.24}$$

ώστε για κάθε $(a_\ell)_{\ell \in E_0}$ να υπάρχει $y \in A$ με

$$\sum_{\ell \in E_0} a_\ell \mathbf{1}_{\{x_\ell \neq y_\ell\}} \leqslant t_0 \sqrt{\sum_{\ell \in E_0} a_\ell^2}.$$

Θέτουμε τώρα $a_\ell = N(G(x), \ell)$ για κάθε $\ell \in E_0$. Υπάρχει $y \in A$ ώστε αν $C = \{\ell \in E_0 \mid x_\ell \neq y_\ell\}$ τότε

$$\sum_{\ell \in C} N(G(x), \ell) \leqslant t_0 \sqrt{\sum_{\ell \in C} N(G(x), \ell)^2} \leqslant \frac{t_0}{u} \sum_{\ell \in E_0} N(G(x), \ell) \tag{3.4.25}$$

λόγω της (3.4.24).

Ο συνολικός αριθμός ανεξάρτητων συνόλων του $G(x)$ πληθικότητας r είναι

$$N = \Big(\frac{r(r-1)}{2}\Big)^{-1} \sum_{\ell \in E_0} N(G(x), \ell). \tag{3.4.26}$$

Πρέπει συνεπώς να έχουμε

$$N \leqslant \sum_{\ell \in C} N(G(x), \ell), \tag{3.4.27}$$

διότι αλλιώς θα υπήρχε ανεξάρτητο σύνολο πληθικότητας r του $G(x)$ που δεν θα περιείχε καμία ακμή του C και συνεπώς θα ήταν ανεξάρτητο σύνολο του $G(y)$ που είναι άτοπο αφού $y \in A$. Από τις (3.4.25), (3.4.26) και (3.4.27) παίρνουμε ότι

$$t_0 \geqslant \frac{2u}{r(r-1)},$$

από όπου έπεται το ζητούμενο. □

Παρατήρηση 3.4.19. Μια μέθοδος για να εκμεταλλευτούμε το Θεώρημα 3.4.18 είναι η εξής: Βρίσκουμε $u_1, u_2 > 0$ ώστε αν

$$A_1 = \Big\{ \sum_{\ell \in E_0} N(G(n,p)), \ell) \geqslant u_1 \Big\}$$

και

$$A = \Big\{ \sum_{\ell \in E_0} N(G(n,p), \ell)^2 \leqslant u_2^2 \Big\}$$

να ισχύει ότι $P(A_1) > \frac{3}{4}$ και $P(A_2) > \frac{3}{4}$, οπότε $P(A_1 \cap A_2) > \frac{1}{2}$. Τότε, αν $u = u_1/u_2$ έχουμε ότι

$$P\Big(u \sqrt{\sum_{\ell \in E_0} N(G(n,p), \ell)^2} \leqslant \sum_{\ell \in E_0} N(G(m.p), \ell) \Big) \geqslant P(A_1 \cap A_2) > \frac{1}{2}.$$

3.4.4 Η μεγαλύτερη αύξουσα υπακολουθία

Ορίζουμε $L_n : [0,1]^n \to \mathbb{N}$ με $L_n(x) = L_n(x_1, \dots, x_n) = s$, όπου s είναι ο μεγαλύτερος φυσικός αριθμός για τον οποίο μπορούμε να βρούμε $1 \leqslant i_1 < \cdots < i_s \leqslant n$ ώστε $x_{i_1} \leqslant \cdots \leqslant x_{i_s}$, δηλαδή το $L_n(x_1, \dots, x_n)$ συμβολίζει το μήκος της μεγαλύτερης αύξουσας υπακολουθίας της $x_1, \dots, x_n$.

Θέτουμε ακόμη, για $a > 0$,

$$A(a) = \{x \in \Omega \mid L_n(x) \leqslant a\}.$$

Λήμμα 3.4.20. *Για κάθε $x \in \Omega^n$ ισχύει ότι*

$$a \geqslant L_n(x) - f_c(A(a), x)\sqrt{L_n(x)}.$$

Ειδικότερα, για κάθε $v > 0$, αν $L_n(x) \geqslant a + v$, τότε $f_c(A(a), x) \geqslant \frac{v}{\sqrt{a+v}}$.

Απόδειξη. Για απλότητα στο συμβολισμό θα γράφουμε $b = L_n(x)$. Εξ' ορισμού, μπορούμε να βρούμε $I = \{i_1 < \cdots < i_b\} \subseteq \{1, \dots, n\}$ ώστε $x_{i_1} \leqslant \cdots \leqslant x_{i_b}$. Για κάθε $1 \leqslant i \leqslant n$ θέτουμε $a_i = 1$ αν $i \in I$ και $a_i = 0$ αν $i \in [n] \setminus I$. Από το Λήμμα 3.3.21 υπάρχει $y \in A(a)$ έτσι ώστε αν $J = \{i \in I \mid x_i \neq y_i\}$ να ισχύει ότι $\mathrm{card}(J) \leqslant f_c(A(a), x)\sqrt{b}$, αφού $x \in A(a)_{f_c(A(a),x)}$.

Επομένως, η $(x_i)_{i \in I \setminus J}$ είναι μια αύξουσα υπακολουθία της $y = (y_1, \dots, y_n)$.

Τώρα, αφού $y \in A(a)$, έχουμε ότι $\mathrm{card}(I \setminus J) \leqslant L_n(y) \leqslant a$, όπου

$$\mathrm{card}(I \setminus J) = \mathrm{card}(I) - \mathrm{card}(J) \geqslant L_n(x) - f_c(A(a), x)\sqrt{L_n(x)}.$$

Συνεπώς, πράγματι

$$a \geqslant L_n(x) - f_c(A(a), x)\sqrt{L_n(x)}.$$

Για την άλλη ανισότητα παρατηρούμε ότι

$$f_c(A(a), x) \geqslant \frac{L_n(x) - a}{\sqrt{L_n(x)}}$$

και ότι η συνάρτηση $u \mapsto \frac{u-a}{\sqrt{u}}$ είναι αύξουσα, και έχουμε το ζητούμενο. $\square$

Θεώρημα 3.4.21. *Έστω M μια διάμεσος της L_n. Τότε, για κάθε $\varepsilon > 0$ έχουμε ότι*

$$P(L_n \geqslant M + \varepsilon) \leqslant 2\exp\Big(-\frac{\varepsilon^2}{4(M+\varepsilon)}\Big)$$

και

$$P(L_n \leqslant M - \varepsilon) \leqslant 2\exp\Big(-\frac{\varepsilon^2}{4M}\Big).$$

Απόδειξη. Από το Λήμμα 3.4.20 για $a = M$ και $v = \varepsilon$ παίρνουμε ότι

$$P(L_n \geqslant M + \varepsilon) \leqslant 1 - P\Big(A(M)_{\frac{\varepsilon}{\sqrt{M+\varepsilon}}}\Big) \leqslant \frac{1}{P(A(M))}\exp\Big(-\frac{\varepsilon^2}{4(M+\varepsilon)}\Big) \leqslant 2\exp\Big(-\frac{\varepsilon^2}{4(M+\varepsilon)}\Big),$$

χρησιμοποιώντας το Θεώρημα 3.3.20 και το γεγονός ότι

$$P(A(M)) = P(L_n \leqslant M) \geqslant \frac{1}{2}.$$

Για την δεύτερη ανισότητα θέτουμε $a = M - \varepsilon$ και $v = \varepsilon$. Τότε, από το Λήμμα 3.4.20 παίρνουμε ότι

$$L_n(x) \geqslant M \quad \Longrightarrow \quad f_c(A(M-\varepsilon), x) \geqslant \frac{\varepsilon}{\sqrt{M}}.$$

Επομένως,

$$\frac{1}{2} \leqslant P(L_n \geqslant M) \leqslant P\Big(f_c(A(M-\varepsilon), x) \geqslant \frac{\varepsilon}{\sqrt{M}}\Big).$$

Όπως πριν, από το Θεώρημα 3.3.20 έχουμε ότι

$$\frac{1}{2} \leqslant P\Big(f_c(A(M-\varepsilon), x) \geqslant \frac{\varepsilon}{\sqrt{M}}\Big) \leqslant \frac{1}{P(A(M-\varepsilon))}\exp\Big(-\frac{\varepsilon^2}{4M}\Big).$$

Συνεπώς, $P(A(M-\varepsilon)) \leqslant 2\exp(-\varepsilon^2/(4M))$ και έχουμε το ζητούμενο. □

3.4.5 Η μεγαλύτερη κοινή υπακολουθία

Ορίζουμε $L_{n,m} : [0,1]^n \times [0,1]^m \to \mathbb{N}_0$ με $L_{n,m}(x,y) = L_{n,m}(x_1,\ldots,x_n,y_1,\ldots,y_m) = s$, όπου s είναι ο μεγαλύτερος φυσικός αριθμός για τον οποίο υπάρχουν $1 \leqslant i_1 < \cdots < i_s \leqslant n$ και $1 \leqslant j_1 < \cdots < j_s \leqslant m$ ώστε $x_{i_k} = y_{i_k}$ για κάθε $1 \leqslant k \leqslant s$, και $L_{n,m}(x,y) = 0$ όταν δεν υπάρχει τέτοιος φυσικός.

Θεώρημα 3.4.22. *Έστω M μια διάμεσος της $L_{n,m}$. Τότε, για κάθε $\varepsilon > 0$ έχουμε ότι*

$$P(L_{n,m} \geqslant M + \varepsilon) \leqslant 2\exp\Big(-\frac{\varepsilon^2}{32(M+\varepsilon)}\Big)$$

και

$$P(L_{n,m} \leqslant M - \varepsilon) \leqslant 2\exp\Big(-\frac{\varepsilon^2}{32M}\Big).$$

Απόδειξη. Θεωρούμε $x = (x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+m}) \in [0,1]^{n+m} = \Omega$ και γράφουμε $L(x) = L_{n,m}(x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+m})$. Για κάθε $a > 0$ θέτουμε $A(a) = \{x \in \Omega \mid L(x) \leqslant a\}$.

Ισχυρισμός. Για κάθε $x \in \Omega$ ισχύει $a \geqslant L(x) - 2\sqrt{2}f_c(A(a), L(x))\sqrt{L(x)}$.

Πράγματι, αν $L(x) = 0$ τότε προφανώς ισχύει. Αν $b = L(x) \in \mathbb{N}$ τότε υπάρχουν $1 \leqslant i_1 < \cdots < i_b \leqslant n < i_{b+1} < \cdots < i_{2b} \leqslant n+m$ ώστε $x_{i_k} = x_{i_{b+k}}$ για κάθε $1 \leqslant k \leqslant b$. Θεωρούμε το σύνολο

$$I = \{i_k \mid 1 \leqslant k \leqslant 2b\}.$$

Από το Λήμμα 3.3.21 για $a_i = 1$, $i \in I$, υπάρχει $y \in A(a)$ ώστε

$$\mathrm{card}(\{i \in I \mid x_i \neq y_i\}) \leqslant f_c(A(a), x)\sqrt{2b}. \tag{3.4.28}$$

Θεωρούμε τώρα το σύνολο

$$J = \{1 \leqslant k \leqslant b \mid x_{i_k} = y_{i_k}, x_{i_{k+b}} = y_{i_{k+b}}\}.$$

Παρατηρούμε ότι

$$\mathrm{card}(J) \geqslant b - 2\mathrm{card}(\{i \in I \mid x_i \neq y_i\}) \geqslant b - 2\sqrt{2b} f_c(A(a), x)$$

λόγω της (3.4.28). Όμως,

$$L(y) \geqslant \mathrm{card}(J), \tag{3.4.29}$$

αφού για κάθε $j \in J$ έχουμε ότι $y_{i_j} = y_{i_{j+b}}$. Επίσης,

$$L(y) \leqslant a, \tag{3.4.30}$$

αφού $y \in A(a)$. Συνεπώς, από τις (3.4.28), (3.4.29) και (3.4.30) παίρνουμε $a \geqslant b - 2\sqrt{2b} f_c(A(a), x)$, δηλαδή τον ισχυρισμό.

Οι δύο ανισότητες του θεωρήματος προκύπτουν ακριβώς όπως στην απόδειξη του Θεωρήματος 3.4.21. $\square$

ΚΕΦΑΛΑΙΟ 4

Η L^1-L^2 ανισότητα για την διασπορά

4.1 Εισαγωγή

Το πρόβλημα από το οποίο προέκυψε η L^1-L^2 ανισότητα του Talagrand αφορά μονότονα υποσύνολα του $\{0,1\}^n$. Λέμε ότι το $A \subseteq \{0,1\}^n$ είναι *μονότονο* αν για κάθε $x \in A$ και κάθε $y \in \{0,1\}^n$ ισχύει ότι

$$\text{«αν για κάθε } 1 \leqslant i \leqslant n, \text{ έχουμε } \; y_i \geqslant x_i \; \text{ τότε } \; y \in A\text{».}$$

Θεωρούμε $0 < p < 1$ και το μέτρο γινόμενο μ_p^n στο $\{0,1\}^n$ το οποίο δίνει βάρος 0 με πιθανότητα $1-p$ και βάρος 1 με πιθανότητα p, δηλαδή

$$\mu_p^n(\{x\}) = (1-p)^{n-k}p^k \;\; \text{όπου} \;\; k = \mathrm{card}\{1 \leqslant i \leqslant n : x_i = 1\}.$$

Για ευκολία θα χρησιμοποιούμε το συμβολισμό μ_p αντί για μ_p^1. Αποδεικνύεται (δείτε το Θεώρημα 4.1.1 πιο κάτω) ότι αν το $A \subseteq \{0,1\}^n$ είναι μονότονο τότε η συνάρτηση $p \mapsto \mu_p^n(A)$ είναι αύξουσα. Η συμπεριφορά της συνάρτησης $p \mapsto \mu_p^n(A)$ παρουσιάζει ενδιαφέρον στη θεωρία διήθησης και στη θεωρία των τυχαίων γραφημάτων. Για πολλά σημαντικά σύνολα παρατηρείται ένα «φαινόμενο κατωφλίου», δηλαδή το $\mu_p^n(A)$ κάνει άλμα από τιμές κοντά στο 0 σε τιμές κοντά στο 1 μέσα σε ένα μικρό διάστημα τιμών του p. Συνθήκες κάτω από τις οποίες εμφανίζεται αυτό το φαινόμενο έχουν ανακαλυφθεί από τον Margulis (βλ. [57] και, επίσης, [92]). Διαισθητικά, παρατηρούμε πάντα φαινόμενο κατωφλίου εκτός αν το A προσδιορίζεται από πολύ λίγες συντεταγμένες. Ο Russo απέδειξε στο [81] ότι, ακόμα ισχυρότερα, έχουμε φαινόμενο κατωφλίου αν το A «εξαρτάται λίγο» από κάθε δεδομένη συντεταγμένη. Η δουλειά του απλουστεύτηκε και ισχυροποιήθηκε από τον Talagrand. Για κάθε $x \in \{0,1\}^n$ και κάθε $1 \leqslant i \leqslant n$ συμβολίζουμε με $U_i(x)$ το σημείο που προκύπτει από το x αν αντικαταστήσουμε τη συντεταγμένη x_i με $1-x_i$ και αφήσουμε τις υπόλοιπες συντεταγμένες αμετάβλητες. Ορίζουμε επίσης

$$A_i = \{x \in \{0,1\}^n : x \in A, U_i(x) \notin A\}.$$

Δείχνουμε πρώτα τον τύπο των Margulis-Russo, από τον οποίο ειδικότερα έπεται ότι για μονότονα $A \subseteq \{0,1\}^n$ η συνάρτηση $p \mapsto \mu_p^n(A)$ είναι αύξουσα.

Θεώρημα 4.1.1 (Margulis-Russo). *Αν A είναι ένα μονότονο υποσύνολο του $\{0,1\}^n$ τότε*

$$\frac{d\mu_p^n(A)}{dp} = \frac{1}{p}\sum_{i=1}^n \mu_p^n(A_i).$$

Θα χρειαστούμε ένα λήμμα. Στα επόμενα, αν $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ τότε θέτουμε

$$|x| = \sum_{i=1}^n x_i = \mathrm{card}\{1 \leqslant i \leqslant n : x_i = 1\}.$$

Λήμμα 4.1.2. *Αν $f : \{0,1\}^n \to \{0,1\}$ και $g_f : (0,1) \to \mathbb{R}$ είναι η συνάρτηση*

$$g_f(p) = \int_{\{0,1\}^n} f\, d\mu_p^n,$$

τότε η g_f είναι παραγωγίσιμη και

$$g_f'(p) = \frac{1}{p(1-p)}\sum_{i=1}^n \int_{\{0,1\}^n} (x_i - p) f(x)\, d\mu_p^n(x).$$

Απόδειξη. Έχουμε ότι

$$g_f(p) = \sum_{x \in \{0,1\}^n} f(x) p^{|x|}(1-p)^{n-|x|}.$$

Έπεται ότι η g_f είναι παραγωγίσιμη, και

$$\begin{aligned} g_f'(p) &= \frac{1}{p}\sum_{x\in\{0,1\}^n} |x| f(x) p^{|x|}(1-p)^{n-|x|} - \frac{1}{1-p}\sum_{x\in\{0,1\}^n}(n-|x|)f(x)p^{|x|}(1-p)^{n-|x|} \\ &= \frac{1}{p}\int_{\{0,1\}^n} |x| f(x)\, d\mu_p^n(x) - \frac{1}{1-p}\int_{\{0,1\}^n}(n-|x|)f(x)\, d\mu_p^n(x) \\ &= \frac{1}{p(1-p)}\int_{\{0,1\}^n}(|x| - np) f(x)\, d\mu_p^n(x) \\ &= \frac{1}{p(1-p)}\sum_{i=1}^n \int_{\{0,1\}^n}(x_i - p) f(x)\, d\mu_p^n(x). \end{aligned}$$

□

Απόδειξη του Θεωρήματος 4.1.1. Παρατηρούμε ότι αν $f = \mathbf{1}_A$ τότε $g_f = \mu_p^n(A)$. Για $x \in \{0,1\}^n$ και για κάθε $1 \leqslant i \leqslant n$ ορίζουμε

$$x^{(i)} = (x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) \quad \text{και} \quad x_{(i)} = (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n).$$

Επίσης, για κάθε $1 \leqslant i \leqslant n$ ορίζουμε

$$\mathrm{Piv}_i(A) = \{x \in \{0,1\}^n : x^{(i)} \in A \ \text{ και } \ x_{(i)} \notin A\}.$$

Παρατηρούμε ότι

$$\begin{aligned}\int_{\{0,1\}^n}(x_i-p)f(x)\cdot \mathbf{1}_{\mathrm{Piv}_i(A)^c}(x)\,d\mu_p^n(x) &= \int_{\{0,1\}^n}(x_i-p)\mathbf{1}_{A\cap \mathrm{Piv}_i(A)^c}(x)\,d\mu_p^n(x)\\ = \int_{\{0,1\}^{n-1}}[p(1-p)\mathbf{1}_{A\cap \mathrm{Piv}_i(A)^c}(x^{(i)}) &- p(1-p)\mathbf{1}_{A\cap \mathrm{Piv}_i(A)^c}(x_{(i)})]\,d\mu_p^{n-1}=0,\end{aligned}$$

αφού λόγω της μονοτονίας του A έχουμε

$$\mathbf{1}_{A\cap \mathrm{Piv}_i(A)^c}(x^{(i)}) = \mathbf{1}_{A\cap \mathrm{Piv}_i(A)^c}(x_{(i)}) = \mathbf{1}_A(x_{(i)}).$$

Οπότε, για κάθε $1\leqslant i\leqslant n$ έχουμε ότι

$$\int_{\{0,1\}^n}(x_i-p)f(x)\,d\mu_p^n(x) = \int_{\{0,1\}^n}(x_i-p)f(x)\mathbf{1}_{\mathrm{Piv}_i(A)}(x)\,d\mu_p^n(x).$$

Όμως, αν $x\in \mathrm{Piv}_i(A)$ τότε $f(x)=\mathbf{1}_A(x)=0$ αν $x_i=0$ και $f(x)=1$ αν $x_i=1$. Συνεπώς,

$$\begin{aligned}\int_{\{0,1\}^n}(x_i-p)f(x)\,d\mu_p^n(x) &= \int_{\{0,1\}^n}(1-p)\mathbf{1}_{\{x_i=1\}\cap \mathrm{Piv}_i(A)}(x)\,d\mu_p^n(x)\\ &= (1-p)\mu_p^n(\mathrm{Piv}_i(A)\cap\{x\in\{0,1\}^n : x_i=1\}) = (1-p)\mu_p^n(A_i),\end{aligned}$$

λόγω της μονοτονίας του A.

Οπότε, από το Λήμμα 4.1.2 έχουμε ότι

$$\frac{d\mu_p^n(A)}{dp} = \frac{1}{p(1-p)}\sum_{i=1}^n\int_{\{0,1\}^n}(x_i-p)\mathbf{1}_A(x)\,d\mu_p^n(x) = \frac{1}{p}\sum_{i=1}^n\mu_p^n(A_i),$$

που είναι το ζητούμενο. □

Ο Talagrand απέδειξε την ακόλουθη ανισότητα, η οποία συνδέει το μέτρο $\mu_p^n(A)$ του A με τα μέτρα $\mu_p^n(A_i)$ των $A_1,\dots,A_n$.

Θεώρημα 4.1.3. *Υπάρχει απόλυτη σταθερά $K>0$ τέτοια ώστε, για κάθε $0<p<1$ και κάθε μονότονο υποσύνολο A του $\{0,1\}^n$,*

$$\mu_p^n(A)(1-\mu_p^n(A)) \leqslant K(1-p)\log\Big(\frac{2}{p(1-p)}\Big)\sum_{i=1}^n\frac{\mu_p^n(A_i)}{\log(1/((1-p)\mu_p^n(A_i)))}. \tag{4.1.1}$$

Το Θεώρημα 4.1.3 προκύπτει από μια γενικότερη ανισότητα που έδειξε ο Talagrand για συναρτήσεις $f:\{0,1\}^n\to\mathbb{R}$. Για κάθε τέτοια συνάρτηση ορίζουμε

$$\Delta_i f(x) = \begin{cases}(1-p)(f(x)-f(U_i(x))) & ,\ \text{αν } x_i=1,\\ p(f(x)-f(U_i(x))) & ,\ \text{αν } x_i=0.\end{cases}$$

Θεώρημα 4.1.4 (L^1-L^2 ανισότητα, Talagrand)**.** *Υπάρχει απόλυτη σταθερά $K>0$ ώστε, για κάθε $f:\{0,1\}^n\to\mathbb{R}$ με $\int f\,d\mu_p^n=0$,*

$$\|f\|_2^2 \leqslant K\log\Big(\frac{2}{p(1-p)}\Big)\sum_{i=1}^n\frac{\|\Delta_i f\|_2^2}{\log(e\|\Delta_i f\|_2/\|\Delta_i f\|_1)}. \tag{4.1.2}$$

Πράγματι, χρησιμοποιώντας το Θεώρημα 4.1.4 μπορούμε να πάρουμε το Θεώρημα 4.1.3 ως εξής.

Απόδειξη του Θεωρήματος 4.1.3. Θεωρούμε ένα μονότονο υποσύνολο A του $\{0,1\}^n$ και τη συνάρτηση $f = \mathbf{1}_A - \mu_p^n(A)$. Παρατηρούμε ότι $\int_{\{0,1\}^n} f\, d\mu_p^n = 0$ και

$$\|f\|_2^2 = \int_{\{0,1\}^n} (\mathbf{1}_A - \mu_p^n(A))^2\, d\mu_p^n = \mu_p^n(A) - 2\mu_p^n(A)^2 + \mu_p^n(A)^2 = \mu_p^n(A)(1-\mu_p^n(A)).$$

Ισχυρισμός. Για κάθε $q \geqslant 1$ ισχύει ότι

$$\|\Delta_i f\|_q^q = \frac{1}{p}\mu_p^n(A_i)(p(1-p)^q + (1-p)p^q).$$

Πράγματι, έχουμε ότι

(4.1.3)
$$\begin{aligned}
&\|\Delta_i f\|_q^q = \int_{\{0,1\}^n} |\Delta_i f|^q\, d\mu_p^n \\
&= \int_{\{0,1\}^{n-1}} p(1-p)^q |\mathbf{1}_A(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n) - \mathbf{1}_A(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)| \\
&+ (1-p)p^q |\mathbf{1}_A(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n) - \mathbf{1}_A(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n)|\, d\mu_p^{n-1} \\
&= (p(1-p)^q + (1-p)p^q) \int_{\{0,1\}^{n-1}} \mathbf{1}_{A_i}(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n)\, d\mu_p^{n-1}
\end{aligned}$$

λόγω της μονοτονίας του A. Όμως,

$$\begin{aligned}
\mu_p^n(A_i) &= \int_{\{0,1\}^n} \mathbf{1}_{A_i}\, d\mu_p^n \\
&= \int_{\{0,1\}^{n-1}} \big(p\cdot \mathbf{1}_{A_i}(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n) \\
&\qquad + (1-p)\cdot \mathbf{1}_{A_i}(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)\big)\, d\mu_p^{n-1} \\
&= \int_{\{0,1\}^{n-1}} p\cdot \mathbf{1}_{A_i}(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n)\, d\mu_p^{n-1},
\end{aligned} \tag{4.1.4}$$

αφού από τη μονοτονία του A και τον ορισμό του A_i η $(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n) \in A_i$ δεν μπορεί να συμβαίνει ποτέ. Από τις (4.1.3) και (4.1.4) έπεται ο ισχυρισμός.

Συνεπώς, $\|\Delta_i f\|_1 = 2(1-p)\mu_p^n(A_i)$ και $\|\Delta_i f\|_2^2 = (1-p)\mu_p^n(A_i)$. Οπότε, από το Θεώρημα 4.1.4 παίρνουμε ότι

$$\begin{aligned}
\mu_p^n(A)(1-\mu_p^n(A)) &\leqslant K \log\Big(\frac{2}{p(1-p)}\Big) \sum_{i=1}^n \frac{(1-p)\mu_p^n(A_i)}{\log\Big(\frac{e}{2}\frac{1}{\sqrt{(1-p)\mu_p^n(A_i)}}\Big)} \\
&\leqslant 2K(1-p)\log\Big(\frac{2}{p(1-p)}\Big) \sum_{i=1}^n \frac{\mu_p^n(A_i)}{\log(1/((1-p)\mu_p^n(A_i)))}.
\end{aligned}$$

Δηλαδή, έχουμε το συμπέρασμα του θεωρήματος με σταθερά $K' = 2K$, όπου K η σταθερά στο Θεώρημα 4.1.4. $\square$

Συνέπειες του Θεωρήματος 4.1.3 είναι οι ακόλουθες.

Πόρισμα 4.1.5. *Έστω* $\varepsilon_p = \max_{1\leqslant i\leqslant n} \mu_p^n(A_i)$. *Τότε,*

$$\frac{d\mu_p^n(A)}{dp} \geqslant \frac{\log(1/\varepsilon_p)}{Kp(1-p)\log(2/(p(1-p)))}\,\mu_p^n(A)(1-\mu_p^n(A)).$$

Απόδειξη. Το συμπέρασμα προκύπτει άμεσα από το Θεώρημα 4.1.3 και τον τύπο των Margulis-Russo του Θεωρήματος 4.1.1. □

Πόρισμα 4.1.6. *Έστω* $\varepsilon = \sup_{0<p<1} \varepsilon_p$, *όπου* $\varepsilon_p = \max_{1\leqslant i\leqslant n} \mu_p^n(A_i)$. *Τότε, αν* $p_1 < p_2$ *έχουμε ότι*

$$\mu_{p_1}^n(A)(1-\mu_{p_2}^n(A)) \leqslant \varepsilon_1^{(p_2-p_1)/K_1},$$

όπου $K_1 > 0$ *είναι μια απόλυτη σταθερά.*

Απόδειξη. Αν θέσουμε $g(x) = \log\left(\frac{x}{1-x}\right)$ τότε το Πόρισμα 4.1.5 μας λέει ότι

$$\frac{d}{dp}\Big(g(\mu_p^n(A))\Big) = \frac{1}{\mu_p^n(A)(1-\mu_p^n(A))}\frac{d\mu_p^n(A)}{dp} \geqslant \frac{\log(1/\varepsilon_p)}{Kp(1-p)\log(2/(p(1-p)))} \geqslant \frac{1}{K_1}\log(1/\varepsilon_p)$$

για κάποια απόλυτη σταθερά $K_1 > 0$, διότι η $x \mapsto x\log\frac{K}{x}$ είναι φραγμένη. Θέτοντας $x_1 = \mu_{p_1}^n(A)$ και $x_2 = \mu_{p_2}^n(A)$ βλέπουμε ότι

$$g(x_2) - g(x_1) \geqslant \frac{p_2-p_1}{K_1}\log\frac{1}{\varepsilon_p}$$

για κάποιο $p \in (p_1, p_2)$. Συνεπώς,

$$\begin{aligned} x_1(1-x_2) &\leqslant \frac{x_1}{1-x_1}\,\frac{1-x_2}{x_2} = \exp(g(x_1)-g(x_2)) \leqslant \exp\Big(-\frac{p_2-p_1}{K_1}\log(1/\varepsilon_p)\Big) \\ &= \varepsilon_p^{\frac{p_2-p_1}{K_1}} \leqslant \varepsilon^{\frac{p_2-p_1}{K_1}}. \end{aligned}$$

□

Πόρισμα 4.1.7. *Έστω* $U = \mu_p^n(A)(1-\mu_p^n(A))/n\log(2/p(1-p))$. *Τότε,*

$$\max_{1\leqslant i\leqslant n} \mu_p^n(A_i) \geqslant \frac{1}{K_2(1-p)}U\log\frac{1}{U},$$

όπου $K_2 > 0$ *είναι μια απόλυτη σταθερά.*

Απόδειξη. Παρατηρούμε αρχικά ότι $\mu_p^n(A)(1-\mu_p^n(A)) \leqslant \frac{1}{4}$ και $\log(2/p(1-p)) \geqslant \log 8 = 3\log 2$. Άρα,

$$U = \frac{\mu_p^n(A)(1-\mu_p^n(A))}{n\log(2/p(1-p))} \leqslant \frac{1}{12\log 2\cdot n} \leqslant \frac{1}{6}. \tag{4.1.5}$$

Αν θέσουμε $\varepsilon = \max_{1\leqslant i\leqslant n} \mu_p^n(A_i)$ τότε από το Θεώρημα 4.1.3 και το γεγονός ότι η $h(x) = \frac{x}{\log(1/x)}$ είναι αύξουσα στο $(0,1)$, παίρνουμε

$$\mu_p^n(A)(1-\mu_p^n(A)) \leqslant K\log(2/(p(1-p)))\cdot n\frac{(1-p)\varepsilon}{\log(1/((1-p)\varepsilon))}.$$

Μπορούμε να υποθέσουμε ότι $K \geqslant 1$. Από τον ορισμό του U έχουμε, ισοδύναμα,

$$\frac{U}{K} \leqslant \frac{(1-p)\varepsilon}{\log(1/((1-p)\varepsilon))}, \tag{4.1.6}$$

και από την (4.1.5) και την $K \geqslant 1$ έπεται ότι $U/K \leqslant 1/2$.

Ισχυρισμός. Αν $0 < x \leqslant 1/2$ και $\frac{y}{\log(1/y)} \geqslant x$ τότε $y \geqslant \frac{x}{2}\log\frac{1}{x}$.

Πράγματι, έστω ότι $y < \frac{x}{2}\log\frac{1}{x}$. Παρατηρούμε ότι $\frac{x}{2}\log\frac{1}{x} \leqslant \frac{1}{2e} < 1$, οπότε από τη μονοτονία της h στο $(0,1)$ παίρνουμε

$$\frac{y}{\log(1/y)} \leqslant \frac{x}{2}\frac{\log\frac{1}{x}}{\log\left(\frac{2}{x\log\frac{1}{x}}\right)} = \frac{x}{2}\frac{\log\frac{1}{x}}{\log 2 + \log\frac{1}{x} - \log\log\frac{1}{x}}.$$

Θα δείξουμε ότι $\frac{y}{\log(1/y)} < x$, το οποίο είναι άτοπο από την υπόθεση. Θέτοντας $z = \log\frac{1}{x} \geqslant \log 2$ αρκεί να ελέγξουμε ότι

$$\frac{1}{2}\frac{z}{\log 2 + z - \log z} < 1,$$

δηλαδή

$$z - 2\log z + 2\log 2 \geqslant 0.$$

Η $m(z) = z - 2\log z + 2\log 2$ είναι φθίνουσα στο $[\log 2, 2]$ και αύξουσα στο $[2,\infty)$, άρα $m(z) \geqslant m(2) = 2 > 0$ και έπεται ο ισχυρισμός.

Αφού $U \leqslant K/2$, εφαρμόζοντας τον ισχυρισμό για $x = U/K$ και $y = (1-p)\varepsilon$ παίρνουμε

$$(1-p)\varepsilon \geqslant \frac{U}{2K}\log\frac{K}{U} \geqslant \frac{U}{2K}\log\frac{1}{U}$$

από την υπόθεση ότι $K \geqslant 1$. Έπεται το ζητούμενο με σταθερά $K_2 = 2K$. □

Στην περίπτωση $p = \frac{1}{2} = \mu_p^n(A)$ το Πόρισμα 4.1.7 έχει επίσης αποδειχθεί από τους Kahn, Kalai και Linial με μεθόδους αρμονικής ανάλυσης.

4.2 Η αρχική απόδειξη του Talagrand

Για την απόδειξη του Θεωρήματος 4.1.4 θεωρούμε τις συναρτήσεις $r_i : \{0,1\}^n \to \mathbb{R}$ με

$$r_i(x) = \begin{cases} \sqrt{\frac{1-p}{p}} & , \text{ αν } x_i = 1, \\ -\sqrt{\frac{p}{1-p}} & , \text{ αν } x_i = 0. \end{cases}$$

Από τον ορισμό προκύπτει άμεσα ότι

$$\int_{\{0,1\}^n} r_i\, d\mu_p^n = 0 \quad \text{και} \quad \int_{\{0,1\}^n} r_i^2\, d\mu_p^n = 1.$$

Λήμμα 4.2.1. *Έστω $X = \{a,b\}^n$ και ϱ ένα μέτρο πιθανότητας στον $(X, \mathcal{P}(X))$. Τότε,*

$$\dim_{\mathbb{R}}(L^2(\varrho)) = 2^n.$$

Απόδειξη. Αρχικά παρατηρούμε ότι για κάθε συνάρτηση $f : X \to \mathbb{R}$ ισχύει ότι $f \in L^2(\varrho)$. Θεωρούμε τις συναρτήσεις $\mathbf{1}_{\{x\}}$, $x \in X$. Αυτές είναι 2^n το πλήθος και προφανώς είναι γραμμικώς ανεξάρτητες. Έστω τώρα $g \in L^2(\varrho)$. Τότε,

$$g = \sum_{x \in X} g(x)\mathbf{1}_{\{x\}}.$$

Οπότε, βλέπουμε ότι $L^2(\varrho) = \mathrm{span}\{\mathbf{1}_{\{x\}} : x \in X\}$ και άρα $\dim_{\mathbb{R}}(L^2(\varrho)) = 2^n$. □

Για κάθε $S \subseteq \{1, \dots, n\}$, $S \neq \emptyset$, ορίζουμε

$$r_S(x) = \prod_{i \in S} r_i(x).$$

Ορίζουμε επίσης $r_\emptyset \equiv 1$. Οι συναρτήσεις r_S σχηματίζουν ορθοκανονική βάση του $L^2(\mu_p^n)$, διότι είναι 2^n το πλήθος και εύκολα ελέγχεται ότι $\int r_S^2 \, d\mu_p^n = 1$ για κάθε $S \subseteq \{1, \dots, n\}$ και $\int r_{S_1} r_{S_2} \, d\mu_p^n = 0$ για κάθε $S_1, S_2 \subseteq \{1, \dots, n\}$ με $S_1 \neq S_2$. Κάθε $g : \{0,1\}^n \to \mathbb{R}$ γράφεται στη μορφή $g = \sum\limits_S a_S r_s$ για κάποιους $a_S \in \mathbb{R}$, και $a_\emptyset = \int g \, d\mu_p^n$. Αν η g ικανοποιεί την $a_\emptyset = \int g \, d\mu_p^n = 0$, ορίζουμε

$$M(g)^2 = \sum_{\emptyset \neq S \subseteq \{1, \dots n\}} \frac{a_S^2}{|S|},$$

όπου συμβολίζουμε με $|S|$ τον πληθάριθμο του S. Στόχος μας είναι να δώσουμε κατάλληλα φράγματα για την ποσότητα $M(g)$. Για να δούμε πώς αυτά σχετίζονται με το Θεώρημα 4.1.4, ας θεωρήσουμε μια συνάρτηση $f : \{0,1\}^n \to \mathbb{R}$ με $\int f \, d\mu_p^n = 0$. Τότε, η f γράφεται στη μορφή $f = \sum_S b_S r_S$, με $b_\emptyset = 0$. Ο τελεστής Δ_i έχει οριστεί με τέτοιο τρόπο που να ικανοποιεί τις $\Delta_i(r_S) = 0$ αν $i \notin S$ και $\Delta_i(r_S) = r_S$ αν $i \in S$. Έχουμε λοιπόν

$$\Delta_i f = \sum_{i \in S} b_S r_S,$$

άρα

$$M(\Delta_i f)^2 = \sum_{i \in S} \frac{b_S^2}{|S|}.$$

Έπεται ότι

$$\|f\|_2^2 = \sum_S b_S^2 = \sum_{i=1}^n M(\Delta_i f)^2. \tag{4.2.1}$$

Το επόμενο λήμμα μας δίνει μια κρίσιμη ιδιότητα των συναρτήσεων r_S.

Λήμμα 4.2.2. *Έστω $q \geqslant 2$ και $\vartheta = 1/\sqrt{p(1-p)}$. Για κάθε $1 \leqslant k \leqslant n$ και κάθε ακολουθία $(a_S)_{|S|=k}$ πραγματικών αριθμών ισχύει ότι*

$$\Big\| \sum_{|S|=k} a_S r_S \Big\|_{L^q(\mu_p^n)} \leqslant (q-1)^{k/2} \vartheta^k \Big(\sum_{|S|=k} a_S^2 \Big)^{1/2}. \tag{4.2.2}$$

Απόδειξη. Η απόδειξη θα γίνει σε τρία βήματα.

Βήμα 1. Θεωρούμε τον διακριτό κύβο $\{-1,1\}^n$ εφοδιασμένο με το ομοιόμορφο μέτρο πιθανότητας λ. Για κάθε $S\subseteq\{1,\dots,n\}$ θέτουμε

$$w_S(\epsilon)=\prod_{i\in S}\epsilon_i.$$

Η (w_S) είναι ορθοκανονική βάση του $L^2(\lambda)$, αφού οι w_S είναι 2^n το πλήθος και ικανοποιούν τις $\int w_S^2\,d\lambda=1$ για κάθε $S\subseteq\{1,\dots,n\}$ και $\int w_{S_1}w_{S_2}\,d\lambda=0$ για κάθε $S_1,S_2\subseteq\{1,\dots,n\}$ με $S_1\neq S_2$. Ο Beckner έχει αποδείξει στο [10] ότι, για $\delta=1/\sqrt{q-1}$, ο τελεστής $T_\delta:L^2(\lambda)\to L^q(\lambda)$ με

$$T_\delta\Big(\sum b_Sw_S\Big)=\sum b_S\delta^{|S|}w_S \tag{4.2.3}$$

έχει νόρμα 1. Ειδικότερα, έχουμε ότι

$$\Big\|\sum_{|S|=k}b_Sw_S\Big\|_{L^q(\lambda)}\leqslant(q-1)^{k/2}\Big(\sum_{|S|=k}b_S^2\Big)^{1/2} \tag{4.2.4}$$

αφού

$$T_{\frac{1}{\sqrt{q-1}}}\Big(\sum_{|S|=k}b_Sw_S\Big)=\frac{1}{(q-1)^{k/2}}\sum_{|S|=k}b_Sw_S.$$

Βήμα 2. Εφοδιάζουμε το γινόμενο $G=\{0,1\}^n\times\{0,1\}^n$ με το μέτρο $\mu'=\mu_p^n\otimes\mu_p^n$ και το γινόμενο $H=G\times\{-1,1\}^n$ με το μέτρο $\nu=\mu'\otimes\lambda$. Για δοθέν $S\subseteq\{1,\dots,n\}$, και για $x,y\in\{0,1\}^n$ και $\epsilon\in\{-1,1\}^n$ θεωρούμε τις συναρτήσεις g_S και $g_{S,\epsilon}$ στο G που ορίζονται από τις

$$\begin{aligned}g_S(x,y)&=\prod_{i\in S}(r_i(x)-r_i(y)),\\ g_{S,\epsilon}(x,y)&=\prod_{i\in S}(r_i(x)-r_i(y))\epsilon_i=g_S(x,y)w_S(\epsilon)\end{aligned}$$

και τη συνάρτηση h_S στο H που ορίζεται από την $h_S(x,y,\epsilon)=g_{S,\epsilon}(x,y)$.

Παρατηρούμε ότι για δοθείσα ακολουθία (a_S), για κάθε $\epsilon\in\{-1,1\}^n$ έχουμε

$$\Big\|\sum a_Sg_S\Big\|_{L^q(\mu')}=\Big\|\sum a_Sg_{S,\epsilon}\Big\|_{L^q(\mu')}.$$

Συνεπώς, από το θεώρημα Fubini παίρνουμε

$$\Big\|\sum a_Sg_S\Big\|_{L^q(\mu')}=\Big\|\sum a_Sh_S\Big\|_{L^q(\nu)}. \tag{4.2.5}$$

Τώρα, χρησιμοποιώντας την (4.2.4) με $b_S=a_Sg_S(x,y)$, για κάθε x,y παίρνουμε

$$\int\Big|\sum_{|S|=k}a_Sg_S(x,y)w_S(\epsilon)\Big|^q d\lambda(\epsilon)\leqslant(q-1)^{kq/2}\Big(\sum_{|S|=k}a_S^2g_S^2(x,y)\Big)^{q/2}. \tag{4.2.6}$$

Παρατηρούμε ότι

$$|r_i(x)-r_i(y)|\leqslant\vartheta,$$

άρα $g_S^2(x,y) \leqslant \vartheta^{2k}$. Χρησιμοποιώντας αυτό το φράγμα στην (4.2.6), ολοκληρώνοντας ως προς x, y και παίρνοντας την q-οστή ρίζα, βλέπουμε ότι

$$\Big\| \sum_{|S|=k} a_S h_S \Big\|_{L^q(\nu)} \leqslant \vartheta^k (q-1)^{k/2} \Big(\sum_{|S|=k} a_S^2 \Big)^{1/2}. \tag{4.2.7}$$

Βήμα 3. Αφού $\int r_i(y)\, d\mu_p^n(y) = 0$, από το θεώρημα Fubini, ολοκληρώνοντας ως προς y μέσα στη νόρμα και χρησιμοποιώντας την ανισότητα Hölder, έχουμε ότι

$$\Big\| \sum a_S r_S \Big\|_{L^q(\mu_p^n)} \leqslant \Big\| \sum a_S g_S \Big\|_{L^q(\mu')},$$

και συνδυάζοντας αυτή την ανισότητα με τις (4.2.5) και (4.2.7) παίρνουμε το ζητούμενο. □

Θα χρησιμοποιήσουμε αυτό το λήμμα μέσω ενός επιχειρήματος δυϊσμού.

Πρόταση 4.2.3. *Έστω $g : \{0,1\}^n \to \mathbb{R}$. Για κάθε $S \subseteq \{1,\dots,n\}$ θέτουμε $a_S = \int r_S g\, d\mu_p^n$. Τότε,*

$$\sum_{|S|=k} a_S^2 \leqslant (q-1)^k \vartheta^{2k} \|g\|_{q'}^2 \tag{4.2.8}$$

όπου q' είναι ο συζυγής εκθέτης του q.

Απόδειξη. Από την ανισότητα Hölder και την (4.2.2) έχουμε ότι

$$\sum_{|S|=k} a_S^2 = \int \Big(\sum_{|S|=k} a_S r_S \Big) g\, d\mu_p^n \leqslant \Big\| \sum_{|S|=k} a_S r_S \Big\|_q \|g\|_{q'} \leqslant (q-1)^{k/2} \vartheta^k \Big(\sum_{|S|=k} a_S^2 \Big)^{1/2} \|g\|_{q'},$$

απ' όπου έπεται η (4.2.8). □

Πρόταση 4.2.4. *Υπάρχει απόλυτη σταθερά $K > 0$ ώστε, για κάθε $g : \{0,1\}^n \to \mathbb{R}$ με $\int g\, d\mu_p^n = 0$,*

$$M(g)^2 \leqslant K \log\Big(\frac{2}{p(1-p)} \Big) \frac{\|g\|_2^2}{\log(e\|g\|_2/\|g\|_1)}. \tag{4.2.9}$$

Απόδειξη. Θέτουμε $a_S = \int r_S g\, d\mu_p^n$. Από την Πρόταση 4.2.3, με $q = 3$ και $q' = \frac{3}{2}$, έχουμε ότι

$$\sum_{|S|=k} a_S^2 \leqslant (2\vartheta^2)^k \|g\|_{3/2}^2.$$

Συνεπώς, για κάθε $m \in \mathbb{N}$,

$$\sum_{1 \leqslant |S| \leqslant m} \frac{a_S^2}{|S|} \leqslant \sum_{k=1}^m \frac{(2\vartheta^2)^k}{k} \|g\|_{3/2}^2.$$

Θέτουμε $x_k = (2\vartheta^2)^k/k$. Αφού $\vartheta^2 = \frac{1}{p(1-p)} \geqslant 4$ έχουμε ότι $x_{k+1}/x_k \geqslant 2$, άρα $\sum_{k=1}^m x_k \leqslant 2x_m$. Ισχύει ότι $M(g)^2 \leqslant \|g\|_2^2$, και για κάθε $m \geqslant 1$ έχουμε

$$\begin{aligned} M(g)^2 &= \sum_{1 \leqslant |S| \leqslant m} \frac{a_S^2}{|S|} + \sum_{|S|>m} \frac{a_S^2}{|S|} \leqslant 2\frac{(2\vartheta^2)^m}{m} \|g\|_{3/2}^2 + \frac{1}{m+1} \sum_{|S|>m} a_S^2 \\ &\leqslant \frac{1}{m+1} \Big(4(2\vartheta^2)^m \|g\|_{3/2}^2 + \|g\|_2^2 \Big), \end{aligned} \tag{4.2.10}$$

αφού $\frac{2}{m} \leqslant \frac{4}{m+1}$ και
$$\sum_{|S|>m} a_S^2 \leqslant \sum_S a_S^2 = \|g\|_2^2.$$
Σημειώνουμε ότι η (4.2.10) ισχύει και για $m = 0$, αφού $M(g)^2 \leqslant \|g\|_2^2$.

Επιλέγουμε ως m τον μεγαλύτερο μη αρνητικό ακέραιο για τον οποίο $(2\vartheta^2)^m\|g\|_{3/2}^2 \leqslant e^2\|g\|_2^2$ (σημειώνουμε ότι αυτή η ανισότητα ισχύει αν $m = 0$, αφού $\|g\|_{3/2} \leqslant \|g\|_2 \leqslant e\|g\|_2$). Τότε, $m \geqslant 0$ και $(2\vartheta^2)^{m+1}\|g\|_{3/2}^2 \geqslant e^2\|g\|_2^2$, άρα
$$m + 1 \geqslant \frac{2\log(e\|g\|_2/\|g\|_{3/2})}{\log(2\vartheta^2)}.$$
Αφού $m+1 \geqslant 1$, από την (4.2.10) και την $2\vartheta^2 = \frac{2}{p(1-p)}$ έπεται ότι, αν θέσουμε $K = \frac{4e^2+1}{2}$, έχουμε

$$\begin{aligned} M(g)^2 &\leqslant \frac{\log(2\vartheta^2)}{2\log(e\|g\|_2/\|g\|_{3/2})}(4e^2+1)\|g\|_2^2 = \frac{K\log(2\vartheta^2)}{\log(e\|g\|_2/\|g\|_{3/2})}\,\|g\|_2^2 \\ &= K\log\Big(\frac{2}{p(1-p)}\Big)\,\frac{\|g\|_2^2}{\log(e\|g\|_2/\|g\|_{3/2})}, \end{aligned} \tag{4.2.11}$$

Τώρα, παρατηρούμε ότι
$$\Big(\int |g|^{3/2}d\mu_p^n\Big)^2 \leqslant \int |g|\,d\mu_p^n \cdot \int g^2\,d\mu_p^n$$
από την ανισότητα Cauchy-Schwarz, δηλαδή έχουμε την
$$\|g\|_{3/2}^3 \leqslant \|g\|_1\|g\|_2^2,$$
την οποία ξαναγράφουμε ως
$$\frac{\|g\|_2}{\|g\|_1} \leqslant \Big(\frac{\|g\|_2}{\|g\|_{3/2}}\Big)^3.$$
Επιστρέφοντας στην (4.2.11) και αντικαθιστώντας τον λόγο $\|g\|_2/\|g\|_{3/2}$ με $(\|g\|_2/\|g\|_1)^{1/3}$ παίρνουμε το ζητούμενο. □

Απόδειξη του Θεωρήματος 4.1.4. Από την (4.2.1) έχουμε ότι
$$\|f\|_2^2 = \sum_{i=1}^n M(\Delta_i f)^2,$$
και εφαρμόζοντας την Πρόταση 4.2.4 παίρνουμε άμεσα την (4.1.2). □

4.3 Απόδειξη μέσω της υπερσυσταλτότητας

Υπενθυμίζουμε ότι στον $(X, \mathcal{A}, P) = (\{-1,1\}^n, \mathcal{P}(\{-1,1\}^n), \mu_p^n)$ έχουμε ορίσει για κάθε $f : \{-1,1\}^n \to \mathbb{R}$ και για κάθε $1 \leqslant i \leqslant n$ τη συνάρτηση $f_i : \{-1,1\} \to \mathbb{R}$ για σταθερά $x_1, \ldots, x_{i-1}$, $x_{i+1}, \ldots, x_n$ ως εξής:
$$f_i(x_i) = f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

Έχουμε επίσης ορίσει τους τελεστές $L_i : L^2(\mu_p^n) \to L^2(\mu_p^n)$ με

$$L_i(f) = \int_{\{-1,1\}} f_i(x_i) d\mu_p(x_i) - f$$

και τον τελεστή

$$L = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} \int_{\{-1,1\}} f_i(x_i) d\mu_p(x_i) - nf.$$

Είδαμε ότι ο L είναι αυτοσυζυγής τελεστής Markov και παράγει την ημιομάδα Markov $(P_t)_{t\geqslant 0}$, όπου

$$P_t = e^{Lf} = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k.$$

Σε αυτή την παράγραφο θα χρησιμοποιήσουμε την υπερσυσταλτότητα για την παραπάνω ημιομάδα $(P_t)_{t\geqslant 0}$ ώστε να καταλήξουμε ακριβώς στην ανισότητα που δημοσίευσε ο Talagrand.

Λήμμα 4.3.1. *Αν $f, g \in L^2(\mu_p^n)$ τότε*

$$\mathcal{E}(f,g) = -\int f Lg \, d\mu_p^n = \sum_{i=1}^{n} \int L_i f \cdot L_i g \, d\mu_p^n.$$

Απόδειξη. Γράφουμε

$$\begin{aligned}
\mathcal{E}(f,g) &= -\int f \cdot Lg \, d\mu_p^n = -\int_X f\Big(\sum_{i=1}^{n} \int_{\{-1,1\}} g_i(x_i)\, d\mu_p(x_i) - ng\Big)\, d\mu_p^n \\
&= \sum_{i=1}^{n} \int_X f\Big(g - \int_{\{-1,1\}} g_i(x_i) d\mu_p(x_i)\Big)\, d\mu_p^n(x) \\
&= \sum_{i=1}^{n} \int_{\{-1,1\}^{n-1}} \Big(\int_{\{-1,1\}} f_i(x_i)\Big(g_i(x_i) - \int_{\{-1,1\}} g_i(x_i) d\mu_p(x_i)\Big) d\mu_p(x_i)\Big) \\
&\qquad\qquad \times\, d\mu_p^{n-1}(x_1, \ldots, \hat{x}_i, \ldots, x_n) \\
&= \sum_{i=1}^{n} \int_{\{-1,1\}^{n-1}} \Big[\int_{\{-1,1\}} f_i(x) g_i(x_i)\, d\mu_p(x_i) - \int f_i(x_i)\, d\mu_p(x_i) \int g_i(x_i) d\mu_p(x_i)\Big)\Big] \\
&\qquad\qquad \times\, d\mu_p^{n-1}(x_1, \ldots, \hat{x}_i, \ldots, x_n).
\end{aligned}$$

Τώρα,

$$
\begin{aligned}
&\sum_{i=1}^{n}\int_X L_i f\cdot L_i g\,d\mu_p^n \\
&=\sum_{i=1}^{n}\int_X\Big(f-\int_{\{-1,1\}} f_i(x_i)\,d\mu_p(x_i)\Big)\Big(g-\int_{\{-1,1\}} g_i(x_i)\,d\mu_p(x_i)\Big)\,d\mu_p^n \\
&=\sum_{i=1}^{n}\int_{\{-1,1\}^{n-1}}\int_{\{-1,1\}}\Big(g_i(x_i)-\int g_i(x_i)\,d\mu_p(x_i)\Big) \\
&\qquad\quad\times\Big(f_i(x_i)-\int f_i(x_i)\,d\mu_p(x_i)\Big)\,d\mu_p(x_i)\,d\mu_p^{n-1} \\
&=\sum_{i=1}^{n}\int_{\{-1,1\}^{n-1}}\Big[\int_{\{-1,1\}} f_i(x)g_i(x_i)\,d\mu_p(x_i)-\int f_i(x_i)\,d\mu_p(x_i)\int g_i(x_i)d\mu_p(x_i)\Big)\Big] \\
&\qquad\quad\times\,d\mu_p^{n-1}(x_1,\dots,\hat{x_i},\dots,x_n),
\end{aligned}
$$

και συνεπώς πράγματι έχουμε ισότητα. □

Λήμμα 4.3.2. *Αν $f\in L^2(\mu_p^n)$ τότε*

$$\int_X f^2\,d\mu_p^n-\int_X (P_1f)^2\,d\mu_p^n=2\int_0^1\sum_{i=1}^{n}\int_X (L_iP_tf)^2\,d\mu_p^n\,dt.$$

Απόδειξη. Είναι φανερό ότι όλες οι πραγματικές συναρτήσεις στον X είναι φραγμένες και συνεπώς ολοκληρώσιμες. Οπότε,

$$
\begin{aligned}
\int_X f^2\,d\mu_p^n-\int_X (P_1f)^2\,d\mu_p^n&=\int_X\big((P_0f)^2-(P_1f)^2\big)\,d\mu_p^n=\int_X-\int_0^1\frac{d}{dt}\big((P_tf)^2\big)\,dt\,d\mu_p^n \\
&=-\int_0^1\int_X 2P_tf\cdot LP_tf\,d\mu_p^n\,dt=2\int_0^1\mathcal{E}(P_tf,P_tf)\,dt \\
&=2\int_0^1\sum_{i=1}^{n}\int_X (L_iP_tf)^2\,d\mu_p^n,
\end{aligned}
$$

από το προηγούμενο λήμμα. □

Λήμμα 4.3.3. *Αν $f\in L^2(\mu_p^n)$ με $\int f\,d\mu_p^n=0$ τότε*

$$\int_X f^2\,d\mu_p^n\leqslant 3\int_0^1\sum_{i=1}^{n}\int_X (L_iP_tf)^2\,d\mu_p^n\,dt.$$

Απόδειξη. Θεωρούμε τη συνάρτηση

$$\Lambda(s)=e^{2s}\int_X (P_sf)^2\,d\mu_p^n,\qquad s\geqslant 0.$$

Τότε,

$$\Lambda'(s)=e^{2s}\Big(2\int_X (P_sf)^2d\mu_p^n+2\int_X P_sf\cdot LP_sf\,d\mu_p^n\Big)=2e^{2s}\Big(\int_X (P_sf)^2\,d\mu_p^n-\mathcal{E}(P_sf,P_sf)\Big).$$

Όμως,

$$\begin{aligned}\mathrm{Var}(P_s f) &= \int (P_s f)^2\, d\mu_p^n - \Big(\int P_s f\, d\mu_p^n\Big)^2 = \int (P_s f)^2\, d\mu_p^n - \Big(\int f\, d\mu_p^n\Big)^2 \\ &= \int (P_s f)^2\, d\mu_p^n\end{aligned}$$

λόγω του αναλλοίωτου του μ_p^n ως προς την $(P_t)_{t\geqslant 0}$. Οπότε,

$$\Lambda'(s) = 2e^{2s}\big[\mathrm{Var}(P_s f) - \mathcal{E}(P_s f, P_s f)\big] \leqslant 0$$

από την ανισότητα Poincaré στον διακριτό κύβο (Πρόταση 2.3.11). Άρα, η Λ είναι φθίνουσα. Συνεπώς $\Lambda(0) \geqslant \Lambda(1)$, δηλαδή

$$\int_X f^2\, d\mu_p^n \geqslant e^2 \int_X (P_1 f)^2\, d\mu_p^n,$$

ή ισοδύναμα,

$$\int_X (P_1 f)^2\, d\mu_p^n \leqslant \frac{1}{e^2}\int_X f^2\, d\mu_p^n.$$

Τώρα, από το προηγούμενο λήμμα έχουμε ότι

$$\begin{aligned}\int_X f^2\, d\mu_p^n &= 2\int_0^1 \sum_{i=1}^n \int_X (L_i P_t f)^2\, d\mu_p^n\, dt + \int_X (P_1 f)^2\, d\mu_p^n \\ &\leqslant 2\int_0^1 \sum_{i=1}^n \int_X (L_i P_t f)^2\, d\mu_p^n\, dt + \frac{1}{e^2}\int_X f^2\, d\mu_p^n,\end{aligned}$$

άρα

$$\int_X f^2\, d\mu_p^n \leqslant \frac{2(e^2-1)}{e^2}\int_0^1 \sum_{i=1}^n \int_X (L_i P_t f)^2\, d\mu_p^n \leqslant 3\int_0^1 \sum_{i=1}^n \int_X (L_i P_t f)^2\, d\mu_p^n.$$

□

Λήμμα 4.3.4. *Έστω $f \in L^2(\mu_p^n)$ και $q(t) = 1 + e^{-4rt}$, $t > 0$, όπου*

$$r = \begin{cases} \frac{p-q}{\log p - \log q} & , \text{ αν } p \neq \frac12 \\ \frac12 & , \text{ αν } p = \frac12. \end{cases}$$

και $q = 1 - p$. Τότε

$$\int_X (L_i P_t f)^2\, d\mu_p^n = \int_X (P_t L_i f)^2\, d\mu_p^n \leqslant \Big(\int_X (L_i f)^{q(t)} d\mu_p^n\Big)^{\frac{2}{q(t)}}.$$

Απόδειξη. Αρχικά θα δείξουμε ότι $LL_i g = L_i L g$ για κάθε $g \in L^2(\mu_p^n)$. Τότε,

$$P_t L_i = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k L_i = \sum_{k=0}^{\infty} \frac{t^k}{k!} L_i L^k = L_i\Big(\sum_{k=0}^{\infty} \frac{t^k}{k!} L^k\Big) = L_i P_t$$

για κάθε $1 \leqslant i \leqslant n$ και κάθε $t \geqslant 0$.

Αφού $L = \sum_{j=1}^n L_j$, αρκεί να δείξουμε ότι $L_j L_i g = L_i L_j g$ για κάθε $1 \leqslant i, j \leqslant n$. Για $j = i$ προφανώς ισχύει. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $j > i$ και γράφουμε

$$\begin{aligned}
L_j L_i g(x) &= L_j \Big(\int_{\{-1,1\}} g(x_1, \ldots, y_i, \ldots, x_n)\, d\mu_p^n(y_i) - g(x_1, \ldots, x_n) \Big) \\
&= \int_{\{-1,1\}} \Big(\Big(\int_{\{-1,1\}} g(x_1, \ldots, y_i, \ldots, x_n) d\mu_p(y_i) \Big)(y_j) - g(x_1, \ldots, y_j, \ldots, x_n) \Big)\, d\mu_p(y_j) \\
&\qquad - \int_{\{-1,1\}} g(x_1, \ldots, y_i, \ldots, x_n)\, d\mu_p(y_i) + g(x_1, \ldots, x_n) \\
&= \int_{\{-1,1\}} \int_{\{-1,1\}} g(x_1, \ldots, y_i, \ldots, y_j, \ldots, x_n)\, d\mu_p(y_i)\, d\mu_p(y_j) \\
&\qquad - \int_{\{-1,1\}} g(x_1, \ldots, y_j, \ldots, x_n) \Big)\, d\mu_p(y_j) \\
&\qquad - \int_{\{-1,1\}} g(x_1, \ldots, y_i, \ldots, x_n)\, d\mu_p(y_i) + g(x_1, \ldots, x_n) \\
&= L_i L_j g(x)
\end{aligned}$$

λόγω συμμετρίας της τελευταίας μορφής ως προς i και j.

Τώρα, έχουμε ότι $1 < q(t) < 2$ για κάθε $t > 0$. Θα δείξουμε ότι ισχύει η συνθήκη για να εφαρμόσουμε την υπερσυσταλτότητα στον διακριτό κύβο. Πρέπει, για κάθε $t > 0$, να έχουμε ότι $e^{4rt} \geqslant \frac{2-1}{q(t)-1}$ δηλαδή $e^{4rt} \geqslant 1/e^{-4rt}$, που ισχύει ως ισότητα για κάθε $t > 0$. Οπότε, έχουμε διαδοχικά ότι

$$\int_X (L_i P_t f)^2\, d\mu_p^n = \int_X (P_t L_i f)^2\, d\mu_p^n = \|P_t L_i f\|_2^2 \leqslant \|L_i f\|_{q(t)}^2 = \Big(\int_X |L_i f|^{q(t)} d\mu_p^n \Big)^{\frac{2}{q(t)}}$$

για κάθε $1 \leqslant i \leqslant n$ και κάθε $t > 0$, εφαρμόζοντας την ανισότητα υπερσυσταλτότητας στον διακριτό κύβο (Θεώρημα 2.4.7). $\square$

Πρόταση 4.3.5. *Αν $f \in L^2(\mu_p^n)$ με $\int_X f\, d\mu_p^n = 0$ τότε*

$$\int_X f^2\, d\mu_p^n \leqslant \frac{1}{r} e^{4r} \sum_{i=1}^n \int_1^2 \|L_i f\|_v^2\, dv,$$

όπου $q = 1 - p$ και

$$r = \begin{cases} \frac{p-q}{\log p - \log q} & , \text{ αν } p \neq \frac{1}{2} \\ \frac{1}{2} & , \text{ αν } p = \frac{1}{2}. \end{cases}$$

Απόδειξη. Χρησιμοποιώντας τα Λήμματα 4.3.3 και 4.3.4 έχουμε ότι

$$\int_X f^2\, d\mu_p^n \leqslant 3 \int_0^1 \sum_{i=1}^n \int_X (L_i P_t f)^2\, d\mu_p^n\, dt \leqslant 3 \int_0^1 \sum_{i=1}^n \Big(\int_X |L_i f|^{q(t)} d\mu_p^n \Big)^{\frac{2}{q(t)}} dt,$$

όπου $q(t) = 1 + e^{-4rt}$. Θέτουμε $v = 1 + e^{-4rt}$, οπότε $dt = \frac{dv}{-4r(v-1)}$. Οπότε,

$$\begin{aligned}
3\int_0^1 \sum_{i=1}^n \Big(\int_X |L_i f|^{q(t)} d\mu_p^n \Big)^{\frac{2}{q(t)}} dt &= 3\sum_{i=1}^n \int_{1+e^{-4r}}^2 \Big(\int_X |L_i f|^v d\mu_p^n \Big)^{\frac{2}{v}} \frac{1}{4r(v-1)}\, dv \\
&\leqslant \frac{3e^{4r}}{4r} \sum_{i=1}^n \int_{1+e^{-4r}}^2 \Big(\int_X |L_i f|^v d\mu_p^n \Big)^{\frac{2}{v}} dv \\
&\leqslant \frac{1}{r} e^{4r} \sum_{i=1}^n \int_1^2 \Big(\int_X |L_i f|^v d\mu_p^n \Big)^{\frac{2}{v}} dv,
\end{aligned}$$

που είναι ακριβώς το ζητούμενο. □

Λήμμα 4.3.6. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας, $\vartheta \in (0,1)$ και $p_1, p_2, \kappa \geqslant 1$ ώστε $\frac{1}{\kappa} = \frac{\vartheta}{p_1} + \frac{1-\vartheta}{p_2}$. Αν $f \in L^\infty(\mu)$ τότε*

$$\|f\|_\kappa \leqslant \|f\|_{p_1}^{\vartheta} \|f\|_{p_2}^{1-\vartheta}.$$

Απόδειξη. Γράφουμε

$$\|f\|_\kappa = \Big(\int |f|^\kappa \, d\mu \Big)^{\frac{1}{\kappa}} = \Big(\int |f|^{\vartheta\kappa} |f|^{(1-\vartheta)\kappa} \Big)^{\frac{1}{\kappa}}.$$

Θέτουμε $\alpha = \frac{p_1}{\vartheta\kappa} > 1$. Τότε, ο συζυγής εκθέτης του α είναι ο $\beta = \frac{p_2}{(1-\vartheta)\kappa}$. Οπότε, από την ανισότητα Hölder,

$$\begin{aligned}
\|f\|_\kappa &\leqslant \Big(\int |f|^{\vartheta\kappa\alpha} \, d\mu \Big)^{\frac{1}{\alpha\kappa}} \Big(\int |f|^{(1-\vartheta)\kappa\beta} \Big)^{\frac{1}{\kappa\beta}} \\
&= \Big(\int |f|^{p_1} \, d\mu \Big)^{\frac{\vartheta}{p_1}} \Big(\int |f|^{p_2} \Big)^{\frac{1-\vartheta}{p_2}} \\
&= \|f\|_{p_1}^{\vartheta} \|f\|_{p_2}^{1-\vartheta}.
\end{aligned}$$

□

Θεώρημα 4.3.7 (L^1-L^2 ανισότητα)**.** *Αν $f : X \to \mathbb{R}$ τότε*

$$\mathrm{Var}_{\mu_p^n}(f) \leqslant \frac{2}{r} e^{4r} \sum_{i=1}^n \frac{\|L_i f\|_2^2}{1 + \log\Big(\frac{\|L_i f\|_2}{\|L_i f\|_1}\Big)},$$

όπου

$$r = \begin{cases} \frac{p-q}{\log p - \log q} & , \text{ αν } p \neq \frac{1}{2} \\ \frac{1}{2} & , \text{ αν } p = \frac{1}{2}. \end{cases}$$

και $q = 1 - p$.

Απόδειξη. Αρκεί να δείξουμε το αποτέλεσμα για τις $f : X \to \mathbb{R}$ με $\int_X f \, d\mu_p^n = 0$, αφού

$$\mathrm{Var}_{\mu_p^n}\Big(f - \int_X f \, d\mu_p^n \Big) = \mathrm{Var}_{\mu_p^n}(f) \quad \text{και} \quad L_i\Big(f - \int_X f \, d\mu_p^n \Big) = L_i f$$

για κάθε $1 \leqslant i \leqslant n$.

Έστω λοιπόν ότι $\int_X f\,d\mu_p^n = 0$. Αν $\kappa \in (1,2)$ και $\vartheta = \vartheta(\kappa)$ ώστε $\frac{1}{\kappa} = \frac{\vartheta}{1} + \frac{1-\vartheta}{2}$, δηλαδή $\vartheta(k) = \frac{2}{\kappa} - 1$, από το προηγούμενο λήμμα παίρνουμε ότι

$$\|L_i f\|_\kappa \leqslant \|L_i f\|_1^{\vartheta(\kappa)} \|L_i f\|_2^{1-\vartheta(\kappa)}.$$

Επομένως,

$$\int_1^2 \|L_i f\|_v^2\,dv \leqslant \int_1^2 \|L_i f\|_1^{2\vartheta(v)} \|L_i f\|_2^{2-2\vartheta(v)}\,dv = \|L_i f\|_2^2 \int_1^2 b^{2\vartheta(v)} dv,$$

όπου $b = \frac{\|L_i f\|_1}{\|L_i f\|_2} \leqslant 1$ αν $\|L_i f\|_2 \neq 0$ και $b = 1$ αν $\|L_i f\|_2 = 0$. Σε κάθε περίπτωση, $0 < b \leqslant 1$.

Τώρα, αν $0 < b < 1$ έχουμε

$$\begin{aligned}\int_1^2 b^{2\vartheta(v)} dv &= \int_0^2 \frac{4b^s}{(s+2)^2}\,ds \leqslant \int_0^2 b^s\,ds \\ &= \frac{1-b^2}{\log\frac{1}{b}} \leqslant \frac{2}{1+\log\frac{1}{b}}.\end{aligned}$$

[Για την τελευταία ανισότητα θέλουμε ισοδύναμα

$$1 - b^2 - \log b + b^2 \log b \leqslant -2\log b,$$

δηλαδή, θέτοντας $t = b^2$ ζητάμε $u(t) \geqslant 1$ για $t \in (0,1)$, όπου $u(t) = t - \frac{1}{2}\log t - \frac{t}{2}\log t$. Παρατηρούμε ότι

$$2u'(t) = \log\frac{1}{t} - \Big(\frac{1}{t} - 1\Big) \leqslant 0$$

για $t \in (0,1)$ και $\lim\limits_{t\to 1^-} u(t) = 1$, το οποίο αποδεικνύει το ζητούμενο.] Παρατηρούμε ότι η ίδια ανισότητα ισχύει και για $b = 1$.

Συνεπώς, από την Πρόταση 4.3.5 παίρνουμε ότι

$$\mathrm{Var}_{\mu_p^n}(f) = \int_X f^2\,d\mu_p^n \leqslant \frac{1}{r}e^{4r} \sum_{i=1}^n \int_1^2 \|L_i f\|_v^2\,dv \leqslant \frac{2}{r}e^{4r} \sum_{i=1}^n \frac{\|L_i f\|_2^2}{1 + \log\Big(\frac{\|L_i f\|_2}{\|L_i f\|_1}\Big)}.$$

□

Παρατήρηση 4.3.8. Μένει να δείξουμε ότι η ανισότητα του τελευταίου θεωρήματος είναι όντως η L^1-L^2 ανισότητα που απέδειξε ο Talagrand.

Πράγματι, έχουμε ότι αν $x = (x_1, \dots, x_n)$ τότε

$$L_i f(x) = (1-p)[f(x_1, \dots, x_{i-1}, -1, x_{i+1}, \dots, x_n) - f(x)]$$

αν $x_i = +1$ και

$$L_i f(x) = p[f(x_1, \dots, x_{i-1}, +1, x_{i+1}, \dots, x_n) - f(x)]$$

αν $x_i = -1$, δηλαδή $L_i f(x) = -\Delta_i f(x)$, όπως αυτό ορίστηκε στην ανισότητα του Talagrand, αν αντικαταστήσουμε το $\{-1,1\}$ με το $\{0,1\}$.

Παρατηρούμε τέλος ότι

$$\lim_{p\to 0^+} \frac{e^{4r}}{r\log\Big(\frac{1-p}{p}\Big)} = \lim_{p\to 0^+}\Bigg(-\frac{e^{\frac{4(2p-1)}{\log\left(\frac{p}{1-p}\right)}}}{2p-1}\Bigg) = 1,$$

άρα το $\frac{1}{r}e^{4r}$ είναι της τάξης του $\log\frac{1}{p}$ καθώς $p \to 0^+$.

4.3.1 Η L^1-L^2 ανισότητα στο χώρο του Gauss

Σε αυτή την παράγραφο περιγράφουμε μια απόδειξη που έδωσαν οι Cordero-Erausquin και Ledoux για την L^1-L^2 ανισότητα του Talagrand, χρησιμοποιώντας υπερσυσταλτότητα. Η απόδειξη δουλεύει σε ένα γενικότερο πλαίσιο ημιομάδων Markov, το οποίο συμπεριλαμβάνει τον διακριτό κύβο, το μέτρο Gauss και πιο γενικά μοντέλα που δεν είναι χώροι γινόμενο. Εκτός από την υπερσυσταλτότητα, βασικό ρόλο παίζει εδώ η υπόθεση ότι η μορφή Dirichlet έχει ένα είδος ανάλυσης σε «διευθύνσεις» που μετατίθενται με τους τελεστές της ημιομάδας Markov. Αυτό συμβαίνει σίγουρα στον διακριτό κύβο και στον χώρο του Gauss. Το ποιες είναι οι διευθύνσεις είναι σαφές σε έναν χώρο γινόμενο, αλλά σε πιο γενικά μοντέλα χρειάζεται κανείς πρόσθετη δομή.

Θα δουλέψουμε λοιπόν στο αφηρημένο πλαίσιο μιας ημιομάδας Markov $\{P_t\}_{t\geqslant 0}$ με γεννήτορα L, και θα υποθέσουμε ότι η αντίστοιχη μορφή Dirichlet $\mathcal{E}$ αναλύεται σε «διευθύνσεις» Γ_i που δρουν στις συναρτήσεις στον X, ως

$$\mathcal{E}(f,f)=\sum_{i=1}^{n}\int(\Gamma_i f)^2 d\mu \tag{4.3.1}$$

με τέτοιο τρόπο ώστε, για κάθε $i=1,\dots,n$, η Γ_i «μετατίθεται» με την $\{P_t\}_{t\geqslant 0}$ με την έννοια ότι, για κάποια σταθερά $\kappa\in\mathbb{R}$, για κάθε $t\geqslant 0$ και για κάθε f που ανήκει σε κατάλληλη κλάση συναρτήσεων,

$$\Gamma_i(P_t f)\leqslant e^{\kappa t}P_t(\Gamma_i f). \tag{4.3.2}$$

Υποθέτουμε επίσης ότι το μ ικανοποιεί ανισότητα υπερσυσταλτότητας με σταθερά $\beta>0$, το οποίο σημαίνει ότι για κάθε $p\geqslant 1+e^{-2t/\beta}$ και για κάθε μη αρνητική $f\in L^p(\mu)$ έχουμε ότι

$$\|P_t f\|_2\leqslant\|f\|_p.$$

Αυτή η υπόθεση της υπερσυσταλτότητας είναι ισοδύναμη με το ότι το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev με σταθερά β.

Θεώρημα 4.3.9 (Talagrand). *Έστω $\{P_t\}_{t\geqslant 0}$ μια ημιομάδα Markov με γεννήτορα L ο οποίος δρα σε κατάλληλη κλάση συναρτήσεων στον χώρο πιθανότητας X. Υποθέτουμε ότι η $\{P_t\}_{t\geqslant 0}$ και ο L έχουν ένα αναλλοίωτο, αντιστρέψιμο και εργοδικό μέτρο πιθανότητας μ. Υποθέτουμε επίσης ότι το ζεύγος (L,μ) ικανοποιεί την ανισότητα υπερσυσταλτότητας με σταθερά $\beta>0$ και ότι ισχύουν οι (4.3.1) και (4.3.2). Τότε, για κάθε f στον $L^2(\mu)$,*

$$\mathrm{Var}_\mu(f)\leqslant C(\beta,\kappa)\sum_{i=1}^{n}\frac{\|\Gamma_i f\|_2^2}{1+\log(\|\Gamma_i f\|_2/\|\Gamma_i f\|_1)},$$

όπου $C(\beta,\kappa)=4\beta e^{(1+\kappa\beta)^+}$.

Απόδειξη. Ξεκινάμε από την αναπαράσταση της διασποράς, κατά μήκος της ημιομάδας $\{P_t\}_{t\geqslant 0}$, μιας συνάρτησης f που ανήκει στο $L^2(\mu)$-πεδίο της ημιομάδας. Έχουμε

$$\mathrm{Var}_\mu(f)=-\int_0^\infty\Big(\frac{d}{dt}\int(P_tf)^2d\mu\Big)\,dt=-2\int_0^\infty\Big(\int P_tf\cdot LP_tf\,d\mu\Big)\,dt.$$

Γνωρίζουμε ότι το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev, άρα και την ανισότητα Poincaré, με σταθερά β, άρα,

$$\|P_t f\|_2 \leqslant e^{-t/\beta}\|f\|_2$$

για κάθε $t \geqslant 0$ και $f \in L^2(\mu)$. Έπεται ότι

$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{1-e^{-t/\beta}}\big(\,\|f\|_2^2 - \|P_t f\|_2^2\,\big) \tag{4.3.3}$$

για κάθε $t > 0$. Επιλέγοντας $T = \beta/2$ παίρνουμε

$$\mathrm{Var}_\mu(f) \leqslant 2\big(\,\|f\|_2^2 - \|P_T f\|_2^2\,\big).$$

Συνεπώς, αρκεί να δώσουμε ένα άνω φράγμα για την ποσότητα

$$\|f\|_2^2 - \|P_T f\|_2^2 = -2\int_0^T \Big(\int P_t f \cdot LP_t f\, d\mu\Big)\, dt = 2\int_0^T \mathcal{E}(P_t f, P_t f)\, d\mu.$$

Χρησιμοποιώντας την (4.3.1) ξαναγράφουμε αυτή τη σχέση ως εξής:

$$\|f\|_2^2 - \|P_T f\|_2^2 = 2\sum_{i=1}^n \int_0^T \Big(\int (\Gamma_i(P_t f))^2\, d\mu\Big)\, dt.$$

Μένει τώρα να δώσουμε ένα άνω φράγμα για καθένα από αυτά τα ολοκληρώματα. Έστω $1 \leqslant i \leqslant n$. Αφού υποθέσαμε ότι ισχύει η (4.3.2), έχουμε

$$\int (\Gamma_i(P_t f))^2 d\mu \leqslant e^{\kappa t}\int (P_t(\Gamma_i f))^2\, d\mu.$$

Επίσης, αφού η $\{P_t\}_{t\geqslant 0}$ ικανοποιεί την ανισότητα υπερσυσταλτότητας με σταθερά $\beta > 0$ βλέπουμε ότι

$$\|P_t(\Gamma_i f)\|_2 \leqslant \|\Gamma_i f\|_p$$

όπου $p = p(t) = 1 + e^{-2t/\beta} \leqslant 2$. Με την αλλαγή μεταβλητής $s = p(t)$, παίρνουμε

$$\mathrm{Var}_\mu(f) \leqslant 2\beta e^{(1+\kappa\beta)^+}\sum_{i=1}^n \int_1^2 \|\Gamma_i f\|_s^2\, ds.$$

Από την ανισότητα Hölder,

$$\|\Gamma_i f\|_s \leqslant \|\Gamma_i f\|_1^{\vartheta}\|\Gamma_i f\|_2^{1-\vartheta},$$

όπου ο $\vartheta = \vartheta(s) \in [0,1]$ ορίζεται από την $\frac{1}{s} = \frac{\vartheta}{1} + \frac{1-\vartheta}{2}$. Έπεται ότι

$$\int_1^2 \|\Gamma_i f\|_s^2 ds \leqslant \|\Gamma_i f\|_2^2 \int_1^2 b^{2\vartheta(s)} ds,$$

όπου $b = \|\Gamma_i f\|_1/\|\Gamma_i f\|_2 \leqslant 1$. Υπολογίζουμε τώρα το τελευταίο ολοκλήρωμα με $r = 2\vartheta(s)$: παίρνουμε

$$\int_1^2 b^{2\vartheta(s)} ds \leqslant \int_0^2 b^r dr \leqslant \frac{2}{1+\log(1/b)},$$

και έχουμε το θεώρημα. $\square$

4.4 Εφαρμογή: ανισότητες συγκέντρωσης για νόρμες

Το κλασικό θεώρημα του Dvoretzky [33] ισχυρίζεται ότι κάθε χώρος πεπερασμένης διάστασης με νόρμα έχει υπόχωρο «μεγάλης διάστασης» που είναι C-ισομορφικός με τον Ευκλείδειο χώρο, όπου C είναι μια απόλυτη σταθερά. Ο V. Milman [65] βελτίωσε την αρχική εκτίμηση του Dvoretzky δείχνοντας το ακόλουθο θεώρημα.

Θεώρημα 4.4.1 (Dvoretzky-Milman). *Έστω X ένας n-διάστατος χώρος με νόρμα και $\varepsilon \in (0,1)$. Μπορούμε να βρούμε ακέραιο $k \geqslant c\varepsilon^2 \log n$ και έναν k-διάστατο υπόχωρο F του X τέτοιον ώστε $d(F, \ell_2^k) \leqslant 1+\varepsilon$, όπου d είναι η απόσταση Banach-Mazur.*

Σε γεωμετρική γλώσσα, το παραπάνω διατυπώνεται ισοδύναμα ως εξής: Αν K είναι ένα συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$, τότε για κάθε $\varepsilon \in (0,1)$ μπορούμε να βρούμε $k \geqslant c\varepsilon^2 \log n$, έναν υπόχωρο $F \in G_{n,k}$, όπου $G_{n,k}$ είναι η πολλαπλότητα Grassmann όλων των k-διάστατων υποχώρων του $\mathbb{R}^n$, και ένα ελλειψοειδές $\mathcal{E}$ στον F τέτοια ώστε

$$\mathcal{E} \subseteq K \cap F \subseteq (1+\varepsilon)\mathcal{E}.$$

Το παράδειγμα του ℓ_∞^n δείχνει ότι η λογαριθμική εξάρτηση του k από το n είναι βέλτιστη αν σταθεροποιήσουμε μια μικρή τιμή του ε.

Η απόδειξη του Milman για το Θεώρημα 4.4.1 δείχνει ότι, για κάθε $\varepsilon \in (0,1)$, κάθε n-διάστατος χώρος με νόρμα $X = (\mathbb{R}^n, \|\cdot\|)$ έχει υπόχωρο F διάστασης

$$k \geqslant c_1 \varepsilon^2 n (M/b)^2 \tag{4.4.1}$$

με $d(F, \ell_2^k) \leqslant 1+\varepsilon$, όπου $\overline{c}_1 > 0$ είναι μια απόλυτη σταθερά, $b = \max\{\|x\| : x \in S^{n-1}\}$ και $M = \int_{S^{n-1}} \|x\| d\sigma(x)$. Η παράμετρος

$$k(X) = n(M/b)^2$$

είναι η λεγόμενη «κρίσιμη διάσταση» του X. Ο ακριβής τρόπος με τον οποίο συνδέονται μεταξύ τους οι τρεις παράμετροι n, ε και k, οι οποίες εμπλέκονται στο Θεώρημα 4.4.1, δεν έχει αποσαφηνιστεί. Είναι χρήσιμο να ορίσουμε μία νέα παράμετρο: Για δεδομένο $\varepsilon \in (0,1)$ και για κάθε n-διάστατο χώρο με νόρμα $X = (\mathbb{R}^n, \|\cdot\|)$ συμβολίζουμε με $k(X,\varepsilon)$ τον μεγαλύτερο ακέραιο k για τον όποίο υπάρχει k-διάστατος υπόχωρος του X ο οποίος είναι $(1+\varepsilon)$-Ευκλείδειος, και για κάθε $0 < \varepsilon < 1$ ορίζουμε

$$k(n,\varepsilon) = \inf\{k(X,\varepsilon) : \dim(X) = n\}.$$

Το θεώρημα του Milman μας δίνει $k(n,\varepsilon) \geqslant c\varepsilon^2 \log n$, και ο Schechtman [86] απέδειξε το ακόλουθο ισχυρότερο αποτέλεσμα.

Θεώρημα 4.4.2 (Schechtman). *Υπάρχει απόλυτη σταθερά $c_0 > 0$ με την εξής ιδιότητα: για κάθε $\varepsilon \in (0,1)$, κάθε $n \geqslant 1$ και κάθε n-διάστατο χώρο με νόρμα X, υπάρχουν*

$$k \geqslant c_0 \big(\varepsilon / \log^2(\varepsilon^{-1})\big) \log n$$

και k-διάστατος υπόχωρος F του X τέτοια ώστε $d(F, \ell_2^k) \leqslant 1+\varepsilon$.

Σε αυτή την παράγραφο παρουσιάζουμε ένα πιο πρόσφατο αποτέλεσμα των Βαλέττα και Παούρη από το [76]. Στη γλώσσα του μέτρου Gauss και των κανονικών τυχαίων διανυσμάτων, το θεώρημα Dvoretzky-Milman μας λέει ότι για κάθε νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ μπορούμε να βρούμε $T \in GL(n)$ τέτοιον ώστε

$$\mathbb{P}(|\,\|TG\| - \mathbb{E}\|TG\|\,| > t\mathbb{E}\|TG\|) \leqslant C\exp(-ct^2\log n) \tag{4.4.2}$$

για κάθε $t > 0$, όπου $G \sim N(\mathbf{0}, I_n)$. Αυτό προκύπτει από το γεγονός ότι αν επιλέξουμε τον $T \in GL(n)$ έτσι ώστε το $T^{-1}(B_X)$ να έχει ως ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα, τότε η κρίσιμη διάσταση $k(T^{-1}(B_X))$ είναι τουλάχιστον της τάξης του $\log n$. Αυτό είναι μάλιστα ακριβές, διότι $k(B_\infty^n) \approx \log n$. Όμως, σε αυτή την ειδική ακραία περίπτωση του ℓ_∞^n είναι γνωστό ότι

$$ce^{-Ct\log n} \leqslant \mathbb{P}(|\,\|G\|_\infty - \mathbb{E}\|G\|_\infty\,| > t\mathbb{E}\|G\|_\infty) \leqslant Ce^{-ct\log n} \tag{4.4.3}$$

για κάθε $t \in (0,1)$. Τίθεται έτσι, φυσιολογικά, το ακόλουθο ερώτημα.

Ερώτημα 4.4.3. Είναι σωστό ότι για κάθε νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ υπάρχει $T \in GL(n)$ τέτοιος ώστε

$$\mathbb{P}(|\,\|TG\|_\infty - \mathbb{E}\|TG\|_\infty\,| > t\mathbb{E}\|TG\|_\infty) \leqslant C\exp(-c\max\{t^2, t\}\log n)$$

για κάθε $t > 0$;

Θα δούμε ότι το Ερώτημα 4.4.3 έχει καταφατική απάντηση αν υποθέσουμε ότι ο χώρος με νόρμα $X = (\mathbb{R}^n, \|\cdot\|)$ δεν έχει ακραία *unconditional σταθερά* (ο ορισμός δίνεται στην Παράγραφο 4.4.2).

Θεώρημα 4.4.4 (Βαλέττας-Παούρης). *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα. Τότε, υπάρχει $T \in GL(n)$ τέτοιος ώστε, για κάθε $t > 0$,*

$$\mathbb{P}\Big(|\,\|TG\| - \mathbb{E}\|TG\|\,| > t\mathbb{E}\|TG\|\Big) \leqslant C\exp\Big(-c\max\{t^2, t\}\log\Big(\frac{en}{(\mathrm{unc}(X))^2}\Big)\Big).$$

Η απόδειξη του Θεωρήματος 4.4.4 χρησιμοποιεί την L^1-L^2 ανισότητα του Talagrand και ένα τοπολογικό εργαλείο, το κλασικό αντιποδικό θεώρημα Borsuk-Ulam (βλ. [20]).

Χρησιμοποιώντας ένα άλλο κλασικό εργαλείο, το θεώρημα Alon-Milman [2], μπορούμε να συμπεράνουμε ότι σε κάθε n-διάστατο χώρο με νόρμα υπάρχει καλή τοπική unconditional δομή. Χρησιμοποιώντας το Θεώρημα 4.4.4 και την διχοτομία που περιγράψαμε πιο πάνω, οι Βαλέττας και Παούρης απέδειξαν το επόμενο θεώρημα.

Θεώρημα 4.4.5 (Βαλέττας-Παούρης). *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα, που η μοναδιαία του μπάλα B_X έχει ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα, και έστω $0 < \delta < 1/2$. Τότε έχουμε την ακόλουθη διχοτομία:*

- *Είτε ο τυχαίος υπόχωρος E με $\dim(E) = k \geqslant n^{1/2-\delta}$ ικανοποιεί την*

$$\mathbb{P}(|\,\|G\|_{E\cap B_X} - \mathbb{E}\|G\|_{E\cap B_X}\,| > t\mathbb{E}\|G\|_{E\cap B_X}) \leqslant Ce^{-ct^2k}$$

  για κάθε $t > 0$,

- *Ή υπάρχουν υπόχωρος F με $\dim(F) = m \geqslant c\sqrt{n}$ και αντιστρέψιμος γραμμικός μετασχηματισμός $T : F \to F$ ώστε*

$$\mathbb{P}(|\, \|TG\|_{F\cap B_X} - \mathbb{E}\|TG\|_{F\cap B_X}\,| > t\mathbb{E}\|TG\|_{F\cap B_X}) \leqslant Ce^{-c\delta \max\{t^2,t\}\log m}$$

  για κάθε $t > 0$, όπου G είναι ένα τυπικό κανονικό τυχαίο διάνυσμα και $c, C > 0$ είναι απόλυτες σταθερές.

Αυτό το θεώρημα, το οποίο μπορούμε να σκεφτόμαστε ως πιθανοθεωρητική εκδοχή του θεωρήματος Alon-Milman για το μέτρο Gauss, είναι αρκετό για να δώσουμε καταφατική απάντηση στο Ερώτημα 4.4.3.

Θεώρημα 4.4.6. *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα. Υπάρχει $T \in GL(n)$ τέτοιος ώστε*

$$\mathbb{P}(|\, \|TG\| - \mathbb{E}\|TG\|\,| > t\mathbb{E}\|TG\|) \leqslant Ce^{-c\max\{t^2,t\}\log n}$$

για κάθε $t > 0$, όπου G είναι ένα τυπικό κανονικό τυχαίο διάνυσμα και $c, C > 0$ είναι απόλυτες σταθερές.

Το Θεώρημα 4.4.6 έχει άμεση σχέση με την απόδειξη του Milman για το θεώρημα του Dvoretzky. Αν ορίσουμε $k_r(X, \varepsilon)$ τον μέγιστο k για τον οποίο ο τυχαίος k-διάστατος υπόχωρος του X είναι $(1+\varepsilon)$-σφαιρικός με πιθανότητα, ας πούμε, μεγαλύτερη από $1/2$, τότε το επιχείρημα του Milman δείχνει ότι $k_r(X, \varepsilon) \geqslant c\varepsilon^2 k(X)$. Από το Θεώρημα 4.4.6 και ένα γνωστό επιχείρημα διακριτοποίησης μέσω δικτύου, έπεται το εξής αποτέλεσμα.

Πόρισμα 4.4.7 (Βαλέττας-Παούρης)**.** *Για κάθε χώρο με νόρμα $X = (\mathbb{R}^n, \|\cdot\|)$ υπάρχει γραμμική εικόνα B της B_X τέτοια ώστε για κάθε $\varepsilon \in (0,1)$ να έχουμε*

$$k_r(B, \varepsilon) \geqslant c\varepsilon \log n / \log(1/\varepsilon).$$

Με άλλα λόγια, η τυχαία k-διάστατη τομή του B με $k \leqslant c\varepsilon \log n/\log(1/\varepsilon)$ είναι $(1+\varepsilon)$-σφαιρική με πιθανότητα μεγαλύτερη από $1 - n^{-c\varepsilon}$.

Το Πόρισμα 4.4.7 δείχνει επίσης ότι

$$k(n, \varepsilon) \geqslant c\varepsilon \log n/\log(1/\varepsilon),$$

κάτι που βελτιώνει το θεώρημα του Schechtman κατά έναν παράγοντα της τάξης του $\log(1/\varepsilon)$. Αυτό είναι το καλύτερο γνωστό αποτέλεσμα, και το πρόβλημα να δοθεί η βέλτιστη εκτίμηση από κάτω για την συνάρτηση $k(n, \varepsilon)$ παραμένει ανοικτό.

4.4.1 Ανισότητα δύο επιπέδων για την συγκέντρωση στο χώρο του Gauss

Ξεκινάμε από την κλασική ανισότητα συγκέντρωσης στον χώρο του Gauss. Για κάθε συνάρτηση Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ με $|f(x) - f(y)| \leqslant L\|x - y\|_2$ για κάθε $x, y \in \mathbb{R}^n$, έχουμε ότι

$$\mathbb{P}(|f(G) - \mathbb{E}(f(G))| > t) \leqslant 2\exp(-t^2/2L^2) \tag{4.4.4}$$

για κάθε $t > 0$, όπου G είναι ένα τυπικό n-διάστατο κανονικό τυχαίο διάνυσμα. Αν $\|\cdot\|$ είναι μια νόρμα στον $\mathbb{R}^n$ τότε από την (4.4.4) συμπεραίνουμε ότι

$$\mathbb{P}(|\,\|G\| - \mathbb{E}\|G\|\,| > t\mathbb{E}\|G\|) \leqslant 2\exp(-t^2k/2) \tag{4.4.5}$$

για κάθε $t > 0$, όπου

$$k = k(X) = k(B_X) := (\mathbb{E}\|G\|/b)^2$$

είναι η *κρίσιμη διάσταση* του $X = (\mathbb{R}^n, \|\cdot\|)$ και

$$b = b(X) = b(B_X) := \max\{\|x\| : \|x\|_2 = 1\}$$

είναι η σταθερά Lipschitz της νόρμας $\|\cdot\|$. Για την περιοχή των μεγάλων αποκλίσεων είναι γνωστό ότι αυτή η εκτίμηση είναι ακριβής: για κάθε $t \geqslant 1$ έχουμε

$$\mathbb{P}(\|G\| \geqslant (1+t)\mathbb{E}\|G\|) \geqslant c\exp(-Ct^2k), \tag{4.4.6}$$

όπου $c, C > 0$ είναι απόλυτες σταθερές. Για την περιοχή των μικρών αποκλίσεων, όταν $0 < t < 1$, υπάρχουν πολλά παραδείγματα που δείχνουν ότι αυτές οι εκτιμήσεις δεν είναι βέλτιστες. Θα αποδείξουμε την ακόλουθη διχοτομία:

Θεώρημα 4.4.8 (Βαλέττας-Παούρης). *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα, που η μοναδιαία του μπάλα B_X έχει ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα, και έστω $0 < \delta < 1/2$. Τότε, ισχύει η ακόλουθη διχοτομία:*

- *Είτε ο τυχαίος υπόχωρος E με $\dim(E) = k \geqslant n^{1/2-\delta}$ ικανοποιεί την*

  $$\mathbb{P}(|\,\|G\|_{E\cap B_X} - \mathbb{E}\|G\|_{E\cap B_X}\,| > t\mathbb{E}\|G\|_{E\cap B_X}) \leqslant Ce^{-ct^2k}$$

  για κάθε $t > 0$,

- *Ή υπάρχουν υπόχωρος F με $\dim(F) = m \geqslant c\sqrt{n}$ και αντιστρέψιμος γραμμικός μετασχηματισμός $T : F \to F$ ώστε*

  $$\mathbb{P}(|\,\|TG\|_{F\cap B_X} - \mathbb{E}\|TG\|_{F\cap B_X}\,| > t\mathbb{E}\|TG\|_{F\cap B_X}) \leqslant Ce^{-c\delta\max\{t^2,t\}\log m}$$

  για κάθε $t > 0$, όπου G είναι τυπικό κανονικό τυχαίο διάνυσμα και $c, C > 0$ είναι απόλυτες σταθερές.

Το πρώτο εργαλείο που χρειαζόμαστε είναι η L^1-L^2 ανισότητα του Talagrand για τον χώρο του Gauss, την οποία συζητήσαμε, σε ένα γενικότερο πλαίσιο, στην προηγούμενη παράγραφο. Στον χώρο του Gauss παίρνει την ακόλουθη μορφή.

Θεώρημα 4.4.9. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια απολύτως συνεχής συνάρτηση. Τότε,*

$$\mathrm{Var}_{\gamma_n}(f) \leqslant C\sum_{i=1}^n \frac{\|\partial_i f\|_{L^2}^2}{1+\log(\|\partial_i f\|_{L^2}/\|\partial_i f\|_{L^1})}, \tag{4.4.7}$$

όπου $C > 0$ είναι μια απόλυτη σταθερά και $\partial_i f$ είναι η i-οστή μερική παράγωγος της f.

Είναι φυσιολογικό να περιμένουμε ότι αυτό το αποτέλεσμα μπορεί να οδηγήσει σε βελτίωση των γνωστών εκθετικών ανισοτήτων συγκέντρωσης που προκύπτουν από την ανισότητα Poincaré. Αρχίζουμε με μια πρώτη εφαρμογή του Θεωρήματος 4.4.9 σε αυτή την κατεύθυνση. Για κάθε συνάρτηση Lipschitz $f:\mathbb{R}^n\to\mathbb{R}$ ορίζουμε

$$b=b(f)=\inf\{t>0:|f(x)-f(y)|\leqslant t\|x-y\|_2 \text{ για κάθε } x,y\in\mathbb{R}^n\}$$

και

$$a=a(f)=\inf\{t>0:|f(x)-f(y)|\leqslant t\|x-y\|_\infty \text{ για κάθε } x,y\in\mathbb{R}^n\}.$$

Αφού $\|x\|_\infty\leqslant\|x\|_2\leqslant\sqrt{n}\|x\|_\infty$ για κάθε $x\in\mathbb{R}^n$, έχουμε

$$b\leqslant a\leqslant b\sqrt{n}. \tag{4.4.8}$$

Παρατηρήστε ότι αν $\|\cdot\|$ είναι μια νόρμα στον $\mathbb{R}^n$ και η $f:\mathbb{R}^n\to\mathbb{R}$ ικανοποιεί την $|f(x)-f(y)|\leqslant L\|x-y\|$ για κάθε $x,y\in\mathbb{R}^n$, τότε η δυϊκή νόρμα $\|y\|_*=\sup\{\langle x,y\rangle:\|x\|\leqslant 1\}$ της $\|\cdot\|$ ικανοποιεί την

$$\|\nabla f(x)\|_*\leqslant L. \tag{4.4.9}$$

Σημειώνουμε ότι η $\nabla f(x)$ είναι καλά ορισμένη σχεδόν παντού από το θεώρημα του Rademacher.

Υποθέτουμε ότι η f είναι αναλλοίωτη ως προς μεταθέσεις συντεταγμένων. Αυτό σημαίνει ότι $f\circ P_\tau=f$ για κάθε μετάθεση $\tau:[n]\to[n]$, όπου P_τ είναι ο πίνακας μεταθέσεων που ορίζεται από τις $P_\tau(e_i)=e_{\tau(i)}$, $i=1,\ldots,n$. Για κάθε $\lambda>0$ βλέπουμε ότι η $h=e^{\lambda f}$ είναι επίσης αναλλοίωτη ως προς μεταθέσεις συντεταγμένων, και ικανοποιεί την

$$\partial_i h=\partial_i(h\circ P_\tau)=\langle P_\tau^*\circ\nabla h\circ P_\tau,e_i\rangle=(\partial_{\tau(i)}h)\circ P_\tau$$

για κάθε $i=1,\ldots,n$. Αφού ο P_τ είναι ορθογώνιος, βλέπουμε ότι

$$\lambda^p\mathbb{E}\left(e^{p\lambda f}\right)|\partial_i f|^p=\|\partial_i h\|_{L^p}^p=\|\partial_{\tau(i)}h\|_{L^p}^p$$

για κάθε $i=1,\ldots,n$ και κάθε $p>0$. Αφού όλες οι μερικές παράγωγοι της h έχουν την ίδια L^2-νόρμα, για $p=2$ παίρνουμε

$$\|\partial_i h\|_{L^2}^2=\frac{1}{n}\sum_{i=1}^n\|\partial_i h\|_{L^2}^2=\frac{\lambda^2}{n}\mathbb{E}(e^{2\lambda f})\|\nabla f\|_{L^2}^2\leqslant\frac{\lambda^2 b(f)^2}{n}\mathbb{E}(e^{2\lambda f})$$

για κάθε $i=1,\ldots,n$. Παρόμοιο επιχείρημα δείχνει ότι

$$\|\partial_i h\|_{L^1}\leqslant\frac{\lambda a(f)}{n}\mathbb{E}(e^{\lambda f}),\qquad i=1,\ldots,n.$$

Εισάγοντας αυτές τις εκτιμήσεις στο Θεώρημα 4.4.9 βλέπουμε ότι

$$\mathbb{E}(e^{2\lambda f})-(\mathbb{E}(e^{\lambda f}))^2\leqslant\frac{C\lambda^2 b^2}{1+\log(nb^2/a^2)}\mathbb{E}(e^{2\lambda f})$$

για κάθε $\lambda>0$. Θέτοντας

$$\varrho^2=\frac{Cb^2}{1+\log(nb^2/a^2)}$$

και εφαρμόζοντας τα παραπάνω για τη συνάρτηση $F := f - \mathbb{E}(f)$ με $\lambda = s/\varrho$, βλέπουμε ότι η συνάρτηση $\psi(s) := \mathbb{E}(e^{2sF/\varrho})$ έχει τις ακόλουθες δύο ιδιότητες:

$$\lim_{s\to 0}\frac{\psi(s)-1}{s} = 0,$$

και

$$(1-s^2)\psi(s) \leqslant \psi^2(s/2) \quad \text{για κάθε } 0 < s < 1.$$

Τότε, δεν είναι δύσκολο να ελέγξουμε ότι $\psi(s) \leqslant 1/(1-s^2)^{-2}$ για κάθε $0 < s < 1$, και ειδικότερα ότι $\psi(1/\sqrt{2}) \leqslant 4$. Από την κλασική ανισότητα συγκέντρωσης (4.4.4) και την ανισότητα Markov παίρνουμε:

Πρόταση 4.4.10. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση αναλλοίωτη ως προς μεταθέσεις συντεταγμένων, η οποία ικανοποιεί τις $|f(x) - f(y)| \leqslant b\|x - y\|_2$ και $|f(x) - f(y)| \leqslant a\|x - y\|_\infty$ για κάθε $x, y \in \mathbb{R}^n$. Τότε,*

$$\mathbb{P}\big(|f(G) - \mathbb{E}(f(G))| > t\big) \leqslant 4\exp\Big(-c\max\Big\{\frac{t^2}{b^2}, \frac{t}{b}\sqrt{\log(nb^2/a^2)}\Big\}\Big) \tag{4.4.10}$$

για κάθε $t > 0$, όπου G είναι ένα τυπικό n-διάστατο κανονικό τυχαίο διάντσμα.

Θα δείξουμε ότι παρόμοια ανισότητα συγκέντρωσης εξακολουθεί να ισχύει αν αφαιρέσουμε την υπόθεση ότι η f είναι αναλλοίωτη ως προς μεταθέσεις συντεταγμένων. Θα χρειαστούμε το επόμενο λήμμα, το οποίο είναι συνέπεια του Θεωρήματος 4.4.9.

Λήμμα 4.4.11. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια απολύτως συνεχής συνάρτηση. Θέτουμε*

$$R(f) = \frac{\mathbb{E}\|\nabla f(G)\|_2^2}{\sum_{i=1}^n (\mathbb{E}|\partial_i f(G)|)^2}.$$

Τότε,

$$\mathrm{Var}(f(G)) \leqslant C\frac{\mathbb{E}\|\nabla f(G)\|_2^2}{1 + \log R(f)}. \tag{4.4.11}$$

Θα δείξουμε την ακόλουθη ανισότητα απόκλισης δύο επιπέδων, στο πνεύμα της αντίστοιχης ανισότητας του Talagrand για την εκθετική κατανομή.

Πρόταση 4.4.12. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση Lipschitz που ικανοποιεί τις $|f(x)-f(y)| \leqslant b\|x - y\|_2$ και $|f(x) - f(y)| \leqslant a\|x - y\|_\infty$ για κάθε $x, y \in \mathbb{R}^n$, και*

$$\|\partial_i f\|_{L^1} \leqslant A, \qquad i = 1, \dots, n.$$

Αν $F = f - \mathbb{E}(f)$ τότε, για κάθε $\lambda > 0$,

$$\mathrm{Var}(e^{\lambda F}) \leqslant \frac{C\lambda^2 b^2}{\log\big(e + \frac{b^2}{aA}\big)}\mathbb{E}(e^{2\lambda F}). \tag{4.4.12}$$

Επιπλέον, για κάθε $t > 0$,

$$\mathbb{P}\big(|f(G) - \mathbb{E}(f(G))| > t\big) \leqslant 4\exp\Big(-c\max\Big\{\frac{t^2}{b^2}, \frac{t}{b}\sqrt{\log(e + b^2/(aA))}\Big\}\Big). \tag{4.4.13}$$

Απόδειξη. Για κάποιο σταθερό $\lambda > 0$ εφαρμόζουμε το Λήμμα 4.4.11 για τη συνάρτηση $e^{\lambda F}$. Παίρνουμε

$$\mathrm{Var}(e^{\lambda F}) \leqslant C\lambda^2 \frac{\mathbb{E}(\|\nabla f\|_2^2 e^{2\lambda F})}{1 + \log\left(\frac{\mathbb{E}\|\nabla f\|_2^2 e^{2\lambda F}}{w}\right)},$$

όπου

$$w = \sum_{i=1}^{n} (\mathbb{E}|\partial_i f| e^{\lambda F})^2.$$

Αφού

$$\mathbb{E}\|\nabla f\|_2^2 e^{2\lambda F} \leqslant b^2 \mathbb{E}(e^{2\lambda F})$$

και η συνάρτηση $z \mapsto \frac{z}{1+\log(z/w)}$ είναι αύξουσα στο $[w, \infty)$, έχουμε

$$\mathrm{Var}(e^{\lambda F}) \leqslant \frac{C\lambda^2 b^2}{1 + \log(b^2 \mathbb{E}(e^{2\lambda F})/w)} \mathbb{E}(e^{2\lambda F}).$$

Χρησιμοποιώντας την ανισότητα Cauchy-Schwarz, το άνω φράγμα για τις L^1-νόρμες $\|\partial_i f\|_{L^1}$ και το γεγονός ότι $\|\nabla f(x)\|_1 \leqslant a$ σχεδόν παντού, παίρνουμε την (4.4.12). Χρησιμοποιώντας αυτή την ανισότητα μπορούμε να πάρουμε την (4.4.13). □

Παρατήρηση 4.4.13. Ο λογάριθμος στην (4.4.13) είναι όμοιος με αυτόν στην (4.4.10) χωρίς την παράμετρο A. Σημειώνουμε ότι η μικρότερη δυνατή τιμή της A ικανοποιεί την $A \leqslant a$. Στόχος μας είναι να συνθέσουμε τη συνάρτηση με κατάλληλο διαγώνιο πίνακα και να βελτιώσουμε αυτό το φράγμα σε $A \leqslant a/n$, κάτι που είναι απαραίτητο για να επιτύχουμε την εκτίμηση (4.4.10).

4.4.2 Unconditional νόρμες

Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας n-διάστατος χώρος με νόρμα. Υπενθυμίζουμε ότι αν $\{b_i\}_{i=1}^n$ είναι μια βάση του X, τότε η unconditional σταθερά της βάσης $\{b_i\}_{i=1}^n$ ορίζεται από την

$$\mathrm{unc}\{b_i\}_{i=1}^n = \sup\Big\{ \Big\| \sum_{i=1}^n \epsilon_i t_i b_i \Big\| : \Big\| \sum_{i=1}^n t_i b_i \Big\| = 1 \Big\},$$

όπου το supremum είναι πάνω από όλες τις επιλογές προσήμων $\epsilon_i = \pm 1$ και όλες τις ακολουθίες συντελεστών $\{t_i\}_{i=1}^n$. Τότε, η unconditional σταθερά του X είναι η ποσότητα

$$\mathrm{unc}(X) = \inf\{\mathrm{unc}\{b_i\}_{i=1}^n : \{b_i\}_{i=1}^n \text{ είναι μια βάση του } X\}.$$

Λέμε ότι ο X είναι 1-unconditional αν $\mathrm{unc}(X) = 1$. Μια απλή συνέπεια του θεωρήματος του John είναι το γεγονός ότι $\mathrm{unc}(X) \leqslant \sqrt{n}$ για κάθε n-διάστατο χώρο με νόρμα X. Αυτό το φράγμα είναι βέλτιστο υπό την έννοια ότι υπάρχουν n-διάστατοι χώροι με νόρμα E (ένα παράδειγμα μας δίνουν οι τυχαίοι υπόχωροι του ℓ_∞^m που έχουν διάσταση ανάλογη του m) που ικανοποιούν την $\mathrm{unc}(E) \geqslant c\sqrt{n}$ για μια απόλυτη σταθερά $c > 0$.

Θα χρειαστούμε επίσης τον ορισμό της θέσης που ελαχιστοποιεί την παράμετρο M. Λέμε ότι μια νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ ελαχιστοποιεί την παράμετρο M αν

$$\mathbb{E}\|G\| \leqslant \mathbb{E}\|TG\|$$

για κάθε $T \in SL(n)$, όπου G είναι ένα τυπικό n-διάστατο τυχαίο διάνυσμα. Αυτή η θέση που ελαχιστοποιεί την παράμετρο M-χαρακτηρίζεται από την ισοτροπική συνθήκη

$$\int_{\mathbb{R}^n} \langle \nabla \|x\|, \vartheta\rangle \langle x, \vartheta\rangle \, d\gamma_n(x) = \frac{\mathbb{E}\|G\|}{n} \tag{4.4.14}$$

για κάθε $\vartheta \in S^{n-1}$.

Πρόταση 4.4.14. *Έστω $\|\cdot\|$ μια 1-unconditional νόρμα στον $\mathbb{R}^n$ που είναι στη θέση που ελαχιστοποιεί την παράμετρο M. Τότε, έχουμε τις ακόλουθες ανισότητες:*

(α) *Για κάθε $t > 0$,*

$$\mathbb{P}(|\,\|G\| - \mathbb{E}\|G\|\,| > t\mathbb{E}\|G\|) \leqslant C \exp\Big(- c \max\{t^2 k, t\sqrt{k \log(en/k)}\Big), \tag{4.4.15}$$

όπου $k = k(X)$ και $X = (\mathbb{R}^n, \|\cdot\|)$.

(β) *Ειδικότερα $k \geqslant c \log n$, άρα για κάθε $t > 0$,*

$$\mathbb{P}(|\,\|G\| - \mathbb{E}\|G\|\,| > t\mathbb{E}\|G\|) \leqslant C \exp\Big(- c \max\{t^2, t\} \log n\Big), \tag{4.4.16}$$

όπου $C, c > 0$ είναι απόλυτες σταθερές.

Απόδειξη. Θεωρούμε την συνάρτηση $f(x) = \|x\|$. Αφού η f είναι unconditional και κυρτή, βλέπουμε ότι η $x_i \mapsto \partial_i f(x)$ είναι αύξουσα συνάρτηση του $|x_i|$. Τότε, από την ανισότητα αναδιάταξης του Chebyshev παίρνουμε

$$\mathbb{E}|\partial_i f(G)| \cdot \mathbb{E}|g_i| \leqslant \mathbb{E}|g_i \partial_i f(G)| = \frac{\mathbb{E}(f(G))}{n},$$

άρα

$$\|\partial_i f\|_{L^1(\gamma_n)} \leqslant \frac{c_1 \mathbb{E}(f(G))}{n}$$

για κάθε $i = 1, \ldots, n$, όπου $c_1 > 0$ είναι μια απόλυτη σταθερά. Στη συνέχεια παρατηρούμε ότι

$$a(f) = \max\{\|x\| : \|x\|_\infty \leqslant 1\} \quad \text{και} \quad \mathbb{E}\|G\| \geqslant c \cdot a(f). \tag{4.4.17}$$

Για να το δούμε αυτό, σημειώνουμε ότι

$$\mathbb{E}\Big\| \sum_{i=1}^n g_i e_i \Big\| = \mathbb{E}_\epsilon \mathbb{E} \Big\| \sum_{i=1}^n \epsilon_i |g_i| e_i \Big\| \geqslant \mathbb{E}_\epsilon \Big\| \sum_{i=1}^n \epsilon_i \mathbb{E}|g_i| e_i \Big\| = \sqrt{2/\pi} \Big\| \sum_{i=1}^n e_i \Big\|,$$

από την ανισότητα Jensen και την υπόθεση ότι η $\|\cdot\|$ είναι unconditional. Για τον άλλο ισχυρισμό παρατηρούμε ότι

$$\|i : \ell_\infty^m \to X\| = \max_{\|x\|_\infty \leqslant 1} \|x\| = \max_{\epsilon_i = \pm 1} \Big\| \sum_{i=1}^n \epsilon_i e_i \Big\| = \Big\| \sum_{i=1}^n e_i \Big\|.$$

Από την Πρόταση 4.4.12 και αυτές τις εκτιμήσεις, προκύπτει άμεσα η (4.4.15).

(β) Εφαρμόζοντας την (4.4.15) με $t \approx 1$ και παίρνοντας υπόψιν μας την (4.4.6) βλέπουμε ότι $k \geqslant c \log(en/k)$, και έπεται η (4.4.16). $\square$

4.4.3 Συγκέντρωση για νόρμες με καλή unconditional δομή

Δείχνουμε τώρα ότι το Ερώτημα 4.4.3 έχει καταφατική απάντηση για χώρους με νόρμα που δεν έχουν κακή σταθερά unconditional βάσης. Το κύριο αποτέλεσμα είναι το ακόλουθο.

Θεώρημα 4.4.15. *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα. Τότε, υπάρχει $T \in GL(n)$ τέτοιος ώστε, για κάθε $t > 0$,*

$$\mathbb{P}\Big(\big|\,\|TG\| - \mathbb{E}\|TG\|\,\big| > t\mathbb{E}\|TG\|\Big) \leqslant C\exp\Big(-c\max\{t^2,t\}\log\Big(\frac{en}{(\mathrm{unc}(X))^2}\Big)\Big).$$

Η απόδειξη βασίζεται στην Πρόταση 4.4.12 και το θεώρημα Borsuk-Ulam.

Λήμμα 4.4.16. *Έστω $f : \mathbb{R}^m \to \mathbb{R}$ μια C^1-ομαλή συνάρτηση με φραγμένες μερικές παραγώγους, και έστω $q > 0$. Υπάρχει διαγώνιος πίνακας $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_m)$ τέτοιος ώστε*

$$\|\Lambda\|_{\mathrm{HS}} = 1$$

και

$$\|\partial_i(f\circ\Lambda)\|_{L^q(\gamma_m)} = \|\partial_j(f\circ\Lambda)\|_{L^q(\gamma_m)} \quad \text{για κάθε } i,j = 1,\dots,m.$$

Απόδειξη. Για κάθε $1 \leqslant j \leqslant m-1$ θεωρούμε τη συνάρτηση $h_j : S^{m-1} \to \mathbb{R}$ με

$$h_j(\lambda) := \lambda_j\|(\partial_j f)\circ\Lambda\|_{L^q} - \lambda_{j+1}\|(\partial_{j+1} f)\circ\Lambda\|_{L^q},$$

όπου $\Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_m)$. Από το θεώρημα κυριαρχημένης σύγκλισης και την συνέχεια των $\partial_j f$ βλέπουμε ότι η h_j είναι συνεχής. Από την συμμετρία του γ_m προκύπτει ότι η h_j είναι περιττή: για κάθε $\lambda \in S^{m-1}$ έχουμε ότι $h_j(-\lambda) = -h_j(\lambda)$. Τώρα, θεωρούμε την απεικόνιση $H : S^{m-1} \to \mathbb{R}^{m-1}$ που ορίζεται από την

$$H(\lambda_1,\dots,\lambda_m) := (h_1(\lambda),\dots,h_{m-1}(\lambda)).$$

Τότε, η H είναι συνεχής και περιττή. Από το αντιποδικό θεώρημα Borsuk-Ulam συμπεραίνουμε ότι υπάρχει $\lambda \in S^{m-1}$ τέτοιο ώστε $H(\lambda) = 0$, το οποίο σημαίνει ότι

$$\|(\partial_i f)\circ\Lambda\|_{L^q(\gamma_m)} = \|(\partial_j f)\circ\Lambda\|_{L^q(\gamma_m)} \quad \text{για κάθε } i,j = 1,\dots,m. \tag{4.4.18}$$

Έπεται ότι $\|\partial_i(f\circ\Lambda)\|_{L^q} = \|\partial_j(f\circ\Lambda)\|_{L^q}$ για κάθε $i,j = 1,\dots,m$, και έχουμε το λήμμα. □

Παρατήρηση 4.4.17. Αν η f δεν είναι σταθερή σε κανέναν γνήσιο υπόχωρο του $\mathbb{R}^m$ τότε $\lambda_i > 0$ για κάθε $i = 1,\dots,m$. Για να το δούμε αυτό, παρατηρούμε πρώτα ότι το σύνολο $\sigma := \{i : \lambda_i \neq 0\}$ είναι μη κενό. Αν υποθέσουμε ότι $\sigma^c \neq \emptyset$ τότε από την (4.4.18) συμπεραίνουμε ότι $\|(\partial_i f)\circ\Lambda\|_{L^q} = 0$ για κάθε $i \in \sigma$. Λόγω συνέχειας, παίρνουμε $(\partial_i f)\circ\Lambda \equiv 0$ για κάθε $i \in \sigma$, άρα $\partial_i f \equiv \mathbf{0}$ στον $\Lambda(\mathbb{R}^m) = \mathbb{R}^\sigma \equiv \mathrm{span}\{e_i : i \in \sigma\}$. Με άλλα λόγια, η $f\big|_{\mathbb{R}^\sigma}$ είναι σταθερή. Από την (4.4.18) βλέπουμε επίσης ότι όλοι οι λ_j πρέπει να έχουν το ίδιο πρόσημο. Αφού $H(\lambda) = H(-\lambda) = 0$, μπορούμε λοιπόν να υποθέσουμε ότι $\lambda_j \geqslant 0$ για κάθε $j = 1,\dots,m$.

Λήμμα 4.4.18. *Έστω $f : \mathbb{R}^m \to \mathbb{R}$ μια C^1-ομαλή συνάρτηση Lipschitz που δεν είναι σταθερή σε κανέναν γνήσιο υπόχωρο. Τότε, υπάρχουν $\lambda_1,\dots,\lambda_m > 0$ τέτοιοι ώστε $\sum_{j=1}^m \lambda_j^2 = 1$ και*

$$\|\partial_j(f\circ\Lambda)\|_{L^1(\gamma_m)} \leqslant \frac{1}{m}a(f\circ\Lambda), \qquad j = 1,\dots,m$$

όπου $\Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_m)$.

Απόδειξη. Αφού η f έχει υποτεθεί C^1-ομαλή, μπορούμε να θεωρήσουμε έναν διαγώνιο πίνακα Λ ο οποίος ικανοποιεί το συμπέρασμα του Λήμματος 4.4.16. Χρησιμοποιώντας επίσης την Παρατήρηση 4.4.17 μπορούμε να υποθέσουμε ότι $\lambda_j > 0$ για κάθε $1 \leqslant j \leqslant m$. Τώρα, παρατηρούμε ότι

$$\|\partial_j(f\circ\Lambda)\|_{L^1(\gamma_m)} = \frac{1}{m}\sum_{j=1}^m \|\partial_j(f\circ\Lambda)\|_{L^1(\gamma_m)} = \frac{1}{m}\int_{\mathbb{R}^m}\|\nabla(f\circ\Lambda)\|_1\, d\gamma_m \leqslant \frac{a(f\circ\Lambda)}{m}$$

για κάθε $1 \leqslant j \leqslant m$, και έχουμε το λήμμα. □

Για το επόμενο κύριο αποτέλεσμα εισάγουμε την σταθερά (random unconditional divergence) $\mathrm{rud}(X)$ ενός χώρου με νόρμα $X = (\mathbb{R}^m, \|\cdot\|)$, που είναι ο μικρότερος $L > 0$ για τον οποίο μπορούμε να βρούμε μια βάση $\{x_i\}_{i=1}^m$ του X με την ιδιότητα ότι

$$\Big\|\sum_{i=1}^m a_i x_i\Big\| \leqslant L\cdot\mathbb{E}_\epsilon\Big\|\sum_{i=1}^m \epsilon_i a_i x_i\Big\| \tag{4.4.19}$$

για κάθε επιλογή συντελεστών $\{a_i\}_{i=1}^m$. Σημειώνουμε ότι $\mathrm{rud}(X) \leqslant \mathrm{unc}(X)$.

Θεώρημα 4.4.19. *Έστω $X = (\mathbb{R}^m, \|\cdot\|)$ ένας χώρος με νόρμα και έστω $L = \mathrm{rud}(X)$. Υπάρχει $T \in GL(n)$ τέτοιος ώστε*

$$\mathbb{P}\Big(\big|\,\|TG\| - \mathbb{E}\|TG\|\,\big| > t\mathbb{E}\|TG\|\Big) \leqslant C\exp\Big(-c\max\{t^2,t\}\log\Big(e+\frac{m}{L^2}\Big)\Big)$$

για κάθε $t > 0$, όπου G είναι ένα τυπικό m-διάστατο κανονικό τυχαίο διάνυσμα και $c, C > 0$ είναι απόλυτες σταθερές.

Απόδειξη. Εφαρμόζοντας κατάλληλο αντιστρέψιμο γραμμικό μετασχηματισμό μπορούμε να υποθέσουμε ότι

$$\|y\| \leqslant L\cdot\mathbb{E}_\epsilon\Big\|\sum_{i=1}^m \epsilon_i y_i e_i\Big\|$$

για κάθε $y \in \mathbb{R}^m$, όπου $\{e_i\}_{i=1}^m$ είναι η συνήθης βάση του $\mathbb{R}^m$. Ισοδύναμα,

$$\max_{\epsilon_i=\pm1}\Big\|\sum_{i=1}^m \epsilon_i y_i e_i\Big\| \leqslant L\cdot\mathbb{E}_\epsilon\Big\|\sum_{i=1}^m \epsilon_i y_i e_i\Big\| \tag{4.4.20}$$

για κάθε $y \in \mathbb{R}^m$. Υποθέτουμε αρχικά ότι η $\|\cdot\|$ είναι λεία νόρμα. Εφαρμόζοντας το Λήμμα 4.4.18 βρίσκουμε έναν διαγώνιο πίνακα $\Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_m)$ τέτοιον ώστε $\|\Lambda\|_{\mathrm{HS}} = 1$ ανδ $\|\partial_i(f\circ\Lambda)\|_{L^q(\gamma_m)} = \|\partial_j(f\circ\Lambda)\|_{L^q(\gamma_m)}$ για κάθε $i,j = 1,\dots,m$. Θέτουμε

$$a_\Lambda := \|\Lambda : \ell_\infty^m \to X\|, \quad b_\Lambda := \|\Lambda : \ell_2^m \to X\|, \quad k_\Lambda := \frac{(\mathbb{E}\|\Lambda G\|)^2}{b_\Lambda^2}.$$

Συνδυάζοντας την (4.4.13) με το Λήμμα 4.4.18 παίρνουμε

$$\mathbb{P}\Big(\big|\,\|\Lambda G\| - \mathbb{E}\|\Lambda G\|\,\big| > t\mathbb{E}\|\Lambda G\|\Big) \leqslant 4\exp\Big(-ct\sqrt{k_\Lambda\log\Big(e+\frac{mb_\Lambda^2}{a_\Lambda^2}\Big)}\Big) \tag{4.4.21}$$

για κάθε $t > 0$. Χρησιμοποιώντας επίσης την (4.4.20), και στο τέλος την αρχή της συστολής, παίρνουμε

$$a_\Lambda = \max_{\epsilon_i=\pm 1} \Big\| \sum_{i=1}^m \epsilon_i \lambda_i e_i \Big\| \leqslant L \cdot \mathbb{E}_\epsilon \Big\| \sum_{i=1}^m \epsilon_i \lambda_i e_i \Big\| \leqslant L\sqrt{\frac{\pi}{2}}\, \mathbb{E}\|\Lambda G\|.$$

Εισάγοντας αυτή την εκτίμηση στην (4.4.21) έχουμε

$$\mathbb{P}\Big(| \, \|\Lambda G\| - \mathbb{E}\|\Lambda G\| \, | > t\mathbb{E}\|\Lambda G\| \Big) \leqslant 4\exp\Big(-ct\sqrt{k_\Lambda \log\Big(e + \frac{m}{L^2 k_\Lambda}\Big)}\Big)$$

για κάθε $t > 0$. Εφαρμόζοντας αυτή την ανισότητα με $t \approx 1$ και παίρνοντας υπόψιν την (4.4.6) βλέπουμε ότι $k_\Lambda \geqslant c\log(e + m/L^2)$. Αυτό ολοκληρώνει την απόδειξη στη λεία περίπτωση.

Για την γενική περίπτωση χρησιμοποιούμε το γεγονός ότι για κάθε νόρμα $\|\cdot\|$ και κάθε $\delta \in (0,1)$ μπορούμε να βρούμε μια λεία νόρμα $\|\cdot\|_\delta$ τέτοια ώστε

$$(1-\delta)\|x\| \leqslant \|x\|_\delta \leqslant (1+\delta)\|x\| \tag{4.4.22}$$

για κάθε $x \in \mathbb{R}^m$. Σταθεροποιούμε $0 < \delta \leqslant (7 + \log m)^{-1}$ και εφαρμόζουμε το κύριο αποτέλεσμα (το οποίο έχουμε ήδη αποδείξει στη λεία περίπτωση) για τη νόρμα $\|\cdot\|_\delta$: βρίσκουμε $T = T_\delta \in GL(m)$ τέτοιον ώστε

$$\mathbb{P}\Big(| \, \|TG\|_\delta - \mathbb{E}\|TG\|_\delta \, | > t\mathbb{E}\|TG\|_\delta \Big) \leqslant 4\exp(-ct\log(e + m/L_\delta^2)$$

για κάθε $t > 0$, όπου $L_\delta = \mathrm{rud}(X_\delta)$ και $X_\delta = (\mathbb{R}^m, \|\cdot\|_\delta)$. Στη συνέχεια, χρησιμοποιώντας την (4.4.22) ελέγχουμε ότι αν $t > 8\delta$ τότε

$$\mathbb{P}\Big(| \, \|TG\| - \mathbb{E}\|TG\| \, | > t\mathbb{E}\|TG\| \Big) \leqslant \mathbb{P}\Big(| \, \|TG\|_\delta - \mathbb{E}\|TG\|_\delta \, | > \frac{t}{2}\mathbb{E}\|TG\|_\delta \Big).$$

Έπεται ότι

$$\mathbb{P}\Big(| \, \|TG\| - \mathbb{E}\|TG\| \, | > t\mathbb{E}\|TG\| \Big) \leqslant 4\exp\Big(-\frac{c}{2}t\log(e + m/L_\delta^2\Big)$$

για κάθε $t > 8\delta$. Αλλάζοντας λίγο τις απόλυτες σταθερές βλέπουμε ότι αυτή η εκτίμηση εξακολουθεί να ισχύει για κάθε $t > 0$. Παρατηρώντας ότι

$$L_\delta \leqslant \frac{1+\delta}{1-\delta} L \leqslant 2L,$$

ολοκληρώνουμε την απόδειξη. □

4.4.4 Πιθανοθεωρητική διχοτομία και το θεώρημα του Dvoretzky

Μπορούμε τώρα να αποδείξουμε το Θεώρημα 4.4.8 και τα πορίσματά του. Θα εκμεταλλευτούμε δύο σημαντικά εργαλεία:

(i) *Λήμμα Dvoretzky-Rogers* [34]. Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα τέτοιος ώστε η μοναδιαία μπάλα B_X του X να έχει ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα. Τότε υπάρχει ορθοκανονική βάση $\{v_i\}_{i=1}^n$ τέτοια ώστε

$$1 = \|v_k\|_2 \geqslant \|v_k\| \geqslant \sqrt{\frac{n-k+1}{n}}$$

για κάθε $k = 1, \ldots, n$. Ειδικότερα, $\|v_k\| \geqslant 1/\sqrt{2}$ για κάθε $1 \leqslant k \leqslant \lfloor n/2 \rfloor$.

Μπορούμε μάλιστα να τροποποιήσουμε την ορθοκανονική βάση $\{v_k\}_{k=1}^n$ και να πάρουμε μια νέα ορθοκανονική βάση $\{w_k\}_{k=1}^n$ με την ιδιότητα ότι $\|w_k\| \geqslant 1/4$ για κάθε $1 \leqslant k \leqslant n$.

(ii) *Θεώρημα Alon-Milman* [2]. Έστω X ένας χώρος με νόρμα και έστω $T : \ell_\infty^n \to X$ με $\|Te_i\| \geqslant 1$ για κάθε $i = 1, \ldots, n$. Ορίζουμε

$$a = \|T : \ell_\infty^n \to X\| \quad \text{και} \quad M_n = \mathbb{E}_\epsilon \Big\| \sum_{i=1}^n \epsilon_i T e_i \Big\|.$$

Τότε, υπάρχει $\sigma \subset [n]$ με $|\sigma| \geqslant cn/a$ τέτοιο ώστε

$$\frac{1}{2} \max_{i\in\sigma} |s_i| \leqslant \Big\| \sum_{i\in\sigma} s_i T e_i \Big\| \leqslant 4M_n \max_{i\in\sigma} |s_i|$$

για κάθε ακολουθία συντελεστών $\{s_i\}_{i\in\sigma}$.

Η απόδειξη των Alon και Milman έδινε $\sigma \subset [n]$ με $|\sigma| \geqslant c\sqrt{n}/M_n$. Η ισχυρότερη διατύπωση του θεωρήματος, όπως το περιγράψαμε παραπάνω, οφείλεται στον Talagrand.

Το Θεώρημα 4.4.8 είναι συνέπεια του ακόλουθου πιο γενικού θεωρήματος.

Θεώρημα 4.4.20. *Έστω $X = (\mathbb{R}^n, \|\cdot\|)$ ένας χώρος με νόρμα που η μοναδιαία του μπάλα B_X έχει ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα, και έστω $0 < \delta < 1/2$. Τότε ισχύει τουλάχιστον ένα από τα ακόλουθα:*

- *Είτε $k(X) \geqslant n^{1/2-\delta}$,*
- *Ή υπάρχουν υπόχωρος F με $\dim(F) = m \geqslant c\sqrt{n}$ και γραμμικός ισομορφισμός $T : F \to F$ τέτοιος ώστε για κάθε $t > 0$*

$$\mathbb{P}\Big(\big| \|TZ\| - \mathbb{E}\|TZ\| \big| > t\mathbb{E}\|TZ\| \Big) \leqslant C \exp\Big(-c\delta \max\{t^2, t\} \log m \Big)$$

όπου $Z \sim N(\mathbf{0}, I_F)$ και $c, C > 0$ είναι απόλυτες σταθερές.

Θα χρειαστούμε ένα ακόμα βασικό εργαλείο, την αρχή της συστολής.

Θεώρημα 4.4.21 (αρχή της συστολής)**.** *Έστω $\{\epsilon_i\}_{i=1}^n$ μια ακολουθία ανεξάρτητων συμμετρικών τυχαίων μεταβλητών Bernoulli και έστω $F : [0,\infty) \to [0,\infty)$ μια αύξουσα κυρτή συνάρτηση. Αν $x_1, \ldots, x_n$ είναι διανύσματα σε έναν χώρο με νόρμα X και a_i, b_i είναι πραγματικοί συντελεστές με $|a_i| \leqslant |b_i|$ για κάθε $i = 1, \ldots, n$, τότε*

$$\mathbb{E}F\Big(\Big\| \sum_{i=1}^n \epsilon_i \alpha_i x_i \Big\| \Big) \leqslant \mathbb{E}F\Big(\Big\| \sum_{i=1}^n \epsilon_i b_i x_i \Big\| \Big).$$

Απόδειξη. Θεωρούμε την συνάρτηση $\varphi : \mathbb{R}^n \to \mathbb{R}$ που ορίζεται από την

$$\varphi(\alpha_1, \ldots, \alpha_n) = \mathbb{E}F\Big(\Big\| \sum_i \epsilon_i \alpha_i x_i \Big\| \Big).$$

Παρατηρούμε ότι η φ είναι κυρτή συνάρτηση, διότι οι F, $\|\cdot\|$ είναι κυρτές και η F είναι αύξουσα. Τότε, ο περιορισμός της φ στο συμπαγές κυρτό σύνολο $\prod_{i=1}^n[-b_i,b_i]$ παίρνει τη μέγιστη τιμή του σε κάποιο ακραίο σημείο $b_\zeta=(\zeta_1b_1,\dots,\zeta_nb_n)$ όπου $\zeta=\{\zeta_i\}_{i=1}^n\in E_2^n$. Από την συμμετρία των τυχαίων μεταβλητών $(\epsilon_i)_{i\leqslant n}$ έχουμε $\varphi(b_\zeta)=\varphi(b_1,\dots,b_n)$ για κάθε $\zeta\in E_2^n$ και έπεται το ζητούμενο. □

Απόδειξη του Θεωρήματος 4.4.20. Έστω $0<\delta<1/2$. Μπορούμε να βρούμε ορθοκανονική βάση $\{w_j\}_{j=1}^n$ με $\|w_j\|\geqslant 1/4$ για κάθε $j=1,\dots,n$. Ας υποθέσουμε ότι $k(X)\leqslant n^{1/2-\delta}$. Τότε, έχουμε

$$\sqrt{\frac{2}{\pi}}\mathbb{E}\Big\|\sum_{i=1}^n\epsilon_iw_i\Big\|\leqslant\mathbb{E}\Big\|\sum_{i=1}^n g_iw_i\Big\|=\mathbb{E}\|G\|=\sqrt{k(X)},$$

χρησιμοποιώντας και την αρχή της συστολής στην πρώτη ανισότητα. Από το θεώρημα Alon-Milman μπορούμε να βρούμε υποσύνολο $\sigma\subset[n]$ με $|\sigma|\geqslant cn/a\geqslant c\sqrt{n}$, όπου $a=\|i:\ell_\infty^n\to X\|$, τέτοιο ώστε

$$\frac{1}{8}\max_{i\in\sigma}|s_i|\leqslant\Big\|\sum_{i\in\sigma}s_iw_i\Big\|\leqslant 4M_n\max_{i\in\sigma}|s_i|$$

για κάθε ακολουθία συντελεστών $\{s_i\}_{i\in\sigma}$, όπου $M_n=\mathbb{E}\left\|\sum_i\epsilon_iw_i\right\|$. Αυτό δείχνει ότι ο υπόχωρος $(F,\|\cdot\|)$, όπου $F=\mathrm{span}\{w_i:i\in\sigma\}$, ικανοποιεί την $d(F,\ell_\infty^\sigma)\leqslant 32M_n\leqslant C\sqrt{k(X)}$. Από το Θεώρημα 4.4.19 μπορούμε να βρούμε γραμμικό ισομορφισμό $T:F\to F$ τέτοιον ώστε

$$\begin{aligned}\mathbb{P}\Big(\big|\,\|TZ\|-\mathbb{E}\|TZ\|\,\big|>t\mathbb{E}\|TZ\|\Big)&\leqslant C\exp\Big(-c\max\{t^2,t\}\log(e+c|\sigma|/k(X))\Big)\\&\leqslant C\exp\Big(-c'\delta\max\{t^2,t\}\log|\sigma|\Big)\end{aligned}$$

για κάθε $t>0$. Αυτό ολοκληρώνει την απόδειξη. □

Το επόμενο θεώρημα είναι σχεδόν άμεση συνέπεια του Θεωρήματος 4.4.20 και μας δίνει το Θεώρημα 4.4.6 και το Πόρισμα 4.4.7.

Θεώρημα 4.4.22. *Έστω $X=(\mathbb{R}^n,\|\cdot\|)$ ένας χώρος με νόρμα. Υπάρχει γραμμικός μετασχηματισμός $S\in GL(n)$ τέτοιος ώστε, για κάθε $t>0$,*

$$\mathbb{P}\Big(\big|\,\|SG\|-\mathbb{E}\|SG\|\,\big|>t\mathbb{E}\|SG\|\Big)\leqslant C\exp\Big(-c\max\{t^2,t\}\log n\Big),$$

όπου $G\sim N(\mathbf{0},I_n)$ και $c,C>0$ είναι απόλυτες σταθερές. Ειδικότερα, για κάθε $\varepsilon\in(0,1)$, η τυχαία k-διάστατη τομή του $S^{-1}(B_X)$ με $k\leqslant c\varepsilon\log n/\log(1/\varepsilon)$ είναι $(1+\varepsilon)$-σφαιρική με πιθανότητα μεγαλύτερη από $1-n^{-c\varepsilon}$.

Απόδειξη. Μπορούμε να υποθέσουμε ότι η B_X έχει ελλειψοειδές μέγιστου όγκου την Ευκλείδεια μοναδιαία μπάλα και $k(X)\leqslant n^{1/3}$, αλλιώς δεν έχουμε τίποτα να αποδείξουμε. Από το Θεώρημα 4.4.20 μπορούμε να βρούμε έναν υπόχωρο F του $\mathbb{R}^n$ με $\dim(F)=m\approx\sqrt{n}$ και $T:F\to F$ με $T\in GL(F)$ έτσι ώστε, για κάθε $t>0$,

$$\mathbb{P}\Big(\big|\,\|TZ\|-\mathbb{E}\|TZ\|\,\big|>t\mathbb{E}\|TZ\|\Big)\leqslant C\exp\Big(-ct\log n\Big)$$

όπου $Z\sim N(\mathbf{0},I_F)$ και $c,C>0$ είναι απόλυτες σταθερές.

Ορίζουμε έναν τελεστή $S:\mathbb{R}^n\to\mathbb{R}^n$ θέτοντας

$$S(x,y)=Tx+\lambda y,\qquad x\in F, y\in F^{\perp},$$

όπου

$$\lambda=\frac{\mathbb{E}\|TZ\|}{\mathbb{E}\|W\|_2\log n},\qquad W\sim N(\mathbf{0},I_{F^{\perp}}).$$

Από την αρχή της συστολής, την τριγωνική ανισότητα και το γεγονός ότι $\|\cdot\|\leqslant\|\cdot\|_2$, βλέπουμε ότι

$$\mathbb{E}\|TZ\|\leqslant\mathbb{E}\|SG\|\leqslant\Big(1+\frac{1}{\log n}\Big)\mathbb{E}\|TZ\|, \tag{4.4.23}$$

όπου $G\sim N(\mathbf{0},I_n)$ και $Z\sim N(\mathbf{0},I_F)$. Θα χρησιμοποιήσουμε τον επόμενο ισχυρισμό.

Ισχυρισμός. Για κάθε $0<t<1$ έχουμε ότι

$$\mathbb{P}(|\,\|SG\|-\mathbb{E}\|SG\|\,|>t\mathbb{E}\|SG\|)\leqslant Ce^{-ct\log n}. \tag{4.4.24}$$

Απόδειξη του Ισχυρισμού. Έστω $Z\sim N(\mathbf{0},I_F)$, $W\sim N(\mathbf{0},I_{F^{\perp}})$ και $G=Z+W$. Μπορούμε να γράψουμε

$$\begin{aligned}
&\mathbb{P}(\|SG\|>(1+t)\mathbb{E}\|SG\|)\\
&\quad\leqslant\mathbb{P}(\|TZ\|>(1+t)\mathbb{E}\|SG\|-\lambda\|W\|)\\
&\quad\leqslant\mathbb{P}(\|TZ\|>(1+t)\mathbb{E}\|TZ\|-10\lambda\mathbb{E}\|W\|_2)+\mathbb{P}(\|W\|>10\mathbb{E}\|W\|_2)\\
&\quad\leqslant\mathbb{P}\Big(\|TZ\|>(1+t)\mathbb{E}\|TZ\|-\frac{10\mathbb{E}\|TZ\|}{\log n}\Big)+\mathbb{P}(\|W\|_2>10\mathbb{E}\|W\|_2)\\
&\quad\leqslant\mathbb{P}\Big(\|TZ\|>\Big(1+\frac{t}{2}\Big)\mathbb{E}\|TZ\|\Big)+\mathbb{P}(\|W\|_2>10\mathbb{E}\|W\|_2)\\
&\quad\leqslant Ce^{-ct\log n},
\end{aligned}$$

για κάθε $\frac{20}{\log n}<t<1$, χρησιμοποιώντας το γεγονός ότι

$$\mathbb{P}(\|W\|_2>10\mathbb{E}\|W\|_2)\leqslant e^{-cn},\qquad W\sim N(\mathbf{0},I_{F^{\perp}}).$$

Για την απόκλιση κάτω από τον μέσο, γράφουμε

$$\begin{aligned}
\mathbb{P}(\|SG\|<(1-t)\mathbb{E}\|SG\|)&\leqslant\mathbb{P}\Big(\|TG\|<(1-t)\Big(1+\frac{1}{\log n}\Big)\mathbb{E}\|TG\|\Big)\\
&\leqslant\mathbb{P}\Big(\|TG\|<\Big(1-\frac{t}{2}\Big)\mathbb{E}\|TG\|\Big)\\
&\leqslant Ce^{-ct\log n}
\end{aligned}$$

για κάθε $\frac{2}{\log n}<t<1$, παίρνοντας υπόψιν μας την (4.4.23). Τώρα, έχουμε τον ισχυρισμό για κάθε $t\in(0,1)$ επιλέγοντας κατάλληλα τις απόλυτες σταθερές. Έχοντας αποδείξει την (4.4.24) βλέπουμε από την (4.4.6) ότι $k(S^{-1}(B_X))\geqslant c\log n$, και έπεται το ζητούμενο. □

ΚΕΦΑΛΑΙΟ 5

Ανισότητα μεταφοράς με τετραγωνικό κόστος

Σε αυτό το κεφάλαιο παρουσιάζουμε την απόδειξη της ανισότητας μεταφοράς με τετραγωνικό κόστος, του Talagrand, χρησιμοποιώντας την εντροπία και τη λογαριθμική ανισότητα Sobolev. Υπενθυμίζουμε ότι αν $(X, \mathcal{A})$ και $(Y, \mathcal{B})$ είναι δύο μετρήσιμοι χώροι, και αν θεωρήσουμε τον χώρο γινόμενο $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ και ένα μέτρο $\pi : \mathcal{A} \otimes \mathcal{B} \to [0, \infty]$ στον $(X \times Y, \mathcal{A} \otimes \mathcal{B})$, τότε:

(α) το πρώτο περιθώριο μέτρο του π είναι το μέτρο $\pi_X : \mathcal{A} \to [0, \infty]$ με $\pi_X(A) = \pi(A \times Y)$, και

(β) το δεύτερο περιθώριο μέτρο του π είναι το μέτρο $\pi_Y : \mathcal{B} \to [0, \infty]$ με $\pi_Y(B) = \pi(X \times B)$.

Έστω μ, ν δύο Borel μέτρα πιθανότητας στον $\mathbb{R}^n$. Θεωρούμε την $w : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ με $w(x, y) = \sum_{i=1}^n (x_i - y_i)^2$. Ορίζουμε

$$T_w(\mu, \nu) = \inf \int_{\mathbb{R}^n \times \mathbb{R}^n} w(x, y)\, d\pi(x, y)$$

όπου το infimum είναι πάνω από όλα τα Borel μέτρα πιθανότητας π στον $\mathbb{R}^n \times \mathbb{R}^n$ για τα οποία έχουμε $\pi_X = \mu$ και $\pi_Y = \nu$. Έστω επίσης $d\gamma(x) = \frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2} dx$ το τυπικό μέτρο Gauss στα Borel υποσύνολα του $\mathbb{R}^n$.

Η ανισότητα του Talagrand ισχυρίζεται ότι για κάθε μέτρο πιθανότητας μ το οποίο είναι απολύτως συνεχές ως προς το γ (θα γράφουμε $\mu \ll \gamma$) με $f = \frac{d\mu}{d\gamma}$ ισχύει ότι

$$T_w(\mu, \gamma) \leqslant 2 \int_{\mathbb{R}^n} f \, \log f \, d\gamma = 2 \int_{\mathbb{R}^n} \log f \, d\mu.$$

5.1 Μεταφορά και συγκέντρωση του μέτρου

Πριν αποδείξουμε την ανισότητα του Talagrand θα δούμε πώς αυτή σχετίζεται με την συγκέντρωση του μέτρου. Έστω (X, d) ένας πλήρης και διαχωρίσιμος μετρικός χώρος. Συμβολίζουμε με $\mathcal{P}(X)$ το σύνολο των Borel μέτρων πιθανότητας στον X και με $\mathcal{M}(X)$ το σύνολο των πεπερασμένων, προσημασμένων μέτρων στον X, δηλαδή τον διανυσματικό χώρο που παράγεται από τον $\mathcal{P}(X)$ εφοδιασμένο με τη νόρμα της ολικής κύμανσης:

$$\|\mu\|_{\mathrm{TV}} = \inf\{\mu_+(X) + \mu_-(X)\}$$

όπου το infimum παίρνεται πάνω από όλα τα μη αρνητικά μέτρα μ_+, μ_- για τα οποία το μ γράφεται στην μορφή $\mu = \mu_+ - \mu_-$. Συμβολίζουμε επίσης με $H(\nu \mid \mu)$ την σχετική εντροπία του ν ως προς το μ που ορίζεται από την

$$H(\nu \mid \mu) = \mathrm{Ent}_\mu\Big(\frac{d\nu}{d\mu}\Big) = \int_X \log\frac{d\nu}{d\mu}\, d\mu,$$

οποτεδήποτε το ν είναι απολύτως συνεχές ως προς το μ ($\nu \ll \mu$) με παράγωγο Radon-Nikodym την $\frac{d\nu}{d\mu}$.

Έστω $\mathcal{P}_1(X)$ η κλάση όλων των Borel μέτρων πιθανότητας μ στον (X,d) που ικανοποιούν την

$$\int_X d(x,x_0)\, d\mu(x) < \infty$$

για κάθε, ισοδύναμα για κάποιο, $x_0 \in X$. Μια κλασική ανισότητα των Pinsker,Csizsár και Kullback είναι η ακόλουθη.

Θεώρημα 5.1.1 (Pinsker-Csizsár-Kullback)**.** *Έστω $\mu, \nu \in \mathcal{P}_1(X)$ με $\nu \ll \mu$. Τότε,*

$$\|\mu - \nu\|_{\mathrm{TV}}^2 \leqslant \frac{1}{2} H(\nu \,|\, \mu).$$

Απόδειξη. Θέτουμε $f = \frac{d\nu}{d\mu}$. Παρατηρούμε ότι

$$\frac{1}{2}\int_X |1-f|\, d\mu = \sup_A |\mu(A) - \nu(A)| =: \|\mu - \nu\|_{\mathrm{TV}}, \tag{5.1.1}$$

όπου το supremum είναι πάνω από όλα τα Borel υποσύνολα A του X. Συνεπώς, αρκεί να δείξουμε ότι

$$\Big(\int_X |1-f|\, d\mu\Big)^2 \leqslant 2\mathrm{Ent}_\mu(f).$$

Για κάθε $s \in [0,1]$ ορίζουμε $f_s = sf + 1 - s$, οπότε $\int f_s d\mu = 1$, και θεωρούμε την

$$\Lambda(s) = 2\mathrm{Ent}_\mu(f_s) - \Big(\int_X |1-f_s|\, d\mu\Big)^2 = 2\mathrm{Ent}_\mu(f_s) - s^2\Big(\int_X |1-f|\, d\mu\Big)^2.$$

Θέλουμε να δείξουμε ότι $\Lambda(1) \geqslant 0$. Παρατηρούμε ότι $\Lambda(0) = \Lambda'(0) = 0$ και

$$\Lambda''(s) = 2\int_X \frac{(1-f)^2}{f_s}\, d\mu - 2\Big(\int_X |1-f|\, d\mu\Big)^2$$

για κάθε $s \in [0,1]$. Εφαρμόζοντας την ανισότητα Cauchy-Schwarz στο ολοκλήρωμα

$$\int_X (|1-f|/\sqrt{f_s})\sqrt{f_s} d\mu,$$

βλέπουμε ότι $\Lambda'' \geqslant 0$, απ' όπου έπεται ότι $\Lambda(s) \geqslant 0$ για κάθε $s \in [0,1]$. $\square$

Έστω (X,d) ένας πλήρης και διαχωρίσιμος μετρικός χώροος. Έστω $c : X \times X \to [0,\infty]$ μια κάτω ημισυνεχής συνάρτηση. Υποθέτουμε ότι η c είναι συμμετρική, δηλαδή $c(y,x) = c(x,y)$ για κάθε $x, y \in X$. Για δεδομένα $\mu, \nu \in \mathcal{P}(X)$, το βέλτιστο κόστος μεταφοράς του μ στο ν με συνάρτηση κόστους την c ορίζεται από την

$$W_c(\mu,\nu) = \inf \int_{X\times X} c(x,y)\, d\pi(x,y), \tag{5.1.2}$$

όπου το infimum είναι πάνω από όλα τα μέτρα πιθανότητας $\pi \in \mathcal{P}(X\times X)$ που έχουν περιθώρια τα μ και ν (τότε λέμε ότι το π είναι ένα «σχέδιο μεταφοράς» από το μ στο ν). Αυτό είναι ισοδύναμο με το να ισχύει η

$$\int_{X\times X} (\varphi(x) + \psi(y))\, d\pi(x,y) = \int_X \varphi(x)\, d\mu(x) + \int_X \psi(y)\, d\nu(y)$$

για όλες τις φραγμένες μετρήσιμες συναρτήσεις $\varphi, \psi : X \to \mathbb{R}$. Ο αριθμός $W_c(\mu,\nu)$ ονομάζεται μερικές φορές «απόσταση Wasserstein» των μ και ν ως προς την c. Άλλοι συγγραφείς κρατούν τον όρο «απόσταση Wasserstein» για την περίπτωση $c(x,y) = d(x,y)^p$, για κάποιο $p \geqslant 1$, και τότε ορίζουν $W_p(\mu,\nu) = W_c^{1/p}(\mu,\nu)$. Ιδιαίτερα σημαντικές είναι οι περιπτώσεις $p = 1$ και $p = 2$.

Αποδεικνύεται, με χρήση του θεωρήματος του Prokhorov, ότι αν η c είναι μη αρνητική και κάτω ημισυνεχής τότε υπάρχει πάντα ένα σχέδιο μεταφοράς π_0 από το μ στο ν το οποίο είναι βέλτιστο, δηλαδή

$$W_c(\mu,\nu) = \int_{X\times X} c(x,y)\, d\pi_0(x,y).$$

Ένα βασικό εργαλείο στη θεωρία της μεταφοράς του μέτρου είναι το θεώρημα δυϊσμού του Kantorovich, το οποίο δίνει μια δυϊκή αναπαράσταση του βέλτιστου κόστους μεταφοράς.

Θεώρημα 5.1.2 (Kantorovich). *Έστω X ένας πλήρης και διαχωρίσιμος μετρικός χώρος και έστω $\mu, \nu \in P(X)$ μέτρα πιθανότητας. Έστω $c : X \times X \to [0,\infty]$ κάτω ημισυνεχής συνάρτηση. Τότε,*

$$\begin{aligned} W_c(\mu,\nu) \;&=\; \sup \Big\{ \int_X \varphi d\mu + \int_X \psi d\nu : \\ &\quad (\varphi,\psi) \in L^1(\mu) \times L^1(\nu), \varphi(x) + \psi(y) \leqslant c(x,y) \text{ ισχύει } (\mu,\nu)\text{-σχεδόν παντού} \Big\}. \end{aligned}$$

Επιπλέον, το infimum στον ορισμό της $W_c(\mu,\nu)$ πιάνεται, και το supremum του δεξιού μέλους δεν αλλάζει αν περιοριστούμε σε ζεύγη (φ,ψ) φραγμένων και συνεχών συναρτήσεων.

Στην περίπτωση όπου η συνάρτηση κόστους δίνεται από κάποια μετρική $\tilde{d}$, όχι αναγκαστικά την ίδια μετρική με την αρχική, η συνθήκη $\varphi(x) + \psi(y) \leqslant \tilde{d}(x,y)$ προφανώς ικανοποιείται αν η $\psi = -\varphi$ και είναι 1-Lipschitz, και μάλιστα αρκεί να θεωρήσουμε μόνο τέτοια ζευγάρια. Έτσι, το θεώρημα του Kantorovich παίρνει την ακόλουθη μορφή.

Θεώρημα 5.1.3 (Kantorovich-Rubinstein). *Έστω (X,d) ένας πλήρης και διαχωρίσιμος μετρικός χώρος και έστω $\tilde{d}$ μια άλλη μετρική στον X, η οποία είναι κάτω ημισυνεχής ως προς την d. Τότε*

$$W_{\tilde{d}}(\mu,\nu) = \sup \left(\int_X \varphi\, d\mu - \int_X \varphi\, d\nu \right), \tag{5.1.3}$$

όπου το supremum είναι πάνω από όλες τις συναρτήσεις φ που είναι 1-Lipschitz ως προς την $\tilde{d}$.

Θεωρούμε την τετριμμένη μετρική

$$d_0(x,y) = \mathbf{1}_{x\neq y}(x,y)$$

στο $X \times X$. Τότε, μια 1-Lipschitz συνάρτηση ως προς αυτή τη μετρική πρέπει να έχει εικόνα μέσα σε ένα διάστημα μήκους 1, συνεπώς το Θεώρημα 5.1.3 δείχνει ότι η $W_{d_0}(\mu,\nu)$ είναι η απόσταση ολικής κύμανσης $\|\mu-\nu\|_{\mathrm{TV}}$ των μ και ν.

Ανισότητες όπως αυτή του Θεωρήματος 5.1.1 σχετίζονται άμεσα με την συγκέντρωση του μέτρου. Έστω (X,d) μετρικός χώρος και μ και ν δύο Borel μέτρα πιθανότητας στην κλάση $\mathcal{P}_1(X)$. Η απόσταση Wasserstein των μ και ν ορίζεται από την

$$W_1(\mu,\nu) = \inf \iint d(x,y)\, d\pi(x,y),$$

όπου το infimum είναι πάνω από όλα τα μέτρα πιθανότητας π στον $X\times X$ που έχουν περιθώρια τα μ και ν. Η απόσταση της ολικής κύμανσης αντιστοιχεί στην τετριμμένη μετρική.

Σταθεροποιούμε $\mu \in \mathcal{P}_1(X)$ και υποθέτουμε ότι υπάρχει σταθερά $C>0$ τέτοια ώστε

$$W_1(\mu,\nu) \leqslant \sqrt{2CH(\nu \mid \mu)} \tag{5.1.4}$$

για κάθε $\nu \in \mathcal{P}_1(X)$. Έστω A και B δύο Borel υποσύνολα του X με $\mu(A)>0$ και $\mu(B)>0$. Θεωρούμε τις δεσμευμένες πιθανότητες $\mu_A = \mu(\cdot \mid A)$ και $\mu_B = \mu(\cdot \mid B)$. Από την τριγωνική ανισότητα για την W_1 και από την (5.1.4) παίρνουμε

$$\begin{aligned} W_1(\mu_A,\mu_B) &\leqslant W_1(\mu,\mu_A) + W_1(\mu,\mu_B) \leqslant \sqrt{2CH(\mu_A \mid \mu)} + \sqrt{2CH(\mu_B \mid \mu)} \\ &= \sqrt{2C\log\frac{1}{\mu(A)}} + \sqrt{2C\log\frac{1}{\mu(B)}}. \end{aligned} \tag{5.1.5}$$

Παρατηρούμε ότι κάθε μέτρο π που έχει περιθώρια μέτρα τα μ_A και μ_B πρέπει να έχει φορέα το $A\times B$, οπότε, από τον ορισμό της W_1,

$$W_1(\mu_A,\mu_B) \geqslant d(A,B) = \inf\{d(x,y) : x\in A, y\in B\}.$$

Αν λοιπόν έχουμε $d(A,B) \geqslant r > 0$ τότε

$$r \leqslant \sqrt{2C\log\frac{1}{\mu(A)}} + \sqrt{2C\log\frac{1}{1-\mu(A_r)}}, \tag{5.1.6}$$

όπου $A_r = \{x\in X : d(x,A) < r\}$. Αν τώρα $\mu(A) \geqslant \frac{1}{2}$ βλέπουμε ότι

$$r \leqslant \sqrt{2C\log 2} + \sqrt{2C\log\frac{1}{1-\mu(A_r)}}, \tag{5.1.7}$$

και έπεται ότι, για κάθε $r \geqslant 2\sqrt{2C\log 2}$ ισχύει η ανισότητα

$$1-\mu(A_r) \leqslant e^{-r^2/8C}.$$

Αποδεικνύεται ότι η υπόθεση (5.1.4) είναι ισοδύναμη με κανονικά φράγματα για το συναρτησοειδές Laplace του (X, d, μ),

$$E_\mu(\lambda) := \sup \int_X e^{\lambda F}\, d\mu,$$

όπου το supremum είναι πάνω από όλες τις 1-Lipschitz συναρτήσεις $F : X \to \mathbb{R}$ που έχουν μέση τιμή $\int_X F\, d\mu = 0$.

Θεώρημα 5.1.4 (Bobkov-Götze). *Έστω $\mu \in \mathcal{P}_1(X)$. Τα ακόλουθα είναι ισοδύναμα:*

(α) *Για κάθε $\nu \in \mathcal{P}_1(X)$,*

$$W_1(\mu, \nu) \leqslant \sqrt{2\sigma^2 H(\nu \,|\, \mu)}.$$

(β) *Το συναρτησοειδές Laplace του μ ικανοποιεί, για κάθε $\lambda \geqslant 0$, την*

$$E_\mu(\lambda) \leqslant \exp(\sigma^2\lambda^2/2).$$

Απόδειξη. Γράφουμε $W_1(\mu, \nu) = \sup\left(\int_X u\, d\mu - \int_X v\, d\nu\right)$, όπου το supremum είναι πάνω από όλες τις φραγμένες συνεχείς συναρτήσεις $u, v : X \to \mathbb{R}$ που ικανοποιούν την $u(x) \leqslant v(y) + d(x, y)$ για κάθε $x, y \in X$. Συμβολίζουμε το σύνολο όλων αυτών των ζυγαριών (u, v) με Φ_d.

Έστω $\nu \in \mathcal{P}_1(X)$ και έστω $f = \frac{d\nu}{d\mu}$. Υποθέτοντας το (α), για κάθε ζυγάρι $(u, v) \in \Phi_d$ έχουμε

$$\int_X v\, d\nu - \int_X u\, d\mu \leqslant \sqrt{2\sigma^2 \mathrm{Ent}_\mu(f)}.$$

Ισοδύναμα, για κάθε $\lambda > 0$,

$$\int_X v\, d\nu - \int_X u\, d\mu \leqslant \left(\sigma^2\lambda\, \frac{2}{\lambda}\mathrm{Ent}_\mu(f)\right)^{1/2} \leqslant \frac{\sigma^2\lambda}{2} + \frac{1}{\lambda}\mathrm{Ent}_\mu(f).$$

Θέτοντας $\psi = \lambda v - \lambda\mathbb{E}_\mu(u) - \frac{\sigma^2\lambda^2}{2}$ και αναδιατάσσοντας, γράφουμε την τελευταία ανισότητα στη μορφή

$$\int_X \psi f\, d\mu \leqslant \mathrm{Ent}_\mu(f).$$

Αυτή ισχύει για κάθε ν, άρα για κάθε f. Επιλέγοντας $f = e^\psi/\mathbb{E}_\mu(e^\psi)$ και διαγράφοντας κάποιους όρους, παίρνουμε $\log(\mathbb{E}_\mu(e^\psi)) \leqslant 0$, άρα $\mathbb{E}_\mu(e^\psi) \leqslant 1$. Επιστρέφοντας στον ορισμό της ψ συμπεραίνουμε ότι, για κάθε $\lambda \geqslant 0$,

$$\int e^{\lambda v}\, d\mu \leqslant e^{\lambda\mathbb{E}_\mu(u) + \frac{\sigma^2\lambda^2}{2}}.$$

Έστω h μια 1-Lipschitz συνάρτηση με $\mathbb{E}_\mu(h) = 0$. Παίρνοντας $u = v = h$, οπότε ικανοποιείται η συνθήκη $u(x) \leqslant v(y) + d(x, y)$ διότι η h είναι 1-Lipschitz, έχουμε ότι

$$\int e^{\lambda h}\, d\mu \leqslant e^{\frac{\sigma^2\lambda^2}{2}}, \qquad \lambda \geqslant 0$$

και αυτό μας δίνει αμέσως το (β) από τον ορισμό του συναρτησοειδούς Laplace.

Για την αντίστροφη συνεπαγωγή χρησιμοποιούμε το γεγονός ότι

$$H(\nu \,|\, \mu) = \mathrm{Ent}_\mu(f) = \sup\left\{\int_X f\psi\, d\mu : \int_X e^\psi\, d\mu \leqslant 1\right\}.$$

Επαναλαμβάνοντας το προηγούμενο επιχείρημα μπορούμε να επιστρέψουμε από το (β) στο (α). $\square$

Έστω μ, ν δύο Borel μέτρα πιθανότητας στον $\mathbb{R}^n$. Ορίζουμε την *τετραγωνική απόσταση Wasserstein* των μ και ν ως

$$W_2(\mu, \nu) = \inf \Big(\int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 \, d\pi(x,y) \Big)^{1/2},$$

όπου το infimum παίρνεται πάνω από όλα τα Borel μέτρα πιθανότητας π στον $\mathbb{R}^n \times \mathbb{R}^n$ με περιθώρια μέτρα $\pi_x = \mu$ και $\pi_y = \nu$.

Στη συνέχεια, για την απόδειξη της ανισότητας θα χρησιμοποιήσουμε έναν διαφορετικό τύπο για την τετραγωνική απόσταση Wasserstein. Ωστόσο, η απόδειξη αυτού του τύπου ξεφεύγει του σκοπού αυτής της εργασίας και γι' αυτό παραλείπεται.

Θεώρημα 5.1.5 (τύπος δυϊσμού). *Αν μ, ν είναι δύο Borel μέτρα πιθανότητας στον $\mathbb{R}^n$, τότε*

$$\frac{1}{2} W_2(\mu, \nu)^2 = \sup \Big(\int_{\mathbb{R}^n} Q_1 \varphi \, d\mu - \int_{\mathbb{R}^n} \varphi \, d\nu \Big),$$

όπου

$$Q_s \varphi(x) = \inf_{y \in \mathbb{R}^n} \Big[\varphi(y) + \frac{1}{2s} |x-y|^2 \Big], \qquad s > 0, \ x \in \mathbb{R}^n$$

και το supremum παίρνεται πάνω από όλες τις συνεχείς και φραγμένες $\varphi : \mathbb{R}^n \to \mathbb{R}$.

Η $(Q_s)_{s \geqslant 0}$ είναι η ημιομάδα ελαχιστικής συνέλιξης Hopf-Lax, για την οποία είναι γνωστό ότι λύνει την εξίσωση Hamilton-Jacobi

$$\frac{d}{ds} Q_s \varphi = -\frac{1}{2} |\nabla Q_s \varphi|^2$$

στο $(0, \infty) \times \mathbb{R}$ με αρχική συνθήκη φ.

5.2 Η αρχική απόδειξη του Talagrand

Σε αυτή την παράγραφο περιγράφουμε την αρχική απόδειξη της ανισότητας από το [95], στην οποία ο Talagrand χρησιμοποιεί μεταφορά του μέτρου στη μία διάσταση, και το πέρασμα στις μεγαλύτερες διαστάσεις γίνεται με ένα επιχείρημα tensorization.

Λήμμα 5.2.1. *Έστω $\mu : \mathcal{B}(\mathbb{R}) \to [0,1]$ Borel μέτρο πιθανότητας με $\mu \ll \gamma$ και $f = \frac{d\mu}{d\gamma} > 0$. Τότε, το μ είναι συνεχές μέτρο πιθανότητας και υπάρχει μοναδική συνάρτηση $\varphi : \mathbb{R} \to \mathbb{R}$ που ικανοποιεί την $\gamma((-\infty, x]) = \mu((-\infty, \varphi(x)])$ για κάθε $x \in \mathbb{R}$.*

Απόδειξη. Αρχικά, για κάθε $x \in \mathbb{R}$ έχουμε ότι $\mu(\{x\}) = \int_{\{x\}} f \, d\gamma = 0$, αφού $\gamma(\{x\}) = 0$. Άρα, το μ είναι συνεχές μέτρο πιθανότητας.

Τώρα, έστω $x \in \mathbb{R}$. Τότε, $\gamma((-\infty, x]) \in (0,1)$. Συνεπώς, αφού το μ είναι συνεχές, υπάρχει $y_x \in \mathbb{R}$ ώστε $\mu((-\infty, y_x]) = \gamma((-\infty, x])$. Το y_x μάλιστα είναι μοναδικό, διότι αν είχαμε $y_x < y'_x$ με $\mu((-\infty, y_x]) = \mu((-\infty, y'_x])$ τότε θα είχαμε

$$\int_{y_x}^{y'_x} f \, d\gamma = \mu((y_x, y'_x]) = 0,$$

το οποίο είναι άτοπο αφού $\gamma((y_x, y'_x]) > 0$ και $f > 0$. Θέτουμε $\varphi(x) = y_x$. $\square$

Λήμμα 5.2.2. *Έστω $\mu : \mathcal{B}(\mathbb{R}) \to [0,1]$ Borel μέτρο πιθανότητας με $\mu \ll \gamma$ και $f = \frac{d\mu}{d\gamma} > 0$. Τότε, η $\varphi : \mathbb{R} \to \mathbb{R}$ με $\gamma((-\infty,x]) = \mu((-\infty,\varphi(x)])$ για κάθε $x \in \mathbb{R}$ είναι $1-1$, επί, γνησίως αύξουσα, συνεχής και λ-σχεδόν παντού παραγωγίσιμη, όπου λ είναι το μέτρο Lebesgue στο $\mathbb{R}$. Επίσης, $\mu = \varphi * \gamma = \gamma \circ \varphi^{-1}$.*

Απόδειξη. Έστω $x_1, x_2 \in \mathbb{R}$ με $x_1 \leqslant x_2$ και $\varphi(x_1) = \varphi(x_2)$. Τότε, $\gamma((-\infty,x_1]) = \gamma((-\infty,x_2])$, δηλαδή $\gamma([x_1,x_2]) = 0$, άρα $x_1 = x_2$. Αυτό αποδεικνύει ότι η φ είναι $1-1$.

Έστω τώρα $y \in \mathbb{R}$. Τότε $\mu((-\infty,y]) \in (0,1)$. Συνεπώς, υπάρχει $x_y \in \mathbb{R}$ με $\gamma((-\infty,x_y]) = \mu((-\infty,y])$. Τότε, $\mu((-\infty,y]) = \mu((-\infty,\varphi(x_y)])$, άρα $\varphi(x_y) = y$, και αυτό δείχνει ότι η φ είναι επί.

Θεωρούμε τώρα τις

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du, \quad G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} f(u) e^{-u^2/2} du.$$

Παρατηρούμε ότι $F(x) = \gamma((-\infty,x])$ και $G(x) = \mu((-\infty,x])$. Οπότε, οι F, G είναι συνεχείς και γνησίως αύξουσες, η F είναι παραγωγίσιμη με $F'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} > 0$ και η G είναι λ-σχεδόν παντού παραγωγίσιμη με $G'(x) = \frac{1}{\sqrt{2\pi}} f(x) e^{-x^2/2} > 0$ λ-σχεδόν παντού.

Έχουμε λοιπόν ότι $F(x) = G(\varphi(x))$ για κάθε $x \in \mathbb{R}$, δηλαδή $\varphi(x) = G^{-1} \circ F(x)$, και άρα η φ είναι συνεχής και γνησίως αύξουσα. Επίσης, $\varphi^{-1}(x) = F^{-1} \circ G(x)$. Οπότε, από το θεώρημα παραγώγισης αντίστροφης συνάρτησης έπεται ότι η φ^{-1} είναι παραγωγίσιμη λ-σχεδόν παντού, με παράγωγο

$$(\varphi^{-1})'(x) = (F^{-1})'(G(x)) G'(x) > 0.$$

Άρα, η φ είναι επίσης λ-σχεδόν παντού παραγωγίσιμη.

Τώρα, για κάθε $x \in \mathbb{R}$ έχουμε ότι

$$\gamma \circ \varphi^{-1}((-\infty,x]) = \gamma((-\infty,\varphi^{-1}(x)]) = \mu((-\infty,x]).$$

Οπότε, πράγματι, $\mu = \gamma \circ \varphi^{-1}$. □

Πρόταση 5.2.3. *Έστω μ Borel μέτρο πιθανότητας στο $\mathbb{R}$ με $\mu \ll \gamma$. Τότε,*

$$W_2(\mu,\gamma) \leqslant 2 \int_{\mathbb{R}} f \log f \, d\gamma.$$

Απόδειξη. Έστω αρχικά ότι $f = \frac{d\mu}{d\gamma} > 0$ και η f είναι απλή. Θεωρούμε την απεικόνιση $g : \mathbb{R} \to \mathbb{R}^2$, $g(x) = (x, \varphi(x))$, όπου φ είναι η συνάρτηση του Λήμματος 5.2.1 για το μ. Θεωρούμε επίσης το $\pi = \gamma \circ g^{-1}$. Τότε,

$$\pi_x((-\infty,x)) = \pi((-\infty,x) \times \mathbb{R}) = \gamma((-\infty,x])$$

για κάθε $x \in \mathbb{R}$, άρα $\pi_x = \gamma$. Επίσης,

$$\pi_y((-\infty,y)) = \pi(\mathbb{R} \times (-\infty,y)) = \gamma \circ \varphi^{-1}((-\infty,y])$$

για κάθε $y \in \mathbb{R}$, άρα $\pi_y = \gamma \circ \varphi^{-1} = \mu$. Συνεπώς,

$$W_2(\gamma,\mu) \leqslant \int_{\mathbb{R}^2} |x-y|^2 \, d\pi(x,y) = \int_{\mathbb{R}} |x-\varphi(x)|^2 \, d\gamma(x)$$

από την Πρόταση 2.5.5. Αρκεί λοιπόν να δείξουμε ότι

$$\int_{\mathbb{R}} |x-\varphi(x)|^2\,d\gamma(x) \leqslant 2\int_{\mathbb{R}} f\log f\,d\gamma = 2\int_{\mathbb{R}} \log f\,d\mu = 2\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x),$$

δηλαδή ότι

$$\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x) \geqslant \int_{\mathbb{R}} \frac{|x-\varphi(x)|^2}{2}\,d\gamma(x).$$

Στο Λήμμα 5.2.2 έχουμε δείξει ότι

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}du = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\varphi(x)} f(u)e^{-u^2/2}du.$$

Παραγωγίζοντας βλέπουμε ότι $e^{-x^2/2} = f(\varphi(x))e^{-\varphi(x)^2/2}\varphi'(x)$ λ-σχεδόν παντού, δηλαδή

$$f(\varphi(x)) = \frac{1}{\varphi'(x)}e^{\frac{\varphi(x)^2-x^2}{2}}$$

λ-σχεδόν παντού. Συνεπώς,

$$\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x) = \int_{\mathbb{R}} \Big(\frac{\varphi(x)^2}{2} - \frac{x^2}{2} - \log\varphi'(x)\Big)\,d\gamma(x).$$

Έστω $\xi(x) = \frac{x^2}{2}$. Έχουμε λοιπόν ότι

$$\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x) = \int_{\mathbb{R}} (\xi(\varphi(x)) - \xi(x) - \log\varphi'(x))\,d\gamma(x). \tag{5.2.1}$$

Παρατηρούμε, ολοκληρώνοντας κατά μέρη, ότι

$$\int_{\mathbb{R}} \xi'(x)(\varphi(x)-x)e^{-\xi(x)}dx = \int_{\mathbb{R}} (\varphi'(x)-1)e^{-\xi(x)}dx,$$

αφού η f είναι απλή. Άρα, η (5.2.1) μας δίνει

$$\begin{aligned}\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x) &= \int_{\mathbb{R}} (\xi(\varphi(x)) - \xi(x) - (\varphi(x)-x)\xi'(x))\,d\gamma(x)\\ &\quad + \int_{\mathbb{R}} (\varphi'(x) - 1 - \log\varphi'(x))\,d\gamma(x).\end{aligned}$$

Όμως,

$$\xi(\varphi(x)) - \xi(x) - (\varphi(x)-x)\xi'(x) = \frac{\varphi(x)^2}{2} - \frac{x^2}{2} - x\cdot(\varphi(x)-x) = \frac{(\varphi(x)-x)^2}{2}$$

και $\varphi'(x) - 1 - \log\varphi'(x) \geqslant 0$ για κάθε $x\in\mathbb{R}$. Οπότε,

$$\int_{\mathbb{R}} \log(f\circ\varphi(x))\,d\gamma(x) \geqslant \int_{\mathbb{R}} \frac{(\varphi(x)-x)^2}{2}\,d\gamma(x),$$

και δείξαμε αυτό που θέλαμε.

Το γενικό αποτέλεσμα έπεται μέσω ενός επιχειρήματος προσέγγισης. □

Θεώρημα 5.2.4 (ανισότητα μεταφοράς με τετραγωνικό κόστος). *Αν μ είναι ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$ με $\mu \ll \gamma_n$, τότε*

$$W_2(\mu,\gamma_n) \leqslant 2\int_{\mathbb{R}^n} f\log f\, d\gamma_n.$$

Απόδειξη. Θα δείξουμε το αποτέλεσμα χρησιμοποιώντας επαγωγή. Για $n=1$ δείξαμε την ανισότητα στην προηγούμενη πρόταση.

Έστω ότι το θεώρημα ισχύει στον $\mathbb{R}^n$ και έστω μ Borel μέτρο πιθανότητας στον $\mathbb{R}^{n+1}$ με $\mu \ll \gamma_{n+1}$ και $f = \frac{d\mu}{d\gamma_{n+1}}$. Για κάθε $x \in \mathbb{R}^n$ θέτουμε

$$g(x) = \int_{\mathbb{R}} f(x,y)\, d\gamma(y).$$

Τότε, $g \geqslant 0$ και

$$\int_{\mathbb{R}^n} g(x)\, d\gamma_n(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}} f(x,y)\, d\gamma(y)\, d\gamma_n(x) = \int_{\mathbb{R}^{n+1}} f(x,y)\, d\gamma_{n+1}(x,y) = 1$$

από το θεώρημα Tonelli. Για κάθε $x \in \mathbb{R}$ θεωρούμε $f_x : \mathbb{R} \to \mathbb{R}$ ώστε

$$f_x(y)g(x) = f(x,y), \quad y \in \mathbb{R}.$$

Παρατηρούμε ότι αν $g(x)=0$ τότε $f(x,y)=0$ για γ_n-σχεδόν κάθε $y \in \mathbb{R}$, και μπορούμε, χωρίς περιορισμό της γενικότητας, να θεωρήσουμε ότι $f(x,y)=0$ για κάθε $y \in \mathbb{R}$. Σε αυτή την περίπτωση θέτουμε $f_x(y)=1$. Αλλιώς, αν $g(x) \neq 0$ έχουμε ότι

$$f_x(y) = \frac{f(x,y)}{g(x)}.$$

Τώρα έχουμε ότι

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}} f\log f\, d\gamma_{n+1} &= \int_{\mathbb{R}^{n+1}} f(x,y)\log(g(x)f_x(y))\, d\gamma_{n+1}(x,y)\\
&= \int_{\mathbb{R}^{n+1}} f(x,y)\log g(x)\, d\gamma_{n+1}(x,y) + \int_{\mathbb{R}^{n+1}} f(x,y)\log f_x(y)\, d\gamma_{n+1}(x,y)\\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}} f(x,y)\log g(x)\, d\gamma(y)\, d\gamma_n(x) + \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}} f_x(y)\log f_x(y)\, d\gamma(y)\Big)g(x)\, d\gamma_n(x)\\
&= \int_{\mathbb{R}^n} g\log g\, d\gamma_n + \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}} f_x\log f_x\, d\gamma\Big)g(x)\, d\gamma_n(x).
\end{aligned}$$

Από την περίπτωση $n=1$ γνωρίζουμε ότι αν μ_x είναι το μέτρο με πυκνότητα f_x στο $\mathbb{R}$ τότε

$$W_2(\delta_x\otimes\mu_x, \delta_x\otimes\gamma) \leqslant 2\int_{\mathbb{R}} f_x\log f_x\, d\gamma.$$

Έπεται ότι, αν η είναι το μέτρο πιθανότητας στον $\mathbb{R}^n$ με πυκνότητα g τότε

$$W_2(\mu, \eta\otimes\gamma) \leqslant 2\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}} f_x\log f_x\, d\gamma\Big)g(x)\, d\gamma_n(x).$$

Παρατηρούμε ότι το κόστος της μεταφοράς εξαρτάται μόνο από την $(n+1)$-οστή συντεταγμένη.

Από την επαγωγική υπόθεση έχουμε επίσης

$$W_2(\eta,\gamma_n) \leqslant 2\int_{\mathbb{R}^n} g\log g\, d\gamma_n,$$

άρα

$$W_2(\eta\otimes\gamma,\gamma_{n+1}) \leqslant 2\int_{\mathbb{R}^n} g\log g\, d\gamma_n,$$

και το κόστος της μεταφοράς εξαρτάται μόνο από τις πρώτες n συντεταγμένες. Συνεπώς,

$$\begin{aligned} W_2(\mu,\gamma_{n+1}) &\leqslant W_2(\mu,\eta\otimes\gamma) + W_2(\eta\otimes\gamma,\gamma_{n+1}) \\ &\leqslant 2\Big(\int_{\mathbb{R}^n} g\log g\, d\gamma_n + \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}} f_x\log f_x\, d\gamma\Big) g(x)\, d\gamma_n(x)\Big) \\ &= 2\int_{\mathbb{R}^{n+1}} f\log f\, d\gamma_{n+1}, \end{aligned}$$

και έχουμε ολοκληρώσει το επαγωγικό βήμα. □

5.3 Απόδειξη μέσω της λογαριθμικής ανισότητας Sobolev

Η δεύτερη απόδειξη που περιγράφουμε βασίζεται στον τύπο δυϊσμού του Θεωρήματος 5.1.5: Αν μ,ν είναι δύο Borel μέτρα πιθανότητας στον $\mathbb{R}^n$, τότε

$$\frac{1}{2}W_2(\mu,\nu)^2 = \sup\Big(\int_{\mathbb{R}^n} Q_1\varphi\, d\mu - \int_{\mathbb{R}^n}\varphi\, d\nu\Big),$$

όπου

$$Q_1\varphi(x) = \inf_{y\in\mathbb{R}^n}\Big[\varphi(y) + \frac{1}{2}|x-y|^2\Big],\ x\in\mathbb{R}^n$$

και το supremum παίρνεται πάνω από όλες τις συνεχείς και φραγμένες $\varphi:\mathbb{R}^n\to\mathbb{R}$.

Λήμμα 5.3.1. *Αν μ,ν είναι Borel μέτρα πιθανότητας στον $\mathbb{R}^n$ με $d\mu = f\,d\nu$ τότε για κάθε $\varphi\in C_b(\mathbb{R}^n)$ ισχύει ότι*

$$\int_{\mathbb{R}^n} Q_1\varphi\, d\mu - \int_{\mathbb{R}^n}\varphi\, d\nu \leqslant \int_{\mathbb{R}^n} f\log f\, d\nu + \log\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\nu - \int_{\mathbb{R}^n}\varphi\, d\nu.$$

Απόδειξη. Έχουμε ότι

$$\begin{aligned} \int_{\mathbb{R}^n} Q_1\varphi\, d\mu - \int_{\mathbb{R}^n}\varphi\, d\nu &= \int_{\mathbb{R}^n} Q_1\varphi\cdot f\, d\nu - \int_{\mathbb{R}^n}\varphi\, d\nu \\ &\leqslant \mathrm{Ent}_\nu(f) + \log\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\nu - \int_{\mathbb{R}^n}\varphi\, d\nu \\ &= \int_{\mathbb{R}^n} f\log f\, d\nu + \log\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\nu - \int_{\mathbb{R}^n}\varphi\, d\nu, \end{aligned}$$

όπου χρησιμοποιήσαμε την Πρόταση 2.1.8 και το γεγονός ότι $\int_{\mathbb{R}^n} f\,d\nu = \mu(\mathbb{R}^n) = 1$. □

Λήμμα 5.3.2. *Αν $\varphi\in C_b(\mathbb{R}^n)$ και γ_n είναι το τυπικό μέτρο Gauss στον $\mathbb{R}^n$, τότε*

$$\log\Big(\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\gamma_n\Big) \leqslant \int_{\mathbb{R}^n}\varphi\, d\gamma_n.$$

Απόδειξη. Αρχικά, αφού $\varphi \in C_b(\mathbb{R}^n)$, υπάρχει $M > 0$ ώστε $|\varphi(y)| \leqslant M$ για κάθε $y \in \mathbb{R}^n$. Οπότε, αν $x \in \mathbb{R}^n$ και $s > 0$, τότε

$$-M \leqslant \varphi(y) + \frac{1}{2s}|x-y|^2$$

για κάθε $y \in \mathbb{R}^n$, και συμπεραίνουμε ότι

$$-M \leqslant Q_s\varphi(x).$$

Προφανώς, επίσης,

$$Q_s\varphi(x) \leqslant \varphi(x) \leqslant M.$$

Άρα, η Q_s είναι φραγμένη για κάθε $s > 0$. Ορίζουμε λοιπόν

$$\Lambda(s) = \int_{\mathbb{R}^n} e^{sQ_s\varphi}\, d\gamma_n > 0, \qquad s > 0.$$

Η Λ είναι παραγωγίσιμη, με

$$\begin{aligned}\Lambda'(s) &= \int_{\mathbb{R}^n} Q_s\varphi \cdot e^{sQ_s\varphi}\, d\gamma_n - \int_{\mathbb{R}^n} s\frac{d}{ds}(Q_s\varphi)\, e^{sQ_s\varphi}\, d\gamma_n \\ &= \int_{\mathbb{R}^n} Q_s\varphi \cdot e^{sQ_s\varphi}\, d\gamma_n - \frac{s}{2}\int_{\mathbb{R}^n} |\nabla Q_s\varphi|^2 e^{sQ_s\varphi}\, d\gamma_n.\end{aligned}$$

Εφαρμόζοντας τη λογαριθμική ανισότητα Sobolev για το τυπικό μέτρο του Gauss στην $f = e^{sQ_s\varphi/2}$ παίρνουμε ότι

$$\mathrm{Ent}_{\gamma_n}(f^2) \leqslant 2\int_{\mathbb{R}^n} |\nabla f|^2\, d\gamma_n,$$

ή, ισοδύναμα,

$$\int_{\mathbb{R}^n} sQ_s\varphi \cdot e^{sQ_s\varphi}\, d\gamma_n - \Lambda(s)\log\Lambda(s) \leqslant \frac{s^2}{2}\int_{\mathbb{R}^n} |\nabla Q_s\varphi|^2 e^{sQ_s\varphi}\, d\gamma_n,$$

και καταλήγουμε στην

$$s\Lambda'(s) - \Lambda(s)\log\Lambda(s) \leqslant 0. \tag{5.3.1}$$

Θεωρώντας τώρα τη συνάρτηση $H : [0,\infty) \to \mathbb{R}$ με

$$H(s) = \frac{\log\Lambda(s)}{s}, \qquad s > 0$$

και

$$H(0) = \int_{\mathbb{R}^n} \varphi\, d\gamma_n,$$

βλέπουμε ότι η H είναι συνεχής και παραγωγίσιμη στο $(0,\infty)$, με $H'(s) \leqslant 0$ λόγω της (5.3.1). Οπότε, η H είναι φθίνουσα στο $[0,\infty)$ και από την $H(1) \leqslant H(0)$ βλέπουμε ότι

$$\log\Big(\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\gamma_n\Big) \leqslant \int_{\mathbb{R}^n} \varphi\, d\gamma_n.$$

$\square$

Θεώρημα 5.3.3 (ανισότητα μεταφοράς με τετραγωνικό κόστος)**.** *Έστω μ Borel μέτρο πιθανότητας στον $\mathbb{R}^n$ με $\mu \ll \gamma_n$ και $d\mu = f\, d\gamma_n$. Τότε,*

$$W_2(\mu,\gamma_n)^2 \leqslant 2\int_{\mathbb{R}^n} f\log f\, d\gamma_n.$$

Απόδειξη. Έστω $\varphi \in C_b(\mathbb{R}^n)$. Τότε, από το Λήμμα 5.3.1 και το Λήμμα 5.3.2 έχουμε ότι

$$\begin{aligned}\int_{\mathbb{R}^n} Q_1\varphi\, d\mu - \int_{\mathbb{R}^n} \varphi\, d\gamma_n &\leqslant \int_{\mathbb{R}^n} f\log f\, d\gamma_n + \log\int_{\mathbb{R}^n} e^{Q_1\varphi}\, d\gamma_n - \int_{\mathbb{R}^n}\varphi\, d\gamma_n \\ &\leqslant \int_{\mathbb{R}^n} f\log f\, d\gamma_n.\end{aligned}$$

Παίρνοντας supremum ως προς τις $\varphi \in C_b(\mathbb{R}^n)$, από τον τύπο δυϊσμού (Θεώρημα 5.1.5) έχουμε ότι

$$\frac{1}{2}W_2(\mu,\gamma_n)^2 \leqslant \int_{\mathbb{R}^n} f\log f\, d\gamma_n,$$

που είναι το ζητούμενο. □

ΚΕΦΑΛΑΙΟ 6

Supremum εμπειρικών διαδικασιών

Έστω $(\Omega, \mathcal{A})$ ένας μετρήσιμος χώρος και $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές με τιμές στον Ω. Έστω $\mathcal{F}$ μια αριθμήσιμη οικογένεια μετρήσιμων συναρτήσεων ορισμένων στο Ω. Θεωρούμε την τυχαία μεταβλητή

$$Z = \sup_{f\in\mathcal{F}} \sum_{i=1}^n f(X_i)$$

και θέτουμε

$$U = \sup_{f\in\mathcal{F}} \|f\|_\infty \quad \text{και} \quad V = \mathbb{E}\Big(\sup_{f\in\mathcal{F}} \sum_{i=1}^n f(X_i)^2\Big).$$

Η ανισότητα του Talagrand ισχυρίζεται ότι υπάρχει σταθερά $K > 0$ ώστε για κάθε $t > 0$ να ισχύει ότι

$$\mathbb{P}\Big(|Z - \mathbb{E}(Z)| \geqslant t\Big) \leqslant K \exp\Big[-\frac{t}{KU}\log\Big(1 + \frac{tU}{V}\Big)\Big].$$

Σε αυτό το κεφάλαιο παρουσιάζουμε την αρχική απόδειξη αυτής της ανισότητας από τον Talagrand, και μια δεύτερη απόδειξη μέσω εντροπίας.

6.1 Η αρχική απόδειξη του Talagrand

Έστω $(\Omega, \mathcal{A}, \mu)$ χώρος πιθανότητας. Για να αποφύγουμε προβλήματα μετρησιμότητας θα υποθέτουμε ότι το Ω είναι πεπερασμένο σύνολο, $\mathcal{A} = \mathcal{P}(\Omega)$ και $\mu(\{\omega\}) > 0$ για κάθε $\omega \in \Omega$. Θεωρούμε επίσης τον χώρο γινόμενο $(\Omega^n, \mathcal{P}(\Omega^n), \mu^n)$, όπου θα γράφουμε και $P = \mu^n = \mu \otimes \cdots \otimes \mu$.

Ορισμός 6.1.1. (i) Για κάθε $1 \leqslant i \leqslant n$ ορίζουμε την $h_i : \Omega^n \times \Omega^n \to \{0, 1\}$ με $h_i(x, y) = 1 - \delta_{x_i, y_i}$ (δηλαδή 0 αν $x_i = y_i$ και 1 αν $x_i \neq y_i$).

(ii) Αν ν είναι ένα μέτρο πιθανότητας στον $(\Omega^n, \mathcal{P}(\Omega^n))$ τότε ορίζουμε

$$f(\nu, x) = \sum_{i=1}^n \Big(\int_{\Omega^n} h_i(x, y)\, d\nu(y)\Big)^2, \qquad x \in \Omega^n.$$

Λήμμα 6.1.2. *Αν ϱ είναι πεπερασμένο μέτρο στον Ω, τότε υπάρχει η $\frac{d\varrho}{d\mu}$.*

Απόδειξη. Θεωρούμε την $f : \Omega \to \mathbb{R}$, $f(\omega) = \frac{\varrho(\{\omega\})}{\mu(\{\omega\})}$. Τότε, αν $A = \{\omega_1, \dots, \omega_k\} \subseteq \Omega$ έχουμε ότι

$$\int_A f\, d\mu = \sum_{i=1}^{k} f(\omega_i) \cdot \mu(\{\omega_i\}) = \sum_{i=1}^{k} \varrho(\{\omega_i\}) = \varrho(A).$$

Άρα, $f = \frac{d\varrho}{d\mu}$. □

Ορισμός 6.1.3. Έστω ν ένα μέτρο πιθανότητας στον Ω^n και $x \in \Omega^n$. Θέτουμε

$$C_i = \{y \in \Omega^n \mid y_i \neq x_i\}, \qquad 1 \leqslant i \leqslant n.$$

(i) Θεωρούμε το μέτρο ϱ_i στον Ω, που είναι η εικόνα του περιορισμού του ν στο C_i μέσω της $y \mapsto y_i$, δηλαδή για κάθε $A \subseteq \Omega$ έχουμε

$$\varrho_i(A) = \nu(\{y \in C_i \mid y_i \in A\}).$$

Θέτουμε $d_i = \frac{d\varrho_i}{d\mu}$ την Radon-Nikodym παράγωγο του ϱ_i ως προς το μ.
(ii) Ορίζουμε

$$\psi(x) = \begin{cases} \tau x^2 & , \text{ αν } \text{ αν } x \leqslant 2 \\ x \log x & , \text{ αν } \text{ αν } x \geqslant 2, \end{cases}$$

όπου $\tau = \frac{\log 2}{2}$.
(iii) Τέλος, θέτουμε

$$m(\nu, x) = \sum_{i=1}^{n} \int \psi(d_i)\, d\mu.$$

Παρατήρηση 6.1.4. (α) Με τον παραπάνω συμβολισμό, για κάθε συνάρτηση $g : \Omega \to \mathbb{R}$ ισχύει ότι

$$\int_{C_i} g(y_i)\, d\nu(y) = \int_\Omega g d_i\, d\mu$$

για κάθε $1 \leqslant i \leqslant n$. Ισοδύναμα ισχύει επίσης ότι

$$\int_{\Omega^n} g(y) h_i(x, y)\, d\nu(y) = \int_\Omega g d_i\, d\mu$$

για κάθε $1 \leqslant i \leqslant n$.

(β) $\varrho_i(\Omega) \leqslant 1$, και άρα $\int_\Omega d_i\, d\mu = \varrho_i(\Omega) \leqslant 1$ για κάθε $1 \leqslant i \leqslant n$.

Λήμμα 6.1.5. *Ισχύει ότι $f(\nu, x) \leqslant K m(\nu, x)$ για κάποια απόλυτη σταθερά $K > 0$.*

Απόδειξη. Παρατηρούμε ότι

$$f(\nu, x) = \sum_{i=1}^{n} \Big(\int_{\Omega^n} h_i(x, y)\, d\nu(y) \Big)^2 = \sum_{i=1}^{n} \Big(\int_\Omega d_i\, d\mu \Big)^2$$

από την προηγούμενη παρατήρηση. Θέτουμε $d_i' = d_i \cdot \mathbf{1}_{\{d_i \leqslant 2\}}$ και $d_i'' = d_i - d_i'$. Τότε,

$$\Big(\int_\Omega d_i\, d\mu\Big)^2 = \Big(\int_\Omega d_i'\, d\mu + \int_\Omega d_i''\, d\mu\Big)^2 \leqslant 2\Big(\int_\Omega d_i'\, d\mu\Big)^2 + 2\Big(\int_\Omega d_i''\, d\mu\Big)^2.$$

Τώρα,

$$\int_\Omega d_i''\, d\mu = \int_\Omega d_i \cdot \mathbf{1}_{\{d_i>2\}}\, d\mu \leqslant \int_\Omega d_i\, d\mu = m_i(\Omega) \leqslant 1.$$

Συνδυάζοντας αυτή την ανισότητα με την ανισότητα Cauchy-Schwarz έχουμε ότι

$$\begin{aligned}
2\Big(\int_\Omega d_i'\, d\mu\Big)^2 + 2\Big(\int_\Omega d_i''\, d\mu\Big)^2 &\leqslant 2\int_\Omega (d_i')^2\, d\mu + 2\int_\Omega d_i''\, d\mu \\
&= \frac{2}{\tau}\int_\Omega \psi(d_i)\, d\mu + 2\int_\Omega d_i \cdot \mathbf{1}_{\{d_i>2\}}\, d\mu \\
&\leqslant \frac{2}{\tau}\int_\Omega \psi(d_i)\, d\mu + 2\int_\Omega d_i \cdot \mathbf{1}_{\{d_i>2\}}\, d\mu \\
&\leqslant \frac{2}{\tau}\int_\Omega \psi(d_i)\, d\mu + \frac{2}{\log 2}\int_\Omega \log d_i \cdot d_i \cdot \mathbf{1}_{\{d_i>2\}}\, d\mu \\
&= \frac{2}{\tau}\int_\Omega \psi(d_i)\, d\mu + \frac{2}{\log 2}\int_\Omega \psi(d_i) \cdot \mathbf{1}_{\{d_i>2\}}\, d\mu \\
&\leqslant \Big(\frac{2}{\tau} + \frac{2}{\log 2}\Big)\int_\Omega \psi(d_i)\, d\mu,
\end{aligned}$$

αφού $\psi \geqslant 0$. Συνεπώς,

$$f(\nu, x) \leqslant \Big(\frac{2}{\tau} + \frac{2}{\log 2}\Big) m(\nu, x).$$

□

Ορισμός 6.1.6. Για κάθε $A \subseteq \Omega^n$ και $x \in \Omega^n$ θέτουμε

$$m(A, x) = \inf\{m(\nu, x) \mid \nu \text{ μέτρο πιθανότητας στον } \Omega^n \text{ με } \nu(A) = 1\}.$$

Ο πρώτος μας στόχος είναι να δείξουμε την ακόλουθη πρόταση.

Πρόταση 6.1.7. *Υπάρχει $L > 0$ με την εξής ιδιότητα: Αν $g : \Omega \times \Omega \to \mathbb{R}$ συνάρτηση, τότε υπάρχει $\kappa : \Omega \times \Omega \to [0, +\infty)$ με τις ακόλουθες ιδιότητες:*

(i) *Για κάθε $\omega \in \Omega$,*

(6.1.1)
$$\int_\Omega \kappa(\omega', \omega)\, d\mu(\omega') \leqslant 1.$$

(ii) *Ισχύει ότι*

(6.1.2)
$$\int_\Omega \exp(I(\omega))\, d\mu(\omega) \leqslant \frac{1}{\int_\Omega g\, d\mu},$$

όπου

$$\begin{aligned}
I(\omega) = \log\Big[\frac{1}{g(\omega)}\Big(1 - \int_\Omega \kappa(\omega', \omega)\, d\mu(\omega')\Big)\Big] &- \int_\Omega \log\Big[\frac{1}{g(\omega')}\kappa(\omega', \omega)\Big]\, d\mu(\omega') \\
&+ \frac{1}{L}\int_\Omega \psi(\kappa(\omega', \omega))\, d\mu(\omega').
\end{aligned}$$

Σημειώνουμε ότι μέρος της διατύπωσης είναι ότι η $I(\omega)$ είναι καλά ορισμένη. Ζητάμε δηλαδή: αν $g(\omega) = 0$ τότε να ισχύει $\int_\Omega \kappa(\omega', \omega)\, d\mu(\omega') = 1$, και αν $g(\omega') = 0$ τότε $\kappa(\omega', \omega) = 0$. Για την Πρόταση 6.1.7 αρκεί να δείξουμε την επόμενη πρόταση.

Πρόταση 6.1.8. *Υπάρχει $L > 0$ με την εξής ιδιότητα: Αν η $g : [0,1] \to [0,1]$ είναι φθίνουσα συνάρτηση, τότε υπάρχει $\kappa : [0,1] \times [0,1] \to [0, +\infty)$ ώστε*

$$\int_0^t \kappa(s,t)\, ds \leqslant 1, \qquad t \in [0,1]$$

και αν

$$I(t) = \log\Big[\frac{1}{g(t)}\Big(1 - \int_0^t \kappa(s,t)\, ds\Big)\Big] + \int_0^t \log\Big[\Big(\frac{1}{g(s)}\Big)\kappa(s,t)\Big]\, ds + \frac{1}{L}\int_0^t \psi(\kappa(s,t))\, ds,$$

τότε

$$\int_0^1 \exp(I(t))\, dt \leqslant \frac{1}{\int_0^1 g(t)\, dt}.$$

Για την απόδειξη της Πρότασης 6.1.8 θα χρειαστούμε μια σειρά από λήμματα.

Λήμμα 6.1.9. *Αν $L > 0$ αρκετά μεγάλο, τότε ισχύει το εξής: Έστω $0 < \beta < \frac{1}{2}$ και $g : [0,1] \to [\beta/2, 2\beta]$ μια φθίνουσα συνάρτηση. Τότε,*

$$\int_0^1 \frac{1}{g(t)} \exp\Big(-\frac{L}{8}\int_0^t \Big(\log\frac{1}{g(s)} - \log\frac{1}{g(t)}\Big)^2 ds\Big)\, dt \leqslant \frac{1}{\int_0^1 g(t)\, dt}.$$

Απόδειξη. Θέτουμε $a = \int_0^1 g(t)\, dt$ και

$$t_0 = \sup\Big\{t \in [0,1] \mid \frac{L}{8}\int_0^t \Big(\log\frac{g(t)}{g(s)}\Big)^2 ds \leqslant 2\Big\}.$$

Αν $1 \geqslant t > t_0$ έχουμε ότι

$$\frac{L}{8}\int_0^t \Big(\log\frac{g(t)}{g(s)}\Big)^2 ds > 2.$$

Οπότε,

$$\frac{1}{g(t)} \exp\Big(-\frac{L}{8}\int_0^t \Big(\log\frac{1}{g(s)} - \log\frac{1}{g(t)}\Big)^2 ds \leqslant \frac{2e^{-2}}{\beta} \leqslant \frac{1}{2\beta} \leqslant \frac{1}{a},$$

αφού

$$a = \int_0^1 g(t)\, dt \leqslant \int_0^1 2\beta\, dt = 2\beta.$$

Συνεπώς,

$$\begin{aligned}(6.1.3)\qquad &\int_0^1 \frac{1}{g(t)} \exp\Big(-\frac{L}{8}\int_0^t \Big(\log\frac{1}{g(s)} - \log\frac{1}{g(t)}\Big)^2 ds\Big)\, dt \\ &\quad\leqslant \int_0^{t_0} \frac{1}{g(t)} \exp\Big(-\frac{L}{8}\int_0^t \Big(\log\frac{1}{g(s)} - \log\frac{1}{g(t)}\Big)^2 ds\Big)\, dt + \frac{1-t_0}{a}.\end{aligned}$$

Ισχυρισμός 1. Αν $c = \frac{1}{t_0}\int_0^{t_0} g(t)\, dt$, τότε $c \geqslant a$.

Απόδειξη. Πράγματι, αφού η g είναι φθίνουσα έχουμε ότι

$$\begin{aligned}\int_0^{t_0} g(t)\,dt - t_0 a &= \int_0^{t_0}(1-t_0)g(t)\,dt - t_0\int_{t_0}^1 g(t)\,dt\\ &\geqslant \int_0^{t_0}(1-t_0)g(t)\,dt - t_0(1-t_0)g(t_0)\\ &= \int_0^{t_0}(1-t_0)(g(t)-g(t_0))\,dt \geqslant 0.\end{aligned}$$

□

Οπότε, από την (6.1.3) και τον Ισχυρισμό 1, αν δείξουμε ότι

$$\int_0^{t_0}\frac{1}{g(t)}\exp\Big(-\frac{L}{8}\int_0^t\Big(\log\frac{1}{g(s)}-\log\frac{1}{g(t)}\Big)^2ds\Big)\,dt \leqslant \frac{t_0}{c} \tag{6.1.4}$$

έχουμε τελειώσει.

Τώρα, αφού για $0 \leqslant t < t_0$ έχουμε ότι

$$\frac{L}{8}\int_0^t\Big(\log\frac{g(t)}{g(s)}\Big)^2 ds \leqslant 2$$

και για $x \leqslant 2$ ισχύει ότι $e^{-x} \leqslant 1-\frac{x}{4}$, αρκεί να δείξουμε ότι

$$\int_0^{t_0}\Big(\frac{1}{g(t)}-\frac{L}{32g(t)}\int_0^t\Big(\log\frac{g(t)}{g(s)}\Big)^2ds\Big)\,dt \leqslant \frac{t_0}{c}.$$

Όμως $\frac{1}{g(t)} \geqslant \frac{1}{2\beta} \geqslant \frac{1}{4c}$, αφού

$$c = \frac{1}{t_0}\int_0^{t_0} g(t)\,dt \geqslant \int_0^{t_0}\frac{\beta}{2t_0}\,dt = \frac{\beta}{2}.$$

Συνεπώς, αρκεί να δείξουμε ότι

$$\int_0^{t_0}\Big(\frac{1}{g(t)}-\frac{L}{128c}\int_0^t\Big(\log\frac{g(t)}{g(s)}\Big)^2ds\Big)\,dt \leqslant \frac{t_0}{c}.$$

Γράφουμε $g(t) = c(1+u(t))$, $t \in [0,1]$. Τότε,

$$\int_0^{t_0} u(t)\,dt = 0.$$

Επομένως έχουμε να δείξουμε ότι

$$\int_0^{t_0}\Big(\frac{1}{1+u(t)}-\frac{L}{128}\int_0^t\Big(\log\Big(\frac{1+u(t)}{1+u(s)}\Big)\Big)^2ds\Big)\,dt \leqslant t_0. \tag{6.1.5}$$

Ισχυρισμός 2. Για κάθε $0 \leqslant s \leqslant t \leqslant t_0$ ισχύει ότι

$$\Big(\log\Big(\frac{1+u(t)}{1+u(s)}\Big)\Big)^2 \geqslant \frac{1}{16}(u(s)-u(t))^2.$$

Απόδειξη. Πράγματι, αν $u(t) = u(s)$ τότε ο ισχυρισμός προφανώς ισχύει ως ισότητα. Έστω ότι $u(t) \neq u(s)$. Έχουμε ότι $u(t) + 1 = \frac{g(t)}{c}$ για κάθε $t \in [0,1]$. Όμως,

$$\frac{1}{4} = \frac{\frac{\beta}{2}}{2\beta} \leqslant \frac{g(t)}{c} \leqslant \frac{2\beta}{\frac{\beta}{2}} = 4$$

για κάθε $t \in [0,1]$. Αν λοιπόν $r(x) = \log x$, $x \in [1/4, 4]$, τότε $r'(x) = \frac{1}{x} \geqslant \frac{1}{4}$ για κάθε $x \in [1/4, 4]$. Συνεπώς, από το θεώρημα μέσης τιμής υπάρχει ξ μεταξύ των $1 + u(t)$ και $1 + u(s)$ ώστε

$$\frac{r(1+u(t)) - r(1+u(s))}{u(t) - u(s)} = r'(\xi) \geqslant \frac{1}{4},$$

δηλαδή

$$\Big(\log\Big(\frac{1+u(t)}{1+u(s)}\Big)\Big)^2 \geqslant \frac{1}{16}(u(s) - u(t))^2.$$

□

Επιπλέον ισχύει ότι $\frac{1}{1+u} = 1 - u + \frac{u^2}{1+u} \leqslant 1 - u + u^2$ για κάθε $u > 0$. Οπότε, από την (6.1.5), τον Ισχυρισμό 2 και αυτή την παρατήρηση, αρκεί να δείξουμε ότι

$$\int_0^{t_0} \Big((1 - u(t) + u^2(t)) - \frac{L}{2048}\int_0^t (u(t) - u(s))^2\, ds\Big)\, dt \leqslant t_0.$$

Ισοδύναμα, αρκεί να δείξουμε ότι

$$\int_0^{t_0} u^2(t)dt - \frac{L}{2048}\int_0^{t_0}\int_0^t (u(t) - u(s))^2 ds\, dt \leqslant 0,$$

αφού $\int_0^{t_0} u(t)\, dt = 0$. Όμως,

$$\int_0^{t_0}\int_0^t (u(t) - u(s))^2\, ds\, dt = \frac{1}{2}\int_0^{t_0}\int_0^{t_0} (u(s) - u(t))^2\, ds\, dt = t_0\int_0^{t_0} u^2(t)dt \geqslant \frac{1}{2}\int_0^{t_0} u^2(t)\, dt.$$

Οπότε τελικά για $L \geqslant 4096$ έχουμε το ζητούμενο. □

Απόδειξη της Πρότασης 6.1.8. Έστω $g : [0,1] \to [0,1]$ φθίνουσα συνάρτηση. Θέτουμε

$$a := \int_0^1 g(t)\, dt$$

και

$$t_1 = \sup\{t \in [0,1] : g(t) \geqslant 2a\}$$

αν $\{t \in [0,1] : g(t) \geqslant 2a\} \neq \emptyset$, και $t_1 = 0$ διαφορετικά. Ορίζουμε επίσης

$$a_1 := \int_0^{t_1} g(s)\, ds.$$

Ισχυρισμός. Ισχύει ότι $t_1 \leqslant \frac{1}{2}$ και $g(s) \geqslant 2a_1$ για κάθε $s \in [0, t_1)$.

Απόδειξη. Πράγματι, αν $t_1 > \frac{1}{2}$ τότε

$$a_1 = \int_0^{t_1} g(s)\,ds \geqslant \int_0^{t_1} 2a\,ds = 2at_1 > a,$$

το οποίο είναι άτοπο. Για την άλλη ανισότητα, έχουμε ότι

$$\frac{g(s)}{a_1} \geqslant \frac{2a}{a_1} \geqslant 2$$

για κάθε $s \in [0, t_1]$. □

Λήμμα 6.1.10. *Αν $L \geqslant 2$ και $a_1 \geqslant 2^{-L/2}a$ τότε η Πρόταση 6.1.8 ισχύει.*

Απόδειξη. Θέτουμε

$$\kappa(s,t) = \begin{cases} 0 & , \text{ αν } t \leqslant t_1 \text{ ή } s \geqslant t_1 \\ \frac{g(s)}{a_1} & , \text{ αν } s < t_1 < t. \end{cases}$$

Από τον ισχυρισμό έχουμε $\kappa(s,t) \geqslant 2$ για $s < t_1 < t$. Επίσης, για $t \leqslant t_1$ έχουμε ότι $I(t) = \log \frac{1}{g(t)} \leqslant \log \frac{1}{a}$. Αν τώρα $t > t_1$ τότε έχουμε ότι

$$I(t) = \frac{1}{a_1} \int_0^{t_1} g(s) \log \frac{1}{g(s)}\,ds + \frac{1}{L} \int_0^{t_1} \psi\Big(\frac{g(s)}{a_1}\Big)\,ds.$$

Από τον προηγούμενο ισχυρισμό, αφού $\psi(x) = x\log x$, για $x \geqslant 2$ έχουμε ότι

$$I(t) = \frac{1}{a_1} \int_0^{t_1} g(s) \log \Big(\frac{1}{g(s)} \Big(\frac{g(s)}{a_1}\Big)^{1/L}\Big)\,ds.$$

Τώρα, αφού $g(s) \geqslant 2a$ για κάθε $s < t_1$, $L \geqslant 2$ και $\frac{a}{a_1} \leqslant 2^{L/2}$, έχουμε ότι

$$\frac{1}{g(s)}\Big(\frac{g(s)}{a_1}\Big)^{1/L} = \frac{1}{g(s)^{1-\frac{1}{L}}}\,\frac{1}{a_1^{\frac{1}{L}}} \leqslant \frac{1}{(2a)^{1-\frac{1}{L}}}\,\frac{1}{a_1^{\frac{1}{L}}} = \frac{1}{a}\Big(\frac{a}{a_1}\Big)^{\frac{1}{L}}\,\frac{1}{2^{1-\frac{1}{L}}} \leqslant \frac{1}{a}\,\frac{\sqrt{2}}{2^{1-\frac{1}{L}}} \leqslant \frac{1}{a}.$$

Οπότε,

$$I(t) \leqslant \frac{1}{a_1} \int_0^{t_1} g(s) \log \frac{1}{a}\,ds = \log \frac{1}{a}.$$

Επομένως, δείξαμε σε κάθε περίπτωση ότι $I(t) \leqslant \log \frac{1}{a}$. Συνεπώς,

$$\int_0^1 \exp(I(t))\,dt \leqslant \frac{1}{a} = \frac{1}{\int_0^1 g(t)\,dt}.$$

Τέλος, λαμβάνοντας υπ' όψιν ότι για $t \leqslant t_1$ έχουμε $\int_0^t \kappa(s,t)\,ds = 0$ και για $t > t_1$ έχουμε

$$\int_0^t \kappa(s,t)\,ds = \int_0^{t_1} \frac{g(s)}{a_1} ds = 1,$$

παίρνουμε το ζητούμενο. □

Για τη συνέχεια της απόδειξης θα υποθέσουμε ότι $a_1 \leqslant 2^{-L/2}a$, διότι αλλιώς από το Λήμμα 6.1.10 έχουμε τελειώσει. Ορίζουμε

$$t_2 = \sup\Big\{t \in [0,1] : \frac{L}{2}\int_{t_1}^{t} \log \frac{g(s)}{g(t)}\, ds \leqslant 1\Big\}$$

και $d \in \mathbb{R}$ ώστε

$$\frac{L}{2}\int_{t_1}^{t_2} \log \frac{g(s)}{d}\, ds = 1,$$

στην περίπτωση που $t_2 > t_1$.

Λήμμα 6.1.11. *Αν* $L \geqslant 16$ *τότε* $d \geqslant \frac{a}{2}$.

Απόδειξη. Αρχικά παρατηρούμε ότι αν $t > t_2$ τότε $g(t) \leqslant d$. Πράγματι, αν υπάρχει $\xi > t_2$ ώστε $g(\xi) > d$ τότε η

$$h(t) = \frac{L}{2}\int_{t_1}^{t} \log \frac{g(s)}{g(t)}\, ds$$

είναι συνεχής στο $[0,\xi]$. Έχουμε ότι $t_2 \in [0,\xi)$ και $g(t_2) > d$. Οπότε

$$h(t_2) = \frac{L}{2}\int_{t_1}^{t_2} \log \frac{g(s)}{g(t_2)}\, ds < \frac{L}{2}\int_{t_1}^{t_2} \log \frac{g(s)}{d}\, ds = 1.$$

Οπότε από συνέχεια, υπάρχει $\zeta > t_2$ ώστε $h(\zeta) < 1$, το οποίο είναι άτοπο από τον ορισμό του t_2. Επομένως,

$$\int_{t_1}^{t_2} g(t)\, dt = \int_0^1 g(t)\, dt - \int_0^{t_1} g(t)\, dt - \int_{t_2}^{1} g(t)\, dt \geqslant a - a_1 - (1-t_2)d. \tag{6.1.6}$$

Ισχυρισμός. Αν $t_1 < t < t_2$ τότε $d \leqslant g(t) < 2a$.

Απόδειξη. Πράγματι, έστω ότι υπάρχει $t \in (t_1, t_2)$ με $g(t) < d$. Τότε,

$$\int_{t_1}^{t} \log\Big(\frac{g(s)}{g(t)}\Big)\, ds > \int_{t_1}^{t} \log\Big(\frac{g(s)}{d}\Big)\, ds \geqslant \int_{t_1}^{t_2} \log\Big(\frac{g(s)}{d}\Big)\, ds = 1,$$

το οποίο είναι άτοπο από τον ορισμό του t_2. Η $g(t) < 2a$ ισχύει για κάθε $t \in (t_1, t_2)$ από τον ορισμό του t_1. $\square$

Οπότε υπάρχουν $u, v \geqslant 0$ με $u + v = 1$ ώστε

$$u(2a) + vd = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} g(t)\, dt.$$

Από την ανισότητα Jensen έπεται ότι

$$\begin{aligned}\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \log \frac{1}{g(s)}\, ds &\leqslant \log\left(\frac{1}{\frac{1}{t_2-t_1}\int_{t_1}^{t_2} g(s)\, ds}\right)\\ &= \log\Big(\frac{1}{u(2a)+vd}\Big) \leqslant u \log\Big(\frac{1}{2a}\Big) + v\log\Big(\frac{1}{d}\Big).\end{aligned}$$

Συνεπώς,

$$\tag{6.1.7}\begin{aligned}\frac{2}{(t_2-t_1)L} &= \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\log\frac{g(s)}{d}\,ds = \log\frac{1}{d}-\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\log\frac{1}{g(s)}\,ds\\ &\geqslant \log\frac{1}{d}-\Big(u\log\frac{1}{2a}+v\log\frac{1}{d}\Big) = u\Big(\log\frac{1}{d}-\log\frac{1}{2a}\Big)\\ &= u\log\frac{2a}{d},\end{aligned}$$

αφού $1-v=u$. Τώρα, λόγω της (6.1.6) έχουμε ότι

$$\begin{aligned}u(2a-d) &= 2ua+vd-d = -d+\frac{1}{t_2-t_1}\int_{t_1}^{t_2} g(t)\,dt \geqslant -d+\frac{a-a_1-(1-t_2)d}{t_2-t_1}\\ &= \frac{1}{t_2-t_1}(a-a_1-(1-t_1)d).\end{aligned}$$

Ας υποθέσουμε τώρα ότι $d<\frac{a}{2}$. Έχουμε όμως ότι $a_1\leqslant 2^{-L/2}a\leqslant \frac{a}{4}$, αφού $L\geqslant 16$. Οπότε,

$$\begin{aligned}u &\geqslant \frac{1}{t_2-t_1}\frac{a-a_1-(1-t_1)d}{2a} \geqslant \frac{1}{t_2-t_1}\frac{a-\frac{a}{4}-(1-t_1)\frac{a}{2}}{2a} \geqslant \frac{1}{t_2-t_1}\frac{a-\frac{a}{4}-\frac{a}{2}}{2a}\\ &= \frac{1}{8(t_2-t_1)}.\end{aligned}$$

Όμως η (6.1.7) μας δίνει ότι

$$\frac{\log 4}{8(t_2-t_1)} \leqslant u\log 4 \leqslant \frac{2}{L(t_2-t_1)},$$

που είναι άτοπο αφού $L\geqslant 16$. □

Ορίζουμε τώρα

$$\tilde{g}(t) = \begin{cases} g(t) & ,\ \text{αν } t<t_2\\ d & ,\ \text{αν } t\geqslant t_2,\end{cases}$$

και βλέπουμε ότι $\tilde{g}\geqslant g$ σχεδόν παντού και $\tilde{g}\geqslant \frac{a}{2}$ με βάση όσα έχουμε δείξει στα προηγούμενα λήμματα. Θα κατασκευάσουμε την κ έτσι ώστε

$$\int_0^1 \exp(I(t))\,dt \leqslant \frac{1}{a}.$$

Έστω $\xi\in(0,1)$, που θα προσδιοριστεί στη συνέχεια. Ορίζουμε

$$\kappa(s,t) = \begin{cases} \frac{\xi g(s)}{a_1} & ,\ \text{αν } s\leqslant t_1\leqslant t\\ \frac{L(1-\xi)}{2}\log\frac{\tilde{g}(s)}{\tilde{g}(t)} & ,\ \text{αν } t_1<s\leqslant t,\end{cases}$$

και $\kappa(s,t)=0$ αλλιώς.

Έστω τώρα $t\in[0,1]$ με $t\geqslant t_1$, διότι αλλιώς $\int_0^t\kappa(s,t)\,ds=0$. Τότε,

$$\begin{aligned}\int_0^t \kappa(s,t)\,ds &= \int_0^{t_1}\frac{\xi g(s)}{a_1}\,ds+\int_{t_1}^t\frac{L(1-\xi)}{2}\log\frac{\tilde{g}(s)}{\tilde{g}(t)}\,ds\\ &= \xi+\frac{L(1-\xi)}{2}\int_{t_1}^t\log\frac{\tilde{g}(s)}{\tilde{g}(t)}\,ds.\end{aligned}$$

Παρατηρούμε ότι, αφού η $\tilde{g}$ είναι φθίνουσα, η συνάρτηση $t \mapsto \int_{t_1}^t \log \frac{\tilde{g}(s)}{\tilde{g}(t)}\, ds$ είναι αύξουσα στο $[t_1, 1]$. Οπότε,

$$\begin{aligned}\int_{t_1}^t \log \frac{\tilde{g}(s)}{\tilde{g}(1)}\, ds \leqslant \int_{t_1}^1 \log \frac{\tilde{g}(s)}{\tilde{g}(1)}\, ds &= \int_{t_1}^{t_2} \log \frac{g(s)}{d}\, ds + \int_{t_2}^1 \log \frac{d}{d}\, ds \\ &= \frac{2}{L} + 0 = \frac{2}{L}.\end{aligned}$$

Άρα, τελικά,

$$\int_0^t \kappa(s,t)\, ds \leqslant \xi + \frac{L(1-\xi)}{2} \cdot \frac{2}{L} = 1.$$

Παρατηρούμε ότι αν $t < t_1$ τότε $I(t) = \log \frac{1}{g(t)}$, και αν $t \geqslant t_1$ τότε

$$I(t) = \xi I_1 + (1-\xi) I_2(t),$$

όπου

$$I_1 = \int_0^{t_1} \Big(\frac{g(s)}{a_1} \log \frac{1}{g(s)} + \frac{1}{L\xi} \psi\Big(\frac{\xi g(s)}{a_1}\Big)\Big)\, ds$$

και

$$I_2(t) = \log \frac{1}{g(t)} - \frac{L}{2} \int_{t_1}^t \log \frac{g(s)}{g(t)} \log \frac{\tilde{g}(s)}{\tilde{g}(t)}\, dt + \frac{1}{L(1-\xi)} \int_{t_1}^t \psi\Big(\frac{L}{2}(1-\xi) \log \frac{\tilde{g}(s)}{\tilde{g}(t)}\Big)\, dt.$$

Οπότε,

$$\begin{aligned}\int_0^1 \exp(I(t))\, dt &= \int_0^{t_1} \frac{1}{g(s)}\, ds + \int_{t_1}^1 \exp(\xi I_1 + (1-\xi) I_2(t))\, dt \\ &\leqslant \frac{t_1}{2a} + \Big(\int_{t_1}^1 \exp(I_1)\, dt\Big)^{\xi} \Big(\int_{t_1}^1 \exp(I_2(t))\, dt\Big)^{1-\xi},\end{aligned} \tag{6.1.8}$$

από τον ορισμό του t_1 και την ανισότητα Hölder.

Λήμμα 6.1.12. *Ισχύει ότι*

$$I_1 \leqslant J + \log \frac{1}{a}, \quad \text{όπου } J = \log \Big[\Big(\frac{1}{2}\Big)^{1-\frac{1}{L}} \Big(\frac{\xi a}{2a_1}\Big)^{\frac{1}{L}}\Big].$$

Απόδειξη. Παρατηρούμε ότι αν εξαρχής υποθέσουμε ότι $\xi \in (0,1)$ ώστε $\xi a \geqslant a_1$, τότε έχουμε ότι

$$\frac{1}{\xi} \psi\Big(\frac{\xi g(s)}{a_1}\Big) = \frac{g(s)}{a_1} \log \frac{\xi g(s)}{a_1},$$

αφού $\xi g(s) \geqslant 2a\xi \geqslant 2a_1$, άρα $\frac{\xi g(s)}{a_1} \geqslant 2$ για κάθε $s < t_1$. Οπότε,

$$I_1 = \int_0^{t_1} \frac{g(s)}{a_1} \log \Big[\frac{1}{g(s)} \Big(\frac{\xi g(s)}{a_1}\Big)^{\frac{1}{L}}\Big]\, ds = \int_0^{t_1} \frac{g(s)}{a_1} \log \Big[\Big(\frac{1}{g(s)}\Big)^{1-\frac{1}{L}} \Big(\frac{\xi}{a_1}\Big)^{\frac{1}{L}}\Big]\, ds. \tag{6.1.9}$$

Όμως, για κάθε $0 \leqslant s < t_1$,

$$\log \Big[\Big(\frac{1}{g(s)}\Big)^{1-\frac{1}{L}} \Big(\frac{\xi}{a_1}\Big)^{\frac{1}{L}}\Big] \leqslant \log \Big[\Big(\frac{1}{2a}\Big)^{1-\frac{1}{L}} \Big(\frac{\xi}{a_1}\Big)^{\frac{1}{L}}\Big] = \log \frac{1}{a} + \log \Big[\Big(\frac{1}{2}\Big)^{1-\frac{1}{L}} \Big(\frac{\xi a}{2a_1}\Big)^{\frac{1}{L}}\Big].$$

Από την (6.1.9) έπεται το ζητούμενο, αφού $t_1 \leqslant 1$. □

Λήμμα 6.1.13. *Αν $t \geqslant t_1$, ισχύει ότι*

$$I_2(t) \leqslant \log \frac{1}{\tilde{g}(t)} - \frac{L}{4} \int_{t_1}^{t} \Big(\log \frac{\tilde{g}(t)}{\tilde{g}(s)} \Big)^2 \, ds.$$

Απόδειξη. Παρατηρούμε αρχικά ότι

$$\log \frac{g(s)}{g(t)} \log \frac{\tilde{g}(s)}{\tilde{g}(t)} \geqslant \Big(\log \frac{\tilde{g}(s)}{\tilde{g}(t)} \Big)^2 \tag{6.1.10}$$

για κάθε $t \geqslant t_1$ και $t_1 \leqslant s \leqslant t$. Πράγματι, αν $s > t_2$ έχουμε ότι $\log \frac{\tilde{g}(s)}{\tilde{g}(t)} = \log \frac{d}{d} = 0$ και άρα έχουμε ισότητα. Αν $s < t_2$ τότε έχουμε ότι $g(s) = \tilde{g}(s)$ και $g(t) \leqslant \tilde{g}(t)$ και έχουμε την ανισότητα.

Παρατηρούμε επίσης ότι $\psi(x) \leqslant x^2$ για κάθε $x \geqslant 0$. Συνεπώς,

$$\psi\Big(\frac{L}{2}(1-\xi) \log \frac{\tilde{g}(s)}{\tilde{g}(t)}\Big) \leqslant \frac{L^2}{4}(1-\xi)^2 \Big(\log \frac{\tilde{g}(s)}{\tilde{g}(t)} \Big)^2. \tag{6.1.11}$$

Από τις (6.1.10), (6.1.11), και αφού $\tilde{g} \geqslant g$ και $1 - \xi \leqslant 1$, έπεται το ζητούμενο. □

Επιστρέφουμε στην απόδειξη της Πρότασης 6.1.8.

Ισχυρισμός 1. Ισχύει ότι

$$\int_{t_1}^{1} \exp\Big(\log \frac{1}{\tilde{g}(t)} - \frac{L}{4} \int_{t_1}^{t} \Big(\log \frac{\tilde{g}(t)}{\tilde{g}(s)} \Big)^2 ds \Big) \, dt \leqslant \frac{(1-t_1)^2}{\int_{t_1}^1 \tilde{g}(t)\, dt}.$$

Απόδειξη. Πράγματι, παρατηρούμε αρχικά ότι $\frac{a}{2} \leqslant \tilde{g}(t) \leqslant 2a$ για κάθε $t \in [t_1, 1]$. Χρησιμοποιώντας το μετασχηματισμό $t \mapsto \frac{t-t_1}{1-t_1}$ που απεικονίζει το $[t_1, 1]$ στο $[0, 1]$ παίρνουμε ότι

$$\begin{aligned} &\int_{t_1}^{1} \exp\Big(\log \frac{1}{\tilde{g}(t)} - \frac{L}{4} \int_{t_1}^{t} \Big(\log \frac{\tilde{g}(t)}{\tilde{g}(s)} \Big)^2 ds \Big) \, dt \\ &\qquad = \int_0^1 (1-t_1) \frac{1}{\tilde{g}((1-t_1)u + t_1)} \exp\Big(- \frac{L(1-t_1)}{4} \int_0^{\frac{t-t_1}{1-t_1}} \frac{\tilde{g}((1-t_1)u + t_1)}{\tilde{g}((1-t_1)v + t_1)} dv \Big) \, du \\ &\qquad \leqslant \frac{1 - t_1}{\int_0^{\frac{t-t_1}{1-t_1}} \tilde{g}((1-t_1)u + t_1)\, du} = \frac{(1-t_1)^2}{\int_{t_1}^1 \tilde{g}(t)\, dt}, \end{aligned}$$

χρησιμοποιώντας την

$$-\frac{L}{4}(1-t_1) \leqslant -\frac{L}{4}\Big(1 - \frac{1}{2}\Big) = -\frac{L}{8}$$

και το Λήμμα 6.1.9. □

Τώρα, αν $\tilde{a} = \int_0^1 \tilde{g}(t)\, dt$ έχουμε ότι

$$\int_{t_1}^{1} \tilde{g}(t)\, dt = \tilde{a} - \int_0^{t_1} \tilde{g}(t)\, dt = \tilde{a} - a_1 = \tilde{a}\Big(1 - \frac{a_1}{\tilde{a}}\Big)$$

λόγω του ορισμού των a_1, g. Αφού $a_1 \leqslant \frac{a}{4} \leqslant \frac{\tilde{a}}{4}$ έχουμε ότι

$$\Big(\int_{t_1}^{1} \tilde{g}(t)\, dt \Big)^{-1} = \frac{1}{\tilde{a}}\Big(1 - \frac{a_1}{\tilde{a}}\Big)^{-1} \leqslant \frac{1}{\tilde{a}}\Big(1 + \frac{2a_1}{\tilde{a}}\Big). \tag{6.1.12}$$

Από την (6.1.8), τα Λήμματα 6.1.12 και 6.1.13, τον ισχυρισμό και την (6.1.12) έχουμε ότι

$$\begin{aligned}\int_0^1 \exp(I(t))\,dt &\leqslant \frac{t_1}{2a} + \Big(\frac{1-t_1}{a}\Big)^{\xi} \exp(\xi J)\cdot\Big[(1-t_1)^2\frac{1}{\tilde a}\Big(1+\frac{2a_1}{\tilde a}\Big)\Big]^{1-\xi} \\ &\leqslant \frac{t_1}{2a} + (1-t_1)\Big(\frac{1}{a}\Big)^{\xi}\exp(\xi J)\cdot\Big[\frac{1}{\tilde a}\Big(1+\frac{2a_1}{\tilde a}\Big)\Big]^{1-\xi}.\end{aligned}$$

Ισχυρισμός 2. Για κατάλληλη επιλογή του ξ ισχύει ότι

$$\Big(1+\frac{2a_1}{\tilde a}\Big)^{1-\xi}\exp(\xi J)\leqslant 1.$$

Απόδειξη. Πράγματι, έχουμε ότι

$$\Big(1+\frac{2a_1}{\tilde a}\Big)^{1-\xi}\exp(\xi J)\leqslant \exp\Big(\frac{2a_1}{\tilde a}+\xi J\Big).$$

Αρκεί λοιπόν να δείξουμε ότι

$$\frac{2a_1}{\tilde a}+\xi J\leqslant 0.$$

Ψάχνουμε ξ της μορφής $\frac{\beta a_1}{\tilde a}$. Τότε,

$$\frac{2a_1}{\tilde a}+\xi J=(2+b)\log\Big(\frac{\beta^{\frac{1}{L}}}{2}\Big)\leqslant 0$$

για $\beta=4$ και L κατάλληλα μεγάλο ώστε να ικανοποιείται η συνθήκη $\xi a\geqslant a_1$. □

Συνδυάζοντας τα παραπάνω παίρνουμε

$$\int_0^1 \exp(I(t))\,dt\leqslant \frac{t_1}{2a}+(1-t_1)\frac{1}{\xi a}\Big(\frac{1}{\tilde a}\Big)^{1-\xi}\leqslant \frac{t_1}{2a}+\frac{1-t}{a}\leqslant\frac{1}{a}.$$

□

Θεώρημα 6.1.14. *Υπάρχει $L>0$ με την ακόλουθη ιδιότητα: Για κάθε $n\in\mathbb{N}$ και για κάθε $A\subseteq\Omega^n$ ισχύει ότι*

$$\int \exp\Big(\frac{1}{L}m(A,x)\Big)\,dP(x)\leqslant\frac{1}{P(A)}. \tag{6.1.13}$$

Απόδειξη. Θα το δείξουμε με επαγωγή.

Για $n=1$: Έστω $A\subseteq\Omega$. Αν $x\in A$, τότε ορίζοντας $\nu=\delta_x$, το μέτρο Dirac στον Ω με βάρος 1 στο x, έχουμε ότι

$$\varrho(B)=\delta_x(\{y\in\Omega: y\neq x, y\in B\})=0$$

για κάθε $B\subseteq\Omega$. Οπότε, $\frac{d\varrho}{d\mu}=0$, άρα $m(\delta_x,x)=0$. Οπότε, αφού $\delta_x(A)=1$ καθώς $x\in A$, έχουμε ότι $m(A,x)=0$.

Έστω τώρα ότι $x\notin A$. Θεωρούμε το μέτρο πιθανότητας ν στον Ω με $\nu(B)=\frac{\mu(A\cap B)}{\mu(A)}$ για κάθε $B\subseteq\Omega$. Επίσης, $\nu(A)=1$. Παρατηρούμε ότι $\nu=\varrho$ και

$$\frac{d\varrho}{d\mu}=\frac{d\nu}{d\mu}=\frac{1}{\mu(B)}\mathbf{1}_A.$$

Πράγματι,

$$\varrho(B) = \nu(\{y \in \Omega : y \neq x, y \in B\}) = \frac{\mu(A \cap B \setminus \{x\})}{\mu(A)} = \frac{\mu(A \cap B)}{\mu(A)} = \nu(B),$$

αφού $x \notin A$, και αυτό για κάθε $B \subseteq \Omega$. Για το άλλο, αν $B \subseteq \Omega$ τότε

$$\int_B \frac{\mathbf{1}_A}{\mu(A)}\, d\mu = \frac{\mu(A \cap B)}{\mu(A)} = \nu(B).$$

Οπότε,

$$m(\nu, x) = \int_\Omega \psi(d)\, d\mu = \int_A \psi\Big(\frac{1}{\mu(A)}\Big)\, d\mu = \frac{\tau}{\mu(A)}$$

αν $\mu(A) \geqslant \frac{1}{2}$, και

$$m(\nu, x) = \int_\Omega \psi(d)\, d\mu = \int_A \psi\Big(\frac{1}{\mu(A)}\Big)\, d\mu = \log\Big(\frac{1}{\mu(A)}\Big)$$

αν $\mu(A) \leqslant \frac{1}{2}$. Όμως,

$$\begin{aligned}\int_\Omega \exp\big(\frac{1}{L} m(A, x)\big)\, d\mu(x) &= \mu(A) + \int_{\Omega \setminus A} \exp\Big(\frac{1}{L} m(A, x)\Big)\, d\mu(x)\\ &\leqslant \mu(A) + \int_{\Omega \setminus A} \exp\big(\frac{1}{L} m(\nu, x)\Big)\, d\mu(x).\end{aligned}$$

(i) Αν $\mu(A) \leqslant \frac{1}{2}$, τότε

$$\int_\Omega \exp\big(\frac{1}{L} m(A, x)\big)\, d\mu(x) \leqslant \mu(A) + \frac{(1 - \mu(A))}{\mu(A)^L} \leqslant \frac{1}{\mu(A)}$$

αν υποθέσουμε ότι $L \geqslant 1$.

(ii) Αν $\mu(A) \geqslant \frac{1}{2}$ τότε

$$\int_\Omega \exp\big(\frac{1}{L} m(A, x)\big)\, d\mu(x) \leqslant \mu(A) + (1 - \mu(A)) \exp\Big(\frac{\tau}{L\mu(A)}\Big) \leqslant \frac{1}{\mu(A)}$$

αν υποθέσουμε ότι $L \geqslant 4$.

Υποθέτουμε τώρα ότι το αποτέλεσμα ισχύει για τον Ω^n και θα το δείξουμε για τον Ω^{n+1}. Συμβολίζουμε με P το μέτρο γινόμενο στον Ω^n και με P' το μέτρο γινόμενο στον Ω^{n+1}. Ταυτίζουμε τον Ω^{n+1} με το $\Omega^n \times \Omega$ και το P' με το $P \otimes \mu$. Για κάθε $\omega \in \Omega$ θέτουμε

$$A_\omega = \{x \in \Omega^n : (x, \omega) \in A\}$$

και $g(\omega) = P(A_\omega)$. Θα εφαρμόσουμε την Πρόταση 6.1.7. Ο αριθμός L σε αυτή την πρόταση θα είναι ο αριθμός L που ζητάμε στο Θεώρημα 6.1.14. Θεωρούμε έναν πυρήνα κ που ικανοποιεί τις (6.1.1) και (6.1.2).

Για κάθε $\omega' \in \Omega$ και για κάθε $x \in \Omega^n$ θεωρούμε ένα μέτρο $\nu_{x,\omega'}$ στον Ω^n τέτοιον ώστε $\nu_{x,\omega'}(A_{\omega'}) = 1$.

Θεωρούμε το μέτρο $\eta_{x,\omega}$ που ορίζεται από την

$$\eta_{x,\omega} = \Big(1 - \int_\Omega \kappa(\omega', \omega)\, d\mu(\omega')\Big)\nu_{x,\omega} \otimes \delta_\omega + \int \kappa(\omega', \omega)\nu_{x,\omega'} \otimes \delta_{\omega'}\, d\mu(\omega').$$

Το $\eta_{x,\omega}$ είναι μέτρο πιθανότητας με φορέα το A. Από την κυρτότητα της ψ έπεται ότι

$$\begin{aligned} m(\eta_{x,\omega}(x,\omega)) \leqslant &\int \psi(\kappa(\omega', \omega))\, d\mu(\omega') + \Big(1 - \int \kappa(\omega', \omega)\, d\mu(\omega')\Big) m(\nu_{x,\omega}, x) \\ &+ \int \kappa(\omega', \omega) m(\nu_{x,\omega'}, x)\, d\mu(\omega'). \end{aligned}$$

Παίρνοντας infimum πάνω από όλες τις δυνατές επιλογές του $\nu_{x,\omega'}$, βλέπουμε ότι

$$\begin{aligned} m(A, (x,\omega)) \leqslant &\int \psi(\kappa(\omega', \omega))\, d\mu(\omega') + \Big(1 - \int \kappa(\omega', \omega)\, d\mu(\omega')\Big) m(A_{\omega'}, x) \\ &+ \int \kappa(\omega', \omega) m(A_{\omega'}, x)\, d\mu(\omega'). \end{aligned}$$

Τώρα, χρησιμοποιώντας την ανισότητα Hölder, παίρνουμε

$$\int \exp\Big(\frac{1}{L} m(A, (x,\omega))\Big)\, dP(x) \leqslant U \exp\Big(\frac{1}{L}\int \psi(\kappa(\omega', \omega))\, d\mu(\omega')\Big),$$

όπου

$$\begin{aligned} U = &\Big(\int \exp\Big(\frac{1}{L} m(A_\omega, x)\Big)\, dP(x)\Big)^{1 - \int \kappa(\omega', \omega)\, d\mu(\omega')} \\ &\times \prod_{\omega'} \Big(\int \exp\Big(\frac{1}{L} m(A_{\omega'}, x)\Big)\, dP(x)\Big)^{\mu(\{\omega'\})\kappa(\omega', \omega)}. \end{aligned}$$

Από την επαγωγική υπόθεση,

$$U \leqslant \Big(\frac{1}{g(\omega)}\Big)^{1 - \int \kappa(\omega', \omega)\, d\mu(\omega')} \prod_{\omega'} \Big(\frac{1}{g(\omega')}\Big)^{\mu(\{\omega'\})h(\omega', \omega)},$$

άρα

$$\int \exp\Big(\frac{1}{L} m(A, (x,\omega))\Big)\, dP(x) \leqslant \exp(I(\omega)).$$

Ολοκληρώνοντας ως προς ω, χρησιμοποιώντας το θεώρημα Fubibi, και παρατηρώντας ότι

$$\int g(\omega)\, d\mu(\omega) = P'(A),$$

ολοκληρώνουμε την απόδειξη. □

Μπορούμε τώρα να αποδείξουμε την ανισότητα του Talagrand.

Θεώρημα 6.1.15. *Υπάρχει σταθερά $K > 0$ με την ακόλουθη ιδιότητα. Θεωρούμε n ανεξάρτητες τυχαίες μεταβλητές X_i με τιμές σε κάποιον μετρήσιμο χώρο Ω. Θεωρούμε επίσης μια αριθμήσιμη οικογένεια $\mathcal{F}$ μετρήσιμων συναρτήσεων στον Ω, και την τυχαία μεταβλητή*

$$Z = \sup_{f \in \mathcal{F}} \sum_{i=1}^n f(X_i).$$

Θέτουμε

$$U = \sup_{f\in\mathcal{F}} \|f\|_\infty \quad \textit{και} \quad V = \mathbb{E}\sup_{f\in\mathcal{F}} \sum_{i=1}^n f^2(X_i).$$

Τότε, για κάθε $t>0$*, έχουμε ότι*

$$\mathbb{P}(|Z-\mathbb{E}(Z)|\geqslant t) \leqslant K\exp\Big(-\frac{1}{K}\frac{t}{U}\log\Big(\frac{tU}{V}\Big)\Big). \tag{6.1.14}$$

Απόδειξη. Για την απόδειξη, με ένα επιχείρημα προσέγγισης αναγόμαστε στην περίπτωση που η κλάση $\mathcal{F}$ είναι πεπερασμένη. Μπορούμε επίσης να αναχθούμε μετά στην περίπτωση που το Ω είναι πεπερασμένο. Συνεπώς, μπορούμε να χρησιμοποιήσουμε το Θεώρημα 6.1.14.

Στον Ω^n θεωρούμε τη συνάρτηση

$$Z(x) = \sup_{f\in\mathcal{F}} \sum_{i=1}^n f(x_i).$$

Θεωρούμε έναν αριθμό a και το σύνολο $A=\{y : Z(y)\leqslant a\}$. Αν για κάποιο $x\in\Omega^n$ και κάποια $f\in\mathcal{F}$ έχουμε $Z(x)=\sum_{i=1}^n f(x_i)$, τότε

$$Z(x) - a \leqslant \int \sum_{i=1}^n (f(x_i)-f(y))\,d\nu(y) = \sum_{i=1}^n \int (f(x_i)-f(y_i))d_i(y_i)\,d\mu(y_i). \tag{6.1.15}$$

Λήμμα 6.1.16. *Αν* $u,v\geqslant 0$ *και* $u\leqslant 1$ *τότε*

$$uv \leqslant u^2 + 2\psi(v). \tag{6.1.16}$$

Απόδειξη. Έχουμε $uv-u^2\leqslant v^2/4$, άρα η ανισότητα ισχύει αν $v\leqslant 2$, γιατί τότε $\psi(v)=\tau v^2$, και $\tau\geqslant 1/8$. Όμως, αν $v\geqslant 2$ τότε έχουμε $2\psi(v)\geqslant 2v\log 2\geqslant v$. Συνεπώς, αρκεί να δείξουμε ότι $u^2+v\geqslant uv$, ή ισοδύναμα, $v(1-u)\geqslant -u^2$ που ισχύει αφού $u\leqslant 1$. □

Θεωρούμε έναν αριθμό $\delta\geqslant 2U$ τον οποίο θα επιλέξουμε αργότερα. Χρησιμοποιώντας την (6.1.16) με $u=\frac{1}{\delta}|f(x_i)-f(\omega)|$, παίρνουμε

$$\begin{aligned}\int_\Omega (f(x_i)-f(\omega))\,d_i(\omega)\,d\mu(\omega) &\leqslant \delta\int \frac{|f(x_i)-f(\omega)|}{\delta}\,d_i(\omega)\,d\mu(\omega)\\ &\leqslant \frac{1}{\delta}\int_\Omega (f(x_i)-f(\omega))^2\,d\mu(\omega) + 2\delta\int\psi(d_i)\,d\mu.\end{aligned} \tag{6.1.17}$$

Παρατηρούμε ότι

$$n\int f^2(\omega)\,d\mu(\omega) = \mathbb{E}\Big[\sum_{i=1}^n f^2(X_i)\Big] \leqslant V,$$

οπότε

$$Z(x) \leqslant a + \frac{2}{\delta}\Big(V + \sum_{i=1}^n f^2(x_i)\Big) + 2\delta m(\nu,x),$$

και παίρνοντας infimum, χρησιμοποιώντας τις (6.1.15), (6.1.17) και την $(x-y)^2\leqslant 2x^2+2y^2$, έχουμε

$$Z(x) \leqslant a + \frac{2}{\delta}\Big(V + \sum_{i=1}^n f^2(x_i)\Big) + 2\delta m(A,x),$$

Έτσι, χρησιμοποιώντας την (6.1.13) βλέπουμε ότι, για κάθε $u, w > 0$,

$$\begin{aligned} P(Z(u) \geqslant a + \frac{2}{\delta}(V+w) + 2\delta u) \\ \leqslant \frac{1}{P(A)} \exp\Big(-\frac{u}{K}\Big) + P\Big(\sup_{f\in\mathcal{F}} \sum_{i=1}^{n} f^2(x_i) \geqslant w\Big). \end{aligned} \tag{6.1.18}$$

Λήμμα 6.1.17. *Ισχύει ότι*

$$P\Big(\sup_{f\in\mathcal{F}} \sum_{i=1}^{n} f^2(x_i) \geqslant 4V + kU^2\Big) \leqslant 4\cdot 2^{-k}. \tag{6.1.19}$$

Απόδειξη. Εφαρμόζουμε την ανισότητα ελέγχου με 2 σημεία. □

Στην (6.1.18) παίρνουμε $w = 4V + kU^2$ και χρησιμοποιώντας την (6.1.19) παίρνουμε, για k της τάξης του u,

$$P(Z \geqslant a + 2\delta^{-1}(5V + uU^2) + 2\delta u) \leqslant \frac{5}{P(A)} \exp\big(-\frac{u}{K}\big).$$

Τώρα επιλέγουμε $\delta = \min\{2U, \sqrt{V/U}\}$ και έχουμε

$$P(Z \geqslant a + K\max\{uU, \sqrt{uV}\}) \leqslant \frac{5}{P(A)} \exp\big(-\frac{u}{K}\big),$$

οπότε, για $v > 0$,

$$P(Z \geqslant a + v) \leqslant \frac{5}{P(A)} \exp\Big(-\frac{1}{K}\min\Big\{\frac{v^2}{V}, \frac{v}{U}\Big\}\Big).$$

Επιλέγοντας $a = M - v$, όπου M μια διάμεσος της Z, παίρνουμε

$$\frac{1}{2} \leqslant \frac{5}{P(Z \leqslant M - v)} \exp\Big(-\frac{1}{K}\min\Big\{\frac{v^2}{V}, \frac{v}{U}\Big\}\Big).$$

Συνδυάζοντας τις δύο εκτιμήσεις έχουμε ότι

$$P(|Z - M| \geqslant v) \leqslant K\exp\Big(-\frac{1}{K}\min\Big\{\frac{v^2}{V}, \frac{v}{U}\Big\}\Big). \tag{6.1.20}$$

Είναι τώρα απλό να δούμε ότι

$$\mathbb{E}(|Z - M|) \leqslant K(U + \sqrt{V}),$$

και εισάγοντας αυτήν την ανισότητα στην (6.1.20), μετά από απλές πράξεις παίρνουμε

$$P(|Z - \mathbb{E}(Z)| \geqslant v) \leqslant K\exp\Big(-\frac{1}{K}\min\Big\{\frac{v^2}{V}, \frac{v}{U}\Big\}\Big). \tag{6.1.21}$$

Η ανισότητα αυτή είναι ασθενέστερη από αυτήν που ζητάμε, αν και μας δίνει την (6.1.14) όταν $t \leqslant 2V/U$. Όταν όμως $v \geqslant V/U$, θέλουμε έναν εκθέτη που να συμπεριφέρεται ως $\frac{v}{V}\log\frac{vU}{V}$ και όχι ως $\frac{v}{V}$. Για το σκοπό αυτό, θεωρούμε μια παράμετρο $\varrho < U$, η οποία θα προσδιοριστεί στη συνέχεια, και για $x \in \Omega^n$ ορίζουμε

$$W(x) = \sup_{f\in\mathcal{F}} \sum_{i=1}^{n} |f(x_i)|_{\{|f(x_i)|\geqslant\varrho\}}.$$

Τότε,

$$W(x) \leqslant \frac{1}{\varrho} \sup_{f\in\mathcal{F}} \sum_{i=1}^{n} f^2(x_i),$$

άρα

$$\mathbb{E}(W) \leqslant \frac{V}{\varrho}.$$

Τώρα χρησιμοποιούμε την ανισότητα ελέγχου με q σημεία, και για $t \geqslant K\,\mathbb{E}(W)$ έχουμε

$$P(W \geqslant t) \leqslant K \exp\Big(-\frac{t}{KU} \log \frac{t}{K\mathbb{E}(W)}\Big) \leqslant K \exp\Big(-\frac{t}{KU} \log \frac{t\varrho}{KV}\Big). \tag{6.1.22}$$

Βλέπουμε έτσι ότι αν επιλέξουμε $\varrho = \sqrt{UV/t}$ θα μπορέσουμε να διατηρήσουμε τον λογαριθμικό όρο. Θεωρούμε λοιπόν την κλάση

$$\mathcal{F}_\varrho = \{f \cdot \mathbf{1}_{\{|f|\leqslant\varrho\}}\}$$

στην οποία θα εφαρμόσουμε την (6.1.22). Αν ορίσουμε την Z_ϱ κατά τον προφανή τρόπο, έχουμε ότι

$$|Z - Z_\varrho| \leqslant W.$$

Ειδικότερα έπεται ότι

$$|\mathbb{E}(Z) - \mathbb{E}(Z_\varrho)| \leqslant \mathbb{E}(W) \leqslant \frac{V}{\varrho} = \sqrt{\frac{tV}{U}}. \tag{6.1.23}$$

Επίσης έχουμε ότι

$$P(Z \geqslant \mathbb{E}(Z) + t) \leqslant P\Big(W \geqslant \frac{t}{2}\Big) + P\Big(Z_\varrho \geqslant \mathbb{E}(Z) + \frac{t}{2}\Big).$$

Χρησιμοποιούμε την (6.1.23) για τον πρώτο όρο. Για τον δεύτερο όρο παρατηρούμε ότι

$$P\Big(Z_\varrho \geqslant \mathbb{E}(Z) + \frac{t}{2}\Big) \leqslant P\Big(Z_\varrho \geqslant \mathbb{E}(Z_\varrho) + \frac{t}{4}\Big),$$

με την προϋπόθεση ότι $t \geqslant 4\sqrt{tV/U}$, δηλαδή $t \geqslant 16V/U$. Για να φράξουμε αυτόν τον τελευταίο όρο χρησιμοποιούμε την (6.1.22) και έχουμε

$$\begin{aligned} P\Big(Z_\varrho) \geqslant \mathbb{E}(Z_\varrho) + \frac{t}{4}\Big) &\leqslant K \exp\Big(-\frac{1}{K} \min\Big\{\frac{t^2}{V}, \frac{t}{\varrho}\Big\}\Big) \\ &\leqslant K \exp\Big(-\frac{1}{K} \min\Big\{\frac{t^2}{V}, \frac{t^{3/2}}{\sqrt{UV}}\Big\}\Big) = K \exp\Big(-\frac{1}{K} \frac{t^{3/2}}{\sqrt{UV}}\Big). \end{aligned}$$

Από τα παραπάνω έχουμε το θεώρημα. □

6.2 Απόδειξη μέσω εντροπίας

Δίνουμε τώρα μια δεύτερη απόδειξη της ανισότητα μέσω εντροπίας. Εισάγουμε αρχικά το πλαίσιο στο οποίο θα δουλέψουμε παρακάτω. Έστω $(\Omega, \mathcal{A}, P)$ χώρος πιθανότητας και $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές με τιμές σε κάποιον μετρήσιμο χώρο $(S, \mathcal{B})$.

Θεωρούμε μια πεπερασμένη οικογένεια $\mathcal{F} = \{g_1, \ldots, g_k\}$ μετρήσιμων πραγματικών συναρτήσεων $g_j : (S, \mathcal{B}) \to \mathbb{R}$, $1 \leqslant j \leqslant k$, με την ιδιότητα $|g_j| \leqslant 1$ για κάθε $1 \leqslant j \leqslant k$. Υποθέτουμε επιπλέον ότι $\max\limits_{1\leqslant j\leqslant k} \|g_j\|_\infty = 1$. Έστω

$$Z = \sup_{g\in\mathcal{F}} \sum_{i=1}^{n} g(X_i),$$

όπου υποθέτουμε ότι $P(Z \neq 0) > 0$. Τότε $|Z| \leqslant n$.

Για να χρησιμοποιήσουμε τα εργαλεία του Κεφαλαίου 2 μπορούμε ισοδύναμα να θεωρήσουμε τις κατανομές πιθανότητας $\mu_1, \ldots, \mu_n$ των ανεξάρτητων τυχαίων μεταβλητών $X_1, \ldots, X_n$ και να θέσουμε $P = \mu_1 \otimes \cdots \otimes \mu_n$ στον χώρο γινόμενο S^n. Με αυτόν τον συμβολισμό,

$$Z = Z(x) = Z(x_1, \ldots, x_n) = \max_{1\leqslant j\leqslant k} \sum_{i=1}^{n} g_j(x_i),$$

όπου $x = (x_1, \ldots, x_n) \in S^n$.

Συμβολισμός. Αν X είναι μια πραγματική τυχαία μεταβλητή στον χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$, θα συμβολίζουμε με $\mathbb{E}(X) = \int_\Omega X\,dP$ τη *μέση τιμή* της X, όποτε το $\int_\Omega X\,dP$ υπάρχει στο $(-\infty, +\infty)$.

Λήμμα 6.2.1. *Αν $0 \leqslant g_j \leqslant 1$ για κάθε $1 \leqslant j \leqslant k$ και $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(\lambda) = e^\lambda - 1 - \lambda$, τότε*

$$\lambda\,\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z}) \log \mathbb{E}(e^{\lambda Z}) \leqslant \varphi(-\lambda)\mathbb{E}(Ze^{\lambda Z}).$$

Απόδειξη. Για κάθε $1 \leqslant i \leqslant n$ και $x = (x_1, \ldots, x_n) \in S^n$ θέτουμε

$$Z^i(x) = \max_{1\leqslant j\leqslant k} \sum_{\substack{m=1\\ m\neq i}}^{n} g_j(x_m) \quad \text{και} \quad d_i = e^{\lambda Z^i(x)} > 0.$$

Από τον ορισμό του, το $Z^i(x)$ εξαρτάται από τα $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ και προφανώς ισχύει ότι $0 \leqslant Z_i(x_i) - Z^i(x) \leqslant 1$ για κάθε $1 \leqslant i \leqslant n$.

Παρατηρούμε τώρα ότι η $\varphi(\lambda) = e^\lambda - 1 - \lambda$, $\lambda \in \mathbb{R}$, είναι κυρτή με $\varphi(0) = 0$ και άρα ικανοποιεί την

$$\varphi(-\lambda u) \leqslant u\varphi(-\lambda), \quad \lambda \in \mathbb{R},\ 0 \leqslant u \leqslant 1.$$

Επίσης,

$$\begin{aligned}\varphi(-\lambda(Z_i(x_i) - Z^i(x)))e^{\lambda Z_i(x)} &= (e^{-\lambda(Z_i(x_i)-Z^i(x))} - 1 + \lambda(Z_i(x_i) - Z^i(x))\, e^{\lambda Z_i(x_i)}\\ &= \lambda(Z_i(x_i) - Z^i(x)) - (e^{\lambda Z_i(x_i)} - e^{\lambda Z^i(x)}).\end{aligned}$$

Οπότε,

$$\begin{aligned}\sum_{i=1}^{n} \lambda(Z_i(x_i) - Z^i(x)) - (e^{\lambda Z_i(x_i)} - e^{\lambda Z^i(x)}) &= \sum_{i=1}^{n} \varphi(-\lambda(Z_i(x_i) - Z^i(x)))e^{\lambda Z_i(x_i)}\\ &\leqslant \sum_{i=1}^{n} (Z_i(x_i) - Z^i(x))\varphi(-\lambda)e^{\lambda Z_i(x_i)}\\ &= \sum_{i=1}^{n} (Z(x) - Z^i(x))e^{\lambda Z(x)}\varphi(-\lambda).\end{aligned}$$

Ισχυρισμός. $\sum_{i=1}^{n}(Z(x) - Z^i(x)) \leqslant Z(x)$.
Πράγματι, έστω ότι $Z(x) = \sum_{m=1}^{n} g_\ell(x_m)$ για κάποιο $1 \leqslant \ell \leqslant k$. Τότε,

$$\begin{aligned} Z(x) - Z^i(x) &= \sum_{m=1}^{n} g_\ell(x_m) - \max_{1\leqslant j\leqslant k} \sum_{\substack{m=1\\ m\neq i}}^{n} g_j(x_m) \leqslant \sum_{m=1}^{n} g_\ell(x_m) - \sum_{\substack{m=1\\ m\neq i}}^{n} g_\ell(x_m) \\ &= g_\ell(x_i). \end{aligned}$$

Οπότε πράγματι

$$\sum_{i=1}^{n}(Z(x) - Z^i(x)) \leqslant \sum_{i=1}^{n} g_\ell(x_i) = Z(x).$$

Συνεπώς έχουμε ότι

$$\sum_{i=1}^{n} \lambda(Z_i(x) - Z^i(x)) - (e^{\lambda Z_i(x_i)} - e^{\lambda Z^i(x)}) \leqslant \varphi(-\lambda) Z(x) e^{\lambda Z(x)}.$$

Τώρα, από την Πρόταση 2.1.13 (ii) για την $f = e^{\lambda Z}$ έχουμε ότι

$$\begin{aligned} \mathrm{Ent}_P(f) &\leqslant \sum_{i=1}^{n} \mathbb{E}\Big(\inf_{c_i>0} \mathbb{E}(f_i(\log f_i - \log c_i) - (f_i - c_i))\Big) \\ &\leqslant \sum_{i=1}^{n} \mathbb{E}\Big(\mathbb{E}(f_i(\log f_i - \log d_i) - (f_i - d_i))\Big). \end{aligned}$$

Έπεται ότι

$$\begin{aligned} \mathbb{E}(\lambda Z e^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log \mathbb{E}(e^{\lambda Z}) &\leqslant \sum_{i=1}^{n} \mathbb{E}(e^{\lambda Z^i}(\lambda Z_i - \lambda Z^i) - (e^{\lambda Z_i} - e^{\lambda Z^i})) \\ &\leqslant \mathbb{E}(\varphi(-\lambda) Z e^{\lambda Z}) = \varphi(-\lambda)\mathbb{E}(Z e^{\lambda Z}). \end{aligned}$$

□

Πρόταση 6.2.2. *Αν $g \geqslant 0$ για κάθε $g \in \mathcal{F}$ τότε*

$$\mathbb{E}(e^{\lambda Z}) \leqslant e^{\mathbb{E}(Z)(e^{\lambda}-1)}$$

για κάθε $\lambda \geqslant 0$.

Απόδειξη. Θεωρούμε την $\Lambda(\lambda) = \mathbb{E}(e^{\lambda Z})$, $\lambda \geqslant 0$. Έχουμε $\Lambda(\lambda) > 0$ για κάθε $\lambda \geqslant 0$. Παρατηρούμε ότι, αφού η Z είναι φραγμένη, η Λ είναι παραγωγίσιμη με $\Lambda'(\lambda) = \mathbb{E}(Z e^{\lambda Z})$.

Ορίζουμε $H : (0, \infty) \to \mathbb{R}$ με $H(0) = \mathbb{E}(Z)$ και

$$H(\lambda) = \frac{\log \Lambda(\lambda)}{\lambda}, \qquad \lambda > 0.$$

Τότε, η H είναι συνεχής και παραγωγίσιμη στο $(0, \infty)$ με

$$\begin{aligned} H'(\lambda) &= \frac{\Lambda'(\lambda)\lambda - \Lambda(\lambda)\log\Lambda(\lambda)}{\lambda^2\Lambda(\lambda)} = \frac{\lambda\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z})}{\lambda^2\Lambda(\lambda)} \\ &\leqslant \frac{\varphi(-\lambda)\mathbb{E}(Ze^{\lambda Z})}{\lambda^2\Lambda(\lambda)} = \frac{\varphi(-\lambda)}{\lambda^2}\,\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}, \end{aligned}$$

λόγω του Λήμματος 6.2.1. Παρατηρούμε επίσης ότι

$$H'(\lambda) = -\frac{1}{\lambda}H(\lambda) + \frac{1}{\lambda}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)}. \tag{6.2.1}$$

Τώρα, αφού $g \geqslant 0$ για κάθε $g \in \mathcal{F}$, έχουμε ότι $Z \geqslant 0$ και $P(Z > 0) > 0$. Συνεπώς, $\Lambda(\lambda) > 1$ για κάθε $\lambda > 0$, απ' όπου έπεται ότι $H(\lambda) > 0$ για κάθε $\lambda > 0$. Άρα,

$$H'(\lambda) \leqslant \frac{\varphi(-\lambda)}{\lambda^2}\frac{\Lambda'(\lambda)}{\Lambda(\lambda)} = \frac{\varphi(-\lambda)}{\lambda}\Big(H'(\lambda) + \frac{1}{\lambda}H(\lambda)\Big)$$

και έχουμε ότι

$$\frac{H'(\lambda)}{H(\lambda)} \leqslant \frac{\varphi(-\lambda)}{\lambda(\lambda-\varphi(-\lambda))} = \frac{1}{\lambda-\varphi(-\lambda)} - \frac{1}{\lambda}$$

για κάθε $\lambda > 0$. Τώρα, για $\lambda > 0$,

$$\begin{aligned}\log\Big(\frac{H(\lambda)}{H(0)}\Big) &= \int_0^\lambda \frac{H'(x)}{H(x)}\,dx \leqslant \int_0^\lambda \Big(\frac{1}{x-\varphi(-x)} - \frac{1}{x}\Big)\,dx\\ &= \int_0^\lambda \Big(\frac{1}{1-e^{-x}} - \frac{1}{x}\Big)\,dx = \Big[x - \log x + \log(1-e^{-x})\Big]_0^\lambda\\ &= \lambda - \log\lambda - \log(1-e^{-\lambda}),\end{aligned}$$

αφού

$$\lim_{x\to 0^+}(x - \log x + \log(1-e^{-x})) = \lim_{x\to 0^+}\log\frac{e^x-1}{x} = \log 1 = 0.$$

Επομένως,

$$\log\Big(\frac{H(\lambda)}{H(0)}\Big) \leqslant \log\Big(\frac{e^\lambda - 1}{\lambda}\Big),$$

δηλαδή

$$H(\lambda) \leqslant H(0)\,\frac{e^\lambda-1}{\lambda}$$

για κάθε $\lambda > 0$. Ισοδύναμα, $\log\Lambda(\lambda) \leqslant H(0)(e^\lambda - 1)$ για κάθε $\lambda > 0$, δηλαδή

$$\mathbb{E}(e^{\lambda Z}) \leqslant e^{\mathbb{E}(Z)(e^\lambda-1)}$$

για $\lambda > 0$, και προφανώς η ίδια ανισότητα ισχύει και για $\lambda = 0$. □

Πόρισμα 6.2.3. *Αν $g \geqslant 0$ για κάθε $g \in \mathcal{F}$, τότε για κάθε $r \geqslant 0$ ισχύει ότι*

$$P(Z \geqslant \mathbb{E}(Z) + r) \leqslant \exp\Big(-\mathbb{E}(Z)h\Big(\frac{r}{\mathbb{E}(Z)}\Big)\Big),$$

όπου $h(x) = (1+x)\log(1+x) - x$.

Απόδειξη. Από την ανισότητα Markov και την Πρόταση 6.2.2 έχουμε

$$P(Z \geqslant \mathbb{E}(Z) + r) \leqslant \frac{\mathbb{E}(e^{\lambda Z})}{e^{\lambda(\mathbb{E}(Z)+r)}} \leqslant e^{-\lambda(\mathbb{E}(Z)+r)+\mathbb{E}(Z)(e^\lambda-1)}$$

για κάθε $\lambda \geqslant 0$. Το ζητούμενο έπεται παρατηρώντας ότι η τελευταία ποσότητα ελαχιστοποιείται για $\lambda = \log\Big(1 + \frac{r}{\mathbb{E}(Z)}\Big)$. □

Λήμμα 6.2.4. *Αν ορίσουμε* $W = W(x) = \max_{1\leqslant j\leqslant k} \sum_{i=1}^n g_j(x_i)^2$ *για κάθε* $x \in S^n$ *και* $V = \mathbb{E}(W)$, *τότε*

$$\lambda\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z}) \leqslant 4V\lambda^2\mathbb{E}(e^{\lambda Z}) + 4\lambda^2\mathbb{E}(We^{\lambda Z})$$

για κάθε $\lambda \in \left(-\frac{1}{4}, \frac{1}{4}\right)$.

Απόδειξη. Έστω αρχικά $1 \leqslant i \leqslant n$ και $x_i, y_i \in S$ με $Z_i(x_i) > Z_i(y_i)$, και $\lambda > 0$. Τότε, χρησιμοποιώντας το θεώρημα μέσης τιμής για την $e^{\lambda y}$ βρίσκουμε $\xi \in (Z_i(y_i), Z_i(x_i))$ ώστε

$$\frac{e^{\lambda Z_i(x_i)} - e^{\lambda Z_i(y_i)}}{Z_i(x_i) - Z_i(y_i)} = \lambda e^{\lambda\xi} \leqslant \lambda e^{\lambda Z_i(x_i)},$$

δηλαδή

$$\lambda(Z_i(x_i) - Z_i(y_i))\,(e^{\lambda Z_i(x_i)} - e^{\lambda Z_i(y_i)}) \leqslant \lambda^2(Z_i(x_i) - Z_i(y_i))^2 e^{\lambda Z_i(x_i)}, \tag{6.2.2}$$

και παρατηρούμε ότι η τελευταία ανισότητα ισχύει και στην περίπτωση όπου $\lambda = 0$ ή $Z_i(x_i) = Z_i(y_i)$.

Για $x, y \in S^n$ ορίζουμε

$$\tilde{W} = \tilde{W}(x,y) = \max_{1\leqslant j\leqslant k} \sum_{i=1}^n (g_j(x_i) - g_j(y_i))^2.$$

Ισχυρισμός. Ισχύει ότι

$$\sum_{i=1}^n (Z_i(x_i) - Z_i(y_i))^2 \mathbf{1}_{\{Z_i(x_i)\geqslant Z_i(y_i)\}} \leqslant \max_{1\leqslant j\leqslant k} \sum_{i=1}^n (g_j(x_i) - g_j(y_i))^2 = \tilde{W}(x,y). \tag{6.2.3}$$

Πράγματι, έχοντας σταθεροποιήσει τα $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ έχουμε ότι

$$\begin{aligned} Z_i(x_i) - Z_i(y_i) = \max_{1\leqslant j\leqslant k} &(g_j(x_1) + \cdots + g_j(x_n)) \\ &- \max_{1\leqslant j\leqslant k} (g_j(x_1) + \cdots + g_j(x_{i-1}) + g_j(y_i) + g_j(x_{i+1}) + \cdots + g_j(x_n)). \end{aligned}$$

Έστω ότι

$$\sum_{m=1}^n g_p(x_m) = \max_{1\leqslant j\leqslant k} \sum_{m=1}^n g_j(x_m)$$

για κάποιο $1 \leqslant p \leqslant k$. Τότε $Z_i(x_i) - Z_i(y_i) \leqslant g_p(x_i) - g_p(y_i)$ και το p δεν εξαρτάται από το i. Οπότε,

$$\begin{aligned} \sum_{i=1}^n (Z_i(x_i) - Z_i(y_i))^2 \mathbf{1}_{\{Z_i(x_i)\geqslant Z_i(y_i)\}} &\leqslant \sum_{i=1}^n (g_p(x_i) - g_p(y_i))^2 \\ &\leqslant \max_{1\leqslant j\leqslant k} \sum_{i=1}^n (g_j(x_i) - g_j(y_i))^2 = \tilde{W}(x,y). \end{aligned}$$

Επομένως, εφαρμόζοντας την Πρόταση 2.1.13 (i) στην $e^{\lambda Z}$ και τις (6.2.2) και (6.2.3), παίρνουμε ότι

$$\lambda\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z}) \leqslant \lambda^2\mathbb{E}(\tilde{W}e^{\lambda Z}) \tag{6.2.4}$$

για κάθε $\lambda \geqslant 0$.

Τώρα, για $\lambda < 0$ και $Z_i(x_i) > Z_i(y_i)$, εφαρμόζοντας το θεώρημα μέσης τιμής για την $e^{\lambda y}$ παίρνουμε ότι υπάρχει $\xi \in (Z_i(y_i), Z_i(x_i))$ ώστε

$$\frac{e^{\lambda Z_i(x_i)} - e^{\lambda Z_i(y_i)}}{Z_i(x_i) - Z_i(y_i)} = \lambda e^{\lambda\xi} \geqslant \lambda e^{\lambda Z_i(y_i)} = \lambda e^{\lambda(Z_i(y_i)-Z_i(x_i))} e^{\lambda Z_i(x_i)} \geqslant \lambda e^{-2\lambda} e^{\lambda Z_i(x_i)}$$

αφού $Z_i(y_i) - Z_i(x_i) \leqslant -1-1 = -2$ και $\lambda < 0$. Οπότε, πολλαπλασιάζοντας με $\lambda(Z_i(x_i)-Z_i(y_i))^2 < 0$ παίρνουμε ότι

$$\lambda(e^{\lambda Z_i(x_i)} - e^{\lambda Z_i(y_i)})\,(Z_i(x_i) - Z_i(y_i)) \leqslant \lambda^2 e^{-2\lambda} e^{\lambda Z_i(x_i)} (Z_i(x_i) - Z_i(y_i))^2.$$

Όπως πριν, καταλήγουμε στην

$$\lambda \mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z}) \log \mathbb{E}(e^{\lambda Z}) \leqslant \lambda^2 e^{-2\lambda} e^{\lambda Z_i(x_i)}. \tag{6.2.5}$$

Συνεπώς, από τις (6.2.4) και (6.2.5) παίρνουμε ότι για $\lambda \in \left(-\frac{1}{4}, \frac{1}{4}\right)$ ισχύει ότι

$$\lambda \mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z}) \log \mathbb{E}(e^{\lambda Z}) \leqslant 2\lambda^2 \mathbb{E}(\tilde{W} e^{\lambda Z}), \tag{6.2.6}$$

αφού $e^{1/2} \leqslant 2$. Τώρα,

$$\begin{aligned} \tilde{W}(x,y) &= \max_{1 \leqslant j \leqslant k} \sum_{i=1}^n (g_j(x_i) - g_j(y_i))^2 \leqslant \max_{1 \leqslant j \leqslant k} \Big(\sum_{i=1}^n (2g_j(x_i)^2 + 2g_j(y_i)^2) \Big) \\ &\leqslant 2W(x) + 2W(y). \end{aligned}$$

Από την ανεξαρτησία λοιπόν των x, y παίρνουμε ότι

$$\mathbb{E}(\tilde{W} e^{\lambda Z}) \leqslant 2\mathbb{E}(We^{\lambda Z}) + 2\mathbb{E}(W)\mathbb{E}(e^{\lambda Z}) = 2V\mathbb{E}(e^{\lambda Z}) + 2\mathbb{E}(We^{\lambda Z}).$$

Η τελευταία ανισότητα σε συνδυασμό με την (6.2.6) μας δίνει το ζητούμενο. □

Πρόταση 6.2.5. *Με τον προηγούμενο συμβολισμό ισχύει ότι*

$$\mathbb{E}(e^{\lambda Z}) \leqslant e^{\lambda \mathbb{E}(Z) + 20\lambda^2 V}$$

για κάθε $0 \leqslant \lambda \leqslant \frac{1}{8}$.

Απόδειξη. Εφαρμόζοντας την Πρόταση 2.1.8 για $f = e^{\lambda Z}/\mathbb{E}(e^{\lambda Z})$ και $g = \lambda W$, παίρνουμε ότι

$$\mathbb{E}(fg) \leqslant \mathrm{Ent}_P(f) + \log \mathbb{E}(e^g)$$

δηλαδή

$$\begin{aligned} \frac{\lambda \mathbb{E}(e^{\lambda Z} W)}{\mathbb{E}(e^{\lambda Z})} &\leqslant \frac{\mathbb{E}((\lambda Z - \log \mathbb{E}(e^{\lambda Z}))e^{\lambda Z})}{\mathbb{E}(e^{\lambda Z})} + \log \mathbb{E}(e^{\lambda W}) \\ &= \frac{\lambda \mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z}) \log \mathbb{E}(e^{\lambda Z})}{\mathbb{E}(e^{\lambda Z})} + \log \mathbb{E}(e^{\lambda W}). \end{aligned}$$

Ισοδύναμα,

$$\lambda\mathbb{E}(e^{\lambda Z}W) \leqslant \lambda\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z}) + \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda W}) \tag{6.2.7}$$

για κάθε $\lambda \in \mathbb{R}$.

Θεωρούμε τώρα τις

$$\Lambda(\lambda) = \mathbb{E}(e^{\lambda Z}) \quad \text{και} \quad R(\lambda) = \mathbb{E}(e^{\lambda W}).$$

Τότε, για κάθε $\lambda \in (0, 1/4)$, χρησιμοποιώντας την (6.2.7) και το Λήμμα 6.2.4 έχουμε ότι

$$\begin{aligned}
&(1-4\lambda)(\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda)) \\
&= (1-4\lambda)(\lambda\mathbb{E}(Ze^{\lambda Z}) - \mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z})) \\
&\leqslant (1-4\lambda)(4V\lambda^2\mathbb{E}(e^{\lambda Z}) + 4\lambda^2\mathbb{E}(We^{\lambda Z})) \\
&\leqslant (1-4\lambda)(4V\lambda^2\mathbb{E}(e^{\lambda Z}) + 4\lambda^2\mathbb{E}(Ze^{\lambda Z}) - 4\lambda\mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda Z}) + 4\lambda\mathbb{E}(e^{\lambda Z})\log\mathbb{E}(e^{\lambda W})) \\
&= (1-4\lambda)(4V\lambda^2\Lambda(\lambda) + 4\Lambda(\lambda)\log R(\lambda) + 4\lambda(\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda))),
\end{aligned}$$

άρα

$$(1-4\lambda)^2(\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda)) \leqslant (1-4\lambda)(4V\lambda^2\Lambda(\lambda) + 4\lambda\Lambda(\lambda)\log R(\lambda))$$

απ' όπου έπεται ότι

$$(1-4\lambda)(\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda)) \leqslant 4V\lambda^2\Lambda(\lambda) + 4\lambda\Lambda(\lambda)\log R(\lambda), \tag{6.2.8}$$

και η τελευταία ανισότητα προφανώς ισχύει και για $\lambda = 0, \frac{1}{4}$.

Εφαρμόζοντας τώρα την Πρόταση 6.2.2 στην W, αφού $0 \leqslant g^2 \leqslant 1$ για κάθε $g \in \mathcal{F}$, έχουμε ότι

$$\log R(\lambda) = \log\mathbb{E}(e^{\lambda W}) \leqslant \mathbb{E}(W)(e^{\lambda} - 1) = V(e^{\lambda} - 1) \tag{6.2.9}$$

για κάθε $\lambda \geqslant 0$. Οπότε η (6.2.8) μέσω της (6.2.9) γίνεται

$$(1-4\lambda)(\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda)) \leqslant 4V(\lambda^2 + \lambda(e^{\lambda} - 1))\Lambda(\lambda). \tag{6.2.10}$$

Θέτοντας τώρα $H(0) = \mathbb{E}(Z)$ και

$$H(\lambda) = \frac{1}{\lambda}\log\Lambda(\lambda), \quad 0 < \lambda < \frac{1}{4}$$

έχουμε ότι η H είναι συνεχής και παραγωγίσιμη στο $(0, 1/4)$ με

$$H'(\lambda) = \frac{\lambda\Lambda'(\lambda) - \Lambda(\lambda)\log\Lambda(\lambda)}{\lambda^2\Lambda(\lambda)} \leqslant \frac{4V}{1-4\lambda}\Big(1 + \frac{e^{\lambda} - 1}{\lambda}\Big)$$

χρησιμοποιώντας την (6.2.10). Τώρα, για $\lambda \in (0, 1/8]$ έχουμε ότι

$$\frac{4V}{1-4\lambda}\Big(1 + \frac{e^{\lambda} - 1}{\lambda}\Big) \leqslant 8V(1 + 8(e^{1/8} - 1)) \leqslant 20V.$$

Οπότε, $H'(\lambda) \leqslant 20V$ για κάθε $\lambda \in (0, 1/8]$, ή ισοδύναμα, $\log\Lambda(\lambda) \leqslant 20\lambda^2 V + \lambda\mathbb{E}(Z)$. Δηλαδή,

$$\mathbb{E}(e^{\lambda Z}) \leqslant e^{\lambda\mathbb{E}(Z) + 20\lambda^2 V},$$

και η ανισότητα προφανώς ισχύει και για $\lambda = 0$. $\square$

Πόρισμα 6.2.6. *Για κάθε $r \geqslant 0$ ισχύει ότι*

$$P(Z \geqslant \mathbb{E}(Z) + r) \leqslant e^{-\min\left\{\frac{r}{16}, \frac{r^2}{80V}\right\}}.$$

Απόδειξη. Αν $r \leqslant 5V$, εφαρμόζοντας την ανισότητα Markov για $\lambda_0 = \frac{r}{40V} \leqslant \frac{1}{8}$ έχουμε ότι

$$P(Z \geqslant \mathbb{E}(Z) + r) \leqslant \frac{\mathbb{E}(e^{\lambda_0 Z})}{e^{\lambda_0(\mathbb{E}(Z)+r)}} \leqslant e^{20\lambda_0^2 V - \lambda_0 r} = e^{-\frac{r^2}{80V}}.$$

Αν $r \geqslant 5V$, εφαρμόζοντας την ανισότητα Markov για $\lambda_0 = \frac{1}{8}$ παίρνουμε ότι

$$P(Z \geqslant \mathbb{E}(Z) + r) \leqslant e^{20\lambda_0^2 V - \lambda_0 r} \leqslant e^{\frac{r}{16} - \frac{r}{8}} = e^{-\frac{r}{16}}.$$

Αφού όμως $\frac{r}{16} \leqslant \frac{r^2}{80V}$ αν και μόνο αν $r \geqslant 5V$, από τις παραπάνω σχέσεις έπεται το ζητούμενο. □

Πρόταση 6.2.7. *Για κάθε $r \geqslant 0$ ισχύει ότι*

$$P(|Z - \mathbb{E}(Z)| \geqslant r) \leqslant 2e^{-\min\left\{\frac{r}{16}, \frac{r^2}{80V}\right\}}.$$

Απόδειξη. Έχουμε ότι

$$\begin{aligned} P(|Z - \mathbb{E}(Z)| \geqslant r) &= P(Z - \mathbb{E}(Z) \geqslant r) + P(Z - \mathbb{E}(Z) \leqslant -r) \\ &= P(Z - \mathbb{E}(Z) \geqslant r) + P((-Z) - \mathbb{E}(-Z) \geqslant r). \end{aligned}$$

Τώρα, εφόσον το Λήμμα 6.2.4 ισχύει και στο $(-1/4, 0)$, η υπόλοιπη επιχειρηματολογία που προηγήθηκε μπορεί να εφαρμοστεί αυτούσια στην $-Z$. Οπότε το ζητούμενο έπεται εφαρμόζοντας το Πόρισμα 6.2.6 για την Z και την $-Z$. □

Λήμμα 6.2.8. *Για $a > 0$ θέτουμε*

$$Z_a^1 = \max_{1 \leqslant j \leqslant k} \sum_{i=1}^n g_j(x_i) \mathbf{1}_{\{|g_j(x_i)| \leqslant a\}} \quad \textit{και} \quad Z_a^2 = \max_{1 \leqslant j \leqslant k} \sum_{i=1}^n |g_j(x_i)| \mathbf{1}_{\{|g_j(x_i)| > a\}}.$$

Τότε,

$$|Z - Z_a^1| \leqslant Z_a^2 \quad \textit{και} \quad |\mathbb{E}(Z) - \mathbb{E}(Z_a^1)| \leqslant \mathbb{E}(Z_a^2).$$

Απόδειξη. Έστω ότι

$$\max_{1 \leqslant j \leqslant k} \sum_{i=1}^n g_j(x_i) = \sum_{i=1}^n g_{m_1}(x_i)$$

και

$$\max_{1 \leqslant j \leqslant k} \sum_{i=1}^n g_j(x_i) \mathbf{1}_{\{|g_j(x_i)| \leqslant a\}} = \sum_{i=1}^n g_{m_2}(x_i) \mathbf{1}_{\{|g_{m_2}(x_i)| \leqslant a\}}.$$

Τότε,

$$\begin{aligned} Z - Z_a^1 &\leqslant \sum_{i=1}^n g_{m_1}(x_i) - \sum_{i=1}^n g_{m_1}(x_i) \mathbf{1}_{\{|g_{m_1}(x_i)| \leqslant a\}} \\ &= \sum_{i=1}^n g_{m_1}(x_i) \mathbf{1}_{\{|g_{m_1}(x_i)| > a\}} \leqslant Z_a^2. \end{aligned} \tag{6.2.11}$$

Επίσης,

$$Z_a^1 - Z \leqslant -\sum_{i=1}^{n} g_{m_2}(x_i)\mathbf{1}_{\{|g_{m_2}(x_i)|>a\}} \leqslant Z_a^2. \tag{6.2.12}$$

Οπότε, πράγματι $|Z - Z_a^1| \leqslant Z_a^2$.

Παίρνοντας μέσες τιμές στις (6.2.11) και (6.2.12) προφανώς έχουμε και την άλλη ανισότητα. □

Λήμμα 6.2.9. *Με το συμβολισμό του Λήμματος* 6.2.8, *για κάθε* $a > 0$ *και* $r \geqslant 0$ *ισχύει ότι*

$$P(|Z_a^1 - \mathbb{E}(Z_a^1)| \geqslant r) \leqslant 2e^{-\min\left\{\frac{r}{16a}, \frac{r^2}{80a^2V}\right\}}$$

και

$$P(Z_a^2 \geqslant \mathbb{E}(Z_a^2) + r) \leqslant e^{-\frac{r}{2}\log\left(1+\frac{a}{\mathbb{E}(Z_a^2)}\right)}.$$

Απόδειξη. Για την πρώτη ανισότητα έχουμε ότι

$$P(|Z_a^1 - \mathbb{E}(Z_a^1)| \geqslant r) = P\Big(\Big|\frac{Z_a^1}{a} - \mathbb{E}\Big(\frac{Z_a^1}{a}\Big)\Big| \geqslant \frac{r}{a}\Big) \leqslant 2\exp\Big(-\min\Big\{\frac{r}{16a}, \frac{r^2}{30a^2V}\Big\}\Big)$$

εφαρμόζοντας την Πρόταση 6.2.7 στην $\frac{Z_a^1}{a}$.

Για την δεύτερη ανισότητα έχουμε ότι

$$P(Z_a^2 \geqslant \mathbb{E}(Z_a^2) + r) \leqslant \exp\Big(-\mathbb{E}(Z_a^2)h\Big(\frac{r}{\mathbb{E}(Z_a^2)}\Big)\Big), \tag{6.2.13}$$

όπου $h(x) = (x+1)\log(x+1) - x$, $x \geqslant 0$, εφαρμόζοντας το Πόρισμα 6.2.3 για την Z_a^2.

Ισχυρισμός. Για κάθε $x \geqslant 0$ ισχύει ότι $h(x) \geqslant \frac{x}{2}\log(x+1)$.

Πράγματι, θεωρούμε την $\kappa(x) = \left(1 + \frac{x}{2}\right)\log(x+1) - x$, $x \geqslant 0$ και έχουμε ότι

$$\kappa'(x) = \frac{\log(x+1)}{2} + \frac{1+\frac{x}{2}}{1+x} - 1 = \frac{\log(x+1)}{2} - \frac{x}{2(x+1)} = \frac{(x+1)\log(x+1) - x}{2(x+1)}.$$

Θεωρώντας την $s(x) = (x+1)\log(x+1) - x$, $x \geqslant 0$ έχουμε ότι $s'(x) = \log(x+1) \geqslant 0$, άρα η s είναι αύξουσα στο $[0,\infty)$. Ειδικότερα, $s(x) \geqslant s(0)$ για κάθε $x \geqslant 0$, άρα $\kappa'(x) \geqslant 0$ για κάθε $x \geqslant 0$. Δηλαδή, η κ είναι αύξουσα στο $[0,\infty)$ και, ειδικότερα, $\kappa(x) \geqslant \kappa(0)$ για κάθε $x \geqslant 0$. Έτσι, έχουμε δείξει τον ισχυρισμό.

Εφαρμόζοντας τον ισχυρισμό στην (6.2.13) για $x = r/\mathbb{E}(Z_a^2) \geqslant 0$ παίρνουμε τη ζητούμενη ανισότητα. □

Λήμμα 6.2.10. *Αν* $4r > 5V$ *τότε*

$$P(|Z - \mathbb{E}(Z)| \geqslant 4r) \leqslant 3e^{-\frac{r}{75}\log\left(1+\frac{4r}{V}\right)}.$$

Απόδειξη. Έχουμε αρχικά ότι

$$Z_a^2 = \max_{1\leqslant j\leqslant k}\sum_{i=1}^{n}|g_j(x_i)|\mathbf{1}_{\{|g_j(x_i)|>a\}} \leqslant \max_{1\leqslant j\leqslant k}\frac{1}{a}\sum_{i=1}^{n}|g_j(x_i)|^2 = \frac{W}{a},$$

άρα
$$\mathbb{E}(Z_a^2) \leqslant \frac{V}{a}.$$
Ορίζουμε τώρα $a = \sqrt{\frac{4V}{5r}} < 1$ αφού $r > 5V/4 > 0$. Έχουμε λοιπόν ότι

$$\mathbb{E}(Z_a^2) \leqslant \frac{V}{a} = \sqrt{\frac{5}{4}Vr} < r. \tag{6.2.14}$$

Συνεπώς,

(6.2.15)
$$\begin{aligned} P(|Z-\mathbb{E}(Z)| \geqslant 4r) &\leqslant P(|Z-Z_a^1| \geqslant 2r) + P(|Z_a^1-\mathbb{E}(Z_a^1)| \geqslant r) + P(|\mathbb{E}(Z_a^1)-\mathbb{E}(Z)| \geqslant r) \\ &\leqslant P(Z_a^2 \geqslant r + \mathbb{E}(Z_a^2)) + P(\mathbb{E}(Z_a^2) \geqslant r) + P(|Z_a^1-\mathbb{E}(Z_a^1)| \geqslant r) \\ &= P(Z_a^2 \geqslant r + \mathbb{E}(Z_a^2)) + P(|Z_a^1-\mathbb{E}(Z_a^1)| \geqslant r), \end{aligned}$$

λόγω της (6.2.14).

Τώρα,

$$\frac{r}{2}\log\Big(1+\frac{r}{\mathbb{E}(Z_a^2)}\Big) \geqslant \frac{r}{2}\log\Big(1+\sqrt{\frac{4r}{5V}}\Big) \geqslant \frac{r}{6}\log\Big(1+\frac{4r}{V}\Big) \geqslant \frac{r}{75}\log\Big(1+\frac{4r}{V}\Big). \tag{6.2.16}$$

Επίσης, με απειροστικό λογισμό προκύπτει ότι

$$\min\Big\{\frac{r}{16a}, \frac{r^2}{80a^2V}\Big\} \geqslant \frac{r}{75}\log\Big(1+\frac{4r}{V}\Big). \tag{6.2.17}$$

Επομένως, εφαρμόζοντας στην (6.2.15) τις ανισότητες του Λήμματος 6.2.9 και τις (6.2.16) και (6.2.17), παίρνουμε ότι
$$P(|Z-\mathbb{E}(Z)| \geqslant 4r) \leqslant 3e^{-\frac{r}{75}\log\left(1+\frac{4r}{V}\right)}.$$

□

Πρόταση 6.2.11. *Για κάθε $r \geqslant 0$ ισχύει ότι*
$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300}\log\Big(1+\frac{r}{V}\Big)\Big).$$

Απόδειξη. Για $r > 5V$, θέτοντας $r_1 = \frac{r}{4}$ έχουμε ότι $4r_1 > 5V$. Οπότε, εφαρμόζοντας το Λήμμα 6.2.10 παίρνουμε ότι
$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300}\log\Big(1+\frac{r}{V}\Big)\Big).$$
Τώρα, για $r \leqslant 5V$, από την Πρόταση 6.2.7 παίρνουμε ότι
$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 2e^{-\frac{r^2}{80V}} \leqslant 2e^{-\frac{r}{80}\log\left(1+\frac{r}{V}\right)},$$
αφού $\log(1+x) \leqslant x$ για κάθε $x \geqslant 0$.

Συνεπώς,
$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300}\log\Big(1+\frac{r}{V}\Big)\Big).$$
Τελικά λοιπόν έχουμε τη ζητούμενη ανισότητα για κάθε $r \geqslant 0$. □

Παρατήρηση 6.2.12. Παρατηρούμε ότι η ανισότητα της Πρότασης 6.2.11 είναι αυτή του Talagrand για το supremum των εμπειρικών ανελίξεων, με την υπόθεση ότι η $\mathcal{F}$ είναι πεπερασμένη οικογένεια και ότι $U = \sup_{f\in\mathcal{F}} \|f\|_\infty = 1$. Θα πάρουμε λοιπόν τώρα τη γενική περίπτωση με αναγωγή σε αυτή την τελευταία πρόταση.

Θεώρημα 6.2.13 (ανισότητα Talagrand για το supremum εμπειρικών διαδικασιών)**.** *Έστω $\mathcal{F}$ μια αριθμήσιμη οικογένεια ομοιόμορφα φραγμένων μετρήσιμων συναρτήσεων $f : (S,\mathcal{B}) \to \mathbb{R}$. Θέτουμε*

$$Z = \sup_{f\in\mathcal{F}} \sum_{i=1}^n f(X_i), \quad U = \sup_{f\in\mathcal{F}} \|f\|_\infty \quad \textit{και} \quad V = \mathbb{E}\Big(\sup_{f\in\mathcal{F}} \sum_{i=1}^n f(x_i)^2\Big).$$

Υποθέτουμε ότι $P(Z\neq 0) > 0$. Τότε,

$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300V}\log\Big(1+\frac{rU}{V}\Big)\Big)$$

για κάθε $r \geqslant 0$.

Απόδειξη. Αφού $P(Z\neq 0) > 0$ έχουμε ότι $V > 0$. Για $r = 0$ το ζητούμενο προφανώς ισχύει. Υποθέτουμε λοιπόν στη συνέχεια ότι $r > 0$.

Περίπτωση 1: Αν η $\mathcal{F}$ είναι πεπερασμένη τότε έχουμε ότι $U = \max_{f\in\mathcal{F}} \|f\|_\infty$. Θεωρούμε το

$$\tilde{\mathcal{F}} = \Big\{\frac{f}{U} : f\in\mathcal{F}\Big\}$$

το οποίο είναι πεπερασμένο, και έχουμε

$$\tilde{U} = \max_{\tilde{f}\in\tilde{\mathcal{F}}} \|\tilde{f}\|_\infty = 1.$$

Τότε, από την Πρόταση 6.2.11 έχουμε ότι

$$P\Big(|\tilde{Z}-\mathbb{E}(\tilde{Z})| \geqslant \frac{r}{U}\Big) \leqslant \exp\big(-\frac{r}{300U}\log\Big(1+\frac{r}{U\tilde{V}}\Big)\Big).$$

Όμως,

$$\tilde{Z} = \sup_{f\in\mathcal{F}} \frac{1}{U}\sum_{i=1}^n f(X_i) = \frac{Z}{U} \quad \text{και} \quad \tilde{V} = \frac{V}{U^2}.$$

Συνεπώς, ισοδύναμα παίρνουμε ότι

$$P(|Z-\mathbb{E}(Z)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300U}\log\Big(1+\frac{rU}{V}\Big)\Big), \qquad r\geqslant 0.$$

Περίπτωση 2: Έστω τώρα ότι η $\mathcal{F}$ είναι άπειρη αριθμήσιμη, και $\mathcal{F} = \{f_n : n\in\mathbb{N}\}$ μια αρίθμηση της. Για κάθε $k\in\mathbb{N}$ θέτουμε

$$Z_k = \max_{1\leqslant j\leqslant k} \sum_{i=1}^n f_j(X_i), \quad U_k = \max_{1\leqslant j\leqslant k} \|f_j\|_\infty \quad \text{και} \quad V_k = \mathbb{E}\Big(\max_{1\leqslant j\leqslant k} \sum_{i=1}^n f_j(X_i)^2\Big).$$

Προφανώς, για κάθε $k\in\mathbb{N}$ έχουμε

$$Z_k \leqslant Z, \quad U_k \leqslant U, \quad \text{και} \quad V_k \leqslant V.$$

Επίσης, $Z_k \nearrow Z$, $U_k \nearrow U$ και $\max_{1\leqslant j\leqslant k}\sum_{i=1}^n f_j(X_i)^2 \nearrow \sup_{f\in\mathcal{F}}\sum_{i=1}^n f_j(X_i)^2$, άρα $V_k \nearrow V$.

Αφού $|Z_k| \leqslant nU$ για κάθε $k \in \mathbb{N}$, έχουμε ότι $\mathbb{E}(Z_k) \nearrow \mathbb{E}(Z)$. Όμως, από την Περίπτωση 1 έχουμε ότι

$$P(|Z_k - \mathbb{E}(Z_k)| \geqslant r) \leqslant 3\exp\Big(-\frac{r}{300V_k}\log\Big(1+\frac{rU_k}{V_k}\Big)\Big), \qquad r \geqslant 0.$$

Έστω $\varepsilon > 0$. Τότε,

$$\begin{aligned}
&P(|Z - \mathbb{E}(Z)| \geqslant r+\varepsilon)\\
&\leqslant P(|Z - Z_k| \geqslant \varepsilon/2) + P(|Z_k - \mathbb{E}(Z_k)| \geqslant r) + P(|\mathbb{E}(Z) - \mathbb{E}(Z_k)| \geqslant \varepsilon/2)\\
&\leqslant P(|Z - Z_k| \geqslant \varepsilon/2) + 3\exp\Big(-\frac{r}{300V_k}\log\Big(1+\frac{rU_k}{V_k}\Big)\Big) + P(|\mathbb{E}(Z) - \mathbb{E}(Z_k)| \geqslant \varepsilon/2).
\end{aligned}$$

Τώρα, παίρνοντας $k \to \infty$ έχουμε ότι

$$P(|Z - \mathbb{E}(Z)| \geqslant r+\varepsilon) \leqslant 3\exp\Big(-\frac{r}{300V}\log\Big(1+\frac{rU}{V}\Big)\Big),$$

αφού $\mathbb{E}(Z_k) \to \mathbb{E}(Z)$ και $Z_k \to Z$ κατά πιθανότητα, $U_k \to U$, $V_k \to V$. Συνεπώς, αφήνοντας το $\varepsilon \to 0^+$ έχουμε ότι

$$P(|Z - \mathbb{E}(Z)| > r) \leqslant 3\exp\Big(-\frac{r}{300V}\log\Big(1+\frac{rU}{V}\Big)\Big)$$

για κάθε $r > 0$. Αν λοιπόν $r > 0$, τότε παίρνοντας μια ακολουθία $0 < r_n \nearrow r$, θέτοντας r_n στην παραπάνω ανισότητα και παίρνοντας $n \to \infty$ έχουμε τελικά το ζητούμενο. $\square$

Βιβλιογραφία

[1] S. Aida, T. Masuda and I. Shigekawa, *Logarithmic Sobolev inequalities and exponential integrability*. J. Funct. Anal. **126**, 83-101 (1994).

[2] N. Alon and V. D. Milman, *Embedding of ℓ_∞^k in finite-dimensional Banach spaces*, Israel J. Math. **45** (1983), 265–280.

[3] N. Alon and J. Spencer, *The probabilistic method*. Fourth edition. Wiley Series in Discrete Mathematics and Optimization. Wiley (2016).

[4] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*. Panoramas et Synthèses vol. 10. Soc. Math. de France (2000).

[5] S. Artstein-Avidan, A. Giannopoulos and V. Milman, *Asymptotic geometric analysis. Part I*. Mathematical Surveys and Monographs 202. American Mathematical Society (2015).

[6] A. Auffinger, M. Damron and J. Hanson. *50 years of first-passage percolation*. University Lecture Series 68. American Mathematical Society (2018).

[7] D. Bakry, *L' hypercontractivité et son utilisation en théorie des semigroupes*. École d' Été de Probabilités de Saint-Flour 1992, Lecture Notes in Math. 1581, 1-114 (1994). Springer.

[8] D. Bakry and M. Émery, *Diffusions hypercontractives*. Séminaire de Probabilités XIX, Lecture Notes in Math. 1123, 177-206 (1985). Springer.

[9] D. Bakry, I. Gentil and M. Ledoux, *Analysis and geometry of Markov diffusion operators*. Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[10] W. Beckner, *Inequalities in Fourier analysis*. Ann. of Math. **102**, 159-182 (1975).

[11] M. Benaïm and R. Rossignol, *Exponential concentration for first passage percolation through modified Poincaré inequalities*. Ann. Inst. Henri Poincaré Probab. Stat. **44**, 544-573 (2008).

[12] I. Benjamini, G. Kalai and O. Schramm, *Noise sensitivity of Boolean functions and applications to percolation*. Inst. Hautes Études Sci. Publ. Math. **90**, 5-43 (1999).

[13] I. Benjamini, G. Kalai and O. Schramm, *First passage percolation has sublinear distance variance*. Ann. Probab. **31**, 1970-1978 (2003).

[14] G. Blower, *The Gaussian isoperimetric inequality and transportation*. Positivity **7**, 203-224 (2003).

[15] S. Bobkov, I. Gentil and M. Ledoux. *Hypercontractivity of Hamilton-Jacobi equations*. J. Math. Pures Appl. **80**, 669-696 (2001).

[16] S. Bobkov and F. Götze, *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*. J. Funct. Anal. **163**, 1-28 (1999).

[17] S. Bobkov and C. Houdré, *A converse Gaussian Poincaré-type inequality for convex functions*. Statist. Probab. Lett. **44**, 281-290 (1999).

[18] V. Bogachev, *Gaussian measures*. Mathematical Surveys and Monographs 62. American Mathematical Society (1998).

[19] A. Bonami, *Étude des coefficients de Fourier des fonctions de $L_p(G)$*. Ann. Inst. Fourier **20**, 335-402 (1971).

[20] K. Borsuk, *Drei sätze über die n-dimensionale euklidische Sphäre*, Fund. Math. **20** (1933) 177–190.

[21] S. Boucheron, G. Lugosi and P. Massart, *Concentration inequalities using the entropy method.* Ann. Probab. **31**, 1583-1614 (2003).

[22] S. Boucheron, G. Lugosi and P. Massart, *Concentration inequalities. A nonasymptotic theory of independence.* Oxford University Press (2013).

[23] S. Chatterjee, *Superconcentration and related topics.* Springer Monographs in Mathematics. Springer (2014).

[24] D. Cordero-Erausquin, *Some applications of mass transport to Gaussian-type inequalities.* Arch. Ration. Mech. Anal. **161**, 257-269 (2002).

[25] D. Cordero-Erausquin and M. Ledoux, *Hypercontractive measures, Talagrand's inequality, and influences.* Geometric aspects of functional analysis, Lecture Notes in Math. 2050, 169-189. Springer (2012).

[26] E. B. Davies and B. Simon, *Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians.* J. Funct. Anal. **59**, 335-395 (1984).

[27] A. Dembo, *Information inequalities and concentration of measure.* Ann. Probab. **25**, 927-939 (1997).

[28] A. Dembo and O. Zeitouni, *Transportation approach to some concentration inequalities in product spaces.* Electron. Comm. Probab. **1**, 83-90 (1996).

[29] A. Dembo and O. Zeitouni, *Large deviations techniques and applications.* Corrected reprint of the second (1998) edition. Stochastic Modelling and Applied Probability 38. Springer (2010).

[30] J.-D. Deuschel and D. Stroock, *Large Deviations.* Pure and Applied Mathematics 137. Academic Press (1989).

[31] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains.* Ann. Appl. Probab. **6**, 695-750 (1996).

[32] D. P. Dubhashi and A. Panconesi, *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press (2009).

[33] A. Dvoretzky, *Some results on convex bodies and Banach spaces*, Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960) pp. 123–160, Jerusalem Academic Press, Jerusalem; Pergamon, Oxford.

[34] A. Dvoretzky and C. A. Rogers, *Absolute and unconditional convergence in normed linear spaces*, Proc. Nat. Acad. Sci. USA **36** (1950), 192–197.

[35] B. Efron and C. Stein, *The jackknife estimate of variance.* Ann. Statist. **9**, 586-596 (1981).

[36] L. Evans, *Partial differential equations.* Graduate Studies in Mathematics 19. American Mathematical Society (1998).

[37] D. Falik and A. Samorodnitsky, *Edge-isoperimetric inequalities and influences.* Combin. Probab. Comput. **16**, 693-712 (2007).

[38] C. Garban and J. Steif, *Noise sensitivity of Boolean functions and percolation.* Institute of Mathematical Statistics Textbooks 5. Cambridge University Press (2015).

[39] E. Giné and R. Nickl, *Mathematical foundations of infinite-dimensional statistical models.* Cambridge Series in Statistical and Probabilistic Mathematics 40. Cambridge University Press (2015).

[40] N. Gozlan and C. Léonard, *Transport inequalities. A survey.* Markov Process. Related Fields **16**, 635-736 (2010).

[41] N. Gozlan, C. Roberto, P.-M. Samson and P. Tetali, *Kantorovich duality for general transport costs and applications.* J. Funct. Anal. **273**, 3327-3405 (2017).

[42] M. Gromov, *Metric Structures for Riemannian and non-Riemannian Spaces.* Birkhäuser (1998).

[43] L. Gross, *Logarithmic Sobolev inequalities.* Amer. J. Math. **97**, 1061-1083 (1975).

[44] L. Gross, *Logarithmic Sobolev inequalities and contractivity properties of semigroups.* Dirichlet Forms, Lecture Notes in Math. 1563, 54-88. Springer (1993).

[45] A. Guionnet and B. Zegarlinski. *Lectures on logarithmic Sobolev inequalities.* Séminaire de Probabilités XXXVI, Lecture Notes in Math. 1801, 1-134. Springer (2003).

[46] W. Johnson and G. Schechtman, *Remarks on Talagrand's deviation inequality for Rademacher's functions.* Functional Analysis - Longhorn Notes, Lecture Notes in Math. 1470, 72-77. Springer (1991).

[47] J.-P. Kahane, *Some random series of functions.* Heath Math. Monographs (1968). Second Edition Cambridge University Press (1985).

[48] J. Kahn, G. Kalai and N. Linial, *The influence of variables on boolean functions.* 29th Symposium on the Foundations of Computer Science, 68-80 (1988).

[49] M. Ledoux, *Remarks on logarithmic Sobolev constants, exponential integrability and bounds on the diameter.* J. Math. Kyoto Univ. **35**, 211-220 (1995).

[50] M. Ledoux, *Isoperimetry and Gaussian Analysis.* École d' Été de Probabilités de Saint-Flour 1994, Lecture Notes in Math. 1648, 165-294 (1996). Springer.

[51] M. Ledoux, *On Talagrand's deviation inequalities for product measures.* ESAIM Probab. Statist. **1**, 63-87 (1996).

[52] M. Ledoux, *Concentration of measure and logarithmic Sobolev inequalities.* Séminaire de Probabilités XXXIII, Lecture Notes in Math. 1709, 120-216. Springer (1999).

[53] M. Ledoux, *The concentration of measure phenomenon.* Mathematical Surveys and Monographs 89. American Mathematical Society (2001).

[54] M. Ledoux, *Four Talagrand inequalities under the same umbrella.* Preprint (2019).

[55] M. Ledoux and M. Talagrand, *Probability in Banach spaces* (*Isoperimetry and processes*). Ergebnisse der Mathematik und ihrer Grenzgebiete 23. Springer (1991).

[56] M. Lifshits, *Gaussian random functions.* Mathematics and its Applications 322. Kluwer (1995).

[57] G. A. Margulis, *Probabilistic characteristic of graphs with large connectivity.* Problems Info. Transmission. Plenum Press, New York (1977).

[58] C. McDiarmid, *Concentration.* Probabilistic Methods for Algorithmic Discrete Mathematics, Algorithms Combin. **16**, 195-248 (1998). Springer.

[59] K. Marton, *A simple proof of the blowing-up lemma.* IEEE Trans. Inform. Theory **32**, 445-446 (1986).

[60] K. Marton, *Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration.* Ann. Probab. **24**, 857-866 (1996).

[61] K. Marton, *A measure concentration inequality for contracting Markov chains.* Geom. Funct. Anal. **6**, 556-571 (1996).

[62] P. Massart, *About the constants in Talagrand's concentration inequalities for empirical processes.* Ann. Probab. **28**, 863-884 (2000).

[63] P. Massart, *Some applications of concentration inequalities to statistics.* Ann. Fac. Sci. Toulouse Math. **9**, 245-303 (2000).

[64] P. Massart, *Concentration inequalities and model selection.* École d'été de Probabilités de Saint-Flour 2003, Lecture Notes in Math. 1896. Springer (2007).

[65] V. D. Milman, *New proof of the theorem of Dvoretzky on sections of convex bodies*, Funct. Anal. Appl. **5** (1971), 28–37.

[66] V. D. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces.* Lecture Notes in Math. 1200. Springer (1986).

[67] M. Molloy and B. Reed, *Graph colouring and the probabilistic method.* Algorithms and Combinatorics 23. Springer (2002).

[68] E. Nelson, *A quartic interaction in two dimensions.* Mathematical Theory of Elementary Particles, Dedham 1965. Proc. Conf., 69-73. MIT Press, Cambridge (1966).

[69] E. Nelson, *The free Markoff field.* J. Funct. Anal. **12**, 211-227 (1973).

[70] R. O'Donnell, *Analysis of Boolean functions*. Cambridge University Press (2014).

[71] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*. Comm. Partial Differential Equations **26**, 101-174 (2001).

[72] F. Otto and C. Villani, *Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality*. J. Funct. Anal. **173**, 361-400 (2000).

[73] D. Panchenko, *A note on Talagrand's concentration inequality*. Electron. Comm. Probab. **6**, 55-65 (2001).

[74] D. Panchenko, *Symmetrization approach to concentration inequalities for empirical processes*. Ann. Probab. **31**, 2068-2081 (2003).

[75] G. Paouris and P. Valettas, *A Gaussian small deviation inequality for convex functions*, Ann. Probab. **46** (2018), no. 3, 1441–1454.

[76] G. Paouris and P. Valettas, *Dichotomies, structure, and concentration in normed spaces*, Adv. Math. **332** (2018), 438–464.

[77] G. Pisier, *Probabilistic methods in the geometry of Banach spaces*. Probability and Analysis, Lecture Notes in Math. 1206, 167-241 (1986). Springer.

[78] W. Rhee and M. Talagrand, *Martingale inequalities and the jackknife estimate of variance*. Statist. Probab. Lett. **4**, 5-6 (1986).

[79] R. Rossignol, *Threshold for monotone symmetric properties through a logarithmic Sobolev inequality*. Ann. Probab. **34**, 1707-1725 (2006).

[80] G. Royer, *An initiation to logarithmic Sobolev inequalities*. SMF/AMS Texts and Monographs 14, American Mathematical Society (2007). (Translated from the 1999 French original).

[81] L. Russo, *An approximate zero-one law*. Z. Wahrsch. Verw. Gebiete **61**, 129-139 (1982).

[82] P.-M. Samson, *Concentration of measure inequalities for Markov chains and φ-mixing processes*. Ann. Probab. **28**, 416-461 (2000).

[83] P.-M. Samson, *Concentration inequalities for convex functions on product spaces*. Stochastic inequalities and applications, Progr. Probab. **56**, 33-52. Birkhäuser (2003).

[84] P.-M. Samson, *Infimum-convolution description of concentration properties of product probability measures, with applications*. Ann. Inst. H. Poincaré Probab. Statist. **43**, 321-338 (2007).

[85] G. Schechtman, *Concentration results and applications*. Handbook of the geometry of Banach spaces Vol. 2, 1603-1634. North-Holland (2003).

[86] G. Schechtman, *Two observations regarding embedding subsets of Euclidean spaces in normed spaces*, Adv. Math. **200** (2006) 125–135.

[87] P. Sosoe, *Fluctuations in first-passage percolation*. Random growth models, Proc. Sympos. Appl. Math. **75**, 69-93. American Mathematical Society (2018)

[88] M. Steele, *An Efron-Stein inequality for nonsymmetric statistics*. Ann. Statist. **14**, 753-758 (1986).

[89] M. Steele, *Probability theory and combinatorial optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics 69. SIAM (1997).

[90] M. Talagrand, *An isoperimetric theorem on the cube and the Khintchine-Kahane inequalities*. Proc. Amer. Math. Soc. **104**, 905-909 (1988).

[91] M. Talagrand, *A new isoperimetric inequality for product measure and the tails of sums of independent random variables*. Geom. Funct. Anal. **1**, 211-223 (1991).

[92] M. Talagrand, *Logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem*. Geom. Func. Anal. **3** (1993), 295-314.

[93] M. Talagrand, *On Russo's approximate zero-one law*. Ann. Probab. **22**, 1576-1587 (1994).

[94] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*. Inst. Hautes Études Sci. Publ. Math. **81**, 73-205 (1995).

[95] M. Talagrand, *Transportation cost for Gaussian and other product measures*. Geom. Funct. Anal. **6**, 587-600 (1996).

[96] M. Talagrand, *New concentration inequalities in product spaces*. Invent. Math. **126**, 505-563 (1996).

[97] M. Talagrand, *A new look at independence*. Ann. Probab. **24**, 1-34 (1996).

[98] T. Tao, *Topics in random matrix theory*. Graduate Studies in Mathematics 132. American Mathematical Society (2012).

[99] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*. Compressed sensing, 210-268. Cambridge University Press (2012).

[100] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*. Cambridge Series in Statistical and Probabilistic Mathematics 47. Cambridge University Press (2018).

[101] C. Villani, *Topics in optimal transportation*. Graduate Studies in Mathematics 58. American Mathematical Society (2003).

[102] C. Villani, *Optimal transport. Old and new*. Grundlehren der mathematischen Wissenschaften 338. Springer (2009).

[103] M. Wainwright, *High-dimensional statistics. A non-asymptotic viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics 48. Cambridge University Press (2019).