Καμπυλότητα Ricci για Αλυσίδες Markov σε Μετρικούς Χώρους

Γεώργιος Παπαδάκης

Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
Αθήνα – 2018

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Θετική καμπυλότητα Ricci στη γεωμετρία Riemann

Η καμπυλότητα Ricci παίζει κεντρικό ρόλο στη γεωμετρία Riemann, με πολλά και σημαντικά αποτελέσματα να προέρχονται από τη μελέτη της. Οι πολλαπλότητες με θετική καμπυλότητα είναι πολύ διαφορετικοί κόσμοι από αυτές με αρνητική. Πολύ συχνά, σε θεωρήματα αρνητικής καμπυλότητας, υποθέτουμε ότι όλες οι τμηματικές καμπυλότητες είναι αρνητικές, ενώ σε θεωρήματα θετικής καμπυλότητας υποθέτουμε ότι η καμπυλότητα Ricci $\mathrm{Ric}(v,v)$ είναι θετική για κάθε εφαπτόμενο διάνυσμα v, το οποίο είναι λιγότερο ισχυρή υπόθεση από το να έχουμε όλες τις τμηματικές καμπυλότητες $K(v,w)$ μη-αρνητικές. Παρακάτω θα αναφέρουμε κάποια από τα πιο γνωστά αποτελέσματα σε χώρους θετικής καμπυλότητας.

Το θεώρημα Bonnet-Myers είναι ένα από τα κλασικά θεωρήματα για χώρους θετικής καμπυλότητας. Η πρώτη μορφή του θεωρήματος δόθηκε το 1855 από τον Pierre Ossian Bonnet και σχετίζεται με το ερώτημα πότε μια γεωδαισιακή επιτυγχάνει την απόσταση των τελικών της σημείων. Το 1935 ο Sumner Byron Myers απέδειξε την ακόλουθη ισχυρή εκδοχή του θεωρήματος [Ber12]:

Θεώρημα 1.1.1 (Bonnet-Myers)**.** *Έστω (M,g) μια πλήρης πολλαπλότητα Riemann διάστασης n. Υποθέτουμε ότι η καμπυλότητα Ricci της M είναι κάτω φραγμένη από $(n-1)\kappa$, όπου $\kappa>0$. Τότε η M είναι συμπαγής, με τη διάμετρο της να ικανοποιεί την:*

$$\mathrm{diam}(M) := \sup_{p,q\in M} d(p,q) \leqslant \frac{\pi}{\sqrt{\kappa}}.$$

Η πρώτη μη-μηδενική ιδιοτιμή της Λαπλασιανής είναι μεγάλης σημασίας, καθώς από αυτήν εξαρτώνται πολλές σημαντικές ποσότητες σε μια πολλαπλότητα Riemann. Ο υπολογισμός κάτω φραγμάτων γι' αυτήν είναι πολύ σημαντικό ζήτημα, με τα άνω φράγματα να επιτυγχάνονται ευκολότερα αλλά να δίνουν λιγότερο ισχυρά αποτελέσματα. Ένα από τα πρώτα σχετικά αποτελέσματα είναι το παρακάτω θεώρημα που απέδειξε το 1958 ο Andé Lichnerowicz [Ber12]:

Θεώρημα 1.1.2 (Lichnerowicz)**.** *Έστω (M,g) μια πολλαπλότητα Riemann διάστασης n με καμπυλότητα Ricci μεγαλύτερη ή ίση από $n-1$. Τότε η πρώτη μη-μηδενική ιδιοτιμή της Λαπλασιανής* *της πολλαπλότητας είναι τουλάχιστον τόσο μεγάλη όσο η αντίστοιχη της n-διάστατης σφαίρας, δηλαδή n. Επιπλέον, ισότητα ισχύει μόνο για πολλαπλότητες ισοπεριμετρικές με την σφαίρα.*

Το 1917 ο Paul Lévy απέδειξε την κλασική του ισοπεριμετρική ανισότητα στην n-διάστατη σφαίρα. Τη δεκαετία του 1970 ο Mikhail Gromov γενίκευσε αυτήν την ανισότητα για όλες τις πολλαπλότητες Riemann με καμπυλότητα Ricci μεγαλύτερη ή ίση από αυτήν της n-διάστατης σφαίρας. Οι ανισότητες αυτού του είδους συνδέονται άμεσα με το φαινόμενο της συγκέντρωσης του μέτρου [Ben21]:

Θεώρημα 1.1.3 (Ισοπεριμετρική ανισότητα Lévy-Gromov)**.** *Έστω (M, g) μια κλειστή πολλαπλότητα Riemann διάστασης n με καμπυλότητα Ricci μεγαλύτερη ή ίση από $n-1 (= \mathrm{Ricc}(S^n))$. Έστω $M_0 \subset M$ με λείο σύνορο και B μπάλα στην S^{n+1} ώστε να ισχύει η σχέση*

$$\frac{\mathrm{vol}(M_0)}{\mathrm{vol}(M)} = \frac{\mathrm{vol}(B)}{\mathrm{vol}(S^n)}.$$

Τότε ισχύει ότι

$$\frac{\mathrm{vol}(\partial M_0)}{\mathrm{vol}(M)} \geqslant \frac{\partial \mathrm{vol}(B)}{\mathrm{vol}(S^n)}.$$

Η παραπάνω ισοπεριμετρική ανισότητα μπορεί να μας δώσει μια επέκταση του φαινομένου της συγκέντρωσης του μέτρου που παρατηρείται στην n-διάστατη σφαίρα σε μια πολλαπλότητα Riemann με καμπυλότητα Ricci μεγαλύτερη από αυτήν της $n-$διάστατης σφαίρας:

Πόρισμα 1.1.4. *(Συγκέντρωση του Μέτρου σε χώρους Θετικής Καμπυλότητας) Έστω M μια $n-$διάστατη πολλαπλότητα Riemann με καμπυλότητα Ricci τουλάχιστον ίση με αυτήν της $n-$διάστατης σφαίρας S^n. Έστω επίσης $f : M \to \mathbb{R}$ μια 1-Lipschitz συνάρτηση. Τότε υπάρχει $m \in \mathbb{R}$ τέτοιος ώστε για κάθε $t \geqslant 0$ να ισχύει*

$$\nu\left(\{x \in S^n : \ |f(x) - m| \geqslant t\}\right) \leqslant 2\mathrm{exp}\left\{-\frac{t^2}{2D^2}\right\}$$

όπου $D = 1/\sqrt{n-1}$ και ν είναι το φυσιολογικό μέτρο στην M, κανονικοποιημένο ώστε να είναι μέτρο πιθανότητας.

Το εύρος αυτών των θεωρημάτων επεκτάθηκε σε σημαντικό βαθμό το 1985 από την θεωρία Bakry-Émery, η οποία αναδεικνύει την αναλυτική και στοχαστική σημασία της καμπυλότητας Ricci. Πιο συγκεκριμένα αποδεικνύουν μια λογαριθμική ανισότητα Sobolev για χώρους θετικής καμπυλότητας [BE85].

Θεώρημα 1.1.5 (Bakry-Émery)**.** *Έστω M μια n-διάστατη πολλαπλότητα Riemann με καμπυλότητα Ricci τουλάχιστον ίση με $K > 0$. Έστω επίσης $(P_t)_{t \geqslant 0}$ η ημιομάδα της εξίσωσης θερμότητας στην M. Τότε για κάθε $t \geqslant 0$ και για κάθε λεία συνάρτηση $f : M \to \mathbb{R}$ ισχύουν τα εξής:*

(i) $\sup_M \|\nabla P_t f\| \leqslant e^{-Kt} \sup_M \|\nabla f\|$

(ii) $\|\nabla P_t f\|(x) \leqslant e^{-Kt} \left(P_t \|\nabla f\|\right)(x)$

(iii) $\mathrm{Ent}\ f := \int f \ \ln \frac{f}{\int f} d\nu \leqslant \frac{1}{2K} \int \frac{\|\nabla f\|^2}{f} d\nu, \ f > 0$

όπου ν είναι το φυσιολογικό μέτρο στην M, κανονικοποιημένο ώστε να είναι μέτρο πιθανότητας.

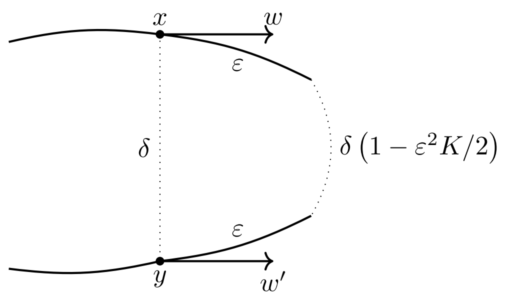
1.2 Καμπυλότητα Ricci σε μετρικούς χώρους

Τα αποτελέσματα της προηγούμενης παραγράφου έχουν οδηγήσει σε πολλές προσπάθειες για μια γενίκευση της έννοιας της καμπυλότητας Ricci σε χώρους με λιγότερο ισχυρή γεωμετρία. Ο ορισμός τον οποίο θα μελετήσουμε οφείλεται στον Yann Ollivier και βασίζεται στην παρατήρηση ότι σε χώρους θετικής καμπυλότητας «μικρές σφαίρες απέχουν λιγότερο από ότι τα κέντρα τους». Πιο συγκεκριμένα, ισχύει η ακόλουθη πρόταση [Oll09]:

Πρόταση 1.2.1. *Έστω (X, d) μια πλήρης, λεία πολλαπλότητα Riemann. Έστω v, w μοναδιαία εφαπτόμενα διανύσματα στο $x \in X$. Έστω $\varepsilon, \delta > 0$. Έστω $y = \exp_x \delta v$ και w' το εφαπτόμενο διάνυσμα στο y που προκύπτει από παράλληλη μετατόπιση του w κατά μήκος της γεωδαισιακής $\exp_x \delta v$. Τότε:*

$$d(\exp_x \varepsilon w, \exp_y \varepsilon w') = \delta \left(1 - \frac{\varepsilon^2}{2} K(v, w) + O(\varepsilon^3 + \varepsilon^2 \delta) \right)$$

καθώς $(\varepsilon, \delta) \to 0$. Η ποσότητα $K(v, w)$ είναι η τμηματική καμπυλότητα του επιπέδου που σχηματίζουν τα v, w.



Για να κατασκευαστεί μια μπάλα με κέντρο x και ακτίνα ε σε μια πολλαπλότητα Riemann αρκεί να θεωρήσουμε όλες τις γεωδαισιακές ευθείες με αρχή το x και μήκος ε. Δηλαδή από το x μετακινούμαστε στο y μέσω της γεωδαισιακής που τα συνδέει. Αυτός ο ορισμός της μπάλας βασίζεται με πολύ ισχυρό τρόπο στη γεωμετρία των πολλαπλοτήτων Riemann. Πρέπει λοιπόν να ορίσουμε τις μπάλες, με έναν τρόπο που δε θα εξαρτάται από την έννοια της γεωδαισιακής ευθείας. Η ιδέα είναι να ανακατασκευάσουμε μια μπάλα μέσω μιας στοχαστικής διαδικασίας κατά την οποία σε κάθε βήμα θα μεταπηδούμε από το x σε ένα σημείο κοντά στο x, σε αντίθεση με το να μετακινούμαστε σε κοντινά σημεία μέσω γεωδαισιακών:

Έστω (M, g) μια λεία πολλαπλότητα Riemann, διάστασης n, και d η μετρική Riemann. Για κάποιο $\varepsilon > 0$, θεωρούμε την αλυσίδα Markov m^ε:

$$dm_x^\varepsilon(y) := \chi_{B(x,\varepsilon)} \frac{d\mathrm{vol}(y)}{\mathrm{vol}(B(x, \varepsilon))}.$$

Σταθεροποιούμε $x \in X$ και θεωρούμε ένα μοναδιαίο εφαπτόμενο διάνυσμα v με αρχή το x. Έστω επίσης y ένα σημείο στη γεωδαισιακή με αρχική ταχύτητα v, και ας υποθέσουμε ότι η απόσταση $d(x, y)$ είναι αρκετά μικρή. Τότε

$$\frac{W_1(m_x^\varepsilon, m_y^\varepsilon)}{d(x, y)} = 1 - \frac{\varepsilon^2 \mathrm{Ric}(v, v)}{2(n+2)} - O(\varepsilon^3 + \varepsilon^2 d(x, y)),$$

όπου με $W_1((m_x^\varepsilon, m_y^\varepsilon)$ συμβολίζουμε την απόσταση Wasserstein των m_x^ε και m_y^ε. Όταν η καμπυλότητα Ricci είναι θετική έχουμε ότι $W_1(m_x^\varepsilon, m_y^\varepsilon) < d(x, y)$. Με αυτόν τον τρόπο μπορούμε να δούμε ότι αν χρησιμοποιήσουμε την αλυσίδα Markov m^ε στη θέση της μπάλας με ακτίνα ε θα διατηρηθεί η ιδιότητα που περιγράψαμε προηγουμένως: «μικρές σφαίρες απέχουν λιγότερο από ότι τα κέντρα τους».

Αν λοιπόν θεωρήσουμε έναν μετρικό χώρο (X, d) και έναν τυχαίο περίπατο $m = \{m_x\}_{x\in X}$ στον μετρικό χώρο, ο οποίος θα παίζει τον ρόλο της μικρής μπάλας, ορίζουμε με φυσιολογικό, ως προς την παραπάνω ανάλυση, τρόπο την coarse καμπυλότητα Ricci του τυχαίου περίπατου στα $x, y \in X$ ως:

$$\kappa(x, y) = 1 - \frac{W_1(m_x, m_y)}{d(x, y)}.$$

Ο ορισμός αυτός θα οδηγήσει σε (ικανοποιητικές) γενικεύσεις των θεωρημάτων που αναφέραμε παραπάνω. Σημαντικό ρόλο στις αποδείξεις αυτών των γενικεύσεων θα παίξει η μελέτη του χώρου των μέτρων πιθανότητας μ στον X με πεπερασμένη πρώτη ροπή, δηλαδή με $\int_x d(o, x)d\mu(x) < \infty$ για κάποιο $o \in X$ (και συνεπώς για όλα τα $o \in X$), τον οποίο συμβολίζουμε με $\mathscr{P}(X)$ και ο οποίος εφοδιασμένος με την απόσταση Wasserstein γίνεται μετρικός χώρος. Μπορούμε να ορίσουμε μια πράξη $*$ μεταξύ μέτρων στον $\mathscr{P}(X)$ και τυχαίων περιπάτων στον μετρικό χώρο, και το αποτέλεσμα αυτής είναι ένα καινούργιο μέτρο πιθανότητας το οποίο δεν ανήκει απαραίτητα στον $\mathscr{P}(X)$. Αν γνωρίζουμε ότι η coarse καμπυλότητα Ricci του μετρικού χώρου είναι ομοιόμορφα κάτω φραγμένη για όλα τα ζευγάρια σημείων του X τότε το καινούργιο μέτρο θα ανήκει πράγματι στον $\mathscr{P}(X)$, ενώ αν επιπλέον το κάτω φράγμα είναι θετικός αριθμός θα υπάρχει μια αναλλοίωτη κατανομή για τον τυχαίο περίπατο, δηλαδή θα υπάρχει μέτρο πιθανότητας ν που ικανοποιεί την $\nu * m = \nu$.

Θα περάσουμε τώρα στις διατυπώσεις των αντίστοιχων των θεωρημάτων που αναφέραμε παραπάνω:

Πρόταση 1.2.2 (L^1 θεώρημα Bonnet-Myers)**.** *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και $\kappa(x, y) \geqslant \kappa > 0$ για κάθε $x, y \in X$ με $x \neq y$. Τότε για κάθε $x, y \in X$ ισχύει ότι*

$$d(x, y) \leqslant \frac{J(x) + J(y)}{\kappa(x, y)}$$

και ειδικότερα

$$\operatorname{diam}(X) \leqslant \frac{2\sup\{J(x) : x \in X\}}{\kappa},$$

όπου $J(x) = W_1(\delta_x, m_x)$.

Το φράγμα που δίνει αυτό το θεώρημα θα είναι το σωστό για όλα τα παραδείγματα που θα αναφερθούν παρακάτω, εκτός από αυτό της κίνησης Brown σε πολλαπλότητα Riemann. Αυτό γιατί, σε αυτήν την περίπτωση, $J \approx \varepsilon$ και $\kappa \approx \varepsilon^2 \mathrm{Ricc}/n$, οπότε η προηγούμενη πρόταση υστερεί κατά ένα όρο $1/\varepsilon$ σε σχέση με το κλασικό θεώρημα Bonnet-Myers. Στην επόμενη πρόταση θα βάλουμε περισσότερες υποθέσεις για να πετύχουμε το σωστό φράγμα, αλλά δυστυχώς το συγκεκριμένο αποτέλεσμα εφαρμόζεται μονάχα στις πολλαπλότητες Riemann.

Πρόταση 1.2.3 (ισχυρό θεώρημα Bonnet-Myers στον L^2)**.** *Έστω (X, d) μετρικός χώρος εφοδιασμένος με έναν τυχαίο περίπατο συνεχούς χρόνου m^{*t}. Έστω επίσης ότι ο X είναι $\varepsilon-$γεωδαισιακός,* *και ότι υπάρχουν σταθερές* $\kappa > 0$, $C \geqslant 0$ *ώστε για κάθε δύο αρκετά μικρά* t, s *και για κάθε δύο* $x, y \in X$ *με* $\varepsilon \leqslant d(x,y) \leqslant 2\varepsilon$ *να ισχύει η ανισότητα*

$$W_1(m_x^{*t}, m_y^{*s}) \leqslant e^{-\kappa \min\{t,s\}} d(x,y) + \frac{C\left(\sqrt{t}-\sqrt{s}\right)^2}{2d(x,y)}.$$

Αν επιπλέον υποθέσουμε ότι $\varepsilon \leqslant \frac{1}{2}\sqrt{C/2\kappa}$, *τότε*

$$\operatorname{diam}(X) \leqslant \pi \sqrt{\frac{C}{2\kappa}} \left(1 + \frac{4\varepsilon}{\sqrt{C/2\kappa}}\right).$$

Περνάμε τώρα σε ένα φασματικό θεώρημα, το οποίο αντιστοιχεί στο θεώρημα του Lichnerowicz:

Ορισμός 1.2.4 (τελεστής τυχαίου περίπατου, Λαπλασιανός τελεστής)**.** Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και έστω ν η αναλλοίωτη κατανομή του m. Ορίζουμε ως *(coarse) σταθερά διάχυσης* του τυχαίου περίπατου στο $x \in X$ την ποσότητα

$$\sigma(x) := \left(\frac{1}{2}\int\int d(y,z)^2 dm_x(y) dm_x(z)\right)^{1/2}$$

και ως *μέση σταθερά διάχυσης* την ποσότητα

$$\sigma := \|\sigma(x)\|_{L^2(X,\nu)}.$$

Για κάθε $f \in L^2(X,\nu)$ ορίζουμε τον **τελεστή τυχαίου περίπατου** ως τον τελεστή M με

$$Mf(x) := \int_{y \in X} f(y) dm_x(y).$$

Επίσης ορίζουμε τον **Λαπλασιανό τελεστή** ως τον $\Delta := M - Id$.

Πρόταση 1.2.5 (φασματικό κενό)**.** *Έστω* (X, d, m) *μετρικός χώρος με έναν τυχαίο περίπατο και έστω ότι* $\kappa(x,y) \geq \kappa > 0$ *για κάθε* $x, y \in X$. *Έστω επίσης* ν *η μοναδική αναλλοίωτη κατανομή του* m. *Υποθέτουμε ότι* $\sigma < \infty$ *και ότι η* ν *είναι αντιστρέψιμη, δηλαδή ότι* $d\nu(x)dm_x(y) = d\nu(y)dm_y(x)$, *ή ότι ο* X *είναι πεπερασμένος. Τότε, η φασματική ακτίνα του τελεστή μέσου όρου στον χώρο* $L^2(X,\nu)/\sim$ *είναι το πολύ* $1-\kappa$.

Αυτού του είδους τα αποτελέσματα είναι ιδιαιτέρως χρήσιμα γιατί συνήθως έχουν ως συνέπεια ανισότητες τύπου Poincaré:

Πόρισμα 1.2.6 (ανισότητα Poincaré)**.** *Έστω* (X, d, m) *ένας μετρικός χώρος με έναν εργοδικό τυχαίο περίπατο και έστω* $\kappa(x,y) \geq \kappa > 0$ *για κάθε* $x, y \in X$. *Έστω επίσης* ν *η μοναδική αναλλοίωτη κατανομή του* m. *Υποθέτουμε ότι* $\sigma < \infty$ *και ότι η* ν *είναι αντιστρέψιμη. Τότε, το φάσμα του* $-\Delta$ *στον* $L^2(X\nu)/\sim$ *περιέχεται στο* $[\kappa, \infty)$. *Επιπλέον ισχύουν οι παρακάτω διακριτές ανισότητες Poincaré για κάθε* $f \in L^2(X,\nu)$:

$$\operatorname{Var}_\nu f \leqslant \frac{1}{\kappa(2-\kappa)} \int \operatorname{Var}_{m_x} f d\nu(x)$$

και

$$\operatorname{Var}_\nu f \leqslant \frac{1}{2\kappa} \int\int \left(f(x) - f(y)\right)^2 d\nu(x) d\nu(y).$$

Όσον αφορά την ισοπεριμετρική ανισότητα Lévy-Gromov αυτό που μας ενδιαφέρει ουσιαστικά είναι το πόρισμα της, δηλαδή η συγκέντρωση μέτρου σε χώρους θετικής καμπυλότητας. Έχουμε λοιπόν το εξής θεώρημα:

Θεώρημα 1.2.7 (Gaussian συγκέντρωση). *Έστω (X, d, m) τυχαίος περίπατος σε έναν μετρικό χώρο με coarse καμπυλότητα Ricci τουλάχιστον $\kappa > 0$. Έστω ν η μοναδική αναλλοίωτη κατανομή.*

Θέτουμε

$$D_x^2 := \frac{\sigma(x)^2}{n_x \kappa}$$

και

$$D^2 := \mathbb{E}_\nu D_x^2.$$

Υποθέτουμε ότι η συνάρτηση $x \mapsto D_x^2$ είναι C-Lipschitz. Ορίζουμε

$$t_{\max} := \frac{8}{9} D^2 \min\{3/2C, 1/\sigma_\infty\}.$$

Τότε, για κάθε 1-Lipschitz συνάρτηση f, αν $t \leqslant t_{\max}$ έχουμε

$$\nu(\{x : f(x) \geqslant t + \mathbb{E}_\nu f\}) \leqslant \exp\left\{-\frac{3t^2}{16D^2}\right\},$$

και αν $t \geqslant t_{\max}$ έχουμε

$$\nu(\{x : f(x) \geqslant t + \mathbb{E}_\nu f\}) \leqslant \exp\left\{-\frac{3t_{\max}}{16D^2} t\right\}.$$

Για να δειχθεί μια λογαριθμική ανισότητα Sobolev θα πρέπει να δώσουμε πρώτα έναν ορισμό για την κλίση, λαμβάνοντας υπόψιν τα καινούργια εργαλεία που έχουν εισαχθεί:

Ορισμός 1.2.8 (νόρμα της κλίσης). Επιλέγουμε $\lambda > 0$ και για κάθε συνάρτηση $f : X \to \mathbb{R}$ ορίζουμε την *λ-επιπέδου κλίση* της f θέτοντας

$$|\nabla_\lambda f|(x) := \sup_{y, y' \in X} \frac{|f(y) - f(y')|}{d(y, y')} e^{-\lambda d(x,y) - \lambda d(y,y')}.$$

Θεώρημα 1.2.9 (ανισότητα Log-Sobolev). *Έστω (X, d, m) τυχαίος περίπατος σε έναν μετρικό χώρο με coarse καμπυλότητα Ricci τουλάχιστον $\kappa > 0$. Έστω ν η μοναδική αναλλοίωτη κατανομή. Έστω $\lambda \leqslant \frac{1}{20\sigma_\infty(U+1)}$, όπου $\sigma_\infty = \frac{1}{2}\sup_x \operatorname{diam}(\operatorname{supp}(m_x))$, και θεωρούμε την λ-επιπέδου κλίση $|\nabla_\lambda|$. Τότε για κάθε $f : X \to \mathbb{R}$ με $|\nabla_\lambda f| < \infty$ έχουμε ότι*

$$\operatorname{Var}_\nu(f) \leqslant \left(\sup_x \frac{4\sigma^2(x)}{\kappa n_x}\right) \int_x |\nabla_\lambda f|^2(x) d\nu(x)$$

και αν η f είναι θετική,

$$\operatorname{Ent}_\nu(f) \leqslant \left(\sup_x \frac{4\sigma^2(x)}{\kappa n_x}\right) \int_x \frac{|\nabla_\lambda f|^2(x)}{f(x)} d\nu(x).$$

Αν επιπλέον ο τυχαίος περίπατος m είναι αναλλοίωτος ως προς τη ν, τότε

$$\operatorname{Var}_\nu(f) \leqslant \int_x V(x) |\nabla_\lambda f|^2(x) d\nu(x)$$

και

$$\mathrm{Ent}_\nu(f) \leqslant \int_x V(x)\frac{|\nabla_\lambda f|^2(x)}{f(x)}d\nu(x),$$

όπου

$$V(x) = 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t} M^{t+1}\left(\frac{\sigma^2(x)}{n_x}\right).$$

Επιπλέον, αν υποθέσουμε ότι ο μετρικός χώρος είναι γεωδαισιακός, ότι υπάρχει ένα «σημείο έλξης» και ότι ικανοποιείται η Gaussian-τύπου ανισότητα μετασχηματισμού Laplace, τότε επιτυγχάνεται ένα θεώρημα εκθετικής συγκέντρωσης:

Θεώρημα 1.2.10 (εκθετική συγκέντρωση)**.** *Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο. Έστω επίσης ότι ο (X, d) είναι r-γεωδαισιακός, $r > 0$, και υπάρχει $o \in X$ ώστε να ισχύουν οι παρακάτω υποθέσεις:*

(i) *Για κάθε $x, y \in X$ έχουμε $\kappa(x, y) \geqslant 0$ (μη-αρνητική καμπυλότητα).*

(ii) *Για κάθε $x \in X$ με $r \leqslant d(x, o) < 2r$ ισχύει ότι $W_1(m_x, \delta_o) < d(x, o)$ (το o «έλκει τα κοντινά σημεία»).*

(iii) *Υπάρχει $s > 0$ ώστε για κάθε m_x, $x \in X$ να ικανοποιείται μια Gaussian-τύπου ανισότητα μετασχηματισμού Laplace: Για κάθε $\lambda > 0$ και για κάθε 1-Lipschitz συνάρτηση $f : \mathrm{supp}(m_x) \to \mathbb{R}$ ισχύει*

$$\mathbb{E}_{m_x}\left(e^{\lambda f}\right) \leqslant e^{\lambda^2 s^2/2} e^{\mathbb{E}_{m_x}(f)}.$$

Θέτουμε $\rho = \inf\{d(x, o) - W_1(m_x, \delta_o) : x \in X \text{ με } r \leqslant d(x, o) < r\}$ και υποθέτουμε ότι $\rho > 0$. Τότε υπάρχει αναλλοίωτη κατανομή για τον τυχαίο περίπατο. Επιπλέον αν $D = s^2/\rho$ και $M = r + 2s^2/\rho + \rho\left(1 + J^2(o)/4s^2\right)$, τότε για κάθε αναλλοίωτη κατανομή ν για τον τυχαίο περίπατο ισχύει ότι

$$\int_x e^{d(x,o)/D} d\nu(x) \leqslant \left(4 + J^2(o)/s^2\right) e^{M/D}$$

και συνεπώς για κάθε 1-Lipschitz συνάρτηση $f : X \to \mathbb{R}$ και για κάθε $t \geqslant 0$ έχουμε την

$$\nu\left(|f - f(o)| \geqslant t + M\right) \leqslant \left(8 + 2J^2(o)/s^2\right) e^{-t/D}.$$

Δηλαδή, έχουμε εκθετική συγκέντρωση. Η τρίτη συνθήκη ισχύει πάντα για $s = \sigma_\infty$, λόγω της Πρότασης 1.16 στο [Led01].

Το δεύτερο κεφάλαιο ολοκληρώνεται με μια γενίκευση του ορισμού της coarse καμπυλότητας Ricci, ώστε να συνδεθεί με μια τοπολογία τύπου Gromov-Hausdorff. Με αυτόν τον τρόπο μπορεί κανείς να μιλήσει για όρια ακολουθιών μετρικών χώρων και των αντίστοιχων καμπυλοτήτων, ενώ επίσης μπορεί να αλλάξει τον τυχαίο περίπατο ή και τον μετρικό χώρο έτσι ώστε η καμπυλότητα να αλλάζει με προβλεπόμενο τρόπο.

Στο πρώτο κομμάτι του τρίτου κεφαλαίου αναφέρουμε τα παραδείγματα που αποτελούν την βάση της όλης θεωρίας, και εκεί ακριβώς φαίνεται η δύναμη του ορισμού. Από την μία πλευρά τα παραδείγματα προερχονται από μεγάλο εύρος διαφορετικών περιοχών των μαθηματικών, ενώ από την άλλη ο υπολογισμός της καμπυλότητας και των σχετικών ποσοτήτων είναι σχετικά εφικτός. Στο δεύτερο κομμάτι του κεφαλαίου, επικεντρωνόμαστε σε γραφήματα, εφοδιασμένα με έναν lazy τυχαίο περίπατο. Η δομή του χώρου μας επιτρέπει να απαλλαγούμε από τον τυχαίο περίπατο, και να μιλάμε πλέον για coarse καμπυλότητα Ricci του γραφήματος [LLY11].

Στο τέταρτο κεφάλαιο ασχολούμαστε με ανισότητες μεταφοράς της εντροπίας και αποδεικνύουμε ότι σε χώρους θετικής καμπυλότητας, αν όλα τα στοιχεία του τυχαίου περίπατου ικανοποιούν μια ανισότητα Gaussian συγκέντρωσης, η οποία είναι ισοδύναμη με την ανισότητα (T_1), τότε το ίδιο θα συμβεί και στην οριακή αναλλοίωτη κατανομή [ERL17]:

Θεώρημα 1.2.11. *Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο και υποθέτουμε ότι η coarse καμπυλότητα Ricci του τυχαίου περίπατου είναι $\kappa > 0$. Επίσης υποθέτουμε ότι για κάθε $x \in X$ τα μέτρα m_x ικανοποιούν την (T_1) με την ίδια σταθερά C. Τότε η αναλλοίωτη κατανομή του τυχαίου περίπατου ικανοποιεί την (T_1) με σταθερά $\frac{C}{\kappa(2-\kappa)}$.*

Τέλος στο πέμπτο κεφάλαιο μελετάμε μια καμπυλωμένη ανισότητα Brunn-Minkowski στον διακριτό κύβο, η οποία δίνει μια ισχυρή επιβεβαίωση του ότι ο ορισμός της coarse καμπυλότητας Ricci είναι «σωστός», αφού παρόμοιου τύπου ανισότητα ισχύει σε πολλαπλότητες Riemann θετικής καμπυλότητας.

ΚΕΦΑΛΑΙΟ 2

Coarse καμπυλότητα Ricci

2.1 Coarse καμπυλότητα Ricci

Ορισμός 2.1.1. Έστω (X, d) Πολωνικός μετρικός χώρος, εφοδιασμενος με την Borel σ-άλγεβρα $\mathscr{B}(X)$. Ένας **τυχαίος περίπατος** m στον X είναι μια οικογένεια $\{m_x(\cdot)\}_{x\in X}$ μέτρων πιθανότητας στην $\mathscr{B}(X)$, η οποία ικανοποιεί τις εξής (τεχνικές) συνθήκες:

(i) Το μέτρο m_x εξαρτάται με μετρήσιμο τρόπο από το σημείο $x \in X$.

(ii) Για κάθε $x \in X$ το μέτρο m_x έχει πεπερασμένη πρώτη ροπή: δηλαδή για κάποιο (και συνεπώς για όλα τα) $o \in X$ και για κάθε $x \in X$ έχουμε ότι

$$\int d(o, y)dm_x(y) < \infty.$$

Η μέτρηση της απόστασης των «σφαιρών» γύρω από δύο σημεία θα γίνει μέσω της απόστασης μεταφοράς:

Ορισμός 2.1.2. Έστω (X, d) μετρικός χώρος και ν_1, ν_2 δύο μέτρα πιθανότητας στον X. Η $\mathbf{L^1}$ **απόσταση μεταφοράς** μεταξύ των ν_1, ν_2 είναι η ποσότητα:

$$W_1(\nu_1, \nu_2) := \inf_{\xi\in\Pi(\nu_1,\nu_2)} \int\limits_{(x,y)\in X\times X} d(x, y)d\xi(x, y),$$

όπου $\Pi(\nu_1, \nu_2)$ είναι το σύνολο όλων των ταιριασμάτων των ν_1, ν_2, δηλαδή το σύνολο όλων των μέτρων πιθανότητας ξ στον $X \times X$ για τα οποία ισχύουν οι $\xi(A \times X) = \nu_1(A)$ για κάθε $A \in \mathscr{B}(X)$ και $\xi(X \times B) = \nu_2(B)$ για κάθε $B \in \mathscr{B}(X)$.

Διαισθητικά το $d\xi(x, y)$ αντιπροσωπεύει τη μάζα που ταξιδεύει από το x στο y έτσι ώστε το αρχικό μέτρο να είναι ν_1 και το τελικό ν_2. Το infimum επιτυγχάνεται, αλλά το βέλτιστο ταίριασμα δεν είναι εν γένει μοναδικό. Στα παρακάτω μας αρκεί μια επιλογή βέλτιστου ταιριάσματος.

Είμαστε τώρα έτοιμοι για τη γενίκευση της καμπυλότητας Ricci:

Ορισμός 2.1.3. Έστω (X, d) μετρικός χώρος και m ένας τυχαίος περίπατος στον X. Έστω $x, y \in X,\ x \neq y$. Η **coarse καμπυλότητα Ricci** της τριάδας (X, d, m) κατά μήκος του (xy) είναι η ποσότητα

$$\kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)}.$$

Μπορούμε να δούμε το μέτρο πιθανότητας m_x με δύο τρόπους. Η οπτική της γεωμετρίας είναι να φανταστούμε το m_x ως μια μπάλα με κέντρο το x. Με όρους θεωρίας πιθανοτήτων τα m_x ορίζουν μια αλυσίδα Markov της οποίας η πιθανότητα μετάβασης από το σημείο x στο σημείο y σε n το πλήθος βήματα είναι

$$dm_x^{*n}(y) := \int\limits_{z \in X} dm_x^{*(n-1)}(z) dm_z(y),$$

όπου $m_x^{*1} = m_x$.

Ορισμός 2.1.4. Ένα μέτρο πιθανότητας ν είναι **αναλλοίωτο** για τον τυχαίο περίπατο m αν $d\nu(x) = \int_X d\nu(y) dm_y(x)$ και είναι **αντιστρέψιμο** αν επιπλέον ισχύει

$$d\nu(x) dm_x(y) = d\nu(y) dm_y(x).$$

2.2 Γεωδαισιακοί χώροι

Η ιδέα πίσω από την καμπυλότητα είναι να χρησιμοποιήσουμε τοπικές ιδιότητες για να συμπεράνουμε ολικές. Ακολουθεί μια απλή πρόταση που μας λέει ότι σε σχεδόν γεωδαισιακούς χώρους, όπως είναι τα γραφήματα, ή πολλαπλότητες, αρκεί να υπολογίσουμε την coarse καμπυλότητα Ricci για σημεία που είναι «κοντά»:

Πρόταση 2.2.1. *Έστω (X, d) ε-γεωδαισιακός χώρος, δηλαδή για κάθε δύο $x, y \in X$, υπάρχει $n \in \mathbb{N}$ και πεπερασμένη ακολουθία σημείων $x_0 = x, x_1, \ldots, x_n = y$ ώστε*

$$d(x, y) = \sum_{i=0}^{n-1} d(x_i, x_{i+1}) \text{ και } d(x_i, x_{i+1}) \leqslant \varepsilon.$$

Αν $\kappa(x, y) \geqslant \kappa$ για κάθε δύο σημεία με $d(x, y) \leqslant \varepsilon$, τότε $\kappa(x, y) \geqslant \kappa$ για κάθε $x, y \in X$.

Απόδειξη. Έστω $x, y \in X$. Τότε υπάρχει $n \in \mathbb{N}$ και ακολουθία σημείων $x_0 = x, x_1, \ldots, x_n = y$ ώστε

$$d(x, y) = \sum_{i=0}^{n-1} d(x_i, x_{i+1}) \text{ και } d(x_i, x_{i+1}) \leqslant \varepsilon.$$

Έχουμε ότι:

$$W_1(m_x, m_y) \leqslant \sum_{i=0}^{n-1} W_1(m_{x_i}, m_{x_{i+1}}) \leqslant (1 - \kappa) \sum_{i=0}^{n-1} d(x_i, x_{i+1}) = (1 - \kappa) d(x, y).$$

□

2.3 Συστολή στον χώρο $\mathscr{P}(X)$

Έστω (X,d) μετρικός χώρος. Ορίζουμε ως $\mathscr{P}(X)$ τον χώρο των μέτρων πιθανότητας μ στον X με πεπερασμένη πρώτη ροπή, δηλαδή υποθέτουμε ότι για κάποιο (και συνεπώς για όλα τα) $o \in X$ έχουμε $\int d(o,x)d\mu(x) < \infty$. Σε αυτόν τον χώρο η απόσταση μεταφοράς W_1 είναι πεπερασμένη και συνεπώς είναι μετρική.

Έστω $\mu \in \mathscr{P}(X)$ και m τυχαίος περίπατος στον X. Ορίζουμε το μέτρο:

$$\mu * m := \int\limits_{x\in X} m_x d\mu(x),$$

το οποίο είναι η εικόνα του μ μέσω του τυχαίου περίπατου m. Γενικά το μέτρο αυτό δεν ανήκει πάντα στον $\mathscr{P}(X)$.

Πρόταση 2.3.1. *Έστω (X,d,m) μετρικός χώρος με έναν τυχαίο περίπατο. Έστω $\kappa \in \mathbb{R}$. Τότε ισχύει ότι $\kappa(x,y) \geqslant \kappa$ για κάθε $x,y \in X$ αν και μόνο αν για κάθε δύο κατανομές πιθανότητας $\mu, \mu' \in \mathscr{P}(X)$ ισχύει ότι*

$$W_1(\mu * m, \mu' * m) \leqslant (1-\kappa)W_1(\mu,\mu'). \tag{2.3.1}$$

Σε αυτήν την περίπτωση, αν $\mu \in \mathscr{P}(X)$ τότε $\mu * m \in \mathscr{P}(X)$.

Απόδειξη. Έστω αρχικά ότι ισχύει η σχέση (2.3.1). Τότε για κάθε $x,y \in X$ ορίζουμε $\mu = \delta_x$ και $\mu' = \delta_y$, όπου δ_x και δ_y τα μέτρα Dirac στα x και y αντίστοιχα.

Έχουμε ότι $\mu * m = \delta_x * m = m_x$ και $\mu' * m = \delta_y * m = m_y$ και επίσης ισχύει ότι $W_1(\delta_x,\delta_y) = d(x,y)$.

Πράγματι, το μοναδικό ταίριασμα των δ_x και δ_y είναι το $\delta_{(x,y)}$: Το $\delta_{(x,y)}$ είναι προφανώς ταίριασμα γιατί, για κάθε ζευγάρι Borel συνόλων $A, B \subset X$ έχουμε

$$\delta_{(x,y)}(A \times X) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

και

$$\delta_{(x,y)}(X \times B) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

δηλαδή $\delta_{(x,y)}(A \times X) = \delta_x(A)$ και $\delta_{(x,y)}(X \times B) = \delta_y(B)$.

Επίσης, αν ξ είναι ένα ταίριασμα των μ και μ' τότε για κάθε ζευγάρι Borel συνόλων $A, B \subset X$ έχουμε $\xi(A \times X) = \delta_x(A)$ και $\xi(X \times B) = \delta_y(B)$.

- Αν $(x,y) \in A \times B$ έχουμε ότι $\xi(A \times B) = 1$.
- Αν $(x,y) \notin A \times B$ και για παράδειγμα $x \notin A$ τότε $\xi(A \times B) \leqslant \xi(A \times X) = \delta_x(A) = 0$.

Συνεπώς έχουμε ότι $\xi = \delta_{(x,y)}$. Τελικά,

$$W_1(\mu * m, \mu' * m) = W_1(m_x, m_y) \leqslant (1-\kappa)W_1(\delta_x,\delta_y) = (1-\kappa)d(x,y).$$

Αυτό είναι ισοδύναμο με το ότι $\kappa(x,y) \geqslant \kappa$.

Αντίστροφα, αν $x, y \in X$ θεωρούμε το ταίριασμα ξ_{xy} των m_x και m_y που επιτυγχάνει την απόσταση $W_1(m_x, m_y)$ (μπορούμε να επιλέξουμε την ξ_{xy} μετρήσιμη από το Πόρισμα 5.22 στο [Vill08]):

$$W_1(m_x, m_y) = \int\limits_{x',y'} d(x', y') d\xi_{xy}(x', y').$$

Από την υπόθεση έχουμε ότι $\kappa(x', y') \geqslant \kappa$ και συνεπώς

$$\int\limits_{x',y'} d(x', y') d\xi_{xy}(x', y') \leqslant (1-\kappa) d(x, y).$$

Θεωρούμε επίσης το βέλτιστο ταίριασμα Ξ μεταξύ των μ και μ':

$$W_1(\mu, \mu') = \int\limits_{x,y} d(x,y) d\Xi(x,y).$$

Τότε το

$$\Xi^* = \int\limits_{x,y} \xi_{xy} d\Xi(x,y)$$

είναι ταίριασμα των $\mu * m$ και $\mu' * m$ διότι

$$\Xi^*(A \times X) = \int_{X \times X} \xi_{xy}(A \times X) d\Xi(x,y) = \int_X m_x(A) d\mu(x) = (\mu * m)(A)$$

για κάθε Borel σύνολο $A \subset X$, και ομοίως ισχύει ότι $\Xi^*(X \times B) = (\mu' * m)(B)$ για κάθε Borel σύνολο $B \subset X$.

Συνεπώς, έχουμε ότι:

$$\begin{aligned}
W_1(\mu * m, \mu' * m) &\leqslant \int\limits_{x,y} d(x,y) d\Xi^*(x,y) \\
&= \int\limits_{x,y} d(x,y) d\left\{ \int\limits_{x',y'} \xi_{x'y'} d\Xi(x', y') \right\} \\
&= \int\limits_{x,y,x',y'} d(x,y) d\Xi(x',y') d\xi_{x'y'}(x,y) \\
&= \int\limits_{x',y'} \left(\int\limits_{x,y} d(x,y) d\xi_{x'y'} \right) d\Xi(x',y') \\
&\leqslant \int\limits_{x',y'} d(x',y')(1 - \kappa(x',y')) d\Xi(x',y') \\
&\leqslant (1-\kappa) \int\limits_{x',y'} d(x',y') d\Xi(x',y') \\
&= (1-\kappa) W_1(\mu, \mu').
\end{aligned}$$

Τέλος αρκεί να δείξουμε ότι $\mu * m \in \mathscr{P}(X)$ στην περίπτωση που $\mu \in \mathscr{P}(X)$:

Έστω $o \in X$. Τότε,

$$W_1(\delta_o, \mu * m) \leqslant W_1(\delta_o, m_o) + W_1(m_o, \mu * m) \leqslant W_1(\delta_o, m_o) + (1-\kappa)W_1(\delta_o, \mu) < \infty$$

καθώς $W_1(\delta_o, \mu), W_1(\delta_o, m_o) < \infty$. □

Ορισμός 2.3.2. Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο. Ορίζουμε ως **άλμα** του τυχαίου περίπατου στο σημείο $x \in X$ την ποσότητα:

$$J(x) := \mathbb{E}_{m_x}(d(x, \cdot)) = W_1(\delta_x, m_x).$$

Πόρισμα 2.3.3. *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και $\kappa(x, y) \geqslant \kappa > 0$ για κάθε $x, y \in X$ με $x \neq y$. Τότε ο τυχαίος περίπατος m έχει μοναδική αναλλοίωτη κατανομή $\nu \in \mathscr{P}(X)$.*

Επιπλέον, για κάθε μέτρο πιθανότητας $\mu \in \mathscr{P}(X)$, η ακολουθία $\{\mu * m^{*n}\}_{n \in \mathbb{N}}$ τείνει εκθετικά γρήγορα στο ν ως προς την W_1 απόσταση:

$$W_1(\mu * m^{*n}, \nu) \leqslant (1-\kappa)^n W_1(\mu, \nu)$$

και ειδικότερα

$$W_1(m_x^{*n}, \nu) \leqslant (1-\kappa)^n J(x)/\kappa.$$

Απόδειξη. Γνωρίζουμε ότι αν ο (X, d) είναι πλήρης χώρος τότε ο $(\mathscr{P}(X), W_1)$ είναι και αυτός πλήρης. Επίσης η σχέση $\kappa > 0$ συνεπάγεται ότι $1-\kappa < 1$. Ορίζουμε συνάρτηση $f : \mathscr{P}(X) \to \mathscr{P}(X)$ με $f(\mu) = \mu * m$. Λόγω της προηγούμενης πρότασης έχουμε ότι $\mu * m \in \mathscr{P}(X)$, συνεπώς η f είναι καλά ορισμένη και επιπλέον ισχύει η σχέση

$$W_1(f(\mu), f(\mu')) \leqslant (1-\kappa)W_1(\mu, \mu')$$

πάλι από την προηγούμενη πρόταση. Συνεπώς η f είναι γνήσια συστολή σε πλήρη χώρο και άρα από το θεώρημα σταθερού σημείου του Banach η f έχει μοναδικό σταθερό σημείο. Δηλαδή υπάρχει $\nu \in \mathscr{P}(X)$ με $\nu * m = \nu$ και αυτή είναι η αναλλοίωτη κατανομή που ψάχνουμε.

Η ν είναι αναλλοίωτη κατανομή για τον m συνεπώς ισχύει $\nu * m = \nu$. Από την προηγούμενη πρόταση έχουμε ότι

$$W_1(\mu * m, \nu) = W_1(\mu * m, \nu * m) \leqslant (1-\kappa)W_1(\mu, \nu)$$

και συνεπώς

$$W_1(\mu * m^{*2}, \nu) = W_1(\mu * (m * m), \nu) \leqslant (1-\kappa)W_1(\mu * m, \nu) \leqslant (1-\kappa)^2 W_1(\mu, \nu).$$

Επαγωγικά ισχύει η ζητούμενη ανισότητα.

Επίσης για $\mu = \delta_x$ έχουμε ότι

$$W_1(\delta_x, \nu) \leqslant W_1(\delta_x, m_x) + W_1(m_x, \nu) \leqslant J(x) + (1-\kappa)W_1(\delta_x, \nu),$$

το οποίο σημαίνει ότι

$$W_1(\delta_x, \nu) \leqslant J(x)/\kappa,$$

και επαγωγικά έχουμε το ζητούμενο. □

Πόρισμα 2.3.4. *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και $\kappa(x, y) \geqslant \kappa > 0$ για κάθε $x, y \in X$ με $x \neq y$. Έστω ν η αναλλοίωτη κατανομή του m. Τότε για κάθε $f : X \to \mathbb{R}$ 1-Lipschitz συνάρτηση και για κάθε κατανομή μ ισχύει ότι*

$$|\mathbb{E}_\nu(f) - \mathbb{E}_\mu(f)| \leqslant \frac{W_1(\mu, \mu * m)}{\kappa}.$$

Ειδικότερα έχουμε ότι για κάθε $x \in X$ ισχύει

$$|f(x) - \mathbb{E}_\nu(f)| \leqslant \frac{J(x)}{\kappa}.$$

Απόδειξη. Έχουμε ότι

$$W_1(\mu * m, \nu) \leqslant (1 - \kappa) W_1(\mu, \nu)$$

και από την τριγωνική ανισότητα παίρνουμε ότι

$$W_1(\mu * m, \nu) \geqslant W_1(\mu, \nu) - W_1(\mu * m, \mu).$$

Άρα,

$$W_1(\mu, \nu) \leqslant \frac{W_1(\mu * m, \mu)}{\kappa}.$$

Αν τώρα η f είναι 1-Lipschitz από το Θεώρημα Δυϊσμού του Kantorovich έχουμε ότι

$$|\mathbb{E}_\nu(f) - \mathbb{E}_\mu(f)| \leqslant W_1(\mu, \nu)$$

και μαζί με την προηγούμενη ανισότητα έχουμε το ζητούμενο.

Τέλος, για $\mu = \delta_x$ έχουμε

$$|\mathbb{E}_{\delta_x}(f) - \mathbb{E}_\nu(f)| \leqslant \frac{W_1(m_x, \delta_x)}{\kappa}$$

και συνεπώς

$$|f(x) - \mathbb{E}_\nu(f)| \leqslant \frac{J(x)}{\kappa}.$$

□

2.4 Θεωρήματα Bonnet-Myers στον L^1

Σε αυτήν την παράγραφο θα δοθεί ένα λιγότερο ισχυρό ανάλογο του θεωρήματος Bonnet-Myers, το οποίο, ειδικότερα, δείχνει ότι η θετική coarse καμπυλότητα Ricci είναι πολύ ισχυρότερη ιδιότητα από το κάτω φράγμα για το φασματικό κενό.

Πρόταση 2.4.1 (L^1 Bonnet-Myers). *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και $\kappa(x, y) \geqslant \kappa > 0$ για κάθε $x, y \in X$ με $x \neq y$. Τότε για κάθε $x, y \in X$ ισχύει ότι*

$$d(x, y) \leqslant \frac{J(x) + J(y)}{\kappa(x, y)},$$

και ειδικότερα

$$\operatorname{diam}(X) \leqslant \frac{2 \sup\{J(x) : x \in X\}}{\kappa}.$$

Απόδειξη. Έχουμε ότι

$$d(x,y) = W_1(\delta_x,\delta_y) \leqslant W_1(\delta_x,m_x) + W_1(m_x,m_y) + W_1(m_y,\delta_y) \leqslant J(x) + (1-\kappa)d(x,y) + J(y)$$

και συνεπώς ισχύει το ζητούμενο. □

Παρατήρηση 2.4.2. Η εκτίμηση της παραπάνω πρότασης δεν είναι η καλύτερη δυνατή για όλες τις κινήσεις Brown πάνω σε πολλαπλότητες Riemann (όπως θα δούμε παρακάτω $J \approx \varepsilon$ και $\kappa \approx \varepsilon^2 \mathrm{Ric}/N$, συνεπώς η εκτίμηση υστερεί κατά έναν όρο $1/\varepsilon$ σε σχέση με το θεώρημα Bonnet-Myers). Παρόλα αυτά, σε αρκετά άλλα παραδείγματα (όπως στον διακριτό κύβο, στην διακριτή και συνεχή ανέλιξη Ornstein-Uhlenbeck) επιτυγχάνουμε την εκτίμηση του Θεωρήματος Bonnet-Myers.

Πρόταση 2.4.3. *Έστω (X,d,m) μετρικός χώρος με έναν τυχαίο περίπατο και $\kappa(x,y) \geqslant \kappa > 0$ για κάθε $x,y \in X$ με $x \neq y$ και έστω ν η αναλλοίωτη κατανομή του m. Τότε για κάθε $x \in X$ έχουμε ότι*

$$\int_X d(x,y)d\nu(y) \leqslant \frac{J(x)}{\kappa},$$

συνεπώς

$$\int_{X\times X} d(x,y)d\nu(x)d\nu(y) \leqslant \frac{2\inf_x J(x)}{\kappa}.$$

Απόδειξη. Η πρώτη σχέση προκύπτει από το Πόρισμα 2.3.4 για την $f = d(x,\cdot)$ (η συνάρτηση απόστασης είναι 1-Lipschitz).

Για την δεύτερη, για κάθε $\varepsilon > 0$ επιλέγουμε $x_0 \in X$ με $J(x_0) < \inf_x J(x) + \varepsilon$ και έχουμε

$$\begin{aligned}\int_{X\times X} d(x,y)d\nu(y)d\nu(y) &\leqslant \int_{X\times X} \left(d(x,x_0) + d(x_0,y)\right) d\nu(x)d\nu(y)\\ &= 2W_1(\delta_{x_0},\nu) \leqslant 2\frac{J(x_0)}{\kappa} < \frac{2\inf_x J(x)+\varepsilon}{\kappa},\end{aligned}$$

άρα ισχύει το ζητούμενο. □

2.5 Τρεις κατασκευές

Η πρώτη κατασκευή που θα δώσουμε περιγράφει τη σύνθεση δύο τυχαίων περιπάτων.

Πρόταση 2.5.1 (σύνθεση)**.** *Έστω (X,d) μετρικός χώρος εφοδιασμένος με δύο τυχαίους περίπατους $m = (m_x)_{x\in X}$ και $m' = (m'_x)_{x\in X}$. Υποθέτουμε ότι η coarse καμπυλότητα Ricci του m (αντίστοιχα του m') είναι τουλάχιστον κ (αντίστοιχα κ'). Έστω m'' η σύνθεση των m και m', δηλαδή ο τυχαίος περίπατος που απεικονίζει το μέτρο πιθανότητας μ στο $\mu * m * m'$. Τότε η coarse καμπυλότητα Ricci του m'' είναι τουλάχιστον $\kappa + \kappa' - \kappa\kappa'$.*

Απόδειξη. Έστω $x,y \in X$. Έχουμε ότι

$$\begin{aligned}W_1(m''_x, m''_y) &= W_1(\delta_x * m'', \delta_y * m'')\\ &= W_1(\delta_x * m * m', \delta_y * m * m')\\ &\leqslant W_1(\delta_x * m, \delta_y, *m)(1-k')\\ &\leqslant W_1(\delta_x,\delta_y)(1-k)(1-k')\\ &= d(x,y)(1-k)(1-k')\end{aligned}$$

και συνεπώς

$$\kappa''(x,y) = 1 - \frac{W_1(m''_x, m''_y)}{d(x,y)} \geqslant 1 - (1-\kappa)(1-\kappa') = \kappa + \kappa' - \kappa\kappa'.$$

□

Η δεύτερη κατασκευή λέγεται *υπέρθεση* (superposition) και περιγράφει τον εξής τυχαίο περίπατο: Αν μας έχουν δοθεί δύο τυχαίοι περίπατοι, κατασκευάζουμε την υπέρθεση τους επιλέγοντας σε κάθε βήμα με ρίψη ενός νομίσματος ποιον από τους δύο θα ακολουθήσουμε.

Πρόταση 2.5.2 (υπέρθεση)**.** *Έστω (X,d) μετρικός χώρος εφοδιασμένος με μια οικογένεια τυχαίων περιπάτων $(m^i)_{i\in I}$. Υποθέτουμε ότι για κάθε $i \in I$ η coarse καμπυλότητα Ricci του m^i είναι τουλάχιστον κ_i. Έστω $(\alpha_i)_{i\in I}$ μια οικογένεια μη-αρνητικών πραγματικών αριθμών με $\sum \alpha_i = 1$. Ορίζουμε τυχαίο περίπατο m στον X με $m_x := \sum \alpha_i m^i$. Τότε η coarse καμπυλότητα Ricci του m είναι τουλάχιστον $\sum \alpha_i \kappa_i$.*

Απόδειξη. Έστω $x, y \in X$. Για κάθε $i \in I$ θεωρούμε ένα ταίριασμα ξ_i μεταξύ των μέτρων m^i_x, m^i_y. Τότε το $\sum \alpha_i \xi_i$ είναι ταίριασμα μεταξύ των μέτρων $\sum \alpha_i m^i_x$ και $\sum \alpha_i m^i_y$. Συνεπώς ισχύει ότι

$$\begin{aligned} W_1(m_x, m_y) &\leqslant \sum_{i\in I} \alpha_i W_1(m^i_x, m^i_y) \\ &= \sum_{i\in I} \alpha_i W_1(\delta_x * m^i, \delta_y * m^i) \\ &\leqslant \sum_{i\in I} (1-\kappa_i) d(x,y) \\ &= \left(1 - \sum_{i\in I} \alpha_i \kappa_i\right) d(x,y), \end{aligned}$$

και συνεπώς, αν $\kappa(x,y)$ είναι η coarse καμπυλότητα Ricci του m έχουμε ότι

$$\kappa(x,y) \geqslant \sum_{i\in I} \alpha_i \kappa_i.$$

□

Πρόταση 2.5.3 (L^1 tensorization)**.** *Έστω $(X_1,d_1), (X_2,d_2), \dots, (X_k,d_k)$ μετρικοί χώροι και έστω οτι κάθε ένας από τους X_i είναι εφοδιασμένος με έναν τυχαίο περίπατο m^i, $i = 1,2,\dots,k$. Έστω $X = X_1 \times X_2 \times \ldots \times X_k$ ο χώρος γινόμενο των X_i εφοδιασμένος με τη μετρική $d := \sum d_i$. Έστω επίσης $\alpha_1, \alpha_2, \dots, \alpha_k$ μια οικογένεια μη-αρνητικών πραγματικών αριθμών με $\sum \alpha_i = 1$. Θεωρούμε τον τυχαίο περίπατο m στον X που ορίζεται ως εξής:*

$$m_{(x_1,x_2,\dots,x_k)} = \sum_{i=1}^k \alpha_i \delta_{x_1} \otimes \delta_{x_2} \otimes \ldots \otimes m_{x_i} \otimes \ldots \otimes \delta_k.$$

Υποθέτουμε ότι για κάθε $i = 1,2,\dots,k$ η coarse καμπυλότητα Ricci του m^i είναι τουλάχιστον κ_i. Τότε η coarse καμπυλότητα Ricci του m είναι τουλάχιστον $\min \alpha_i \kappa_i$.

Απόδειξη. Για κάθε $x = (x_1, x_2, \ldots, x_k) \in X$ ορίζουμε

$$\widetilde{m}_x^i := \delta_{x_1} \otimes \ldots \otimes m_{x_i} \otimes \ldots \otimes \delta_{x_k},$$

και συνεπώς έχουμε ότι

$$m_x = \sum_{i=1}^{k} \alpha_i \widetilde{m}_x^i.$$

Έστω $x = (x_1, x_2, \ldots, x_k),\ y = (y_1, y_2, \ldots, y_k) \in X$. Τότε έχουμε ότι

$$\begin{aligned}
W_1(m_x, m_y) &= W_1\left(\sum_{i=1}^{k} \alpha_i \widetilde{m}_x^i, \sum_{i=1}^{k} \alpha_i \widetilde{m}_y^i\right) \\
&\leqslant \sum_{i=1}^{k} \alpha_i W_1\left(\widetilde{m}_x^i, \widetilde{m}_y^i\right) \\
&\leqslant \sum_{i=1}^{k} \alpha_i \left(W_1\left(m_x^i, m_y^i\right) + \sum_{j \neq i} d_j(x_j, y_j)\right) \\
&\leqslant \sum_{i=1}^{k} \alpha_i \left((1-\kappa_i) d_i(x_i, y_i) + \sum_{j \neq i} d_j(x_j, y_j)\right) \\
&= \sum_{i=1}^{k} \alpha_i \left(-\kappa_i d_i(x_i, y_i) + \sum_{j=1}^{k} d_i(x_j, y_j)\right) \\
&= \sum_{i=1}^{k} d_i(x_i, y_i) - \sum_{i=1}^{k} \alpha_i \kappa_i d_i(x_i, y_i) \\
&\leqslant (1 - \min \alpha_i \kappa_i) \sum_{i=1}^{k} d_i(x_i, y_i) \\
&\leqslant (1 - \min a_i \kappa_i) d(x, y),
\end{aligned}$$

και συνεπώς ισχύει το ζητούμενο. □

2.6 Συναρτήσεις Lipschitz και φασματικό κενό

Ορισμός 2.6.1 (τελεστής τυχαίου περίπατου, Λαπλασιανός τελεστής)**.** Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και έστω ν η αναλλοίωτη κατανομή του m. Για κάθε $f \in L^2(X, \nu)$ ορίζουμε τον **τελεστή τυχαίου περίπατου** ως τον τελεστή M με

$$Mf(x) := \int_{y \in X} f(y) dm_x(y)$$

Επίσης ορίζουμε τον **Λαπλασιανό τελεστή** ως τον $\Delta := M - Id$.

Πρόταση 2.6.2 (συστολή Lipschitz)**.** *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και έστω $\kappa \in \mathbb{R}$. Τότε η coarse καμπυλότητα Ricci του m είναι τουλάχιστον κ αν και μόνο αν για κάθε ρ-Lipschitz συνάρτηση $f : X \to \mathbb{R}$ η συνάρτηση Mf είναι $\rho(1-\kappa)$-Lipschitz.*

Απόδειξη. Έστω $x, y \in X$ και έστω $\kappa(x,y)$ η coarse καμπυλότητα Ricci του m.

Υποθέτουμε αρχικά ότι $\kappa(x,y) \geqslant \kappa$. Αν ξ είναι το βέλτιστο ταίριασμα των μέτρων m_x και m_y τότε έχουμε ότι

$$\begin{aligned}
|Mf(y) - Mf(x)| &= \left| \int_z f(z) dm_y(z) - \int_w f(w) dm_x(w) \right| \\
&= \left| \int_{z,w} (f(z) - f(w))\, d\xi(z,w) \right| \\
&\leqslant \int_{z,w} |f(z) - f(w)|\, d\xi(z,w) \\
&\leqslant \rho \int_{z,w} d(z,w) d\xi(z,w) \\
&= \rho W_1(m_x, m_y) \\
&= \rho d(x,y)(1 - \kappa(x,y)) \\
&\leqslant \rho(1-\kappa) d(x,y),
\end{aligned}$$

και συνεπώς η Mf είναι $\rho(1-\kappa)$-Lipschitz.

Αντίστροφα, αν η Mf είναι $\rho(1-\kappa)$-Lipschitz για κάθε f ρ-Lipschitz, τότε το Θεώρημα Δυϊσμού του Kantorovich μας δίνει ότι

$$\begin{aligned}
W_1(m_x, m_y) &= \sup_{f \ 1-\text{Lipschitz}} \int f d(m_x - m_y) \\
&= \sup_{f \ 1-\text{Lipschitz}} \{Mf(x) - Mf(y)\} \\
&\leqslant (1-\kappa) d(x,y).
\end{aligned}$$

□

Έστω ν μια αναλλοίωτη κατανομή του τυχαίου περίπατου m. Θεωρούμε την σχέση ισοδυναμίας $\sim$ στον $L^2(X,\nu)$ με $f \sim g$ αν και μόνο αν $f = g + c$ για κάποια σταθερά c. Επίσης θεωρούμε τον χώρο πηλίκο $L^2(X,\nu)/\sim$ εφοδιασμένο με τη νόρμα

$$\|f\|^2_{L^2(X,\nu)/\sim} := \|f - \mathbb{E}_\nu(f)\|^2_{L^2(X,\nu)} = \mathrm{Var}_\nu f = \frac{1}{2} \int_{X \times X} (f(x) - f(y))^2 \, d\nu(x) d\nu(y).$$

Πρόταση 2.6.3. *Ισχύουν τα εξής:*

(i) *Οι τελεστές M και Δ είναι αυτοσυζυγείς αν και μόνο αν η ν είναι αντιστρέψιμη για τον τυχαίο περίπατο m.*

(ii) *Ισχύει η σχέση*
$$\mathrm{Var}_\nu f = \int \mathrm{Var}_{m_x} f d\nu(x) + \mathrm{Var}_\nu Mf,$$
και συνεπώς $\|Mf\|_2 \leqslant \|f\|_2$.

(iii) *Ισχύει ότι $\|Mf\|_\infty \leqslant \|f\|_\infty$.*

Απόδειξη. (i) Έχουμε ότι ο M είναι αυτοσυζυγής αν για κάθε $g \in L^2(X,\nu)$ ισχύει

$$\langle Mf, g\rangle_{L^2(X,\nu)/\sim} = \langle g, Mf\rangle_{L^2(X,\nu)/\sim}.$$

Αν η ν είναι αντιστρέψιμη τότε ισχύει ότι $d\nu(x)dm_x(y) = d\nu(y)dm_y(x)$, άρα

$$\begin{aligned}
\langle Mf, g\rangle_{L^2(X,\nu)/\sim} &= \int_x Mf(x)g(x)d\nu(x) \\
&= \int_x \left(\int_y f(y)dm_x(y)\right) g(x)d\nu(x) \\
&= \int_y \int_x f(y)g(x)dm_y(x)d\nu(y) \\
&= \int_y \left(\int_x g(x)dm_y(x)\right) f(y)d\nu(y) \\
&= \int_y Mg(y)f(y)d\nu(y) \\
&= \langle f, Mg\rangle_{L^2(X,\nu)/\sim}.
\end{aligned}$$

Συνεπώς, ο M είναι αυτοσυζυγής.

Αντίστροφα, αν ο M είναι αυτοσυζυγής, τότε έχουμε το ζητούμενο για $f = g = Id$.

(ii) Έχουμε ότι:

$$\mathbb{E}_\nu(Mf) = \int_x Mf(x)d\nu(x) = \int_x \int_y f(y)dm_x(y)d\nu(x) \overset{\nu \text{ αναλ.}}{=} \int_y f(y)d\nu(y) = \mathbb{E}_\nu(f),$$

και επιπλέον

$$\mathbb{E}_{m_x}(f) = \int_y f(y)dm_x(y) = Mf(x).$$

Άρα έχουμε ότι

$$\begin{aligned}
\int_x \mathrm{Var}_{m_x}(f)d\nu(x) + \mathrm{Var}_\nu(Mf) &= \int_x \left(\mathbb{E}_{m_x}(f)^2 - \mathbb{E}_{m_x}(f^2)\right) d\nu(x) + \mathbb{E}_\nu(Mf)^2 - \mathbb{E}_\nu(Mf^2) \\
&= \int_x \int_y f^2(y)dm_x(y)d\nu(y) - \int_x (Mf(x))^2 d\nu(x) \\
&\quad + \mathbb{E}_\nu(f)^2 - \mathbb{E}_\nu(Mf^2) \\
&= \int_y f(y)^2 d\nu(y) - \mathbb{E}_\nu(Mf^2) + \mathbb{E}_\nu(f)^2 - \mathbb{E}_\nu(Mf^2) \\
&= \mathbb{E}_\nu(f^2) - \mathbb{E}_\nu(f)^2 \\
&= \mathrm{Var}_\nu(f).
\end{aligned}$$

(iii) Αυτό είναι προφανές καθώς για κάθε $x \in X$ έχουμε ότι

$$\begin{aligned} |Mf(x)| &= \left| \int_y f(y) dm_x(y) \right| \\ &\leqslant \int_y |f(y)| dm_x(y) \\ &\leqslant \|f\|_\infty \int_y dm_x(y) \\ &= \|f\|_\infty . \end{aligned}$$

□

Η παρακάτω πρόταση δείχνει ότι η φασματική ακτίνα (και συνεπώς το φασματικό κενό) του τελεστή M στον χώρο $L^2(X,\nu)/\sim$ είναι το πολύ $1-\kappa$. Ο χώρος των συναρτήσεων Lipschitz είναι πυκνός υπόχωρος του $L^2(X,\nu)$ και για να ισχύει το ίδιο φράγμα για τον M σε όλο τον χώρο θα πρέπει είτε ο τελεστής να είναι αυτοσυζυγής ή ο χώρος να είναι πεπερασμένης διάστασης.

Υπενθύμιση: Το **φασματικό κενό** ενός τελεστή είναι η τιμή της μικρότερης μη-μηδενικής ιδιοτιμής του τελεστή.

Η **φασματική ακτίνα** του τελεστή είναι η ποσότητα $\max|\lambda_i|$ όπου λ_i οι ιδιοτιμές του τελεστή. Επίσης ισχύει ο φασματικός τύπος του Gelfand:

$$\lim_{t\to\infty} \|T\|^{1/t} = \max|\lambda_i|.$$

Ορισμός 2.6.4. Έστω (X,d) μετρικός χώρος με έναν τυχαίο περίπατο m. Ορίζουμε ως **(coarse) σταθερά διάχυσης** του τυχαίου περίπατου στο $x \in X$ την ποσότητα

$$\sigma(x) := \left(\frac{1}{2} \int\int d(y,z)^2 dm_x(y) dm_x(z) \right)^{1/2},$$

και αν ν είναι μια αναλλοίωτη κατανομή του m ορίζουμε ως **μέση σταθερά διάχυσης** την ποσότητα

$$\sigma := \|\sigma(x)\|_{L^2(X,\nu)}.$$

Επιπλέον, έστω $\sigma_\infty(x) := \frac{1}{2}\mathrm{diam}(\mathrm{supp}(m_x))$ και $\sigma_\infty := \sup_x \sigma_\infty(x)$. Ορίζουμε ως **τοπική διάσταση** στο x την ποσότητα

$$n_x := \frac{\sigma(x)^2}{\sup\{\mathrm{Var}_{m_x} f : \ f : \mathrm{supp}(m_x) \to \mathbb{R} \ 1-\mathrm{Lipschitz}\}},$$

και τέλος θέτουμε $n := \inf_x n_x$.

Πρόταση 2.6.5 (διασπορά Lipschitz συναρτήσεων)**.** *Έστω (X,d,m) τυχαίος περίπατος σε έναν μετρικό χώρο με coarse καμπυλότητα Ricci τουλάχιστον $\kappa > 0$. Έστω ν η μοναδική αναλλοίωτη κατανομή. Υποθέτουμε ότι $\sigma < \infty$. Τότε, η διασπορά μιας 1-Lipschitz συνάρτησης είναι το πολύ $\frac{\sigma^2}{n\kappa(2-\kappa)}$.*

Παρατηρήστε ότι, αφού $\kappa \leqslant 1$, ισχύει η ανισότητα

$$\frac{\sigma^2}{n\kappa(2-\kappa)} \leqslant \frac{\sigma^2}{n\kappa}.$$

Έπεται ότι όλες οι Lipschitz συναρτήσεις ανήκουν στον $L^2/\sim$. Ειδικότερα, έχουμε

$$\int d(x,y)^2 d\nu(x)d\nu(y) < \infty.$$

Το γεγονός ότι η Lipschitz νόρμα φράσσει την L^2-νόρμα χρησιμοποιήθηκε παραπάνω στη συζήτηση των φασματικών ιδιοτήτων του τελεστή του τυχαίου περίπατου.

Η υπόθεση ότι $\sigma < \infty$ είναι απαραίτητη. Ένα αντιπαράδειγμα μας δίνει ο τυχαίος περίπατος στο $\mathbb{N}$ που στέλνει κάθε $x \in \mathbb{N}$ σε κάποια σταθερή κατανομή ν στο $\mathbb{N}$ που έχει άπειρη δεύτερη ροπή. Η coarse καμπυλότητα Ricci είναι ίση με 1, όμως η ταυτοτική συνάρτηση δεν ανήκει στον L^2.

Απόδειξη. Υποθέτουμε αρχικά ότι η $|f|$ είναι φραγμένη από κάποιον $A \in \mathbb{R}$, οπότε $\mathrm{Var}_\nu(f) < \infty$. Θα δείξουμε ότι $\mathrm{Var}_\nu(M^t f) \to 0$. Έστω B_r η μπάλα ακτίνας r στον X με κέντρο κάποιο σημείο βάσης. Έχουμε δει ότι αν μια g είναι $\rho-$Lipschitz τότε η Mg είναι $\rho(1-\kappa)$-Lipschitz και επομένως η $M^t g$ είναι $\rho^t(1-\kappa)^t$-Lipschitz. Χρησιμοποιώντας αυτό το γεγονός έχουμε ότι η $M^t f$ είναι $(1-\kappa)^t$-Lipschitz στην B_r και φραγμένη από A στο $X \setminus B_r$. Έτσι βλέπουμε ότι

$$\begin{aligned}
\mathrm{Var}_\nu(M^t f) &= \frac{1}{2}\iint (M^t f(x) - M^t f(y))^2 d\nu(x)d\nu(y) \\
&= \frac{1}{2}\iint_{B_r \times B_r} (M^t f(x) - M^t f(y))^2 d\nu(x)d\nu(y) + \\
&\quad + \frac{1}{2}\iint_{B_r^c \times B_r^c} (M^t f(x) - M^t f(y))^2 d\nu(x)d\nu(y) \\
&\leqslant \frac{1}{2}\iint_{B_r \times B_r} 4(1-\kappa)^{2t} d(x,y)^2 d\nu(x)d\nu(y) + \\
&\quad + \frac{1}{2}\iint_{B_r^c \times B_r^c} 4A^2 d\nu(x)d\nu(y) \\
&\leqslant 2(1-\kappa)^{2t} r^2 + 2A^2\nu^2(A \setminus B_r).
\end{aligned}$$

Επιλέγοντας $r = 1/(1-\kappa)^{t/2}$ (ώστε να μηδενίζονται και οι δύο προσθετέοι όταν το r τείνει στο άπειρο) συμπεραίνουμε ότι $\mathrm{Var}_\nu(M^t f) \to 0$.

Γνωρίζουμε ότι $\mathrm{Var}_\nu(f) = \mathrm{Var}_\nu(Mf) + \int \mathrm{Var}_{m_x}(f) d\nu(x)$ και συνεπώς βάζοντας στη θέση της f την Mf επαγωγικά έχουμε, για κάθε $N \in \mathbb{N}$, ότι

$$\mathrm{Var}_\nu(f) = \sum_{t=0}^{N} \int \mathrm{Var}_{m_x}(M^t f)\, d\nu(x) + \mathrm{Var}_\nu M^N f,$$

άρα, αν αφήσουμε το N να πάει στο άπειρο έχουμε ότι

$$\mathrm{Var}_\nu(f) = \sum_{t=0}^{\infty} \int \mathrm{Var}_{m_x}(M^t f)\, d\nu(x).$$

Αφού η f είναι 1-Lipschitz, από τον ορισμό της τοπικής διάστασης n_x έχουμε $\mathrm{Var}_{m_x}(f) \leqslant \sigma(x)^2/n_x$. Αφού η $M^t f$ είναι $(1-\kappa)^t$-Lipschitz, έχουμε $\mathrm{Var}_{m_x}(M^t f) \leqslant (1-\kappa)^{2t}\sigma(x)^2/n_x$, άρα έχουμε ότι

$$\mathrm{Var}_\nu(f) = \sum_{t=0}^{\infty} \int \mathrm{Var}_{m_x}(M^t f)\, d\nu(x) \leqslant \frac{\sigma^2}{n\kappa(2-\kappa)}.$$

Για την περίπτωση που η f δεν είναι φραγμένη χρησιμοποιούμε ένα απλό οριακό επιχείρημα. □

Πρόταση 2.6.6 (φασματικό κενό). *Έστω (X, d, m) μετρικός χώρος με έναν τυχαίο περίπατο και έστω $\kappa(x, y) \geq \kappa > 0$ για κάθε $x, y \in X$. Έστω επίσης ν η μοναδική αναλλοίωτη κατανομή του m και ότι $\sigma < \infty$. Υποθέτουμε επίσης ότι ν αντιστρέψιμος ή ότι ο X είναι πεπερασμένος. Τότε η φασματική ακτίνα του τελεστή μέσου όρου στον χώρο $L^2(X,\nu)/\sim$ είναι το πολύ $1-\kappa$.*

Απόδειξη. Στην περίπτωση που ο χώρος X είναι πεπερασμένος έχουμε ότι οι Lipschitz συναρτήσεις συμπίπτουν με τις L^2 συναρτήσεις, και οι νόρμες $\|\cdot\|_{L^2(X\nu)}$ και $\|\cdot\|_{L^2(X\nu)/\sim}$ είναι ισοδύναμες. Συνεπώς η πρόταση ισχύει.

Αν τώρα η ν είναι αντιστρέψιμη, έχουμε ότι ο M είναι αυτοσυζυγής και αν η $f : X \to \mathbb{R}$ είναι ρ-Lipschitz τότε, όπως δείχνουμε παρακάτω, η f ανήκει στον $L^2(X\nu)/\sim$ και ισχύει ότι

$$\|f\|^2_{L^2(X\nu)/\sim} = \mathrm{Var}_\nu f \leqslant \rho^2\sigma^2/n\kappa(2-\kappa) < \infty$$

καθώς υποθέσαμε ότι $\sigma < \infty$.

Επίσης έχουμε δει ότι η $M^t f$ είναι $\rho(1-\kappa)^t$-Lipschitz και συνεπώς από την Πρόταση 2.6.5 έχουμε ότι

$$\mathrm{Var}_\nu M^t f \leqslant C\rho^2(1-\kappa)^{2t},$$

όπου $C = \frac{\sigma^2}{n\kappa(2-\kappa)}$. Έπεται ότι

$$\lim_{t\to\infty}\left(\sqrt{\mathrm{Var}_\nu M^t f}\right)^{1/t} \leqslant 1-\kappa,$$

το οποίο δείχνει ότι η φασματική ακτίνα της M είναι το πολύ $1-\kappa$ στον υπόχωρο των συναρτήσεων Lipschitz.

Τέλος, οι συναρτήσεις Lipschitz είναι πυκνές στον $L^2(X,\nu)$ (αυτό προκύπτει από το γεγονός ότι τα μέτρα πιθανότητας σε μετρικούς χώρους είναι κανονικά, επομένως οι χαρακτηριστικές συναρτήσεις προσεγγίζονται από συναρτήσεις Lipschitz) και το γεγονός ότι ο M είναι αυτοσυζυγής δίνει το ζητούμενο φράγμα της φασματικής ακτίνας σε όλο τον χώρο $L^2(X,\nu)$. □

Πόρισμα 2.6.7 (ανισότητα Poincaré). *Έστω (X, d, m) μετρικός χώρος με έναν εργοδικό τυχαίο περίπατο και έστω $\kappa(x, y) \geq \kappa > 0$ για κάθε $x, y \in X$. Έστω επίσης ν η μοναδική αναλλοίωτη κατανομή του m και ότι $\sigma < \infty$. Υποθέτουμε επίσης ότι η ν είναι αντιστρέψιμη. Τότε το φάσμα του $-\Delta$ στον $L^2(X,\nu)/\sim$ περιέχεται στο $[\kappa,\infty)$. Επιπλέον ισχύουν οι παρακάτω διακριτές ανισότητες Poincaré για κάθε $f \in L^2(X,\nu)$:*

$$\mathrm{Var}_\nu f \leqslant \frac{1}{\kappa(2-\kappa)} \int \mathrm{Var}_{m_x} f d\nu(x)$$

και

$$\mathrm{Var}_\nu f \leqslant \frac{1}{2\kappa} \int\int (f(x)-f(y))^2\, d\nu(x)d\nu(y).$$

Απόδειξη. Οι δύο ανισότητες είναι ισοδύναμες εκφράσεις των $\mathrm{Var}_\nu Mf \leqslant (1-\kappa)^2 \mathrm{Var}_\nu f$ και $\langle f, Mf\rangle_{L^2(X\nu)/\{\sigma\tau\alpha\vartheta.\}} \leqslant (1-\kappa)\mathrm{Var}_\nu f$ αντίστοιχα: Για την πρώτη ανισότητα έχουμε ότι $Mf \leqslant (1-\kappa)f$ καθώς το $1-\kappa$ είναι φράγμα της φασματικής ακτίνας του M και επομένως

$$\mathrm{Var}_\nu(Mf) \leqslant \mathrm{Var}_\nu((1-\kappa)f) = (1-\kappa)^2 \mathrm{Var}_\nu(f).$$

Άρα, εφόσον

$$\mathrm{Var}_\nu(f) = \int \mathrm{Var}_{m_x}(f) d\nu + \mathrm{Var}_\nu Mf,$$

η ανισότητα γίνεται

$$\mathrm{Var}_\nu(f) - \int \mathrm{Var}_{m_x}(f) d\nu \leqslant (1-\kappa)^2 \mathrm{Var}_\nu(f),$$

και συνεπώς

$$\mathrm{Var}_\nu(f) \leqslant \frac{1}{\kappa(2-\kappa)} \int \mathrm{Var}_{m_x} f d\nu.$$

Για την δεύτερη ανισότητα, πάλι επειδή το $1-\kappa$ είναι φράγμα της φασματικής ακτίνας του M έχουμε ότι

$$\langle f, Mf\rangle_{L^2/\sim} \leqslant \langle f, (1-\kappa)f\rangle_{L^2/\sim} = (1-\kappa)\|f\|^2_{L^2/\sim} = (1-\kappa)\mathrm{Var}_\nu(f),$$

και συνεπώς

$$\mathrm{Var}_\nu(f) \leqslant \frac{\mathrm{Var}_\nu(f) - \langle f, Mf\rangle_{L^2/\sim}}{\kappa} \leqslant \frac{\mathrm{Var}_\nu(f)}{\kappa} = \frac{1}{2\kappa} \iint (f(x) - f(y))^2 d\nu(x) dm_x(y).$$

□

2.7 Ανισότητες Gaussian συγκέντρωσης

Λήμμα 2.7.1. *Έστω $\varphi : X \to \mathbb{R}$ μια α-Lipschitz συνάρτηση με $\alpha \leqslant 1$. Υποθέτουμε ότι $\lambda \leqslant \frac{1}{3\sigma_\infty}$. Τότε, για κάθε $x \in X$,*

$$(Me^{\lambda\varphi})(x) \leqslant e^{\lambda M\varphi(x) + \lambda^2\alpha^2 \frac{\sigma(x)^2}{n_x}}.$$

Απόδειξη. Για κάθε ομαλή συνάρτηση g και κάθε πραγματική τυχαία μεταβλητή Y, παίρνοντας ανάπτυγμα Taylor γύρω από το $\mathbb{E}(Y)$ με υπόλοιπο Lagrange έχουμε

$$g(Y) = g(\mathbb{E}(Y)) + g'(\mathbb{E}(Y))(Y - \mathbb{E}(Y)) + g''(\xi)(Y - \mathbb{E}(Y))^2$$

για κάποιο $\xi \in \mathbb{R}$, και συνεπώς

$$g(Y) \leqslant g(\mathbb{E}(Y)) + g'(\mathbb{E}(Y))(Y - \mathbb{E}(Y)) + \sup(g'')(Y - \mathbb{E}(Y))^2.$$

Άρα έχουμε ότι

$$\begin{aligned}\mathbb{E}(g(Y)) &\leqslant g(\mathbb{E}(Y)) + g'(\mathbb{E}(Y))(\mathbb{E}(Y) - \mathbb{E}(Y)) + \sup(g'')(Y - \mathbb{E})^2 \\ &= g(\mathbb{E}(Y)) + \sup(g'')\mathrm{Var}(Y).\end{aligned}$$

Εφαρμόζοντας αυτήν την ανισότητα για την $g(Y) = e^{\lambda Y}$ παίρνουμε

$$(Me^{\lambda\varphi})(x) = \mathbb{E}_{m_x}(e^{\lambda\varphi}) \leqslant e^{\lambda M\varphi(x)} + \frac{\lambda^2}{2}\left(\sup_{\mathrm{supp}(m_x)} e^{\lambda\varphi}\right)\mathrm{Var}_{m_x}(\varphi)$$

και παρατηρούμε ότι, αφού $\mathrm{diam}(\mathrm{supp}(m_x)) \leqslant 2\sigma_\infty$ και η φ είναι 1-Lipschitz, ισχύει η

$$\sup_{\mathrm{supp}(m_x)} \varphi \leqslant \mathbb{E}_{m_x}(\varphi) + 2\sigma_\infty,$$

συνεπώς

$$(Me^{\lambda\varphi})(x) \leqslant e^{\lambda M\varphi(x)} + \frac{\lambda^2}{2} e^{\lambda M\varphi(x)+2\lambda\sigma_\infty}\mathrm{Var}_{m_x}(\varphi).$$

Τώρα, από τον ορισμό της n_x έχουμε ότι $\mathrm{Var}_{m_x}(\varphi) \leqslant \alpha^2\frac{\sigma(x)^2}{n_x}$ καθώς η φ είναι α-Lipschitz και συνεπώς η φ/ρ είναι 1-Lipschitz. Επιπλέον, αν $\lambda \leqslant \frac{1}{3\sigma_\infty}$ τότε $e^{2\lambda\sigma_\infty} \leqslant 2$, απ' όπου έπεται ότι

$$\begin{aligned}\left(Me^{\lambda\varphi}\right)(x) &\leqslant e^{\lambda M\varphi(x)} + \lambda^2 e^{\lambda M\varphi(x)}\frac{\alpha^2\sigma^2(x)}{n_x} \\ &= e^{\lambda M\varphi(x)}\left(1 + \frac{\lambda^2\alpha^2\sigma^2(x)}{n_x}\right) \\ &\leqslant e^{\lambda M\varphi(x)+\lambda^2\alpha^2\frac{\sigma^2(x)}{n_x}}.\end{aligned}$$

□

Θεώρημα 2.7.2 (Gaussian συγκέντρωση)**.** *Έστω (X, d, m) τυχαίος περίπατος σε έναν μετρικό χώρο με coarse καμπυλότητα Ricci τουλάχιστον $\kappa > 0$. Έστω ν η μοναδική αναλλοίωτη κατανομή.*

Θέτουμε

$$D_x^2 := \frac{\sigma(x)^2}{n_x\kappa}$$

και

$$D^2 := \mathbb{E}_\nu D_x^2.$$

Υποθέτουμε ότι η συνάρτηση $x \mapsto D_x^2$ είναι C-Lipschitz. Ορίζουμε

$$t_{\max} := \frac{8}{9}D^2\min\{3/2C, 1/\sigma_\infty\}.$$

Τότε, για κάθε 1-Lipschitz συνάρτηση f, αν $t \leqslant t_{\max}$ έχουμε

$$\nu(\{x : f(x) \geqslant t + \mathbb{E}_\nu f\}) \leqslant \exp\left\{-\frac{3t^2}{16D^2}\right\},$$

και αν $t \geqslant t_{\max}$ έχουμε

$$\nu(\{x : f(x) \geqslant t + \mathbb{E}_\nu f\}) \leqslant \exp\left\{-\frac{3t_{\max}}{16D^2}t\right\}.$$

Απόδειξη. Για την απόδειξη χρησιμοποιούμε μια παραλλαγή της μεθόδου των martingales για την απόδειξη Gaussian συγκέντρωσης.

Σημειώνουμε ότι η άμεση εφαρμογή της γνωστής μεθόδου των martingales θα έδινε

$$(Me^{\lambda\varphi})(x) \leqslant e^{\lambda M\varphi(x)+2\lambda^2\alpha^2\sigma_\infty^2},$$

μια εκτίμηση που είναι πολύ ασθενής για να δώσει λογικές εκτιμήσεις για τη διασπορά.

Επιστρέφουμε τώρα στην απόδειξη του Θεωρήματος 2.7.2. Έστω f μια 1-Lipschitz συνάρτηση και $\lambda \geqslant 0$. Ορίζουμε $f_0 := f$ και επαγωγικά

$$f_{N+1}(x) := Mf_N(x) + \lambda\frac{\sigma(x)^2}{n_x}\left(1-\frac{\kappa}{2}\right)^{2N}.$$

Ας υποθέσουμε ότι $\lambda \leqslant \frac{1}{2C}$. Τότε η $\lambda\frac{\sigma(x)^2}{n_x}$ είναι $\kappa/2$-Lipschitz, καθώς από την υπόθεση έχουμε ότι η D_x^2 είναι C-Lipschitz και άρα η $\frac{\sigma^2(x)}{n_x}$ είναι κC-Lispchitz. Συνεπώς,

$$\left|\lambda\frac{\sigma^2(x)}{n_x} - \lambda\frac{\sigma(y)^2}{n_y}\right| \leqslant \lambda C\kappa d(x,y) \leqslant \frac{\kappa}{2}d(x,y).$$

Χρησιμοποιώντας την Πρόταση 2.6.2 δείχνουμε ότι η f_N είναι $(1-\kappa/2)^N$-Lipschitz: Έχουμε ότι η Mf είναι $(1-\kappa)$-Lipschitz και συνεπώς η $f_1(x) = Mf(x) + \lambda\frac{\sigma^2(x)}{n_x}$ είναι $(1-\kappa)+\kappa/2$-Lipschitz, δηλαδή $(1-\kappa/2)$-Lipschitz. Επαγωγικά έχουμε το ζητούμενο.

Υποθέτουμε επιπλέον ότι ισχύει το $\lambda \leqslant 1/3\sigma_\infty$ και συνεπώς από το προηγούμενο λήμμα για $\alpha = (1-\kappa/2)^{2N}$ έχουμε ότι

$$(Me^{\lambda f_N})(x) \leqslant e^{\lambda Mf_N(x)+\lambda^2\frac{\sigma(x)^2}{n_x}(1-\kappa/2)^{2N}} = e^{\lambda f_{N+1}(x)},$$

οπότε, επαγωγικά,

$$(M^N e^{\lambda f})(x) \leqslant e^{\lambda f_N(x)}.$$

Θέτουμε $g(x) := \frac{\sigma(x)^2}{n_x}$ και χρησιμοποιώντας τον ορισμό της f_N παίρνουμε

$$f_N(x) = (M^N f)(x) + \lambda\sum_{i=1}^{N}(M^{N-i}g)(x)(1-\kappa/2)^{2(i-1)}.$$

Ισχύουν τώρα τα εξής:

- $(M^N f)(x) \to \mathbb{E}_\nu(f)$ για $N\to\infty$: Είδαμε προηγουμένως ότι $\mathrm{Var}_\nu M^N f \to 0$ για $N\to\infty$ και επομένως $(M^N f)(x) - \mathbb{E}_\nu(M^N f) \to 0$ με πιθανότητα 1. Επιπλέον $\mathbb{E}_\nu(Mf) = \mathbb{E}_\nu(f)$ (λόγω του αναλλοίωτου του μέτρου ν για τον τυχαίο περίπατο m), το οποίο δίνει ότι $\mathbb{E}_\nu(M^N f) = \mathbb{E}_\nu(f)$ για κάθε $N \in \mathbb{N}$.

- $\displaystyle\sum_{i=1}^{N}(M^{N-i}g)(x)(1-\kappa/2)^{2(i-1)} \to \sum_{i=1}^{\infty}\mathbb{E}_\nu(g)(1-\kappa/2)^{2(i-1)}$ για $N\to\infty$: Έχουμε ότι

$$\begin{aligned}\sum_{i=1}^{N}(M^{N-i}g)(x)(1-\kappa/2)^{2(i-1)} &= \sum_{i=1}^{N}\left((M^{N-i}g)(x) - \mathbb{E}_\nu(g)\right)(1-\kappa/2)^{2(i-1)} \\ &\quad + \sum_{i=1}^{N}\mathbb{E}_\nu(g)(1-\kappa/2)^{2(i-1)}\end{aligned}$$

και συνεπώς αρκεί να δείξουμε ότι $\sum_{i=1}^{N}\left((M^{N-i}g)(x)-\mathbb{E}_\nu(g)\right)(1-\kappa/2)^{2(i-1)}$ όταν $N\to\infty$.

Γενικότερα ισχύει ότι αν $(\alpha_{n,m})_{n,m\in\mathbb{N}}$ και $(\beta_m)_{m\in\mathbb{N}}$ είναι ακολουθίες πραγματικών αριθμών, με $\beta_m\geqslant 0$ για κάθε $m\in\mathbb{N}$,

$$\lim_{n\to\infty}\alpha_{n,m}=0 \text{ και } \sum_{m=1}^{\infty}\beta_m<+\infty,$$

τότε

$$\lim_{n\to\infty}\sum_{m=1}^{n}\alpha_{n,m}\beta_m=0.$$

Πράγματι, εφόσον η $(\alpha_{n,m})$ συγκλίνει, υπάρχει $A>0$ ώστε $|\alpha_{n,m}|\leqslant A$ για κάθε $n,m\in\mathbb{N}$. Επίσης,

$$\sum_{m=1}^{\infty}\beta_m<+\infty$$

και άρα έχουμε ότι για κάθε $\varepsilon>0$ υπάρχει $N\in\mathbb{N}$ ώστε για κάθε $n\geqslant N$

$$\sum_{m=N+1}^{n}\beta_m<\frac{\varepsilon}{A}.$$

Επομένως έχουμε ότι για κάθε $n\geq N$ ισχύει ότι

$$\begin{aligned}\left|\sum_{m=1}^{n}\alpha_{n,m}\beta_m\right| &= \left|\sum_{m=1}^{N}\alpha_{n,m}\beta_m+\sum_{m=N+1}^{n}\alpha_{n,m}\beta_m\right|\\ &\leqslant \sum_{m=1}^{N}|\alpha_{n,m}\beta_m|+\sum_{m=N+1}^{n}|\alpha_{n,m}\beta_m|\\ &\leqslant \sum_{m=1}^{N}|\alpha_{n,m}\beta_m|+\varepsilon.\end{aligned}$$

Τώρα, για $n\to\infty$ η παραπάνω σχέση δίνει

$$\left|\lim_{n\to\infty}\sum_{m=1}^{n}\alpha_{n,m}\beta_m\right|\leqslant\sum_{m=1}^{n}\left(\lim_{n\to\infty}|a_{n,m}|\right)\beta_m+\varepsilon=\varepsilon,$$

και εφόσον αυτό ισχύει για κάθε $\varepsilon>0$ έχουμε ότι

$$\lim_{n\to\infty}\sum_{m=1}^{n}\alpha_{n,m}\beta_m=0.$$

Συνεπώς έχουμε ότι

$$\begin{aligned}\lim_{N\to\infty} f_N(x) &= \mathbb{E}_\nu(f) + \lambda\sum_{i=1}^{\infty}\mathbb{E}_\nu(g)(1-\kappa/2)^{2(i-1)}\\ &= \mathbb{E}_\nu(f) + \lambda\mathbb{E}_\nu(g)\frac{1}{1-(1-\kappa/2)^2}\\ &= \mathbb{E}_\nu(f) + \lambda\mathbb{E}_\nu(g)\frac{4}{4-(2-\kappa)^2}\\ &= \mathbb{E}_\nu(f) + \lambda\mathbb{E}_\nu(g)\frac{4}{\kappa(4-\kappa)} \qquad (\kappa \leqslant 1)\\ &\leqslant \mathbb{E}_\nu(f) + \lambda\mathbb{E}_\nu(g)\frac{4}{3\kappa}.\end{aligned}$$

Όμως, $(M^N e^{\lambda f})(x) \to \mathbb{E}_\nu(e^{\lambda f})$, άρα

$$\mathbb{E}_\nu(e^{\lambda f}) \leqslant \lim_{N\to\infty} e^{\lambda f_N} \leqslant \exp\left\{\lambda\mathbb{E}_\nu(f) + \frac{4\lambda^2}{3\kappa}\mathbb{E}_\nu\frac{\sigma(x)^2}{n_x}\right\} = \exp\left\{\lambda\mathbb{E}_\nu(f) + \frac{4\lambda^2 D^2}{3}\right\}.$$

Τέλος, έχουμε ότι

$$\begin{aligned}\nu\left(\{f(x) \geqslant \mathbb{E}_\nu(f)+t\}\right) &= \nu\left(\left\{e^{\lambda f(x)} \geqslant e^{\lambda\mathbb{E}_\nu(f)}e^{\lambda t}\right\}\right)\\ &\leqslant \mathbb{E}_\nu(e^{\lambda f})e^{-\lambda\mathbb{E}_\nu(f)}e^{-\lambda t}\\ &\leqslant \exp\left\{\lambda\mathbb{E}_\nu(f) + \frac{4\lambda^2 D^2}{3}\right\}e^{-\lambda\mathbb{E}_\nu(f)}e^{-\lambda t}\\ &= \exp\left\{\frac{4\lambda^2 D^2}{3} - \lambda t\right\}.\end{aligned}$$

Ελαχιστοποιούμε ως προς λ λαμβάνοντας υπόψιν τους περιορισμούς $\lambda \leqslant 1/2C$ και $\lambda \leqslant 1/3\sigma_\infty$: Η συνάρτηση

$$h(\lambda) = \frac{4\lambda^2 D^2}{3} - \lambda t$$

έχει παράγωγο

$$h'(\lambda) = \frac{8\lambda D^2}{3} - t.$$

Συνεπώς παρουσιάζει ελάχιστο στο $\lambda_{\min} = 3t/8D^2$. Πρέπει επιπλέον

$$\lambda \leqslant \lambda_{\max} := \min\{1/2C, 1/3\sigma_\infty\} = \frac{1}{3}\min\{3/2C, 1/\sigma_\infty\},$$

δηλαδή

$$t \leqslant t_{\max} = \frac{8}{9}D^2\min\{3/2C, 1/\sigma_\infty\}.$$

Άρα, για $t \leqslant t_{\max}$ έχουμε

$$\nu\left(\{f(x) \geqslant \mathbb{E}_\nu(f)+t\}\right) \leqslant \exp\left\{-\frac{3t^2}{16D^2}\right\},$$

ενώ αν $t \geqslant t_{\max}$ έχουμε ότι η ελάχιστη τιμή της h λαμβάνεται για $\lambda = \lambda_{\max}$ και τότε έχουμε

$$\nu\left(\{f(x) \geqslant \mathbb{E}_\nu(f)+t\}\right) \leqslant \exp\left\{\frac{3t_{\max}}{8D^2}\left(\frac{t_{\max}}{2} - t\right)\right\}$$

καθώς $\lambda_{\max} = 3t_{\max}/8D^2$. □

Παρατήρηση 2.7.3. Μέσω της Πρότασης 2.4.3 και του Πορίσματος 2.3.4 μπορούμε να βρούμε φράγματα για την D^2, τα οποία είναι αρκετά κομψά, χωρίς να ξέρουμε, εκ των προτέρων, κάτι για την αναλλοίωτη κατανομή ν.

Παρατήρηση 2.7.4. Μπορεί κανείς να δει από την απόδειξη του θεωρήματος Gaussian συγκέντρωσης πως η συνθήκη ότι η D_x^2 είναι συνάρτηση Lipschitz μπορεί να αντικατασταθεί από την ασθενέστερη συνθήκη ότι η D_x^2 είναι φραγμένη από κάποια συνάρτηση Lipschitz. Για παράδειγμα αν η $\sigma^2(x)$ είναι φραγμένη, τότε μπορούμε να θέσουμε $D^2 := \sup\limits_x \frac{\sigma^2(x)}{\kappa n_x}$ και $C = 0$.

Παρατήρηση 2.7.5. Μπορούμε να αντικαταστήσουμε την συνθήκη ότι η σ_∞ είναι ομοιόμορφα φραγμένη με τη συνθήκη: για κάθε $x \in X$ υπάρχει s_x ώστε: $\mathbb{E}_{m_x}(e^{\lambda f}) \leqslant \exp\left\{\frac{\lambda^2 s_x^2}{2} + \lambda \mathbb{E}_{m_x}(f)\right\}$ για κάθε f 1-Lipschitz. Τότε ισχύει ένα παρόμοιο υεώρημα με αυτό που δείξαμε, αλλάζοντας όπου πρέπει την $\sigma^2(x)/n_x$ με την s_x^2. Το καινούργιο υεώρημα δεν ταιριάζει, όμως, στους διακριτούς χώρους, γιατί όταν οι πιθανότητες μετάβασης, από ένα σημείο σε κάποιο άλλο είναι μικρές, τότε το βέλτιστο s_x^2 για το οποίο η ανισότητα ισχύει είναι πολύ μεγαλύτερο από την $\sigma^2(x)$. Για παράδειγμα αν $d(x,y) = 1$, $m_x(y) = \varepsilon$ η s_x πρέπει να ικανοποιεί την $s_x^2 \geqslant \frac{1}{2}\ln\frac{1}{\varepsilon} >> \varepsilon$. Έτσι με αυτή την υπόθεση θα έχουμε κακές εκτιμήσεις για την D^2.

2.8 Τοπικός έλεγχος και λογαριθμική ανισότητα Sobolev

Θα προσπαθήσουμε τώρα να ελέγξουμε την κλίση της Mf σε κάποιο σημείο, συναρτήσει της κλίσης της f στα γειτονικά του σημεία. Εδώ ερχόμαστε πιο κοντά στη θεωρία Bakry-Émery, και θα μπορέσουμε να πάρουμε μια μορφή λογαριθμικής ανισότητας Sobolev.

Ορισμός 2.8.1 (νόρμα της κλίσης)**.** Επιλέγουμε $\lambda > 0$ και για κάθε συνάρτηση $f : X \to \mathbb{R}$ ορίζουμε την *λ-επιπέδου κλίση* της f θέτοντας

$$|\nabla_\lambda f|(x) := \sup_{y,y' \in X} \frac{|f(y) - f(y')|}{d(y,y')} e^{-\lambda d(x,y) - \lambda d(y,y')}.$$

Παρατήρηση 2.8.2. (i) Όταν τα y, y' απέχουν «πολύ» από το x τότε δεν συνεισφέρουν στο supremum, συνεπώς μας ενδιαφέρουν y, y' κοντά στο x.

(ii) Αν η f είναι λεία συνάρτηση και η X είναι συμπαγής πολλαπλότητα Riemann τότε η λ-επιπέδου κλίση τείνει στο γνωστό $|\nabla f|(x)$ για $\lambda \to \infty$.

(iii) Σε μετρικούς χώρους η κλίση μιας συνάρτησης f ορίζεται ως

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)}$$

Η σύνδεση των δύο ορισμών είναι η εξής: Έχουμε για $y' = x$ και για κάθε $y \in X$, ότι

$$\frac{|f(x) - f(y)|}{d(x,y)} e^{-2\lambda d(x,y)} \leqslant |\nabla_\lambda f|(x)$$

και για $\lambda \to 0$ παίρνουμε

$$\frac{|f(x) - f(y)|}{d(x,y)} \leqslant \lim_{\lambda \to 0} |\nabla_\lambda f|(x).$$

Άρα,

$$|\nabla f|(x) = \limsup_{y\to x} \frac{|f(x)-f(y)|}{d(x,y)} \leqslant \lim_{\lambda\to 0} |\nabla_\lambda f|(x).$$

(iv) Η $\ln|\nabla_\lambda f|$ είναι λ-Lipschitz. Πράγματι, για κάθε $x, x' \in X$ έχουμε ότι

$$\begin{aligned}
|\ln|\nabla_\lambda f|(x) - \ln|\nabla_\lambda f|(x')| &\leqslant \left| \sup_{y,y'} \left\{ \ln \frac{|f(y)-f(y')|}{d(y,y')} + (-\lambda d(x,y) - \lambda d(y,y')) \right.\right. \\
&\qquad \left.\left. - \ln \frac{|f(y)-f(y')|}{d(y,y')} + (-\lambda d(x',y) - \lambda d(y,y')) \right\} \right| \\
&= \left| \lambda \sup_{y,y'} \{d(x,y) - d(y,x')\} \right| \\
&\leqslant \lambda d(x,x').
\end{aligned}$$

Ορισμός 2.8.3 (αστάθεια)**.** Έστω $x, y \in X$ και ξ_{xy} το βέλτιστο ταίριασμα μεταξύ των m_x και m_y, δηλαδή αυτό για το οποίο

$$\kappa(x,y) = \frac{1}{d(x,y)} \int_{z,w} (d(x,y) - d(z,w))\, d\xi_{xy}(z,w).$$

Θέτουμε

$$\kappa_+(x,y) = \frac{1}{d(x,y)} \int_{z,w} (d(x,y) - d(z,w))_+\, d\xi_{xy}(z,w)$$

και

$$\kappa_-(x,y) = \frac{1}{d(x,y)} \int_{z,w} (d(x,y) - d(z,w))_-\, d\xi_{xy}(z,w),$$

όπου α_+ και α_- είναι το θετικό και το αρνητικό μέρος ενός πραγματικού αριθμού α. Έχουμε λοιπόν $\kappa(x,y) = \kappa_+(x,y) - \kappa_-(x,y)$. Η αστάθεια $U(x,y)$ ορίζεται ως

$$U(x,y) := \frac{\kappa_-(x,y)}{\kappa(x,y)} \quad \text{και} \quad U := \sup_{x,y\in X, x\neq y} U(x,y).$$

Παρατήρηση 2.8.4. Αν ο X είναι ε-γεωδαισιακός και $U(x,y) \leqslant A$ για κάθε $x, y \in X$ με $d(x,y) \leqslant \varepsilon$ τότε $U \leqslant A$. Αυτό προκύπτει εύκολα από το γεγονός ότι αν $\kappa(x,y) \geqslant \kappa$ για κάθε $x, y \in X$ με $d(x,y) \leqslant \varepsilon$ τότε $\kappa(x,y) \geqslant \kappa$ για κάθε $x, y \in X$.

Στα περισσότερα διακριτά παραδείγματα η αστάθεια θα είναι 0, δηλαδή το βέλτιστο ταίριασμα μεταξύ των m_x και m_y δεν αυξάνει τις αποστάσεις. Αυτός θα μπορούσε να είναι ο ορισμός της μη-αρνητικής τμηματικής καμπυλότητας Ricci για αλυσίδες Markov. Σε πολλαπλότητες Riemann η αστάθεια ελέγχεται από την μεγαλύτερη τμηματική καμπυλότητα.

Τα θεωρήματα που ακολουθούν έχουν ενδιαφέρον μόνο αν μπορούμε να πετύχουμε μια λογική εκτίμηση για την κλίση $|\nabla_\lambda f|$ βασιζόμενοι σε τοπικά δεδομένα. Αυτό δεν μπορεί να συμβεί όταν η f δεν είναι $\lambda -$ log-Lipschitz, κάτι το οποίο είναι απολύτως συνεπές με αυτά που είδαμε στα προηγούμενα: η Gaussian συγκέντρωση συμβαίνει σε πεπερασμένη κλίμακα, ακολουθούμενη από εκθετική συγκέντρωση, και συνεπώς δεν μπορεί να ισχύει η ανισότητα Log-Sobolev γενικά.

Λήμμα 2.8.5. *Έστω $x, y \in X$ με $\kappa(x,y) > 0$. Έστω A μια θετική συνάρτηση στον $X \times X$ τέτοια ώστε $\sup A / \inf A \leqslant e^{\varrho}$, όπου $\varrho \leqslant \frac{1}{2(1+U)}$. Έστω επίσης ξ_{xy} το βέλτιστο ταίριασμα μεταξύ των m_x και m_y. Τότε,*

$$\int_{z,w} A(z,w)\frac{d(z,w)}{d(x,y)}d\xi_{xy}(z,w) \leqslant (1-\kappa(x,y)/2)\int_{z,w} A(z,w)d\xi_{xy}(z,w),$$

και ειδικότερα

$$\int_{z,w} A(z,w)(d(z,w)-d(x,y))d\xi_{xy}(z,w) \leqslant 0.$$

Απόδειξη. Η ιδέα είναι η εξής: όταν η A είναι σταθερή, το αποτέλεσμα προφανώς ισχύει διότι, από τον ορισμό, $\int_{z,w} d(z,w)/d(x,y)\xi_{xy}(z,w) = 1-\kappa(x,y)$. Τώρα, όταν η A είναι αρκετά κοντά σε σταθερή συνάρτηση, ισχύει το ίδιο με κάποια αριθμητική απώλεια. Θέτουμε $F = \sup_{z,w} A(z,w)$. Τότε,

$$\int_{z,w} A(z,w)\frac{d(z,w)}{d(x,y)}d\xi_{xy}(z,w) = \int_{z,w} A(z,w) + F\int_{z,w}\frac{A(z,w)}{F}\left(\frac{d(z,w)}{d(x,y)}-1\right)d\xi_{xy}(z,w).$$

Ορίζουμε $L_- = \{(z,w) \in X \times X : d(x,y) < d(z,w)\}$ και $L_+ = X \times X \setminus L_-$. Συνεπώς έχουμε ότι για κάθε (z,w) στο L_+ ισχύει ότι $\frac{d(z,w)}{d(x,y)} - 1 < 0$ και για κάθε (z,w) στο L_+ αντίστοιχα ισχύει ότι $1 - \frac{d(z,w)}{d(x,y)} \geqslant 0$. Άρα, οι ορισμοί των k_- και k_+ μας δίνουν:

$$\kappa_-(x,y) = \int_{L_-}(d(z,w)/d(x,y)-1)d\xi_{xy}(z,w)$$

και

$$\kappa_+(x,y) = \int_{L_+}(1-d(z,w)/d(x,y))d\xi_{xy}(z,w).$$

Έχουμε επίσης ότι $A(z,w) \leqslant F$ για κάθε $(z,w) \in X \times X$ και ειδικότερα για κάθε (z,w) στον L_- και ότι $A(z,w) \leqslant e^{-\rho}F$ για κάθε (z,w) στον $X \times X$, άρα και για κάθε (z,w) στο L_+ καθώς η $\sup A/\inf A \leqslant e^{\rho}$ και συνεπώς $\inf A \geqslant e^{-\rho}F$. Έπεται ότι

$$\int_z A(z)\frac{d(x+z,y+z)}{d(x,y)} \leqslant \int_z A(z) + F(\kappa_-(x,y) - e^{-\varrho}\kappa_+(x,y)).$$

Τώρα, από τον ορισμό της U έχουμε $\kappa_-(x,y) \leqslant U\kappa(x,y)$. Σε συνδυασμό με τη σχέση $\varrho \leqslant \frac{1}{2(1+U)}$ θα δείξουμε ότι $e^{-\varrho}\kappa_+(x,y) - \kappa_-(x,y) \geqslant \kappa(x,y)/2$, και άρα θα έχουμε

$$\begin{aligned}\int_{z,w} A(z,w)\frac{d(z,w)}{d(x,y)}d\xi_{xy}(z,w) &\leqslant \int_{z,w} A(z,w)d\xi_{xy}(z,w) - F\kappa(x,y)/2\\ &\leqslant (1-\kappa(x,y)/2)\int_{z,w} A(z,w)d\xi_{xy}(z,w)\end{aligned}$$

όπως θέλαμε. Αρκεί λοιπόν να δείξουμε ότι $\inf A \geqslant e^{-\rho}F$. Εδώ θα φανεί γιατί επιλέξαμε την συνθήκη $\rho \leqslant 1/2(U+1)$. Έχουμε ότι

$$e^{-\varrho}\kappa_+ - \kappa_- \geqslant \frac{\kappa}{2}$$

αν και μόνο αν

$$e^{-\rho}(\kappa+\kappa_-)-\kappa_- - \frac{\kappa}{2} \geqslant 0$$

ή, ισοδύναμα,

$$2e^{-\rho}\kappa + 2e^{-\rho}\kappa_- - \kappa - 2\kappa_- \geqslant 0,$$

δηλαδή

$$(*) \qquad (2e^{-\rho}-1)\kappa + (2e^{-\rho}-2)\kappa_- \geqslant 0.$$

Έχουμε τώρα ότι $2e^{-\rho}-1 \geqslant 0$ καθώς αυτό είναι ισοδύναμο με την $\rho \leqslant e^{1/2}$ η οποία ισχύει από την υπόθεση ότι $\rho \leqslant e^{1/2(U+1)}$ και το γεγονός ότι $U \geqslant 0$. Συνεπώς, για να δείξουμε την ανισότητα $(*)$ αρκεί να δείξουμε την $(2e^{-\rho}-1)\kappa_- + U(2e^{-\rho}-2)\kappa_- \geqslant 0$ αφού $\kappa_- \leqslant U\kappa$. Όμως,

$$(2e^{-\rho}-1)\kappa_- + U(2e^{-\rho}-2)\kappa_- \geqslant 0$$

αν και μόνο αν

$$2e^{-\rho} - 1 + 2\left(e^{-\rho}-2\right)U \geqslant 0$$

ή, ισοδύναμα,

$$e^{-\rho} - \frac{2U+1}{2(U+1)} \geqslant 0,$$

και αν χρησιμοποιήσουμε και την $e^{\alpha} \geqslant \alpha+1$ έχουμε ότι

$$\begin{aligned}\frac{2U+1}{2(U+1)} &= 1 + \frac{2U+1}{2(U+1)} - 1\\ &= 1 - \frac{1}{2(U+1)}\\ &\leqslant \exp\left\{-\frac{1}{2(U+1)}\right\}.\end{aligned}$$

Συνεπώς, αν $\rho \leqslant 1/2(U+1)$ τότε $e^{-\rho} \geqslant \exp\left\{-\frac{1}{2(U+1)}\right\}$ και τελικά

$$e^{-\rho} - \frac{2U+1}{2(U+1)} \geqslant \exp\left\{-\frac{1}{2(U+1)}\right\} - \frac{2U+1}{2(U+1)} \geqslant 0,$$

το οποίο δείχνει το ζητούμενο. □

Θεώρημα 2.8.6 (συστολή κλίσης)**.** *Υποθέτουμε ότι η coarse καμπυλότητα Ricci είναι τουλάχιστον $\kappa > 0$. Θεωρούμε $\lambda \leqslant \frac{1}{20\sigma_\infty(1+U)}$ και την λ-επιπέδου κλίση $|\nabla_\lambda|$. Τότε, για κάθε συνάρτηση $f : X \to \mathbb{R}$ με $|\nabla_\lambda f| < \infty$ έχουμε*

$$|\nabla_\lambda(Mf)|(x) \leqslant (1-\kappa/2)M(|\nabla_\lambda f|)(x)$$

για κάθε $x \in X$.

Απόδειξη. Έστω $x, y, y' \in X$ και ξ_{xy} και $\xi_{xy'}$ τα βέλτιστα ταιριάσματα μεταξύ των m_x, m_y και $m_y, m_{y'}$ αντίστοιχα. Εφαρμόζοντας το gluing λήμμα (βλέπε [Vill03, Λήμμα 7.6]) μπορούμε να επιλέξουμε βέλτιστο ταίριασμα $\xi_{xyy'}$ μεταξύ των ξ_{xy} και $\xi_{yy'}$. Έστω $\Lambda = e^{-\lambda(d(x,y)+d(y,y'))}$. Έχουμε

$$
\begin{aligned}
\text{(2.8.1)} \qquad \frac{|Mf(y) - Mf(y')|}{d(y,y')}\Lambda &= \frac{\Lambda}{d(y,y')}\left|\int_z f(z)dm_y(z) - \int_{z'} f(z')dm_{y'}(z')\right| \\
&= \frac{\Lambda}{d(y,y')}\left|\int_{z,z'} (f(z) - f(z'))\, d\xi_{yy'}(z,z')\right| \\
&\leqslant \frac{\Lambda}{d(y,y')}\int_{z,z'} |f(z) - f(z')|\, d\xi_{yy'}(z,z') \\
&= \int_{w,z,z'} \frac{|f(z) - f(z')|\, e^{-\lambda(d(w,z)+d(z,z'))}}{d(z,z')} \\
&\quad \times \frac{\Lambda d(z,z')}{d(y,y')e^{-\lambda(d(w,z)+d(z,z'))}} d\xi_{xyy'}(w,z,z') \\
&\leqslant \int_{w,z,z'} \sup_{z,z'}\left\{\frac{|f(z) - f(z')|\, e^{-\lambda(d(w,z)+d(z,z'))}}{d(z,z')}\right\} \\
&\quad \times \frac{\Lambda d(z,z')}{d(y,y')e^{-\lambda(d(w,z)+d(z,z'))}} d\xi_{xyy'} \\
&= \int_{w,z,z'} |\nabla_\lambda f|(w) \cdot \frac{\Lambda d(z,z')}{d(y,y')e^{-\lambda(d(w,z)+d(z,z'))}} d\xi_{xyy'}(w,z,z') \\
&= \int_{w,z,z'} B(w,z,z')\Gamma(w,z,z') d\xi_{xyy'}(w,z,z'),
\end{aligned}
$$

όπου $B(w,z,z') = |\nabla_\lambda f|(w)$ και $\Gamma(w,z,z') = \exp\{\lambda(d(w,z) + d(z,z') - d(x,y) - d(y,y'))\}$. Έχουμε ότι $\operatorname{diam}(\operatorname{supp}(m_x)) \leqslant 2\sigma_\infty$ και $\operatorname{diam}(\operatorname{supp}(m_y)) \leqslant 2\sigma_\infty$ και συνεπώς για κάθε $w, z \in X$ με $w \in \operatorname{supp}(m_x)$, $z \in \operatorname{supp}(m_y)$ ισχύει ότι

$$
\begin{aligned}
|d(w,z) - d(x,y) + d(z,z') - d(y,y')| &\leqslant |d(w,z) - d(x,y)| + |d(z,z') - d(y,y')| \\
&\leqslant 4\sigma_\infty + 4\sigma_\infty = 8\sigma_\infty,
\end{aligned}
$$

άρα $|\Gamma(w,z,z')| \leqslant e^{8\sigma_\infty}$. Επίσης η $\ln|\nabla_\lambda f|$ είναι $\lambda-$ Lipschitz και επομένως παίρνουμε ότι

$$|B(w,z,z')| = |\nabla_\lambda f|(w) \leqslant e^{2\lambda\sigma_\infty}.$$

Αν τώρα θέσουμε $A(w,z,z') = B(w,z,z')\Gamma(w,z,z')$ έχουμε ότι $|A(w,z,z')| \leqslant e^{10\lambda\sigma_\infty}$. Θα εφαρμόσουμε το Λήμμα 2.8.5 για $\rho = 10\lambda\sigma_\infty$ και επομένως πρέπει να ισχύει επιπλέον ότι $\rho \leqslant 1/2(U+1)$ δηλαδή ότι $\lambda \leqslant 1/20\sigma_\infty(U+1)$, το οποίο έχουμε από την υπόθεση. Άρα παίρνουμε ότι

$$\text{(2.8.2)} \qquad \int_{w,z,z'} A(w,z,z')\frac{d(z,z')}{d(y,y')} d\xi_{xyy'}(w,z,z') \leqslant \left(1 - \frac{\kappa}{2}\right)\int_{w,z,z'} A(w,z,z') d\xi_{xyy'}(w,z,z').$$

Επίσης,

(2.8.3)

$$
\begin{aligned}
\int_{w,z,z'} B(w,z,z')\Gamma(w,z,z') d\xi_{xyy'}(w,z,z') &= \int_{w,z,z'} B(w,z,z') d\xi_{xyy'}(w,z,z') \\
&\quad + \int_{w,z,z'} B(w,z,z')(\Gamma(w,z,z') - 1) d\xi_{xyy'}(w,z,z')
\end{aligned}
$$

και η ανισότητα $e^\alpha - 1 \leqslant \alpha e^\alpha$ για κάθε $\alpha \in \mathbb{R}$ δίνει

$$\begin{aligned}\Gamma(w,z,z') - 1 &= \exp\{\lambda(d(w,z) + d(z,z') - d(x,y) - d(y,y'))\} - 1 \\ &\leqslant \lambda(d(w,z) + d(z,z') - d(x,y) - d(y,y'))\exp\{\lambda(d(w,z) + d(z,z') - d(x,y) - d(y,y'))\},\end{aligned}$$

δηλαδή

$$\Gamma(w,z,z') - 1 \leqslant \lambda(d(w,z) + d(z,z') - d(x,y) - d(y,y'))\Gamma(w,z,z').$$

Έπεται ότι

$$\begin{aligned}\int_{w,z,z'} B(w,z,z')(\Gamma(w,z,z') - 1)d\xi_{xyy'} &\leqslant \lambda \int_{w,z,z'} B(w,z,z')\Gamma(w,z,z')(d(w,z) + d(z,z') \\ &\qquad - d(x,y) - d(y,y'))d\xi_{xyy'} \\ &= \lambda \int_{w,z,z'} A(z,z,w')(d(w,z) - d(x,y))d\xi_{xyy'} \\ &+ \lambda \int_{w,z,z'} A(z,z,w')(d(z,z') - d(y,y'))d\xi_{xyy'}.\end{aligned}$$

Πάλι από το Λήμμα 2.8.5 έχουμε ότι κάθε ένας από τους δύο προσθετέους είναι μικρότερος ή ίσος του 0 οπότε

$$\int_{w,z,z'} B(w,z,z')(\Gamma(w,z,z') - 1)d\xi_{xyy'}(w,z,z') \leqslant 0,$$

το οποίο δίνει την

$$\begin{aligned}\int_{w,z,z'} B(w,z,z')\Gamma(w,z,z')d\xi_{xyy'}(w,z,z') &\leqslant \int_{w,z,z'} B(w,z,z')d\xi_{xyy'}(w,z,z') \\ &= \int_{w,z,z'} |\nabla_\lambda f|(w)d\xi_{xyy'}(w,z,z') \\ &= \int_w |\nabla_\lambda f|(w)dm_x(w) = M|\nabla_\lambda f|(x).\end{aligned}$$

Από τις ανισότητες (2.8.1), (2.8.2) και (2.8.3) έχουμε τελικά ότι για κάθε $y, y' \in X$ ισχύει

$$\frac{|Mf(y) - Mf(y')|}{d(y,y')} e^{-\lambda(d(x,y)+d(y,y'))} \leqslant \left(1 - \frac{\kappa}{2}\right) M\left(|\nabla_\lambda f|\right)(x),$$

επομένως

$$\sup_{y,y'} \left\{ \frac{|Mf(y) - Mf(y')|}{d(y,y')} e^{-\lambda(d(x,y)+d(y,y'))} \right\} \leqslant \left(1 - \frac{\kappa}{2}\right)\left(M|\nabla_\lambda f|\right)(x),$$

το οποίο δίνει το ζητούμενο αποτέλεσμα:

$$|\nabla_\lambda Mf|(x) \leqslant \left(1 - \frac{\kappa}{2}\right) M\left(|\nabla_\lambda f|\right)(x).$$

□

Περνάμε τώρα στη λογαριθμική ανισότητα Sobolev. Η εντροπία μιας συνάρτησης ως προς το μέτρο πιθανότητας ν ορίζεται ως

$$\mathrm{Ent}_\nu(f) := \mathbb{E}_\nu(f \ln f) - \mathbb{E}_\nu(f)\ln\left(\mathbb{E}_\nu(f)\right).$$

Αν (X, d, m) είναι ένας μετρικός χώρος με έναν τυχαίο περίπατο και ν αναλλοίωτη κατανομή για τον m, τότε ισχύει η σχέση

$$\operatorname{Ent}_\nu(f) = \int_x \operatorname{Ent}_{m_x}(f) d\nu(x) + \operatorname{Ent}_\nu(Mf),$$

σε αντιστοιχία με τον τύπο

$$\operatorname{Var}_\nu(f) = \int_x \operatorname{Var}_{m_x}(f) d\nu(x) + \operatorname{Var}_\nu(Mf).$$

Πράγματι, έχουμε ότι για κάθε $x \in X$ ισχύει

$$\begin{aligned}\operatorname{Ent}_{m_x}(f) &= \mathbb{E}_{m_x}(f \ln f) - \mathbb{E}_{m_x}(f) \ln\left(\mathbb{E}_{m_x}(f)\right) \\ &= M(f \ln f)(x) - Mf(x) \ln\left(Mf(x)\right),\end{aligned}$$

και συνεπώς, αφού η ν είναι αναλλοίωτη για τον m έχουμε για κάθε g ότι $E_\nu(Mg) = E_\nu(g)$, άρα

$$\begin{aligned}\int_x \operatorname{Ent}_{m_x}(f) d\nu(x) &= \mathbb{E}_\nu(M(f \ln f)) - \mathbb{E}_\nu(Mf \cdot M(f \ln f)) \\ &= \mathbb{E}_\nu(f \ln f) - \mathbb{E}_\nu(Mf \ln\left(M(f)\right)).\end{aligned}$$

Από την άλλη πλευρά,

$$\begin{aligned}\operatorname{Ent}_\nu(Mf) &= \mathbb{E}_\nu(Mf \ln(Mf)) - \mathbb{E}_\nu(Mf) \ln\left(\mathbb{E}_\nu((Mf))\right) \\ &= \mathbb{E}_\nu(Mf \ln(Mf)) - \mathbb{E}_\nu(f) \ln\left(\mathbb{E}_\nu(f)\right).\end{aligned}$$

Επομένως,

$$\operatorname{Ent}_\nu(f) = \int_x \operatorname{Ent}_{m_x}(f) d\nu(x) + \operatorname{Ent}_\nu(Mf).$$

Από αυτή τη σχέση με επαγωγή προκύπτει ότι

$$\operatorname{Ent}_\nu(f) = \sum_{t=0}^{\infty} \int_x \operatorname{Ent}_{m_x}(M^t f) d\nu(x),$$

διότι

$$\lim_{t \to \infty} \operatorname{Ent}_\nu(M^t f) = 0,$$

αφού $M^t f \to \mathbb{E}_\nu(f) = c$ και $\operatorname{Ent}_\nu(c) = \mathbb{E}_\nu(c \ln c) - \mathbb{E}_\nu(c) \ln \mathbb{E}_\nu(c) = c \ln c - c \ln c = 0$. Θα χρειαστούμε επίσης το παρακάτω λήμμα:

Λήμμα 2.8.7. *Έστω $f : X \to \mathbb{R}$ με $|\nabla_\lambda f| < \infty$ και έστω $x \in X$. Τότε η f είναι Lipschitz με σταθερά $e^{4\lambda\sigma_\infty} M\left(|\nabla_\lambda f|\right)(x)$ στο $\operatorname{supp}(m_x)$.*

Απόδειξη. Για κάθε $y, y' \in X$ έχουμε ότι

$$|f(y) - f(y')| \leqslant |\nabla_\lambda f|(z) d(y, y') e^{\lambda(d(y,y') + d(z,y))}$$

και για $z = y$ παίρνουμε ότι

$$|f(y) - f(y')| \leqslant |\nabla_\lambda f|(y) d(y, y') e^{\lambda d(y,y')}.$$

Αν τώρα $y, y' \in \operatorname{supp}(m_x)$ έχουμε ότι

$$|f(y) - f(y')| \leqslant |\nabla_\lambda f|(y) d(y,y') e^{\lambda d(y,y')} \leqslant |\nabla_\lambda f|(y) d(y,y') e^{2\lambda\sigma_\infty}.$$

Το γεγονός ότι η $\ln|\nabla_\lambda f|$ είναι $\lambda-$Lipschitz δίνει ότι

$$|\nabla_\lambda f|\,(y) \leqslant e^{2\lambda d(y,y')}\,|\nabla_\lambda f|\,(y'),$$

και συνεπώς

$$|\nabla_\lambda f|(y) \leqslant e^{2\lambda\sigma_\infty} \inf_{\operatorname{supp}(m_x)} |\nabla_\lambda f| \leqslant e^{2\lambda\sigma_\infty} M\left(|\nabla_\lambda f|\right)(x).$$

Τελικά,

$$|f(y) - f(y')| \leqslant d(y,y') M\left(|\nabla_\lambda f|\right)(x) e^{4\lambda\sigma_\infty}.$$

□

Θεώρημα 2.8.8 (ανισότητα Log-Sobolev). *Υποθέτουμε ότι η coarse καμπυλότητα Ricci είναι τουλάχιστον $\kappa > 0$ και έστω ν η αναλλοίωτη κατανομή του m. Υποθέτουμε ότι $\lambda \leqslant 1/20\sigma_\infty(U+1)$ και θεωρούμε την λ-επιπέδου κλίση $|\nabla_\lambda|$. Τότε, για κάθε $f : X \to \mathbb{R}$ με $|\nabla_\lambda f| < \infty$, έχουμε ότι*

$$\operatorname{Var}_\nu(f) \leqslant \left(\sup_x \frac{4\sigma^2(x)}{\kappa n_x}\right) \int_x |\nabla_\lambda f|^2(x) d\nu(x),$$

και, αν η f είναι θετική,

$$\operatorname{Ent}_\nu(f) \leqslant \left(\sup_x \frac{4\sigma^2(x)}{\kappa n_x}\right) \int_x \frac{|\nabla_\lambda f|^2(x)}{f(x)} d\nu(x).$$

Αν επιπλέον ο τυχαίος περίπατος m είναι αναλλοίωτος ως προς τη ν, τότε

$$\operatorname{Var}_\nu(f) \leqslant \int_x V(x) |\nabla_\lambda f|^2(x) d\nu(x)$$

και

$$\operatorname{Ent}_\nu(f) \leqslant \int_x V(x) \frac{|\nabla_\lambda f|^2(x)}{f(x)} d\nu(x),$$

όπου

$$V(x) = 2\sum_{t=0}^{\infty} \left(1 - \frac{\kappa}{2}\right)^{2t} M^{t+1}\left(\frac{\sigma^2(x)}{n_x}\right).$$

Απόδειξη. Το Λήμμα 2.8.7, για σταθερό $x \in X$, δίνει ότι η f είναι $M(|\nabla_\lambda f|)(x) e^{4\lambda\sigma_\infty}-$Lipschitz στο $\operatorname{supp}(m_x)$. Άρα, από τον ορισμό του n_x έχουμε ότι

$$\operatorname{Var}_{m_x}\left(\frac{f}{M(|\nabla_\lambda f|(x)\sigma^2(x) e^{4\sigma_\infty}}\right) \leqslant \frac{\sigma^2(x)}{n_x},$$

δηλαδή

$$\operatorname{Var}_{m_x}(f) \leqslant \frac{e^{8\sigma_\infty}\left(M(|\nabla_\lambda f|(x)\right)^2 \sigma^2(x)}{n_x}.$$

Επιπλέον ισχύει ότι

$$\begin{aligned} e^{8\lambda\sigma_\infty} &\leqslant \exp\left\{\frac{8\sigma_\infty}{20\sigma_\infty(1+U)}\right\} \\ &\leqslant \exp\left\{\frac{8\sigma_\infty}{40\sigma_\infty}\right\} \\ &\leqslant e^{1/5} \leqslant 2, \end{aligned}$$

και άρα

$$\mathrm{Var}_{m_x}(f) \leqslant \frac{\left(M(|\nabla_\lambda f|(x)\right)^2 \sigma^2(x)}{n_x}.$$

Θα δείξουμε ότι $\mathrm{Ent}_{m_x}(f) \leqslant \mathrm{Var}_{m_x}(f)/Mf(x)$. Πράγματι, αν υποθέσουμε ότι $\mathbb{E}_{m_x}(f) = 1$, τότε η σχέση $\alpha - 1 \leqslant \alpha \ln \alpha \leqslant \alpha^2 - \alpha$ για $\alpha > 0$ δίνει ότι

$$\begin{aligned} \mathrm{Ent}_{m_x}(f) &= \mathbb{E}_{m_x}(f \ln f) - \mathbb{E}_{m_x}(f) \ln\left(\mathbb{E}_{m_x}(f)\right) \\ &\leqslant \mathbb{E}_{m_x}(f^2) - \mathbb{E}_{m_x}(f) - \mathbb{E}_{m_x}(f)\ln\left(\mathbb{E}_{m_x}(f)\right) \\ &= \mathbb{E}_{m_x}(f^2) - 1 \\ &= \frac{\mathbb{E}_{m_x}(f^2) - \mathbb{E}_{m_x}(f)^2}{\mathbb{E}_{m_x}(f)^2} \\ &= \frac{\mathrm{Var}_{m_x}(f^2)}{Mf(x)}. \end{aligned}$$

Αν τώρα η f είναι τυχούσα, θέτοντας $g = f/\mathbb{E}_{m_x}(f)$ έχουμε ότι $\mathbb{E}_{m_x}(g) = 1$, άρα $\mathrm{Ent}_{m_x}(g) \leqslant \mathrm{Var}_{m_x}(g)/Mg(x) = \mathrm{Var}_{m_x}(g)$, και εφόσον $\mathrm{Ent}_{m_x}\left(f/\mathbb{E}_{m_x}(f)\right) = \mathrm{Ent}_{m_x}(f)/\mathbb{E}_{m_x}(f)$, ισχύει ότι

$$\frac{\mathrm{Ent}_{m_x}}{\mathbb{E}_{m_x}(f)} \leqslant \frac{\mathrm{Var}_{m_x}(f)}{\mathbb{E}_{m_x}(f)^2},$$

το οποίο δίνει το ζητούμενο. Τελικά παίρνουμε

$$\mathrm{Ent}_\nu(f) \leqslant \frac{2(M|\nabla_\lambda f|)(x))^2\sigma^2(x)}{n_x Mf(x)}.$$

Συνεπώς,

$$\mathrm{Var}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\int_x \frac{\sigma^2(x)}{n_x}\left(M\left(|\nabla_\lambda M^t f|(x)\right)\right)^2 d\nu(x)$$

και

$$\mathrm{Ent}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\int_x \frac{\sigma^2(x)}{n_x}\frac{\left(M\left(|\nabla_\lambda M^t f|(x)\right)\right)^2}{M^{t+1}f(x)} d\nu(x).$$

Το Θεώρημα 2.8.6 δίνει ότι $|\nabla_\lambda M^t f|(y) \leqslant (1-\kappa/2)^t M^t\left(|\nabla_\lambda f|\right)(y)$, άρα

$$\mathrm{Var}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\int_x \frac{\sigma^2(x)}{n_x}\left(M^{t+1}|\nabla_\lambda f|\right)^2\left(1-\frac{\kappa}{2}\right)^{2t} d\nu(x)$$

και

$$\mathrm{Ent}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\int_x \frac{\sigma^2(x)}{n_x}\frac{\left(M\left(|\nabla_\lambda M^{t+1} f|(x)\right)\right)^2}{M^{t+1}f(x)}\left(1-\frac{\kappa}{2}\right)^{2t} d\nu(x).$$

Χρησιμοποιώντας την κυρτότητα των συναρτήσεων $\alpha \mapsto \alpha^2$ και $(\alpha,\beta) \mapsto \alpha^2$, καταλήγουμε στις ανισότητες

$$\left(M^{t+1}|\nabla_\lambda f|\right)^2 \leqslant M^{t+1}\left(|\nabla_\lambda f|^2\right)(x)$$

και

$$\frac{\left(M^{t+1}|\nabla_\lambda f|\right)^2}{M^{t+1}f(x)} \leqslant M^{t+1}\left(\frac{(\nabla_\lambda f|^2}{f}\right)(x),$$

οι οποίες τελικά δίνουν τις ανισότητες

$$\mathrm{Var}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x \frac{\sigma^2(x)}{n_x}M^{t+1}\left(|\nabla_\lambda f|^2\right)(x)d\nu(x) \tag{2.8.4}$$

και

$$\mathrm{Ent}_\nu(f) \leqslant 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x \frac{\sigma^2(x)}{n_x}M^{t+1}\left(\frac{|\nabla_\lambda f|^2}{f}(x)\right)d\nu(x). \tag{2.8.5}$$

Επίσης,

$$\begin{aligned}\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^2 &= \frac{1}{1-(1-\kappa/2)^2}\\ &= \frac{4}{4-(2-\kappa)^2}\\ &= \frac{4}{\kappa(4-\kappa)}\\ &\leqslant \frac{4}{3\kappa}.\end{aligned} \tag{2.8.6}$$

Αν ο τυχαίος περίπατος m δεν είναι αντιστρέψιμος ως προς την ν, ισχύει η ταυτότητα

$$\int_x Mh(x)d\nu(x) = \int_x\int_y h(y)dm_x(y)d\nu(x) = \int_y\int_x h(x)dm_y(x)d\nu(y) = \int_x h(x)d\nu(x)$$

και, με επαγωγή,

$$\int_x g(x)M^{t+1}h(x)d\nu(x) \leqslant (\sup g)\int_x M^{t+1}h(x)d\nu(x) = (\sup g)\int_x h(x)d\nu(x).$$

Εφαρμόζοντας το παραπάνω (δύο φορές) για $g(x) = \frac{\sigma^2(x)}{n_x}$ και $h(x) = |\nabla_\lambda f|^2(x)$ και $h(x) = |\nabla_\lambda f|^2(x)/f(x)$, σε συνδυασμό με τις (2.8.4), (2.8.5) και (2.8.6), έχουμε

$$\mathrm{Var}_\nu(f) \leqslant \left(\sup_x \frac{8\sigma^2(x)}{3\kappa n_x}\right)\int_x \frac{|\nabla_\lambda f|^2(x)}{f(x)}d\nu(x)$$

και

$$\mathrm{Ent}_\nu(f) \leqslant \left(\sup_x \frac{8\sigma^2(x)}{3\kappa n_x}\right)\int_x M^{t+1}\left(\frac{|\nabla_\lambda f|^2}{f}(x)\right)d\nu(x).$$

Στην περίπτωση που ο m είναι αντιστρέψιμος ως προς την ν ισχύει η ταυτότητα

$$\begin{aligned}\int g(x)Mh(x)d\nu(x) &= \int_x\int_y g(x)h(y)dm_x(y)d\nu(x)\\ &= \int_y\int_x g(x)h(y)dm_y(x)d\nu(y)\\ &= \int_y h(y)Mg(y)d\nu(y)\end{aligned}$$

και συνεπώς, επαγωγικά,

$$\int_x g(x)M^{t+1}h(x)d\nu(x) = \int_x h(x)M^{t+1}g(x)d\nu(x).$$

Άρα, οι ανισότητες (2.8.4) και (2.8.5) δίνουν

$$\begin{aligned}\mathrm{Var}_\nu(f) &\leqslant 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x \frac{\sigma^2(x)}{n_x}M^{t+1}\left(|\nabla_\lambda f|^2\right)(x)d\nu(x)\\ &= 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x M^{t+1}\left(\frac{\sigma^2(x)}{n_x}\right)|\nabla_\lambda f|^2(x)d\nu(x)\\ &= \int_x V(x)|\nabla_\lambda f|^2(x)d\nu(x)\end{aligned}$$

και

$$\begin{aligned}\mathrm{Ent}_\nu(f) &\leqslant 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x \frac{\sigma^2(x)}{n_x}M^{t+1}\left(\frac{|\nabla_\lambda f|^2}{f}\right)(x)d\nu(x)\\ &= 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}\int_x M^{t+1}\left(\frac{\sigma^2(x)}{n_x}\right)\frac{|\nabla_\lambda f|^2(x)}{f(x)}d\nu(x)\\ &= \int_x V(x)\frac{|\nabla_\lambda f|^2(x)}{f(x)}d\nu(x),\end{aligned}$$

όπου

$$V(x) = 2\sum_{t=0}^{\infty}\left(1-\frac{\kappa}{2}\right)^{2t}M^{t+1}\left(\frac{\sigma^2(x)}{n_x}\right).$$

□

Παρατήρηση 2.8.9. Μπορούμε να δούμε την ποσότητα $V(x)|\nabla_\lambda f|^2(x)$ ως μια εκδοχή της μορφής Dirichlet του τυχαίου περίπατου. Αν επιπλέον η $\sigma^2(x)/\kappa n_x$ είναι $C-$Lipschitz (όπως στο Θεώρημα 2.5.8) τότε $V(x) \leqslant \frac{4\sigma^2}{\kappa n} + 2C\frac{J(x)}{\kappa}$.

2.9 Εκθετική συγκέντρωση σε χώρους μη-αρνητικής coarse καμπυλότητας Ricci

Σε αυτήν την παράγραφο θα μελετήσουμε μια περίπτωση όπου επιτυγχάνεται εκθετική συγκέντρωση, χωρίς να υποθέσουμε ότι η coarse καμπυλότητα Ricci είναι ομοιόμορφα κάτω φραγμένη από θετική σταθερά. Θα αρκεστούμε στην ασθενέστερη υπόθεση της μη αρνητικής καμπυλότητας σε κάθε δύο σημεία του χώρου, και την ύπαρξη ενός «σημείου έλξης». Το τελευταίο μπορεί να γίνει πιο κατανοητό στο παρακάτω βασικό παράδειγμα.

Θεωρούμε το σύνολο των φυσικών αριθμών εφοδιασμένο με την συνήθη απόσταση. Ο τυχαίος περίπατος που θα επιλέξουμε είναι ο απλός τυχαίος περίπατος, όπου αν κάποια χρονική στιγμή βρισκόμαστε σε έναν φυσικό N μετακινούμαστε στον $N-1$ με πιθανότητα $p>0$ και στον $N+1$ με πιθανότητα $1-p$. Αν βρισκόμαστε στο 0 παραμένουμε εκεί με πιθανότητα p και μετακινούμαστε στο 1 με πιθανότητα $1-p$. Δηλαδή έχουμε $m_N = p\delta_{N-1} + (1-p)\delta_{N+1}$, αν $N \geqslant 1$, και $m_0 = p\delta_0 + (1-p)\delta_1$. Εδώ η ύπαρξη αναλλοίωτης κατανομής ισοδυναμεί με το να έχουμε ότι $p > 1/2$. Στην περίπτωση αυτή έχουμε εκθετική συγκέντρωση με decay distance $\frac{1}{\ln(p/(1-p))}$. Για $p = \frac{1}{2} + \varepsilon$ με ε κοντά στο μηδέν αυτή η ποσότητα συμπεριφέρεται σαν $1/(4\varepsilon)$. Σε αντίθετη περίπτωση η μη-αρνητική καμπυλότητα δεν αρκεί για να πετύχουμε εκθετική συγκέντρωση. Γεωμετρικά αυτό που συμβαίνει στην περίπτωση που το $p > \frac{1}{2}$ είναι ότι το σημείο 0 «τραβάει» τον γείτονα του, και, μέσω της μη-αρνητικής καμπυλότητας, αυτό μεταδίδεται σε όλους τους φυσικούς. Αυτό το παράδειγμα θα είναι υπόδειγμα για τις υποθέσεις που θα βάλουμε στο παρακάτω θεώρημα.

Θεώρημα 2.9.1 (Εκθετική Συγκέντρωση)**.** *Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο. Έστω επίσης ότι ο (X, d) είναι r-γεωδαισιακός, $r > 0$, και υπάρχει $o \in X$ ώστε να ισχύουν οι παρακάτω υποθέσεις:*

(i) *Για κάθε $x, y \in X$ έχουμε $\kappa(x, y) \geqslant 0$ (μη-αρνητική καμπυλότητα).*

(ii) *Για κάθε $x \in X$ με $r \leqslant d(x, o) < 2r$ ισχύει ότι $W_1(m_x, \delta_o) < d(x, o)$ (το ο «έλκει τα κοντινά σημεία»).*

(iii) *Υπάρχει $s > 0$ ώστε για κάθε m_x, $x \in X$ να ικανοποιείται μια Gaussian-τύπου ανισότητα μετασχηματισμού Laplace: Για κάθε $\lambda > 0$ και για κάθε 1-Lipschitz συνάρτηση $f : \operatorname{supp}(m_x) \to \mathbb{R}$ ισχύει*

$$\mathbb{E}_{m_x}\left(e^{\lambda f}\right) \leqslant e^{\lambda^2 s^2/2} e^{\mathbb{E}_{m_x}(f)}.$$

Θέτουμε $\rho = \inf\{d(x, o) - W_1(m_x, \delta_o) : x \in X \text{ με } r \leqslant d(x, o) < r\}$ και υποθέτουμε ότι $\rho > 0$. Τότε υπάρχει αναλλοίωτη κατανομή για τον τυχαίο περίπατο. Επιπλέον αν $D = s^2/\rho$ και $M = r + 2s^2/\rho + \rho\left(1 + J^2(o)/4s^2\right)$, τότε για κάθε αναλλοίωτη κατανομή ν για τον τυχαίο περίπατο ισχύει ότι

$$\int_x e^{d(x,o)/D} d\nu(x) \leqslant \left(4 + J^2(o)/s^2\right) e^{M/D}$$

και συνεπώς για κάθε 1-Lipschitz συνάρτηση $f : X \to \mathbb{R}$ και για κάθε $t \geqslant 0$ έχουμε την

$$\nu\left(|f - f(o)| \geqslant t + M\right) \leqslant \left(8 + 2J^2(o)/s^2\right) e^{-t/D}.$$

Δηλαδή, έχουμε εκθετική συγκέντρωση. Η τρίτη συνθήκη ισχύει πάντα για $s = \sigma_\infty$, λόγω της Πρότασης 1.16 στο [Led01].

Θα χρειαστούμε το παρακάτω λήμμα, το οποίο δείχνει ότι «η μη-αρνητική καμπυλότητα μεταδίδει την έλξη», δηλαδή ότι για σημεία τα οποία δεν είναι απαραίτητα $2r$-κοντά στο o η δεύτερη συνθήκη του θεωρήματος ισχύει με σφάλμα ρ.

Λήμμα 2.9.2. *Έστω (X,d,m) όπως στο προηγούμενο θεώρημα. Τότε, για κάθε $x \in X$ με $d(x,o) \geqslant r$ ισχύει ότι $W_1(m_x,\delta_o) \leqslant d(x,o) - \rho$.*

Απόδειξη του Λήμματος. Έστω $x \in X$ με $d(x,o) \geqslant r$. Αν επιπλέον $d(x,o) < 2r$, τότε η δεύτερη συνθήκη του θεωρήματος δίνει ότι $W_1(m_x,\delta_o) < d(x,o)$ και από τον ορισμό του ρ έχουμε ότι $\rho \leqslant W_1(m_x,\delta_o) - d(x,o)$, το οποίο δίνει το ζητούμενο. Υποθέτουμε, λοιπόν, ότι $d(x,o) \geqslant 2r$. O (X,d) είναι $r-$γεωδαισιακός, οπότε υπάρχει πεπερασμένη ακολουθία σημείων $0 = x_0, x_1, \ldots, x_n = x$, ώστε $d(x_i,x_{i+1}) \leqslant r$ και $\sum d(x_i,x_{i+1}) = d(x,o)$. Επιπλέον, μπορούμε να υποθέσουμε ότι $d(o,x_2) > r$ (σε αντίθετη περίπτωση αφαιρούμε το x_1). Θέτουμε $z = x_1$ αν $d(x_1,o) = r$, ή $z = x_2$ αν $d(x_1,o) < r$. Έτσι έχουμε ότι $r \leqslant d(z,o) < 2r$ και επομένως ισχύει ότι $W_1(m_z,\delta_o) \leqslant d(z,o) - \rho$. Επίσης η μη-αρνητική καμπυλότητα δίνει ότι $W_1(m_z,m_x) \leqslant d(z,x)$. Τελικά $W_1(m_x,\delta_o) \leqslant W_1(m_z,\delta_o) + W_1(m_z,m_x) \leqslant d(z,o) - \rho + d(z,x)$. Από την κατασκευή του z έχουμε ότι $d(z,o) + d(z,x) = d(x,o)$ και άρα έχουμε το ζητούμενο. $\square$

Απόδειξη του Θεωρήματος. Η ιδέα της απόδειξης είναι να χρησιμοποιήσουμε τη συνάρτηση $f(x) = e^{\lambda d(x,o)}$ (με μια μικρή τροποποίηση). Σε μακρινά σημεία από το o, λόγω του ότι υπό τον τυχαίο περίπατο η μέση απόσταση από το o μειώνεται κατά ρ (Λήμμα 2.9.2), αναμένουμε η συνάρτηση να πολλαπλασιαστεί με $e^{-\lambda\rho}$ υπό τον τελεστή του τυχαίου περίπατου. Κοντά στο o, η εξέλιξη της συνάρτησης υπό τον τυχαίο περίπατο, δηλαδή η μέση τιμή της ως προς m_x, ελέγχεται από την ποσότητα s^2 και το άλμα $J(o)$ του o (τρίτη υπόθεση). Επιπλέον το ολοκλήρωμα της συνάρτησης ως προς μια αναλλοίωτη κατανομή του τυχαίου περίπατου διατηρείται από τον τελεστή του τυχαίου περίπατου, και πολλαπλασιάζεται με μια ποσότητα μικρότερη της μονάδας σε μακρινά σημεία, και αυτό δείχνει ότι τα βάρη σημείων που απέχουν μακριά από το o δεν είναι πολύ μεγάλα.

Θεωρούμε τη συνάρτηση $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ με τύπο

$$\varphi(t) = \begin{cases} 0 & \text{αν } t < r \\ \frac{(t-r)^2}{qr} & \text{αν } r \leqslant t < r\left(\frac{q}{2}+1\right) \\ t - r - \frac{qr}{4} & \text{αν } t \geqslant r\left(\frac{q}{2}+1\right) \end{cases},$$

για κάποιο $q > 0$ που θα επιλέξουμε αργότερα. Παρατηρούμε ότι η φ είναι 1-Lipschitz, αύξουσα, και $\varphi'' \leqslant 2/qr$. Έστω Y θετική τυχαία μεταβλητή. Παίρνοντας ανάπτυγμα Taylor με υπόλοιπο Lagrange (όπως στην απόδειξη του Λήμματος 2.7.1) έχουμε

$$\mathbb{E}(\varphi(Y)) \leqslant \varphi(\mathbb{E}(Y)) + \frac{1}{2}\mathrm{Var}(Y)\sup\varphi'' \leqslant \varphi(\mathbb{E}(Y)) + \frac{1}{qr}\mathrm{Var}(Y).$$

Για $\lambda > 0$ θεωρούμε τη συνάρτηση $f(x) = e^{\lambda\varphi(d(x,o))}$, η οποία είναι 1-Lipschitz, και άρα η τρίτη υπόθεση δίνει ότι

$$Mf(x) \leqslant e^{\lambda^2 s^2/2} e^{\lambda M\varphi(d(x,o))}. \tag{2.9.1}$$

Η υπόθεση της Gaussian-type ανισότητας μετασχηματισμού Laplace δίνει εκτίμηση για τη διασπορά $\mathrm{Var_{m_x}(f) \leqslant s^2}$ για κάθε 1-Lipschitz συνάρτηση f: η ανισότητα

$$\mathbb{E}_{m_x}\left(e^{\lambda f}\right) \leqslant e^{\lambda^2 s^2/2} e^{\mathbb{E}_{m_x}(f)}$$

είναι ισοδύναμη με την

$$\mathbb{E}_{m_x}(e^{\lambda(f-\mathbb{E}(f))}) \leqslant e^{\lambda^2 s^2/2}$$

και συνεπώς αρκεί να δείξουμε το ζητούμενο για f με $\mathbb{E}_{m_x}(f)=0$. Σε αυτήν την περίπτωση έχουμε ότι $\mathbb{E}_{m_x}(e^{\lambda f}) \leqslant e^{\lambda^2 s^2/2}$ και από το θεώρημα Taylor στα δύο μέλη έχουμε ότι για κάθε $\lambda>0$

$$1+\lambda\mathbb{E}_{m_x}(f)+\frac{\lambda^2}{2}\mathbb{E}_{m_x}(f^2)+O(\lambda^3) \leqslant 1+\frac{\lambda^2 s^2}{2}+O(\lambda^3),$$

δηλαδή

$$\mathrm{Var}_{m_x}(f)+O(\lambda) \leqslant s^2+O(\lambda),$$

και για $\lambda \to 0$ έχουμε το ζητούμενο.

Συνεπώς, η παραπάνω εκτίμηση μαζί με την (2.9.1) δίνουν την $Mf(x) \leqslant \varphi\left(Md(x,o)\right)+\frac{s^2}{qr}$ και το ανάπτυγμα Taylor με υπόλοιπο Lagrange δίνει

$$M\varphi(x,o) \leqslant \varphi\left(Md(x,o)\right)+\frac{s^2}{qr} = \varphi\left(W_1(m_x,\delta_o)\right)+\frac{s^2}{qr}$$

και τελικά

$$Mf(x) \leqslant e^{\lambda^2 s^2/2+\lambda^2 s^2/qr} e^{\lambda\varphi(W_1(m_x,\delta_o))}.$$

Για τις διάφορες τιμές του $d(x,o)$ θα βρούμε διαφορετικά φράγματα για το $\varphi(W_1\left(m_x,\delta_o\right))$:

(i) Αν $d(x,o)<r$ έχουμε ότι $W_1(m_x,\delta_o) \leqslant W_1(m_x,m_o)+W_1(m_o,\delta_o) \leqslant d(x,o)+J(o) < r+J(o)$ και συνεπώς $\varphi\left(W_1(m_x,\delta_o)\right) \leqslant \varphi(r+J(o)) \leqslant J^2(o)/qr$, καθώς $r \leqslant J(o)+r \leqslant r\left(\frac{l}{2}+1\right)$. Τελικά έχουμε ότι

$$Mf(x) \leqslant \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}+\frac{\lambda J(o)^2}{qr}\right\} = \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}+\frac{\lambda J(o)^2}{qr}\right\} f(x),$$

αφού όταν $d(x,o)<r$ ισχύει ότι $f(x)=1$.

(ii) Αν $r \leqslant d(x,o)$ τότε το Λήμμα 2.9.2 δίνει ότι $W_1(m_x,\delta_o) \leqslant d(x,o)-\rho$ και συνεπώς

$$Mf(x) \leqslant \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}\right\} \exp\left\{\lambda\varphi(d(x,o)-\rho)\right\}.$$

Αν τώρα $d(x,o) \geqslant r\left(\frac{N}{2}+1\right)+\rho$ έχουμε ότι $\varphi(d(x,o)-\rho)=\varphi(d(x,o))-\rho$ και άρα

$$Mf(x) \leqslant \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}-\lambda\rho\right\} e^{\lambda\varphi(d(o,x))} = \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}-\lambda\rho\right\} f(x),$$

ενώ αν $r \leqslant d(x,o) \leqslant r\left(\frac{l}{2}+1\right)+\rho$, έχουμε ότι $\varphi(d(x,o)-\rho) \leqslant \varphi(d(x,o)$ και έτσι

$$Mf(x) \leqslant \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}\right\} e^{\lambda\varphi(d(x,o))} = \exp\left\{\frac{\lambda^2 s^2}{2}+\frac{\lambda s^2}{qr}\right\} f(x).$$

Έστω ν μέτρο πιθανότητας με $\int f d\nu < \infty$ και έστω $\widetilde{X} = B\left(o, r\left(\frac{q}{2}+1\right)\right)$. Ορίζουμε συναρτήσεις $A(\nu) = \int_{\widetilde{X}} f d\nu$ και $B(\nu) = \int_{X \setminus \widetilde{X}} f d\nu$. Οι εκτιμήσεις για το Mf που βρήκαμε παραπάνω μας δίνουν ότι

$$\begin{aligned} A(\nu * m) + B(\nu * m) &= \int f d(\nu * m) \\ &= \int_{\widetilde{X}} M f d\nu + \int_{X \setminus \widetilde{X}} M f d\nu \\ &\leqslant \exp\left\{\frac{\lambda^2 s^2}{2} + \frac{\lambda s^2}{qr} + \frac{\lambda J(o)^2}{qr}\right\} \int_{\widetilde{X}} f d\nu \\ &\quad + \exp\left\{\frac{\lambda^2 s^2}{2} + \frac{\lambda s^2}{qr} - \lambda\rho\right\} \int_{X \setminus \widetilde{X}} f d\nu \\ &= \alpha A(\nu) + \beta B(\nu) \end{aligned}$$

όπου $\alpha = \exp\left\{\frac{\lambda^2 s^2}{2} + \frac{\lambda s^2}{qr} + \frac{\lambda J(o)^2}{qr}\right\}$ και $\beta = \exp\left\{\frac{\lambda^2 s^2}{2} + \frac{\lambda s^2}{Nr} - \lambda\rho\right\}$. Επιλέγοντας $\lambda = \rho/s^2$ και $q = 4s^2/r\rho$, θα δείξουμε ότι $\beta < 1$. Αντικαθιστώντας τα λ και q στο β έχουμε ότι

$$\beta = \exp\left\{\frac{\rho^2}{2s^2} + \frac{s^2}{4} - \frac{\rho^2}{s^2}\right\} = \exp\left\{\frac{s^2}{4} - \frac{\rho^2}{s^2}\right\} = \exp\left\{\frac{s^2 - 4\rho^2}{s^2}\right\}$$

και συνεπώς $\beta < 1$ αν και μόνο αν $s^2/2 < \rho$. Αυτό όμως ισχύει λόγω των υποθέσεων μας. Τώρα, έχουμε ότι

$$A(\nu) = \int_{\widetilde{X}} f d\nu = \int_{\widetilde{X}} e^{\varphi(d(x,o))} d\nu \leqslant e^{\lambda qr/4},$$

και συνεπώς

$$\alpha A(\nu) + \beta B(\nu) \leqslant (\alpha - \beta) e^{\lambda qr/4} + \beta(A(\nu) + B(\nu)).$$

Αν υποθέσουμε ότι

$$A(\nu) + B(\nu) \leqslant \frac{(\alpha - \beta)e^{\lambda qr/4}}{1 - \beta},$$

παίρνουμε ότι

$$\alpha A(\nu) + \beta B(\nu) \leqslant \frac{(\alpha - \beta)e^{4\lambda qr/4}}{1 - \beta}.$$

Δηλαδή αν

$$\int f d\nu \leqslant \frac{(\alpha - \beta)e^{4\lambda qr/4}}{1 - \beta},$$

τότε

$$\int M f d\nu \leqslant \frac{(\alpha - \beta)e^{4\lambda qr/4}}{1 - \beta}.$$

Θέτουμε $R = \frac{(\alpha-\beta)e^{4\lambda qr/4}}{1-\beta}$ και έχουμε ότι το σύνολο $C = \{\nu : \int f d\nu \leqslant R\}$ παραμένει αναλλοίωτο από τον τυχαίο περίπατο m. Επιπλέον, αν $A(\nu) + B(\nu) > R$ τότε $\alpha A(\nu) + \beta B(\nu) < A(\nu) + B(\nu)$, καθώς αν αυτό δεν ίσχυε θα είχαμε

$$\begin{aligned}
R &< A(\nu) + B(\nu) \leqslant \alpha A(\nu) + \beta B(\nu) \\
&\leqslant (\alpha - \beta) e^{\lambda q r/4} + \beta(A(\nu) + B(\nu)) \\
&< (\alpha - \beta) e^{\lambda q r/4} + \beta R \\
&= (\alpha - \beta) e^{\lambda q r/4} + \beta \frac{(\alpha - \beta) e^{4\lambda q r/4}}{1 - \beta} \\
&= \frac{(\alpha - \beta) e^{4\lambda q r/4}}{1 - \beta} = R.
\end{aligned}$$

Αυτό σημαίνει ότι αν το ν είναι αναλλοίωτη κατανομή για τον τυχαίο περίπατο, δηλαδή $\int f d\nu = \int M f d\nu$, τότε $\nu \in C$ (αλλιώς, θα είχαμε $\int M f d\nu \leqslant \alpha A(\nu) + \beta B(\nu) < A(\nu) + B(\nu) = \int f d\nu$).

Είμαστε τώρα έτοιμοι να δείξουμε την ύπαρξη αναλλοίωτης κατανομής για τον τυχαίο περίπατο m. Είναι προφανές ότι το C είναι κλειστό και κυρτό σύνολο. Επιπλέον είναι tight, καθώς αν K συμπαγές και $K \subset B(y, \varepsilon)$ τότε $\nu(C \setminus K) \leqslant R e^{-\lambda \varepsilon}$. Από το θεώρημα του Prokhorov έχουμε ότι το C είναι συμπαγές στην τοπολογία της ασθενούς σύγκλισης. Επομένως το C είναι κυρτό και συμπαγές σύνολο στον τοπολογικό γραμμικό χώρο των προσημασμένων Borel μέτρων πιθανότητας στον X, και είναι αναλλοίωτο από τον τελεστή του τυχαίου περίπατου M, ο οποίος είναι affine απεικόνιση. Από το Θεώρημα Markov-Kakutani έχει σταθερό σημείο, δηλαδή υπάρχει $\nu \in C$ με $\int M f d\nu = \int f d\nu$. Μένει να εκτιμήσουμε το R για να βρούμε το φράγμα του θεωρήματος. Έχουμε

$$\begin{aligned}
R = \frac{\alpha/\beta - 1}{1/\beta - 1} e^{\lambda q r/4} &= \frac{e^{\lambda J(o)^2/qr + \lambda \rho} - 1}{e^{\lambda \rho - \lambda s^2/qr - \lambda^2 s^2/2} - 1} e^{\lambda q r/4} \\
&\leqslant \frac{\rho + J(o)^2/qr}{\rho - s^2/qr - \lambda s^2/2} e^{\lambda J(o)^2/qr + \lambda \rho + \lambda q r/4}
\end{aligned}$$

χρησιμοποιώντας τις ανισότητες $e^t - 1 \leqslant t e^t$ και $e^t - 1 \geqslant t$. Έχουμε $\lambda = \rho/s^2$ και $q = 4s^2/r\rho$ και επομένως

$$R \leqslant \left(4 + J(o)^2/s^2\right) e^{\lambda(s^2/\rho + \rho(1 + J(o)^2/4s^2))}.$$

Αν λοιπόν θεωρήσουμε μια αναλλοίωτη κατανομή ν έχουμε ότι $\int f d\nu \leqslant R$ και εφόσον $d(x, o) \leqslant \varphi(d(x, o)) + r(1 + q/4)$ έχουμε ότι

$$\begin{aligned}
\int e^{\lambda d(x,o)} d\nu &\leqslant e^{\lambda r(1+q/4)} \int f d\nu \leqslant R e^{\lambda r(1+q/4)} \\
&\leqslant \left(4 + J(o)^2/s^2\right) e^{\lambda(s^2/\rho + \rho(1 + J(o)^2/4s^2))} e^{\lambda r(1+q/4)} \\
&= \left(4 + J(o)^2/s^2\right) \exp\left\{ \frac{\rho}{s^2} \left(\frac{s^2}{\rho} + \rho \left(1 + \frac{J(o)^2}{4s^2} \right) + r + \frac{s^2}{\rho} \right) \right\} \\
&= \left(4 + J(o)^2/s^2\right) e^{m/D},
\end{aligned}$$

όπου $D = s^2/\rho$ και $m = r + 2s^2/\rho + \rho(1 + J(o)^2/4s^2)$. Τέλος έχουμε ότι για κάθε f 1-Lipschitz ισχύει ότι

$$\begin{aligned}\nu\left(|f-f(o)|\geqslant t+M\right)&=2\nu\left(f-f(o)\geqslant t+M\right)\\&=2\nu\left(e^{(f-f(o))/D}\geqslant e^{(t+M)/D}\right)\\&\leqslant 2e^{-(t+m)/D}\int e^{(f(x)-f(o)/D}d\nu(x)\\&\leqslant 2e^{-(t+m)/D}\int e^{d(x,o)/D}d\nu\\&\leqslant e^{-(t+m)/D}\left(8+2J(o)^2/s^2\right)e^{m/D}\\&=\left(8+2J(o)^2/s^2\right)e^{-t/D}\end{aligned}$$

και η απόδειξη ολοκληρώθηκε. □

2.10 Θεωρήματα Bonnet-Myers στον L^2

Γενικά δεν είναι εφικτό να δοθεί φράγμα για τη διάμετρο ενός χώρου θετικής καμπυλότητας, αντίστοιχο με εκείνο στην περίπτωση των πολλαπλοτήτων Riemann, το οποίο να είναι αντιστρόφως ανάλογο προς την τετραγωνική ρίζα της καμπυλότητας. Το απλούστερο παράδειγμα στο οποίο αυτό είναι αδύνατο να συμβεί είναι ο διακριτός κύβος. Σε αυτήν την παράγραφο περιγράφουμε επιπλέον συνθήκες οι οποίες μας δίνουν τέτοιο φράγμα σε δύο διαφορετικές καταστάσεις. Αρχικά θα δοθεί φράγμα της μέσης απόστασης μεταξύ δύο σημείων, αντί για τη διάμετρο. Θα χρειαστεί να υποθέσουμε την ύπαρξη «σημείου έλξης». Ύστερα θα δώσουμε μια γενίκευση του θεωρήματος Bonnet-Myers για πολλαπλότητες Riemann, αν και υπάρχει έλλειψη εφαρμογών σε συγκεκριμένα παραδείγματα.

Πρόταση 2.10.1 (L^2 Bonnet-Myers για μέσες αποστάσεις)**.** *Έστω (X,d,m) μετρικός χώρος με τυχαίο περίπατο και θετική καμπυλότητα τουλάχιστον $\kappa>0$. Έστω ότι για κάποιο $o\in X$ ισχύει ότι $W_1(m_x,\delta_o)\leqslant d(x,o)$ για κάθε $x\in X$ με $r\leqslant d(x,o)<2r$. Έστω επίσης ότι ο X είναι $r-$γεωδαισιακός. Τότε*

$$\int d(x,o)d\nu(x)\leqslant\sqrt{\frac{1}{\kappa}\int\frac{\sigma(x)^2}{n_x}d\nu(x)}+5r,$$

όπου ν η αναλλοίωτη κατανομή του τυχαίου περίπατου.

Απόδειξη. Ορίζουμε $\varphi:\mathbb{R}\to\mathbb{R}$ με

$$\varphi(t)=\begin{cases}0 & \text{αν } t\leqslant 2r\\(t-2r)^2 & \text{αλλιώς}\end{cases}.$$

Για κάθε τυχαία μεταβλητή Y ισχύει ότι

$$\mathbb{E}_{m_x}(\varphi(Y))\leqslant\varphi(\mathbb{E}_{m_x}(Y))+\frac{1}{2}\mathrm{Var}_{m_x}(Y)\sup\varphi''=\varphi(\mathbb{E}_{m_x}(Y))+\mathrm{Var}_{m_x}(Y).$$

Θεωρούμε τη συνάρτηση $f:X\to\mathbb{R}$ με $f(x)=\varphi(d(x,o))$. Θα δείξουμε ότι για κάθε $x\in X$ ισχύει ότι $Mf(x)\leqslant(1-\kappa)^2f(x)+\sigma(x)^2/n_x+9r^2$. Ύστερα, εφόσον ισχύει ότι $\int fd\nu=\int Mfd\nu$, θα έχουμε ότι

$$\int f d\nu \leqslant (1-\kappa)^2 f(x) + \sigma(x)^2/n_x + 9r^2$$

άρα

$$\int f d\nu \leqslant \frac{1}{\kappa(2-\kappa)} \int \frac{\sigma(x)^2}{n_x} d\nu(x) + 9r^2 \leqslant \frac{1}{\kappa} \int \frac{\sigma(x)^2}{n_x} d\nu(x) + 9r^2,$$

και αφού η φ είναι κυρτή συνάρτηση, από την ανισότητα Jensen θα έχουμε ότι

$$\int \varphi(d(x,o)) d\nu \geqslant \varphi\left(\int d(x,o) d\nu\right) = \left(\int d(x,o) - 2r\right)^2.$$

Συνδυάζοντας αυτά τα δύο παίρνουμε

$$\int d(o,x) d\nu(x) - 2r \leqslant \sqrt{\frac{1}{\kappa} \int \frac{\sigma(x)^2}{n_x} d\nu(x)} + 3r,$$

δηλαδή

$$\int d(o,x) d\nu(x) \leqslant \sqrt{\frac{1}{\kappa} \int \frac{\sigma(x)^2}{n_x} d\nu(x)} + 5r.$$

Μένει λοιπόν να δείξουμε ότι για κάθε $x \in X$ ισχύει ότι $Mf(x) \leqslant (1-\kappa)^2 f(x) + \sigma(x)^2/n_x + 9r^2$. Έστω $x \in X$. Αρχικά υποθέτουμε ότι $r \leqslant d(x,o) < 2r$. Τότε, από την υπόθεση έχουμε ότι $\int d(y,o) dm_x(y) \leqslant d(x,o)$ και συνεπώς $\varphi\left(\int d(y,o) dm_x(y)\right) \leqslant \varphi(d(x,o)) = 0$. Άρα $\varphi\left(\int d(y,o) dm_x(y)\right) = 0$ και έπεται ότι

$$Mf(x) = \int \varphi(d(y,o)) dm_x(y) \leqslant \varphi\left(\int d(y,o) dm_x(y)\right) + \frac{\sigma(x)^2}{n_x} = \frac{\sigma(x)^2}{n_x}.$$

Αν τώρα $d(x,o) \geqslant 2r$ χρησιμοποιούμε το γεγονός ότι ο X είναι $r-$γεωδαισιακός και βρίσκουμε $z \in X$ με $d(x,o) = d(x,z) + d(z,o)$ και $r \leqslant d(z,o) < 2r$ (όπως στην απόδειξη του Λήμματος 2.9.2). Τότε,

$$\begin{aligned} \int d(y,o) dm_x(y) &= W_1(m_x, \delta_o) \leqslant W_1(m_x, m_z) + W_1(m_z, \delta_o) \\ &\leqslant (1-\kappa) d(x,z) + W_1(m_z, \delta_o) \\ &\leqslant (1-\kappa) d(x,z) + d(z,o) \leqslant (1-\kappa) d(x,z) + 2\kappa r, \end{aligned}$$

άρα

$$\begin{aligned} Mf(x) &\leqslant \varphi\left(\int d(y,o) dm_x(y)\right) + \frac{\sigma(x)^2}{n_x} \\ &\leqslant ((1-\kappa) d(x,o) + 2\kappa r - 2r)^2 + \frac{\sigma(x)^2}{n_x} \\ &= (1-\kappa)^2 (d(x,o) - 2r)^2 + \frac{\sigma(x)^2}{n_x} \\ &= (1-\kappa)^2 \varphi(d(x,o)) + \frac{\sigma(x)^2}{n_x}. \end{aligned}$$

Τέλος, στην περίπτωση που $d(x,o) \leqslant r$ έχουμε ότι $\int d(y,o)dm_x(y) = W_1(m_x,\delta_o) \leqslant W_1(m_x,m_o)+ W_1(m_o,\delta_o) \leqslant (1-\kappa)d(x,o)+J(o) \leqslant J(o)+r(1-\kappa) \leqslant J(o)+r$. Θα φράξουμε το άλμα $J(o)$. Αν ισχύει ότι $X \subset B(o,r)$ τότε το ζητούμενο είναι προφανές, διότι $\int d(x,o)d\nu(x) \leqslant r$. Έστω λοιπόν ότι υπάρχει $y \in X$ με $d(y,o) \geqslant r$. Ο X είναι $r-$γεωδαισιακός, συνεπώς μπορούμε να υποθέσουμε ότι $d(y,o) \leqslant 2r$. Τότε,

$$\begin{aligned} J(o) &= W_1(m_o,\delta_o) \leqslant W_1(m_o,m_y)+W_1(m_y,\delta_o) \\ &\leqslant (1-\kappa)d(y,o)+W_1(m_y,\delta_o) \leqslant (1-\kappa)d(y,o)+d(y,o) \\ &\leqslant 2d(y,o) \leqslant 4r. \end{aligned}$$

Επομένως, $\int d(y,o)dm_x(y) \leqslant r+J(o) \leqslant 5r$ και έχουμε $\varphi\left(\int d(y,o)dm_x(y)\right) \leqslant 9r^2$, το οποίο δίνει ότι $Mf(x) \leqslant 9r^2 + \frac{\sigma(x)^2}{n_x}$. Αν συνδυάσουμε τις τρεις περιπτώσεις έχουμε ότι

$$Mf(x) \leqslant (1-\kappa)^2 f(x) + \frac{\sigma(x)^2}{n_x} + 9r^2.$$

□

Η επόμενη πρόταση βασίζεται στην αντίστοιχη ιδιότητα που ικανοποιεί η συνήθης κίνηση Brown σε μια πολλαπλότητα Riemann, η οποία εξασφαλίζει ένα θεώρημα τύπου Bonnet-Myers. Δυστυχώς, οι μόνοι χώροι που δίνουν τέτοιο παράδειγμα είναι οι πολλαπλότητες Riemann, συνεπώς η χρήση του φαίνεται να είναι περιορισμένη.

Ο ορισμός της καμπυλότητας που δώσαμε ελέγχει την απόσταση μεταφοράς μεταξύ δύο μέτρων, που σχετίζονται αντίστοιχα με δύο σημεία x και y, σε κάποια δοθείσα χρονική στιγμή t. Η συνθήκη που χρειαζόμαστε για να ισχύει η επόμενη πρόταση ελέγχει την απόσταση μεταφοράς δύο μέτρων, που σχετίζονται αντίστοιχα με δύο σημεία x και y, σε δύο δοθείσες χρονικές στιγμές t και s. Θα βασιστούμε σε αυτά που ισχύουν για Gaussian μέτρα στον $\mathbb{R}^N$. Για κάθε δύο $x,y \in \mathbb{R}^N$ και για $t,s>0$, έστω m_x^{*t} και m_y^{*s} οι νόμοι των συνήθων κινήσεων Brown οι οποίες ξεκινούν από τα x και y τις χρονικές στιγμές t και s αντίστοιχα. Για την L^2 απόσταση μεταφοράς μεταξύ των δύο μέτρων ισχύει

$$W_2(m_x^{*t}, m_y^{*s})^2 = d(x,y) + N\left(\sqrt{t}-\sqrt{s}\right)^2,$$

συνεπώς

$$W_1(m_x^{*t}, m_y^{*s}) \leqslant d(x,y) + \frac{N\left(\sqrt{t}-\sqrt{s}\right)^2}{2d(x,y)}.$$

Θα θεωρήσουμε ως δεδομένη αυτήν την ανισότητα και θα αντιγράψουμε την παραδοσιακή απόδειξη του θεωρήματος Bonnet-Myers.

Πρόταση 2.10.2 (ισχυρό θεώρημα Bonnet-Myers στον L^2)**.** *Έστω (X,d) μετρικός χώρος εφοδιασμένος με έναν τυχαίο περίπατο συνεχούς χρόνου m^{*t}. Υποθέτουμε επίσης ότι ο X είναι $\varepsilon-$γεωδαισιακός, και ότι υπάρχουν σταθερές $\kappa>0$, $C \geqslant 0$ ώστε για κάθε δύο αρκετά μικρά t,s και για κάθε δύο $x,y \in X$ με $\varepsilon \leqslant d(x,y) \leqslant 2\varepsilon$ να ισχύει η ανισότητα*

$$W_1(m_x^{*t}, m_y^{*s}) \leqslant e^{-\kappa \min\{t,s\}} d(x,y) + \frac{C\left(\sqrt{t}-\sqrt{s}\right)^2}{2d(x,y)}.$$

Επιπλέον υποθέτουμε ότι $\varepsilon \leqslant \frac{1}{2}\sqrt{C/2\kappa}$. *Τότε*

$$\operatorname{diam}(X) \leqslant \pi\sqrt{\frac{C}{2\kappa}}\left(1+\frac{4\varepsilon}{\sqrt{C/2\kappa}}\right).$$

Παρατήρηση 2.10.3. Όταν $t = s$, η ανισότητα της υπόθεσης γίνεται

$$W_1(m_x^{*t}, m_y^{*t}) \leqslant e^{-\kappa t} d(x,y),$$

και μπορεί να θεωρηθεί ως η εκδοχή της θετικής καμπυλότητας για τυχαίους περίπατους συνεχούς χρόνου. Η σταθερά C παίζει το ρόλο της σταθεράς διάχυσης. Η υπόθεση $d(x,y) \geqslant \varepsilon$ υπάρχει για να αποφύγουμε προβλήματα στη σύγκλιση της ποσότητας $\frac{C(\sqrt{t}-\sqrt{s})^2}{2d(x,y)}$ όταν το x τείνει στο y.

Απόδειξη. Έστω $x, y \in X$. Ο X είναι $\varepsilon-$γεωδαισιακός και συνεπώς, όπως στην απόδειξη του Λήμματος 2.9.2, μπορούμε να θεωρήσουμε ότι υπάρχει πεπερασμένη ακολουθία σημείων του X, $x = x_0, x_1, \ldots, x_{K-1}, x_K = y$ με $\varepsilon \leqslant d(x_i, x_{i+1} \leqslant 2\varepsilon$ για $i = 0, 1, \ldots, K-1$ και $\sum d(x_i, x_{i+1}) = d(x,y)$.

Θέτουμε $t_i = \eta \sin^2\left(\frac{\pi d(x,x_i)}{d(x,y)}\right)$ για κάποιο η που θα επιλέξουμε αργότερα. Έχουμε ότι $t_0 = t_K = 0$, επομένως

$$\begin{aligned} d(x,y) = W_1(\delta_x, \delta_y) &\leqslant \sum_{i=0}^{K-1} W_1\left(m_{x_i}^{*t_i}, m_{x_{i+1}}^{*t_{i+1}}\right) \\ &\leqslant \sum_{i=0}^{K-1}\left(e^{-\kappa\min\{t_i,t_{i+1}\}} d(x_i,x_{i+1}) + \frac{C\left(\sqrt{t_{i+1}}-\sqrt{t_i}\right)^2}{2d(x_i,x_{i+1})}\right) \end{aligned}$$

από την υπόθεση. Ισχύει τώρα ότι

$$|\sin\beta - \sin\alpha| = \left|2\sin\left(\frac{\beta-\alpha}{2}\right)\cos\left(\frac{\alpha+\beta}{2}\right)\right| \leqslant |\beta-\alpha|\left|\cos\left(\frac{\alpha+\beta}{2}\right)\right|,$$

το οποίο δίνει

$$\frac{C\left(\sqrt{t_{i+1}}-\sqrt{t_i}\right)^2}{2d(x_i,x_{i+1})} \leqslant \frac{C\eta\pi^2 d(x_i,x_{i+1})}{2d(x,y)^2}\cos^2\left(\pi\frac{d(x,x_i)+d(x,x_{i+1})}{2d(x_i,x_{i+1})}\right).$$

Για η αρκετά μικρό ισχύει ότι $e^{-\kappa\min\{t_i,t_{i+1}\}} = 1 - \kappa\min\{t_i,t_{i+1}\} + O(\eta^2)$ και παίρνουμε

$$\begin{aligned} d(x,y) \leqslant \sum_{i=0}^{K-1} & d(x_i,x_{i+1}) - \kappa\min\{t_i,t_{i+1}\} d(x_i,x_{i+1}) \\ &+ \frac{C\eta\pi^2 d(x_i,x_{i+1})}{2d(x,y)^2}\cos^2\left(\pi\frac{d(x,x_i)+d(x,x_{i+1})}{2d(x,y)}\right) + O(\eta^2). \end{aligned}$$

Οι όροι $\sum d(x_i,x_{i+1})\cos^2\left(\pi\frac{d(x,x_i)+d(x,x_{i+1})}{2d(x,y)}\right)$ και $\sum \min\{t_i,t_{i+1}\} d(x_i,x_{i+1})$ είναι κοντά στα ολοκληρώματα $d(x,y)\int_{[0,1]}\cos^2(\pi u)du$ και $d(x,y)\eta\int_{[0,1]}\sin^2(\pi u)du$ αντίστοιχα, με το σφάλμα του ολοκληρώματος Riemann να φράσσεται από $\pi\varepsilon/d(x,y)$. Συνεπώς καταλήγουμε στην

$$\begin{aligned} d(x,y) \leqslant\; & d(x,y) - \kappa\eta d(x,y)\left(\frac{1}{2} - \frac{\pi\varepsilon}{d(x,y)}\right) \\ &+ \frac{C\eta\pi^2}{2d(x,y)^2} d(x,y)\left(\frac{1}{2} + \frac{\pi\varepsilon}{d(x,y)}\right) + O(\eta^2), \end{aligned}$$

και για μικρό η,

$$d(x,y)^2 \leqslant \frac{C\pi^2}{2\kappa}\frac{1+2\pi\varepsilon/d(x,y)}{1-2\pi\varepsilon/d(x,y)}.$$

Τέλος, η υπόθεση $\varepsilon \leqslant \frac{1}{2}\sqrt{C/2\kappa}$ μας δίνει είτε ότι $d(x,y) \leqslant \pi\sqrt{C/2\kappa}$ ή ότι $2\pi\varepsilon/d(x,y) \leqslant 2\pi\varepsilon/\pi\sqrt{C/2\kappa} \leqslant 1/2$, οπότε και χρησιμοποιούμε την ανισότητα $(1+\alpha)/(1-\alpha) \leqslant 1+4\alpha$, για $\alpha \leqslant 1/2$, το οποίο δίνει το ζητούμενο. □

2.11 Καμπυλότητα Ricci και τοπολογία Gromov-Hausdorff

Σε αυτήν την παράγραφο θα μελετήσουμε τη συμπεριφορά της καμπυλότητας αφενός όταν αλλάζει η μορφή του μετρικού χώρου, δηλαδή η μετρική, και αφετέρου όταν χρησιμοποιούμε διαφορετικό τυχαίο περίπατο. Τελικά, για να καταφέρουμε να αλλάξουμε και τον ίδιο τον χώρο, θα ορίσουμε μια τοπολογία τύπου Gromov-Hausdorff για μετρικό χώρο με τυχαίο περίπατο.

Ορισμός 2.11.1. Έστω X μη κενό σύνολο και d, ρ μετρικές στον X. Οι d και ρ θα ονομάζονται **bi-Lipschitz ισοδύναμες** αν υπάρχει C-Lipschitz ομοιομορφισμός $f : (X,d) \to (X,\rho)$ με

$$\frac{1}{C}d(x,y) \leqslant \rho(f(x),f(y)) \leqslant Cd(x,y),$$

για κάθε $x, y \in X$.

Πρόταση 2.11.2 (αλλαγή μετρικής). *Έστω (X,d,m) μετρικός χώρος με τυχαίο περίπατο και έστω ρ μια μετρική στον X, η οποία είναι bi-Lipschitz ισοδύναμη με την d, με σταθερά $C \geqslant 1$. Υποθέτουμε ότι η coarse καμπυλότητα Ricci του τυχαίου περίπατου m στον (X,d) είναι τουλάχιστον κ. Τότε η coarse καμπυλότητα Ricci του τυχαίου περίπατου m στον (X,ρ) είναι τουλάχιστον κ', όπου $1-\kappa' = C^2(1-\kappa)$.*

Απόδειξη. Έστω $x, y \in X$ και ξ_{xy} το βέλτιστο ταίριασμα μεταξύ των m_x και m_y. Τότε, εφόσον η coarse καμπυλότητα Ricci του m στον (X,d) είναι τουλάχιστον κ, έχουμε ότι

$$1-\kappa \geqslant \frac{\int d(z,w)d\xi_{xy}(z,w)}{d(x,y)} \geqslant \frac{\int \rho(f(z),f(w))d\xi_{xy}(z,w)}{C^2\rho(f(x),f(y))},$$

και συνεπώς

$$1-\kappa' = C^2(1-\kappa) \geqslant \frac{\int \rho(f(z),f(w))d\xi_{xy}(z,w)}{C^2\rho(f(x),f(y))},$$

το οποίο δίνει ότι $\kappa'(f(x),f(y)) \geqslant \kappa'$, όπου $\kappa'(f(x),f(y))$ είναι η coarse καμπυλότητα Ricci του m στον (X,ρ) στα σημεία $f(x)$ και $f(y)$. Το γεγονός ότι η f είναι $1-1$ και επί δίνει το ζητούμενο. □

Έστω (X,d) μετρικός χώρος. Θεωρούμε τον χώρο των προσημασμένων μέτρων με μηδενική μάζα $\mathscr{P}_0(X) = \{\mu_+ - \mu_- : \mu_+, \mu_- \in \mathscr{P}(X)\}$. Στον χώρο αυτό θεωρούμε τη νόρμα

$$\|\mu_+ - \mu_-\| := \sup_{f\ 1-\text{Lipschitz}} \int f d(\mu_+ - \mu_-) = W_1(\mu_+,\mu_-).$$

Πρόταση 2.11.3 (αλλαγή τυχαίου περίπατου). *Έστω (X, d) μετρικός χώρος και m, m' δύο τυχαίοι περίπατοι στον X. Υποθέτουμε ότι η coarse καμπυλότητα Ricci του m είναι τουλάχιστον κ, και ότι η απεικόνιση $x \mapsto m_x - m'_x \in \mathscr{P}_0(X)$ είναι C-Lipschitz. Τότε η coarse καμπυλότητα Ricci του m' είναι τουλάχιστον $\kappa - C$.*

Απόδειξη. Έστω $x, y \in X$. Η $x \mapsto m_x - m'_x \in \mathscr{P}_0(X)$ είναι C-Lipschitz, συνεπώς έχουμε ότι

$$W_1(m_x, m_y) - W_1(m'_x, m'_y) \leqslant \|m_x - m'_x - (m_y - m'_y)\| \leqslant Cd(x, y).$$

Έπεται ότι

$$\kappa - C \leqslant 1 - \frac{W_1(m_x, m_y)}{d(x, y)} - C \leqslant 1 - \frac{W_1(m'_x, m'_y)}{d(x, y)},$$

το οποίο δίνει το ζητούμενο. □

Τώρα, θα ασχοληθούμε με το τι συμβαίνει όταν αλλάζει ο ίδιος ο χώρος. Γι' αυτό το λόγο, θα χρειαστεί να εισάγουμε μια γενίκευση της τοπολογίας Gromov-Hausdorff, λαμβάνοντας υπόψιν τον τυχαίο περίπατο.

Ορισμός 2.11.4 (τοπολογία Gromov-Hausdorff). Έστω (X, d_X) και (Y, d_Y) δύο μετρικοί χώροι. Θα λέμε ότι οι (X, d_X) και (Y, d_Y) βρίσκονται σε απόσταση Gromov-Hausdorff το πολύ $\varepsilon \in [0, +\infty]$ αν υπάρχουν ημι-μετρικός χώρος (Z, d_Z) και ισομετρικές εμφυτεύσεις των X και Y, $f_X : (X, d_X) \to (Z, d_Z)$ και $f_Y : (Y, d_Y) \to (Z, d_Z)$ αντίστοιχα, ώστε για κάθε $x \in X$ να υπάρχει $y \in Y$ με $d_Z(f_X(x), f_Y(y)) \leqslant \varepsilon$, και ομοίως για κάθε $y \in Y$.

Ορισμός 2.11.5 (τοπολογία Gromov-Hausdorff για μετρικούς χώρους με τυχαίο περίπατο). Έστω (X, d_X, m^X) και (Y, d_Y, m^Y) δύο μετρικοί χώροι. Έστω $\varepsilon \in [0, +\infty]$. Θα λέμε ότι οι (X, d_X, m^X) και (Y, d_Y, m^Y) βρίσκονται $\varepsilon-$ κοντά αν υπάρχουν ημι-μετρικός χώρος (Z, d_Z) και ισομετρικές εμφυτεύσεις των X και Y, $f_X : (X, d_X) \to (Z, d_Z)$ και $f_Y : (Y, d_Y) \to (Z, d_Z)$ αντίστοιχα, ώστε για κάθε $x \in X$ να υπάρχει $y \in Y$ με $d_Z(f_X(x), f_Y(y)) \leqslant \varepsilon$ και $W_1(f_X(m^X_x), f_Y(m^Y_y)) \leqslant 2\varepsilon$, και ομοίως για κάθε $y \in Y$.

Παρατήρηση 2.11.6. Η σχέση «$\varepsilon-$κοντά» ορίζει μια ημιμετρική στην κλάση των μετρικών χώρων με τυχαίο περίπατο. Η απόσταση δύο τέτοιων χώρων είναι το infimum πάνω από όλα τα $\varepsilon > 0$ που ικανοποιούν τον Ορισμό 2.11.5.

Ορισμός 2.11.7. Έστω $\left\{(X^N, d^N_X, m^{X,N})\right\}$ ακολουθία μετρικών χώρων με τυχαίο περίπατο και (X, d_X, m^X) μετρικός χώρος με τυχαίο περίπατο. Θα λέμε ότι η $\left\{(X^N, d^N_X, m^{X,N})\right\}$ τείνει στον (X, d_X, m^X) αν η ημι-απόσταση τους τείνει στο 0. Επιπλέον, θα λέμε ότι η ακολουθία (x^N) με $x^N \in X^N$ τείνει στο $x \in X$ αν τα x^N και το x ικανοποιούν τον Ορισμό 2.11.5.

Πρόταση 2.11.8 (συνέχεια coarse καμπυλότητας Ricci). *Έστω $\left\{(X^N, d^N_X, m^{X,N})\right\}$ ακολουθία μετρικών χώρων με τυχαίο περίπατο που συγκλίνει στον μετρικό χώρο με τυχαίο περίπατο (X, d_X, m^X). Έστω $x, y \in X$ και $(x^N, y^N) \in X^N \times X^N$ για κάθε $N \in \mathbb{N}$ μια ακολουθία που συγκλίνει στο (x, y). Τότε $\kappa(x^N, y^N) \to (x, y)$. Ειδικότερα, για κάθε $N \in \mathbb{N}$, ο $(X^N, d^N_X, m^{X,N})$ έχει coarse καμπυλότητα Ricci τουλάχιστον κ, τότε το ίδιο ισχύει και για τον (X, d_X, m^X).*

Απόδειξη. Για $x, y \in X$ έχουμε $\kappa(x,y) = 1 - \frac{W_1(m_x,m_y)}{d(x,y)}$ και ομοίως για κάθε $N \in \mathbb{N}$ και για $x^N, y^N \in X^N$, $\kappa^N(x^N, y^N) = 1 - \frac{W_1(m_x^N, m_y^N)}{d(x^N,y^N)}$. Η υπόθεση δίνει ότι $d(x^N, y^N) \to d(x,y)$ και $W_1(m_x^N, m_y^N) \to W_1(m_x, m_y)$ για $N \to \infty$. □

Τα παραπάνω μας λένε ότι το να έχει ένας μετρικός χώρος coarse καμπυλότητα Ricci τουλάχιστον κ είναι κλειστή ιδιότητα. Όμως ισχύει ότι η coarse καμπυλότητα Ricci ενός μετρικού χώρου (X, d_X, m) μπορεί να είναι μεγαλύτερη από το $\limsup$ των coarse καμπυλοτήτων Ricci των (X^N, d_X^N, m^N), μιας και σημεία του X^N, που συνεισφέρουν στην καμπυλότητα του X^N, μπορεί να τείνουν σε ένα κοινό σημείο του X. Συνεπώς η θετική καμπυλότητα δεν είναι ανοικτή ιδιότητα. Θα δώσουμε έναν καινούργιο ασθενέστερο ορισμό για την coarse καμπυλότητα Ricci, εμπνευσμένο από το πώς μπορεί να περάσει κανείς από δέντρα σε $\delta-$υπερβολικούς χώρους, ώστε να κάνουμε την ιδιότητα ανοικτή.

Ορισμός 2.11.9. Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο. Έστω $\delta \geqslant 0$. Η coarse καμπυλότητα Ricci μέχρι δ κατά μήκος των $x, y \in X$ είναι η μεγαλύτερη τιμή του $\kappa \leqslant 1$ για την οποία

$$W_1(m_x, m_y) \leqslant (1-\kappa)d(x,y) + \delta.$$

Από τον ορισμό έπεται εύκολα το εξής.

Πρόταση 2.11.10. *Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο, για τον οποίο η coarse καμπυλότητα Ricci μέχρι δ να είναι τουλάχιστον κ. Έστω $\delta' > 0$. Τότε υπάρχει περιοχή V_X του X, ώστε για κάθε $Y \in V_X$ να έχει coarse καμπυλότητα Ricci μέχρι $\delta + \delta'$ τουλάχιστον κ.*

ΚΕΦΑΛΑΙΟ 3

Εφαρμογές και παραδείγματα

3.1 Κλασικά παραδείγματα

Σε αυτήν την ενότητα θα αναφερθούμε σε κάποια παραδείγματα συγκεκριμένων χώρων και θα συγκρίνουμε τα αποτελέσματα της θεωρίας της coarse καμπυλότητας Ricci με τα κλασικά αποτελέσματα του κάθε παραδείγματος, όπου αυτό είναι εφικτό.

Παράδειγμα 3.1.1 (τυχαίος περίπατος με βήμα ε)**.** Έστω (X, d, μ) μετρικός χώρος με μέτρο. Υποθέτουμε ότι οι μπάλες του (X, d) έχουν πεπερασμένο μέτρο και ότι $\mathrm{supp}(\mu) = \mathrm{X}$. Για κάθε $\varepsilon > 0$ ορίζουμε τον τυχαίο περίπατο με βήμα ε στον X, ο οποίος ξεκινάει από κάποιο σημείο $x \in X$, να είναι η οικογένεια μέτρων πιθανότητας $m_x = \mu|_{B(x,\varepsilon)}/\mu(B(x, \varepsilon))$, δηλαδή από το x «πηδάμε» σε κάποιο σημείο της μπάλας $B(x, \varepsilon)$ με πιθανότητα ανάλογη προς το μέτρο της. Με αυτόν τον τρόπο μπορούμε να δημιουργήσουμε με φυσιολογικό τρόπο έναν τυχαίο περίπατο αν μας έχει δοθεί ένα αρχικό μέτρο για τον χώρο.

Περνάμε τώρα σε πιο συγκεκριμένα παραδείγματα.

Παράδειγμα 3.1.2 ($\mathbb{Z}^N$ και $\mathbb{R}^N$)**.** Έστω m ο απλός τυχαίος περίπατος στο $\mathbb{Z}^N$ εφοδιασμένο με τη μετρική του γραφήματος. Δηλαδή, έχουμε ότι $m_x(z) = 1/2N$ αν x, z γειτονικές κορυφές και 0 αλλιώς. Τότε για x, y γειτονικές κορυφές έχουμε ότι η coarse καμπυλότητα Ricci στα x, y είναι 0. Επιπλέον, εφόσον ο χώρος μας είναι γεωδαισιακός το ίδιο θα ισχύει για κάθε δύο σημεία του (Πρόταση 2.2.1). Πράγματι έχουμε ότι ένα ταίριασμα μεταξύ των x, y είναι το

$$\xi_{xy}(z, w) = \begin{cases} 1/2N & \text{αν } x = x \pm i, y = y \pm i \\ 0 & \text{αλλιώς} \end{cases}$$

όπου $i = (0, 0, \ldots, 1, \ldots, 0)$ με το 1 να βρίσκεται στην θέση i, και συνεπώς

$$\begin{aligned} W_1(m_x, m_y) &\leqslant \int_{z,w} d(z,w) d\xi_{xy}(z,w) \\ &= \sum_{z,w} d(z,w)\xi_{xy}(z,w) \\ &= \sum_{i=1}^{N} d(x \pm i, y \pm i)\frac{1}{2N} \\ &= \frac{1}{2N}\sum_{i=1}^{N} d(x,y) = d(x,y). \end{aligned}$$

Επιπλέον έχουμε ότι $W_1(m_x, m_y) = \sup\left\{\left|\int f dm_x - \int g dm_y\right|\right\}$, με το supremum να παίρνεται πάνω από όλες τις $f \in L^1(m_x))$, $g \in L^1(m_y)$ με $f(z) + g(w) \leqslant d(z,w)$ και συνεπώς για $f(z) = d(x,z)$ και $g(w) = d(y,w)$ έχουμε ότι $W_1(m_x, m_y) \geqslant d(x, x^i) + d(y, y^i) = 2 \geqslant 1 = d(x,y)$.

Ομοίως μπορούμε να δούμε ότι αυτό γενικεύεται στις περιπτώσεις των $\mathbb{Z}^N$ και $\mathbb{R}^N$ αν τους θεωρήσουμε εφοδιασμένους με οποιαδήποτε μετρική η οποία είναι αναλλοίωτη στις μεταφορές.

Παράδειγμα 3.1.3 (πολλαπλότητες Riemann)**.** Έστω (X,d) μια λεία πολλαπλότητα Riemann διάστασης N. Έστω $\varepsilon > 0$ και m^ε η αλυσίδα Markov που ορίζει ο τύπος

$$dm_x^\varepsilon(y) = \frac{1}{\mathrm{vol}\,(B(x,\varepsilon))} d\mathrm{vol}(y),$$

αν $y \in B(x,\varepsilon)$ και 0 αλλιώς. Έστω $x \in X$ και v ένα μοναδιαίο διάνυσμα στο x. Θεωρούμε y στη γεωδαισιακή που ορίζει το v, με $d(x,y)$ αρκετά μικρή. Τότε ισχύει ότι

$$\kappa(x,y) = \frac{\varepsilon^2 \mathrm{Ric}(\mathrm{v},\mathrm{v})}{2(N+2)} + O\left(\varepsilon^3 + \varepsilon^2 d(x,y)\right).$$

Έτσι, βλέπουμε ότι με τον ορισμό της coarse καμπυλότητας Ricci που δώσαμε πετυχαίνουμε μια σωστή τιμή σε σχέση με τη συνήθη καμπυλότητα Ricci μιας πολλαπλότητας Riemann και συνεπώς «κερδίζουμε» το δικαίωμα να μιλάμε για γενίκευση της έννοιας της καμπυλότητας Ricci.

Το μέτρο $d\nu(x) = \frac{B(x,\varepsilon)}{B_{Eucl}(\varepsilon)} d\mathrm{vol}(x)$ είναι αντιστρέψιμο για τον τυχαίο περίπατο. Έστω επίσης $\inf \mathrm{Ric}$ ο μεγαλύτερος $K > 0$ ώστε $\mathrm{Ric}(v,v) \geqslant K$ για κάθε v εφαπτόμενο διάνυσμα. Ισχύουν τα εξής:

- Για κάθε $x \in X$ $\sigma(x)^2 \sim \varepsilon^2 \frac{N}{N+2}$.
- $n \approx N$.
- Η Gaussian διασπορά του Θεωρήματος 2.7.2 είναι $D^2 \approx 1/\inf \mathrm{Ric}$.
- Το Gaussian εύρος $t_{\max} \approx 1/(\varepsilon \inf \mathrm{Ric}) \to \infty$.

Παράδειγμα 3.1.4 (διακριτός κύβος)**.** Έστω $X = \{0,1\}^N$ ο διακριτός κύβος εφοδιασμένος με τη μετρική Hamming. Έστω m ο τυχαίος περίπατος με $m_x(x) = 1/2$ και $m_x(x^i) = 1/2N$, όπου x^i

είναι η γειτονική κορυφή του x με αλλαγμένη την συντεταγμένη i. Τότε για $x, y \in X$ έχουμε ότι $\kappa(x,y) = 1/N$. Έστω $x = (0,0,\dots,0), y = (1,0,\dots,0) \in X$. Θεωρούμε το ταίριασμα

$$\xi_{xy}(z,w) = \begin{cases} \frac{1}{2} - \frac{1}{2N} & \text{αν } z = x, w = y \\ \frac{1}{2N} & \text{αν } z = x, w = y^1 \\ \frac{1}{2N} & \text{αν } z = x^1, w = y \\ \frac{1}{2N} & \text{αν } z = x^i, w = y^i \text{ για } i \geqslant 2 \end{cases}$$

Το ξ_{xy} είναι πράγματι ταίριασμα των m_x, m_y γιατί έχουμε ότι:

$$\sum_w \xi_{xy}(x,w) = \xi_{xy}(x,y) + \xi_{xy}(x,y^1) = \frac{1}{2N}$$

και

$$\sum_w \xi_{xy}(x^i,w) = \frac{1}{2N}, \text{ για κάθε } i = 1,2,\dots,N,$$

το οποίο δίνει ότι $\xi_{xy}(A \times X) = m_x(A)$ και ομοίως έχουμε ότι $\xi_{xy}(X \times B) = m_y(B)$. Συνεπώς,

$$\begin{aligned} W_1(m_x,m_y) &\leqslant \sum_{z,w} d(z,w)\xi_{xy}(z,w) \\ &= d(x,y)\left(\frac{1}{2} - \frac{1}{2N}\right) + d(x,y^1)\frac{1}{2N} + d(x^1,y)\frac{1}{2N} + \sum_{i=2}^{N} d(x^i,y^i)\frac{1}{2N} \\ &= \frac{1}{2} - \frac{1}{2N} + \frac{N-1}{2N} = 1 - \frac{1}{N}, \end{aligned}$$

καθώς έχουμε ότι $x^1 = y$, $y^1 = x$ και x^i, y^i γειτονικές για κάθε $i \geqslant 2$. Τελικά έχουμε

$$\kappa(x,y) = 1 - \frac{W_1(m_x,m_y)}{d(x,y)} \geqslant 1 - \left(1 - \frac{1}{N}\right) = \frac{1}{N}.$$

Για την αντίστροφη ανισότητα θεωρούμε την $f : X \to \{0,1\}$ με $f(x_1,\dots,x_N) = x_1$. Αυτή είναι 1-Lipschitz και ισχύουν οι

$$\sum_z f(z)m_x(z) = f(x)m_x(x) + \sum_{i=1}^{N} f(x^i)m_x(x^i) = m_x(x^1) = \frac{1}{2N}$$

και

$$\sum_w f(w)m_y(w) = f(y)m_y(y) + \sum_{i=1}^{N} f(y^i)m_y(y^i) = \frac{1}{2} + \frac{N-1}{2N} = 1 - \frac{1}{2N}.$$

Επομένως, από το θεώρημα του Kantorovich έχουμε ότι

$$W_1(m_x,m_y) \geqslant \left|\sum_z f(z)m_x(z) - \sum_w f(w)m_y(w)\right| = \left|\frac{1}{2N} - 1 + \frac{1}{2N}\right| = 1 - \frac{1}{N}.$$

Αυτό δίνει την αντίστροφη ανισότητα $\kappa(x,y) \leqslant 1/N$.

Μια συντομότερη απόδειξη μπορεί να δοθεί αν χρησιμοποιήσουμε την Πρόταση 2.5.1: έστω $x = (x_1, x_2, \ldots, x_k)$, $x_1, x_2, \ldots, x_k \in \{0, 1\}$ και m^i ο τυχαίος περίπατος με $m_{x_i}(0) = m_{x_i}(1) = 1/2$. Τότε, η coarse καμπυλότητα Ricci του m^i είναι $\kappa^i(x, y) = 1$. Επίσης,

$$m_{(x_1,x_2,\ldots,x_k)} = \sum_{i=1}^{k} \alpha_i \delta_{x_1} \otimes \delta_{x_2} \otimes \ldots \otimes m_{x_i} \otimes \ldots \otimes \delta_k,$$

με $\alpha_i = 1/N$ και τελικά έχουμε ότι $\kappa(x, y) = \min \alpha_i \kappa_i = 1/N$.

Για κάθε γράφημα και για κάθε τυχαίο περίπατο m στο γράφημα ισχύει ότι $\sigma^2(x)/n_x \leqslant 1 - m_x(\{x\})$. Πράγματι, έχουμε ότι $\sigma^2(x)n_x = \sup\{\mathrm{Var}_{m_x}(f)\}$ πάνω από τις 1-Lipschitz συναρτήσεις f. Αν θεωρήσουμε μια 1-Lipschitz συνάρτηση f, και υποθέσουμε ότι $f(x) = 0$ (μπορούμε καθώς η διασπορά παραμένει αναλλοίωτη αν προσθέσουμε μια σταθερά) έχουμε ότι $|f(y)| \leqslant 1$ για y γειτονική κορυφή της x και $m_x(\{y : \ y \sim x|\}) = 1 - m_x(\{x\})$, συνεπώς $\mathrm{Var}_{m_x}(f) = \mathbb{E}_{m_x}(f^2) - (\mathbb{E}_{m_x}(f))^2 \leqslant \mathbb{E}_{m_x}(f^2) \leqslant 1 - m_x(\{x\})$. Στον διακριτό κύβο ισχύουν επομένως τα εξής:

- Για κάθε $x \in X$, $\sigma^2(x)/n_x \leqslant 1/2$.
- Η προσέγγιση της διασποράς της Πρότασης 2.6.5 είναι $\sigma^2/n\kappa(2-\kappa) \sim N/4$, όσο δηλαδή και η πραγματική διασπορά.
- Η Gaussian διασπορά είναι $D^2 \leqslant N/2$.
- Το Gaussian εύρος είναι $t_{max} = N/2$.

Παράδειγμα 3.1.5 (διαδικασία Ornstein-Uhlenbeck)**.** Έστω $s \geqslant 0$, $\alpha > 0$ και θεωρούμε τη διαδικασία Ornstein-Uhlenbeck στον $\mathbb{R}^N$ η οποία δίνεται από τη στοχαστική διαφορική εξίσωση

$$dX_t = -\alpha X_t dt + s dB_t,$$

όπου B_t η συνήθης N-διάστατη κίνηση Brown. Η αναλλοίωτη κατανομή είναι Gaussian με διασπορά $s^2/2\alpha$. Έστω $t > 0$ και m ο τυχαίος περίπατος που ακολουθεί τη στοχαστική διαδικασία για χρόνο t. Δηλαδή το m_x είναι ένα Gaussian μέτρο πιθανότητας με κέντρο $e^{-\alpha t}x$ και διασπορά $s^2\left(1 - e^{-2\alpha t}\right)/2\alpha$. Τότε η coarse καμπυλότητα Ricci $\kappa(x, y)$ του m είναι $1 - e^{-\alpha t}$ για κάθε $x, y \in \mathbb{R}^N$.

Αρκεί να δείξουμε ότι $W_1(m_x, m_y) = e^{-\alpha t}|x - y|$. Έχουμε ότι

$$m_x(A) = \frac{1}{(2\pi)^N} \int_A e^{-\|z - e^{-t}x\|^2/2} dz$$

και για κάθε f 1-Lipschitz ισχύει ότι

$$\begin{aligned}
W_1(m_x, m_y) &\geqslant \left| \int f(z) m_x(z) - \int f(w) dm_y(w) \right| \\
&= \frac{1}{(2\pi)^N} \left| \int f(z) e^{\|z - e^{-t}x\|^2} dz - \int f(w) e^{\|w - e^{-t}y\|^2} dw \right| \\
&= \frac{1}{(2\pi)^N} \left| \int \left(f(z) - f(z + e^{-t}y - e^{-t}x)\right) e^{\|z - e^{-t}x\|^2} dz \right|,
\end{aligned}$$

συνεπώς για $f(z) = z$ έχουμε $W_1(m_x, m_y) \geqslant e^{-t}\|x - y\|$. Για την αντίστροφη ανισότητα θα χρησιμοποιήσουμε το γεγονός ότι $W_1(m_x, m_y) \leqslant W_2(m_x, m_y)$ και ότι αν μ, ν είναι Gaussian μέτρα με μέσες τιμές μ_1 και μ_2 και πίνακες συνδιακυμάνσεων Σ_1 και Σ_2 αντίστοιχα, τότε

$$W_2^2(\mu, \nu) = \|\mu_1 - \mu_2\|^2 + \mathrm{Tr}\left(\Sigma_1 + \Sigma_2 - 2\left(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2}\right)\right)^{1/2}.$$

Τα m_x και m_y έχουν την ίδια διασπορά, συνεπώς έχουμε ότι $W_2(m_x, m_y) = e^{-t}\|x-y\|$ και επομένως παίρνουμε την αντίστροφη ανισότητα

$$W_1(m_x, m_y) \leqslant W_2(m_x, m_y) = e^{-t}\|x - y\|.$$

Παράδειγμα 3.1.6 (πολυωνυμική κατανομή)**.** Θεωρούμε τον χώρο $X = \{(x_0, x_1, \ldots, x_d) : x_i \in \mathbb{N},\ \sum x_i = N\}$, τον οποίο βλέπουμε σαν τον χώρο που περιγράφει N το πλήθος σφαιρίδια σε $d+1$ κάλπες. Θεωρούμε τη στοχαστική διαδικασία, κατά την οποία επιλέγουμε ένα από τα N το πλήθος σφαιρίδια τυχαία, το αφαιρούμε από την κάλπη που βρίσκεται και το τοποθετούμε τυχαία σε μία από τις $d+1$ κάλπες. Δηλαδή έχουμε ότι η πιθανότητα μετάβασης από το $(x_0, x_1, \ldots, x_d)$ στο $(x_0, \ldots, x_i - 1, \ldots, x_j + 1, \ldots, x_d)$ είναι $x_i/N(d+1)$. Η πολυωνυμική κατανομή

$$\nu((x_0, x_1, \ldots, x_d)) = \frac{N!}{(d+1)^N \prod x_i!}$$

είναι αντιστρέψιμη γι' αυτήν την αλυσίδα Markov. Αν επιπλέον, για τον X, θεωρήσουμε τη μετρική $d((x_i), (y_i)) = \frac{1}{2}\sum |x_i - y_i|$ (δηλαδή τη μετρική του γραφήματος που δημιουργουν οι παραπάνω κινήσεις) θα έχουμε ότι η coarse καμπυλότητα Ricci του m είναι $1/N$.

Παράδειγμα 3.1.7 (γεωμετρική κατανομή)**.** Θεωρούμε τον τυχαίο περίπατο στο σύνολο των φυσικών αριθμών (μαζί με το 0) με πιθανότητες μετάβασης $p_n(n+1) = 1/3$, $p_{n+1}(n) = 2/3$ και $p_0(0) = 2/3$). Η γεωμετρική κατανομή $\nu(n) = 2^{-(n+1)}$ είναι αντιστρέψιμη για τον τυχαίο περίπατο και για $n \geqslant 1$ ισχύει $\kappa(n, n+1) = 0$.

Επίσης αν $0 < \alpha < 1$ θεωρούμε τον τυχαίο περίπατο με πιθανότητες μετάβασης $p_n(0) = \alpha$, $p_n(n+1) = 1 - \alpha$. Τότε η γεωμετρική κατανομή $\nu(n) = \alpha(1-\alpha)^n$ είναι αναλλοίωτη, αλλά όχι αντιστρέψιμη, για τον τυχαίο περίπατο, και $\kappa(n, m) = \alpha$.

Παράδειγμα 3.1.8 (δ-υπερβολικές ομάδες)**.** Έστω X το γράφημα Cayley μιας δ-υπερβολικής ομάδας που παράγεται από κάποιο πεπερασμένο σύνολο γεννητόρων. Θεωρούμε $N \in \mathbb{N}$ που εξαρτάται από την ομάδα και θεωρούμε τον τυχαίο περίπατο που αποτελείται από N βήματα του απλού τυχαίου περίπατου στο γράφημα. Τότε, $\kappa(x, y) = -2N/d(x, y)(1 + 0(1))$ όταν τα $d(x, y)$ και k τείνουν στο άπειρο.

Αν θεωρήσουμε $z \in B(x, k)$ και $w \in B(y, k)$ έχουμε την σχέση $d(z, w) = d(x, y) + d(x, z) + d(y, w) - (y, z)_x - (x, w)_y$, όπου $(\alpha, \beta)_\gamma = \frac{1}{2}\left(d(\gamma, \alpha) + d(\beta, \gamma) - d(\alpha, \beta)\right)$ είναι το γινόμενο Gromov των α και β με βάση το γ. Αυτό καταλήγει σε ένα άθροισμα ενός όρου που εξαρτάται μόνο από το z και ενός δεύτερου όρου που εξαρτάται μονάχα από το w. Συνεπώς, για να υπολογίσουμε την απόσταση Wasserstein των m_x και m_y αρκεί να υπολογίσουμε τη μέση τιμή του $(y, z)_x$ για z στην μπάλα $B(x, k)$ και ομοίως για το $(x, w)_y$. Ισχύει ότι οι μπάλες μεγαλώνουν με εκθετικό ρυθμό (Πρόταση 21 στο [Oll04], και επομένως έχουμε ότι οι μέσες τιμές που ψάχνουμε είναι φραγμένες από κάποια σταθερά, το οποίο δίνει το ζητούμενο.

Παράδειγμα 3.1.9 (διωνυμική κατανομή)**.** Έστω $X = \{0,1\}^N$ με μετρική την Hamming μετρική και θεωρούμε την εξής αλυσίδα Markov στον X, για κάθε $x \in X$: για κάποιο $p \in (0,1)$, σε κάθε βήμα, επιλέγουμε τυχαία μία από τις N το πλήθος συντεταγμένες του x. Αν ισούται με 0, την αλλάζουμε σε 1 με πιθανότητα p. Αλλιώς, αν ισούται με 1 την αλλάζουμε με 0 με πιθανότητα $1-p$. Η διωνυμική κατανομή

$$\nu((x_1,\dots,x_N)) = \prod_{i=1}^{N} p^{x_i}(1-p)^{1-x_i}$$

είναι αντιστρέψιμη γι' αυτόν τον τυχαίο περίπατο.

Η Πρόταση 2.5.3 μας δίνει ότι η coarse καμπυλότητα Ricci του τυχαίου περίπατου είναι $1/N$.

Έστω k το πλήθος των συντεταγμένων του x που ισούνται με 1. Τότε ο k ακολουθεί μια αλυσίδα Markov στο $\{0,1,\dots,N\}$ με πιθανότητες μετάβασης

$$\begin{aligned} p_k(k-1) &= p(1-k/N), \\ p_k(k-1) &= (1-p)k/N, \\ p_k(k) &= pk/N + (1-p)(1-k/N). \end{aligned}$$

Η διωνυμική κατανομή

$$\mu(N,k) = \binom{N}{k} p^k (1-p)^{N-k}$$

είναι αντιστρέψιμη γι' αυτήν την αλυσίδα Markov, και μάλιστα η καμπυλότητα Ricci παραμένει $1/N$.

Αν τώρα σταθεροποιήσουμε κάποιο $\lambda > 0$ και θεωρήσουμε $p = \lambda/N$, για $N \to \infty$ γνωρίζουμε ότι η αναλλοίωτη κατανομή του τυχαίου περίπατου τείνει σε κατανομή Poisson με σταθερά λ. Ισχύουν τα εξής:

- Η σταθερά διάχυσης του τυχαίου περιπάτου είναι $\sigma^2(k) = (\lambda + k)/N + O(1/N^2)$.
- Η εκτίμηση της μέσης τιμής της Πρότασης 2.4.3 είναι $\mathbb{E}(k) \leqslant L(0)/\kappa = \lambda$, με την μέση τιμή να είναι ακριβώς λ.
- Η μέση σταθερά διάχυσης είναι $\sigma^2 = \mathbb{E}(\sigma^2(k) = 2\lambda/N + O(1/N^2$.
- Η εκτίμηση της διασποράς της Πρότασης 2.6.5 είναι $\sigma^2/n\kappa(2-\kappa) \leqslant \lambda + O(1/N)$, με τη διασπορά να ισούται με λ.
- Η Gaussian διασπορά είναι $D^2 \leqslant 2\lambda + O(1/N)$.
- Η σταθερά Lipschitz της D_x^2 του θεωρήματος 2.7.2 είναι $C = 1 + O(1/N)$.
- Το Gaussian εύρος είναι $t_{\max} = 4\lambda/3$.

3.2 Coarse καμπυλότητα Ricci σε γραφήματα

Ορισμός 3.2.1 (γράφημα)**.** Ένα γράφημα G είναι μια τριάδα (V, w, m), όπου V είναι ένα αριθμήσιμο σύνολο, $w : V \times V \to [0, +\infty)$ μια συμμετρική συνάρτηση, η οποία μηδενίζεται στη διαγώνιο, και $m : V \to (0, +\infty)$. Το σύνολο V λέγεται το σύνολο των κορυφών, η w το βάρος των ακμών, και η m το μέτρο των κορυφών. Θα γράφουμε $x \sim y$ αν $w(x,y) > 0$ και θα λέμε ότι το ζεύγος $\{x,y\}$ είναι μια ακμή του γραφήματος. Το γράφημα θα λέγεται συνδυαστικό αν $w(x,y) \in \{0,1\}$, για κάθε $x, y \in V$, και $m = 1$.

Ορισμός 3.2.2 (βαθμός κορυφής)**.** Ορίζουμε τη συνάρτηση $\mathrm{Deg} : V \to \mathbb{R}$ με

$$\mathrm{Deg}(x) := \frac{1}{m(x)} \sum_{y \in V} w(x,y).$$

Το $\mathrm{Deg}(x)$ θα λέγεται ο βαθμός της κορυφής x.

Ορίζουμε τους χώρους:

$$\begin{aligned} C(V) &:= \{f : V \to \mathbb{R}\} = \mathbb{R}^V, \\ l_\infty(V) &:= \{f \in C(V) : f \text{ φραγμένη}\}, \\ C_c(V) &:= \{f \in C(V) : \text{το στήριγμα της } f \text{ είναι πεπερασμένο}\}. \end{aligned}$$

Ορισμός 3.2.3 (Λαπλασιανός τελεστής)**.** Ο Λαπλασιανός τελεστής του γραφήματος G είναι ο τελεστής $\Delta : C(V) \to C(V)$ με

$$\Delta f(x) := \frac{1}{m(x)} \sum_{y \sim x} w(x,y) \left(f(y) - f(x)\right),$$

για κάθε $f \in C(V)$.

Ορισμός 3.2.4 (ημιομάδα θερμότητας)**.** Η εξίσωση

$$\begin{aligned} \Delta u(x,t) &= \partial_t u(x,t) \; x \in V, t \geqslant 0 \\ u(x,0) &= f(x) \; x \in V \end{aligned}$$

είναι η εξίσωση της θερμότητας και θα συμβολίζουμε με $P_t f$ τη μικρότερη μη-αρνητική, φραγμένη και συνεχή λύση $u(x,t)$ της εξίσωσης. Η $\{P_t\}_t$ θα λέγεται ημιομάδα θερμότητας.

Ορισμός 3.2.5 (ε-lazy τυχαίος περίπατος)**.** Έστω $\varepsilon > 0$. Ο ε-lazy τυχαίος περίπατος είναι η οικογένεια μέτρων $\{m_x^\varepsilon\}_{x \in V}$ με

$$m_x^\varepsilon : \begin{cases} 1 - \varepsilon \mathrm{Deg}(x) & \text{αν } y = x \\ \varepsilon w(x,y)/m(x) & \text{αλλιώς} \end{cases}.$$

Ορισμός 3.2.6 (τελεστής καμπυλότητας Ricci)**.** Έστω $\Gamma : C(V) \times C(V) \to C(V)$ ο διγραμμικός τελεστής:

$$\Gamma(f,g)(x) = \frac{1}{2} \left(\Delta(f(x)g(x)) - f(x)\Delta g(x) - g(x)\Delta f(x)\right).$$

Ο τελεστής καμπυλότητας Ricci ορίζεται ως ο διγραμμικός τελεστής $R : C(V) \times C(V) \to C(V)$ με

$$R(f,g)(x) = \frac{1}{2} \left(\Delta\Gamma(f,g)(x) - \Gamma(f,\Delta g)(x) - \Gamma(g,\Delta f)(x)\right).$$

Ορισμός 3.2.7 (ανισότητα καμπυλότητας - διάστασης)**.** Ο Λαπλασιανός τελεστής Δ ικανοποιεί την ανισότητα καμπυλότητας - διάστασης $CD(m,K)(m \in (1,+\infty]$ (ο συμβολισμός οφείλεται στους Bakry και Émery) αν:

$$R(f,g)(x) \geqslant \frac{1}{m}\Big((\Delta f(x))^2 + k(x)\Gamma(f,f)(x)\Big).$$

Το m λέγεται διάσταση του τελεστή Δ και η συνάρτηση k κάτω φράγμα της καμπυλότητας Ricci του τελεστή Δ.

Θα ασχοληθούμε με την coarse καμπυλότητα Ricci για lazy τυχαίους περίπατους σε γραφήματα. Παρακάτω θα θεωρούμε γραφήματα με $w : V \times V \to \{0,1\}$ και $m(x) = d_x := \#\{y \in V : y \sim x\}$, $\alpha := 1 - \varepsilon$, οπότε και θα έχουμε:

$$m_x^\alpha(y) = \begin{cases} \alpha & \text{αν } y = x \\ (1-\alpha)/d_x & \text{αν } y \sim x \\ 0 & \text{αλλιώς} \end{cases}$$

Λήμμα 3.2.8. *Έστω $x, y \in V$. Η coarse καμπυλότητα Ricci κ_α είναι κοίλη συνάρτηση του α.*

Απόδειξη. Έστω $0 \leqslant \alpha \leqslant \beta \leqslant \gamma \leqslant 1$, και θέτουμε $\lambda = (\gamma-\beta)/(\gamma-\alpha)$. Τότε $\beta = \lambda\alpha + (1-\lambda)\gamma$. Με αυτούς τους ορισμούς η κ^α θα είναι κοίλη αν

$$\kappa_\beta \geqslant \lambda\kappa_\alpha + (1-\lambda)\kappa_\gamma.$$

Έστω ξ_{xy}^α το βέλτιστο ταίριασμα μεταξύ των m_x^α και m_y^α, το οποίο επιτυγχάνει την coarse καμπυλότητα Ricci κ_α. Ομοίως, έστω ξ_{xy}^γ το βέλτιστο ταίριασμα μεταξύ των m_x^γ και m_y^γ, το οποίο επιτυγχάνει την coarse καμπυλότητα Ricci κ_γ. Συνεπώς, ισχύουν οι σχέσεις:

$$W(m_x^\alpha, m_y^\alpha) = \sum_{z,w\in V} \xi_{xy}^\alpha(z,w)d(z,w),$$

$$W(m_x^\gamma, m_y^\gamma) = \sum_{z,w\in V} \xi_{xy}^\gamma(z,w)d(z,w).$$

Θέτουμε $\xi_{xy}^\beta = \lambda\xi_{xy}^\alpha + (1-\lambda)\xi_{xy}^\gamma$ και θα δείξουμε ότι αυτό είναι ταίριασμα μεταξύ των m_x^β και m_y^β. Για κάθε $w \in V$ ισχύει ότι

$$\begin{aligned}\sum_{z\in V} \xi_{xy}^\beta(z,w) &= \sum_{z\in V}\left(\lambda\xi_{xy}^\alpha(z,w) + (1-\lambda)\xi_{xy}^\gamma(z,w)\right) \\ &= \lambda\sum_{z\in V}\xi_{xy}^\alpha(z,w) + (1-\lambda)\sum_{z\in V}\xi_{xy}^\gamma(z,w) \\ &= \lambda m_y^\alpha(w) + (1-\lambda)m^\gamma(w),\end{aligned}$$

καθώς τα ξ_{xy}^α και ξ_{xy}^γ είναι ταιριάσματα. Θα δείξουμε τώρα ότι

$$\lambda m_y^\alpha(w) + (1-\lambda)m^\gamma(w) = m_y^\beta.$$

- Αν $w = y$ έχουμε ότι:

$$\begin{aligned}\lambda m_y^\alpha(y) + (1-\lambda)m_y^\gamma(y) &= \lambda\alpha + (1-\lambda)\gamma \\ &= \beta \\ &= m_y^\beta(y).\end{aligned}$$

- Αν $z \sim y$ έχουμε ότι:

$$\begin{aligned}\lambda m_y^\alpha(z) + (1-\lambda)m_y^\gamma(z) &= \lambda\frac{1-\alpha}{d_y} + (1-\lambda)\frac{1-\gamma}{d_y} \\ &= \frac{1-\beta}{d_x} \\ &= m_y^\beta(z).\end{aligned}$$

- Τέλος, σε διαφορετική περίπτωση η ισότητα είναι τετριμμένη, καθώς $m_y^\alpha(z) = m_y^\beta(z) = m_y^\gamma(z) = 0$.

Με όμοιο τρόπο δείχνουμε ότι για κάθε $w \in V$ ισχύει η σχέση:

$$\sum_{z\in V} \xi_{xy}^\beta(z,w) = m_x^\beta(w).$$

Επομένως δείξαμε ότι το ξ_{xy}^β είναι ταίριασμα των m_x^β και m_y^β και αυτό μας δίνει:

$$\begin{aligned}W_1(m_x^\beta, m_y^\beta) &\leqslant \sum_{z,w} \xi_{xy}^\beta(z,w)d(z,w) \\ &= \lambda \sum_{z,w\in V} \xi_{xy}^\alpha(z,w)d(z,w) + (1-\lambda) \sum_{z,w\in V} \xi_{xy}^\gamma(z,w)d(z,w) \\ &= \lambda W_1(m_x^\alpha, m_y^\alpha) + (1-\lambda)W_1(m_x^\gamma, m_y^\gamma).\end{aligned}$$

Τελικά καταλήγουμε στο ότι:

$$\begin{aligned}\kappa_\beta(x,y) &= 1 - \frac{W_1(m_x^\beta, m_y^\beta}{d(x,y)} \\ &\geqslant \lambda\left(1 - \frac{W_1(m_x^\alpha, m_y^\alpha}{d(x,y)}\right) + (1-\lambda)\left(1 - \frac{W_1(m_x^\gamma, m_y^\gamma}{d(x,y)}\right) \\ &= \lambda\kappa_\alpha(x,y) + (1-\lambda)\kappa_\gamma,\end{aligned}$$

το οποίο ολοκληρώνει την απόδειξη. □

Λήμμα 3.2.9. *Για κάθε $\alpha \in [0,1]$ και για κάθε δύο κορυφές $x, y \in V$, ισχύει η ανισότητα*

$$\kappa_\alpha(x,y) \leqslant (1-\alpha)\frac{2}{d(x,y)}.$$

Απόδειξη. Έχουμε

$$\begin{aligned}W_1(m_x^\alpha, m_y^\alpha) &\geqslant W_1(\delta_x, \delta_y) - W_1(d_x, m_x^\alpha) - W_1(\delta_y, m_y^\alpha) \\ &= d(x,y) - 2(1-\alpha).\end{aligned}$$

Συνεπώς,

$$\begin{aligned}\kappa_\alpha(x,y) &= 1 - \frac{W_1(m_x^\alpha, m_y^\alpha)}{d(x,y)}\\ &\leqslant (1-\alpha)\frac{2}{d(x,y)}.\end{aligned}$$

□

Παρατήρηση 3.2.10. Θεωρούμε τη συνάρτηση $h : [0,1] \to \mathbb{R}$ με $h(\alpha) = \kappa_\alpha(x,y)/(1-\alpha)$. Το Λήμμα 3.2.8 δείχνει ότι η συνάρτηση h είναι αύξουσα στο $[0,1)$, ενώ το Λήμμα 3.2.9 δείχνει ότι η h είναι φραγμένη. Αυτές οι δύο συνθήκες εγγυώνται ότι το όριο της h καθώς το $\alpha \to 1$ υπάρχει.

Ορισμός 3.2.11 (coarse καμπυλότητα Ricci)**.** Η coarse καμπυλότητα Ricci των κορυφών $x, y \in V$ του γραφήματος G είναι η ποσότητα

$$\kappa(x,y) = \lim_{\alpha\to 1} \frac{\kappa_\alpha(x,y)}{(1-\alpha)}.$$

Δείχνουμε τώρα τα βασικά θεωρήματα αυτής της παραγράφου:

Θεώρημα 3.2.12 (Bonnet-Myers)**.** *Για κάθε δύο κορυφές $x, y \in V$ του γραφήματος G με $\kappa(x,y) > 0$, ισχύει ότι*

$$d(x,y) \leqslant \frac{2}{\kappa(x,y)}.$$

Επιπλέον, αν για κάθε $x, y \in V$ ισχύει ότι $\kappa(x,y) \geqslant \kappa > 0$, τότε η διάμετρος του G είναι φραγμένη:

$$\mathrm{diam}(G) \leqslant \frac{2}{\kappa}.$$

Απόδειξη. Από το Λήμμα 3.2.9 έχουμε ότι

$$\frac{\kappa_\alpha(x,y)}{1-\alpha} \leqslant \frac{2}{d(x,y)},$$

και συνεπώς για $\alpha \to 1$ έχουμε

$$\kappa(x,y) \leqslant \frac{2}{d(x,y)}.$$

Αν η καμπυλότητα $\kappa(x,y)$ είναι θετική έχουμε το ζητούμενο:

$$d(x,y) \leqslant \frac{2}{d(x,y)}.$$

Στην περίπτωση που έχουμε ότι $\kappa(x,y) \geqslant \kappa > 0$, τότε το παραπάνω μας δίνει το ζητούμενο φράγμα για τη διάμετρο του γραφήματος. □

Ορισμός 3.2.13. Έστω G πεπερασμένο γράφημα με n κορυφές. Ορίζουμε τον πίνακα γειτνίασης του γραφήματος ως τον πίνακα $A = (a_{xy})_{1\leqslant x,y\leqslant n}$ με $a_{xy} = 1$ αν $x \sim y$ και $a_{xy} = 0$ αλλιώς. Επίσης ορίζουμε τον πίνακα των βαθμών των κορυφών ως τον διαγώνιο πίνακα $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Τέλος ορίζουμε τον κανονικοποιημένο Λαπλασιανό πίνακα του γραφήματος ως τον πίνακα $L = I - D^{-1/2}AD^{-1/2}$. Οι ιδιοτιμές του L λέγονται Λαπλασιανές ιδιοτιμές και γράφουμε

$$0 = \lambda_0 \leqslant \lambda_1 \leqslant \ldots \leqslant \lambda_{n-1}.$$

Παρατήρηση 3.2.14. Ο Λαπλασιανός πίνακας είναι συμμετρικός, συνεπώς οι ιδιοτιμές του υπάρχουν και είναι μη-αρνητικοί πραγματικοί αριθμοί, με τη μικρότερη ιδιοτιμή να είναι ίση με 0.

Θεώρημα 3.2.15 (Lichnerowicz)**.** *Έστω G πεπερασμένο γράφημα με n κορυφές και λ_1 η μικρότερη μη-μηδενική Λαπλασιανή ιδιοτιμή. Αν για κάθε $x, y \in V$ ισχύει ότι $\kappa(x, y) \geqslant \kappa \geqslant 0$, τότε $\lambda_1 \geqslant \kappa$.*

Απόδειξη. Το γράφημα G είναι πεπερασμένο, συνεπώς το όριο

$$\lim_{\alpha \to 1} \frac{\kappa_\alpha(x, y)}{1 - \alpha}$$

είναι ομοιόμορφο ως προς $x, y \in V$. Δηλαδή, για κάθε $\varepsilon > 0$ υπάρχει $\alpha_\varepsilon \in [0, 1)$ τέτοιο ώστε για κάθε $\alpha \in (\alpha_\varepsilon, 1)$ και για κάθε $x, y \in V$, να ισχύει ότι

$$\frac{\kappa_\alpha(x, y)}{1 - \alpha} > (1 - \varepsilon)\kappa > 0.$$

Έστω M_α ο τελεστής του τυχαίου περίπατου $\{m_x^\alpha\}_{x \in V}$. Από την Πρόταση 2.6.2 έχουμε ότι για κάθε ρ-Lipschitz συνάρτηση $f : V \to \mathbb{R}$ η $M_\alpha f$ είναι $\rho(1 - (1 - \varepsilon)(1 - \alpha)\kappa)$-Lipschitz. Αυτό μας δίνει ότι ο χρόνος μίξης του M_α είναι το πολύ $1 - (1 - \varepsilon)(1 - \alpha)\kappa$. Επίσης ο M_α μπορεί να γραφτεί ως ένας $n \times n$ πίνακας:

$$M_\alpha = \alpha I_n + (1 - \alpha)D^{-1}A.$$

Οι ιδιοτιμές του είναι $1, 1 - (1 - \alpha)\lambda_1, 1 - (1 - \alpha)\lambda_2, \ldots, 1 - (1 - \alpha)\lambda_{n-1}$. Επομένως, όταν $\alpha \to 1$, ο χρόνος μίξης του M_α είναι ακριβώς $1 - (1 - \alpha)\lambda_1$. Συνδυάζοντας αυτά τα δύο καταλήγουμε στην

$$1 - (1 - \alpha)\lambda_1 \leqslant 1 - (1 - \varepsilon)(1 - \alpha)\kappa,$$

δηλαδή

$$\lambda_1 \geqslant (1 - \varepsilon)\kappa.$$

Το ε ήταν τυχόν, οπότε όταν τείνει στο 0 έχουμε το ζητούμενο:

$$\lambda_1 \geqslant \kappa.$$

□

Επιστρέφουμε τώρα σε γενικότερα γραφήματα $G = (V, w, m)$. Εδώ η coarse καμπυλότητα Ricci δίνεται μέσω του τύπου

$$\kappa(x, y) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \kappa_\varepsilon(x, y).$$

Σκοπός μας είναι να εκφράσουμε την καμπυλότητα με έναν τρόπο που δεν θα περιέχει το όριο για $\alpha \to 1$. Αυτό θα επιτευχθεί μέσω της έννοιας της κλίσης:

Ορισμός 3.2.16 (κλίση)**.** Για κάθε $x, y \in V, x \neq y$ και για κάθε $f \in C(V)$ η κλίση της f στα x, y είναι η ποσότητα

$$\nabla_{xy} f := \frac{f(x) - f(y)}{x - y}.$$

Επίσης ορίζουμε

$$\|\nabla f\|_\infty := \sup_{x \neq y} |\nabla_{xy} f| = \sup_{x \sim y} |\nabla_{xy} f| \in [0, +\infty].$$

Τέλος, για $K \geqslant 0$, θέτουμε $\mathrm{Lip}(K) = \{f \in C(V) : \|\nabla f\|_\infty \leqslant K\}$.

Θεώρημα 3.2.17 (καμπυλότητα μέσω της Λαπλασιανής). *Έστω G γράφημα και $x \neq y$ κορυφές του G. Τότε,*

$$\kappa(x,y) = \inf \nabla_{xy}\Delta f,$$

με το infimum να παίρνεται πάνω από όλες τις $f \in \mathrm{Lip}(1) \cap C_c(V)$ με $\nabla_{xy} f = 1$.

Απόδειξη. Από τους ορισμούς έχουμε ότι

$$\begin{aligned} W_1(m_x^\varepsilon, m_y^\varepsilon) &= \sup_{f \in \mathrm{Lip}(1)} \sum_{z \in G} f(z)(m_y^\varepsilon(z) - m_x^\varepsilon(z)) \\ &= \sup_{f \in \mathrm{Lip}(1)} \{(f(y) + \varepsilon\Delta f(y)) - f(x) + \varepsilon\Delta f(x)\} \\ &= d(x,y) \sup_{f \in \mathrm{Lip}(1)} \nabla_{xy}(f + \varepsilon\Delta). \end{aligned}$$

Επομένως, το θεώρημα δυϊσμού του Kantorovich δίνει

$$\begin{aligned} \frac{\kappa_\varepsilon(x,y)}{\varepsilon} &= \frac{1}{\varepsilon}\left(1 - \frac{W_1(m_x^\varepsilon, m_y^\varepsilon)}{d(x,y)}\right) \\ &= \frac{1}{\varepsilon}\left(\inf_{f \in \mathrm{Lip}(1)} (1 - \nabla_{xy}(f + \varepsilon\Delta f))\right) \\ &= \inf_{f \in \mathrm{Lip}(1)} \left(\frac{1}{\varepsilon}(1 - \nabla_{xy} f) + \nabla_{xy}\Delta f\right) \\ &\leqslant \inf \nabla_{xy}\Delta f, \end{aligned}$$

με το τελευταίο infimum να παίρνεται πάνω από όλες τις $f \in \mathrm{Lip}(1) \cap C_c(V)$ με $\nabla_{xy} f = 1$.

Για να δείξουμε την αντίστροφη ανισότητα πρέπει να βρούμε μια $f_\varepsilon \in \mathrm{Lip}(1) \cap C_c(V)$ η οποία θα ελαχιστοποιεί την $\frac{1}{\varepsilon}(1 - \nabla_{xy} f) + \nabla_{xy}\Delta f$. Αρχικά, για κάθε $f \in \mathrm{Lip}(1)$ κατασκευάζουμε μια $\widetilde{f} \in \mathrm{Lip}(1)$, με $\mathrm{supp}(\widetilde{f}) = B_{2r}(x)$, όπου $r := d(x,y) + 1$, η οποία ικανοποιεί την

$$\frac{1}{\varepsilon}(1 - \nabla_{xy} f) + \nabla_{xy}\Delta f = \frac{1}{\varepsilon}(1 - \nabla_{xy}\widetilde{f}) + \nabla_{xy}\Delta\widetilde{f}.$$

Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι $f(x) = 0$ (αλλιώς δουλεύουμε με την $f^* = f - f(x)$). Αυτό δίνει $|f(z)| \leqslant r$ στο $B_1(x) \cup B_1(y)$, διότι η f είναι 1-Lipschitz, και άρα αν $x \in B_1(x)$ έχουμε $|f(z)| = |f(z) - f(x)| \leqslant d(x,z) \leqslant 1 \leqslant r$, ενώ αν $x \in B_1(y)$ έχουμε $|f(z)| = |f(z) - f(x)| \leqslant d(x,z) \leqslant d(z,y) + d(x,y) \leqslant 1 + d(x,y) = r$. Θέτουμε τώρα $\varphi : V \to \mathbb{R}$ με

$$\varphi(z) = [r \wedge (3r - d(x,z))]_+,$$

όπου $f \wedge g = \min\{f, g\}$. Παρατηρούμε ότι $\varphi(z) = r$ για κάθε $z \in B_1(x) \cup B_1(y)$ και επομένως η $\widetilde{f} := -\varphi \vee f \wedge \varphi$ ικανοποιεί τη ζητούμενη συνθήκη:

$$\frac{1}{\varepsilon}(1 - \nabla_{xy} f) + \nabla_{xy}\Delta f = \frac{1}{\varepsilon}(1 - \nabla_{xy}\widetilde{f}) + \nabla_{xy}\Delta\widetilde{f}.$$

Επιπλέον η φ έχει φορέα το $B_{2r}(x)$, άρα και η $\widetilde{f}$. Μπορούμε να περιορίσουμε, λοιπόν, το infimum σε συναρτήσεις με στήριγμα το συμπαγές $\widehat{B}_{2r}(x)$, το οποίο, λόγω συνέχειας, δίνει την ύπαρξη της f_ε.

Η συμπάγεια του $\widehat{B}_{2r}(x)$, το γεγονός ότι $f_\varepsilon \in \mathrm{Lip}(1)$ και η σχέση $f_\varepsilon(x) = 0$ για όλα τα ε συνεπάγονται την ύπαρξη ακολουθίας ε_n η οποία συγκλίνει στο 1 ώστε η οριακή συνάρτηση

$$f_0 := \lim_{n\to\infty} f_{\varepsilon_n}$$

να υπάρχει. Επίσης,

$$\frac{\kappa_\varepsilon(x,y)}{\varepsilon} = \frac{1}{\varepsilon}(1 - \nabla_{xy} f) + \nabla_{xy}\Delta f_\varepsilon,$$

και το όριο

$$\lim_{\varepsilon\to 0^+} \frac{\kappa_\varepsilon(x,y)}{\varepsilon}$$

υπάρχει, άρα $\nabla_{xy} f_\varepsilon \to 1$ για $\varepsilon \to 0^+$. Όλα τα παραπάνω δίνουν ότι $f_0 \in \mathrm{Lip}(1) \cap C_c(V)$, και $\nabla_{xy} f_0 = 1$ και εφόσον $\nabla_{xy} f_\varepsilon \leqslant 1$, παίρνουμε ότι

$$\begin{aligned}
\kappa(x,y) &= \lim_{\varepsilon\to 0^+} \left(\frac{1}{\varepsilon}(1 - \nabla_{xy} f_\alpha) + \nabla_{xy}\Delta f_\varepsilon \right) \\
&\geqslant \lim_{n\to\infty} \nabla_{xy}\Delta f_{\varepsilon_n} \\
&= \nabla_{xy}\Delta f_0 \\
&\geqslant \inf \nabla_{xy}\Delta f,
\end{aligned}$$

με το infimum να παίρνεται πάνω από όλες τις $f \in \mathrm{Lip}(1) \cap C_c(V)$ με $\nabla_{xy} f = 1$.

Οι δύο ανισότητες για το άνω και κάτω φράγμα της $\kappa(x,y)$ δίνουν την ζητούμενη ισότητα. □

Πόρισμα 3.2.18. *Για κάθε δύο κορυφές $x \neq y$ ισχύει ότι*

$$\kappa(x,y) = \inf \nabla_{xy}\Delta f,$$

με το infimum να παίρνεται πάνω από όλες τις συναρτήσεις $f : B_1(x) \cup B_1(y) \to \mathbb{Z},\ f \in \mathrm{Lip}(1)$ και $\nabla_{xy} f = 1$.

ΚΕΦΑΛΑΙΟ 4

Ανισότητες μεταφοράς της εντροπίας σε διακριτούς χώρους

4.1 Coarse καμπυλότητα Ricci και ανισότητες μεταφοράς της εντροπίας

Έστω (X, d, m) αριθμήσιμος μετρικός χώρος με τυχαίο περίπατο και έστω μ μέτρο πιθανότητας στον X. Η εντροπία μιας συνάρτησης $f : X \to \mathbb{R}_+$ ως προς το μέτρο μ ορίζεται ως

$$\mathrm{Ent}_\mu = \mathbb{E}_\mu \left[f \ln \left(\frac{f}{\mathbb{E}_\mu(f)} \right) \right].$$

Παρακάτω θα λέμε ότι η coarse καμπυλότητα Ricci του (X, d, m) είναι ο μεγαλύτερος $\kappa \in [-\infty, 1]$ για τον οποίο ισχύει η ανισότητα

$$W_1(m_x, m_y) \leqslant (1 - \kappa) d(x, y),$$

για κάθε $x, y \in X$.

Ορισμός 4.1.1. Θα λέμε ότι το μέτρο πιθανότητας μ ικανοποιεί την ιδιότητα Gaussian συγκέντρωσης με σταθερά C αν για κάθε συνάρτηση Lipschitz $f : X \to \mathbb{R}$ ισχύει η ανισότητα

$$\int_X e^f d\mu \leqslant \exp \left(\int_X f d\mu + C \|f\|_{\mathrm{Lip}}^2 \right).$$

Παρατήρηση 4.1.2. Για να δείξει κανείς ότι ένα μέτρο μ ικανοποιεί την ιδιότητα Gaussian συγκέντρωσης με σταθερά C, αρκεί να δείξει την ανισότητα

$$\int_X e^f d\mu \leqslant \exp \left(\int_X f d\mu + C \|f\|_{\mathrm{Lip}}^2 \right),$$

για κάθε $f : X \to \mathbb{R}$ με $\|f\|_{\mathrm{Lip}} \leqslant 1$ και $\int_X f d\mu = 0$.

Θα περάσουμε στην δυϊκή αναπαράσταση της ιδιότητας Gaussian συγκέντρωσης με όρους ανισότητας μεταφοράς. Θα χρειαστούμε την έννοια της σχετικής εντροπίας δύο μέτρων μ, ν στον X:

Ορισμός 4.1.3 (σχετική εντροπία)**.** Έστω μ, ν μέτρα στον X. Η σχετική εντροπία ή απόκλιση Kullback ορίζεται ως εξής:

$$D(\nu \parallel \mu) = \mathrm{Ent}_\nu\left(\frac{d\nu}{d\mu}\right) = \int_X \ln\left(\frac{d\nu}{d\mu}\right) d\nu$$

αν το ν είναι απολύτως συνεχές ως προς το μ, και $D(\nu\|\mu) = +\infty$ αλλιώς. Αν f, g είναι τυχαίες μεταβλητές με κατανομές μ και ν αντίστοιχα, με $D(f\|g)$ θα εννοούμε την σχετική εντροπία $D(\nu\|\mu)$.

Ορισμός 4.1.4. Θα λέμε ότι το μ ικανοποιεί την (T_1) με σταθερά C αν για κάθε μέτρο πιθανότητας ν στον X ισχύει η ανισότητα

$$W_1(\mu, \nu)^2 \leqslant C \cdot D(\nu \parallel \mu).$$

Λήμμα 4.1.5. *Ένα μέτρο πιθανότητας μ ικανοποιεί την ιδιότητα Gaussian συγκέντρωσης με σταθερά C αν και μόνο αν ικανοποιεί την (T_1) με σταθερά $4C$.*

Απόδειξη. Θα δείξουμε ότι η (T_1) με σταθερά $4C$ είναι ισοδύναμη με το εξής: Για κάθε $f : X \to \mathbb{R}$ με $\|f\|_{\mathrm{Lip}} \leqslant 1$ και $\int_X f d\mu = 0$ και για κάθε $t \in \mathbb{R}$ ισχύει η ανισότητα

$$\int_X e^{tf} d\mu \leqslant e^{Ct^2}. \tag{4.1.1}$$

Μπορούμε να υποθέσουμε ότι $t \geqslant 0$ και παρατηρούμε ότι η ανισότητα (4.1.1) είναι ισοδύναμη με το να δείξουμε ότι για κάθε μη-αρνητική Borel μετρήσιμη συνάρτηση $g : X \to \mathbb{R}$,

$$\int_X (tf - Ct^2) g d\mu \leqslant \mathrm{Ent}_\mu(g).$$

Αυτό ισχύει γιατί, $\mathrm{Ent}_\mu(g) = \sup \int ghd\mu$ με το supremum να είναι πάνω από όλες τις $h : X \to \mathbb{R}$ με $\int_X e^h d\mu \leqslant 1$. Συνεπώς, αν ισχύει η (4.1.1) τότε για την $h = tf - Ct^2$ έχουμε ότι $\int e^h d\mu = \int e^{tf - Ct^2} d\mu \leqslant 1$ και συνεπώς $\int (tf - Ct^2) g d\mu = \int hgd\mu \leqslant \mathrm{Ent}_\mu(g)$. Αν τώρα έχουμε ότι $\int (tf - Ct^2) g d\mu \leqslant \mathrm{Ent}_\mu(g)$, είναι προφανές ότι ισχύει το ζητούμενο. Η ανισότητα αυτή είναι ομογενής ως προς την g, επομένως μπορούμε να θεωρήσουμε ότι η g είναι πυκνότητα πιθανότητας για κάποιο μέτρο ν. Επιπλέον, η μέση τιμή της f είναι 0 και συνεπώς η ανισότητα (4.1.1) ισχύει αν και μόνο αν για κάθε g συνάρτηση πυκνότητας πιθανότητας ενός μέτρου ν ισχύει ότι

$$\int_X (fg - f) d\mu \leqslant Ct + \frac{1}{t}\int_X g \ln g d\mu. \tag{4.1.2}$$

Ελαχιστοποιούμε το δεξιό μέλος, ως προς $t > 0$ και έχουμε την ισοδύναμη μορφή της (4.1.2)

$$\int_X (fg - f) d\mu \leqslant \sqrt{4C \int_X g \ln g d\mu}. \tag{4.1.3}$$

Τώρα, αν χρησιμοποιήσουμε το γεγονός ότι $g = \frac{d\nu}{d\mu}$ έχουμε ότι η (4.1.3) ισοδυναμεί με την

$$\int_X f d\nu - \int_X f d\mu \leqslant \sqrt{4C \int_X \ln \frac{d\nu}{d\mu} d\nu} = \sqrt{4CD(\nu \| \mu)}. \tag{4.1.4}$$

Τέλος, το θεώρημα δυϊσμού του Kantorovich δίνει ότι η (4.1.4) είναι ισοδύναμη με την

$$W_1(\mu, \nu) \leqslant \sqrt{4C \cdot D(\nu \parallel \mu)},$$

δηλαδή ισχύει η (T_1) με σταθερά $4C$. □

Θεώρημα 4.1.6. *Έστω (X, d, m) μετρικός χώρος με τυχαίο περίπατο και υποθέτουμε ότι η coarse καμπυλότητα Ricci του τυχαίου περίπατου είναι $\kappa > 0$. Επίσης υποθέτουμε ότι για κάθε $x \in X$ τα μέτρα m_x ικανοποιούν την (T_1) με την ίδια σταθερά C. Τότε η αναλλοίωτη κατανομή του τυχαίου περίπατου ικανοποιεί την (T_1) με σταθερά $\frac{C}{\kappa(2-\kappa)}$.*

Θα δώσουμε δύο αποδείξεις του Θεωρήματος 4.1.6. Η πρώτη είναι αρκετά σύντομη και χρησιμοποιεί το θεώρημα δυϊσμού του Kantorovich. Στη δεύτερη κατασκευάζουμε ένα ταίριασμα βασιζόμενοι σε μια διαδικασία η οποία είναι ελάχιστης εντροπίας.

Απόδειξη μέσω δυϊσμού. Έστω $f : X \to \mathbb{R}$ μια Lipschitz συνάρτηση. Χρησιμοποιώντας την υπόθεση ότι τα m_x ικανοποιούν την (T_1) με σταθερά C και το Λήμμα 4.1.5 έχουμε ότι τα m_x ικανοποιούν την ιδιότητα Gaussian συγκέντρωσης με σταθερά $C/4$:

$$Me^f(x) \leqslant \exp\left(Mf(x) + \frac{C}{4}\|f\|_{\mathrm{Lip}}^2\right),$$

για κάθε $x \in X$. Με επαγωγή έχουμε ότι

$$M^N e^f(x) \leqslant \exp\left(M^N f(x) + \frac{C}{4}\sum_{i=0}^{N-1}\|M^i f\|_{\mathrm{Lip}}^2\right), \tag{4.1.5}$$

για κάθε $N \in \mathbb{N}$ και για κάθε $x \in X$. Παρατηρούμε ότι:

$$\begin{aligned} 1 - \kappa &= \sup_{x \neq y}\left\{\frac{W_1(m_x, m_y)}{d(x,y)}\right\} \\ &= \sup_{x \neq y, g}\left\{\frac{Mg(x) - Mg(y)}{\|g\|_{\mathrm{Lip}} d(x,y)}\right\} \\ &= \sup_{g}\left\{\frac{\|Mg\|_{\mathrm{Lip}}}{\|g\|_{\mathrm{Lip}}}\right\}, \end{aligned}$$

επομένως $\|Mg\|_{\mathrm{Lip}} \leqslant (1-\kappa)\|g\|_{\mathrm{Lip}}$ για κάθε $g : X \to \mathbb{R}$ συνάρτηση Lipschitz. Συνεπώς, έχουμε με επαγωγή ότι

$$\|M^N f\|_{\mathrm{Lip}} \leqslant (1-\kappa)^N \|f\|_{\mathrm{Lip}}$$

για κάθε $N \in \mathbb{N}$. Η ανισότητα (4.1.5) δίνει ότι

$$M^N e^f(x) \leqslant \exp\left(M^N f(x) + \frac{C}{4\kappa(2-\kappa)}\|f\|_{\mathrm{Lip}}^2\right).$$

Για $N \to \infty$ έχουμε ότι

$$\int_X e^f d\nu \leqslant \exp\left(\int_X f d\nu + \frac{C}{4\kappa(2-\kappa)}\|f\|_{\mathrm{Lip}}^2\right),$$

το οποίο σημαίνει ότι πράγματι η αναλλοίωτη κατανομή ν ικανοποιεί την ιδιότητα Gaussian συγκέντρωσης με σταθερά $\frac{C}{4\kappa(2-\kappa)}$. □

Απόδειξη με κατασκευή ταιριάσματος.

Πρόταση 4.1.7. *Έστω (X,d,m) που ικανοποιεί τις υποθέσεις του θεωρήματος. Σταθεροποιούμε μια χρονική στιγμή T και ένα σημείο $x_0 \in X$. Έστω $\{B_0 = x_0, B_1, \dots, B_T\}$ ο τυχαίος περίπατος διακριτού χρόνου που αντιστοιχεί στον πυρήνα Markov m και ξεκινάει από το x_0. Θεωρούμε επίσης μια τυχαία διαδικασία $\{X_0 = x_0, \dots, X_T\}$ στον X, η οποία επίσης ξεκινάει από το x_0. Τότε, υπάρχει ταίριασμα των δύο διαδικασιών ώστε:*

$$\mathbb{E}[d(X_T, B_T)] \leqslant \sqrt{\frac{C}{\kappa(2-\kappa)} D(\{X_0,\dots,X_T\}||\{B_0,\dots,B_T\})}.$$

Έχοντας την παραπάνω πρόταση, βλέπουμε ότι για να αποδείξουμε την ανισότητα (T_1) για τον (X,d,m) αρκεί, για κάθε μέτρο πιθανότητας ν στον X, να κατασκευάσουμε μια στοχαστική διαδικασία $\{X_t\}$ η οποία θα έχει τις εξής ιδιότητες:

(i) $X_T \sim \nu$.

(ii) Η σχετική εντροπία μεταξύ των $\{X_0,\dots,X_T\}$ και $\{B_0,\dots,B_T\}$ να είναι όσο μικρότερη γίνεται, δηλαδή να ισχύει η σχέση

$$D(\{X_0,\dots,X_T\}||\{B_0,\dots,B_T\}) = D(X_T\|B_T).$$

Περνάμε τώρα στην κατασκευή:

Για $t \in \{1,\dots,T\}$ και για $x_1,\dots,x_{t-1} \in X$ θεωρούμε τη δεσμευμένη κατανομή της X_t δεδομένων των $X_0 = x_0,\dots,X_{t-1} = x_{t-1}$, την συμβολίζουμε με $\nu(t,x_0,\dots,x_{t-1},\cdot)$. Κατασκευάζουμε το ταίριασμα των $\{X_t\}$ και $\{B_t\}$ ως εξής. Θέτουμε $X_0 = B_0 = x_0$ και δεδομένων των $(X_1,B_1),\dots,(X_{t-1},B_{t-1})$ θέτουμε (X_t,B_t) να είναι το ταίριασμα των $\nu(t,X_0,\dots,X_{t-1},\cdot)$ και $m_{B_{t-1}},\cdot)$. Αυτό από την κατασκευή του είναι βέλτιστο.

Για την απόδειξη της πρότασης θα χρειαστούμε το παρακάτω λήμμα:

Λήμμα 4.1.8. *Για κάθε $t \in \{1,\dots,T\}$ ισχύει ότι*

$$\mathbb{E}_{t-1}[d(X_t,B_t)] \leqslant \sqrt{C \cdot D(\nu(t,X_0,\dots,X_{t-1},\cdot)\|m_{X_{t-1}}(\cdot))} + (1-\kappa)d(X_{t-1},B_{t-1}),$$

όπου $\mathbb{E}_{t-1}$ είναι η δεσμευμένη μέση τιμή δεδομένων των $(X_0,B_0),\dots,(X_{t-1},B_{t-1})$.

Απόδειξη. Ο ορισμός του ταιριάσματος, η τριγωνική ανισότητα για την W_1, η (T_1) για τα m_x και η θετική καμπυλότητα δίνουν:

$$\begin{aligned}
\mathbb{E}_{t-1}[d(X_t,B_t)] &= W_1(\nu(t,X_0,\dots,X_{t-1},\cdot), m_{B_{t-1}}(\cdot)) \\
&\leqslant W_1(\nu(t,X_0,\dots,X_{t-1},\cdot), m_{X_{t-1}}(\cdot)) + W_1(m_{X_{t-1}}(\cdot), m_{B_{t-1}}(\cdot))) \\
&\leqslant \sqrt{C \cdot D(\nu(t,X_0,\dots,X_{t-1},\cdot)\|m_{X_{t-1}}(\cdot))} + (1-\kappa)d(X_{t-1},B_{t-1}).
\end{aligned}$$

□

Απόδειξη της πρότασης. Ο κανόνας της αλυσίδας για τη σχετική εντροπία δίνει ότι

$$\sum_{t=1}^{T} \mathbb{E}[D(\nu(t, X_0, \ldots, X_{t-1}, \cdot) \| m_{X_{t-1}}(\cdot))] = D(\{X_0, \ldots, X_T\} \| \{B_0, \ldots, B_T\}).$$

Χρησιμοποιώντας το προηγούμενο λήμμα επαγωγικά, και μετά την ανισότητα Cauchy-Schwarz, έχουμε ότι

$$\begin{aligned}
\mathbb{E}[d(X_T, B_T)] &\leqslant \sum_{t=1}^{T} (1-\kappa)^{T-t} \mathbb{E}\left[\sqrt{C \cdot D(\nu(t, X_0, \ldots, X_{t-1}, \cdot) \| m_{X_{t-1}}(\cdot))}\,\right] \\
&\leqslant \sqrt{\sum_{t=1}^{T} (1-\kappa)^{2(T-t)}} \sqrt{\sum_{t=1}^{T} C \cdot \mathbb{E}[D(\nu(t, X_o, \ldots, X_{t-1}, \cdot) \| m_{X_{t-1}}(\cdot))} \\
&\leqslant \sqrt{\frac{1}{\kappa(2-\kappa)}} \sqrt{C \cdot D(\{X_0, \ldots, X_T\} \| \{B_0, \ldots, B_T\})},
\end{aligned}$$

το οποίο είναι το ζητούμενο. □

Θα κατασκευάσουμε τώρα μια στοχαστική διαδικασία $X_0 = x_0, X_1, \ldots, X_T$ που θα ικανοποιεί την συνθήκη (ii). Έστω ν μέτρο πιθανότητας στον X, $x_0 \in X$ και $T \geqslant 1$. Έστω μ_T η κατανομή της $B_T(x_0)$ και f η πυκνότητα του ν ως προς το μ_T. Έστω $\{X_t\}_{t=0}^{T}$ η αλυσίδα Markov στον X με πιθανότητες μετάβασης την χρονική στιγμή t να δίνονται από τον τύπο

$$q_t(x, y) := \mathbb{P}(X_t = y | X_{t-1} = x) = \frac{M^{T-t} f(y)}{M^{T-t+1} f(x)} m_x(y).$$

Λήμμα 4.1.9. *Αν $m^T(x_0, x) > 0$ για κάθε $x \in \operatorname{supp}(\mu)$ τότε η $\{X_t\}$ είναι καλά ορισμένη. Επιπλέον, για κάθε $x_1, \ldots, x_T \in X$ έχουμε ότι*

$$\mathbb{P}((X_1, \ldots, X_T)) = (x_1, \ldots, x_T)) = \mathbb{P}((B_1, \ldots, B_T) = (x_1, \ldots, x_T)) f(x_T).$$

Ειδικότερα, η X_T έχει κατανομή $d\nu = f d\mu_T$.

Απόδειξη. Από τον ορισμό της (X_T) έχουμε ότι

$$\begin{aligned}
\mathbb{P}((X_1, \ldots, X_T) = (x_1, \ldots, x_T)) &= \prod_{t=1}^{T} \mathbb{P}(X_t = x_t | (X_1, \ldots, X_{t-1}) = (x_1, \ldots, x_{t-1})) \\
&= \prod_{t=1}^{T} \frac{M^{T-t} f(x_t)}{M^{T-t+1} f(x_{t-1})} m_{x_{t-1}}(x_t) \\
&= \frac{f(x_T)}{M^T f(x_0)} \left(\prod_{t=1}^{T} m_{x_{t-1}}, x_t) \right),
\end{aligned}$$

το οποίο είναι το ζητούμενο. □

Αυτό το λήμμα δείχνει ότι η κατανομή της διαδικασίας $\{X_t\}$ έχει πυκνότητα $f(x_T)$ σε σχέση με την κατανομή της διαδικασίας $\{B_t\}$. Επομένως έχουμε ότι

$$D(\{X_0,\ldots,X_T\}\|\{B_0,\ldots,B_T\}) = \mathbb{E}(\ln f(X_T)) = D(\nu\|\mu_T),$$

αφού η X_T έχει κατανομή μ. Μάλιστα η $\{X_t\}$ είναι η μοναδική διαδικασία με αυτήν την ιδιότητα, λόγω της κυρτότητας της σχετικής εντροπίας.

Για να ολοκληρώσουμε την απόδειξη του θεωρήματος, αρκεί να συνδυάσουμε την ιδιότητα (ii) με την Πρόταση 4.1.7. Θα πάρουμε

$$W_1(\nu,\mu_T) = \mathbb{E}[d(X_T,B_T)] \leqslant \sqrt{\frac{C}{\kappa(2-\kappa)}D(\nu\|\mu_T)}.$$

Αν αφήσουμε το T να πάει στο άπειρο έχουμε το ζητούμενο αποτέλεσμα.

□

Η υπόθεση ότι τα m_x ικανοποιούν την (T_1) με σταθερά C για κάθε $x \in X$, αν και φαίνεται ιδιαίτερα περίεργη εκ πρώτης όψεως, τελικά ικανοποιείται αυτόματα για τυχαίους περίπατους σε γραφήματα. Πράγματι, σε κάθε γράφημα ισχύει η ανισότητα του Pinsker: Για κάθε δύο μέτρα πιθανότητας μ,ν ισχύει

$$\mathrm{TV}(\mu,\nu) \leqslant \sqrt{\frac{1}{2}D(\mu \parallel \nu)},$$

όπου $TV(\mu,\nu) = \sup_A|\mu(A)-\nu(A)|$. Αυτή η ανισότητα μας δίνει το επόμενο λήμμα.

Λήμμα 4.1.10. *Έστω μ μέτρο πιθανότητας στον μετρικό χώρο (X,d) και ας υποθέσουμε ότι ο φορέας του μ έχει πεπερασμένη διάμετρο Δ. Τότε το μ ικανοποιεί την (T_1) με σταθερά $\Delta^2/2$.*

Απόδειξη. Έστω ν απολύτως συνεχές ως προς το μ. Τότε τα μ και ν έχουν φορέα με διάμετρο Δ και συνεπώς

$$W_1(\mu,\nu) \leqslant \Delta \cdot \mathrm{TV}(\mu,\nu).$$

Η ανισότητα του Pinsker δίνει το ζητούμενο. □

Μια συγκεκριμένη περίπτωση που παρουσιάζει ενδιαφέρον είναι ο τυχαίος περίπατος σε πεπερασμένα γραφήματα. Έστω $G=(V,E)$ ένα συνεκτικό, μη προσανατολισμένο γράφημα. Θεωρούμε συνάρτηση βάρους $c: E \to \mathbb{R}_+$ στις ακμές του γραφήματος. Μπορούμε να ορίσουμε μια αλυσίδα Markov $\{X_t\}$ στο γράφημα μέσω του τύπου

$$Pr\left[X_{t+1}=y \mid X_t=x\right] = \frac{c\left(\{x,y\}\right)}{\sum_{z\in V} c\left(\{x,z\}\right)}.$$

Μια τέτοια αλυσίδα Markov ορίζει έναν τυχαίο περίπατο m στο γράφημα G. Αν επιπλέον ισχύει ότι για κάθε $x \in V$

$$c\left(\{x,y\}\right) \geqslant \frac{1}{2}\sum_{z\in V} c\left(\{x,z\}\right),$$

τότε ο τυχαίος περίπατος θα λέγεται lazy.

Ο φορέας του m_x για κάθε $x \in V$ έχει διάμετρο 2, συνεπώς το Θεώρημα 4.1.6 σε συνδυασμό με το Λήμμα 4.1.10 δίνουν το ακόλουθο πόρισμα:

Πόρισμα 4.1.11. *Αν ο τυχαίος περίπατος στο γράφημα G έχει θετική coarse καμπυλότητα Ricci κ, τότε το αναλλοίωτο μέτρο ν του τυχαίου περίπατου ικανοποιεί την σχέση*

$$W_1(\mu,\nu)^2 \leqslant \frac{2}{\kappa(2-\kappa)} D(\mu \parallel \nu),$$

για κάθε μέτρο πιθανότητας μ.

Παρατήρηση 4.1.12. Γενικά ισχύει ότι

$$d(x,y) \leqslant W_1(m_x,m_y) + 2,$$

για κάθε $x, y \in V$, αφού μετά από κάθε βήμα του τυχαίου περίπατου η απόσταση από το σημείο εκκίνησης είναι το πολύ 1. Εφόσον η coarse καμπυλότητα Ricci είναι κ, η διάμετρος του γραφήματος είναι το πολύ 2κ. Συνεπώς το Λήμμα 4.1.10 δίνει ότι

$$W_1(\mu,\nu)^2 \leqslant 2/\kappa^2 D(\mu \parallel \nu),$$

δηλαδή ότι το ν ικανοποιεί την (T_1) με σταθερά $2/\kappa^2$. Το Πόρισμα 4.1.11 δίνει το ίδιο αποτέλεσμα με καλύτερη σταθερά, καθώς είναι της τάξης του $1/\kappa$ και όχι του $1/\kappa^2$.

ΚΕΦΑΛΑΙΟ 5

Καμπυλότητα Ricci στον διακριτό κύβο

5.1 Ανισότητα Brunn–Minkowski

Μια γνωστή διατύπωση της ανισότητας Brunn–Minkowski είναι η εξής: αν A_0, A_1 είναι δύο μη κενά συμπαγή υποσύνολα του $\mathbb{R}^n$ τότε

$$\ln \mathrm{vol}(A_t) \geqslant (1-t)\ln \mathrm{vol}(A_0) + t\ln \mathrm{vol}(A_1)$$

για κάθε $0 \leqslant t \leqslant 1$, όπου $A_t = \{(1-t)a_0 + ta_1 : a_0 \in A_0, a_1 \in A_1\}$. Με άλλα λόγια, ο λογάριθμος του όγκου του A_t είναι κοίλη συνάρτηση του t. Από αυτή τη μορφή της ανισότητας Brunn–Minkowski, η οποία είναι «απειροδιάστατη» με την έννοια ότι δεν εμφανίζεται κάπου η διάσταση n, μπορεί κανείς να πάρει την πιο γνωστή μορφή της ανισότητας

$$\mathrm{vol}(A_t)^{1/n} \geqslant (1-t)\mathrm{vol}(A_0)^{1/n} + t\,\mathrm{vol}(A_1)^{1/n}.$$

Αν αντικαταστήσουμε τον $\mathbb{R}^n$ με μια πολλαπλότητα Riemann που έχει θετική καμπυλότητα, τότε η ανισότητα Brun-Minkowski ισχυροποιείται. Αποδεικνύεται ότι αν X είναι μια λεία και πλήρης πολλαπλότητα Riemann με καμπυλότητα Ricci μεγαλύτερη από κάποια σταθερά $\kappa > 0$, τότε για κάθε ζεύγος μη κενών συμπαγών υποσυνόλων $A_0, A_1 \subset X$ ισχύει η ανισότητα

$$\ln \mathrm{vol}(A_t) \geqslant (1-t)\ln \mathrm{vol}(A_0) + t\ln \mathrm{vol}(A_1) + \frac{\kappa}{2}t(1-t)d(A_0,A_1)^2,$$

όπου τώρα A_t είναι το σύνολο όλων των σημείων $\gamma(t)$, όπου γ είναι οποιαδήποτε γεωδαισιακή καμπύλη με $\gamma(0) \in A_0$ και $\gamma(1) \in A_1$, και

$$d(A_0,A_1) = \inf\{d(a_0,a_1) : a_0 \in A_0, a_1 \in A_1\}.$$

Μάλιστα, αυτού του τύπου η ανισότητα έχει εμφανιστεί στη βιβλιογραφία σαν ενδεχόμενος *ορισμός* της θετικής καμπυλότητας Ricci σε πιο γενικούς χώρους. Η ιδέα είναι ότι σε χώρους με θετική καμπυλότητα «τα μέσα εξαπλώνονται» οπότε το σύνολο των μέσων δύο συνόλων είναι μεγαλύτερο απ' ότι στον Ευκλείδειο χώρο. Αυτό έχει οδηγήσει στον ορισμό της *κυρτότητας της εντροπίας ως προς μετατόπιση* για πολλαπλότητες Riemann. Η θεωρία αυτή έχει αναπτυχθεί από τους Sturm, Lott και Villani, αλλά δεν είναι σαφής η μορφή που θα μπορούσε να πάρει σε διακριτούς χώρους.

Μια άλλη απόπειρα ορισμού της καμπυλότητας Ricci είναι η *coarse καμπυλότητα Ricci*, η οποία εισήχθη από τον Ollivier και την μελετήσαμε στα προηγούμενα κεφάλαια. Η βασική ιδέα εδώ ήταν ότι η καμπυλότητα Ricci είναι θετική όταν «οι μπάλες βρίσκοναι πιο κοντά απ' ότι τα κέντρα τους» με την έννοια της απόστασης μεταφοράς.

Σε αυτό το κεφάλαιο συζητάμε τις δύο αυτές προσεγγίσεις στο παράδειγμα του διακριτού κύβου $X=\{0,1\}^N$. Αυτός είναι ο απλούστερος διακριτός χώρος που περιμένουμε ότι θα έχει θετική καμπυλότητα Ricci, για διάφορους λόγους. Το ερώτημα «να υπολογιστεί η καμπυλότητα Ricci του διακριτού κύβου» έχει διατυπωθεί, με αυτήν ακριβώς τη φράση, από τον Stroock το 1998. Ο υπολογισμός της coarse καμπυλότητας Ricci του $\{0,1\}^N$ είναι εφικτός, και η τιμή της είναι $\kappa=\frac{2}{N+1}$. Η ακριβής τιμή της καμπυλότητας με την έννοια της κυρτότητας της εντροπίας ως προς μετατόπιση δεν είναι γνωστή. Είναι όμως γνωστό ότι είναι της τάξης του $\frac{1}{N}$, κάτι που συμφωνεί με την αντίστοιχη τιμή για την coarse καμπυλότητα Ricci. Εκτός από αυτήν την εκτίμηση, θα αποδείξουμε μια συνδυαστική ανισότητα Brunn-Minkowski στον διακριτό κύβο, στην οποία υπεισέρχεται ένας όρος που οφείλεται στην θετική καμπυλότητα.

5.1.1 Ανισότητα Brunn-Minkowski στον διακριτό κύβο

Θεωρούμε τον διακριτό κύβο $X=\{0,1\}^N$, $N\geqslant 1$, εφοδιασμένο με την Hamming μετρική

$$d(x,y)=\mathrm{card}(\{i:x_i\neq y_i\})=\sum_{i=1}^N |x_i-y_i|.$$

Αν A,B είναι μη κενά υποσύνολα του X, ορίζουμε $d(A,B)=\min\{d(a,b):a\in A, b\in B\}$.

Αν a και b είναι δύο σημεία του X, ονομάζουμε *μέσο* των a και b οποιοδήποτε σημείο $m\in X$ για το οποίο

$$d(m,a)=d(m,b)=d(a,b) \quad \text{και} \quad |d(m,b)-d(a,b)/2|<1.$$

Πιο συγκεκριμένα, αν η απόσταση $d(a,b)$ είναι άρτια τότε μέσος είναι το μεσαίο σημείο σε κάθε βέλτιστο μονοπάτι από το a στο b στον X, ενώ αν η απόσταση $d(a,b)$ είναι περιττή τότε μέσος είναι ένα από τα δύο μεσαία σημεία σε αυτό το βέλτιστο μονοπάτι. Στον διακριτό κύβο οι μέσοι δεν ορίζονται μονοσήμαντα: το πλήθος τους είναι ίσο με $\binom{d(a,b)}{d(a,b)/2}$ αν ο $d(a,b)$ είναι άρτιος, και $2\binom{d(a,b)}{(d(a,b)-1)/2}$ αν ο $d(a,b)$ είναι περιττός.

Αν A και B είναι δύο μη κενά υποσύνολα του X, ορίζουμε ως *σύνολο των μέσων των A και B* το σύνολο των μέσων όλων των ζευγαριών $(a,b)\in A\times B$. Με αυτήν την ορολογία ισχύει το ακόλουθο.

Θεώρημα 5.1.1. *Έστω A και B δύο μη κενά υποσύνολα του $X=\{0,1\}^N$ και έστω M το σύνολο των μέσων των A και B. Τότε,*

$$\ln(\mathrm{card}(M))\geqslant \frac{1}{2}\ln(\mathrm{card}(A))+\frac{1}{2}\ln(\mathrm{card}(B))+\frac{\kappa}{8}d(A,B)^2,$$

όπου $\kappa=\frac{1}{2N}$.

Το θεώρημα αυτό είναι ανάλογο με την ανισότητα Brunn–Minkowski (στην περίπτωση $t = 1/2$) που περιγράψαμε για πολλαπλότητες Riemann, με τη σταθερά κ να παίζει το ρόλο του κάτω φράγματος για την καμπυλότητα.

Η τάξη μεγέθους $\frac{1}{N}$ για τη σταθερά κ είναι βέλτιστη: αν τα A και B είναι μονοσύνολα που έχουν απόσταση ίση με N, τότε $d(A,B)^2 = N^2$, ενώ το πλήθος των μέσων είναι $\binom{N}{N/2} \sim 2^N \sqrt{\frac{2}{\pi N}}$, συνεπώς ο $\ln(\mathrm{card}(M))$ είναι της τάξης του N. Στη συνέχεια θα δούμε ότι μπορεί κανείς να βελτιώσει το Θεώρημα 5.1.1, αντικαθιστώντας την απόσταση $d(A,B)$ με μια απόσταση μεταφοράς.

5.1.2 Εντροπία των μέσων στο διακριτό κύβο

Έστω μ ένα μέτρο πιθανοτητας σε ένα διακριτό σύνολο X. Η *εντροπία Shannon* του μ είναι η ποσότητα

$$S(\mu) := -\sum_{x\in X} \mu(x) \ln \mu(x).$$

Θα χρειαστούμε επίσης την *σχετική εντροπία* (ή απόκλιση Kullback–Leibler) ενός μέτρου μ ως προς ένα μέτρο πιθανότητας ν στο X, η οποία ορίζεται ως εξής:

$$H(\mu|\nu) := \sum_{x\in X} \mu(x) \ln \frac{\mu(x)}{\nu(x)} \geqslant 0.$$

Αν το X είναι πεπερασμένο και το ν είναι το ομοιόμορφο μέτρο πιθανότητας στο X, τότε

$$H(\mu|\nu) = \ln(\mathrm{card}(X)) - S(\mu).$$

Για να διατυπώσουμε μια εντροπική εκδοχή του Θεωρήματος 5.1.1 ορίζουμε πρώτα τα μέσα δύο μέτρων μ_0 και μ_1. Η ιδέα είναι η ακόλουθη: επιλέγουμε πρώτα τυχαίο σημείο a με κατανομή το μ_0, στη συνέχεια ένα *ανεξάρτητο* σημείο b με κατανομή μ_1, και τέλος έναν τυχαίο μέσο των a και b, ομοιόμορφα από το σύνολο των μέσων τους. Για να κάνουμε αυτή την ιδέα αυστηρή, για κάθε ζεύγος σημείων a και b του X θεωρούμε το μέτρο $\mathrm{mid}(a,b)$ που είναι το ομοιόμορφο μέτρο στο σύνολο των μέσων των a και b. Αν τώρα μ_0 και μ_1 είναι δύο μέτρα πιθανότητας στο X, το *μέτρο των μέσων* των μ_0 και μ_1 είναι το μέτρο

$$\mathrm{mid}(\mu_0,\mu_1) := \iint \mathrm{mid}(a,b)\, d\mu_0(a)\, d\mu_1(b).$$

Με αυτήν την ορολογία έχουμε το εξής.

Θεώρημα 5.1.2. *Έστω μ_0 και μ_1 δύο μέτρα πιθανότητας στον διακριτό κύβο $X = \{0,1\}^N$. Έστω $\mu_{1/2} = \mathrm{mid}(\mu_0,\mu_1)$ το μέτρο των μέσων τους. Τότε,*

$$S(\mu_{1/2}) \geqslant \frac{1}{2}(S(\mu_0) + S(\mu_1)) + \frac{\kappa}{8} W_1(\mu_0,\mu_1)^2,$$

όπου $\kappa = \frac{1}{2N}$. Ισοδύναμα,

$$H(\mu_{1/2}|\nu) \leqslant \frac{1}{2}(H(\mu_0|\nu) + H(\mu_1|\nu)) - \frac{\kappa}{8} W_1(\mu_0,\mu_1)^2,$$

όπου ν είναι το ομοιόμορφο μέτρο πιθανότητας στον $X = \{0,1\}^N$.

Στο Θεώρημα 5.1.2 χρησιμοποιούμε την απόσταση Wasserstein

$$W_1(\mu, \mu') := \inf_{\xi} \iint d(a,b)\, d\xi(a,b),$$

όπου το infimum παίρνεται πάνω από όλα τα μέτρα ξ στο $X \times X$ για τα οποία $\int_b d\xi(a,b) = d\mu(a)$ και $\int_a d\xi(a,b) = d\mu'(b)$, δηλαδή όλα τα ταιριάσματα των μ και μ'.

Παρατηρήστε ότι αν τα μ_0 και μ_1 εχουν φορείς τα σύνολα A και B, τότε

$$d(A,B) \leqslant W_1(\mu_0, \mu_1)$$

και σε αυτήν την περίπτωση το μέτρο $\mu_{1/2}$ έχει φορέα το σύνολο M των μέσων των A και B, άρα

$$S(\mu_{1/2}) \leqslant \ln(\mathrm{card}(M)).$$

Συνεπώς, αν τα μ_0 και μ_1 είναι τα ομοιόμορφα μέτρα στα A και B, τότε από το Θεώρημα 5.1.2 προκύπτει άμεσα το Θεώρημα 5.1.1.

5.2 Δύο έννοιες διακριτής καμπυλότητας Ricci

Υπάρχουν δύο διαφορετικές εκδοχές της έννοιας της καμπυλότητας Ricci σε διακριτούς χώρους. Δίνουμε μια σύντομη περιγραφή τους, ξεκινώντας από την coarse καμπυλότητα Ricci που συζητήσαμε στα προηγούμενα κεφάλαια.

5.2.1 Coarse καμπυλότητα Ricci

Έστω $x \in X$ και θεωρούμε το ομοιόμορφο μέτρο στις γειτονικές κορυφές της x, συμπεριλαμβανομένης και της ίδιας της x: $m_x(z) = 1/(N+1)$ αν $z = x, x^1, \ldots, x^N$, όπου x^i είναι η γειτονική κορυφή της x με αλλαγμένη την i συντεταγμένη. Για αυτόν τον τυχαίο περίπατο θα δείξουμε ότι η coarse καμπυλότητα Ricci του X είναι $2/(N+1)$: Έστω $x = (0,\ldots,0), y = (1,\ldots,0)$ και θεωρούμε το ταίριασμα

$$\xi_{xy}(z,w) = \begin{cases} 1/(N+1) & \text{αν } z = x, w = y^1 \\ 1/(N+1) & \text{αν } z = x^1, w = y \\ 1/(N+1) & \text{αν } z = x^i, w = y^i,\ i \geqslant 2 \end{cases}$$

Τότε έχουμε ότι

$$W_1(m_x, m_y) \leqslant \sum_{z,w} d(z,w)\xi_{xy}(z,w) = \sum_{i=2}^{N} \frac{1}{N+1} = \frac{N-1}{N+1},$$

επομένως

$$\kappa(x,y) = 1 - \frac{W_1(m_x, m_y)}{d(x,y)} = 1 - \frac{N-1}{N+1} = \frac{2}{N+1}.$$

Για την αντίστροφη ανισότητα θεωρούμε την συνάρτηση $f : X \to \{0,1\}$ με $f(x_1,\ldots,x_N) = x_1$ και έχουμε ότι

$$W_1(m_x, m_y) \geqslant \left| \sum_z f(z) m_x(z) - \sum_w f(w) m_y(w) \right| = \left| \frac{1}{N+1} - \frac{N}{N+1} \right| = \frac{N-1}{N+1},$$

άρα

$$\kappa(x,y) \leqslant 1 - \frac{N-1}{N+1} = \frac{2}{N+1}.$$

5.2.2 Κυρτότητα ως προς μετατόπιση

Σε αυτήν την παράγραφο θα μελετήσουμε τις ιδέες των K.-T.Sturm και M.-K.von Renesse για ομοιόμορφα κάτω φράγματα της καμπυλότητας Ricci σε πολλαπλότητες Riemann. Έστω (M,g) μια λεία και συνεκτική N-διάστατη πολλαπλότητα Riemann, με την Riemannian απόσταση των $x,y \in M$ να συμβολίζεται με $d(x,y)$ και τον Riemannian όγκο με $m(dx) = \mathrm{vol}(dx)$. Θα χρησιμοποιήσουμε την L^2-απόσταση Wasserstein, οπότε ορίζουμε γενικότερα την L^r-απόσταση Wasserstein:

Ορισμός 5.2.1. Έστω m_x και m_y μέτρα στην M. Ορίζουμε την L^r-απόσταση Wasserstein ως

$$W_r(m_x,m_y) = \inf\left\{\int_{M\times M} d(z,w)^r d\xi_{xy}(z,w)\right\}^{1/r},$$

με το infimum να είναι πάνω από όλα τα ταιριάσματα ξ_{xy} των m_x και m_y. Ο χώρος των μέτρων πιθανότητας μ στην M με $\int_M d(z,w)^r\mu(dy) < \infty$ θα συμβολίζεται με $\mathscr{P}^r(M)$.

Παρατήρηση 5.2.2. Ο χώρος $\mathscr{P}^r(M)$ είναι γεωδαισιακός χώρος.

Ορισμός 5.2.3. Έστω $\nu \in \mathscr{P}^r(M)$. Η εντροπία του ν ορίζεται ως η ποσότητα

$$\mathrm{Ent}(\nu) := \int_M \frac{d\nu}{dx}\ln\frac{d\nu}{dx}\mathrm{vol}(dx)$$

όταν το μέτρο ν είναι απολύτως συνεχές ως προς τον όγκο vol και όταν $\int_M \frac{d\nu}{dx}\left[\ln\frac{d\nu}{dx}\right]_+ \mathrm{vol}(dx) < \infty$ και $\mathrm{Ent}(\nu) := +\infty$ αλλιώς.

Ορισμός 5.2.4. Έστω (X,d) γεωδαισιακός χώρος, $K \in \mathbb{R}$ και $f : X \to [-\infty,+\infty]$. Θα λέμε ότι η συνάρτηση f είναι $K-$κυρτή αν και μόνο αν για κάθε γεωδαισιακή $\gamma : [0,1] \to X$ και για κάθε $t \in [0,1]$ ισχύει η σχέση

$$f(\gamma_t) \leqslant (1-t)f(\gamma_0) + tf(\gamma_1) - \frac{K}{2}t(1-t)d^2(\gamma_0,\gamma_1).$$

Οι $K-$κυρτές συναρτήσεις στον $\mathscr{P}^2(M)$ θα λέγονται $K-$κυρτές συναρτήσεις ως προς μετατόπιση.

Θεώρημα 5.2.5. *Έστω (M,g) λεία και συνεκτική πολλαπλότητα Riemann και $K \in \mathbb{R}$. Τα εξής είναι ισοδύναμα:*

(i) $\mathrm{Ric}(M) \geqslant K$, *δηλαδή* $\mathrm{Ric}_x(v,v) \geqslant K|v|^2$ *για κάθε* $x \in M$, $v \in T_xM$.

(ii) *Η εντροπία* $\mathrm{Ent}(\cdot)$ *είναι $K-$κυρτή συνάρτηση ως προς μετατόπιση στον $\mathscr{P}^2(M)$, δηλαδή ισχύει η σχέση:*

$$Ent(\mu_t) \leqslant (1-t)Ent(\mu_0) + tEnt(\mu_1) - \frac{K}{2}t(1-t)W_2(\mu_0,\mu_1)^2.$$

Για κάθε μέτρο πιθανότητας μ ισχύει η σχέση $Ent(\mu) \geqslant -\ln(\mathrm{vol}(\mathrm{supp}(\mu)))$, με την ισότητα να ισχύει όταν το μ είναι το ομοιόμορφο μέτρο στον φορέα του. Παίρνουμε τώρα για μ_0 και μ_1 τα ομοιόμορφα μέτρα πιθανότητας στα σύνολα A_0 και A_1 αντίστοιχα και η παραπάνω σχέση δίνει την ανισότητα

$$\ln(\mathrm{vol}(\mathrm{supp}(\mu_t))) \geqslant (1-t)\ln(\mathrm{vol}(A_0)) + t\,\ln(\mathrm{vol}(A_1)) + \frac{K}{2}t(1-t)W_2(\mu_0,\mu_1)^2.$$

Αυτή η ανισότητα μοιάζει αρκετά με την ανισότητα Brunn-Minkowski που αναφέραμε προηγουμένως.

Η σχέση μεταξύ της coarse καμπυλότητας Ricci και της ιδιότητας της κυρτότητας ως προς μετατόπιση δεν είναι ξεκάθαρη.

5.3 Ανισότητα Brunn–Minkowski χωρίς καμπυλότητα

Δίνουμε πρώτα κάποιους ορισμούς που θα χρειαστούν στη συνέχεια και αποδεικνύουμε το Θεώρημα 5.1.1 στην περίπτωση που $\kappa = 0$. Δηλαδή, θεωρούμε δύο μη κενά υποσύνολα A και B του $X = \{0,1\}^N$, το σύνολο M των μέσων των A και B, και θέλουμε να δείξουμε ότι

$$\ln(\mathrm{card}(M)) \geqslant \frac{1}{2}(\ln(\mathrm{card}(A)) + \ln(\mathrm{card}(B))),$$

ή, ισοδύναμα, ότι

$$\mathrm{card}(M) \geqslant \sqrt{\mathrm{card}(A)\cdot \mathrm{card}(B)}.$$

Έστω $a = (a_1,\dots,a_N) \in A$ και $b = (b_1,\dots,b_N) \in B$. Ένα μέσο $m = (m_1,\dots,m_N)$ των a και b είναι μια N-άδα τέτοια ώστε $m_i = a_i$ αν $a_i = b_i$ και τέτοια ώστε οι μισές από τις υπόλοιπες συντεταγμένες m_i να συμπίπτουν με αυτές του a και οι άλλες μισές με αυτές του b. Συμβολίζουμε με $r = d(a,b)$ το πλήθος των διαφορετικών συντεταγμένων των a και b. Για σταθερά a και b, υπάρχει μια ένα-προς-ένα αντιστοιχία ανάμεσα στα μέσα m των a και b και τα υποσύνολα $c \subset \{1,\dots,r\}$ που έχουν πληθάριθμο $r/2$ αν ο r είναι άρτιος και πληθάριθμο $\frac{r}{2} \pm \frac{1}{2}$ αν ο r είναι περιττός. Το σύνολο c περιγράφει το σύνολο των i για τα οποία $a_i \neq b_i$ που επιλέγονται από το a για να οριστεί ένα μέσο m των a και b.

Ονομάζουμε r-*πέρασμα* ένα $c \subset \{1,\dots,r\}$ που ικανοποιεί την $|\mathrm{card}(c) - r/2| \leqslant 1/2$. Συμβολίζουμε επίσης με $m = \varphi_c(a,b)$ το μέσο των a και b που ορίζεται από το πέρασμα c. Για κάθε πέρασμα c, συμβολίζουμε με $\overline{c}$ το συμπλήρωμά του, $\{1,\dots,r\} \setminus c$, το οποίο είναι επίσης πέρασμα.

Παρατηρούμε ότι, για δοθέν $d(a,b)$-πέρασμα c, το ζεύγος $\Phi_c(a,b) := (\varphi_c(a,b), \varphi_{\overline{c}}(a,b)) = (m,m')$ μας επιτρέπει να προσδιορίσουμε τα a και b. Πράγματι, οι συντεταγμένες των m και m' που συμπίπτουν είναι οι ίδιες με τις αντίστοιχες των a και b, ενώ εκείνες που διαφέρουν είναι επίσης διαφορετικές στα a και b, και γνωρίζοντας το πέρασμα c μπορούμε να προσδιορίσουμε ποιές από αυτές προέρχονται από το a και ποιες από το b. Μάλιστα, οι ρόλοι των (a,b) και (m,m') είναι συμμετρικοί, δηλαδή η διαδικασία αποκωδικοποίησης ισοδυναμεί με την διαδικασία κωδικοποίησης: έχουμε $(a,b) = \Phi_c(\Phi_c(a,b))$.

Τώρα, για κάθε $r \in \{0,\dots,N\}$ ορίζουμε το συγκεκριμένο r-πέρασμα $c_r := \{1,2,\dots,\lfloor r/2 \rfloor\}$. Τότε η απεικόνιση $(a,b) \mapsto \Phi_{c_{d(a,b)}}(a,b)$ από το $A \times B$ στο $M \times M$, όπου M είναι το σύνολο των μέσων των A και B, είναι ένα-προς-ένα. Συνεπώς, $\mathrm{card}(A\times B) \leqslant \mathrm{card}(M \times M)$, όπως θέλαμε.

Θα χρειαστούμε μια χρήσιμη ιδιότητα των απεικονίσεων κωδικοποίησης φ_c και Φ_c. Αν $\Phi_c(a,b) = (m,m')$, γράφουμε $a = \varphi_c^{-1}(m,m')$ και $b = \varphi_{\bar{c}}^{-1}(m,m') = \varphi_v^{-1}(m',m)$. Στο σύνολο C_r θεωρούμε την απόσταση

$$d(c,c') = \mathrm{card}(c \triangle c').$$

Τότε, έχουμε την ακόλουθη πρόταση.

Πρόταση 5.3.1. *Έστω $m,m' \in \{0,1\}^N$ και $c_1,c_2 \in C_{d(m,m')}$. Αν $a_1 = \varphi_{c_1}^{-1}(m,m')$ και $a_2 = \varphi_{c_2}^{-1}(m,m')$, τότε*

$$d(a_1,a_2) = d(c_1,c_2).$$

Απόδειξη. Για δοθέντα m και m', αν αλλάξουμε το πέρασμα c τότε η προεικόνα $\varphi_c^{-1}(m,m')$ αλλάζει εξίσου. Ακριβέστερα, αν θέσουμε $r = d(m,m')$ τότε αφού τα 0 και 1 έχουν τον ίδιο ρόλο στον διακριτό κύβο, μπορούμε να υποθέσουμε ότι $m = 0^N$ αι $m' = 1^r 0^{N-r}$.

Αν $c \in C_r$, έχουμε $a_1 = \varphi_{c_1}^{-1}(m,m') = w(c_1)0^{N-r}$, όπου $w(c_1) \in \{0,1\}^r$ είναι η r-άδα που η i-οστή συντεταγμένη της είναι ίση με 0 αν $i \in c_1$ και ίση με 1 αν $i \notin c_1$. Όμοια, $a_2 = w(c_2)0^{N-r}$. Τότε,

$$d(a_1,a_2) = d(w(c_1),w(c_2)) = \mathrm{card}(c_1 \triangle c_2),$$

όπως θέλαμε. □

5.4 Συγκέντρωση στο σύνολο των περασμάτων

Για να πετύχουμε μια βελτιωμένη ανισότητα υποθέτοντας ότι έχουμε θετική καμπυλότητα κ, χρειάζεται να μελετήσουμε κάποιες γεωμετρικές ιδιότητες του συνόλου των περασμάτων. Ακριβέστερα, θα δείξουμε ότι στο σύνολο αυτό εμφανίζεται συγκέντρωση του μέτρου. Το γεγονός αυτό θα προκύψει από τη συγκέντρωση του μέτρου στην ομάδα των μεταθέσεων. Αρχικά περιγράφουμε τη συγκέντρωση του μέτρου στην ομάδα των μεταθέσεων (η ακριβής μορφή που θα χρησιμοποιήσουμε προέρχεται από το [BHT06]).

Λήμμα 5.4.1. *Έστω S_n η ομάδα των μεταθέσεων στο σύνολο $\{1,\dots,n\}$. Εφοδιάζουμε την S_n με τη μετρική*

$$d(\sigma,\tau) = \mathrm{card}(\{i : \sigma(i) \neq \tau(i)\}), \qquad \sigma,\tau \in S_n.$$

Θεωρούμε επίσης το ομοιόμορφο μέτρο πιθανότητας ν στην S_n. Αν $f : S_n \to \mathbb{R}$ είναι μια 1-Lipschitz συνάρτηση, τότε η f ικανοποιεί την ανισότητα συγκέντρωσης

$$\nu\left(\left\{f \geqslant \int f\,d\nu + t\right\}\right) \leqslant e^{-t^2/2(n-1)}$$

για κάθε $t \geqslant 0$, καθώς και την εκτίμηση Laplace

$$\int e^{\lambda f} d\nu \leqslant e^{\lambda \int f\,d\nu + (n-1)\lambda^2/2}$$

για κάθε $\lambda \in \mathbb{R}$.

Πρόταση 5.4.2. *Έστω $n \geqslant 1$ και C_n το σύνολο των $c \subset \{1,\dots,n\}$ που ικανοποιούν την $|\mathrm{card}(c) - n/2| < 1$. Εφοδιάζουμε το C_n με την απόσταση $d(c,c') = \mathrm{card}(c \triangle c')$ και το ομοιόμορφο μέτρο πιθανότητας μ.*

Αν $f : C_n \to \mathbb{R}$ είναι μια 1-Lipschitz συνάρτηση, τότε η f ικανοποιεί την ανισότητα συγκέντρωσης

$$\mu\left(\left\{f \geqslant \int f\,d\mu + t\right\}\right) \leqslant e^{-t^2/2n}$$

για κάθε $t \geqslant 0$, καθώς και την εκτίμηση Laplace

$$\int e^{\lambda f}\,d\nu \leqslant e^{\lambda \int f\,d\nu + n\lambda^2/2}$$

για κάθε $\lambda \in \mathbb{R}$.

Απόδειξη. Εξετάζουμε πρώτα την περίπτωση που ο n είναι άρτιος. Τότε, η φυσιολογική δράση της S_n στο $\{1,\dots,n\}$ διατηρεί το C_n. Σταθεροποιούμε το $c_0 := \{1,\dots,n/2\} \in C_n$ και ορίζουμε την προβολή $\pi : S_n \to C_n$ με $\pi(\sigma) = \sigma(c_0)$. Όλες οι ίνες της π έχουν τον ίδιο πληθάριθμο $((n/2)!)^2$.

Επιπλέον, αν εφοδιάσουμε τα S_n και C_n με τις αποστάσεις που ορίσαμε παραπάνω, η απεικόνιση π είναι 1-Lipschitz. Πράγματι, για κάθε $\sigma,\tau \in S_n$ έχουμε

$$\begin{aligned} d(\pi(\sigma),\pi(\tau)) = d(\sigma(c_0),\tau(c_0)) &= \sum_{x\in\{1,\dots n\}} |\mathbf{1}_{x\in\sigma(c_0)} - \mathbf{1}_{x\in\tau(c_0)}| \\ &= \sum_{x\in\{1,\dots n\}} |\mathbf{1}_{\sigma^{-1}(x)\in c_0} - \mathbf{1}_{\tau^{-1}(x)\in c_0}| \leqslant \sum_{x\in\{1,\dots n\}} \mathbf{1}_{\sigma^{-1}(x)\neq\tau^{-1}(x)} \\ &= d(\sigma^{-1},\tau^{-1}) = d(\sigma,\tau). \end{aligned}$$

Συνεπώς, αν η $f : C_n \to \mathbb{R}$ είναι 1-Lipschitz, τότε η $f_1 := f \circ \pi$ είναι 1-Lipschitz στην S_n. Τότε, η f_1 ικανοποιεί την ανισότητα συγκέντρωσης

$$\nu\left(\left\{f_1 \geqslant \int f_1\,d\nu + t\right\}\right) \leqslant e^{-t^2/2(r-1)},$$

όπου ν είναι το ομοιόμορφο μέτρο πιθανότητας στην S_n. Αφού όλες οι ίνες της π έχουν τον ίδιο πληθάριθμο, η π απεικονίζει το ν στο ομοιόμορφο μέτρο μ, άρα η ίδια εκτίμηση ισχύει για την $f : C_n \to \mathbb{R}$ ως προς το μ. Το ίδιο ακριβώς επιχείρημα δουλεύει για την εκτίμηση Laplace.

Στην περίπτωση που ο n είναι περιττός, σταθεροποιούμε το $c_0 = \{1,\dots,\lfloor n/2\rfloor\} \in C_n$ και το $c_1 = \{1,\dots,\lceil n/2\rceil\} \in C_n$. Θεωρούμε το σύνολο $S_n^* : (S_n \times \{0\}) \sqcup (S_n \times \{1\})$ και ορίζουμε την απεικόνιση $\pi : S_n^* \to C_n$ με $\pi(\sigma,i) = \sigma(c_i)$ για $i = 0,1$. Τότε, όλες οι ίνες της π έχουν τον ίδιο πληθάριθμο $\lfloor n/2\rfloor!\,\lceil n/2\rceil!$. Εφοδιάζουμε το S_n^* με τη μετρική $d((\sigma,i),(\tau,j)) = |i-j| + d(\sigma,\tau)$. Μπορούμε τότε να ελέγξουμε ότι η $\pi : S_n^* \to C_n$ είναι 1-Lipschitz.

Για δοθείσα 1-Lipschitz συνάρτηση $f : C_n \to \mathbb{R}$ θεωρούμε τη συνάρτηση $f_1 := f \circ \pi : S_n^* \to \mathbb{R}$. Χρησιμοποιώντας κλασικές τεχνικές, δείχνουμε ότι έχουμε συγκέντρωση του μέτρου στην S_n^* και ότι η f_1 ικανοποιεί την εκτίμηση Laplace

$$\int e^{\lambda f_1}\,d\nu \leqslant e^{\lambda\int f_1\,d\nu + (r-1)\lambda^2/2 + \lambda^2/8} \leqslant e^{\lambda\int f_1\,d\nu + r\lambda^2/2},$$

όπου ν είναι το ομοιόμορφο μέτρο πιθανότητας στην S_n^*. Έπεται ότι $\nu(\{f_1 \geqslant \int f_1\,d\nu + t\}) \leqslant e^{-t^2/2r}$. Όπως πριν, βλέπουμε ότι η ίδια εκτίμηση ισχύει για την f στο C_n. $\square$

Πόρισμα 5.4.3. *Έστω A ένα υποσύνολο του συνόλου των περασμάτων C_n και έστω $\overline{A} := \{\overline{c} : c \in A\}$. Υποθέτουμε ότι $d(A, \overline{A}) \geqslant k$. Τότε,*

$$\mathrm{card}(A) \leqslant e^{-k^2/8n}\mathrm{card}(C_n).$$

Απόδειξη. Θεωρούμε τη συνάρτηση $f : C_n \to \mathbb{R}$ με $f(c) := \frac{1}{2}(d(c, \overline{A}) - d(c, A))$. Η f είναι 1-Lipschitz και παίρνει τιμές μεγαλύτερες ή ίσες από $k/2$ στο A. Λόγω συμμετρίας, η μέση τιμή της f είναι 0. Εφαρμόζοντας την Πρόταση 5.4.2 βλέπουμε ότι το σχετικό μέτρο του A στο C_n είναι το πολύ ίσο με $e^{-k^2/8n}$. □

Στη συνέχεια αποδεικνύουμε μια εκλέπτυνση του Πορίσματος 5.4.3, στην οποία αντικαθιστούμε το σύνολο A με ένα μέτρο ξ, τους πληθαρίθμους με εντροπίες, και την απόσταση $d(A, \overline{A})$ με την απόσταση Wasserstein $W_1(\xi, \overline{\xi})$.

Πόρισμα 5.4.4. *Έστω ξ ένα μετρο πιθανότητας στο σύνολο των περασμάτων C_n. Θεωρούμε το συμπλήρωμα $\overline{\xi}$ του ξ, το οποίο ορίζεται από την*

$$\overline{\xi}(c) := \xi(\overline{c}), \qquad c \in C_n.$$

Τότε,

$$S(\xi) \leqslant \ln(\mathrm{card}(C_n)) - \frac{1}{8n} W_1(\xi, \overline{\xi})^2,$$

όπου S είναι η εντροπία Shannon.

Για την απόδειξη χρειαζόμαστε μια αναδιατύπωση της Πρότασης 5.4.2.

Λήμμα 5.4.5. *Για κάθε μέτρο πιθανότητας ξ στο C_n ισχύει*

$$W_1(\xi, \mu)^2 \leqslant 2nH(\xi|\mu),$$

όπου μ είναι το ομοιόμορφο μέτρο πιθανότητας στο C_n και H είναι η σχετική εντροπία.

Απόδειξη του Πορίσματος 5.4.4. Λόγω συμμετρίας έχουμε

$$W_1(\xi, \overline{\xi}) \leqslant W_1(\xi, \mu) + W_1(\mu, \overline{\xi}) = 2W_1(\xi, \mu).$$

Συνεπώς,

$$H(\xi|\mu) \geqslant \frac{1}{8n} W_1(\xi, \overline{\xi})^2.$$

Τέλος, χρησιμοποιώντας την $H(\xi|\mu) = \ln(\mathrm{card}(C_n)) - S(\xi)$ ξαναγράφουμε την τελευταία ανισότητα στην ισοδύναμη μορφή

$$S(\xi) \leqslant \ln(\mathrm{card}(C_n)) - \frac{1}{8n} W_1(\xi, \overline{\xi})^2,$$

συναρτήσει της εντροπίας Shannon. □

Απόδειξη του Θεωρήματος 5.1.1. Θεωρούμε δύο μη κενά υποσύνολα A και B του $X = \{0,1\}^N$, και το σύνολο M των μέσων των A και B. Πρέπει να δείξουμε ότι, για $\kappa = \frac{1}{2N}$,

$$\ln(\mathrm{card}(M)) \geqslant \frac{1}{2}(\ln(\mathrm{card}(A) + \ln(\mathrm{card}(B)) + \frac{\kappa d(A,B)^2}{8}.$$

Η διαφορά με την περίπτωση όπου $\kappa = 0$ είναι ότι τώρα θεωρούμε όλα τα περάσματα ταυτόχρονα. Έστω C_r το σύνολο των r-περασμάτων. Ορίζουμε

$$Y := \{(a,b,c) : a \in A, b \in B, c \in C_{d(a,b)}\}.$$

Θεωρούμε την απεικόνιση $f : Y \to M \times M$ που ορίζεται από την

$$f(a,b,c) = \Phi_c(a,b).$$

Η απεικόνιση f δεν είναι απαραίτητα ένα-προς-ένα. Θα δείξουμε όμως ότι υπάρχει κάποιος έλεγχος στο πόσο απέχει από το να είναι ένα-προς-ένα. Η ιδέα είναι ότι, αν μας δοθεί ένα ζεύγος μέσων (m,m'), η γεωμετρία των A και B μας επιτρέπει να μαντέψουμε κατά κάποιον τρόπο το πέρασμα που χρησιμοποιήθηκε, και έτσι ο πληθάριθμος του $f^{-1}(m,m')$ είναι φραγμένος.

Ορίζουμε $Y_r := \{(a,b,c) \in Y : d(a,b) = r\}$ και $(M \times M)_r := \{(m,m') \in M \times M : d(m,m') = r\}$. Σταθεροποιούμε $(m,m') \in (M \times M)_r$ και παρατηρούμε ότι η ίνα $f^{-1}(m,m')$ είναι σε ένα-προς-ένα αντιστοιχία με το σύνολο E των περασμάτων $c \in C_r$ για τα οποία $\Phi_c^{-1}(m,m') \in A \times B$. Όμοια, για το σύνολο $E' = \{c \in C_r : \Phi_c^{-1}(m,m') \in B \times A\}$. Από τον ορισμό έχουμε $\Phi_c = (\varphi_c, \varphi_{\overline{c}})$, άρα τα στοιχεία του E' είναι τα συμπληρώματα των στοιχείων του E.

Παρατηρούμε ότι $d(E,E') \geqslant d(A,B)$. Πράγματι, αν $c \in E$ και $c' \in E'$ τότε $\varphi_c^{-1}(m,m') \in A$ και $\varphi_{c'}^{-1}(m,m') \in B$. Γνωρίζουμε ότι η αποκωδικοποίηση είναι ισομετρική, άρα $d(c,c') \geqslant d(A,B)$.

Από το Πόρισμα 5.4.3 έχουμε $\mathrm{card}(E) \leqslant \mathrm{card}(C_r) \cdot e^{-d(A,B)^2/8r}$. Αφού ο πληθάριθμος του E ισούται με τον πληθάριθμο της ίνας $f^{-1}(m,m')$, έπεται ότι η αντίστροφη εικόνα οποιουδήποτε σημείου της $f(Y_r)$ έχει το πολύ $\mathrm{card}(C_r) \cdot e^{-d(A,B)^2/8r}$ στοιχεία. Συνεπώς,

$$\mathrm{card}(Y_r) \leqslant \mathrm{card}(C_r) \cdot e^{-d(A,B)^2/8r} \cdot \mathrm{card}((M \times M)_r).$$

Θέτοντας $(A \times B)_r := \{(a,b) \in A \times B : d(a,b) = r\}$ έχουμε

$$\mathrm{card}(Y_r) \leqslant \mathrm{card}((A \times B)_r) \cdot \mathrm{card}(C_r),$$

άρα

$$\mathrm{card}((M \times M)_r) \geqslant e^{d(A,B)^2/8r} \mathrm{card}((A \times B)_r).$$

Τέλος, αθροίζοντας πάνω από όλους τους $r \in \{1,\ldots,N\}$ βλέπουμε ότι

$$\mathrm{card}(M \times M) \geqslant e^{d(A,B)^2/8N} \mathrm{card}(A \times B),$$

το οποίο αποδεικνύει το θεώρημα. □

5.5 Εντροπία στο σύνολο των μέσων

Σε αυτήν την παράγραφο αποδεικνύουμε το Θεώρημα 5.1.2. Υπενθυμίζουμε ότι αν a και b είναι δύο σημεία στον διακριτό κύβο X, το μέτρο των μέσων $\mathrm{mid}(a,b)$ είναι το ομοιόμορφο μέτρο πιθανότητας στο σύνολο όλων των μέσων των a και b. Το μέτρο των μέσων δύο μέτρων πιθανότητας μ_A και μ_B ορίζεται από την

$$\mathrm{mid}(\mu_A,\, u_B) := \iint \mathrm{mid}(a,b)\, d\mu_A(a)\, d\mu_B(b),$$

δηλαδή είναι ο μέσος όρος των $\mathrm{mid}(a,b)$, όπου τα a και b επιλέγονται τυχαία και ανεξάρτητα, με κατανομές μ_A και μ_B αντίστοιχα.

Η απόδειξη του Θεωρήματος 5.1.2 ακολουθεί την ίδια γραμμή με αυτήν του Θεωρήματος 5.1.1, με τη διαφορά ότι την θέση των συνόλων παίρνουν τα μέτρα πιθανότητας. Θα βοηθήσει να σκέφτεται κανείς τα μέτρα πιθανότητας σαν σύνολα των οποίων τα σημεία έχουν βάρη, με την εντροπία Shannon αυτών των μέτρων πιθανότητας να παίζει το ρόλο του πληθαρίθμου τους.

Ένα βασικό σημείο στην προηγούμενη απόδειξη ήταν η εκτίμηση του πληθαρίθμου των ινών της απεικόνισης $(a,b,c) \mapsto (m,m') = \Phi_c(a,b)$. Από αυτήν προέκυψε το κάτω φράγμα για τον πληθάριθμο του συνόλου $\{(m,m')\}$. Εδώ, θα χρησιμοποιηθεί η προσεταιριστικότητα της εντροπίας Shannon μέσω της οποίας θα εκφραστεί η ίδια σχέση, και αυτό θα μας δώσει κάτω φράγμα για την εντροπία του (m,m') αν η εντροπία των ινών είναι γνωστή.

Στο τελευταίο βήμα χρησιμοποιείται η $W_1(\mu_a,\mu_B)$ αντί της $d(A,B)$. Στην προηγούμενη απόδειξη χρησιμοποιήσαμε την απεικόνιση $c \mapsto \overline{c}$ και το γεγονός ότι $\Phi_c(a,b) = \Phi_{\overline{c}}(a,b)$ για να συμπεράνουμε ότι αν $\Phi_c(a,b) = \Phi_{c'}(a',b')$ τότε $d(\overline{c},c') = d(b,a') \geqslant d(A,B)$. Κατόπιν, χρησιμοποιήσαμε το Πόρισμα 5.4.3 για να φράξουμε τον πληθάριθμο του συνόλου E αυτών των περασμάτων c σε μια ίνα. Εδώ, θα χρησιμοποιήσουμε τη σχέση $d(\overline{c},c') = d(b,a')$ για να μετατρέψουμε κάθε ταίριασμα των E και $\overline{E}$ σε ταίριασμα των A και B με την ίδια απόσταση μεταφοράς. Κατόπιν, χρησιμοποιούμε το Πόρισμα 5.4.4 στη θέση του Πορίσματος 5.4.3 και παίρνουμε ένα φράγμα για την εντροπία των περασμάτων c σε μια ίνα.

Απόδειξη του Θεωρήματος 5.1.2. Έστω a και b ανεξάρτητες τυχαίες μεταβλητές με κατανομή μ_A και μ_B αντίστοιχα. Συμβολίζουμε, όπως πριν, με C_r το σύνολο των r-περασμάτων. Έστω c μια τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο $C_{d(a,b)}$, η οποία είναι ανεξάρτητη από τις a και b δεδομένης της $d(a,b)$. Ορίζουμε επίσης τις τυχαίες μεταβλητές $m := \varphi_c(a,b)$ και $m' := \varphi_{\overline{c}}(a,b)$. Συνεπώς, η κατανομή της m, όπως και της m', είναι το $\mathrm{mid}(\mu_A,\mu_B)$.

Συμβολίζουμε με $S((y))$ την εντροπία Shannon της κατανομής μιας τυχαίας μεταβλητής y. Έχουμε $S((m,m')) \leqslant S((m)) + S((m'))$, και αφού οι m και m' έχουν την ίδια κατανομή $\mathrm{mid}(\mu_A,\mu_B)$, παίρνουμε

$$S(\mathrm{mid}(\mu_A,\mu_B)) \geqslant \frac{1}{2} S((m,m')).$$

Θεωρούμε την απεικόνιση Φ που στέλνει το (a,b,c) στο $\Phi_c(m,m')$. Έστω $Y_{(m,m')}$ η κατανομή του (a,b,c) δεδομένου του (m,m'). Από την προσεταιριστικότητα της εντροπίας, η εντροπία Shannon της κατανομής του (m,m') είναι ίση με την εντροπία της κατανομής του (a,b,c) μείον τη μέση εντροπία των ινών της Φ, δηλαδή,

$$S((m,m')) = S((a,b,c)) - \mathbb{E}\, S(Y_{(m,m')}).$$

Ο πρώτος όρος υπολογίζεται ως εξής. Οι τυχαίες μεταβλητές a και b είναι ανεξάρτητες, και με δεδομένη την $d(a,b)$, η τυχαία μεταβλητή c είναι ανεξάρτητη από τις a και b, και έχει κατανομή την ομοιόμορφη κατανομή $U_{d(a,b)}$ στο $C_{d(a,b)}$. Άρα,

$$S((a,b,c)) = S((a)) + S((b)) + \mathbb{E}\, S(U_{d(a,b)}) = S(\mu_A) + S(\mu_B) + \mathbb{E}\, \ln(C_{d(a,b)}).$$

Μελετάμε τώρα τον δεύτερο όρο $\mathbb{E}\, S(Y_{(m,m')})$. Έστω $E_{(m,m')}$ η κατανομή της c με δεδομένο το (m,m'), δηλαδή η τρίτη περιθώρια κατανομή της $Y_{(m,m')}$. Με δεδομένο το (m,m'), η τιμή του c προσδιορίζει τα a και b, άρα

$$S((a,b,c)|(m,m')) = S((c)|(m,m')),$$

δηλαδή

$$S(Y_{(m,m')}) = S(E_{(m,m')}).$$

Συνεπώς,

$$S((m,m')) = S(\mu_A) + S(\mu_B) + \mathbb{E}\, \ln(C_{d(a,b)}) - \mathbb{E}\, S(E_{(m,m')}).$$

Τώρα, χρησιμοποιώντας την απλή ανισότητα $S(E_{(m,m')}) \leqslant \ln(C_{d(m,m')})$ και παίρνοντας υπόψιν την $d(a,b) = d(m,m')$, παίρνουμε

$$S((m,m')) \geqslant S(\mu_A) + S(\mu_B) + \mathbb{E}\, \ln(C_{d(a,b)}) - \mathbb{E}\, \ln(C_{d(m,m')}) = S(\mu_A) + S(\mu_B).$$

Έπεται ότι

$$S((m)) \geqslant \frac{1}{2}(S(\mu_A) + S(\mu_B)),$$

δηλαδή έχουμε την περίπτωση $\kappa = 0$ του θεωρήματος.

Θα δείξουμε ότι η $E_{(m,m')}$ έχει μικρή εντροπία Shannon χρησιμοποιώντας τις ιδιότητες συγκέντρωσης του συνόλου των περασμάτων. Από το Πόρισμα 5.4.4 έχουμε

$$S(E_{(m,m')}) \leqslant \ln(\mathrm{card}(C_{d(m,m')})) - \frac{1}{8d(m,m')} W_1(E_{(m,m'),\overline{E}_{(m,m')}})^2,$$

όπου $\overline{E}_{(m,m')}$ είναι η εικόνα του $E_{(m,m')}$ μέσω της $c \mapsto \overline{c}$. Άρα, χρειάζεται να εκτιμήσουμε την απόσταση των $E_{(m,m')}$ και $\overline{E}_{(m,m')}$. Μάλιστα, χρειάζεται να εκτιμήσουμε μόνο τη μέση τιμή αυτής της απόστασης πάνω από τα (m,m'). Ο ισχυρισμός είναι ότι

$$\mathbb{E}\, W_1(E_{(m,m')}, \overline{E}_{(m,m')})^2 \geqslant W_1(\mu_A, \mu_B)^2.$$

Για την απόδειξη αυτής της ανισότητας, σταθεροποιούμε το (m,m') και συμβολίζουμε με $A_{(m,m')}$ και $B_{(m,m')}$ τις κατανομές των a και b αντίστοιχα, με δεδομένο το (m,m'). Αφού $a = \varphi_c^{-1}(m,m')$ και $b = \varphi_{\overline{c}}^{-1}(m,m')$, κάθε ταίριασμα των $E_{(m,m')}$ και $\overline{E}_{(m,m')}$ προσδιορίζει ένα ταίριασμα των $A_{(m,m')}$ και $A_{(m,m')}$. Επιπλέον, αφού η αποκωδικοποίηση είναι ισομετρική (από την Πρόταση 5.3.1), αυτά τα ταιριάσματα ορίζουν την ίδια απόσταση μεταφοράς. Έτσι προκύπτει η

$$W_1(A_{(m,m')}, B_{(m,m')}) \leqslant W_1(E_{(m,m')}, \overline{E}_{(m,m')}).$$

Αν για κάθε (m,m') έχουμε ένα ταίριασμα των $A_{(m,m')}$ και $B_{(m,m')}$, προσθέτονας παίρνουμε ένα ταίριασμα των μ_A και μ_B. Άρα,

$$W_1(\mu_A, \mu_B) \leqslant \mathbb{E}\, W_1(E_{(m,m')}, \overline{E}_{(m,m')}).$$

Λόγω κυρτότητας έπεται η

$$W_1(\mu_A,\mu_B)^2 \leqslant \mathbb{E}\, W_1(E_{(m,m')},\overline{E}_{(m,m')}).$$

Συνδυάζοντας όλα τα παραπάνω και παίρνοντας υπόψιν την $d(m,m')=d(a,b)$ έχουμε

$$\begin{aligned} S((m,m')) &= S((a,b,c)) - \mathbb{E}\, S(Y_{(m,m')}) \\ &= S(\mu_A)+S(\mu_B)+\mathbb{E}\,\ln(\mathrm{card}(C_{d(a,b)})) - \mathbb{E}\, S(E_{(m,m')}) \\ &\geqslant S(\mu_A)+S(\mu_B)+\mathbb{E}\,\ln(\mathrm{card}(C_{d(a,b)})) - \mathbb{E}\,\ln(\mathrm{card}(C_{d(m,m')})) \\ &\qquad + \mathbb{E}\left[\frac{W_1(E_{(m,m')},\overline{E}_{(m,m')})^2}{8d(m,m')}\right] \\ &\geqslant S(\mu_A)+S(\mu_B)+\frac{1}{8N}\mathbb{E}\, W_1(E_{(m,m')},\overline{E}_{(m,m')})^2 \\ &\geqslant S(\mu_A)+S(\mu_B)+\frac{1}{8N}W_1(\mu_A,\mu_B)^2, \end{aligned}$$

άρα

$$S((m)) \geqslant \frac{1}{2}(S(\mu_A)+S(\mu_B)) + \frac{1}{16N}W_1(\mu_A,\mu_B)^2,$$

το οποίο ολοκληρώνει την απόδειξη. □

Παρατηρήσεις 5.5.1. Ένας πρώτος περιορισμός στα αποτελέσματα που παρουσιάστηκαν σε αυτό το κεφάλαιο είναι ότι αναγκαζόμαστε να μελετήσουμε μόνο την περίπτωση $t=1/2$. Αυτό οφείλεται στη συνδυαστική φύση των επιχειρημάτων που χρησιμοποιούνται, τα οποία στην πιο βασική περίπτωση όπου $\kappa=0$ έχουν ως κέντρο τον ορισμό κατάλληλης ένα-προς-ένα απεικόνισης από το $A\times B$ στο $M\times M$.

Θα μπορούσαμε ισως να ξεπεράσουμε αυτή τη δυσκολία αν υποθέταμε ότι τα σύνολα A και B είναι κυρτά (με την έννοια ότι το μέσο δύο σημείων του A ανήκει στο A, και όμοια για το B). Πράγματι, σε αυτήν την περίπτωση (και αν αγνοήσουμε το τεχνικό πρόβλημα της στρογγυλοποίησης των αποστάσεων στον πλησιέστερο ακέραιο), αν M είναι το σύνολο των $1/2$-μέσων των A και B, τότε ένα μέσο των A και M είναι $1/4$-μέσο των A και B και ούτω καθεξής, οπότε μπορούμε να περιγράψουμε τα t-μέσα των A και B ως επαναλαμβανόμενα $1/2$-μέσα και να χρησιμοποιήσουμε το Θεώρημα 5.1.1 σε κάθε βήμα. Αυτό φαίνεται να δουλεύει καλά αν $\kappa=0$ αλλά δεν είναι καθαρό το τι θα συμβεί με τον όρο της απόστασης όταν $\kappa>0$. Από την άλλη πλευρά, αν τα A και B δεν είναι κυρτά, με τις επαναλήψεις προκύπτουν «μέσα» πολλών σημείων του A και πολλών σημείων του B, κάτι που δεν μπορούμε να χειριστούμε.

Η ένα-προς-ένα απεικόνιση από το $A\times B$ στο $M\times M$ που χρησιμοποιήθηκε στην απόδειξη επεκτείνεται φυσιολογικά σε αντίστοιχη απεικόνιση από το $A\times B$ στο $M_t\times M_{(1-t)}$, όπου M_t είναι το σύνολο των t-μέσων. Αυτό οδηγεί σε ένα κάτω φράγμα για το $\ln(\mathrm{card}(M_t))+\ln(\mathrm{card}(M_{(1-t)}))$ συναρτήσει του $\ln(\mathrm{card}(A))+\ln(\mathrm{card}(B))$ και ενός όρου που εξαρτάται από την καμπυλότητα. Κάτι τέτοιο ισχύει ήδη στην περίπτωση των πολλαπλοτήτων Riemann: προσθέτουμε τις ανισότητες Brunn–Minkowski που έχουμε για το t και το $(1-t)$. Δεν είναι όμως σαφές το πώς μπορεί κανείς να ερμηνεύσει φυσιολογικά μια τέτοια ανισότητα.

ΠΑΡΑΡΤΗΜΑ A

A.1 Γεωμετρία Riemann

Θα εφοδιάσουμε μια γενική διαφορική πολλαπλότητα με μια δομή η οποία θα μας επιτρέψει να μιλάμε για θεμελιώδεις γεωμετρικές έννοιες, όπως γωνίες και μήκη καμπυλών. Η δομή αυτή θα δημιουργήσει μια μετρική στην πολλαπλότητα.

Ορισμός A.1.1 (πολλαπλότητα Riemann)**.** Μια πολλαπλότητα Riemann είναι μια διαφορική πολλαπλότητα M εφοδιασμένη με μια οικογένεια εσωτερικών γινομένων $g = \{\langle\cdot,\cdot\rangle_p\}_{p\in M}$. Η g λέγεται μετρική Riemann. Η μετρική Riemann επάγει τη νόρμα ενός εφαπτόμενου διανύσματος $v \in T_pM$ μέσω του τύπου $\|v\| = \sqrt{\langle v, v\rangle_p}$.

Τώρα, θα εισάγουμε ένα επιπλέον εργαλείο, σε μια γενική διαφορική πολλαπλότητα, το οποίο θα μας οδηγήσει στην έννοια της παράλληλης μετατόπισης.

Ορισμός A.1.2 (γραμμική συνοχή)**.** Έστω M μια διαφορική πολλαπλότητα και $\mathscr{T}(M)$ η εφαπτόμενη δέσμη της M. Μια γραμμική συνοχή στην M είναι μια απεικόνιση $\nabla : \mathscr{T}(M) \times \mathscr{T}(M) \to \mathscr{T}(M)$ με τις ιδιότητες:

(i) Η ∇ είναι γραμμική ως προς την πρώτη μεταβλητή πάνω από το $C^\infty(M)$:

$$\nabla_{fX+gX'}Y = f\nabla_X Y + g\nabla_{X'}Y, \text{ για κάθε } f, g \in C^\infty(M).$$

(ii) Η ∇ είναι γραμμική ως προς την δεύτερη μεταβλητή πάνω από το $\mathbb{R}$:

$$\nabla_X(\lambda Y + \mu Y') = \lambda\nabla_X Y + \mu\nabla_X Y', \text{ για κάθε } \lambda, \mu \in \mathbb{R}.$$

(iii) Ισχύει η σχέση

$$\nabla_X(fY) = f\nabla_X Y + (Xf)\,Y, \text{ για κάθε } f \in C^\infty(M).$$

Πρόταση A.1.3 (ύπαρξη συνοχής)**.** *Κάθε διαφορική πολλαπλότητα επιδέχεται γραμμικής συνοχής.*

Ορισμός A.1.4 (ταχύτητα και συνοχή κατά μήκος καμπύλης)**.** Έστω I διάστημα του $\mathbb{R}$ και $\gamma : I \to M$ μια καμπύλη στην πολλαπλότητα M. Η ταχύτητα της γ την χρονική στιγμή $t \in I$ συμβολίζεται με $\dot{\gamma}(t)$ και ορίζεται ως το εφαπτόμενο διάνυσμα του $T_{\gamma(t)}M$:

$$\dot{\gamma}(t) = \frac{d}{dt}\left(f \circ \gamma\right)(t).$$

Ένα διανυσματικό πεδίο κατά μήκος της γ είναι μια διαφορίσιμη απεικόνιση $V : I \to \mathscr{T}(M)$ ώστε $V(t) \in T_{\gamma(t)}M$, για κάθε $t \in I$. Το σύνολο των διανυσματικών πεδίων κατά μήκος της γ συμβολίζεται με $\mathscr{T}(\gamma)$.

Πρόταση A.1.5 (συναλλοίωτη παράγωγος)**.** *Έστω M διαφορική πολλαπλότητα, $\gamma : I \to M$ καμπύλη, και ∇ γραμμική συνοχή στην M. Τότε η ∇ ορίζει έναν μοναδικό τελεστή*

$$D_t : \mathscr{T}(\gamma) \to \mathscr{T}(\gamma),$$

ο οποίος ικανοποιεί τις εξής ιδιότητες:

(i) *Είναι γραμμικός πάνω από το $\mathbb{R}$:*

$$D_t(\lambda V + \mu W) = \lambda D_t(V) + \mu D_t(W), \text{ για κάθε } \lambda, \mu \in \mathbb{R}.$$

(ii) *Ισχύει ο τύπος:*

$$D_t(fV) = \dot{f}V + fD_t(V), \text{ για κάθε } f \in C^\infty(I).$$

(iii) *Αν το V είναι επεκτάσιμο, τότε για κάθε επέκταση $\widetilde{V}$ του V έχουμε:*

$$D_t(V)(t) = \nabla_{\dot{\gamma}(t)}\widetilde{V}.$$

Το $D_t(V)$ λέγεται συναλλοίωτη παράγωγος του V κατά μήκος της γ.

Ορισμός A.1.6 (γεωδαισιακή ευθεία)**.** Μια καμπύλη $\gamma : I \to M$ λέγεται γεωδαισιακή ως προς την γραμμική συνοχή ∇ αν

$$D_t\dot{\gamma} \equiv 0.$$

Θεώρημα A.1.7 (ύπαρξη και μοναδικότητα γεωδαισιακών)**.** *Έστω M μια διαφορική πολλαπλότητα και ∇ μια γραμμική συνοχή της M. Για κάθε $p \in M$, $v \in T_pM$, και $t_0 \in \mathbb{R}$, υπάρχει ένα ανοικτό διάστημα $I \subset \mathbb{R}$ το οποίο περιέχει το t_0 και γεωδαισιακή $\gamma : I \to M$ η οποία ικανοποιεί τις σχέσεις $\gamma(t_0) = p$ και $\dot{\gamma}(t_0) = v$. Κάθε δύο τέτοιες γεωδαισιακές συμπίπτουν στην τομή των πεδίων ορισμού τους.*

Θεώρημα A.1.8 (παράλληλη μετατόπιση)**.** *Έστω M διαφορική πολλαπλότητα και ∇ γραμμική συνοχή της. Για κάθε καμπύλη $\gamma : I \to M$ και κάθε $t_0 \in I$, και $v_0 \in T_{\gamma(t_0)}M$, υπάρχει μοναδικό παράλληλο διανυσματικό πεδίο V κατά μήκος της γ ώστε $V(t_0) = v_0$. Αυτό το διανυσματικό πεδίο θα ονομάζεται η παράλληλη μετατόπιση του v_0 κατά μήκος της γ.*

Θα συνδέσουμε τώρα τις συνοχές με τη δομή Riemann. Σκοπός μας είναι να καταλήξουμε σε μια συνοχή η οποία θα είναι κανονική για την πολλαπλότητα Riemann.

Ορισμός Α.1.9 (συμβατότητα συνοχής και μετρικής)**.** Έστω (M, g) πολλαπλότητα Riemann. Μια γραμμική συνοχή ∇ της (M, g) θα λέγεται συμβατή με τη μετρική Riemann αν ισχύει η σχέση:

$$\nabla_X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle,$$

για κάθε $X, Y, Z \in \mathscr{T}(M)$.

Ορισμός Α.1.10 (στρέψη)**.** Έστω (M, g) πολλαπλότητα Riemann και ∇ γραμμική συνοχή της πολλαπλότητας. Ο τελεστής στρέψης της πολλαπλότητας είναι ο $T : \mathscr{T}(M) \times \mathscr{T}(M) \to \mathscr{T}(M)$:

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y],$$

για κάθε $X, Y \in \mathscr{T}(M)$, όπου $[\cdot, \cdot]$ είναι το γινόμενο Lie των X, Y. Η συνοχή ∇ θα λέγεται ελευθέρα στρέψης αν ισχύει ότι $T(X, Y) = 0$ για κάθε $X, Y \in \mathscr{T}(M)$.

Θεώρημα Α.1.11 (θεμελιώδες θεώρημα της γεωμετρίας Riemann)**.** *Έστω (M, g) πολλαπλότητα Riemann. Υπάρχει μοναδική γραμμική συνοχή της πολλαπλότητας η οποία είναι συμβατή με την μετρική Riemann και είναι ελευθέρας στρέψης. Αυτή θα ονομάζεται συνοχή Levi-Civita.*

Ορισμός Α.1.12 (εκθετική απεικόνιση)**.** Έστω (M, g) πολλαπλότητα Riemann και $E = \{V \in \mathscr{T}(M) :$ η γεωδαισιακή γ_V ορίζεται σε ένα διάστημα που περιέχει το $[0, 1]\}$. Η εκθετική απεικόνιση ορίζεται ως η συνάρτηση $\exp : E \to M$ με $\exp(V) = \gamma_V(1)$. Για κάθε $p \in M$ συμβολίζουμε με $\exp_p$ τον περιορισμό της εκθετικής απεικόνισης στο σύνολο $E \cup T_p M$.

Ορισμός Α.1.13 (μήκος καμπύλης)**.** Έστω (M, g) πολλαπλότητα Riemann. Το μήκος μιας καμπύλης $\gamma : [\alpha, \beta] \to M$ ορίζεται ως η ποσότητα

$$l(\gamma) = \int_\alpha^\beta \|\dot{\gamma}\| dt.$$

Ορισμός Α.1.14 (μετρική σε πολλαπλότητα Riemann)**.** Έστω (M, g) πολλαπλότητα Riemann. Η απεικόνιση $d : M \times M \to \mathbb{R}$ με:

$$d(p, q) = \inf\{l(\gamma) : \ \gamma \text{ γεωδαισιακή από το } p \text{ στο } q\}$$

είναι μετρική στην πολλαπλότητα M.

Ορισμός Α.1.15 (τελεστής καμπυλότητας)**.** Έστω (M, g) πολλαπλότητα Riemann. Ο τελεστής καμπυλότητας της M είναι ο $R : \mathscr{T}(M) \times \mathscr{T}(M) \times \mathscr{T}(M) \to \mathscr{T}(M)$:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

Ορισμός Α.1.16 (τμηματική καμπυλότητα και καμπυλότητα Ricci)**.** Έστω (M, g) πολλαπλότητα Riemann, $p \in M$ και $v, w \in T_p M$. Η τμηματική καμπυλότητα του επιπέδου που ορίζουν τα διανύσματα v, w είναι η ποσότητα:

$$K(v, w) = -\frac{R(v, w, v, w}{\|v\|^2 \|w\|^2 - \langle v, w\rangle^2}.$$

Έστω $\{e_i\}_{i=1}^n$ ορθοκανονική βάση του $T_p M$, όπου n η διάσταση της πολλαπλότητας. Η καμπυλότητα Ricci του επιπέδου που ορίζουν τα διανύσματα v, w είναι η ποσότητα:

$$Ricc(v, w) = -\sum_{i=1}^n R(v, e_i, w, e_i) = nK(v, w).$$

Ορισμός Α.1.17 (πεδία Jacobi)**.** Έστω (M, g) πολλαπλότητα Riemann, $\gamma : I \to M$ καμπύλη και $V \in \mathscr{T}(\gamma)$ διανυσματικό πεδίο κατά μήκος της γ. Το V Θα λέγεται πεδίο Jacobi αν ικανοποιεί την εξίσωση Jacobi:

$$D_t^2 V + R(V, \dot{\gamma})\dot{\gamma} = 0.$$

Πρόταση Α.1.18 (ύπαρξη και μοναδικότητα πεδίων Jacobi)**.** *Έστω (M, g) πολλαπλότητα Riemann, $\gamma : I \to M$ καμπύλη, $\alpha \in I$ και $p = \gamma(\alpha)$. Για κάθε ζευγάρι εφαπτόμενων διανυσμάτων $v, w \in T_pM$ υπάρχει μοναδικό πεδίο Jacobi $J \in \mathscr{T}(\gamma)$ κατά μήκος της γ το οποίο ικανοποιεί τις αρχικές συνθήκες*

$$J(\alpha) = v, \ D_t J(\alpha) = w.$$

Ορισμός Α.1.19 (συζυγές σημείο)**.** Έστω (M, g) πολλαπλότητα Riemann, $p, q \in M$ και $\gamma : I \to M$ γεωδαισιακή καμπύλη που συνδέει τα p, q. Το q θα λέγεται συζυγές σημείο ως προς το p κατά μήκος της γ αν υπάρχει ένα πεδίο Jacobi κατά μήκος της γ το οποίο μηδενίζεται στα p και q, αλλά δεν είναι ταυτοτικά μηδέν.

Ορισμός Α.1.20 (κανονικά διανυσματικά πεδία κατά μήκος καμπύλης)**.** Έστω (M, g) πολλαπλότητα Riemann, $p, q \in M$ και $\gamma : [\alpha, \beta] \to M$ καμπύλη. Ένα διανυσματικό πεδίο κατά μήκος της γ $V \in \mathscr{T}(\gamma)$ λέγεται κανονικό αν $V(\alpha) = V(\beta) = 0$.

Ορισμός Α.1.21 (index form)**.** Έστω (M, g) πολλαπλότητα Riemann, R ο τελεστής καμπυλότητας της πολλαπλότητας, και $\gamma : I \to M$ καμπύλη. Η index form του συνόλου των κανονικών διανυσματικών πεδίων κατά μήκος της γ δίνεται από τον τύπο

$$I(V, W) = \int_\alpha^\beta \left(\langle D_t V, D_t W \rangle - R(V, \dot{\gamma}, \dot{\gamma}, W) \right) dt.$$

Θεώρημα Α.1.22. *Έστω (M, g) πολλαπλότητα Riemann, $p, q \in M$ και $\gamma : I \to M$ γεωαδαισιακή καμπύλη που συνδέει τα p, q. Αν υπάρχει εσωτερικό σημείο της γεωδαισιακής, το οποίο είναι συζυγές σημείο ως προς το p, τότε υπάρχει κανονικό διανυσματικό πεδίο $V \in \mathscr{T}(\gamma)$ κατά μήκος της γ ώστε $I(V, V) < 0$. Ειδικότερα η γ δεν ελαχιστοποιεί την απόσταση των p, q.*

Α.2 Αλυσίδες Markov

Ορισμός Α.2.1 (στοχαστική διαδικασία)**.** Μια ακολουθία τυχαίων μεταβλητών $\{X_n\}_{n \geqslant 0}$ με τιμές σε ένα σύνολο E θα λέγεται στοχαστική διαδικασία, και το E χώρος καταστάσεων.

Ορισμός Α.2.2 (αλυσίδα Markov)**.** Έστω $\{X_n\}_{n \geqslant 0}$ στοχαστική διαδικασία με χώρο καταστάσεων E. Αν για κάθε φυσικό αριθμό $n \geqslant 0$ και για όλες τις καταστάσεις $i_0, i_1, \dots, i_{n-1}, i, j \in E$ ισχύει η σχέση

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \tag{A.2.1}$$

τότε η στοχαστική διαδικασία θα λέγεται αλυσίδα Markov. Αν επιπλέον το δεξιό μέλος της (Α.2.1) είναι ανεξάρτητο του n, η στοχαστική διαδικασία θα λέγεται ομογενής αλυσίδα Markov.

Η ισότητα (Α.2.1) θα λέγεται ιδιότητα Markov. Ο πίνακας $P = (p_{ij})_{i,j \in E}$, όπου

$$p_{ij} = P(X_{n+1} = j | X_n = i),$$

θα λέγεται ο πίνακας μετάβασης της αλυσίδας Markov και ένας τέτοιος πίνακας θα λέγεται τυχαίος ή στοχαστικός πίνακας. Η X_0 λέγεται αρχική κατάσταση της αλυσίδας Markov και η κατανομή πιθανότητας της ν με τύπο

$$\nu(i) = P(X_0 = i)$$

θα λέγεται αρχική κατανομή της αλυσίδας Markov. Για κάθε $\kappa \geqslant 0$ και για κάθε $i_0, i_1, \ldots, i_k$ οι σχέσεις

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k) = \nu(i_0) p_{i_0 i} \ldots p_{i_{k-1} i_k}$$

λέγονται η κατανομή πιθανότητας της αλυσίδας Markov.

Θεώρημα Α.2.3. *Η κατανομή πιθανότητας μιας ομογενούς αλυσίδας Markov καθορίζεται πλήρως από την αρχική κατανομή και τον πίνακα μετάβασης.*

Α.3 Θεωρία Πληροφορίας

Ορισμός Α.3.1. Έστω $(\Omega, \mathscr{B}, P)$ χώρος πιθανότητας και $(A, \mathscr{B}_A)$ μετρήσιμος χώρος. Αν $f : \Omega \to A$ μετρήσιμη συνάρτηση, ορίζουμε μέτρο πιθανότητας P_f στον $(A, \mathscr{B}_A)$ με

$$P_f(B) = P(f^{-1}(B)),$$

για κάθε $B \in \mathscr{B}$. Το P_f θα λέγεται η κατανομή της f και το σύνολο A θα λέγεται το αλφάβητο της f.

Σκοπός μας είναι να ορίσουμε την εντροπία και τη σχετική εντροπία για όσον το δυνατόν γενικότερα μέτρα και τυχαίες μεταβλητές. Αρχίζουμε από τις πεπερασμένες τυχαίες μεταβλητές:

Ορισμός Α.3.2 (εντροπία πεπερασμένης τυχαίας μεταβλητής)**.** Έστω $(\Omega, \mathscr{B}, P)$ χώρος πιθανότητας και $(A, \mathscr{B}_A)$ πεπερασμένο αλφάβητο της τυχαίας μεταβλητής $f : \Omega \to A$. Η εντροπία της f στον χώρο $(\Omega, \mathscr{B}, P)$ ορίζεται ως η ποσότητα

$$H_P(f) = -\sum_{\alpha \in A} P(f = \alpha) \ln P(f = \alpha),$$

με την σύμβαση ότι $0 \ln 0 = 0$. Αν p_f είναι η συνάρτηση πιθανότητας της f, δηλαδή $p_f(\alpha) = P(f = \alpha)$, γράφουμε

$$H_P(f) = -\sum_{\alpha \in A} p_f(\alpha) \ln p_f(\alpha).$$

Επίσης, αν $A = \{\alpha_1, \alpha_2, \ldots, \alpha_{|A|}\}$ και αν ορίσουμε $Q = \{Q_i : i = 1, 2, \ldots, |A|\}$ όπου $Q_i = \{\omega \in \Omega : f(\omega) = \alpha_i\} = f^{-1}(\{\alpha_i\})$, τότε μπορούμε να θεωρήσουμε την εντροπία σαν συνάρτηση της διαμέρισης Q:

$$H_P(Q) = -\sum_{i=1}^{|A|} P(Q_i) \ln P(Q_i).$$

Τέλος, ορίζοντας την τυχαία μεταβλητή $P(f) : \Omega \to \mathbb{R}$ με $P(f)(\omega) = P(\lambda \in \omega : f(\lambda) = f(\omega))$ μπορούμε να γράψουμε

$$H_P(f) = \mathbb{E}_P(-\ln P(f)).$$

Ορισμός Α.3.3 (σχετική εντροπία μέτρων πιθανότητας σε πεπερασμένο χώρο)**.** Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Έστω $f : \Omega \to A$ τυχαία μεταβλητή με αλφάβητο $(A, \mathscr{B}_A)$, το οποίο είναι πεπερασμένο, και Q η διαμέριση του A που επάγεται από την f. Επίσης συμβολίζουμε με P_f, M_f τις κατανομές της f, που επάγουν τα μέτρα P, M αντίστοιχα. Η σχετική εντροπία της f με μέτρο P σε σχέση με το μέτρο M είναι η ποσότητα:

$$H_{P\|M}(f) \equiv H_{P\|M}(Q) := \sum_{\alpha \in A} p_f(\alpha) \ln \frac{p_f(\alpha)}{m_f(\alpha)} = \sum_{i=1}^{|A|} P(Q_i) \ln \frac{P(Q_i)}{M(Q_i)}.$$

Παρατηρούμε ότι ο ορισμός έχει νόημα όταν ισχύει ότι $p_f(\alpha) = 0$ αν και μόνο αν $m_f(\alpha) = 0$, δηλαδή όταν το P_f είναι απολύτως συνεχές ως προς το M_f. Ορίζουμε λοιπόν $H_{P\|M}(f) = \infty$ αν το P_f δεν είναι απολύτως συνεχές ως προς το M_f.

Θα ορίσουμε, τώρα, την σχετική εντροπία δύο μέτρων πιθανότητας, τα οποία δεν προέρχονται από κάποια τυχαία μεταβλητή.

Ορισμός Α.3.4 (σχετική εντροπία μέτρων πιθανότητας πεπερασμένου χώρου)**.** Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος, Ω πεπερασμένο, και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Η σχετική εντροπία (ή η απόκλιση) του P σε σχέση με το M ορίζεται ως:

$$D(P\|M) = \sum_{\omega \in \Omega} P(\omega) \ln \frac{P(\omega)}{M(\omega)},$$

δηλαδή είναι η σχετική εντροπία της ταυτοτικής απεικόνισης του Ω.

Θεώρημα Α.3.5 (ανισότητα απόκλισης)**.** *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος, Ω πεπερασμένο, και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Τότε*

$$D(P\|M) \geqslant 0,$$

με την ισότητα να ισχύει αν και μόνο αν $P = M$.

Θεώρημα Α.3.6. *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος, Ω πεπερασμένο, και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Τότε ισχύει ότι*

$$D(P\|M) = \sup\{\mathbb{E}_P(\varphi) - \ln \mathbb{E}_M(e^{\varphi})\},$$

με το supremum να παίρνεται πάνω από όλες τις τυχαίες μεταβλητές $\varphi : \Omega \to \mathbb{R}$.

Συνεχίζουμε με τον ορισμό της δεσμευμένης εντροπίας δύο τυχαίων μεταβλητών:

Ορισμός Α.3.7 (δεσμευμένη εντροπία πεπερασμένων τυχαίων μεταβλητών)**.** Έστω $(\Omega, \mathscr{B}, P)$ χώρος πιθανότητας και $X, Y : \Omega \to A$ τυχαίες μεταβλητές με πεπερασμένα αλφάβητα $(A_Q, \mathscr{B}_{A_X})$ και $(A_U, \mathscr{B}_{A_U})$ αντίστοιχα. Η δεσμευμένη εντροπία της X δεδομένης της Y είναι η ποσότητα

$$H(X|Y) = H(X,Y) - H(Y),$$

όπου

$$H(X,Y) = -\sum_{x,y} p_{X,Y}(x,y) \ln p_{X,Y}(x,y),$$

είναι η από κοινού εντροπία των X, Y και $p_{X,Y}$ η από κοινού συνάρτηση πιθανότητας των X, Y. Ισχύει ότι

$$H(X|Y) = -\sum_{x,y} P(X = x, Y = y) \ln P(X = x, Y = y) = -\sum_{x,y} p_{X,Y}(x, y) \ln p_{X|Y}(x|y),$$

όπου $p_{X|Y}(x|y) = p_{X,Y}(x, y)/p_Y(y)$ είναι η δεσμευμένη συνάρτηση πιθανότητας της X δεδομένης της Y. Ομοίως ορίζουμε την δεσμευμένη σχετική εντροπία

$$H_{P\|M}(X|Y) = H_{P\|M}(X, Y) - H_{P\|M}(Y).$$

Θεώρημα Α.3.8 (κανόνας της αλυσίδας)**.** *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος, P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$ και $X_1, X_2, \ldots X_n \Omega \to A$ τυχαίες μεταβλητές με πεπερασμένα αλφάβητα. Τότε ισχύουν τα εξής*

$$\begin{aligned} H(X_1, X_2, \ldots, X_n) &= \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_{i-1}) \\ H_{P\|M}(X_1, X_2, \ldots, X_n) &= \sum_{i=1}^{n} H_{P\|M}(X_i|X_1, \ldots, X_{i-1}) \end{aligned}$$

Προχωράμε στον γενικό ορισμό της σχετικής εντροπίας:

Ορισμός Α.3.9 (σχετική εντροπία μέτρων πιθανότητας)**.** Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος, P, M μέτρα πιθανότητας στον $(\Omega\mathscr{B})$. Η σχετική εντροπία (ή απόκλιση) του P σε σχέση με το M ορίζεται από την σχέση

$$D(P\|M) := \sup_{Q} H_{P\|M}(Q) = \sup_{f} D(P_f\|M_f),$$

με το πρώτο supremum να παίρνεται πάνω από όλες τις πεπερασμένες μετρήσιμες διαμερίσεις του Ω και το δεύτερο να είναι πάνω από όλες τις τυχαίες μεταβλητές $f : \Omega \to A$ με πεπερασμένα αλφάβητα $(A, \mathscr{B}_A)$.

Λήμμα Α.3.10 (ανισότητα απόκλισης)**.** *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Τότε*

$$D(P\|M) \geqslant 0,$$

με την ισότητα να ισχύει αν και μόνο αν $P = M$.

Λήμμα Α.3.11. *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$. Τότε έχουμε ότι*

$$D(P\|M) = \infty,$$

αν το P δεν είναι απολύτως συνεχές ως προς το M και όταν το P είναι απολύτως συνεχές ως προς το M, οπότε και υπάρχει η Radon-Nikodym παράγωγος $f = dP/dM$, τότε

$$D(P\|M) = \int_{\omega\in\Omega} \ln f(\omega) dP(\omega) = \int_{\omega\in\Omega} f(\omega) \ln f(\omega) dM(\omega).$$

Θεώρημα A.3.12. *Έστω $(\Omega, \mathscr{B})$ μετρήσιμος χώρος και P, M μέτρα πιθανότητας στον $(\Omega, \mathscr{B})$, P απολύτως συνεχές ως προς το M. Τότε*

$$D(P\|M) = \sup\{\mathbb{E}_P(\varphi) - \ln \mathbb{E}_M(e^{\varphi})\},$$

με το supremum να παίρνεται πάνω από όλες τις τυχαίες μεταβλητές $\varphi : \Omega \to \mathbb{R}$ για τις οποίες η e^{φ} είναι M-ολοκληρώσιμη και η μέση τιμή $\mathbb{E}_P(\varphi)$ είναι καλά ορισμένη.

Ορισμός A.3.13 (δεσμευμένη σχετική εντροπία)**.** Έστω $(\Omega, \mathscr{B}, P)$ χώρος πιθανότητας και $X, Y : \Omega \to A$ τυχαίες μεταβλητές με αλφάβητα $(A_Q, \mathscr{B}_{A_X}), (A_U, \mathscr{B}_{A_U})$ αντίστοιχα. Έστω P_{XY}, m_{XY} δύο κατανομές στο $(A_X \times A_Y, \mathscr{B}_{A_X \times A_Y})$ και έστω ότι το P_{XY} είναι απολύτως συνεχές ως προς το M_{XY} . Γράφουμε M_Y, P_Y τις περιθώριες κατανομές, για παράδειγμα $M_Y(F) = M_{XY}(A_X \times F)$. Τότε το P_Y είναι απολύτως συνεχές ως προς το M_Y και ορίζουμε τις πυκνότητες (Radon-Nikodym παραγώγους)

$$f_{XY} = \frac{dP_{XY}}{dM_{XY}}, f_Y = \frac{dP_Y}{dM_Y},$$

ώστε να ισχύουν οι

$$P_{XY}(F) = \int_F f_{XY} dM_{XY},\ F \in \mathscr{B}_{A_X \times A_Y}$$
$$P_Y(F) = \int_F f_Y dM_Y,\ F \in \mathscr{B}_{A_Y}.$$

Επίσης ορίζουμε την δεσμευμένη πυκνότητα

$$f_{X|Y}(x|y) = \begin{cases} \frac{f_{XY}(x,y)}{f_Y(y)} & \text{αν } f_Y(y) > 0 \\ 1 & \text{αν } f_Y(y) = 0 \end{cases}.$$

Υποθέτουμε ότι η $h_Y = \ln f_Y$ υπάρχει και ορίζουμε $h_{X|Y} = \ln f_{X|Y}$. Η δεσμευμένη σχετική εντροπία, της X δεδομένης της Y, του μέτρου P σε σχέση με το μέτρο M ορίζεται ως:

$$\begin{aligned} H_{P\|M}(X|Y) := \mathbb{E}_{P_{XY}}(h_{X|Y}) &= \int \ln f_{X|Y}(x|y) dP_{XY}(x,y) \\ &= \int f_{XY}(x,y) \ln f_{X|Y}(x|y) dM_{XY}(x,y), \end{aligned}$$

αν το ολοκλήρωμα υπάρχει.

Λήμμα A.3.14 (κανόνας της αλυσίδας)**.** *Έστω $(\Omega, \mathscr{B}, P)$ χώρος πιθανότητας και $X, Y : \Omega \to A$ τυχαίες μεταβλητές με αλφάβητα $(A_Q, \mathscr{B}_{A_X}), (A_U, \mathscr{B}_{A_U})$ αντίστοιχα. Τότε ισχύει η σχέση*

$$D(P_{XY}\|M_{XY}) = H_{P\|M}(X|Y) + D(P_Y\|M_Y),$$

ή, ισοδύναμα,

$$H_{P\|M}(X, Y) = H_{P\|M}(Y) + H_{P\|M}(X|Y).$$

Αν επιπλέον $H_{P\|M}(Y) < \infty$, τότε

$$H_{P\|M}(X|Y) = H_{P\|M}(X, Y) - H_{P\|M}(Y).$$

A.4 Απόσταση Wasserstein

Το πρόβλημα από το οποίο ξεκινάει η θεωρία της βέλτιστης μεταφοράς είναι το εξής: Θέλουμε να μεταφέρουμε μια ποσότητα υλικού από μια αποθήκη X σε μια δεύτερη Y. Θεωρούμε ότι ο όγκος των δύο αποθηκών είναι ο ίδιος και ότι η μάζα του υλικού είναι 1. Υποθέτουμε επίσης ότι ο χώρος που καταλαμβάνει το υλικό στην αποθήκη X μετριέται μέσω ενός μέτρου πιθανότητας μ, ενώ ο αντίστοιχος χώρος στην αποθήκη Y μέσω ενός μέτρου ν. Τέλος θεωρούμε ότι το κόστος για να μεταφερθεί μια μονάδα μάζας από την θέση $x \in X$ στην θέση $y \in Y$ είναι $c(x, y)$. Επομένως έχουμε δύο χώρους πιθανότητας $(X, \mu), (Y, \nu)$ και μια συνάρτηση κόστους $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$. Το πρόβλημα μεταφοράς είναι να ελαχιστοποιηθεί το κόστος μεταφοράς.

Ορισμός A.4.1 (σχέδιο μεταφοράς ή ταίριασμα δύο μέτρων)**.** Έστω $(X, \mu), (Y, \nu)$ χώροι μέτρου. Ένα σχέδιο μεταφοράς ή ταίριασμα των μ, ν είναι ένα μέτρο π στον χώρο $X \times Y$ με

$$\pi(A \times Y) = \mu(A), \ \pi(X \times B) = \nu(B),$$

για κάθε ζεύγος μετρήσιμων συνόλων $A \subset X$ και $B \subset Y$. Συμβολίζουμε με $\Pi(\mu, \nu)$ τον χώρο των ταιριασμάτων των μ, ν.

Ορισμός A.4.2 (πρόβλημα βέλτιστης μεταφοράς του Kantorovich)**.** Έστω $(X, \mu), (Y, \nu)$ χώροι μέτρου, $c : X \times Y \to \mathbb{R}_+ \cup \{+\infty\}$ μια συνάρτηση κόστους. Να ελαχιστοποιηθεί η ποσότητα

$$I[\pi] = \int_{X \times Y} c(x, y) d\pi(x, y), \ \text{για } \pi \in \Pi(\mu, \nu).$$

Θεώρημα A.4.3 (θεώρημα δυϊσμού του Kantorovich)**.** *Έστω X, Y Πολωνικοί χώροι, μ, ν μέτρα πιθανότητας στους X, Y αντίστοιχα και $c : X \times Y \to \mathbb{R}_+ \cup \{+\infty\}$ μια κάτω ημι-συνεχής συνάρτηση κόστους. Για κάθε $\pi \in P(X \times Y)$ και $(\varphi, \psi) \in L^1(X, \mu) \times L^1(Y, \nu)$ ορίζουμε*

$$I[\pi] = \int_{X \times Y} c(x, y) d\pi(x, y), \ J(\varphi, \psi) = \int_X \varphi d\mu + \int_Y \psi d\nu.$$

Έστω επίσης Φ_c το σύνολο όλων των ζευγαριών $(\varphi, \psi) \in L^1(X, \mu) \times L^1(Y, \nu)$ με

$$\varphi(x) + \psi(y) \leqslant c(x, y),$$

για μ-σχεδόν όλα τα $x \in X$ και ν-σχεδόν όλα τα $y \in Y$. Τότε ισχύει ότι

$$\inf_{\Pi(\mu,\nu)} I[\pi] = \sup_{\Phi_c} J(\varphi, \psi).$$

Μάλιστα το infimum στο αριστερό μέλος της ισότητας επιτυγχάνεται. Επιπλέον, η τιμή του supremum στο δεξιό μέλος δεν αλλάζει αν περιοριστούμε σε συναρτήσεις φ, ψ οι οποίες είναι συνεχείς και φραγμένες.

Θεώρημα A.4.4 (θεώρημα δυϊσμού Kantorovich-Rubinstein)**.** *Έστω X Πολωνικός χώρος και d μια κάτω ημισυνεχής μετρική στον X. Για μ και ν μέτρα στον X θέτουμε*

$$T_d(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y) d\pi(x, y),$$

και $\mathrm{Lip}(X)$ *τον χώρο των Lipschitz συνεχών συναρτήσεων του* X, *εφοδιασμένο με τη νόρμα*

$$\|\varphi\|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)}.$$

Τότε

$$T_d(\mu,\nu) = \sup\left\{\int_X \varphi d(\mu-\nu) : \ \varphi \in \mathrm{Lip}(X) \cap L^1(X, |\mu-\nu|), \ \|\varphi\|_{\mathrm{Lip}} \leqslant 1\right\}.$$

Επιπλέον, η τιμή του supremum δεν αλλάζει αν περιοριστούμε σε φ *οι οποίες είναι φραγμένες.*

Ορισμός A.4.5 (απόσταση Wasserstein)**.** Έστω (X,d) Πολωνικός μετρικός χώρος. Για $p > 0$ ορίζουμε την L^p απόσταση Wasserstein των μ, ν

$$W_p(\mu,\nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d^p(x,y) d\pi(x,y)\right)^{1/p}.$$

Η παραπάνω συνάρτηση είναι μετρική στον χώρο $\mathscr{P}(X)$, ο οποίος είναι ο χώρος των Borel μέτρων πιθανότητας μ, για τα οποία επιπλέον ισχύει η σχέση $\int_X d(x,o)^p d\mu(x) < \infty$ για κάποιο (και συνεπώς για όλα τα) $o \in X$.

Θεώρημα A.4.6. *Έστω* (X,d) *πλήρης και διαχωρίσιμος μετρικός χώρος. Τότε, ο* $(\mathscr{P}(X), W_p)$ *είναι πλήρης και διαχωρίσιμος.*

Θεώρημα A.4.7. *Έστω* (M,g) *συμπαγής πολλαπλότητα Riemann και*

$$\mathscr{P}_2(M) = (\mathscr{P}(M), W_2).$$

Ο $\mathscr{P}_2(M)$ *είναι γεωδαισιακός χώρος.*

A.5 Απόσταση ολικής κύμανσης

Ορισμός A.5.1 (απόσταση ολικής κύμανσης)**.** Έστω (X,d) Πολωνικός μετρικός χώρος και μ, ν μέτρα Borel στον X. Η απόσταση ολικής κύμανσης των μ και ν ορίζεται ως:

$$\mathrm{TV}(\mu,\nu) = \sup_{A \in \mathscr{B}(X)} \{\mu(A) - \nu(A)\}.$$

Ορισμός A.5.2 (σχετική εντροπία)**.** Έστω μ, ν μέτρα στον X. Η σχετική εντροπία ή απόκλιση Kullback των μ και ν ορίζεται ως εξής:

$$D(\nu \parallel \mu) = \mathrm{Ent}_\nu\left(\frac{d\nu}{d\mu}\right) = \int_X \ln\left(\frac{d\nu}{d\mu}\right) d\nu$$

αν το ν είναι απολύτως συνεχές ως προς το μ και $D(\nu\|\mu) = +\infty$ αλλιώς. Αν f, g είναι τυχαίες μεταβλητές με κατανομές μ και ν αντίστοιχα, με $D(f\|g)$ θα εννοούμε την σχετική εντροπία $D(\nu\|\mu)$.

Θεώρημα A.5.3 (ανισότητα Pinsker)**.** *Έστω* (X,d) *Πολωνικός μετρικός χώρος και* μ, ν *δύο μέτρα Borel στον* X. *Τότε ισχύει ότι*

$$\mathrm{TV}(\mu,\nu) \leqslant \sqrt{\frac{1}{2} D(\mu\|\nu)}.$$

A.6 Θεώρημα Prokhorov

Ορισμός A.6.1 (σφιχτότητα)**.** Έστω $(X, \mathscr{T})$ τοπολογικός χώρος. Μια οικογένεια Π Borel μέτρων πιθανότητας θα λέγεται σφιχτή, αν για κάθε $\varepsilon > 0$, υπάρχει συμπαγές $K \subset X$ ώστε

$$\mu(K) > 1 - \varepsilon,$$

για κάθε $\mu \in \Pi$.

Ορισμός A.6.2 (ασθενής σύγκλιση)**.** Έστω $(X, \mathscr{T})$ τοπολογικός χώρος, $\{\mu_n\}_{n\in\mathbb{N}}$ ακολουθία Borel μέτρων πιθανότητας και μ Borel μέτρο πιθανότητας. Η $\{\mu_n\}$ συγκλίνει ασθενώς στο μ, αν για κάθε $f : X \to \mathbb{R}$ συνεχή και φραγμένη ισχύει ότι

$$\int_X f d\mu_n \to \int_X f d\mu.$$

Θεώρημα A.6.3 (Prokhorov)**.** *Έστω (X, d) διαχωρίσιμος μετρικός χώρος, Π μια οικόγενεια Borel μέτρων πιθανότητας και μ Borel μέτρο πιθανότητας. Τότε:*

(i) *Η Π είναι σφιχτή αν και μόνο αν η κλειστότητα της Π είναι ακολουθιακά συμπαγής στον $\mathscr{P}(X)$ εφοδιασμένο με την τοπολογία της ασθενούς σύγκλισης.*

(ii) *Ο $\mathscr{P}(X)$ εφοδιασμένος με την τοπολογία της ασθενούς σύγκλισης είναι μετρικοποιήσιμος.*

(iii) *Αν ο (X, d) είναι πλήρης, τότε υπάρχει πλήρης μετρική ρ στον $\mathscr{P}(X)$ ισοδύναμη με την τοπολογία της ασθενούς σύγκλισης. Επιπλέον, η Π είναι σφιχτή αν και μόνο αν η κλειστότητά της είναι συμπαγής στον χώρο $(\mathscr{P}(X), \rho)$.*

A.7 Θεώρημα σταθερού σημείου Markov-Kakutani

Το θεώρημα σταθερού σημείου Markov-Kakutani εξασφαλίζει *κοινό* σταθερό σημείο για μια οικογένεια αφφινικών απεικονίσεων που ορίζονται σε ένα συμπαγές κυρτό σύνολο και αντιμετατίθενται. Μια απεικόνιση $T : X \to X$ λέγεται αφφινική αν υπάρχει $x_0 \in X$ ώστε η $S : X \to X$ με $S(x) = T(x) - x_0$ να είναι γραμμικός τελεστής.

Θεώρημα A.7.1 (Markov–Kakutani)**.** *Έστω X ένας τοπικά κυρτός χώρος, έστω K ένα μη κενό, κυρτό και συμπαγές υποσύνολο του X και έστω $\{T_i : i \in I\}$ μια οικογένεια συνεχών αφφινικών απεικονίσεων $T_i : K \to K$ οι οποίες αντιμετατίθενται: $T_i \circ T_j = T_j \circ T_i$ για κάθε $i, j \in I$. Τότε, υπάρχει $x_0 \in K$ ώστε $T_i(x_0) = x_0$ για κάθε $i \in I$.*

Απόδειξη. Για κάθε $i \in I$ και $n \in \mathbb{N}$ ορίζουμε $T_i^{(n)} : K \to K$ με

$$T_i^{(n)} = \frac{I + T_i + T_i^2 + \cdots + T_i^{n-1}}{n},$$

όπου $T_i^k = T_i \circ \cdots \circ T_i$ (k φορές). Από την υπόθεση ότι οι T_i αντιμετατίθενται, έπεται ότι

$$T_i^{(n)} \circ T_j^{(m)} = T_j^{(m)} \circ T_i^{(n)} \tag{A.7.1}$$

για κάθε $i, j \in I$ και $n, m \in \mathbb{N}$. Θεωρούμε την οικογένεια

$$\mathcal{K} := \{T_i^{(n)}(K) : i \in I, n \in \mathbb{N}\}.$$

Κάθε σύνολο στην $\mathcal{K}$ είναι συμπαγές και κυρτό. Χρησιμοποιώντας την (A.7.1) βλέπουμε ότι, για κάθε $i_1, \ldots, i_k \in I$ και $n_1, \ldots, n_k \in \mathbb{N}$,

$$T_{i_1}^{(n_1)} \circ \cdots \circ T_{i_k}^{(n_k)}(K) \subseteq \bigcap_{j=1}^{k} T_{i_j}^{(n_j)}(K).$$

Δηλαδή, η $\mathcal{K}$ έχει την ιδιότητα πεπερασμένης τομής. Από τη συμπάγεια του K έπεται ότι υπάρχει $x_0 \in K$ ώστε $x_0 \in T_i^{(n)}(K)$ για κάθε $i \in I$ και κάθε $n \in \mathbb{N}$. Αυτό σημαίνει ότι, για κάθε $i \in I$ και $n \in \mathbb{N}$ υπάρχει $x = x(i, n) \in K$ ώστε

$$x_0 = \frac{x + T_i(x) + \cdots + T_i^{n-1}(x)}{n},$$

δηλαδή

$$\begin{aligned} T_i(x_0) - x_0 &= \frac{T(x) + T_i^2(x) + \cdots + T_i^n(x)}{n} - \frac{x + T_i(x) + \cdots + T_i^{n-1}(x)}{n} \\ &= \frac{1}{n}(T_i^n(x) - x) \in \frac{1}{n}(K - K). \end{aligned}$$

Αφού το K είναι συμπαγές, το $K - K$ είναι επίσης συμπαγές και $0 \in K - K$. Έστω U ανοικτή περιοχή του 0. Υπάρχει $n \in \mathbb{N}$ ώστε $K - K \subseteq nU$. Άρα, $T_i(x_0) - x_0 \in \frac{1}{n}(K - K) \subseteq U$. Έπεται ότι $T_i(x_0) = x_0$ για κάθε $i \in I$ (το $i \in I$ ήταν τυχόν). $\square$

Βιβλιογραφία

[BE85] Bakry, Dominique, and Émery, Michel. “Diffusions hypercontractives.” Séminaire de probabilités, XIX, 1983-84, Springer, Berlin (Lecture Notes in Math.) Tome 1123 (1985), pp. 177-206.

[BGL13] Bakry, Dominique, Ivan Gentil, and Michel Ledoux. “Analysis and geometry of Markov diffusion operators.” Vol. 348. Springer Science and Business Media, 2013.

[Ben21] Benty, Ryan. “Concentration of Measure and Ricci Curvature.” Diss. Concordia University, 2012.

[Ber12] Berger, Marcel. “A panoramic view of Riemannian geometry.” Springer Science and Business Media, 2012.

[BG99] Bobkov, Sergey G., and Friedrich Götze. “Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.” Journal of Functional Analysis 163.1 (1999): 1-28.

[BHT06] Bobkov, Sergey G., Christian Houdré, and Prasad Tetali. “The subgaussian constant and concentration inequalities.” Israel J. Math. 156 (2006): 255-283.

[ERL17] Eldan, Ronen, James R. Lee, and Joseph Lehec. “Transport-entropy inequalities and curvature in discrete-space Markov chains.” A Journey Through Discrete Mathematics. Springer, Cham, 2017. 391-406.

[FY18] Fathi, Max, and Yan Shu. “Curvature and transport inequalities for Markov chains in discrete spaces.” Bernoulli 24.1 (2018): 672-698.

[Gr11] Gray, Robert M. “Entropy and information theory.” Springer Science and Business Media, 2011.

[Grom07] Gromov, Mikhail. “Metric structures for Riemannian and non-Riemannian spaces.” Springer Science and Business Media, 2007.

[Led01] Ledoux, Michel. “The concentration of measure phenomenon”. No. 89. American Mathematical Soc., 2001.

[Lee06] Lee, John M. “Riemannian manifolds: an introduction to curvature.” Vol. 176. Springer Science and Business Media, 2006.

[LLY11] Lin, Yong, Linyuan Lu, and Shing-Tung Yau. “Ricci curvature of graphs.” Tohoku Mathematical Journal, Second Series 63.4 (2011): 605-627.

[Mar96] Marton, Katalin. “Bounding $\bar{d}$-distance by informational divergence: A method to prove measure concentration.” The Annals of Probability 24.2 (1996): 857-866.

[MS09] Milman, Vitali D., and Gideon Schechtman. “Asymptotic theory of finite dimensional normed spaces: Isoperimetric inequalities in Riemannian manifolds.” Vol. 1200. Springer, 2009.

[NR17] Najman, Laurent, and Pascal Romon (Eds.). “Modern approaches to discrete curvature.” Vol. 2184, Springer, 2017.

[Oll04] Ollivier, Yann. “Sharp phase transition theorems for hyperbolicity of random groups.” Geom. Funct. Anal. 14.3 (2004): 595-679.

[Oll07] Ollivier, Yann. “Ricci curvature of metric spaces.” Comptes Rendus Mathematique 345.11 (2007): 643-646.

[Oll09] Ollivier, Yann. "Ricci curvature of Markov chains on metric spaces." Journal of Functional Analysis 256.3 (2009): 810-864.

[Oll10] Ollivier, Yann. "A survey of Ricci curvature for metric spaces and Markov chains." Probabilistic approach to geometry 57 (2010): 343-381.

[Oll11] Ollivier, Yann. "A visual introduction to Riemannian curvatures and some discrete generalizations." Analysis and Geometry of Metric Measure Spaces: Lecture Notes of the 50th Séminaire de Mathématiques Supérieures (SMS), Montr?al 56 (2011): 197-219.

[OV12] Ollivier, Yann, and Cédric Villani. "A Curved Brunn–Minkowski Inequality on the Discrete Hypercube, Or: What Is the Ricci Curvature of the Discrete Hypercube?." Siam Journal on Discrete Mathematics 26.3 (2012): 983-996.

[RS05] von Renesse, Max-K., and Karl-Theodor Sturm. "Transport inequalities, gradient estimates, entropy and Ricci curvature." Communications on pure and applied mathematics 58.7 (2005): 923-940.

[Vill03] Villani, Cédric. "Topics in Optimal Transportation." Grad. Stud. Math., vol. 58, Amer. Math. Soc., Providence, RI, 2003.

[Vill08] Villani, Cédric. "Optimal transport: old and new." Vol. 338. Springer Science and Business Media, 2008.