ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΜΑΘΗΜΑΤΙΚΑ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΤΟΥΣ»

Βέλτιστη Μεταφορά Μάζας

Μεταπτυχιακή Εργασία

Λιανάκης Ιωάννης

Επιβλέπουσα Καθηγήτρια: Σουζάνα Παπαδοπούλου

Ηράκλειο, Φεβρουάριος 2017

Η παρούσα μεταπτυχιακή εργασία εκπονήθηκε στα πλαίσια του διατμηματικού προγράμματος μεταπτυχιακών σπουδών «Μαθηματικά και Εφαρμογές τους» των τμημάτων Μαθηματικών και Εφαρμοσμένων Μαθηματικών του Πανεπιστημίου Κρήτης για την απονομή Μεταπτυχιακού Διπλώματος Ειδίκευσης στην κατεύθυνση «Επιχειρησιακά Μαθηματικά».

Την εξεταστική επιτροπή αποτέλεσαν οι ακόλουθοι καθηγητές του τμήματος Μαθηματικών και Εφαρμοσμένων Μαθηματικών:

Παπαδοπούλου Σουζάνα

Αθανασόπουλος Ιωάννης

Φίλιππας Στάθης

ΠΕΡΙΕΧΟΜΕΝΑ

ΣΥΜΒΟΛΙΣΜΟΙ

Θα δουλέψουμε σε τοπολογικούς χώρους εφοδιασμένους με την Borel $\sigma-$άλγεβρα. Δεδομένου ενός χώρου X, με $\mathcal{B}_X$ συμβολίζουμε την Borel $\sigma-$άλγεβρα στον X, με $P(X)$ συμβολίζουμε το σύνολο των Borel μέτρων πιθανότητας στον X και με $M(X)$ το σύνολο των πεπερασμένων προσημασμένων μέτρων στην $\mathcal{B}_X$. Όταν μιλάμε για μέτρα ορισμένα σε έναν τοπολογικό χώρο, εννοούμε μέτρα Borel. Ο διανυσματικός χώρος $M(X)$ εφοδιάζεται με τη νόρμα της ολικής κύμανσης $\|\mu\|_{TV} = \inf\{\mu_+(X) + \mu_-(X)\}$, όπου το infimum λαμβάνεται πάνω σε όλα τα μη αρνητικά μέτρα μ_+, μ_- τέτοια ώστε $\mu = \mu_+ - \mu_-$. Το infimum πετυχαίνεται όταν τα μέτρα μ_+ και μ_- είναι ιδιόμορφα το ένα ως προς το άλλο, οπότε η σχέση $\mu = \mu_+ - \mu_-$ λέγεται ανάλυση Hahn του μ.

Για ένα δεδομένο μέτρο μ στον X και για $p \in [1, +\infty)$, συμβολίζουμε με $L^p(X)$ ή $L^p(d\mu)$ τον χώρο Lebesgue τάξεως p αναφορικά με το μέτρο μ, δηλαδή:

$$L^p(d\mu) = \left\{ f: X \longrightarrow \mathbb{R} \ : \ \int_X |f(x)|^p d\mu < +\infty \right\}.$$

Ο χώρος $L^p(d\mu)$ αποτελείται από κλάσεις ισοδυναμίας, όπου δύο συναρτήσεις ανήκουν στην ίδια κλάση ισοδυναμίας αν ταυτίζονται σχεδόν παντού.

Αν T είναι μια μετρήσιμη απεικόνιση από έναν χώρο μέτρου (X, μ) σε έναν χώρο $(Y, \mathcal{A})$, όπου $\mathcal{A}$ μια $\sigma-$άλγεβρα υποσυνόλων του Y, συμβολίζουμε με $T \# \mu$ το μέτρο εικόνα του μ μέσω της T. Ειδικότερα, $(T \# \mu)(B) = \mu\big(T^{-1}(B)\big)$ όπου $T^{-1}(B) = \{x \in X : \ T(x) \in B\}, \ B \in \mathcal{A}$.

Έστω X ένας τοπολογικός χώρος. Συμβολίζουμε με $C(X)$ τον χώρο των συνεχών συναρτήσεων $f: X \to \mathbb{R}$, με $C_b(X)$ τον χώρο των συνεχών και φραγμένων συναρτήσεων $f: X \to \mathbb{R}$ και με $C_0(X)$ τον χώρο των συνεχών συναρτήσεων $f: X \to \mathbb{R}$ που στο άπειρο τείνουν στο 0, δηλαδή $f \in C_0(X)$ αν και μόνο αν για κάθε $\varepsilon > 0$, υπάρχει συμπαγές $K_\varepsilon \subseteq X$ τέτοιο ώστε $|f(x)| < \varepsilon$, $x \in X \setminus K_\varepsilon$. Ακόμα, συμβολίζουμε με $C_c(X)$ το χώρο των συνεχών συναρτήσεων $f: X \to \mathbb{R}$ με συμπαγή φορέα, όπου εξ ορισμού ο φορέας μιας συνάρτησης f είναι το σύνολο $Supp(f) = \overline{\{x \in X : f(x) \neq 0\}}$.

Ο χώρος $C_b(X)$ εφοδιάζεται με τη supremum νόρμα, δηλαδή $\|f\|_\infty = \sup\limits_{x \in X} |f(x)|$.

Αν $A \subseteq X$, συμβολίζουμε με $Int(A)$ το εσωτερικό του A δηλ. το μεγαλύτερο ανοικτό σύνολο που περιέχεται στο A, με $\bar{A}$ την κλειστή θήκη του A, δηλ. το μικρότερο κλειστό σύνολο που περιέχει το A και με ∂A το σύνορο του A, δηλαδή $\partial A = \bar{A} \setminus Int(A)$. Εξ ορισμού, ο φορέας ενός μέτρου μ στον X είναι το μικρότερο κλειστό σύνολο $F \subseteq X$ με $\mu(X \setminus F) = 0$, και συμβολίζεται με $Supp(\mu)$. Όταν, όμως, λέμε ότι το μ συγκεντρώνεται στο $A \subseteq X$, αυτό σημαίνει ότι $\mu(X \setminus A) = 0$ χωρίς αναγκαστικά το A να είναι κλειστό.

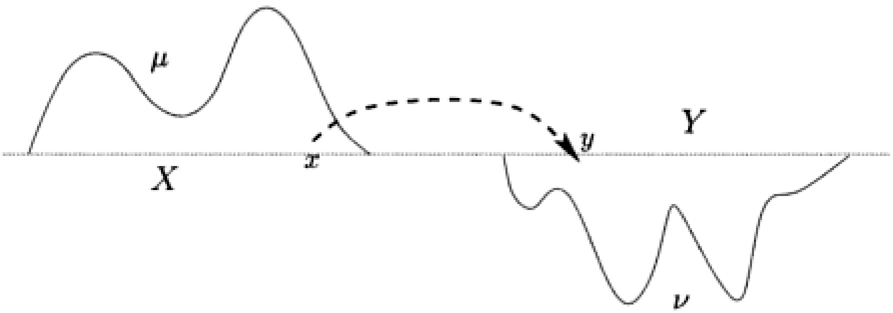
Αν ο (X, ρ) είναι ένας μετρικός χώρος, τότε θα είναι εφοδιασμένος με την τοπολογία που επάγεται από την μετρική του και θα συμβολίζουμε με $B(x, r)$ ή $B_r(x)$ την ανοικτή μπάλα με κέντρο x και ακτίνα r. Αν $F \subseteq X$, η απόσταση του σημείου $x \in X$ από το σύνολο F είναι $\rho(x, F) = \inf_{y \in F}\{\rho(x, y)\}$. Συμβολίζουμε, ακόμα, με $Lip(X)$ το σύνολο των Lipschitz συναρτήσεων $f: X \to \mathbb{R}$.

Αν X είναι ένας χώρος Banach και X^* ο τοπολογικός δυϊκός αυτού, θα συμβολίζουμε με $\langle \cdot \, , \cdot \rangle$ την αγκύλη δυϊκότητας ανάμεσα στους X και X^*. Αν φ είναι μια κυρτή συνάρτηση σε έναν χώρο Banach X, τότε θα συμβολίζουμε με φ^* την δυϊκή κυρτή συνάρτηση της φ, υπό την έννοια της δυϊκότητας Legendre – Fenchel. Θα συμβολίζουμε, ακόμα, με $\partial\varphi$ το υποδιαφορικό της φ και θα το ταυτίζουμε με το γράφημά του $Graph(\partial\varphi)$, το οποίο είναι ένα υποσύνολο του $X \times X^*$.

ΕΙΣΑΓΩΓΗ

Θεμελίωση του προβλήματος βέλτιστης μεταφοράς μάζας

Υποθέτουμε ότι έχουμε μια σωρό άμμου την οποία πρέπει να μεταφέρουμε, με το ελάχιστο δυνατό κόστος, σε άλλη θέση με άλλη κατανομή μάζας. Η μάζα, την οποία θεωρούμε ίση με 1, διατηρείται κατά τη μεταφορά.



Μπορούμε να μοντελοποιήσουμε την παλιά και τη νέα κατανομή μάζας με δύο μέτρα πιθανότητας μ και ν, τα οποία ορίζονται, σε δύο χώρους μέτρου X και Y, αντίστοιχα. Αν A, B είναι δύο μετρήσιμα υποσύνολα των X, Y αντίστοιχα, τότε το $\mu(A)$ εκφράζει τη μάζα της άμμου που βρίσκεται μέσα στο A ενώ το $\nu(B)$ εκφράζει τη μάζα της άμμου που θα τοποθετηθεί μέσα στο B.

Η μεταφορά της άμμου απαιτεί κάποιο κόστος, το οποίο μπορεί να μοντελοποιηθεί μέσω μιας συνάρτησης κόστους ορισμένη στον $X \times Y$. Άτυπα μιλώντας, το $c(x, y)$ εκφράζει το κόστος μεταφοράς μιας μονάδας μάζας από την τοποθεσία x στην τοποθεσία y. Υποθέτουμε ότι η συνάρτηση κόστους είναι μετρήσιμη και μη αρνητική, ενώ δεν μπορούμε να αποκλείσουμε και την περίπτωση να παίρνει την τιμή $+\infty$.

Στην εργασία αυτή θα ασχοληθούμε με το βασικό ερώτημα:

Πώς να πραγματοποιήσουμε τη μεταφορά με το ελάχιστο δυνατό κόστος;

Ένας οποιοσδήποτε τρόπος μεταφοράς λέγεται πλάνο μεταφοράς και μοντελοποιείται από ένα μέτρο πιθανότητας π ορισμένο στον $X \times Y$. Άτυπα μιλώντας και πάλι, το $d\pi(x, y)$ εκφράζει την ποσότητα μάζας που μεταφέρεται από την τοποθεσία x στην τοποθεσία y. Προκειμένου ένα πλάνο μεταφοράς $\pi \in P(X \times Y)$ να είναι αποδεκτό, θα πρέπει η μάζα που παίρνουμε από το σημείο x να είναι ίση με $d\mu(x)$ και η μάζα που μεταφέρουμε στο σημείο y να είναι ίση με $d\nu(y)$. Αυτό σημαίνει ότι

$$\int_Y d\pi(x, y) = d\mu(x) \qquad \text{και} \qquad \int_X d\pi(x, y) = d\nu(y).$$

Με άλλα λόγια, για όλα τα μετρήσιμα υποσύνολα $A \subseteq X$ και $B \subseteq Y$ απαιτούμε να ισχύει:

$$\pi(A \times Y) = \mu(A) \qquad \text{και} \qquad \pi(X \times B) = \nu(B).$$

Για τα μέτρα πιθανότητας π, τα οποία ικανοποιούν τις παραπάνω σχέσεις, λέμε ότι έχουν περιθώρια μέτρα μ, ν και είναι τα αποδεκτά πλάνα μεταφοράς. Συμβολίζουμε το σύνολο των αποδεκτών πλάνων μεταφοράς με $\Pi(\mu, \nu)$.

Έχουμε, πλέον, έναν ξεκάθαρο μαθηματικό ορισμό για το βασικό πρόβλημα που θα μας απασχολήσει σε αυτή την εργασία.

Πρόβλημα βέλτιστης μεταφοράς μάζας του Kantorovich:

$$\text{Ελαχιστοποίηση ως προς } \pi \in \Pi(\mu, \nu) \text{ του συναρτησοειδούς } I[\pi] = \int_{X \times Y} c(x, y)\, d\pi(x, y).$$

Για ένα δεδομένο πλάνο μεταφοράς π, η μη αρνητική (και πιθανώς άπειρη) ποσότητα $I[\pi]$ λέγεται ολικό κόστος μεταφοράς αναφορικά με το π. Το βέλτιστο κόστος μεταφοράς είναι η τιμή

$$T_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} I[\pi].$$

Τα αποδεκτά πλάνα μεταφοράς π για τα οποία ισχύει $I[\pi] = T_c(\mu, \nu)$, θα είναι τα βέλτιστα πλάνα μεταφοράς.

ΕΙΣΑΓΩΓΙΚΟ ΚΕΦΑΛΑΙΟ

Στο κεφάλαιο αυτό θα καταγράψουμε και θα αποδείξουμε προτάσεις και θεωρήματα τα οποία θα χρειαστούμε στο κυρίως μέρος της εργασίας.

1. Το Θεώρημα Hahn – Banach.

Αναλυτική μορφή του Θεωρήματος Hahn – Banach

Επέκταση γραμμικών συναρτησιακών

Έστω P ένα σύνολο εφοδιασμένο με μια σχέση (μερικής) διατάξεως που συμβολίζεται με $\leq$ και $Q \subseteq P$. Το Q λέγεται ολικά διατεταγμένο αν για κάθε $a, b \in Q$ ισχύει $a \leq b$ ή $b \leq a$ (δηλαδή κάθε δύο στοιχεία του Q σχετίζονται). Το $c \in P$ είναι ένα άνω φράγμα του Q, αν για κάθε $a \in Q$ ισχύει $a \leq c$. Το $m \in P$ είναι ένα μεγιστικό (maximal) στοιχείο του P αν για κάθε $x \in P$ τέτοιο ώστε $m \leq x$, έχουμε αναγκαστικά $x = m$. Το P λέγεται επαγωγικό αν κάθε ολικά διατεταγμένο υποσύνολο του P έχει ένα άνω φράγμα.

Για την απόδειξη του Θεωρήματος Hahn - Banach, θα χρειαστούμε το λήμμα του Zorn, του οποίου παραθέτουμε την εκφώνηση:

Λήμμα 1 (Λήμμα Zorn)

Κάθε διατεταγμένο, επαγωγικό, μη κενό σύνολο έχει ένα μεγιστικό στοιχείο.

Θεώρημα 2 (Hahn-Banach – Επέκταση γραμμικών συναρτησοειδών)

Έστω E ένας γραμμικός χώρος πάνω στο $\mathbb{R}$ και $p\colon E \to \mathbb{R}$ μια απεικόνιση που ικανοποιεί τις σχέσεις:

(1) $$p(\lambda x) = \lambda p(x), \quad \forall x \in E, \forall \lambda > 0$$

(2) $$p(x + y) \leq p(x) + p(y), \quad \forall x, y \in E$$

Έστω, επιπλέον, $G \subseteq E$ ένας γραμμικός υπόχωρος και έστω $g\colon G \to \mathbb{R}$ ένα γραμμικό συναρτησοειδές τέτοιο ώστε $g(x) \leq p(x), \quad \forall x \in G$.
Τότε, υπάρχει γραμμικό συναρτησοειδές f ορισμένο στον E που επεκτείνει το g, δηλαδή

$$g(x) = f(x), \quad \forall x \in G.$$

Μια απεικόνιση $p\colon E \to \mathbb{R}$ που ικανοποιεί τις σχέσεις (1) και (2) λέγεται υποπροσθετική.

Απόδειξη:

Θεωρούμε το σύνολο

$$P = \{h: D(h) \to \mathbb{R} \mid D(h) \text{ γραμμικός υπόχωρος του } E, \quad h \text{ γραμμικό συναρτησοειδές,} \\ G \subseteq D(h), \quad \text{το } h \text{ επεκτείνει το } g \quad \text{και} \quad h(x) \leq p(x), \forall x \in D(h)\}.$$

Το P είναι εφοδιασμένο με τη σχέση (μερικής) διάταξης

$$h_1 \leq h_2 \Leftrightarrow D(h_1) \subseteq D(h_2) \text{ και το } h_2 \text{ επεκτείνει το } h_1.$$

Πράγματι, η σχέση $\leq$ ικανοποιεί τις ιδιότητες μιας σχέσης μερικής διάταξης:

Ανακλαστική: προφανώς $h_1 \leq h_1$ για κάθε $h_1 \in P$.

Αντισυμμετρική: για κάθε $h_1, h_2 \in P$ ισχύει:

$$h_1 \leq h_2 \Rightarrow D(h_1) \subseteq D(h_2) \text{ και το } h_2 \text{ επεκτείνει το } h_1 \Rightarrow h_2 \nleq h_1 \text{ εκτός αν } h_1 = h_2.$$

Μεταβατική: για κάθε $h_1, h_2, h_3 \in P$ ισχύει:

$$h_1 \leq h_2 \Rightarrow D(h_1) \subseteq D(h_2) \text{ και το } h_2 \text{ επεκτείνει το } h_1$$

$$h_2 \leq h_3 \Rightarrow D(h_2) \subseteq D(h_3) \text{ και το } h_3 \text{ επεκτείνει το } h_2$$

Άρα, $D(h_1) \subseteq D(h_2) \subseteq D(h_3)$ και $\forall x \in D(h_1)$: $h_3(x) = h_2(x) = h_1(x)$, δηλαδή το h_3 επεκτείνει το h_1, οπότε $h_1 \leq h_3$.

Επίσης, $P \neq \emptyset$, αφού $g \in P$. Ακόμα, το P είναι επαγωγικό.

Πράγματι, έστω $Q \subseteq P$ ένα ολικά διατεταγμένο υποσύνολο του P. Θέτουμε $Q = \{h_i : i \in I\}$ και ορίζουμε $D(h) \coloneqq \bigcup_{i \in I} D(h_i)$ και $h(x) \coloneqq h_i(x)$, αν $x \in D(h_i)$.

Ο ορισμός αυτός, πράγματι, έχει νόημα επειδή το Q είναι ολικά διατεταγμένο. Ο $D(h)$ θα είναι γραμμικός υπόχωρος του E, αφού: αν $x, y \in D(h)$, τότε $x \in D(h_i)$ και $y \in D(h_j)$ για κάποια $i, j \in I$. Θα ισχύει: $h_i \leq h_j$ ή $h_j \leq h_i$. Έστω ότι $h_i \leq h_j$. Τότε $x \in D(h_i)$, άρα $x + y \in D(h_j) \subseteq D(h)$. Ακόμα, αν $x \in D(h)$, τότε $x \in D(h_i)$ για κάποιο $i \in I$, οπότε $\forall \lambda \in \mathbb{R} : \lambda x \in D(h_i) \subset D(h)$. Επίσης, ο ορισμός του $h(x) = h_i(x)$, $\alpha\nu\, x \in D(h_i)$ είναι μονοσήμαντος, αφού αν $x \in D(h_i)$ και $x \in D(h_j)$, τότε $h_i \leq h_j$ ή $h_j \leq h_i$, άρα $h_j(x) = h_i(x)$.

Ακόμα, το συναρτησοειδές $h: D(h) \subseteq E \to \mathbb{R}$ είναι γραμμικό, αφού

$$\forall x, y \in D(h), \quad x \in D(h_i) \text{ για κάποιο } h_i \quad \text{και} \quad y \in D(h_j) \text{ για κάποιο } h_j.$$

Θα ισχύει μια από τις σχέσεις: $h_i \leq h_j$ ή $h_j \leq h_i$, έστω $h_i \leq h_j$. Τότε, $D(h_i) \subseteq D(h_j)$, δηλαδή $x \in D(h_j)$ και επίσης $x + y \in D(h_j)$. Άρα,

$$h(x + y) = h_j(x + y) = h_j(x) + h_j(y) = h(x) + h(y).$$

Επιπλέον, $G \subseteq D(h_i) \quad \forall i \in I$, άρα $G \subseteq \bigcup_{i\in I} D(h_i\,) = D(h)$ και για κάθε $x \in D(h)$, θα έχουμε $x \in D(h_i)$ για κάποιο h_i, άρα $h(x) = h_i(x) \leq p(x)$. Τελικά, $h \in P$.

Επίσης, $\forall h_i \in Q$ ισχύει $h_i \leq h$. Πράγματι, $D(h_i) \subseteq \bigcup_{i\in I} D(h_i) = D(h)$ και $\forall x \in D(h_i) : h_i(x) = h(x)$ εξ' ορισμού. Δηλαδή, το $h \in P$ αποτελεί ένα άνω φράγμα του Q και συνεπώς το P είναι επαγωγικό. Από το λήμμα Zorn, προκύπτει ότι το P έχει ένα μεγιστικό στοιχείο, το οποίο ονομάζουμε f.

Θα δείξουμε, με απαγωγή σε άτοπο, ότι $D(f) = E$. Υποθέτουμε ότι $D(f) \neq E$. Έστω $x_0 \notin D(f)$. Θέτουμε $D(h) = D(f) + \mathbb{R} \cdot x_0$, και για $x \in D(f)$ και $t \in \mathbb{R}, \quad h(x + tx_0) = f(x) + ta$, όπου a είναι μια σταθερά που θα ορίσουμε αργότερα έτσι ώστε $h \in P$.

Πρέπει να εξασφαλίσουμε ότι $f(x) + ta \leq p(x + tx_0), \quad \forall x \in D(f), \forall t \in \mathbb{R}$.

Λόγω της σχέσης (1), αρκεί να επαληθεύσουμε την παραπάνω ανισότητα για $t = 1$ και $t = -1$, δηλαδή:

$$f(x) + a \leq p(x + x_0), \quad \forall x \in D(f)$$

$$f(x) - a \leq p(x - x_0), \quad \forall x \in D(f).$$

Με άλλα λόγια, πρέπει να επιλέξουμε τη σταθερά a, έτσι ώστε

$$\sup_{y\in D(f)} \{f(y) - p(y - x_0)\} \leq a \leq \inf_{x\in D(f)} \{p(x + x_0) - f(x)\}.$$

Μια τέτοια επιλογή είναι δυνατή, αφού λόγω της σχέσης (2), για κάθε $x, y \in D(f)$ ισχύει:

$$f(x) + f(y) = f(x + y) \leq p(x + y) = p(x + x_0 + y - x_0) \leq p(x + x_0) + p(y - x_0),$$

δηλαδή

$$f(y) - p(y - x_0) \leq p(x + x_0) - f(x), \;\; x, y \in D(f).$$

Επομένως, επιλέγοντας μια τέτοια σταθερά a, έχουμε $h \in P$.

Επίσης, $D(f) \subseteq D(h)$ και για κάθε $x \in D(f)$ έχουμε $h(x) = h(x + 0 \cdot a) = f(x) + 0 \cdot a = f(x)$, δηλαδή $f \leq h$, αλλά $f \neq h$. Αυτό αντιφάσκει με το γεγονός ότι το f είναι μεγιστικό στοιχείο. Άρα, $D(f) = E$.

Δηλαδή, $f: E \to \mathbb{R}$, f γραμμικό, $G \subseteq D(f) = E$, το f επεκτείνει το g και για κάθε $x \in D(h) = E$ ισχύει $f(x) \leq p(x)$. Επομένως, η απόδειξη του θεωρήματος είναι πλήρης.

■

Θα αποδείξουμε ότι για να ισχύει η ανισότητα

$$f(x) + ta \leq p(x + tx_0), \quad \forall t \in \mathbb{R}, \forall x \in D(f)$$

αρκεί να την επαληθεύσουμε για $t = 1$ και $t = -1$.

Απόδειξη:

Αν η παραπάνω ανισότητα ισχύει για $t = 1$, τότε θα ισχύει για κάθε $t > 0$. Πράγματι, αν για κάθε $x \in D(f)$ έχουμε $f(x) + a \le p(x + x_0)$, τότε για αυθαίρετο $t > 0$, έχουμε:

$$f(x) + ta = f\left(t\frac{x}{t}\right) + ta = t\left[f\left(\frac{x}{t}\right) + a\right] \le tp\left(\frac{x}{t} + x_0\right) = p(x + tx_0).$$

Αν η παραπάνω ανισότητα ισχύει για $t = -1$, τότε θα ισχύει για κάθε $t < 0$. Πράγματι, αν για κάθε $x \in D(f)$ έχουμε $f(x) - a \le p(x - x_0)$, τότε για αυθαίρετο $t < 0$, έχουμε:

$$f(x) + ta = f\left(t\frac{x}{t}\right) + ta = -t\left[f\left(-\frac{x}{t}\right) - a\right] \le -tp\left(-\frac{x}{t} - x_0\right) = p(x + tx_0).$$

Για $t = 0$, η ανισότητα είναι αληθής αφού για κάθε $x \in D(f)$ έχουμε $f(x) \le p(x)$.

■

Γεωμετρική μορφή του Θεωρήματος Hahn – Banach

Διαχωρισμός κυρτών συνόλων

Σε ότι ακολουθεί, το E συμβολίζει ένα γραμμικό χώρο με νόρμα.

Ορισμός 3

Ένα υπερεπίπεδο είναι ένα σύνολο της μορφής $H = \{x \in E : f(x) = a\}$, όπου f είναι ένα γραμμικό συναρτησοειδές πάνω στον E, $f \not\equiv 0$ και $a \in \mathbb{R}$. Λέμε ότι το H είναι ένα υπερεπίπεδο με εξίσωση $[f = a]$.

Πρόταση 4

Το υπερεπίπεδο H με εξίσωση $[f = a]$ είναι κλειστό αν και μόνο αν το f είναι συνεχές.

Απόδειξη:

Όσον αφορά το αντίστροφο, υποθέτουμε ότι το f είναι συνεχές και $(x_n)_{n=1}^{\infty}$ είναι μια ακολουθία στοιχείων του H τέτοια ώστε $x_n \to x$ καθώς $n \to +\infty$. Τότε, $f(x_n) \to f(x)$ και $f(x_n) = a$, για κάθε $n \in \mathbb{N}$, συνεπώς $f(x) = a$, δηλαδή $x \in H$. Έτσι, λοιπόν, αποδεικνύεται ότι το H είναι κλειστό.

Για το ευθύ, υποθέτουμε ότι το H είναι κλειστό. Τότε, το συμπλήρωμα αυτού, το H^c, είναι ανοικτό και επίσης, $H^c \neq \emptyset$, αφού $f \not\equiv 0$. Έστω, λοιπόν, $x_0 \in H^c$. Θα ισχύει, τότε, μία από τις σχέσεις $f(x_0) < a$ ή $f(x_0) > a$. Έστω ότι $f(x_0) < a$. Θα υπάρχει $r > 0$ τέτοιο ώστε $B_r(x_0) \subset H^c$, όπου $B_r(x_0) = \{x \in E : \|x - x_0\| < r\}$. Τότε, θα ισχύει

$$f(x) < a, \quad \forall x \in B_r(x_0) \tag{3}$$

Πράγματι, ας υποθέσουμε ότι υπάρχει $x_1 \in B_r(x_0)$ τέτοιο ώστε $f(x_1) > a$. Τότε, το ευθύγραμμο τμήμα $\{x_t = tx_0 + (1-t)x_1 : t \in [0,1]\} \subseteq B_r(x_0)$ αφού το $B_r(x_0)$ είναι κυρτό σύνολο, και άρα $\forall t \in [0,1],\ \ f(x_t) \neq a$.
Όμως, μπορούμε να βρούμε $t \in (0,1)$ ώστε $f(x_t) = a$. Πράγματι,

$$f(x_t) = a \iff f[tx_0 + (1-t)x_1] = a \iff tf(x_0) + (1-t)f(x_1) = a \iff$$

$$t = \frac{f(x_1) - a}{f(x_1) - f(x_0)}$$

και επίσης, $t \in (0,1)$, αφού $f(x_0) < a$, άρα $f(x_1) - a < f(x_1) - f(x_0)$.

Έτσι καταλήγουμε σε άτοπο, οπότε η (3) ισχύει.

Από τη σχέση (3), λοιπόν, έχουμε $f(x_0 + rz) < a, \quad \forall z \in B_1(0)$.

Άρα,

$$f(x_0) + rf(z) < a, \quad \forall z \in B_1(0)$$

ή

$$f(z) < \frac{a - f(x_0)}{r}, \quad \forall z \in B_1(0).$$

Οπότε,

$$\sup_{\|z\| \leq 1} f(z) \leq \frac{a - f(x_0)}{r} = M$$

Άρα το γραμμικό συναρτησοειδές f είναι φραγμένο, ισοδύναμα το f είναι συνεχές.

■

Ορισμός 5

Έστω $A \subseteq E,\ B \subseteq E$.
Το υπερεπίπεδο H με εξίσωση $[f = a]$ διαχωρίζει τα A και B με την ευρεία έννοια αν:

$$f(x) \leq a, \quad \forall x \in A \quad \text{και} \quad f(x) \geq a, \quad \forall x \in B.$$

Το H διαχωρίζει τα A και B με τη στενή έννοια αν υπάρχει $\varepsilon > 0$ τέτοιο ώστε

$$f(x) \leq a - \varepsilon, \quad \forall x \in A \quad \text{και} \quad f(x) \geq a + \varepsilon, \quad \forall x \in B.$$

Γεωμετρικά, ο διαχωρισμός εκφράζει ότι τα A και B βρίσκονται «εκατέρωθεν του H».

Θεώρημα 6 (Hahn-Banach, 1η γεωμετρική μορφή)

Έστω $A \subseteq E$, $B \subseteq E$ δύο κυρτά σύνολα, μη κενά και ξένα. Υποθέτουμε ότι το A είναι ανοιχτό.

Τότε, υπάρχει ένα κλειστό υπερεπίπεδο που διαχωρίζει τα A και B με την ευρεία έννοια.

Η απόδειξη του Θεωρήματος θα βασιστεί στα δύο ακόλουθα λήμματα.

Λήμμα 7 (Στάθμη ενός κυρτού συνόλου)

Έστω $C \subseteq E$ ένα κυρτό και ανοιχτό σύνολο, με $0 \in C$. Για κάθε $x \in E$, θέτουμε:

(4) $$p(x) = \inf\left\{a > 0 : \frac{1}{a}x \in C\right\}$$

(το p λέγεται στάθμη του συνόλου C). Το p ικανοποιεί τις ακόλουθες σχέσεις:

(5) $$p(\lambda x) = \lambda p(x), \quad \forall x \in E \; \kappa\alpha\iota \; \lambda > 0$$

(6) $$p(x+y) \leq p(x) + p(y), \quad \forall x, y \in E$$

(7) $$\text{υπάρχει } M > 0 \text{ τ.ω } 0 \leq p(x) \leq M\|x\|, \quad \forall x \in E$$

(8) $$C = \{x \in E : p(x) < 1\}$$

<u>Απόδειξη:</u>

Θα αποδείξουμε πρώτα την (7).

Εφόσον το C είναι ανοιχτό και $0 \in C$, θα υπάρχει $r > 0$ τέτοιος ώστε $B_r(0) \subseteq C$. Παρατηρούμε ότι για κάθε $x \in E \setminus \{0\}$ και $a \in \mathbb{R}$ τέτοιο ώστε $a > \frac{\|x\|}{r}$ ισχύει

$$\left\|\frac{1}{a}x\right\| = \frac{1}{a}\|x\| < \frac{r}{\|x\|}\|x\| = r$$

δηλαδή $\frac{1}{a}x \in B_r(0) \subseteq C$. Άρα,

$$\forall x \in E, \quad p(x) = \inf\left\{a > 0 : \frac{1}{a}x \in C\right\} \leq \frac{1}{r}\|x\|.$$

Επομένως ισχύει η (7) με $M = \frac{1}{r}$. Στη συνέχεια, αποδεικνύουμε την (5).

Για κάθε $x \in E$ και $\lambda > 0$ ισχύει $p(\lambda x) = \inf\{a > 0 : \frac{1}{a}\lambda x \in C\}$. Θέτοντας $b = \frac{a}{\lambda}$ έχουμε:

$$p(\lambda x) = \inf\left\{b\lambda : \frac{1}{b}x \in C\right\} = \lambda \inf\left\{b : \frac{1}{b}x \in C\right\} = \lambda p(x).$$

Συνεχίζουμε με την απόδειξη της (8).

Έστω $x \in C$. Επειδή το C είναι ανοιχτό, υπάρχει $\varepsilon > 0$ τέτοιο ώστε $(1+\varepsilon)x \in C$. Άρα, $p(x) \leq \frac{1}{1+\varepsilon} < 1$, δηλαδή $C \subseteq \{x \in E : p(x) < 1\}$.

Για τον αντίστροφο εγκλεισμό, έστω $x \in E$ τέτοιο ώστε $p(x) < 1$. Δηλαδή, υπάρχει a με $0 < a < 1$ τέτοιο ώστε $\frac{1}{a}x \in C$. Τότε, όμως, επειδή το C είναι κυρτό, έχουμε $x = a\left(\frac{1}{a}x\right) + (1-a)0 \in C$.

Συνεπώς, $\{x \in E : p(x) < 1\} \subseteq C$, δηλαδή η (8) είναι αληθής. Τέλος, θα αποδείξουμε την (6).

Έστω $x, y \in E$ και $\varepsilon > 0$. Από τις σχέσεις (5) και (8), γνωρίζουμε ότι $\frac{1}{p(x)+\varepsilon}x \in C$ και ομοίως $\frac{1}{p(y)+\varepsilon}y \in C$. Άρα, λόγω κυρτότητας του C,

$$\forall t \in [0,1], \qquad \frac{t}{p(x)+\varepsilon}x + \frac{1-t}{p(y)+\varepsilon}y \in C.$$

Ειδικά, για $t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$ παίρνουμε

$$\frac{p(x)+\varepsilon}{(p(x)+p(y)+2\varepsilon)(p(x)+\varepsilon)}x + \frac{p(y)+\varepsilon}{(p(x)+p(y)+2\varepsilon)(p(y)+\varepsilon)}y = \frac{1}{p(x)+p(y)+2\varepsilon}(x+y) \in C.$$

Άρα, από την (8)

$$p\left(\frac{1}{p(x)+p(y)+2\varepsilon}(x+y)\right) < 1$$

από την οποία, λόγω της (5)

$$\frac{p(x+y)}{p(x)+p(y)+2\varepsilon} < 1$$

και επομένως $p(x+y) \leq p(x) + p(y) + 2\varepsilon$.

Επειδή, το ε ήταν ένας τυχαίος θετικός αριθμός θα ισχύει $p(x+y) \leq p(x) + p(y)$ δηλαδή η (6) είναι αληθής.

■

Λήμμα 8

Έστω $C \subseteq E$ ένα κυρτό σύνολο, ανοιχτό και μη κενό. Έστω, επίσης, $x_0 \in E \setminus C$. Τότε, υπάρχει συνάρτηση $f \in E^*$ τέτοια ώστε $f(x) < f(x_0)$ για κάθε $x \in C$. Ειδικά, το υπερεπίπεδο με εξίσωση $[f = f(x_0)]$ διαχωρίζει τα $\{x_0\}$ και C με την ευρεία έννοια.

Απόδειξη:

1η περίπτωση: $0 \in C$

Εισάγουμε τη στάθμη p του C, όπως ορίστηκε προηγουμένως στο λήμμα 7. Θεωρούμε το χώρο $G \coloneqq \mathbb{R} \cdot x_0 = \{tx_0 : t \in \mathbb{R}\}$ και το γραμμικό συναρτησιακό $g: G \to \mathbb{R}$ με $g(tx_0) = t, \quad t \in \mathbb{R}$.

Ισχύει $g(x) \leq p(x), \quad \forall x \in G$.

Πράγματι, αν $x \in G$, τότε $x = tx_0$ για κάποιο $t \in \mathbb{R}$.

Αν $t > 0$, έχουμε:

$$g(x) \leq p(x) \Leftrightarrow g(tx_0) \leq p(tx_0) \Leftrightarrow t \leq \inf\left\{a > 0 : \frac{1}{a} tx_0 \in C\right\}$$

$$\Leftrightarrow t \leq t \cdot \inf\left\{a > 0 : \frac{1}{a} x_0 \in C\right\} \Leftrightarrow 1 \leq p(x_0)$$

το οποίο ισχύει αφού το $x_0 \notin C$.

Αν $t = 0$, τότε $g(0) = 0$ και $p(0) = 0$, οπότε ισχύει $g(0) \leq p(0)$.

Αν $t < 0$, έχουμε:

$$g(x) \leq p(x) \Leftrightarrow g(tx_0) \leq p(tx_0) \Leftrightarrow t \leq \inf\left\{a > 0 : \frac{1}{a} tx_0 \in C\right\}$$

το οποίο ισχύει αφού ένας αρνητικός αριθμός είναι μικρότερος ή ίσος από έναν μη αρνητικό.

Εφόσον, λοιπόν, ο G είναι ένας γραμμικός υπόχωρος του E και $p(x)$ είναι μια υποπροσθετική συνάρτηση που κυριαρχεί το $g(x)$ στο G, σύμφωνα με το Θεώρημα Hahn-Banach, υπάρχει $f \in E^*$ το οποίο επεκτείνει το g και τέτοιο ώστε $f(x) \leq g(x), \quad \forall x \in E$.

Ειδικότερα, $f(x_0) = g(x_0) = 1$ και λόγω της σχέσης (7), για κάθε $x \in E$ έχουμε $f(x) \leq p(x) \leq M\|x\|$, δηλαδή το f είναι φραγμένο, οπότε και συνεχές.

Επίσης, λόγω της σχέσης (8), για κάθε $x \in C$, ισχύει $f(x) \leq p(x) < 1 = f(x_0)$ και έτσι η απόδειξη για την 1η περίπτωση είναι πλήρης.

2η περίπτωση: $0 \notin C$

Παίρνουμε κάποιο $x_1 \in C$ και θεωρούμε το σύνολο $C' = C - x_1$, δηλαδή τη μεταφορά του C κατά το διάνυσμα x_1. Το C' είναι, προφανώς, κυρτό, ανοιχτό και μη κενό και μάλιστα $0 \in C'$, οπότε εργαζόμαστε όπως και στην 1η περίπτωση με το νέο σύνολο C'.

■

Απόδειξη του Θεωρήματος Hahn-Banach, 1η γεωμετρική μορφή:

Θέτουμε $C \coloneqq A - B = \{a - b : a \in A, b \in B\}$.

Το C είναι κυρτό. Πράγματι, αν $c_1 = a_1 - b_1,\ \ c_2 = a_2 - b_2 \in C$ και $t \in (0,1)$ τότε:

$$\begin{aligned} tc_1 + (1-t)c_2 &= t(a_1 - b_1) + (1-t)(a_2 - b_2) = (ta_1 + (1-t)a_2) - (tb_1 + (1-t)b_2) \\ &= a' - b' \in A - B = C. \end{aligned}$$

Το C είναι ανοιχτό. Πράγματι,

$$C = \bigcup_{b \in B} (A - b).$$

και για κάθε $b \in B$, το $(A - b)$ είναι ανοιχτό. Επομένως, το C είναι ανοιχτό, ως ένωση ανοιχτών συνόλων.

Ακόμα, $0 \notin C$. Πράγματι, αν υποθέσουμε ότι το $0 \in C$, τότε $0 = a - b$, για κάποια $a \in A$ και $b \in B$. Δηλαδή, $a = b \in A \cap B$, το οποίο όμως είναι άτοπο γιατί $A \cap B = \emptyset$.

Άρα, από το λήμμα 8, υπάρχει $f \in E^*$ τέτοια ώστε $\ f(z) < f(0) = 0, \quad \forall z \in C$.

Αν $\ z = x - y,\ \ x \in A, y \in B$, τότε $\ f(x) - f(y) = f(x - y) < 0$.

Άρα, για κάθε $x \in A$ και $y \in B$ ισχύει $f(x) < f(y)$. Άρα, $\sup_{x \in A} f(x) \leq \inf_{y \in B} f(y)$ και επομένως για σταθερό $a \in \mathbb{R}$ τέτοιο ώστε $\sup_{x \in A} f(x) \leq a \leq \inf_{y \in B} f(y)$, το υπερεπίπεδο με εξίσωση $[f = a]$ διαχωρίζει τα A και B με την ευρεία έννοια.

■

Θεώρημα 9 (Hahn-Banach, 2η γεωμετρική μορφή)

Έστω $A \subseteq E, B \subseteq E$ δύο κυρτά σύνολα, μη κενά και ξένα. Υποθέτουμε ότι το A είναι κλειστό και το B συμπαγές. Τότε, υπάρχει ένα κλειστό υπερεπίπεδο που διαχωρίζει τα A και B με τη στενή έννοια.

Απόδειξη:

Έστω $\varepsilon > 0$. Θέτουμε $A_\varepsilon = A + B_\varepsilon(0)$ και $B_\varepsilon = B + B_\varepsilon(0)$, οπότε τα σύνολα $A_\varepsilon, B_\varepsilon$ είναι κυρτά, ανοιχτά και μη κενά. Επιπλέον, αν επιλέξουμε το $\varepsilon > 0$ αρκούντως μικρό, μπορούμε να πετύχουμε: $A_\varepsilon \cap B_\varepsilon = \emptyset$.

Πράγματι, υποθέτουμε ότι $\forall \varepsilon > 0, \quad A_\varepsilon \cap B_\varepsilon \neq \emptyset$. Για κάθε $n \in \mathbb{N}$, επιλέγουμε $x_n \in A$ και $y_n \in B$ έτσι ώστε $\|x_n - y_n\| \leq \frac{2}{n}$. Επειδή το B είναι συμπαγές, θα υπάρχει συγκλίνουσα υπακολουθία της

y_n, έστω $y_{n_k} \to y \in B$. Τότε, όμως, $\|x_{n_k} - y_{n_k}\| \leq \frac{2}{n}$ και άρα $x_{n_k} \to y$. Επειδή, όμως, το A είναι κλειστό, θα πρέπει $y \in A$. Τελικά, $y \in A \cap B = \emptyset$, οπότε έχουμε οδηγηθεί σε άτοπο.

Σύμφωνα με την 1η γεωμετρική μορφή του Θεωρήματος Hahn-Banach, υπάρχει κλειστό υπερεπίπεδο με εξίσωση $[f = a]$ το οποίο διαχωρίζει τα A_ε και B_ε με την ευρεία έννοια. Επομένως,

$$f(x + \varepsilon z) \leq a \leq f(y + \varepsilon z), \quad \forall x \in A, \forall y \in B, \forall z \in B_1(0)$$

$$f(x) + \varepsilon f(z) \leq a \leq f(y) + \varepsilon f(z), \quad \forall x \in A, \forall y \in B, \forall z \in B_1(0)$$

και άρα, για κάθε $x \in A, y \in B$, έχουμε:

$$\sup_{\|z\|\leq 1}\{f(x) + \varepsilon f(z)\} \leq a \leq \inf_{\|z\|\leq 1}\{f(y) + \varepsilon f(z)\}$$

$$f(x) + \varepsilon \sup_{\|z\|\leq 1} f(z) \leq a \leq f(y) + \varepsilon \inf_{\|z\|\leq 1} f(z)$$

$$f(x) + \varepsilon \sup_{\|z\|\leq 1} f(z) \leq a \leq f(y) - \varepsilon \sup_{\|z\|\leq 1}\{-f(z)\}$$

$$f(x) + \varepsilon\|f\| \leq a \leq f(y) - \varepsilon\|f\|$$

Άρα,

$$f(x) \leq a - \varepsilon\|f\|, \quad \forall x \in A$$

$$f(y) \geq a + \varepsilon\|f\|, \quad \forall y \in B$$

και αφού $\|f\| \neq 0$, συμπεραίνουμε ότι τα A και B διαχωρίζονται με τη στενή έννοια από το υπερεπίπεδο με εξίσωση $[f = a]$.

■

2. Το Λήμμα του Urysohn.

Ορισμός 10

Έστω X ένας τοπολογικός χώρος.

Ο φορέας μιας συνάρτησης $f: X \to \mathbb{R}$ είναι το σύνολο $supp(f) = \overline{\{x \in X : f(x) \neq 0\}}$.

Το σύνολο των συναρτήσεων $f: X \to \mathbb{R}$ που έχουν συμπαγή φορέα συμβολίζεται με $C_c(X)$.

Ο χώρος X λέγεται τοπικά συμπαγής, αν κάθε σημείο του X έχει μια περιοχή της οποίας η κλειστότητα είναι συμπαγής.

Ο χώρος X λέγεται χώρος Hausdorff, αν κάθε δύο διακεκριμένα σημεία του μπορούν να διαχωριστούν με ξένες περιοχές, δηλαδή για κάθε $p, q \in X$ με $p \neq q$, υπάρχουν U περιοχή του p και V περιοχή του q τέτοιες ώστε $U \cap V = \emptyset$.

Ο χώρος X λέγεται κανονικός χώρος Hausdorff, αν οποιαδήποτε δύο συμπαγή υποσύνολα αυτού μπορούν να διαχωριστούν με ξένα ανοικτά σύνολα, δηλαδή για κάθε $K_1, K_2 \subseteq X$ συμπαγή, υπάρχουν ανοικτά U_1, U_2 με $U_1 \supseteq K_1, U_2 \supseteq U_2$ και $U_1 \cap U_2 = \emptyset$.

Θεώρημα 11

Έστω X ένας χώρος Hausdorff, $K \subseteq X$ συμπαγές και $p \in K^c$. Τότε, υπάρχουν ανοικτά σύνολα $U, W \subseteq X$, τέτοια ώστε $p \in U, K \subseteq W$ και $U \cap W = \emptyset$.

Με άλλα λόγια, σε έναν χώρο Hausdorff, ένα συμπαγές σύνολο και ένα σημείο εκτός αυτού μπορούν να διαχωριστούν με ανοικτά σύνολα.

Απόδειξη:

Εάν $q \in K$, επειδή ο X είναι χώρος Hausdorff, υπάρχουν ξένα ανοικτά σύνολα U_q και V_q τέτοια ώστε $p \in U_q$ και $q \in V_q$. Εφόσον, το K είναι συμπαγές, υπάρχουν σημεία $q_1, q_2, \dots, q_n \in K$ τέτοια ώστε $K \subseteq V_{q_1} \cup V_{q_2} \cup \dots \cup V_{q_n}$.

Τότε, το συμπέρασμα του Θεωρήματος ικανοποιείται από τα (ανοικτά) σύνολα $U = U_{q_1} \cap U_{q_2} \cap \dots \cap U_{q_n}$ και $W = V_{q_1} \cup V_{q_2} \cup \dots \cup V_{q_n}$, τα οποία είναι ξένα, αφού αν υποθέσουμε ότι υπάρχει $x \in U \cap W$, τότε $x \in U_{q_i}$ $\forall i = 1,2, \dots, n$ και $x \in V_{q_j}$ για κάποιο $j = 1,2, \dots, n$. Όμως, τότε, θα έχουμε $x \in U_{q_j} \cap V_{q_j}$, το οποίο αντιφάσκει με την κατασκευή των U_{q_j}, V_{q_j}.

■

Θεώρημα 12

Εάν $\{K_a\}$ είναι μια οποιαδήποτε συλλογή συμπαγών υποσυνόλων ενός τοπολογικού χώρου και

$$\bigcap_a K_a = \emptyset$$

τότε υπάρχει πεπερασμένη υποσυλλογή της $\{K_\alpha\}$ που έχει, επίσης, κενή τομή.

Απόδειξη:

Για κάθε a, έστω $V_a = K_a^c$. Σταθεροποιούμε ένα μέλος K_1 της $\{K_a\}$. Εφόσον για κάθε σημείο $x \in K_1$, υπάρχει a ώστε $x \in K_a^c = V_a$, η συλλογή $\{V_a\}$ είναι μια ανοικτή κάλυψη του K_1. Το K_1 είναι συμπαγές, άρα $K_1 \subseteq V_{a_1} \cup V_{a_2} \cup \dots \cup V_{a_n}$ για κάποια πεπερασμένη συλλογή $\{V_{a_i}\}_{i=1}^{n}$. Απ' αυτό, συνεπάγεται ότι $K_1 \cap K_{a_1} \cap \dots \cap K_{a_n} = \emptyset$.

■

Θεώρημα 13

Έστω X ένας τοπικά συμπαγής χώρος Hausdorff, $U \subseteq X$ ανοικτό και $K \subseteq U$ συμπαγές.

Τότε, υπάρχει ανοικτό σύνολο V με συμπαγή κλειστότητα τέτοιο ώστε $K \subseteq V \subseteq \overline{V} \subseteq U$.

Απόδειξη:

Εφόσον ο X είναι τοπικά συμπαγής, για κάθε $x \in K$ υπάρχει περιοχή $U_x \ni x$ και $\overline{U_x}$ συμπαγής. Μπορούμε να υποθέσουμε ότι $U_x \subseteq U$, αφού σε αντίθετη περίπτωση παίρνοντας $U_x' = U_x \cap U$, η $U_x{}'$ συνεχίζει να αποτελεί περιοχή του x και επίσης, $\overline{U_x{}'} = \overline{U_x \cap U} \subseteq \overline{U_x} \cap \overline{U} \subseteq \overline{U_x}$, δηλαδή $\overline{U_x{}'}$ συμπαγής ως κλειστό υποσύνολο συμπαγούς συνόλου. Ακόμα, λόγω συμπάγειας του K, υπάρχουν $x_1, x_2, \dots, x_n \in K$ τέτοια ώστε $K \subseteq U_{x_1} \cup U_{x_2} \cup \dots \cup U_{x_n}$. Επομένως, αν $G \coloneqq U_{x_1} \cup U_{x_2} \cup \dots \cup U_{x_n}$, τότε $K \subseteq G$ και επίσης το $\overline{G} = \overline{U_{x_1}} \cup \overline{U_{x_2}} \cup \dots \cup \overline{U_{x_n}}$ είναι συμπαγές ως πεπερασμένη ένωση συμπαγών συνόλων.

Αν $U = X$, τότε επιλέγουμε $V = G$. Σε αντίθετη περίπτωση, σύμφωνα με το Θεώρημα 11, σε κάθε $p \in U^c$ αντιστοιχεί ένα ανοικτό σύνολο W_p τέτοιο ώστε $K \subseteq W_p$ και $p \notin \overline{W_p}$.

Άρα, η $\{U^c \cap \overline{G} \cap \overline{W_p} \ : \ p \in U^c\}$ είναι μια συλλαγή συμπαγών συνόλων με κενή τομή. Λόγω του Θεωρήματος 12, υπάρχουν σημεία $p_1, p_2, \dots, p_n \in U^c$ τέτοια ώστε

$$U^c \cap \overline{G} \cap \overline{W_{p_1}} \cap \overline{W_{p_2}} \cap \dots \cap \overline{W_{p_n}} = \emptyset.$$

Τότε, το σύνολο $V \coloneqq G \cap W_{p_1} \cap W_{p_2} \cap \dots \cap W_{p_n}$ έχει όλες τις απαιτούμενες ιδιότητες, αφού $K \subseteq V \subseteq \overline{V}$ και $\overline{V} \subseteq U$ διότι

$$\overline{V} \subseteq \overline{G} \cap \overline{W_{p_1}} \cap \overline{W_{p_2}} \cap \dots \cap \overline{W_{p_n}} \subseteq U$$

αφού

$$U^c \cap \left(\overline{G} \cap \overline{W_{p_1}} \cap \overline{W_{p_2}} \cap \dots \cap \overline{W_{p_n}}\right) = \emptyset.$$

■

Πριν διατυπώσουμε το λήμμα του Urysohn, θα καταγράψουμε μερικούς συμβολισμούς, ορισμούς και ιδιότητες που θα χρησιμοποιήσουμε.

Συμβολισμοί.

Το $K \prec f$ σημαίνει ότι το K είναι ένα συμπαγές υποσύνολο του X, $f \in C_c(X)$, $0 \leq f(x) \leq 1$ για κάθε $x \in X$ και $f(x) = 1$ για κάθε $x \in K$.

Το $f \prec V$ σημαίνει ότι το V είναι ένα ανοικτό υποσύνολο του X, $f \in C_c(X)$, $0 \leq f(x) \leq 1$ για κάθε $x \in X$ και ο φορέας της f βρίσκεται μέσα στο V.

Το $K \prec f \prec V$ δηλώνει και τα δύο: $K \prec f$ και $f \prec V$.

Ορισμοί – Ιδιότητες.

Έστω X ένας τοπολογικός χώρος.

Μια συνάρτηση $f: X \to \mathbb{R}$ λέγεται κάτω ημισυνεχής αν για κάθε $x \in X$, ισχύει $f(x) \leq \liminf\limits_{y \to x} f(y)$ ενώ λέγεται άνω ημισυνεχής αν για κάθε $x \in X$, ισχύει $f(x) \geq \limsup\limits_{y \to x} f(y)$. Η f είναι συνεχής αν είναι ταυτόχρονα άνω και κάτω ημισυνεχής.

Εύκολα μπορούμε να διαπιστώσουμε ότι το supremum κάτω ημισυνεχών συναρτήσεων είναι κάτω ημισυνεχής και αντίστοιχα, το infimum άνω ημισυνεχών συναρτήσεων είναι άνω ημισυνεχής συνάρτηση.

Θεώρημα 14 (Λήμμα του Urysohn)

Υποθέτουμε ότι ο X είναι ένας τοπικά συμπαγής χώρος Hausdorff, $V \subseteq X$ ανοικτό, $K \subseteq V$ συμπαγές.

Τότε, υπάρχει συνάρτηση $f \in C_c(X)$ τέτοια ώστε $K \prec f \prec V$ και ειδικότερα $1_K \leq f \leq 1_V$.

<u>Απόδειξη:</u>

Ορίζουμε $r_1 = 0, r_2 = 1$ και έστω $r_3, r_4, \ldots$ να είναι μια απαρίθμηση των ρητών αριθμών στο διάστημα $(0,1)$. Από το Θεώρημα 13, μπορούμε να βρούμε ανοικτά σύνολα V_0 και V_1 τέτοια ώστε $\overline{V_0}$ συμπαγές και $K \subseteq V_1 \subseteq \overline{V_1} \subseteq V_0 \subseteq \overline{V_0} \subseteq V$. Υποθέτουμε ότι $n \geq 2$ και τα $V_{r_1}, V_{r_2}, \ldots, V_{r_n}$ έχουν επιλεγεί με τέτοιον τρόπο ώστε $r_i < r_j \Longrightarrow \overline{V_{r_j}} \subseteq V_{r_i}$.

Τότε, από τους $r_1, r_2, \ldots, r_n$, έστω r_i ο μεγαλύτερος αριθμός που είναι μικρότερος από τον r_{n+1} και έστω r_j ο μικρότερος που είναι μεγαλύτερος από τον r_{n+1}.

Χρησιμοποιώντας ξανά το Θεώρημα 13, μπορούμε να βρούμε σύνολο $V_{r_{n+1}}$ τέτοιο ώστε $\overline{V_{r_j}} \subseteq V_{r_{n+1}} \subseteq \overline{V_{r_{n+1}}} \subseteq V_{r_i}$.

Συνεχίζοντας έτσι, αποκτούμε μια συλλογή $\{V_r\}$ ανοικτών συνόλων, ένα για κάθε ρητό $r \in [0,1]$, με τις ακόλουθες ιδιότητες: $K \subseteq V_1$, $\overline{V_0} \subseteq V$, κάθε $\overline{V_r}$ είναι συμπαγές και $r < s \Longrightarrow \overline{V_s} \subseteq V_r$.

Ορίζουμε τις συναρτήσεις:

$$f_r(x) = \begin{Bmatrix} r, & \text{αν } x \in V_r \\ 0, & \text{αλλιώς} \end{Bmatrix} \qquad g_s(x) = \begin{Bmatrix} 1, & \text{αν } x \in \overline{V_s} \\ s, & \text{αλλιώς} \end{Bmatrix}$$

και

$$f(x) = \sup_{r \in [0,1] \cap \mathbb{Q}} f_r(x)\ , \quad g(x) = \inf_{s \in [0,1] \cap \mathbb{Q}} g_s(x).$$

Η f είναι το supremum μιας συλλογής κάτω ημισυνεχών συναρτήσεων και η g είναι το infimum μιας συλλογής άνω ημισυνεχών συναρτήσεων.

Πράγματι, το σύνολο

$$\{x : f_r(x) > a\} = \begin{Bmatrix} X, & \text{αν } a < 0 \\ V_r, & \text{αν } 0 \le a < r \\ \emptyset, & \text{αν } a \ge r \end{Bmatrix}$$

είναι ανοικτό σε κάθε περίπτωση. Άρα, για κάθε ρητό $r \in [0,1]$, η f_r είναι μια κάτω ημισυνεχής συνάρτηση.

Αντίστοιχα, το σύνολο

$$\{x : g_s(x) < a\} = \begin{Bmatrix} X, & \text{αν } a > 1 \\ X \setminus \overline{V_s}, & \text{αν } s < a \le 1 \\ \emptyset, & \text{αν } a \ge s \end{Bmatrix}$$

είναι ανοικτό σε κάθε περίπτωση. Άρα, για κάθε ρητό $s \in [0,1]$, η g_s είναι μια άνω ημισυνεχής συνάρτηση. Επομένως, η f είναι κάτω ημισυνεχής, ενώ η g είναι άνω ημισυνεχής.

Ισχύει ότι $0 \le f \le 1$, $f(x) = 1$ για κάθε $x \in K$ και $supp(f) \subseteq \overline{V_0}$.

Όντως, αν $x \in K$, τότε εφόσον $K \subseteq V_1$, έχουμε $f(x) = \sup_r f_r(x) \ge f_1(x) = 1$ και επειδή $f \le 1$ στο X, προκύπτει ότι $f(x) = 1$ για κάθε $x \in K$.

Ακόμα,

$$supp(f) = \overline{\{x : f(x) \neq 0\}} = \overline{\left\{x : \sup_r f_r(x) \neq 0\right\}} = \overline{\bigcup_{0<r\le 1} V_r} \subseteq \overline{V_0},$$

αφού για κάθε ρητό $0 < r \le 1$, $V_r \subseteq \overline{V_r} \subseteq V_0 \subseteq \overline{V_0}$.

Θα ολοκληρώσουμε την απόδειξη δείχνοντας ότι $f = g$.

Η ανισότητα $f_r(x) > g_s(x)$ είναι δυνατή μόνο όταν $r > s$, $x \in V_r$ και $x \notin \overline{V_s}$. Όμως, $r > s \Longrightarrow V_r \subseteq V_s$. Άρα, $f_r \le g_s$ για όλα τα r και s, οπότε $f \le g$. Στη συνέχεια, υποθέτουμε ότι υπάρχει x ώστε $f(x) < g(x)$. Τότε, θα υπάρχουν ρητοί r, s τέτοιοι ώστε $f(x) < r < s < g(x)$. Εφόσον $f(x) < r$, έχουμε ότι $x \notin V_r$. Εφόσον $g(x) > s$, έχουμε ότι $x \in \overline{V_s}$. Τότε, όμως, από την κατασκευή των V_r, αφού $r_i < r_j \Longrightarrow \overline{V_{r_j}} \subseteq V_{r_i}$, υπάρχει αντίφαση. Συνεπώς, $f = g$.

Τελικά, η f είναι μια συνεχής συνάρτηση, αφού είναι ταυτόχρονα άνω και κάτω ημισυνεχής. Επίσης, ο φορέας της f

$$supp(f) = supp(g) = \overline{\left\{x : \inf_s g_s(x) \neq 0\right\}} \subseteq \overline{\{x : g_0(x) \neq 0\}} \subseteq \overline{V_0}$$

είναι συμπαγής ως κλειστό υποσύνολο συμπαγούς συνόλου, και άρα $f \in C_c(X)$. Τέλος, η f ικανοποιεί την ανισότητα $1_K \leq f \leq 1_V$, αφού

$$x \in K \Longrightarrow f(x) = 1 \geq 1_K(x)$$

$$x \in X - V \Longrightarrow f(x) = \inf_{s \in [0,1] \cap \mathbb{Q}} s = 0 \leq 1_V(x).$$

■

3. Το Θεώρημα Ascoli – Arzela.

Έστω X ένας τοπολογικός χώρος και $\mathcal{F} \subseteq C(X)$.

Το $\mathcal{F}$ λέγεται ισοσυνεχές στο σημείο $x \in X$, αν για κάθε $\varepsilon > 0$ υπάρχει περιοχή U_x του x τέτοια ώστε $\sup_{f \in \mathcal{F}} |f(y) - f(x)| < \varepsilon,\ \ y \in U_x$.

Το $\mathcal{F}$ λέγεται ισοσυνεχές αν είναι ισοσυνεχές σε κάθε $x \in X$.

Το $\mathcal{F}$ λέγεται κατά σημείο φραγμένο, αν το σύνολο $\{f(x) : \ f \in \mathcal{F}\}$ είναι ένα φραγμένο υποσύνολο του $\mathbb{R}$ για κάθε $x \in X$.

Το $\mathcal{F}$ λέγεται ολικά φραγμένο, αν για κάθε $\varepsilon > 0$ το $\mathcal{F}$ μπορεί να καλυφθεί από πεπερασμένου πλήθους μπάλες ακτίνας ε, δηλαδή

$$\forall \varepsilon > 0, \text{υπάρχουν } f_1, \dots, f_k \in \mathcal{F} \text{ έτσι ώστε } \mathcal{F} \subseteq \bigcup_{j=1}^{k} B_\varepsilon(f_j),$$

όπου $B_\varepsilon(f) = \{g \in \mathcal{F} : \ \|g - f\|_\infty < \varepsilon\}$.

Λήμμα 15

Έστω X ένας πλήρης μετρικός χώρος και $K \subseteq X$ κλειστό. Τότε, το K είναι συμπαγές αν και μόνο αν είναι ολικά φραγμένο.

Απόδειξη:

Το ευθύ είναι απλό αφού αν το K είναι συμπαγές, τότε μια οποιαδήποτε κάλυψη του K με μπάλες ακτίνας το πολύ ε και κέντρα στο K περιέχει πεπερασμένη υποκάλυψη.

Αντιστρόφως, έστω K ολικά φραγμένο και ακολουθία (x_n) στο K. Για κάθε $m \geq 1$, υπάρχουν πεπερασμένου πλήθους μπάλες ακτίνας το πολύ $\frac{1}{m}$ που καλύπτουν το K και τουλάχιστον μία από αυτές περιέχει τους όρους x_n για άπειρα το πλήθος n.

Για $m = 1$, παίρνουμε μπάλα B_1 ακτίνας το πολύ 1 τέτοια ώστε το $N_1 = \{n \in \mathbb{N} : \ x_n \in B_1\}$ είναι απειροσύνολο και επιλέγουμε $n_1 \in N_1$.

Για $m = 2$, παίρνουμε μπάλα B_2 ακτίνας το πολύ $\frac{1}{2}$ τέτοια ώστε το $N_2 = \{n \in \mathbb{N} : \ n > n_1$ και $x_n \in B_2 \cap B_1\}$ είναι απειροσύνολο και επιλέγουμε $n_2 \in N_2$.

Για $m = 3$, παίρνουμε μπάλα B_3 ακτίνας το πολύ $\frac{1}{3}$ τέτοια ώστε το $N_3 = \{n \in \mathbb{N} : \ n > n_2$ και $x_n \in B_3 \cap B_2 \cap B_1\}$ είναι απειροσύνολο και επιλέγουμε $n_3 \in N_3$.

Συνεχίζουμε με τον ίδιο τρόπο, δημιουργούμε υπακολουθία $\left(x_{n_k}\right)$ της (x_n) η οποία, εφόσον $x_{n_l} \in B_k$ για κάθε $l \geq k$, είναι ακολουθία Cauchy. Αφού ο X είναι πλήρης, η $\left(x_{n_k}\right)$ συγκλίνει στον X και αφού το K είναι κλειστό, το όριο ανήκει στο K. Άρα, η (x_n) περιέχει συγκλίνουσα υπακολουθία, δηλαδή το K είναι συμπαγές.

■

Θεώρημα 16 (Ascoli – Arzela)

Έστω X ένας συμπαγής χώρος Hausdorff. Αν $\mathcal{F}$ είναι ένα ισοσυνεχές, κατά σημείο φραγμένο υποσύνολο του $C(X)$, τότε το $\mathcal{F}$ είναι ολικά φραγμένο (ως προς την ομοιόμορφη μετρική) και η κλειστότητα του $\mathcal{F}$ στο $C(X)$ είναι συμπαγής.

Απόδειξη:

Έστω $\varepsilon > 0$. Αφού το $\mathcal{F}$ είναι ισοσυνεχές, για κάθε $x \in X$ υπάρχει ανοικτή περιοχή U_x του x τέτοια ώστε

$$\sup_{f \in \mathcal{F}} |f(y) - f(x)| < \frac{\varepsilon}{4}, \quad y \in U_x.$$

Αφού ο X είναι συμπαγής, μπορούμε να επιλέξουμε σημεία $x_1, \dots, x_n \in X$ τέτοια ώστε

$$X = \bigcup_{j=1}^{n} U_{x_j}.$$

Αφού το $\mathcal{F}$ είναι κατά σημείο φραγμένο, για κάθε $j = 1,2,..,n$ το $\{f(x_j) : f \in \mathcal{F}\}$ είναι ένα φραγμένο υποσύνολο του $\mathbb{R}$, άρα και το $\Gamma = \{f(x_j) : f \in \mathcal{F}, j = 1,2, \dots, n\}$ είναι ένα φραγμένο υποσύνολο του $\mathbb{R}$.

Συνεπώς, υπάρχει πεπερασμένο σύνολο $\{z_1, \dots, z_m\}$ το οποίο είναι $\frac{\varepsilon}{4}$−πυκνό στο Γ, δηλαδή

$$\forall f \in \mathcal{F}, \forall j \in \{1,2, \dots, n\}, \text{υπάρχει } k \in \{1,2, \dots, m\} \text{ ώστε } \left|f(x_j) - z_k\right| < \frac{\varepsilon}{4}.$$

Πράγματι, αφού το Γ είναι φραγμένο, υπάρχουν $M_1, M_2 \in \mathbb{R}$ τέτοια ώστε

$$M_1 \leq f(x_j) \leq M_2, \quad \forall f \in \mathcal{F}, \forall j \in \{1,2, \dots, n\}.$$

Αν, λοιπόν, χωρίσουμε το διάστημα $[M_1, M_2]$ σε διαστήματα πλάτους το πολύ $\frac{\varepsilon}{4}$ και ορίσουμε ως z_i τα άκρα των διαστημάτων αυτών, παίρνουμε το ζητούμενο αποτέλεσμα.

Έστω $A = \{x_1, \dots, x_n\}$ και $B = \{z_1, \dots, . z_m\}$. Το σύνολο $B^A = \{g : A \longrightarrow B\}$, των συναρτήσεων από το A στο B, είναι πεπερασμένο (έχει πληθάριθμο m^n). Για κάθε $\varphi \in B^A$ ορίζουμε

$$\mathcal{F}_\varphi = \left\{f \in \mathcal{F} : \left|f(x_j) - \varphi(x_j)\right| < \frac{\varepsilon}{4}, \quad j = 1,2, \dots, n\right\}.$$

Τότε, ισχύει

$$\mathcal{F} = \bigcup_{\varphi \in B^A} \mathcal{F}_\varphi.$$

Πράγματι, έστω $f \in \mathcal{F}$. Τότε, για κάθε $j \in \{1,2, \dots, n\}$ υπάρχει $k(j) \in \{1,2, \dots, m\}$ ώστε $\left|f(x_j) - z_{k(j)}\right| < \frac{\varepsilon}{4}$. Επίσης, υπάρχει $\varphi \in B^A$ τέτοια ώστε $\varphi(x_j) = z_{k(j)}$ για κάθε $j \in \{1,2, \dots, n\}$. Άρα, $\left|f(x_j) - \varphi(x_j)\right| < \frac{\varepsilon}{4}$ για κάθε $j \in \{1, 2, \dots, n\}$, δηλαδή $f \in \mathcal{F}_\varphi$.

Στη συνέχεια, ισχυριζόμαστε ότι κάθε $\mathcal{F}_\varphi$ έχει διάμετρο το πολύ ε, δηλαδή

$$\forall \varphi \in B^A, \quad diam(\mathcal{F}_\varphi) = \sup_{f,g \in \mathcal{F}_\varphi} \|f - g\|_\infty \leq \varepsilon.$$

Πράγματι, έστω $f, g \in \mathcal{F}_\varphi$. Αφού για κάθε $j \in \{1,2, \dots, n\}$ έχουμε $\left|f(x_j) - \varphi(x_j)\right| < \frac{\varepsilon}{4}$ και $\left|g(x_j) - \varphi(x_j)\right| < \frac{\varepsilon}{4}$, θα ισχύει

$$\left|f(x_j) - g(x_j)\right| \leq \left|f(x_j) - \varphi(x_j)\right| + \left|\varphi(x_j) - g(x_j)\right| < \frac{\varepsilon}{2}, \quad j = 1,2, \dots, n.$$

Για κάθε $x \in X$, υπάρχει $j \in \{1,2, \dots, n\}$ ώστε $x \in U_{x_j}$ και τότε έχουμε

$$|f(x) - g(x)| \leq \left|f(x) - f(x_j)\right| + \left|f(x_j) - g(x_j)\right| + \left|g(x_j) - g(x)\right| < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon.$$

Αφού, λοιπόν, ο ισχυρισμός είναι αληθής, επιλέγοντας μια f από κάθε μη κενό σύνολο $\mathcal{F}_\varphi$, παίρνουμε ένα ε −πυκνό υποσύνολο του $\mathcal{F}$. Αυτό ακριβώς δείχνει ότι το $\mathcal{F}$ είναι ολικά φραγμένο.

Ακόμα, η κλειστότητα ενός ολικά φραγμένου συνόλου είναι ολικά φραγμένο σύνολο και ο $C(X)$ είναι πλήρης, οπότε σύμφωνα με το προηγούμενο λήμμα, το $\overline{\mathcal{F}}$ είναι συμπαγές στον $C(X)$.

■

4. Το Θεώρημα αναπαράστασης του Riesz.

Έστω X ένας τοπολογικός χώρος και $\mathcal{B}_X$ η Borel σ −άλγεβρα του X. Ένα Borel μέτρο πιθανότητας μ στον X λέγεται κανονικό αν και μόνο αν για κάθε $A \in \mathcal{B}_X$ ισχύει:

$$\mu(A) = \sup\{\mu(K) : \quad K \subseteq A \ \text{ συμπαγές}\} = \inf\{\mu(O) : \quad O \supseteq A \ \text{ ανοικτό}\}.$$

Ένα προσημασμένο μέτρο πιθανότητας μ στον X λέγεται κανονικό αν και μόνο αν το $|\mu|$ είναι κανονικό, όπου το μέτρο $|\mu|$ λέγεται ολική κύμανση του μ και ορίζεται ως εξής:

$$|\mu|(E) = \sup \sum_{i=1}^{\infty} |\mu(E_i)|\,, \quad E \in \mathcal{B}_X$$

και το supremum λαμβάνεται πάνω σε όλες τις δυνατές διαμερίσεις $\{E_i\}$ του E με τα E_i σύνολα Borel.

Θεώρημα 17

Έστω X ένας τοπικά συμπαγής χώρος Hausdorff, $V_1, \dots, V_n$ ανοικτά υποσύνολα του X και $K \subseteq X$ συμπαγές τέτοιο ώστε $K \subseteq V_1 \cup \dots \cup V_n$. Τότε, υπάρχουν συναρτήσεις h_i με $h_i \prec V_i$, $i = 1, 2, \dots, n$ και τέτοιες ώστε $h_1(x) + \cdots + h_n(x) = 1, \quad x \in K$.

<u>Απόδειξη:</u>

Κάθε $x \in K$ έχει μια περιοχή W_x με συμπαγή κλειστότητα και τέτοια ώστε $\overline{W_x} \subseteq V_i$ για κάποιο $i = 1, 2, \dots, n$ (το οποίο εξαρτάται από το x). Επειδή το K είναι συμπαγές, υπάρχουν σημεία $x_1, \dots, x_m$ τέτοια ώστε $K \subseteq W_{x_1} \cup \dots \cup W_{x_m}$. Εάν $i \in \{1, 2, \dots, n\}$, έστω

$$H_i = \bigcup \left\{ \overline{W_{x_j}} : \ \overline{W_{x_j}} \subseteq V_i \right\}$$

Από το λήμμα του Urysohn, υπάρχουν συναρτήσεις g_i τέτοιες ώστε $H_i \prec g_i \prec V_i$, $i = 1, 2, \dots n$.

Ορίζουμε:

$$h_1 = g_1$$

$$h_2 = (1 - g_1)g_2$$

$$\dots$$

$$h_n = (1 - g_1)(1 - g_2) \dots (1 - g_{n-1})g_n$$

Τότε, $h_i \prec V_i$ για $i = 1, 2, \dots, n$.

Εύκολα μπορούμε να δείξουμε με επαγωγή ότι

$$h_1 + h_2 + \dots + h_n = 1 - (1 - g_1)(1 - g_2) \dots (1 - g_n).$$

Πράγματι, για $n = 1$: $h_1 = 1 - (1 - g_1) = g_1$ που ισχύει.

Υποθέτουμε ότι για $n = k$: $h_1 + \dots + h_k = 1 - (1 - g_1) \dots (1 - g_k)$

Θα αποδείξουμε ότι για $n = k + 1$, ισχύει:

$$h_1 + \dots + h_k + h_{k+1} = 1 - (1 - g_1) \dots (1 - g_k)(1 - g_{k+1}).$$

Έχουμε:

$$(h_1 + \dots + h_k) + h_{k+1} = 1 - (1 - g_1) \dots (1 - g_k) + h_{k+1}$$

$$= 1 - (1 - g_1) \dots (1 - g_k) + (1 - g_1) \dots (1 - g_k)g_{k+1} = 1 - (1 - g_1) \dots (1 - g_k)(1 - g_{k+1}).$$

Εφόσον, $K \subseteq H_1 \cup \dots \cup H_n$, σε κάθε $x \in K$ ισχύει $g_i(x) = 1$ για τουλάχιστον ένα $i \in \{1,2, \dots, n\}$. Από τη σχέση $h_1 + h_2 + \dots + h_n = 1 - (1 - g_1)(1 - g_2) \dots (1 - g_n)$ έχουμε ότι για κάθε $x \in K$:

$$h_1(x) + \dots + h_n(x) = 1 - \big(1 - g_1(x)\big) \dots \big(1 - g_i(x)\big) \dots \big(1 - g_n(x)\big)$$

δηλαδή $h_1(x) + \dots + h_n(x) = 1$, αφού $1 - g_i(x) = 0$.

■

Θεώρημα 18 (Θεώρημα Riesz για θετικά γραμμικά συναρτησοειδή)

Έστω X ένας τοπικά συμπαγής χώρος Hausdorff και έστω Λ ένα θετικό γραμμικό συναρτησοειδές στο $C_c(X)$. Τότε, υπάρχει μια σ −άλγεβρα $\mathcal{M}$ στο X η οποία περιέχει όλα τα Borel υποσύνολα του X, τέτοια ώστε να υπάρχει ένα μοναδικό θετικό μέτρο μ στην $\mathcal{M}$ το οποίο έχει τις ακόλουθες ιδιότητες:

A) Το μέτρο μ «αναπαριστά» το συναρτησοειδές Λ, υπό την έννοια ότι

$$\Lambda f = \int_X f \, d\mu, \quad f \in C_c(X).$$

B) Για κάθε συμπαγές σύνολο $K \subseteq X$, $\mu(K) < +\infty$.

Γ) Για κάθε $E \in \mathcal{M}$, έχουμε

$$\mu(E) = \inf\{\mu(V) \ : E \subseteq V, V \text{ ανοικτό}\}.$$

Δ) Για κάθε ανοικτό σύνολο E και για κάθε $E \in \mathcal{M}$ με $\mu(E) < +\infty$ ισχύει:

$$\mu(E) = \sup\{\mu(K) \ : K \subseteq E, K \text{ συμπαγές}\}.$$

Ε) Αν $E \in \mathcal{M}$, $A \subseteq E$ και $\mu(E) = 0$, τότε $A \in \mathcal{M}$.

Απόδειξη:

Καθ' όλη τη διάρκεια της απόδειξης, το γράμμα K δηλώνει ένα συμπαγές υποσύνολο του X και το V δηλώνει ένα ανοικτό υποσύνολο του X.

Μοναδικότητα του μέτρου μ.

Έστω $\mathcal{M}$ σ −άλγεβρα που περιέχει τα Borel υποσύνολα του X και μ_1, μ_2 δύο μέτρα στην $\mathcal{M}$ για τα οποία ισχύει το θεώρημα. Εάν ένα μέτρο μ ικανοποιεί τις ιδιότητες Γ και Δ, τότε το μ καθορίζεται στην $\mathcal{M}$ από τις τιμές του πάνω στα συμπαγή σύνολα. Άρα, αρκεί να αποδείξουμε ότι για κάθε $K \subseteq X$, ισχύει $\mu_1(K) = \mu_2(K)$.

Έστω $\varepsilon > 0$. Από τις ιδιότητες Β και Γ, έχουμε ότι $\mu_2(K) < +\infty$ και $\mu_2(K) = \inf\{\mu_2(V) : \ K \subseteq V\}$. Άρα υπάρχει $V \supseteq K$ τέτοιο ώστε $\mu_2(V) < \mu_2(K) + \varepsilon$. Από το λήμμα του Urysohn, υπάρχει $f \in C_c(X)$ τέτοια ώστε $1_K \leq f \leq 1_V$, άρα

$$\mu_1(K) = \int_X 1_K \, d\mu_1 \leq \int_X f \, d\mu_1 = \Lambda f = \int_X f \, d\mu_2 \leq \int_X 1_V \, d\mu_2 = \mu_2(V) < \mu_2(K) + \varepsilon.$$

Συνεπώς, $\mu_1(K) \leq \mu_2(K)$. Ανταλλάσοντας τους ρόλους των μ_1 και μ_2 παίρνουμε την αντίστροφη ανισότητα, οπότε $\mu_1(K) = \mu_2(K)$.

Κατασκευή των μ και $\mathcal{M}$.

Για κάθε ανοικτό $V \subseteq X$, ορίζουμε:

$$\mu(V) \coloneqq \sup\{\Lambda f : \ f \in C_c(X), f \prec V\} \tag{9}$$

Αν $V_1 \subseteq V_2$, τότε η σχέση (9) συνεπάγεται ότι $\mu(V_1) \leq \mu(V_2)$. Άρα, αν E είναι ένα ανοικτό σύνολο, τότε

$$\mu(E) = \inf\{\mu(V) : \ E \subseteq V, V \ \textit{ανοικτό}\} \tag{10}$$

Επομένως, είναι συνεπές με την (9) να ορίσουμε το $\mu(E)$ για κάθε $E \subseteq X$ μέσω της (10). Σημειώνουμε ότι, παρόλο που έχουμε ορίσει το $\mu(E)$ για κάθε $E \subseteq X$, η αριθμήσιμη αθροισιμότητα του μ θα αποδειχθεί μόνο πάνω σε μια συγκεκριμένη σ −άλγεβρα $\mathcal{M}$ στο X.

Ορίζουμε

$$\mathcal{M}_F \coloneqq \{E \subseteq X : \ \mu(E) < +\infty \ \text{ και}$$
$$\mu(E) = \sup\{\mu(K) : \ K \subseteq E, K \text{ συμπαγές}\}\} \tag{11}$$

$$\mathcal{M} \coloneqq \{E \subseteq X : \ E \cap K \in \mathcal{M}_F \text{ για κάθε συμπαγές σύνολο } K\}.$$

<u>Θα αποδείξουμε ότι τα μ και $\mathcal{M}$ έχουν τις απαιτούμενες ιδιότητες.</u>

Το μ είναι μονότονο, δηλαδή $A \subseteq B \implies \mu(A) \leq \mu(B)$ και επίσης, $\mu(E) = 0 \implies E \in \mathcal{M}_F$ και $E \in \mathcal{M}$. Πράγματι,

$$\mu(E) = 0 \implies \mu(E) < +\infty \text{ και } \mu(E) = \sup\{\mu(K) : \ K \subseteq E, K \text{ συμπαγές}\},$$

όπου το supremum λαμβάνεται για $K = \emptyset$. Άρα, $E \in \mathcal{M}_F$.

Επίσης, για κάθε συμπαγές $K \subseteq X$, έχουμε $E \cap K \in \mathcal{M}_F$, αφού $\mu(E \cap K) = 0$. Άρα, $E \in \mathcal{M}$.

Επομένως, ισχύει η ιδιότητα Ε καθώς επίσης και η Γ εξ ορισμού του μ.

Θα χωρίσουμε την απόδειξη των υπόλοιπων ισχυρισμών σε αρκετά βήματα.

Παρατηρούμε ότι αν Λ θετικό, τότε Λ μονότονο, δηλαδή

$$g \geq f \implies g - f \geq 0 \implies \Lambda(g - f) \geq 0 \implies \Lambda g - \Lambda f \geq 0 \implies \Lambda g \geq \Lambda f.$$

Η μονοτονία θα χρησιμοποιηθεί στα βήματα 2 και 10.

<u>Βήμα 1ο:</u> Αν $E_1, E_2, \ldots$ είναι αυθαίρετα υποσύνολα του X, τότε

$$\mu\left(\bigcup_{i=1}^{+\infty} E_i\right) \leq \sum_{i=1}^{+\infty} \mu(E_i) \tag{12}$$

<u>Απόδειξη:</u>

Θα δείξουμε πρώτα ότι αν V_1, V_2 είναι ανοικτά υποσύνολα του X, τότε

$$\mu(V_1 \cup V_2) \leq \mu(V_1) + \mu(V_2). \tag{13}$$

Επιλέγουμε συνάρτηση $g \in C_c(X)$ με $g \prec V_1 \cup V_2$. Εφόσον το $Supp(g)$ είναι συμπαγές και $Supp(g) \subseteq V_1 \cup V_2$, σύμφωνα με το Θεώρημα 16, υπάρχουν συναρτήσεις h_1, h_2 τέτοιες ώστε

$$h_1 \prec V_1, \qquad h_2 \prec V_2 \ \text{ και } \ h_1(x) + h_2(x) = 1, \ x \in Supp(g).$$

Άρα,

$$h_1 g \prec V_1, \qquad h_2 g \prec V_2 \ \text{ και } \ g = (h_1 + h_2)g = h_1 g + h_2 g,$$

οπότε

$$\begin{aligned} \Lambda g &= \Lambda(h_1 g) + \Lambda(h_2 g) \\ &\leq \sup\{\Lambda f : \ f \leq 1_{V_1}\} + \sup\{\Lambda f : \ f \leq 1_{V_2}\} = \mu(V_1) + \mu(V_2). \end{aligned} \tag{14}$$

Εφόσον η (14) ισχύει για κάθε $g \prec V_1 \cup V_2$, θα ισχύει

$$\mu(V_1 \cup V_2) = \sup\{\Lambda g : \ g \prec V_1 \cup V_2\} \leq \mu(V_1) + \mu(V_2)$$

και έτσι η σχέση (13) απεδείχθη.

Αν $\mu(E_i) = +\infty$, για κάποιο $i \in \mathbb{N}$, η σχέση (12) είναι τετριμμένα αληθής.

Φανερά, αρκεί η (12) να αποδειχθεί στην περίπτωση στην οποία $\mu(E_i) < +\infty$, για κάθε $i \in \mathbb{N}$.

Από τον ορισμό του μ, υπάρχουν ανοικτά σύνολα $V_i \supseteq E_i$ τέτοια ώστε

$$\mu(V_i) < \mu(E_i) + \frac{\varepsilon}{2^i}, \qquad i \in \mathbb{N}.$$

Θέτουμε

$$V := \bigcup_{i=1}^{+\infty} V_i$$

και επιλέγουμε $f \in C_c(X)$ με $f \prec V$. Εφόσον

$$Supp(f) \subseteq \bigcup_{i=1}^{+\infty} V_i,$$

θα υπάρχει $n \in \mathbb{N}$ τέτοιος ώστε $Supp(f) \subseteq V_1 \cup \ldots \cup V_n$, δηλαδή $f \prec V_1 \cup \ldots \cup V_n$.

Εφαρμόζοντας επαγωγή στη σχέση (13), παίρνουμε:

$$\begin{aligned} \Lambda f &\leq \sup\{\Lambda g : \ g \prec V_1 \cup \ldots \cup V_n\} = \mu(V_1 \cup \ldots \cup V_n) \leq \mu(V_1) + \cdots + \mu(V_n) \leq \sum_{i=1}^{+\infty} \mu(V_i) \\ &\leq \sum_{i=1}^{+\infty} \left(\mu(E_i) + \frac{\varepsilon}{2^i}\right) = \sum_{i=1}^{+\infty} \mu(E_i) + \varepsilon. \end{aligned}$$

Αφού αυτό ισχύει για κάθε $f \prec V$ και αφού

$$\bigcup_{i=1}^{+\infty} E_i \subseteq V$$

παίρνουμε ότι

$$\mu\left(\bigcup_{i=1}^{+\infty} E_i\right) \le \mu(V) = \sup\{\Lambda g : \ g \prec V\} \le \sum_{i=1}^{+\infty} \mu(E_i) + \varepsilon$$

και επειδή το ε ήταν ένας αυθαίρετος θετικός αριθμός, η σχέση (12) απεδείχθη.

■

Βήμα 2ο: Εάν το Κ είναι συμπαγές, τότε $K \in \mathcal{M}_F$ και

(15) $$\mu(K) = \inf\{\Lambda f : \ K \prec f\}$$

Αυτό συνεπάγεται τον ισχυρισμό Β του θεωρήματος, αφού $K \in \mathcal{M}_F \implies \mu(K) < +\infty$ από τον ορισμό της $\mathcal{M}_F$.

Απόδειξη:

Εάν $K \prec f$ και $\alpha \in (0,1)$, έστω $V_\alpha = \{x \in X : \ f(x) > \alpha\}$. Τότε, για κάθε $\alpha \in (0,1)$, $K \subseteq V_\alpha$ και αν $g \prec V_\alpha$ τότε $\alpha g \le f$.

Πράγματι, αν $x \in V_\alpha$, $g(x) \le 1$ και $\alpha g(x) \le \alpha$ ενώ $f(x) > \alpha$. Άρα, $\alpha g(x) \le f(x)$. Αν $x \in V_\alpha^c$, $g(x) = 0$ και $\alpha g(x) = 0$ και $f(x) \ge 0$. Άρα, $\alpha g(x) \le f(x)$.

Επομένως,

$$\mu(K) \le \mu(V_\alpha) = \sup\{\Lambda g : \ g \prec V_\alpha\} \le \frac{1}{\alpha} \Lambda f.$$

Για $\alpha \longrightarrow 1$, παίρνουμε

(16) $$\mu(K) \le \Lambda f.$$

Άρα, $\mu(K) < +\infty$ και επειδή το K προφανώς ικανοποιεί τη σχέση (11), έχουμε $K \in \mathcal{M}_F$.

Αν $\varepsilon > 0$, από τον ορισμό του μ, υπάρχει $V \supseteq K$ τέτοιο ώστε $\mu(V) < \mu(K) + \varepsilon$. Από το λήμμα Urysohn, υπάρχει $f_0 \in C_c(X)$ τέτοια ώστε $K \prec f_0 \prec V$. Άρα,

$$\Lambda f_0 \le \sup\{\Lambda g : \ g \prec V\} = \mu(V) < \mu(K) + \varepsilon.$$

Άρα, η σχέση (16) ισχύει για κάθε f τέτοια ώστε $K \prec f$ και επίσης, για κάθε $\varepsilon > 0$ υπάρχει f_0 τέτοια ώστε $K \prec f_0$ και $\Lambda f_0 \le \mu(K) + \varepsilon$.

Επομένως, ισχύει η (15), δηλαδή $\mu(K) = \inf\{\Lambda f : \ K \prec f\}$.

■

Βήμα 3ο: Κάθε ανοικτό σύνολο E ικανοποιεί τη σχέση

$$\mu(E) = \sup\{\mu(K) \ : K \subseteq E, K \text{ συμπαγές}\}.$$

Άρα, η $\mathcal{M}_F$ περιέχει κάθε ανοικτό σύνολο E με $\mu(E) < +\infty$.

Απόδειξη:

Έστω ένα ανοικτό σύνολο E και $\alpha \in \mathbb{R}$ τέτοιο ώστε $\alpha < \mu(E)$. Επειδή $\mu(E) = \sup\{\Lambda f : \ f \prec E\}$, υπάρχει $g \prec V$ τέτοια ώστε $\alpha < \Lambda g$. Εάν W είναι ένα οποιοδήποτε ανοικτό σύνολο με $W \supseteq Supp(g)$, τότε $g \prec W$ και άρα $\Lambda g \leq \sup\{\Lambda f : \ f \prec W\} = \mu(W)$.

Άρα, συμβολίζοντας $Supp(g) = K$ συμπαγές, έχουμε

$$\Lambda g \leq \inf\{\mu(W) : \ W \supseteq K, W \text{ ανοικτό}\} = \mu(K),$$

όπου η τελευταία ισότητα ισχύει από τον ορισμό του μ.

Συνεπώς, έχουμε βρει συμπαγές $K \subseteq E$ τέτοιο ώστε $\alpha < \mu(K)$ και έτσι ισχύει ο ισχυρισμός Δ, δηλαδή $\mu(E) = \sup\{\mu(K) \ : K \subseteq E, K \text{ συμπαγές}\}$.

■

Βήμα 4ο: Έστω

$$E = \bigcup_{i=1}^{+\infty} E_i,$$

όπου $E_1, E_2, \ldots$ είναι ξένα ανά δύο και $E_i \in \mathcal{M}_F$, $i = 1, 2, \ldots$ Τότε,

$$\mu(E) = \sum_{i=1}^{+\infty} \mu(E_i)\,. \tag{17}$$

Εάν, επιπλέον, $\mu(E) < +\infty$, τότε $E \in \mathcal{M}_F$.

Απόδειξη:

Θα δείξουμε πρώτα ότι

$$\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2), \tag{18}$$

όπου K_1, K_2 είναι ξένα συμπαγή σύνολα.

Επιλέγουμε $\varepsilon > 0$. Από το λήμμα του Urysohn, υπάρχει $f \in C_c(X)$ τέτοια ώστε $K_1 \prec f \prec K_2$. Από το 2ο βήμα, έχουμε $\mu(K_1 \cup K_2) = \inf\{\Lambda g : \ K_1 \cup K_2 \prec g\}$, άρα υπάρχει συνάρτηση g τέτοια ώστε $K_1 \cup K_2 \prec g$ και $\Lambda g < \mu(K_1 \cup K_2) + \varepsilon$.

Παρατηρούμε ότι $K_1 \prec fg$ και $K_2 \prec (1-f)g$. Εφόσον το Λ είναι γραμμικό, έπεται από την (16) ότι $\mu(K_1) + \mu(K_2) \leq \Lambda(fg) + \Lambda(g - fg) = \Lambda g < \mu(K_1 \cup K_2) + \varepsilon$.

Εφόσον το ε ήταν αυθαίρετος θετικός αριθμός, ισχύει $\mu(K_1) + \mu(K_2) \leq \mu(K_1 \cup K_2)$.

Η αντίστροφη ανισότητα $\mu(K_1 \cup K_2) \leq \mu(K_1) + \mu(K_2)$ έχει αποδειχθεί στο 1ο βήμα, συνεπώς έπεται η (18).

Εάν $\mu(E) = +\infty$, η σχέση (17)

$$\mu(E) = \sum_{i=1}^{+\infty} \mu(E_i)$$

έπεται από το 1ο βήμα, καθώς

$$\mu(E) \leq \sum_{i=1}^{+\infty} \mu(E_i) \quad \text{και} \quad +\infty = \mu(E) \geq \sum_{i=1}^{+\infty} \mu(E_i).$$

Επομένως, υποθέτουμε ότι $\mu(E) < +\infty$. Έστω $\varepsilon > 0$. Εφόσον $E_i \in \mathcal{M}_F$, $i = 1, 2, \ldots$ ικανοποιούνται οι σχέσεις

$$\mu(E_i) = \sup\{\mu(K): \ K \subseteq E_i, K \text{ συμπαγές}\}, \qquad i = 1,2, \ldots.$$

Άρα, υπάρχουν συμπαγή σύνολα $H_i \subseteq E_i$ τέτοια ώστε

$$\mu(H_i) > \mu(E_i) - \frac{\varepsilon}{2^i}, \quad i = 1, 2, \ldots \tag{19}$$

Θέτουμε $K_n = H_1 \cup \ldots \cup H_n$, $n \in \mathbb{N}$ και χρησιμοποιώντας επαγωγή στη σχέση (18), παίρνουμε (τα H_ι είναι ξένα ανά δύο):

$$\mu(E) \geq \mu(K_n) = \sum_{i=1}^{n} \mu(H_i) > \sum_{i=1}^{n} \left(\mu(E_i) - \frac{\varepsilon}{2^i}\right) > \sum_{i=1}^{n} \mu(E_i) - \varepsilon. \tag{20}$$

Αφού η (20) ισχύει για κάθε $n \in \mathbb{N}$ και για αυθαίρετο $\varepsilon > 0$, έχουμε:

$$\mu(E) \geq \sum_{i=1}^{+\infty} \mu(E_i).$$

Η αντίστροφη ανισότητα έχει αποδειχθεί στο 1ο βήμα, άρα

$$\mu(E) = \sum_{i=1}^{+\infty} \mu(E_i).$$

Επιπλέον, αν $\mu(E) < +\infty$ και $\varepsilon > 0$, από τη σχέση

$$\mu(E) = \sum_{i=1}^{+\infty} \mu(E_i)$$

έχουμε ότι

$$\mu(E) - \sum_{i=1}^{N} \mu(E_i) < \varepsilon,$$

άρα,

$$\mu(E) < \sum_{i=1}^{N} \mu(E_i) + \varepsilon\,, \quad \text{για κάποιο } N \in \mathbb{N}. \tag{21}$$

Όμως, από τη σχέση (19), έχουμε:

$$\mu(E) < \sum_{i=1}^{N} \mu(E_i) + \varepsilon < \sum_{i=1}^{N} \left(\mu(H_i) + \frac{\varepsilon}{2^i}\right) + \varepsilon < \sum_{i=1}^{N} \mu(H_i) + \varepsilon + \varepsilon = \mu(K_N) + 2\varepsilon.$$

Αυτό δείχνει ότι το E ικανοποιεί τη σχέση (11):

$$\mu(E) = \sup\{\mu(K):\ \ K \subseteq E, K \text{ συμπαγές}\}$$

(αφού K_N συμπαγές) και συνεπώς $E \in \mathcal{M}_F$.

■

<u>Βήμα 5ο:</u> Εάν $E \in \mathcal{M}_F$ και $\varepsilon > 0$, υπάρχει συμπαγές K και ανοικτό V έτσι ώστε $K \subseteq E \subseteq V$ και $\mu(V \setminus K) < \varepsilon$.

<u>Απόδειξη:</u>

Από τον ορισμό, έχουμε $\mu(E) = \inf\{\mu(V):\ \ V \supseteq E, V \text{ ανοικτό}\}$.

Επειδή $E \in \mathcal{M}_F$, έχουμε $\mu(E) = \sup\{\mu(K):\ \ K \subseteq E, K \text{ συμπαγές}\}$.

Συνεπώς, υπάρχουν K, V τέτοια ώστε $K \subseteq E, V \supseteq E$ έτσι ώστε

$$\mu(K) > \mu(E) - \frac{\varepsilon}{2} \qquad \text{και} \qquad \mu(V) < \mu(E) + \frac{\varepsilon}{2},$$

και άρα,

$$\mu(V) - \frac{\varepsilon}{2} < \mu(E) < \mu(K) + \frac{\varepsilon}{2}.$$

Το σύνολο $V \setminus K$ είναι ανοικτό, επομένως από το 3ο βήμα, $V \setminus K \in \mathcal{M}_F$.

Άρα, από το 4ο βήμα: $\mu(K) + \mu(V \setminus K) = \mu(V) < \mu(K) + \varepsilon$ δηλαδή, $\mu(V \setminus K) < \varepsilon$.

■

<u>Βήμα 6ο :</u> Εάν $A, B \in \mathcal{M}_F$, τότε τα σύνολα $A \setminus B$, $A \cup B$ και $A \cap B$ ανήκουν στην $\mathcal{M}_F$.

Απόδειξη:

Έστω $\varepsilon > 0$. Από το 5ο βήμα, γνωρίζουμε ότι υπάρχουν σύνολα K_i, V_i, $i = 1,2$ τέτοια ώστε $K_1 \subseteq A \subseteq V_1$, $K_2 \subseteq B \subseteq V_2$ και $\mu(V_i \setminus K_i) < \varepsilon$, $i = 1, 2$.

Εφόσον,

$$A \setminus B \subseteq V_1 \setminus K_2 \subseteq (V_1 \setminus K_1) \cup (K_1 \setminus V_2) \cup (V_2 \setminus K_2),$$

το 1ο βήμα μας δείχνει ότι:

$$\begin{aligned} \mu(A \setminus B) \leq \mu(V_1 \setminus K_1) + \mu(K_1 \setminus V_2) + \mu(V_2 \setminus K_2) \leq \varepsilon + \mu(K_1 \setminus V_2) + \varepsilon \\ = \mu(K_1 \setminus V_2) + 2\varepsilon. \end{aligned} \tag{22}$$

Το $K_1 \setminus V_2$ είναι ένα συμπαγές υποσύνολο του $A \setminus B$, οπότε η σχέση (22) δείχνει ότι το $A \setminus B$ ικανοποιεί τη σχέση (11): $\mu(E) = \sup\{\mu(K) : \ K \subseteq E, K \text{ συμπαγές}\}$, συνεπώς $A \setminus B \in \mathcal{M}_F$.

Εφόσον, $A \cup B = (A \setminus B) \cup B$ και $\mu(A \cup B) < +\infty$, το 4ο βήμα μας δείχνει ότι $A \cup B \in \mathcal{M}_F$.

Εφόσον, $A \cap B = A \setminus (A \setminus B)$, έχουμε επίσης ότι $A \cap B \in \mathcal{M}_F$.

■

Βήμα 7ο: Η $\mathcal{M}$ είναι μια σ −άλγεβρα στο X, η οποία περιέχει όλα τα Borel σύνολα.

Απόδειξη:

Έστω K ένα αυθαίρετο συμπαγές υποσύνολο του X. Εάν $A \in \mathcal{M}$, τότε $A^c \cap K = K \setminus (A \cap K)$ και από το 2ο βήμα $K \in \mathcal{M}_F$, δηλαδή το $A^c \cap K$ είναι η διαφορά δύο στοιχείων της $\mathcal{M}_F$. Άρα, $A^c \cap K \in \mathcal{M}_F$, οπότε $A^c \in \mathcal{M}$. Μόλις αποδείξαμε ότι $A \in \mathcal{M} \implies A^c \in \mathcal{M}$.

Στη συνέχεια, υποθέτουμε ότι

$$A = \bigcup_{i=1}^{+\infty} A_i \ , \quad \text{όπου } A_i \in \mathcal{M}, \qquad i = 1, 2, \dots$$

Θέτουμε $B_1 = A_1 \cap K$ και

$$B_n = (A_n \cap K) \setminus (B_1 \cup \dots \cup B_{n-1}), \quad n = 2, 3, \dots \tag{23}$$

Τότε, η $\{B_n\}_{n=1}^{+\infty}$ είναι μια ακολουθία ξένων ανά δύο συνόλων όπου, λόγω του 6ου βήματος, $B_i \in \mathcal{M}_F$ για κάθε $i = 1, 2, \dots$

Επίσης,

$$A \cap K = \bigcup_{n=1}^{+\infty} B_n.$$

Από το 4ο βήμα έπεται ότι $A \cap K \in \mathcal{M}_F$, οπότε $A \in \mathcal{M}$. Μόλις αποδείξαμε ότι:

$$A_i \in \mathcal{M}\ ,\ \ i = 1, 2, \dots \ \Longrightarrow\ A \coloneqq \bigcup_{i=1}^{+\infty} A_i \in \mathcal{M}.$$

Τέλος, αν C είναι ένα κλειστό υποσύνολο του X, το $C \cap K$ είναι συμπαγές (ως κλειστό υποσύνολο συμπαγούς συνόλου), άρα $C \cap K \in \mathcal{M}_F$ και συνεπώς, $C \in \mathcal{M}$. Ειδικότερα, $X \in \mathcal{M}$ αφού όλος ο χώρος είναι κλειστό σύνολο.

Έχουμε, λοιπόν, αποδείξει ότι η κλάση $\mathcal{M}$ είναι μια σ −άλγεβρα, η οποία περιέχει όλα τα κλειστά υποσύνολα του X. Άρα, η $\mathcal{M}$ θα περιέχει και όλα τα Borel υποσύνολα του X.

■

<u>Βήμα 8ο:</u> Η $\mathcal{M}_F$ αποτελείται ακριβώς από εκείνα τα σύνολα $E \in \mathcal{M}$ για τα οποία $\mu(E) < +\infty$.

Αυτό συνεπάγεται τον ισχυρισμό Δ του θεωρήματος.

<u>Απόδειξη:</u>

Έστω $E \in \mathcal{M}_F$. Τα βήματα 2 και 6 συνεπάγονται ότι $E \cap K \in \mathcal{M}_F$ για κάθε συμπαγές $K \subseteq X$, άρα $E \in \mathcal{M}$. Δηλαδή, $\mathcal{M}_F \subseteq \mathcal{M}$.

Αντίστροφα, έστω $E \in \mathcal{M}$ με $\mu(E) < +\infty$ και έστω $\varepsilon > 0$. Λόγω της σχέσης (10), υπάρχει ανοικτό σύνολο $V \supseteq E$ με $\mu(V) < +\infty$. Από τα βήματα 2 και 5, υπάρχει συμπαγές $K \subseteq V$ με $\mu(V - K) < \varepsilon$.

Εφόσον $E \cap K \in \mathcal{M}_F$, από τον ορισμό της $\mathcal{M}$, υπάρχει συμπαγές σύνολο $H \subseteq E \cap K$ έτσι ώστε $\mu(E \cap K) < \mu(H) + \varepsilon$.

Εφόσον $E \subseteq (E \cap K) \cup (V - K)$, έπεται ότι $\mu(E) \leq \mu(E \cap K) + \mu(V \setminus K) < \mu(H) + 2\varepsilon$.

Άρα, για τυχαίο $\varepsilon > 0$, βρήκαμε συμπαγές $H \subseteq E$ τέτοιο ώστε $\mu(E) \leq \mu(H) + \varepsilon$. Αυτό σημαίνει ότι $\mu(E) = \sup\{\mu(H){:}\ H \subseteq E, H \text{ συμπαγές}\}$, δηλαδή το E ικανοποιεί τη σχέση (11), άρα $E \in \mathcal{M}_F$.

■

<u>Βήμα 9ο:</u> Το μ είναι μέτρο στην $\mathcal{M}$.

<u>Απόδειξη:</u>

Η αριθμήσιμη αθροισιμότητα του μ στην $\mathcal{M}$ έπεται αμέσως από τα βήματα 4 και 8. Πράγματι, αν $\{A_i\}_{i=1}^{+\infty}$ είναι μια ακολουθία ξένων ανά δύο στοιχείων της $\mathcal{M}$ (όπου $\mu(A_i) < +\infty$ για κάθε $i \in \mathbb{N}$, αλλιώς δεν έχουμε τίποτα να αποδείξουμε) και

$$A = \bigcup_{i=1}^{+\infty} A_i\ ,$$

τότε από το 8ο βήμα: $A_i \in \mathcal{M}_F,\ i = 1, 2, \dots$ και από το 4ο βήμα:

$$\mu(A) = \sum_{i=1}^{+\infty} \mu(A_i).$$

■

Βήμα 10ο: Για κάθε $f \in C_c(X)$, ισχύει

$$\Lambda f = \int_X f \, d\mu.$$

Αυτό αποδεικνύει τον ισχυρισμό Α και ολοκληρώνει την απόδειξη του θεωρήματος.

Απόδειξη:

Αρκεί να αποδείξουμε ότι για κάθε $f \in C_c(X)$, ισχύει η ανισότητα

$$\Lambda f \leq \int_X f \, d\mu$$

καθώς, εφόσον αποδειχθεί, η γραμμικότητα του Λ θα μας δώσει την αντίστροφη ανισότητα:

$$-\Lambda f = \Lambda(-f) \leq \int_X (-f) \, d\mu = -\int_X f \, d\mu$$

οπότε

$$\Lambda f \geq \int_X f \, d\mu.$$

Έστω, λοιπόν, μια $f \in C_c(X)$ και $K \coloneqq supp(f)$. Θεωρούμε, ακόμα, ότι $\forall x \in X \;\; f(x) \in [a, b]$, δηλαδή το διάστημα $[a, b]$ περιέχει το σύνολο τιμών της f.

Έστω $\varepsilon > 0$. Επιλέγουμε κατάλληλο $n \in \mathbb{N}$ και στη συνέχεια για $i = 0, 1, 2, \ldots, n$ επιλέγουμε $y_i \in \mathbb{R}$ έτσι ώστε $y_i - y_{i-1} < \varepsilon$ και $y_0 < a < y_1 < \cdots < y_n = b$.

Θέτουμε $E_i = \{x \in X : \;\; y_{i-1} < f(x) \leq y_i\} \cap K, \quad i = 1, 2, \ldots, n$

Εφόσον η f είναι συνεχής, είναι και Borel μετρήσιμη, συνεπώς τα σύνολα E_i, $i = 1, 2, \ldots, n$ είναι Borel, ξένα ανά δύο και ισχύει

$$\bigcup_{i=1}^{n} E_i = K.$$

Λόγω της σχέσης (10), υπάρχουν ανοικτά σύνολα $V_i \supseteq E_i$ τέτοια ώστε

$$\mu(V_i) < \mu(E_i) + \frac{\varepsilon}{n}, \quad i = 1, 2, \ldots, n$$

και επίσης, $y_{i-1} < f(x) < y_i + \varepsilon$, για κάθε $x \in V_i$.

Από το Θεώρημα 17, αφού $K \subseteq V_1 \cup \ldots \cup V_n$, υπάρχουν συναρτήσεις $h_i \prec V_i$ τέτοιες ώστε

$$\sum_{i=1}^{n} h_i(x) = 1, \quad x \in K.$$

Άρα,

$$f(x) = \sum_{i=1}^{n} h_i(x) \cdot f(x) \quad \text{και} \quad K \prec \sum_{i=1}^{n} h_i\,,$$

οπότε από το 2ο βήμα, έχουμε

$$\mu(K) = \inf\{\Lambda g : \ K \prec g\} \leq \Lambda\left(\sum_{i=1}^{n} h_i\right) = \sum_{i=1}^{n} \Lambda h_i.$$

Αφού $h_i f \leq (y_i + \varepsilon)h_i$ και $y_i - \varepsilon < f(x)$ για κάθε $x \in V_i$, άρα $y_i - \varepsilon \leq m_i$, όπου $m_i = \inf_{x \in E_i} f(x)$, έχουμε:

$$\begin{aligned}
\Lambda f &= \Lambda\left(\sum_{i=1}^{n} h_i f\right) = \sum_{i=1}^{n} \Lambda(h_i f) \leq \sum_{i=1}^{n} (y_i + \varepsilon)\Lambda h_i = \sum_{i=1}^{n} (|a| + y_i + \varepsilon)\Lambda h_i - |a| \sum_{i=1}^{n} \Lambda h_i \\
&\leq \sum_{i=1}^{n} (|a| + y_i + \varepsilon)\mu(V_i) - |a|\mu(K) \quad (\text{από την (9) και το 2ο βήμα}) \\
&\leq \sum_{i=1}^{n} (|a| + y_i + \varepsilon)\left[\mu(E_i) + \frac{\varepsilon}{n}\right] - |a|\mu(K) \\
&= \frac{\varepsilon}{n}\sum_{i=1}^{n} (|a| + y_i + \varepsilon) + \sum_{i=1}^{n} (|a| + y_i + \varepsilon)\mu(E_i) - |a|\mu(K) \\
&= \frac{\varepsilon}{n}\sum_{i=1}^{n} (|a| + y_i + \varepsilon) + |a|\sum_{i=1}^{n} \mu(E_i) + \sum_{i=1}^{n} (y_i + \varepsilon)\mu(E_i) - |a|\mu(K) \\
&= \frac{\varepsilon}{n}\sum_{i=1}^{n} (|a| + y_i + \varepsilon) + \sum_{i=1}^{n} (y_i - \varepsilon + 2\varepsilon)\mu(E_i) \\
&\leq \frac{\varepsilon}{n}\sum_{i=1}^{n} (|a| + y_i + \varepsilon) + \sum_{i=1}^{n} (y_i - \varepsilon)\mu(E_i) + 2\varepsilon\sum_{i=1}^{n} \mu(E_i) \\
&\leq \frac{\varepsilon}{n} n(|a| + b + \varepsilon) + \sum_{i=1}^{n} m_i\mu(E_i) + 2\varepsilon\mu(K) \leq \varepsilon(|a| + b + \varepsilon) + 2\varepsilon\mu(K) + \int_K f\, d\mu \\
&= \varepsilon\big(|a| + b + \varepsilon + 2\mu(K)\big) + \int_X f\, d\mu \qquad (\text{αφού } f(x) = 0,\ \ \forall x \in K^c)
\end{aligned}$$

Αφού, όμως, το ε ήταν ένας αυθαίρετος θετικός αριθμός, έχουμε

$$\Lambda f \leq \int_X f\, d\mu$$

και έτσι η απόδειξη του θεωρήματος είναι πλήρης.

■

Το θεώρημα που μόλις αποδείξαμε χαρακτηρίζει τα θετικά γραμμικά συναρτησοειδή στον $C_c(X)$. Με τη βοήθεια των παρακάτω τριών θεωρημάτων, για την απόδειξη των οποίων παραπέμπουμε στο [14] (Θεωρήματα 6.12, 6.13 και 6.16), θα καταφέρουμε να χαρακτηρίσουμε τα φραγμένα γραμμικά συναρτησοειδή στον $C_c(X)$.

Εφόσον ο $C_c(X)$ είναι ένας πυκνός υπόχωρος του $C_0(X)$ αναφορικά με την supremum νόρμα, κάθε φραγμένο γραμμικό συναρτησοειδές Φ στον $C_c(X)$ επεκτείνεται με μοναδικό τρόπο σε ένα φραγμένο γραμμικό συναρτησοειδές στον $C_0(X)$. Συνεπώς, μπορούμε να δουλέψουμε με τον χώρο $C_0(X)$.

Θεώρημα 19

Έστω μ ένα προσημασμένο μέτρο ορισμένο σε μια $\sigma-$άλγεβρα $\mathcal{M}$ στον X. Τότε υπάρχει μετρήσιμη συνάρτηση h στον X με $|h(x)| = 1$ για κάθε $x \in X$ και τέτοια ώστε $d\mu = h\, d|\mu|$.

Θεώρημα 20

Έστω μ ένα μέτρο στην $\sigma-$άλγεβρα $\mathcal{M}$ του X, $g \in L^1(d\mu)$ και

$$\lambda(E) = \int_E g\, d\mu, \qquad E \in \mathcal{M}.$$

Τότε,

$$|\lambda|(E) = \int_E |g|\, d\mu, \qquad E \in \mathcal{M}.$$

Θεώρημα 21

Έστω $1 \leq p < +\infty$, μ ένα $\sigma-$πεπερασμένο μέτρο στον X και Φ ένα φραγμένο γραμμικό συναρτησοειδές στον $L^p(d\mu)$. Τότε, υπάρχει μοναδική συνάρτηση $g \in L^q(d\mu)$, όπου q είναι ο συζυγής εκθέτης του p, τέτοια ώστε

$$\Phi(f) = \int_X fg\, d\mu, \qquad f \in L^p(d\mu).$$

Επιπλέον, $\|\Phi\| = \|g\|_{L^q(d\mu)}$.

Θεώρημα 22 (Θεώρημα Riesz για φραγμένα γραμμικά συναρτησοειδή)

Έστω X ένας τοπικά συμπαγής χώρος Hausdorff. Τότε, ο χώρος $M(X)$ των κανονικών προσημασμένων μέτρων Borel που έχουν πεπερασμένη ολική κύμανση ταυτίζεται με τον τοπολογικό δυϊκό χώρο του $C_0(X)$.

Συγκεκριμένα: αν Φ είναι ένα φραγμένο γραμμικό συναρτησοειδές στον $C_0(X)$, τότε αυτό αναπαρίσταται από ένα μοναδικό κανονικό προσημασμένο μέτρο Borel μ στον X, υπό την έννοια ότι

$$\Phi(f) = \int_X f\, d\mu, \qquad f \in C_0(X). \tag{24}$$

Επιπλέον, η νόρμα του Φ είναι η ολική κύμανση του μ: $\|\Phi\| = |\mu|(X)$.

<u>Απόδειξη:</u>

Όσον αφορά τη μοναδικότητα, έστω μ ένα κανονικό προσημασμένο μέτρο Borel στον X τέτοιο ώστε

$$\int_X f\, d\mu = 0, \qquad f \in C_0(X).$$

Σύμφωνα με το Θεώρημα 19, υπάρχει Borel μετρήσιμη συνάρτηση h, με $|h| = 1$, τέτοια ώστε $d\mu = h\, d|\mu|$. Για οποιαδήποτε ακολουθία $\{f_n\}$ στο $C_0(X)$ έχουμε

$$|\mu|(X) = \int_X (h - f_n) h\, d|\mu| \leq \int_X |h - f_n|\, d|\mu|,$$

και αφού ο $C_c(X)$ είναι πυκνός στον $L^1(d|\mu|)$ ([14], Θεώρημα 3.14), η ακολουθία $\{f_n\}$ μπορεί να επιλεγεί έτσι ώστε

$$\int_X |h - f_n|\, d|\mu| \longrightarrow 0, \quad \text{καθώς } n \to \infty.$$

Άρα, $|\mu|(X) = 0$ και $\mu = 0$.

Ακόμα, η διαφορά δύο κανονικών προσημασμένων μέτρων Borel στον X είναι, επίσης, κανονικό προσημασμένο μέτρο. Επομένως, αν μ_1, μ_2 είναι κανονικά προσημασμένα μέτρα Borel στον X με

$$\int_X f\, d\mu_1 = \int_X f\, d\mu_2, \qquad f \in C_0(X)$$

τότε

$$\int_X f\, d(\mu_1 - \mu_2) = 0, \qquad f \in C_0(X)$$

άρα $\mu_1 - \mu_2 = 0$, άρα $\mu_1 = \mu_2$.

Στη συνέχεια, θεωρούμε ένα φραγμένο γραμμικό συναρτησοειδές Φ στον $C_0(X)$. Χωρίς βλάβη της γενικότητας, υποθέουμε ότι $\|\Phi\| = 1$. Θα κατασκευάσουμε ένα θετικό γραμμικό συναρτησοειδές Λ στον $C_c(X)$ τέτοιο ώστε

$$|\Phi(f)| \leq \Lambda(|f|) \leq \|f\|_\infty, \qquad f \in C_c(X). \tag{25}$$

Τότε, σύμφωνα με το θεώρημα Riesz για θετικά γραμμικά συναρτησοειδή, θα υπάρχει μέτρο Borel λ στον X τέτοιο ώστε

$$\Lambda(f) = \int_X f\, d\lambda, \qquad f \in C_c(X).$$

Ακόμα, από τα συμπεράσματα Γ και Δ του Θεωρήματος 18, προκύπτει ότι το λ είναι κανονικό αν $\lambda(X) < +\infty$. Πράγματι $\lambda(X) < +\infty$, αφού λόγω της (9)

$$\lambda(X) = \sup\{\Lambda(f) : \quad 0 \leq f \leq 1, f \in C_c(X)\}$$

και $|\Lambda(f)| \leq 1$ αν $\|f\|_\infty \leq 1$, οπότε στην πραγματικότητα έχουμε $\lambda(X) \leq 1$.

Από τη σχέση (25), έχουμε ακόμα

$$|\Phi(f)| \leq \Lambda(|f|) = \int_X |f|\, d\lambda = \|f\|_{L^1(d\lambda)}, \qquad f \in C_c(X).$$

Άρα, το Φ είναι ένα γραμμικό συναρτησοειδές στον $C_c(X)$ με νόρμα το πολύ 1, αναφορικά με την $L^1(d\lambda)$ −νόρμα στον $C_c(X)$. Το Φ, λοιπόν, επεκτείνεται σε ένα γραμμικό συναρτησοειδές στον $L^1(d\lambda)$ που διατηρεί τη νόρμα.

Από το Θεώρημα 21 (περίπτωση $p = 1$), υπάρχει Borel μετρήσιμη συνάρτηση g, με $|g| \leq 1$, τέτοια ώστε

$$\Phi(f) = \int_X fg\, d\lambda, \qquad f \in C_c(X). \tag{26}$$

Θα δείξουμε ότι το μέτρο $\mu = gd\lambda$ είναι κανονικό. Από το Θεώρημα 20, $|\mu| = |gd\lambda| = |g|d\lambda$. Έστω $E \in \mathcal{B}_X$. Αφού το λ είναι κανονικό, υπάρχει φθίνουσα ακολουθία ανοικτών συνόλων $O_n \supseteq E$ τέτοια ώστε $\lambda(O_n) \downarrow \lambda(E)$. Ορίζουμε

$$A = \bigcap_{n=1}^{\infty} O_n.$$

Τότε,

$$|\mu|(O_n) = \int_{O_n} |g| d\lambda \downarrow \int_A |g| d\lambda = |\mu|(A).$$

Από την άλλη, για κάθε $n \in \mathbb{N}$, $\lambda(O_n) \geq \lambda(A) \geq \lambda(E)$ και αφού $\lambda(O_n) \downarrow \lambda(E)$, έχουμε $\lambda(A) = \lambda(E)$ και $\lambda(A \setminus E) = 0$. Άρα,

$$|\mu|(A) = |\mu|(E) + |\mu|(A \setminus E) = \int_E |g| d\lambda + \int_{A \setminus E} |g| d\lambda = \int_E |g| d\lambda = |\mu|(E).$$

Τελικά, $|\mu|(O_n) \downarrow |\mu|(E)$. Αντίστοιχα, μπορούμε να δείξουμε ότι υπάρχει ακολουθία συμπαγών συνόλων K_n τέτοια ώστε $|\mu|(K_n) \uparrow |\mu|(E)$. Έτσι, λοιπόν, το $g d\lambda$ είναι κανονικό.

Κάθε μέλος της σχέσης (26) είναι ένα συνεχές συναρτησοειδές στον $C_c(X)$ και ο $C_c(X)$ είναι πυκνός στον $C_0(X)$. Επομένως, η (26) ισχύει για κάθε $f \in C_0(X)$, και έτσι αποδεικνύουμε τη σχέση (24) με $d\mu = g d\lambda$.

Αφού $\|\Phi\| = 1$, από την (26) έχουμε

$$\int_X |g|\, d\lambda \geq \sup\{|\Phi(f)| : \quad f \in C_0(X), \|f\|_\infty \leq 1\} = 1.$$

Γνωρίζουμε, ακόμα, ότι $\lambda(X) \leq 1$ και $|g| \leq 1$. Συνεπώς, $\lambda(X) = 1$ και $|g| = 1$, $d\lambda$ −σχεδόν παντού. Άρα, από το Θεώρημα 20, $d|\mu| = |g| d\lambda = d\lambda$ και $|\mu|(X) = \lambda(X) = 1 = \|\Phi\|$

και έτσι ολοκληρώνεται η απόδειξη του θεωρήματος.

Απομένει, λοιπόν, να κατασκευάσουμε το θετικό γραμμικό συναρτησοειδές Λ που θα ικανοποιεί την (25). Συμβολίζουμε με $C_c^+(X)$ την κλάση των μη αρνητικών συναρτήσεων του $C_c(X)$.

Αν $f \in C_c^+(X)$, ορίζουμε

$$\Lambda(f) = \sup\{|\Phi(h)| : \quad h \in C_c(X), |h| \leq f\}.$$

Τότε, $\Lambda(f) \geq 0$, το Λ ικανοποιεί τη σχέση (25) για $f \in C_c^+(X)$ και η ανισότητα $0 \leq f_1 \leq f_2$ συνεπάγεται $\Lambda(f_1) \leq \Lambda(f_2)$. Επίσης, αν $c > 0$, τότε $\Lambda(cf) = c\Lambda(f)$ για $f \in C_c^+(X)$.

Πρέπει να δείξουμε ότι $\Lambda(f + g) = \Lambda(f) + \Lambda(g)$, $f, g \in C_c^+(X)$ και στη συνέχεια, να επεκτείνουμε το Λ σε ένα γραμμικό συναρτησοειδές στον $C_c(X)$.

Σταθεροποιούμε $f, g \in C_c^+(X)$. Αν $\varepsilon > 0$, υπάρχουν $h_1, h_2 \in C_c(X)$ τέτοιες ώστε $|h_1| \leq f$, $|h_2| \leq g$ και

$$\Lambda(f) \leq |\Phi(h_1)| + \varepsilon, \qquad \Lambda(g) \leq |\Phi(h_2)| + \varepsilon.$$

Υπάρχουν αριθμοί a_i, $|a_i| = 1$, έτσι ώστε $a_i \Phi(h_i) = |\Phi(h_i)|$, $i = 1,2$. Τότε

$$\begin{aligned} \Lambda(f) + \Lambda(g) &\leq |\Phi(h_1)| + |\Phi(h_2)| + 2\varepsilon = a_1\Phi(h_1) + a_2\Phi(h_2) = \Phi(a_1 h_1 + a_2 h_2) + 2\varepsilon \\ &\leq \Lambda(|h_1| + |h_2|) + 2\varepsilon \leq \Lambda(f + g) + 2\varepsilon. \end{aligned}$$

Έτσι, λοιπόν, έχουμε $\Lambda(f + g) \geq \Lambda(f) + \Lambda(g)$.

Στη συνέχεια, επιλέγουμε $h \in C_c(X)$ με $|h| \leq f + g$ και έστω $V = \{x : \quad f(x) + g(x) > 0\}$. Ορίζουμε

$$h_1(x) = \frac{f(x)h(x)}{f(x)+g(x)}, \qquad h_2(x) = \frac{g(x)h(x)}{f(x)+g(x)}, \qquad x \in V$$

$$h_1(x) = h_2(x) = 0, \qquad x \notin V.$$

Η h_1 είναι συνεχής σε κάθε σημείο του V. Αν $x_0 \notin V$, τότε $h(x_0) = 0$. Αφού η h είναι συνεχής και $|h_1(x)| \leq |h(x)|,\ \ x \in X,$ η h_1 είναι συνεχής στο x_0. Άρα, $h_1 \in C_c(X)$ και ομοίως, $h_2 \in C_c(X)$.

Αφού $h_1 + h_2 = h$ και $|h_1| \leq f,\ |h_2| \leq g,$ έχουμε

$$|\Phi(h)| = |\Phi(h_1) + \Phi(h_2)| \leq |\Phi(h_1)| + |\Phi(h_2)| \leq \Lambda(f) + \Lambda(g)$$

Άρα, $\Lambda(f+g) \leq \Lambda(f) + \Lambda(g)$ και έτσι έχουμε αποδείξει την ισότητα $\Lambda(f+g) = \Lambda(f) + \Lambda(g)$.

Αν, τώρα, $f \in C_c(X)$, τότε $f = f^+ - f^-$ και $|f| = f^+ + f^-$, οπότε $f^+ = \frac{1}{2}(|f| + f)$ και $f^- = \frac{1}{2}(|f| - f)$, άρα $f^+, f^- \in C_c^+(X)$.

Για να επεκτείνουμε το Λ στον $C_c(X)$, ορίζουμε: $\Lambda(f) = \Lambda(f^+) - \Lambda(f^-), \quad f \in C_c(X)$.

Αποδεικνύουμε ότι η επέκταση του Λ είναι γραμμικό συναρτησοειδές στον $C_c(X)$.

Αν $f, g \in C_c(X)$, τότε θέτουμε $h = f + g$ και έχουμε $h^+ - h^- = f^+ - f^- + g^+ - g^-$ ή αλλιώς $h^+ + f^- + g^- = f^+ + g^+ + h^-$, άρα

$$\Lambda(h^+ + f^- + g^-) = \Lambda(f^+ + g^+ + h^-)$$

ισοδύναμα, $\Lambda(h^+) + \Lambda(f^-) + \Lambda(g^-) = \Lambda(f^+) + \Lambda(g^+) + \Lambda(h^-)$

ισοδύναμα, $\Lambda(h^+) - \Lambda(h^-) = \Lambda(f^+) - \Lambda(f^-) + \Lambda(g^+) - \Lambda(g^-)$

ισοδύναμα, $\Lambda(h) = \Lambda(f) + \Lambda(g)$.

Αν $f \in C_c(X)$ πραγματική συνάρτηση και $a \in \mathbb{R}$, θέλουμε να δείξουμε ότι $\Lambda(af) = a\Lambda(f)$. Πράγματι, αν $a > 0$, τότε

$$\Lambda(af) = \Lambda(af^+ - af^-) = \Lambda(af^+) - \Lambda(af^-) = a\Lambda(f^+) - a\Lambda(f^-) = a\Lambda(f).$$

Αν $a = -1$, τότε

$$\Lambda(-f) = \Lambda((-f)^+ - (-f)^-) = \Lambda((-f)^+) - \Lambda((-f)^-) = \Lambda(f^-) - \Lambda(f^+) = -\Lambda(f).$$

Αν $a < 0$, τότε

$$\Lambda(af) = \Lambda\big(-(-af)\big) = -\Lambda(-af) = -(-a)\Lambda(f) = a\Lambda(f).$$

Τέλος, θα δείξουμε ότι η σχέση (25) ισχύει για $f \in C_c(X)$. Έστω, λοιπόν, $f \in C_c(X)$. Για την πρώτη ανισότητα έχουμε

$$|\Phi(f)| = |\Phi(f^+) - \Phi(f^-)| \leq |\Phi(f^+)| + |\Phi(f^-)| \leq \Lambda(f^+) + \Lambda(f^-) = \Lambda(f^+ + f^-) = \Lambda(|f|)$$

ενώ για τη δεύτερη, αφού $|f| \in C_c^+(X)$, έχουμε $\Lambda(|f|) \leq \||f|\|_\infty = \|f\|_\infty$.

■

5. Χρήσιμες προτάσεις για την απόδειξη του Θεωρήματος 1.3.

Στην παράγραφο αυτή, θα καταγράψουμε και θα αποδείξουμε βασικές ιδιότητες των Πολωνικών χώρων τις οποίες θα χρησιμοποιήσουμε αργότερα στην απόδειξη του Θεωρήματος 1.3 του 1ου κεφαλαίου.

Ορισμός 23

Ένας μετρικός χώρος X λέγεται Πολωνικός όταν είναι πλήρης (κάθε ακολουθία Cauchy συγκλίνει) και διαχωρίσιμος (περιέχει αριθμήσιμο πυκνό υποσύνολο).

Πρόταση 24

Έστω μ ένα Borel μέτρο πιθανότητας ορισμένο πάνω σε έναν μετρικό χώρο X. Τότε, το μ είναι κανονικό, το οποίο σημαίνει ότι για κάθε Borel σύνολο A, ισχύει

$$\mu(A) = \sup\{\mu(F) : \ F \subseteq A \ \text{κλειστό}\} = \inf\{\mu(G) : \ A \subseteq G \ \text{ανοικτό}\}$$

Απόδειξη:

Θα αποδείξουμε ότι για κάθε Borel σύνολο A και για κάθε $\varepsilon > 0$, υπάρχει ένα κλειστό σύνολο F και ένα ανοικτό σύνολο G έτσι ώστε $F \subseteq A \subseteq G$ και $\mu(G \setminus F) < \varepsilon$. Έστω, λοιπόν, $\varepsilon > 0$. Έστω, επίσης, d η μετρική στον X και $d(x, A)$ η απόσταση του σημείου x από το σύνολο A.

Αν το σύνολο Α είναι κλειστό, τότε μπορούμε να πάρουμε $F = A$ και $G_\delta = \{x \in X : d(x, A) < \delta\}$ για κάποιο $\delta > 0$. Εφόσον $G_\delta \searrow A$ καθώς $\delta \searrow 0$, υπάρχει $\delta_1 > 0$ τέτοιο ώστε $\mu\left(G_{\delta_1} \setminus F\right) < \varepsilon$.

Άρα, το μόνο που χρειάζεται να δείξουμε είναι ότι η κλάση $\mathcal{A}$ των συνόλων που έχουν τη ζητούμενη ιδιότητα, αποτελεί μια σ-άλγεβρα. Τότε, εφόσον η κλάση $\mathcal{A}$ περιέχει τα κλειστά σύνολα, θα περιέχει και όλα τα σύνολα Borel.

Έστω $(A_n)_{n=1}^{\infty}$ μια ακολουθία συνόλων της $\mathcal{A}$. Για κάθε $n \in \mathbb{N}$, επιλέγουμε ένα κλειστό σύνολο F_n και ένα ανοιχτό σύνολο G_n ώστε

$$F_n \subseteq A_n \subseteq G_n \quad \text{και} \quad \mu(G_n \setminus F_n) < \frac{\varepsilon}{2^{n+1}} .$$

Παίρνουμε

$$G = \bigcup_{n=1}^{\infty} G_n \quad \text{και} \quad F = \bigcup_{n=1}^{n_0} F_n$$

με το n_0 κατάλληλα επιλεγμένο έτσι ώστε

$$\mu\left(\bigcup_{n=1}^{\infty} F_n \setminus F\right) < \frac{\varepsilon}{2}.$$

Το F είναι κλειστό ως πεπερασμένη ένωση κλειστών και το G είναι ανοικτό ως ένωση ανοικτών. Τότε

$$F \subseteq \bigcup_{n=1}^{\infty} A_n \subseteq G$$

και

$$\mu(G \setminus F) = \mu\left(\bigcup_{n=1}^{\infty} G_n \setminus F\right) = \mu\left[\left(\bigcup_{n=1}^{\infty} G_n \setminus \bigcup_{n=1}^{\infty} F_n\right) \cup \left(\bigcup_{n=1}^{\infty} F_n \setminus F\right)\right]$$
$$\leq \mu\left(\bigcup_{n=1}^{\infty} (G_n \setminus F_n)\right) + \mu\left(\bigcup_{n=1}^{\infty} F_n \setminus F\right) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Επομένως

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Ακόμα, αν $A \in \mathcal{A}$, τότε υπάρχει κλειστό F και ανοικτό G με $F \subseteq A \subseteq G$ και $\mu(G \setminus F) < \varepsilon$. Εύκολα βλέπουμε ότι $G^c \subseteq A^c \subseteq F^c$, G^c κλειστό, F^c ανοικτό και $\mu(F^c \setminus G^c) = \mu(G \setminus F) < \varepsilon$. Άρα, $A^c \in \mathcal{A}$. Τελικά, η $\mathcal{A}$ είναι κλειστή ως προς αριθμήσιμες ενώσεις και συμπληρώματα, συνεπώς αποτελεί σ-άλγεβρα.

■

Πρόταση 25

Έστω μ ένα Borel μέτρο πιθανότητας ορισμένο σε έναν Πολωνικό χώρο X. Τότε το μ είναι σφικτό, δηλαδή $\forall \varepsilon > 0$, υπάρχει $K \subseteq X$ συμπαγές ώστε $\mu(X \setminus K) \leq \varepsilon$.

Απόδειξη:

Έστω $D = \{a_1, a_2, \dots\}$ ένα αριθμήσιμο πυκνό υποσύνολο του X. Τότε, για κάθε $\delta > 0$, έχουμε

$$\bigcup_{k=1}^{\infty} B_\delta(a_k) = X.$$

Άρα,

$$\mu(X) = \lim_{n\to\infty} \mu\left(\bigcup_{k=1}^{\infty} B_\delta(a_k)\right), \quad \text{για κάθε } \delta > 0.$$

Έστω $\varepsilon > 0$. Τότε για κάθε $m \geq 1$ υπάρχει n_m τέτοιο ώστε

$$\mu\left(\bigcup_{k=1}^{n_m} B_{\frac{1}{m}}(a_k)\right) > \mu(X) - \frac{\varepsilon}{2^m}.$$

Έστω

$$K = \bigcap_{m=1}^{\infty}\bigcup_{k=1}^{n_m} \overline{B_{\frac{1}{m}}}(a_k).$$

Τότε το K είναι κλειστό και για κάθε $\delta > 0$, διαλέγοντας $m > \frac{1}{\delta}$, έχουμε

$$K \subseteq \bigcup_{k=1}^{n_m} \overline{B_{\frac{1}{m}}}(a_k) \subseteq \bigcup_{k=1}^{n_m} B_\delta(a_k).$$

Άρα, το K είναι ολικά φραγμένο, ισοδύναμα το K είναι συμπαγές (λήμμα 15). Ακόμα,

$$\mu(X \setminus K) = \mu\left(\bigcup_{m=1}^{\infty}\left(X \setminus \bigcup_{k=1}^{n_m} \overline{B_{\frac{1}{m}}}(a_k)\right)\right) \leq \sum_{m=1}^{\infty} \mu\left(X \setminus \bigcup_{k=1}^{n_m} \overline{B_{\frac{1}{m}}}(a_k)\right)$$
$$= \sum_{m=1}^{\infty}\left(\mu(X) - \mu\left(\bigcup_{k=1}^{n_m} \overline{B_{\frac{1}{m}}}(a_k)\right)\right) < \sum_{m=1}^{\infty} \frac{\varepsilon}{2^m} = \varepsilon.$$

■

Πρόταση 26 (Το λήμμα του Ulam)

Έστω μ ένα Borel μέτρο πιθανότητας ορισμένο σε έναν Πολωνικό χώρο X. Τότε, το μ συγκεντρώνεται σε ένα $\sigma-$συμπαγές σύνολο, δηλαδή υπάρχει μετρήσιμο σύνολο S, το οποίο μπορεί να γραφεί ως μια αριθμήσιμη ένωση συμπαγών συνόλων, και τέτοιο ώστε $\mu(S) = 1$.

Απόδειξη:

Ήδη γνωρίζουμε ότι ένα τέτοιο μέτρο μ είναι κανονικό. Άρα,

$$1 = \mu(X) = \sup\{\mu(K) : K \subseteq X, K \text{ συμπαγές}\}$$

Επομένως, για κάθε $n \in \mathbb{N}$, μπορούμε να βρούμε συμπαγές σύνολο $K_n \subseteq X$ τέτοιο ώστε

$$\mu(K_n) > 1 - \frac{1}{2^n}.$$

Στη συνέχεια, ορίζουμε

$$S \coloneqq \bigcup_{n=1}^{\infty} K_n.$$

Αυτό είναι ένα σ −συμπαγές σύνολο, και για κάθε $n \in \mathbb{N}$

$$1 \geq \mu(S) = \mu\left(\bigcup_{n=1}^{\infty} K_n\right) \geq 1 - \frac{1}{2^n}.$$

Άρα, $\mu[S] = 1$, και ο πρώτος ισχυρισμός απεδείχθη. Προχωράμε στην απόδειξη της ισοδυναμίας.

Για το ευθύ, έστω $\varepsilon > 0$. Εφόσον $\mu(S) = 1$, όπου $S = \bigcup_{n=1}^{\infty} K_n$, υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε

$$\mu\left(\bigcup_{n=1}^{n_0} K_n\right) \geq 1 - \varepsilon.$$

Παίρνουμε, τότε, $K_\varepsilon = \bigcup_{n=1}^{n_0} K_n$. Το K_ε είναι συμπαγές, ως πεπερασμένη ένωση συμπαγών συνόλων, και $\mu(K_\varepsilon^c) = 1 - \mu(K_\varepsilon) \leq 1 - (1-\varepsilon) = \varepsilon$. Άρα, το μ είναι σφικτό.

Αντίστροφα, έστω ότι το μ είναι σφικτό. Τότε,

$$\forall n \in \mathbb{N}, \qquad \text{υπάρχει συμπαγές σύνολο } K_n \text{ ώστε } \mu(K_n^c) \leq \frac{1}{n}.$$

Ορίζουμε $K \coloneqq \bigcup_{n=1}^{\infty} K_n$. Το K είναι σ-συμπαγές σύνολο και για κάθε $n \in \mathbb{N}$

$$\mu(K) = \mu\left(\bigcup_{n=1}^{\infty} K_n\right) \geq \mu(K_n) = 1 - \mu(K_n^c) \geq 1 - \frac{1}{n}$$

δηλαδή $\mu(K) = 1$.

■

Ορισμός 27

Έστω (X, ρ) ένας μετρικός χώρος και (μ_k) μια ακολουθία μέτρων πιθανότητας στον $P(X)$. Θα λέμε ότι η (μ_k) συγκλίνει ασθενώς στο μέτρο $\mu \in P(X)$ και θα γράφουμε $\mu_k \overset{w}{\to} \mu$, αν

$$\forall \varphi \in C_b(X), \qquad \lim_{k \to +\infty} \int_X \varphi \, d\mu_k = \int_X \varphi \, d\mu.$$

Ένα σύνολο $A \in \mathcal{B}_X$ με την ιδιότητα $\mu(\partial A) = 0$ (το ∂A είναι κλειστό, επομένως $\partial A \in \mathcal{B}_X$) λέγεται σύνολο συνέχειας για το μέτρο μ.

Πρόταση 28

Οι ακόλουθες προτάσεις είναι ισοδύναμες:

(i). $\mu_k \overset{w}{\to} \mu$

(ii). $\limsup\limits_{k\to\infty} \mu_k(F) \leq \mu(F)$, για κάθε κλειστό $F \subseteq X$

(iii). $\liminf\limits_{k\to\infty} \mu_k(G) \geq \mu(G)$, για κάθε ανοικτό $G \subseteq X$

(iv). $\lim\limits_{k\to\infty} \mu_k(A) = \mu(A)$, για κάθε σύνολο συνέχειας A για το μέτρο μ.

Απόδειξη:

$(i) \Rightarrow (ii)$. Έστω $\mu_k \overset{w}{\to} \mu$. Τότε ισχύει

$$\lim_{k\to+\infty} \int_X \varphi \, d\mu_k = \int_X \varphi \, d\mu$$

για κάθε συνεχή και φραγμένη συνάρτηση $\varphi: X \to \mathbb{R}$.

Έστω $F \subseteq X$ κλειστό και $\delta > 0$. Για τα σύνολα $G_\varepsilon = \{x \in X : \ \rho(x, F) < \varepsilon\}$, $\varepsilon > 0$ ισχύει $G_\varepsilon \downarrow F$ καθώς $\varepsilon > 0$, επομένως υπάρχει $\varepsilon > 0$ τέτοιο ώστε $\mu(G_\varepsilon) < \mu(F) + \delta$.

Ορίζουμε τη συνεχή συνάρτηση $f: \mathbb{R} \to \mathbb{R}$ ως εξής:

$$f(t) = \begin{cases} 1, & \text{αν } t \leq 0 \\ 1 - t, & \text{αν } 0 < t < 1 \\ 0, & \text{αν } t \geq 1 \end{cases}$$

και τη συνάρτηση $\varphi: X \to \mathbb{R}$ με $\varphi(x) = f\left(\frac{1}{\varepsilon}\rho(x, F)\right)$.

Τότε, η φ είναι ομοιόμορφα συνεχής στον X, $\varphi(x) = 1$ αν $x \in F$, $\varphi(x) = 0$ αν $x \in G_\varepsilon^c$ και $0 \leq \varphi(x) \leq 1$ για κάθε $x \in X$.

Έχουμε, λοιπόν,

$$\lim_{k\to+\infty} \int_X \varphi \, d\mu_k = \int_X \varphi \, d\mu.$$

Ακόμα,

$$\mu_k(F) = \int_F \varphi \, d\mu_k \leq \int_X \varphi \, d\mu_k$$

και

$$\int_X \varphi\, d\mu = \int_{G_\varepsilon} \varphi\, d\mu \leq \mu(G_\varepsilon) < \mu(F) + \delta.$$

Επομένως,

$$\limsup_{k\to\infty} \mu_k(F) \leq \lim_{k\to\infty} \int_X \varphi\, d\mu_k = \int_X \varphi\, d\mu < \mu(F) + \delta.$$

Εφόσον το δ ήταν τυχαίος θετικός θετικός αριθμός, έχουμε $\limsup_{k\to\infty} \mu_k(F) \leq \mu(F)$.

$(ii) \Rightarrow (i)$. Υποθέτουμε ότι $\limsup_{k\to\infty} \mu_k(F) \leq \mu(F)$ για κάθε $F \subseteq X$ κλειστό και ότι $\varphi \in C_b(X)$. Αρχικά, θα δείξουμε ότι

$$\limsup_{k\to\infty} \int_X \varphi\, d\mu_k \leq \int_X \varphi\, d\mu.$$

Υπάρχουν $M_1, M_2 \in \mathbb{R}$ ώστε $M_1 < \varphi(x) < M_2$, $x \in X$ από όπου προκύπτει $0 < \frac{\varphi(x)-M_1}{M_2-M_1} < 1$. Μετασχηματίζοντας, λοιπόν, την φ γραμμικά, μπορούμε να αναγάγουμε το πρόβλημα στην περίπτωση που $0 < \varphi(x) < 1$, $x \in X$.

Σταθεροποιούμε προσωρινά ένα $n \in \mathbb{Z}$ και ορίζουμε

$$F_i = \left\{x \in X : \frac{i}{n} \leq \varphi(x)\right\}, \quad i = 0, 1, \ldots, n.$$

Εφόσον $0 < \varphi(x) < 1$ έχουμε

$$\sum_{i=1}^{n} \frac{i-1}{n} \mu\left\{x \in X : \frac{i-1}{n} \leq \varphi(x) < \frac{i}{n}\right\} \leq \int_X \varphi\, d\mu \leq \sum_{i=1}^{n} \frac{i}{n} \mu\left\{x \in X : \frac{i-1}{n} \leq \varphi(x) < \frac{i}{n}\right\}.$$

Το άθροισμα στα δεξιά είναι

$$\sum_{i=1}^{n} \frac{i}{n}\big(\mu(F_{i-1}) - \mu(F_i)\big) = \frac{1}{n} + \frac{1}{n}\sum_{i=1}^{n} \mu(F_i)$$

ενώ το άθροισμα στα αριστερά είναι

$$\sum_{i=1}^{n} \frac{i-1}{n}\big(\mu(F_{i-1}) - \mu(F_i)\big) = \frac{1}{n}\sum_{i=1}^{n} \mu(F_i).$$

Συνεπώς, έχουμε

$$\frac{1}{n}\sum_{i=1}^{n} \mu(F_i) \leq \int_X \varphi\, d\mu \leq \frac{1}{n} + \frac{1}{n}\sum_{i=1}^{n} \mu(F_i).$$

Από την υπόθεσή μας, ισχύει

$$\limsup_{k\to\infty} \mu_k(F_i) \leq \mu(F_i), \quad i = 0, 1, \dots, n$$

και άρα

$$\limsup_{k\to\infty} \int_X \varphi \, d\mu_k \leq \frac{1}{n} + \int_X \varphi \, d\mu.$$

Αφήνοντας το $n \to \infty$, παίρνουμε τη σχέση

$$\limsup_{k\to\infty} \int_X \varphi \, d\mu_k \leq \int_X \varphi \, d\mu.$$

Εφαρμόζοντας την παραπάνω σχέση στη $-\varphi$, παίρνουμε

$$\liminf_{k\to\infty} \int_X \varphi \, d\mu_k \geq \int_X \varphi \, d\mu.$$

Συνδυάζοντας τις δύο τελευταίες σχέσεις, προκύπτει η ασθενής σύγκλιση $\mu_k \xrightarrow{w} \mu$.

$(ii) \Leftrightarrow (iii)$. Η ισοδυναμία αυτή προκύπτει εύκολα παίρνοντας τα συμπληρώματα.

$(ii) \Rightarrow (iv)$. Για κάθε $A \in \mathcal{B}_X$, έχουμε

$$\mu(\bar{A}) \geq \limsup_{k\to\infty} \mu_k(\bar{A}) \geq \limsup_{k\to\infty} \mu_k(A) \geq \liminf_{k\to\infty} \mu_k(A) \geq \liminf_{k\to\infty} \mu_k\big(Int(A)\big) \geq \mu\big(Int(A)\big).$$

Αν $\mu(\partial A) = 0$, έχουμε $\mu(\bar{A}) = \mu\big(Int(A)\big)$ και έτσι έπεται ότι $\lim_{k\to\infty} \mu_k(A) = \mu(A)$.

$(iv) \Rightarrow (ii)$. Έστω $F \subseteq X$ κλειστό. Αφού $\partial\{x \in X : \rho(x,F) \leq \delta\} \subseteq \{x \in X : \rho(x,F) = \delta\}$, τα σύνολα αυτά είναι ξένα για διαφορετικά $\delta > 0$, και άρα το πολύ αριθμήσιμο πλήθος από αυτά μπορούν να έχουν θετικό μέτρο ως προς το μ. Επομένως, υπάρχει ακολουθία θετικών αριθμών $\delta_n \searrow 0$ ώστε για τα σύνολα $F_n = \{x \in X : \rho(x,F) \leq \delta_n\}$ να ισχύει $\mu(\partial F_n) = 0$. Έχουμε, λοιπόν,

$$\limsup_{k\to\infty} \mu_k(F) \leq \lim_{k\to\infty} \mu_k(F_n) = \mu(F_n), \text{ για κάθε } n \in \mathbb{N}.$$

Ακόμα, αφού το F είναι κλειστό, $F_n \searrow F$, οπότε η (ii) ισχύει.

■

Ο χώρος $\mathbb{R}^\infty$

Ο $\mathbb{R}^\infty$ είναι το τοπολογικό γινόμενο μιας αριθμήσιμης ακολουθίας αντιγράφων του $\mathbb{R}$, δηλαδή ως σύνολο είναι ο χώρος των ακολουθιών $x = (x_1, x_2, \dots)$ πραγματικών αριθμών. Η τοπολογία έχει ως βασικές περιοχές ενός σημείου x τα σύνολα της μορφής $N_{k,\varepsilon}(x) = \{y \colon |x_i - y_i| < \varepsilon, \quad i = 1,2,\dots,k\}$ με $\varepsilon > 0$ και $k \in \mathbb{N}$. Άρα, η σύγκλιση στον $\mathbb{R}^\infty$ συμπίπτει με τη σύγκλιση κατά συντεταγμένη, δηλαδή $\lim_{n\to\infty} x(n) = x$ αν και μόνο αν $\lim_{n\to\infty} x_k(n) = x_k$ για κάθε k. Με αυτή την τοπολογία, ο $\mathbb{R}^\infty$ είναι ένας πλήρης διαχωρίσιμος μετρικός χώρος. ([4], σελ. 218 – 219)

Συμβολίζουμε με $\pi_n: \mathbb{R}^\infty \longrightarrow \mathbb{R}^n$ την προβολή που ορίζεται ως εξής: $\pi_n(x) = (x_1, \dots, x_n)$. Κάθε σύνολο της μορφής $\pi_n^{-1}H$ με $H \in \mathcal{B}_{\mathbb{R}^n}$ λέγεται σύνολο πεπερασμένης διάστασης. Αφού κάθε προβολή π_n είναι συνεχής και άρα Borel μετρήσιμη, τα πεπερασμένης διάστασης σύνολα είναι Borel υποσύνολα του $\mathbb{R}^\infty$. Έστω $\mathcal{F}$ η κλάση των πεπερασμένης διάστασης συνόλων και $\mathcal{B}$ μια $\sigma-$άλγεβρα υποσυνόλων του $\mathbb{R}^\infty$ με $\mathcal{F} \subseteq \mathcal{B}$. Τα σύνολα $N_{k,\varepsilon}(x) = \{y\colon |x_i - y_i| < \varepsilon, \quad i = 1,2,\dots,k\}$, $\varepsilon > 0, k \in \mathbb{N}, x \in \mathbb{R}^\infty$ ανήκουν στην $\mathcal{F}$ και αποτελούν βάση της τοπολογίας του $\mathbb{R}^\infty$. Αφού ο $\mathbb{R}^\infty$ είναι διαχωρίσιμος, κάθε ανοιχτό σύνολο γράφεται ως αριθμήσιμη ένωση συνόλων της μορφής $N_{k,\varepsilon}(x)$, άρα κάθε ανοιχτό σύνολο ανήκει στην $\mathcal{B}$. Έπεται ότι $\mathcal{B}_{\mathbb{R}^\infty} \subseteq \mathcal{B}$, δηλαδή η $\mathcal{B}_{\mathbb{R}^\infty}$ είναι η ελάχιστη $\sigma-$άλγεβρα που περιέχει την $\mathcal{F}$. Άρα, η κλάση $\mathcal{F}$ γεννάει την $\mathcal{B}_{\mathbb{R}^\infty}$. Αποδεικνύεται ότι η $\mathcal{F}$ αποτελεί μια κλάση που καθορίζει την ασθενή σύγκλιση, δηλαδή $\mu_k \overset{w}{\to} \mu$ αν και μόνο αν $\mu_k(A) \to \mu(A)$ για όλα τα πεπερασμένης διάστασης σύνολα συνέχειας ως προς το μ. ([4], σελ. 19)

Κριτήριο συμπάγειας στον $\mathbb{R}^\infty$

Έστω $A \subseteq \mathbb{R}^\infty$. Το A έχει συμπαγή κλειστότητα αν και μόνο αν το σύνολο $\{x_k : x \in A\}$ είναι, για κάθε $k \in \mathbb{N}$, ένα φραγμένο υποσύνολο του $\mathbb{R}$.

Απόδειξη:

Για το ευθύ, έστω ότι το $\overline{A}$ είναι συμπαγές και $\varepsilon > 0$. Τότε υπάρχουν σημεία $x(1), x(2), \dots, x(n)$ τέτοια ώστε

$$\overline{A} \subseteq \bigcup_{i=1}^{n} N_{k,\varepsilon}\big(x(i)\big).$$

Αν $x \in A$, τότε για την $k-$οστή του συντεταγμένη ισχύει

$$\min_{1\le i\le n}\{x_k(i) - \varepsilon\} \le x_k \le \max_{1\le i\le n}\{x_k(i) + \varepsilon\},$$

δηλαδή το σύνολο $\{x_k : x \in A\}$ είναι φραγμένο ως υποσύνολο του $\mathbb{R}$.

Για το αντίστροφο, θα χρησιμοποιήσουμε την κλασσική διαγώνια μέθοδο. Έστω μια ακολουθία $\{x(n)\}$ στο A. Επιλέγουμε μια ακολουθία υπακολουθιών

$$\begin{cases} x(n_{11}), & x(n_{12}), & x(n_{13}), \dots\dots \\ x(n_{21}), & x(n_{22}), & x(n_{23}), \dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots & & \\ \dots\dots\dots\dots\dots\dots\dots\dots & & \end{cases}$$

με τον ακόλουθο τρόπο. Η 1η γραμμή είναι μια υπακολουθία της $\{x(n)\}$, επιλεγμένη ώστε το $x_1 = \lim_{i\to\infty} x_1(n_{1i})$ να υπάρχει. Υπάρχει μια τέτοια υπακολουθία επειδή το σύνολο $\{x_1 : x \in A\}$ είναι ένα φραγμένο υποσύνολο του $\mathbb{R}$. Η 2η γραμμή είναι μια υπακολουθία της 1ης γραμμής, επιλεγμένη ώστε το $x_2 = \lim_{i\to\infty} x_2(n_{2i})$ να υπάρχει. Υπάρχει μια τέτοια υπακολουθία επειδή το σύνολο $\{x_2 : x \in A\}$ είναι ένα φραγμένο υποσύνολο του $\mathbb{R}$. Συνεχίζοντας με τον ίδιο τρόπο, η $k-$οστή γραμμή είναι μια υπακολουθία της γραμμής $(k-1)$ και το $x_k = \lim_{i\to\infty} x_k(n_{ki})$ υπάρχει. Συμβολίζουμε $x = (x_1, x_2, \dots, x_k, \dots)$ το σημείο του $\mathbb{R}^\infty$ με συντεταγμένες τα όρια x_k. Αν $n_i = n_{ii}$, τότε η $\{x(n_i)\}$ είναι μια υπακολουθία της $\{x(n)\}$. Επιπλέον, για κάθε k, οι όροι $x(n_k), x(n_{k+1}), \dots$ αποτελούν υπακολουθία της $k-$οστής γραμμής, άρα $\lim_{i\to\infty} x_k(n_i) = x_k$. Επομένως, $\lim_{i\to\infty} x(n_i) = x$ και έπεται ότι το $\overline{A}$ είναι συμπαγές.

■

Αν $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$, με τον συμβολισμό $y < x$ $(y \le x)$ εννοούμε ότι $y_i < x_i$ $(y_i \le x_i)$ για κάθε $i = 1,2,\dots,n$. Ένα σύνολο της μορφής $(a, b] = \{x \in \mathbb{R}^n : \ a < x \le b\} = (a_1, b_1] \times \dots \times (a_n, b_n]$ λέγεται $n-$διάστατο ορθογώνιο. Έστω $V = \{a_1, b_1\} \times \dots \times \{a_n, b_n\}$ το σύνολο των 2^n κορυφών του ορθογωνίου $(a, b]$. Αν $v \in V$, ορίζουμε το πρόσημο της κορυφής αυτής: $sgn(v) = (-1)^{\#\text{των } a_i \text{ στην } v}$.

Μια συνάρτηση $F: \mathbb{R}^n \to \mathbb{R}$ λέγεται συνεχής στο σημείο x αν $\forall \varepsilon > 0$, $\exists \delta > 0$ τέτοιο ώστε: $x - \delta e < y < x + \delta e \implies |F(x) - F(y)| < \varepsilon$, όπου $e = (1, 1, \dots, 1)$.

Μια συνάρτηση $F: \mathbb{R}^n \to \mathbb{R}$ λέγεται συνεχής από πάνω στο σημείο x αν $\forall \varepsilon > 0$, $\exists \delta > 0$ τέτοιο ώστε: $x \le y < x + \delta e \implies |F(x) - F(y)| < \varepsilon$.

Μια συνάρτηση κατανομής στον $\mathbb{R}^n$ είναι μια συνάρτηση $F(x) = F(x_1, \dots, x_n)$ με τις ακόλουθες τρεις ιδιότητες:

(i) Η F είναι παντού συνεχής από πάνω.

(ii) $0 \le F(x) \le 1$ για κάθε $x \in \mathbb{R}^n$, η F είναι αύξουσα ως προς κάθε μεταβλητή και για κάθε ορθογώνιο $A = (a, b]$ στον $\mathbb{R}^n$ ισχύει

$$\Delta_A F = \sum_{v \in V} sgn(v)\, F(v) \ge 0.$$

(iii) $F(x) \to 0$ καθώς οποιαδήποτε συντεταγμένη του x τείνει στο $-\infty$ και $F(x) \to 1$ καθώς όλες οι συντεταγμένες του x τείνουν στο $+\infty$.

Αν F είναι μια συνάρτηση κατανομής στον $\mathbb{R}^n$, υπάρχει ακριβώς ένα μέτρο πιθανότητας μ στον $\mathbb{R}^n$ τέτοιο ώστε $\mu(A) = \Delta_A F$ για όλα τα ορθογώνια $A = (a, b]$. ([19], Θεώρημα 1.6, σελ. 446)

Αν η F δεν είναι συνεχής στο x, τότε η ένωση των n υπερεπιπέδων, τα οποία είναι παράλληλα προς τα κύρια υπερεπίπεδα και περιέχουν το x, έχει θετικό μέτρο ως προς το μ. Φανερά, υπάρχει το πολύ αριθμήσιμο πλήθος τέτοιων υπερεπιπέδωνμε θετικό μέτρο ως προς το μ.

Απόδειξη:

Καταρχήν, θα δείξουμε ότι αν $F\left(x+\frac{1}{k}e\right)-F\left(x-\frac{1}{k}e\right)\longrightarrow 0$ καθώς $k\to\infty$, τότε η F είναι συνεχής στο x.

Πράγματι, έστω ακολουθία $x_\rho\to x$ και $\varepsilon>0$. Υπάρχει k_0 ώστε $F\left(x+\frac{1}{k_0}e\right)-F\left(x-\frac{1}{k_0}e\right)<\varepsilon$. Ακόμα, υπάρχει ρ_0 ώστε για $\rho\geq\rho_0$ να έχουμε $x-\frac{1}{k_0}e<x_\rho<x+\frac{1}{k_0}e$. Τότε, $F\left(x-\frac{1}{k_0}e\right)\leq F(x_\rho)\leq F\left(x+\frac{1}{k_0}e\right)$, $\rho\geq\rho_0$ και $F\left(x-\frac{1}{k_0}e\right)\leq F(x)\leq F\left(x+\frac{1}{k_0}e\right)$. Άρα, για $\rho\geq\rho_0$ έχουμε

$$\left|F(x_\rho)-F(x)\right|\leq F\left(x+\frac{1}{k_0}e\right)-F\left(x-\frac{1}{k_0}e\right)<\varepsilon,$$

δηλαδή $F(x_\rho)\to F(x)$.

Στη συνέχεια, υποθέτουμε ότι η F δεν είναι συνεχής στο x. Τότε, $F\left(x+\frac{1}{k}e\right)-F\left(x-\frac{1}{k}e\right)\nrightarrow 0$ καθώς $k\to\infty$. Αν μ είναι το μέτρο που έχει συνάρτηση κατανομής την F, τότε

$$\begin{aligned}F\left(x+\frac{1}{k}e\right)-F\left(x-\frac{1}{k}e\right)&=\mu\left(\left\{y:\ y\leq x+\frac{1}{k}e\right\}\setminus\left\{y:\ y\leq x-\frac{1}{k}e\right\}\right)\\&=\mu\left(\left\{y:\ y\leq x+\frac{1}{k}e\right\}\right)-\mu\left(\left\{y:\ y\leq x-\frac{1}{k}e\right\}\right)\longrightarrow\mu(\{y:\ y\leq x\})-\mu(\{y:\ y<x\})>0.\end{aligned}$$

Αν $y\in\mathbb{R}^n$ με $y\leq x$ και $y\nless x$, τότε $y_i\leq x_i,\ \ i=1,2,\dots n$ και υπάρχει i ώστε $\mathrm{y_i}=x_i$. Ισχύει

$$\{y:\ y\leq x, y\nless x\}\subseteq\bigcup_{i=1}^{n}\{y:\ y_i=x_i\}$$

άρα η ένωση των n υπερεπιπέδων που είναι παράλληλα προς τα κύρια υπερεπίπεδα και περιέχουν το x έχει θετικό μέτρο.

■

Θεώρημα 29 (Θεώρημα επιλογής Helly)

Έστω F_k μια ακολουθία συναρτήσεων κατανομής στον $\mathbb{R}^n$. Τότε, υπάρχει υπακολουθία F_{k_l} και συνάρτηση F στον $\mathbb{R}^n$ που ικανοποιεί τις συνθήκες (i) και (ii) παραπάνω (αλλά πιθανόν όχι την (iii)) και τέτοια ώστε $\lim\limits_{l\to\infty}F_{k_l}(x)=F(x)$ για όλα τα σημεία συνέχειας x της F.

Απόδειξη:

Έστω $\mathbb{R}_0^n$ το σύνολο των σημείων του $\mathbb{R}^n$ με ρητές συντεταγμένες και έστω $r_1, r_2, \dots$ μια απαρίθμηση των στοιχείων του $\mathbb{R}_0^n$. Για κάθε F_k στη δοσμένη ακολουθία συναρτήσεων κατανομής, ισχύει $(F_k(r_1), F_k(r_2), \dots\ )\in\mathbb{R}^\infty$.

Αφού $0 \leq F_k(x) \leq 1$, από το κριτήριο συμπάγειας για τον $\mathbb{R}^\infty$ προκύπτει ότι μια υπακολουθία των παραπάνω σημείων συγκλίνει υπό την έννοια του $\mathbb{R}^\infty$ σε κάποιο σημείο $(z_1, z_2, \dots) \in \mathbb{R}^\infty$. Ορίζουμε $F_0 : \mathbb{R}_0^n \to \mathbb{R}$ με $F_0(r_k) = z_k$. Έτσι, βλέπουμε ότι υπάρχει συνάρτηση F_0 στον $\mathbb{R}_0^n$ και υπακολουθία F_{k_l} της F_k τέτοια ώστε $\lim_{l\to\infty} F_{k_l}(r) = F_0(r),\ \ r \in \mathbb{R}_0^n$.

Αφού κάθε F_k είναι μια συνάρτηση κατανομής, προκύπτει ότι η F_0 ικανοποιεί τη συνθήκη (ii) στον $\mathbb{R}_0^n$, δηλαδή $0 \leq F_0(r) \leq 1$, η F_0 είναι αύξουσα καθώς οποιαδήποτε συντεταγμένη μεταβάλλεται πάνω στους ρητούς αριθμούς και

$$\Delta_A F_0 = \sum_{v\in V} sgn(v)\, F_0(v) \geq 0$$

για οποιοδήποτε ορθογώνιο $A = (a, b]$ με $a, b \in \mathbb{R}_0^n$.

Ορίζουμε $F : \mathbb{R}^n \to \mathbb{R}$ ως εξής: $F(x) = \inf\{F_0(r) :\ \ x < r\,, r \in \mathbb{R}_0^n\},\quad x \in \mathbb{R}^n$.

Η F ικανοποιεί τις συνθήκες (i) και (ii), αλλά πιθανόν όχι την (iii).

Πράγματι, η F φανερά είναι αύξουσα. Για το (i), δείχνουμε αρχικά ότι αν $F(x + \frac{1}{k}e) \longrightarrow F(x)$ καθώς $k \to \infty$ τότε η F είναι συνεχής από πάνω στο x.

Έστω ακολουθία $y_\rho \to x$ με $y_\rho > x$ και $\varepsilon > 0$. Υπάρχει k_0 ώστε $F\left(x + \frac{1}{k_0}e\right) < F(x) + \varepsilon$ και υπάρχει ρ_0 ώστε $x < y_\rho < x + \frac{1}{k_0}e,\ \ \rho \geq \rho_0$. Τότε, $F(x) \leq F(y_\rho) \leq F\left(x + \frac{1}{k_0}e\right)$. Άρα, για $\rho \geq \rho_0$ έχουμε $|F(y_\rho) - F(x)| < F\left(x + \frac{1}{k_0}e\right) - F(x) < \varepsilon$, δηλαδή $F(y_\rho) \to F(x)$.

Έστω $x \in \mathbb{R}^n$. Από τον ορισμό της F, έχουμε $F\left(x + \frac{1}{k}e\right) = \inf\left\{F_0(r):\ \ x + \frac{1}{k}e < r, r \in \mathbb{R}_0^n\right\}$. Αν ορίσουμε $A_k = \left\{F_0(r):\ \ x + \frac{1}{k}e < r, r \in \mathbb{R}_0^n\right\}$, τότε $A_k \subseteq A_{k+1}$ και

$$A = \bigcup_{k=1}^{\infty} A_k = \{F_0(r) :\ \ x < r\,, r \in \mathbb{R}_0^n\}.$$

Ισχύει, λοιπόν $F\left(x + \frac{1}{k}e\right) = \inf A_k \longrightarrow \inf A = F(x)$, δηλαδή η F είναι συνεχής από πάνω στο x.

Έστω, ακόμα, ορθογώνιο $A = (a, b]$. Για κάθε κορυφή $v \in V$, υπάρχει ακολουθία $\left\{r_k^{(v)}\right\}$ στον $\mathbb{R}_0^n$ με $r_k^{(v)} > v$ και $r_k^{(v)} \to v$ καθώς $k \to \infty$. Τότε, $F_0\left(r_k^{(v)}\right) \to F(v)$ και αφού

$$\sum_{v\in V} sgn(v) F_0\left(r_k^{(v)}\right) \geq 0, \text{θα ισχύει και} \sum_{v\in V} sgn(v) F(v) \geq 0.$$

Έστω x σημείο συνέχειας της F και $\varepsilon > 0$. Βρίσκουμε σημεία $r', r'' \in \mathbb{R}_0^n$ τέτοια ώστε $r' < x < r''$ και $F(x) - \varepsilon < F_0(r') \leq F_0(r'') < F(x) + \varepsilon$.

Για κάθε $k_l \in \mathbb{N}$, έχουμε $F_{k_l}(r') \leq F_{k_l}(x) \leq F_{k_l}(r'')$.

Συνδυάζοντας τις παραπάνω σχέσεις παίρνουμε

$$F(x) - \varepsilon < F_0(r') = \lim_{l\to\infty} F_{k_l}(r') \leq \liminf_{l\to\infty} F_{k_l}(x) \leq \limsup_{l\to\infty} F_{k_l}(x)$$
$$\leq \lim_{l\to\infty} F_{k_l}(r'') = F_0(r'') < F(x) + \varepsilon.$$

Αφού το ε ήταν τυχαίος θετικός αριθμός, έχουμε $\lim_{l\to\infty} F_{k_l}(x) = F(x)$ και η απόδειξη του θεωρήματος είναι πλήρης.

■

Ορισμός 30

Μια οικογένεια $\mathcal{P}$ μέτρων πιθανότητας σε έναν τοπολογικό χώρο X ονομάζεται σφικτή αν για κάθε $\varepsilon > 0$ υπάρχει συμπαγές σύνολο $K_\varepsilon \subseteq X$ ώστε $\sup_{\mu\in\mathcal{P}} \mu(X \setminus K_\varepsilon) \leq \varepsilon$.

Θεώρημα 31 (Prokhorov)

Έστω X ένας μετρικός χώρος και $\mathcal{P}$ μια σφικτή οικογένεια μέτρων στον $P(X)$. Τότε, η $\mathcal{P}$ είναι σχετικά ακολουθιακά συμπαγής στον $P(X)$.

Αυτό σημαίνει ότι για οποιαδήποτε ακολουθία μέτρων (μ_k) στην $\mathcal{P}$ υπάρχει μια υπακολουθία $\left(\mu_{n_k}\right)$ και ένα μέτρο πιθανότητας $\mu_* \in P(X)$, έτσι ώστε για κάθε $\varphi \in C_b(X)$,

$$\lim_{k\to+\infty} \int_X \varphi \, d\mu_{n_k} = \int_X \varphi \, d\mu_*.$$

Η απόδειξη του θεωρήματος θα χωριστεί σε 4 βήματα: $X = \mathbb{R}^n$, $X = \mathbb{R}^\infty$, $X = \sigma$ −συμπαγής, $X =$ αυθαίρετος μετρικός χώρος. Θα αποδείξουμε κάθε ένα από τα βήματα 2, 3 και 4 κάνοντας αναγωγή στο αμέσως προηγούμενο βήμα.

Για το 2ο βήμα της απόδειξης θα χρειαστούμε το παρακάτω λήμμα.

Λήμμα 32

Έστω X και X' δύο μετρικοί χώροι. Εάν $\mathcal{P}$ είναι μια σφικτή οικογένεια μέτρων πιθανότητας στον $(X, \mathcal{B}_X)$ και $h: X \longrightarrow X'$ είναι μια συνεχής απεικόνιση, τότε η οικογένεια $\mathcal{P}' = \{\mu h^{-1} : \ \mu \in \mathcal{P}\}$ είναι σφικτή στον $(X', \mathcal{B}_{X'})$.

Απόδειξη:

Έστω $\varepsilon > 0$. Επιλέγουμε συμπαγές $K \subseteq X$ τέτοιο ώστε $\mu(K) > 1 - \varepsilon$ για κάθε $\mu \in \mathcal{P}$. Εάν $K' = h(K)$, τότε το K' είναι, επίσης, συμπαγές (η εικόνα συμπαγούς συνόλου μέσω συνεχούς απεικόνισης είναι συμπαγές) και $h^{-1}(K') \supseteq K$. Συνεπώς, για κάθε $\mu \in \mathcal{P}$ έχουμε

$$\mu h^{-1}(K') = \mu\big(h^{-1}(K')\big) \geq \mu(K) > 1 - \varepsilon$$

δηλαδή, η οικογένεια $\mathcal{P}'$ είναι σφικτή.

■

Θεώρημα 33

Έστω (X, ρ) και (X', ρ') δύο μετρικοί χώροι και $h: X \longrightarrow X'$ μια Borel μετρήσιμη απεικόνιση. Θέτουμε $D_h = \{x \in X :\ \text{η } h \text{ δεν είναι συνεχής στο } x\}$.

Τότε, $D_h \in \mathcal{B}_X$ και ισχύει:

$$\text{αν } \mu_k \xrightarrow{w} \mu \text{ και } \mu(D_h) = 0, \text{ τότε } \mu_k h^{-1} \xrightarrow{w} \mu h^{-1}.$$

Απόδειξη:

Ορίζουμε το σύνολο

$$A_{\varepsilon,\delta} = \big\{x \in X :\ \exists y, z \in X \text{ τ.ω } \rho(x,y) < \delta, \rho(x,z) < \delta \text{ και } \rho'\big(h(y), h(z)\big) \geq \varepsilon\big\}.$$

Τότε, το $A_{\varepsilon,\delta}$ είναι ανοικτό σύνολο. Πράγματι, αν $x \in A_{\varepsilon,\delta}$, τότε βρίσκουμε μπάλα $B_\eta(x) \subseteq A_{\varepsilon,\delta}$ με $\eta = \min\{\delta - \rho(x,y), \delta - \rho(x,z)\}$, όπου y, z όπως στον ορισμό του $A_{\varepsilon,\delta}$.

Επίσης,

$$D_h = \bigcup_{\substack{\varepsilon > 0 \\ \varepsilon:\ \text{ρητός}}} \bigcap_{\substack{\delta > 0 \\ \delta:\ \text{ρητός}}} A_{\varepsilon,\delta},$$

άρα, το D_h είναι ένα Borel σύνολο στον X.

Θα δείξουμε ότι, αν F είναι ένα κλειστό υποσύνολο του X', τότε $\limsup_k \mu_k h^{-1}(F) \leq \mu h^{-1}(F)$.

Αφού $\mu_k \xrightarrow{w} \mu$, έχουμε $\limsup_k \mu_k\big(h^{-1}(F)\big) \leq \limsup_k \mu_k\big(\overline{h^{-1}(F)}\big) \leq \mu\big(\overline{h^{-1}(F)}\big)$.

Άρα, αρκεί να αποδείξουμε ότι $\mu\big(\overline{h^{-1}(F)}\big) = \mu\big(h^{-1}(F)\big)$, το οποίο, όμως, προκύπτει από την υπόθεση ότι $\mu(D_h) = 0$ και τον εγκλεισμό $\overline{h^{-1}(F)} \subseteq D_h \cup h^{-1}(F)$.

Ο παραπάνω εγκλεισμός είναι πράγματι σωστός αφού αν $x \in \overline{h^{-1}(F)}$, τότε υπάρχει ακολουθία (x_n) στο $h^{-1}(F)$ με $x_n \to x$. Άρα, $h(x_n) \in F$. Αν η h είναι συνεχής στο x, τότε $h(x_n) \to h(x)$ και αφού F κλειστό, έπεται ότι $h(x) \in F$, δηλαδή $x \in h^{-1}(F)$. Αν, πάλι, η h δεν είναι συνεχής στο x, τότε $x \in D_h$.

■

Για το 3ο βήμα της απόδειξης όπου υποθέτουμε ότι ο X είναι σ −συμπαγής αλλά και τη γενική περίπτωση του αυθαίρετου μετρικού χώρου, θα χρειαστούμε δύο ακόμα λήμματα.

Έστω X_0 ένα Borel υποσύνολο του X. Ο X_0 είναι και ο ίδιος ένας μετρικός χώρος αναφορικά με τη σχετική τοπολογία. Έστω $\mathcal{B}_{X_0}$ η κλάση των Borel συνόλων για τον X_0. Τότε $\mathcal{B}_{X_0} = \{A \subseteq X_0 : A \in \mathcal{B}_X\}$ και ειδικότερα, $\mathcal{B}_{X_0} \subseteq \mathcal{B}_X$.

Εάν μ είναι ένα μέτρο πιθανότητας στον $(X, \mathcal{B}_X)$ με $\mu(X_0) = 1$, συμβολίζουμε μ^r το μέτρο πιθανότητας στον $(X_0, \mathcal{B}_{X_0})$ που προκύπτει από τον περιορισμό του μ στην κλάση $\mathcal{B}_{X_0}$.

Εάν μ είναι ένα μέτρο πιθανότητας στον $(X_0, \mathcal{B}_{X_0})$, συμβολίζουμε μ^e το μέτρο πιθανότητας στον $(X, \mathcal{B}_X)$ με $\mu^e(A) = \mu(A \cap X_0)$, $A \in \mathcal{B}_X$. Σημειώνουμε ότι $\mu^e(X_0) = 1$.

Εάν μ είναι ένα μέτρο πιθανότητας στον $(X, \mathcal{B}_X)$ με $\mu(X_0) = 1$, τότε $(\mu^r)^e = \mu$, ενώ αν μ είναι ένα μέτρο πιθανότητας στον $(X_0, \mathcal{B}_{X_0})$, τότε $(\mu^e)^r = \mu$.

Λήμμα 34

Έστω X ένας μετρικός χώρος και $X_0 \in \mathcal{B}_X$. Αν $\mathcal{P}$ είναι μια σφικτή οικογένεια μέτρων πιαθνότητας στον $(X_0, \mathcal{B}_{X_0})$, τότε η $\mathcal{P}^e = \{\mu^e : \mu \in \mathcal{P}\}$ είναι μια σφικτή οικογένεια στον $(X, \mathcal{B}_X)$. Αν $\mu_k \overset{w}{\to} \mu$ στον $(X_0, \mathcal{B}_{X_0})$, τότε $\mu_k^e \overset{w}{\to} \mu^e$ στον $(X, \mathcal{B}_X)$.

<u>Απόδειξη:</u>

Στο λήμμα 32 και στο θεώρημα 33 παίρνουμε την h να είναι η ταυτοτική απεικόνιση από τον X_0 στον X.

■

Λήμμα 35

Έστω X ένας μετρικός χώρος και $X_0 \in \mathcal{B}_X$. Αν $\mu_k \overset{w}{\to} \mu$ στον $(X, \mathcal{B}_X)$ και $\mu_k(X_0) = \mu(X_0) = 1$, τότε $\mu_k^r \overset{w}{\to} \mu^r$ στον $(X_0, \mathcal{B}_{X_0})$.

Απόδειξη:

Ένα ανοικτό σύνολο στον X_0 είναι της μορφής $G_0 = G \cap X_0$, όπου G είναι ανοικτό στον X. Αφού $\mu_k^r(G_0) = \mu_k(G)$ και $\mu^r(G_0) = \mu(G)$, έχουμε:

$$\mu_k \xrightarrow{w} \mu \iff \liminf_n \mu_k(G) \geq \mu(G), \text{ για κάθε ανοικτό } G \text{ στον } X \iff \liminf_n \mu_k^r(G_0)$$
$$\geq \mu^r(G_0), \text{ για κάθε ανοικτό } G_0 \text{ στον } X_0 \iff \mu_k^r \xrightarrow{w} \mu^r.$$

■

Τέλος, θα χρειαστούμε και το Θεώρημα Kolmogorov του οποίου τη διατύπωση παραθέτουμε αμέσως παρακάτω. Απόδειξη του θεωρήματος δίνεται στο ([4], σελ. 228).

Έστω π_k η προβολή από τον $\mathbb{R}^\infty$ στον $\mathbb{R}^k$, δηλαδή αν $x = (x_1, x_2, \dots) \in \mathbb{R}^\infty$, τότε $\pi_k(x) = (x_1, \dots, x_k)$. Για $k > 1$, έστω ψ_k η προβολή από τον $\mathbb{R}^k$ στον $\mathbb{R}^{k-1}$, $\psi_k(x_1, \dots, x_k) = (x_1, \dots, x_{k-1})$.

Αν P είναι ένα μέτρο πιθανότητας στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ ορίζουμε τα μέτρα πιθανότητας μ_k στον $\left(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k}\right)$ ως εξής: $\mu_k = P\pi_k^{-1}$.

Εφόσον $\pi_{k-1} = \psi_k \pi_k,\ \ k > 1,$ τα μέτρα μ_k ικανοποιούν τη σχέση $\mu_{k-1} = \mu_k \psi_k^{-1},\ \ k > 1$.

Θεώρημα Kolmogorov

Εάν τα μέτρα πιθανότητας μ_k στον $\left(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k}\right),\ \ \kappa \geq 1,$ ικανοποιούν τη σχέση $\mu_{k-1} = \mu_k \psi_k^{-1},\ k > 1$ τότε, υπάρχει στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ ένα μοναδικό μέτρο πιθανότητας P που ικανοποιεί τη σχέση $\mu_k = P\pi_k^{-1}$.

Απόδειξη του Θεωρήματος του Prokhorov:

Βήμα 1ο: $X = \mathbb{R}^n$

Έστω σφικτή ακολουθία $(\mu_k) \in \mathcal{P}$. Αν (F_k) είναι η ακολουθία των αντίστοιχων συναρτήσεων κατανομής, τότε από το Θεώρημα Helly, γνωρίζουμε ότι υπάρχει υπακολουθία $\left(F_{k_l}\right)$ και συνάρτηση F συνεχής από πάνω τέτοια ώστε $F_{k_l}(x) \longrightarrow F(x)$ για όλα τα σημεία συνέχειας της F.

Στον $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ υπάρχει μέτρο μ τέτοιο ώστε για κάθε ορθογώνιο $(a, b]$ να ισχύει

$$\mu(a, b] = \Delta_{(a,b]}F = \sum_{v \in V} sgn(v)\, F(v),$$

όπου V είναι το σύνολο των 2^n κορυφών του ορθογωνίου $(a, b]$ και το πρόσημο κάθε κορυφής $v \in V$ είναι $sgn(v) = (-1)^{\#\text{ των } a_i \text{ στην } v}$.

Θα αποδείξουμε ότι $\mu(\mathbb{R}^n) = 1$. Έστω, λοιπόν, $\varepsilon > 0$. Αφού η οικογένεια $\mathcal{P}$ είναι σφικτή, υπάρχει συμπαγές σύνολο $K \subseteq \mathbb{R}^n$ τέτοιο ώστε για κάθε $l \in \mathbb{N}$, $\mu_{k_l}(K) > 1 - \varepsilon$. Στη συνέχεια, επιλέγουμε $a, b \in \mathbb{R}^n$ έτσι ώστε $K \subseteq (a, b]$, οι 2^n κορυφές του ορθογωνίου να είναι σημεία συνέχειας της F και τα $2n$ υπερεπίπεδα που περιέχουν τις έδρες του ορθογωνίου να έχουν μέτρο 0 ως προς το μ και τα μ_{k_l}. Αυτό είναι εφικτό, γιατί το πλήθος των παράλληλων υπερεπιπέδων που μπορούν να έχουν θετικό μέτρο ως προς κάποιο μέτρο είναι το πολύ αριθμήσιμο.

Αφού

$$\mu_{k_l}(a, b] = \Delta_{(a,b]} F_{k_l} = \sum_{v \in V} sgn(v)\, F_{k_l}(v),$$

η σχέση $F_{k_l}(x) \longrightarrow F(x)$ συνεπάγεται ότι $\mu_{k_l}(a, b] \longrightarrow \mu(a, b]$.

Αφού για κάθε $l \in \mathbb{N}$, $\mu_{k_l}(a, b] \geq \mu_{k_l}(K) > 1 - \varepsilon$, έπεται ότι $\mu(a, b] \geq 1 - \varepsilon$. Αφού το ε ήταν ένας τυχαίος θετικός αριθμός, έχουμε $\mu(\mathbb{R}^n) = 1$.

Άρα, το μ είναι μέτρο πιθανότητας και αφού η σύγκλιση $F_{k_l}(x) \longrightarrow F(x)$ ισχύει σε όλα τα σημεία συνέχειας της F, προκύπτει η ασθενής σύγκλιση $\mu_{k_l} \overset{w}{\to} \mu$.

Πράγματι, ένα ορθογώνιο $(a, b]$ καθορίζεται από τα $2n$ υπερεπίπεδα διάστασης $(n - 1)$ που περιέχουν τις έδρες του. Έστω $\mathcal{U}$ το σύνολο των ορθογωνίων για τα οποία αυτά τα υπερεπίπεδα έχουν μέτρο 0 ως προς το μ και τα μ_{k_l}. Κάθε κορυφή ενός ορθογωνίου στο $\mathcal{U}$ είναι σημείο συνέχειας της F και το $\mathcal{U}$ είναι κλειστό ως προς πεπερασμένου πλήθους τομές.

Αφού

$$\mu(a, b] = \sum_{v \in V} sgn(v)\, F(v)$$

και ομοίως για το $\mu_{k_l}(a, b]$, η σύγκλιση $F_{k_l}(x) \longrightarrow F(x)$ συνεπάγεται ότι $\mu_{k_l}(A) \to \mu(A)$ για κάθε $A \in \mathcal{U}$. Αν $A_1, \dots, A_m \in \mathcal{U}$, το ίδιο συμβαίνει και με τις τομές αυτών, οπότε

$$\mu_{k_l}\left(\bigcup_{i=1}^{m} A_i\right) = \sum_i \mu_{k_l}(A_i) - \sum_{i,j} \mu_{k_l}\big(A_i \cap A_j\big) + \sum_{i,j,p} \mu_{k_l}\big(A_i \cap A_j \cap A_p\big) - \ \dots$$

$$\longrightarrow \sum_i \mu(A_i) - \sum_{i,j} \mu\big(A_i \cap A_j\big) + \sum_{i,j,p} \mu\big(A_i \cap A_j \cap A_p\big) - \ \dots = \mu\left(\bigcup_{i=1}^{m} A_i\right).$$

Παρατηρούμε ότι τα υπερεπίπεδα που επιτρέπουμε σε κάθε κατεύθυνση αποτελούν πυκνό υποσύνολο όλων των παράλληλων υπερεπιπέδων αφού εξαιρούμε ένα αριθμήσιμο πλήθος υπερεπιπέδων. Αν $G \subseteq \mathbb{R}^n$ ανοικτό, τότε, με βάση την παραπάνω παρατήρηση και αφού ο $\mathbb{R}^n$ είναι διαχωρίσιμος, υπάρχει ακολουθία $\{A_i\}$ στοιχείων του $\mathcal{U}$ τέτοια ώστε

$$G = \bigcup_i A_i.$$

Έστω $\varepsilon > 0$. Επιλέγουμε m τέτοιο ώστε

$$\mu\left(\bigcup_{i\leq m} A_i\right) > \mu(G) - \varepsilon.$$

Έχουμε τότε

$$\mu(G) - \varepsilon < \mu\left(\bigcup_{i\leq m} A_i\right) = \lim_{l\to\infty} \mu_{k_l}\left(\bigcup_{i\leq m} A_i\right) \leq \liminf_{l\to\infty} \mu_{k_l}(G)\,.$$

Αφού το ε ήταν αυθαίρετο, έχουμε $\liminf_{l\to\infty} \mu_{k_l}(G) \geq \mu(G)$ για κάθε ανοικτό $G \subseteq \mathbb{R}^n$, οπότε σύμφωνα με την πρόταση 28, αυτό ισοδυναμεί με την ασθενή σύγκλιση $\mu_{k_l} \overset{w}{\to} \mu$.

<u>Βήμα 2ο:</u> $X = \mathbb{R}^\infty$

Εάν $\mathcal{P}$ είναι μια σφικτή οικογένεια στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$, τότε, από το λήμμα 32 για κάθε $n \in \mathbb{N}$, η οικογένεια $\mathcal{P}' = \{\mu\pi_n^{-1} :\ \mu \in \mathcal{P}\}$ είναι σφικτή στον $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, όπου $\pi_n : \mathbb{R}^\infty \longrightarrow \mathbb{R}^n$ είναι η προβολή $\pi_n(x) = (x_1, \dots, x_n)$.

Από το 1ο βήμα που μόλις αποδείξαμε για τον $\mathbb{R}^n$, προκύπτει ότι από μια δοσμένη ακολουθία $(\mu_k) \in \mathcal{P}$ μπορούμε να βρούμε μια υπακολουθία αυτής (μ_{k_l}) τέτοια ώστε η $\mu_{k_l}\pi_n^{-1}$ να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ_n στον $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ για κάθε $n \in \mathbb{N}$. Η επιλογή της ακολουθίας (μ_{k_l}) μπορεί να γίνει ως εξής:

$$(27)\qquad \begin{cases} \mu_{k_{11}}\pi_1^{-1}\,,\ \ \mu_{k_{12}}\pi_1^{-1}\,,\ \ \mu_{k_{13}}\pi_1^{-1}\,, & \dots \quad \overset{w}{\to} \mu_1 \\ \mu_{k_{21}}\pi_2^{-1}\,,\ \ \mu_{k_{22}}\pi_2^{-1}\,,\ \ \mu_{k_{23}}\pi_2^{-1}\,, & \dots \quad \overset{w}{\to} \mu_2 \\ \mu_{k_{31}}\pi_3^{-1}\,,\ \ \mu_{k_{32}}\pi_3^{-1}\,,\ \ \mu_{k_{33}}\pi_3^{-1}\,, & \dots \quad \overset{w}{\to} \mu_3 \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \end{cases}$$

Στην 1η γραμμή της (27), επιλέγουμε υπακολουθία $(\mu_{k_{1l}})$ της (μ_k) έτσι ώστε η $\mu_{k_{1l}}\pi_1^{-1}$ να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ_1 στον $(\mathbb{R}^1, \mathcal{B}_{\mathbb{R}^1})$. Στην 2η γραμμή της (27), επιλέγουμε υπακολουθία $(\mu_{k_{2l}})$ της $(\mu_{k_{1l}})$ έτσι ώστε η $\mu_{k_{2l}}\pi_2^{-1}$ να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ_2 στον $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$. Συνεχίζοντας με τον ίδιο τρόπο, για κάθε $n \in \mathbb{N}$, στην $n-$οστή γραμμή της (27) επιλέγουμε υπακολουθία $(\mu_{k_{nl}})$ της $\left(\mu_{k_{(n-1)l}}\right)$ έτσι ώστε η $\mu_{k_{nl}}\pi_n^{-1}$ να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ_n στον $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$.

Μέσω του διαγώνιου επιχειρήματος, επιλέγουμε την υπακολουθία $(\mu_{k_l}) = (\mu_{k_{11}}, \mu_{k_{22}}, \mu_{k_{33}}, \dots)$. Συμβολίζουμε $k_l = k_{ll}$, οπότε η (μ_{k_l}) είναι υπακολουθία της (μ_k). Για κάθε $n \in \mathbb{N}$, η ακολουθία $\left(\mu_{k_n}, \mu_{k_{(n+1)}}, \dots\right)$ είναι υπακολουθία της $n-$οστής γραμμής της (27), άρα η $\mu_{k_l}\pi_n^{-1}$ συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ_n στον $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$.

Επομένως, η ασθενής σύγκλιση $\mu_{k_l}\pi_n^{-1} \xrightarrow{w} \mu_n$ ισχύει για κάθε $n \in \mathbb{N}$ ταυτόχρονα.

Έστω $\psi_n: \mathbb{R}^n \to \mathbb{R}^{n-1}$ η προβολή που ορίζεται ως $\psi_n(x_1, \dots, x_n) = (x_1, \dots, x_{n-1})$. Από τη σχέση $\mu_{k_l}\pi_{n-1}^{-1} = \mu_{k_l}\pi_n^{-1}\psi_n^{-1}$ και το Θεώρημα 33, βλέπουμε ότι τα μέτρα μ_n ικανοποιούν τη συνθήκη $\mu_{n-1} = \mu_n\psi_n^{-1}$, $n > 1$. Άρα, από το Θεώρημα Kolmogorov, υπάρχει μοναδικό μέτρο πιθανότητας Q στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ τέτοιο ώστε $Q\pi_n^{-1} = \mu_n$, $n \in \mathbb{N}$.

Τότε, όμως, για κάθε $n \in \mathbb{N}$ έχουμε $\mu_{k_l}\pi_n^{-1} \xrightarrow{w} Q\pi_n^{-1}$ από όπου συνεπάγεται ότι $\mu_{k_l} \xrightarrow{w} Q$.

Πράγματι, από τη συνέχεια της π_n έπεται ότι για $H \subseteq \mathbb{R}^n$ έχουμε $\partial(\pi_n^{-1}H) \subseteq \pi_n^{-1}\partial H$. Για την απόδειξη του αντίστροφου εγκλεισμού, παρατηρούμε ότι αν $x \in \pi_n^{-1}\partial H$, δηλαδή $\pi_n x \in \partial H$, τότε υπάρχουν σημεία $a^{(u)} \in H$ και σημεία $\beta^{(u)} \in H^c$ τέτοια ώστε $a^{(u)} \to \pi_n x$ και $\beta^{(u)} \to \pi_n x$ καθώς $u \to \infty$. Αφού τα σημεία $\left(\alpha_1^{(u)}, \dots, \alpha_n^{(u)}, x_{n+1}, \dots\right)$ ανήκουν στο $\pi_n^{-1}H$ και συγκλίνουν στο x, και αφού τα σημεία $\left(\beta_1^{(u)}, \dots, \beta_n^{(u)}, x_{n+1}, \dots\right)$ ανήκουν στο $(\pi_n^{-1}H)^c$ και επίσης συγκλίνουν στο x, ισχύει $x \in \partial(\pi_n^{-1}H)$. Άρα, $\partial(\pi_n^{-1}H) = \pi_n^{-1}\partial H$. Αφού $\mu_{k_l}\pi_n^{-1} \xrightarrow{w} Q\pi_n^{-1}$, ισχύει $\mu_{k_l}(A) \to Q(A)$ για σύνολα $A = \pi_n^{-1}H$ με $H \in \mathcal{B}_{\mathbb{R}^n}$ και $Q(\pi_n^{-1}\partial H) = 0$. Αφού $Q(\pi_n^{-1}\partial H) = 0$ ισοδυναμεί με $Q\big(\partial(\pi_n^{-1}H)\big) = 0$, η συγκλιση $\mu_{k_l}(A) \to Q(A)$ ισχύει για όλα τα πεπερασμένης διάστασης σύνολα συνέχειας ως προς το Q. Αφού τα πεπερασμένης διάστασης σύνολα σχηματίζουν μια κλάση που καθορίζει την ασθενή σύγκλιση, ισχύει $\mu_{k_l} \xrightarrow{w} Q$.

Τελικά, η σφικτότητα συνεπάγεται τη σχετική συμπάγεια στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$.

Βήμα 3ο: $X = \sigma-$συμπαγής

Εάν ο X είναι $\sigma-$συμπαγής, τότε είναι διαχωρίσιμος και συνεπώς είναι ομοιομορφικός με κάποιο υποσύνολο του $\mathbb{R}^\infty$ (βλ. Παρατήρηση μετά την απόδειξη του Θεωρήματος). Αφού ο X είναι $\sigma-$συμπαγής, το ίδιο συμβαίνει και με την εικόνα του μέσω του ομοιομορφισμού. Ειδικότερα, η εικόνα αυτή είναι ένα Borel υποσύνολο του $\mathbb{R}^\infty$. Άρα, ο X είναι ομοιομορφικός με ένα Borel υποσύνολο του $\mathbb{R}^\infty$. Η ασθενής σύγκλιση παραμένει κάτω από τον ομοιομορφισμό και το ίδιο συμβαίνει με τη σχετική συμπάγεια της $\mathcal{P}$. Μπορούμε, επομένως, να αντικαταστήσουμε τον X με την ομοιομορφική του εικόνα και συνεπώς να υποθέσουμε ότι ο X είναι ένα Borel υποσύνολο του $\mathbb{R}^\infty$.

Εάν η οικογένεια $\mathcal{P}$ είναι σφικτή στον $(X, \mathcal{B}_X)$, τότε εφαρμόζοντας το λήμμα 34 στον $\mathbb{R}^\infty$ και στο υποσύνολο αυτού X, η οικογένεια $\mathcal{P}^e = \{\mu^e : \ \mu \in \mathcal{P}\}$ είναι σφικτή στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$. Όπως έχουμε ήδη αποδείξει στο 2ο βήμα, η $\mathcal{P}^e$ είναι σχετικά συμπαγής. Για κάθε ακολουθία (μ_k) στην $\mathcal{P}$, η αντίστοιχη ακολουθία (μ_k^e) στην $\mathcal{P}^e$ περιέχει υπακολουθία $\left(\mu_{k_l}^e\right)$ η οποία συγκλίνει ασθενώς στον $(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}^\infty})$ σε κάποιο μέτρο Q.

Από τη σφικτότητα της $\mathcal{P}$, έχουμε ότι για κάθε $\varepsilon > 0$ υπάρχει συμπαγές υποσύνολο $K \subseteq X$ τέτοιο ώστε $\mu_{k_l}^e(K) = \mu_{k_l}(K) > 1 - \varepsilon$, $l \in \mathbb{N}$ και επομένως,

$$Q(X) \geq Q(K) \geq \limsup_{l\to\infty} \mu^e_{k_l}(K) > 1-\varepsilon.$$

Άρα, $Q(X)=1$ και $\mu^e_{k_l}(X)=1,\ \ l\in\mathbb{N}$ και άρα από το λήμμα 35 και τη σχέση $(\mu^e)^r=\mu$ όταν μ είναι μέτρο πιθανότητας στον $(X_0,\mathcal{B}_{X_0})$, έχουμε $\mu_{k_l} \overset{w}{\to} Q^r$ στον $(X,\mathcal{B}_X)$.

Άρα, η σφικτότητα συνεπάγεται τη σχετική συμπάγεια και στην περίπτωση που ο X είναι σ −συμπαγής.

<u>Βήμα 4ο :</u> Ο X είναι αυθαίρετος μετρικός χώρος (η γενική περίπτωση).

Αφού η οικογένεια $\mathcal{P}$ είναι σφικτή στον X, για κάθε $i\in\mathbb{N}$ υπάρχει συμπαγές σύνολο $K_i\subseteq X$ τέτοιο ώστε

$$\mu(K_i) > 1-\frac{1}{i}\ ,\ \ \text{για κάθε } \mu\in\mathcal{P}.$$

Θέτουμε

$$X_0 = \bigcup_{i\in\mathbb{N}} K_i.$$

Τότε $\mu(X_0)=1,\ \ \forall\mu\in\mathcal{P}$ και η οικογένεια $\mathcal{P}^r=\{\mu^r:\ \ \mu\in\mathcal{P}\}$ είναι σφικτή στον $(X_0,\mathcal{B}_{X_0})$. Από το 3ο βήμα που μόλις αποδείξαμε, η $\mathcal{P}^r$ είναι σχετικά συμπαγής. Άρα, για κάθε ακολουθία (μ_k) στην $\mathcal{P}$, για την αντίστοιχη ακολουθία (μ^r_k) υπάρχει υπακολουθία $(\mu^r_{k_l})$ και μέτρο Q έτσι ώστε $\mu^r_{k_l} \overset{w}{\to} Q$ στον $(X_0,\mathcal{B}_{X_0})$.

Από το λήμμα 34 και τη σχέση $(\mu^r)^e$ όταν μ είναι ένα μέτρο πιθανότητας στον $(X,\mathcal{B}_X)$ με $\mu(X_0)=1$, έχουμε ότι $\mu_{k_l} \overset{w}{\to} Q^e$ στον $(X,\mathcal{B}_X)$.

Άρα, η $\mathcal{P}$ είναι σχετικά συμπαγής και το θεώρημα έχει αποδειχτεί πλήρως.

■

Παρατήρηση. Κάθε διαχωρίσιμος μετρικός χώρος (X,ρ) είναι ομοιομορφικός με ένα υποσύνολο του $\mathbb{R}^\infty$.

<u>Απόδειξη:</u>

Έστω $\{d_1,d_2,\dots\}$ μια ακολουθία σημείων πυκνή στον X. Ορίζουμε την απεικόνιση $h: X\longrightarrow\mathbb{R}^\infty$ ως εξής: $h(x)=(\rho(x,d_1),\rho(x,d_2),\dots),\quad x\in X,$

Έστω $x\in X$ και (x_n) μια ακολουθία με $x_n\to x$. Τότε, $\lim_{n\to\infty}\rho(x_n,d_k)=\rho(x,d_k),\ \ k\in\mathbb{N}$

οπότε, εφόσον η τοπολογία του $\mathbb{R}^\infty$ είναι αυτή της σύγκλισης κατά συντεταγμένη, έχουμε $h(x_n)\to h(x)$. Άρα, η h είναι συνεχής.

Στη συνέχεια, υποθέτουμε ότι (x_n) είναι μια ακολουθία που δεν συγκλίνει στο x. Τότε, υπάρχει $\varepsilon > 0$ και υπακολουθία $\left(x_{n_l}\right)$ έτσι ώστε $\rho\left(x_{n_l}, x\right) > \varepsilon$. Αφού η ακολουθία (d_k) είναι πυκνή στον X, υπέρχει $k_0 \in \mathbb{N}$ έτσι ώστε $\rho\left(x, d_{k_0}\right) < \frac{\varepsilon}{2}$. Τότε,

$$\rho\left(x_{n_l}, d_{k_0}\right) > \frac{\varepsilon}{2}, \quad \text{για κάθε } l \in \mathbb{N}.$$

Άρα, η ακολουθία $\rho\left(x_n, d_{k_0}\right)$ δεν μπορεί να συγκλίνει στο $\rho\left(x, d_{k_0}\right)$ και άρα η $h(x_n)$ δεν μπορεί να συγκλίνει στο $h(x)$. Αποδείξαμε, λοιπόν, ότι: $h(x_n) \to h(x) \implies x_n \to x$.

Ομοίως (παίρνοντας $x_n \equiv y$), προκύπτει ότι $h(x) = h(y) \implies x = y$.

Επομένως, η $h : X \longrightarrow h(X)$ είναι συνεχής αμφιμονοσήμαντη απεικόνιση της οποίας η αντίστροφη είναι, επίσης, συνεχής. Δηλαδή, ο X είναι ομοιομορφικός με το υποσύνολο $h(X)$ του $\mathbb{R}^\infty$.

■

Πρόταση 36

Αν ο (X, d) είναι ένας μετρικός χώρος και F είναι μια μη αρνητική κάτω ημισυνεχής συνάρτηση στον X, τότε αυτή μπορεί να γραφεί ως το supremum μιας αύξουσας ακολουθίας ομοιόμορφα συνεχών μη αρνητικών συναρτήσεων.

<u>Απόδειξη:</u>

Ορίζουμε $F_n(x) = \inf_{y \in X}[F(y) + nd(x, y)]$.

Κάθε F_n είναι μια καλώς ορισμένη (αφού η F είναι κάτω φραγμένη από το 0) μη αρνητική συνάρτηση. Επίσης, $\forall x \in X: \ F_n(x) \le F_{n+1}(x)$, δηλαδή η ακολουθία $(F_n)_{n=1}^{\infty}$ είναι αύξουσα.

Κάθε F_n είναι ομοιόμορφα συνεχής συνάρτηση. Πράγματι, έστω $\varepsilon > 0$. Ζητάμε να βρούμε $\delta > 0$ τέτοιο ώστε $d(x_1, x_2) < \delta \implies |F_n(x_1) - F_n(x_2)| < \varepsilon$. Λόγω της τριγωνικής ανισότητας, για κάθε $y \in X$ και $x_1 \neq x_2$ ισχύει $d(x_1, y) - d(x_2, y) \le d(x_1, x_2)$. Άρα, για κάθε $n \in \mathbb{N}$, $nd(x_1, y) - nd(x_2, y) \le nd(x_1, x_2)$ και προσθαφαιρώντας το $F(y)$,

$$\left(F(y) + nd(x_1, y)\right) - \left(F(y) + nd(x_2, y)\right) \le nd(x_1, x_2).$$

Επίσης, εξ ορισμού του infimum, υπάρχει $y_0 \in X$ τέτοιο ώστε

$$F(y_0) + nd(x_2, y_0) < \inf_{y \in X}\left(F(y) + nd(x_2, y)\right) + nd(x_1, x_2)$$

Άρα,

$$\begin{aligned}
&\inf_{y\in X}\big(F(y)+nd(x_1,y)\big)-\inf_{y\in X}\big(F(y)+nd(x_2,y)\big)\\
&\qquad< \inf_{y\in X}\big(F(y)+nd(x_1,y)\big)-\big(F(y_0)+nd(x_2,y_0)\big)+nd(x_1,x_2)\\
&\qquad\le \big(F(y_0)+nd(x_1,y_0)\big)-\big(F(y_0)+nd(x_2,y_0)\big)+nd(x_1,x_2)\\
&\qquad= n[d(x_1,y_0)-d(x_2,y_0)+d(x_1,x_2)]\le 2nd(x_1,x_2).
\end{aligned}$$

Δηλαδή, για κάθε $x_1,x_2\in X$ με $x_1\neq x_2$ έχουμε $F_n(x_1)-F_n(x_2)<2nd(x_1,x_2)$, οπότε, προκειμένου να ισχύει $|F_n(x_1)-F_n(x_2)|<\varepsilon$, αρκεί να επιλέξουμε $\delta=\frac{\varepsilon}{2n}$.

Τότε, για $x_1,x_2\in X$ με $d(x_1,x_2)<\delta$, θα ισχύει

$$|F_n(x_1)-F_n(x_2)|<2n\delta=2n\cdot\frac{\varepsilon}{2n}=\varepsilon.$$

Μένει, ακόμα, να αποδείξουμε ότι $F(x)=\sup_{n\in\mathbb{N}}F_n(x)$, για κάθε $x\in X$. Πράγματι, έστω $x\in X$. Τότε, για κάθε $n\in\mathbb{N}$, $F_n(x)=\inf_{y\in X}[F(y)+nd(x,y)]\le F(x)$ αρκεί να επιλέξουμε $y=x$ στο infimum. Συνεπώς, $\sup_{n\in\mathbb{N}}F_n(x)\le F(x)$. Έστω, ακόμα, $\varepsilon>0$. Για την απόδειξη της αντίστροφης ανισότητας, θα βρούμε $n_0\in\mathbb{N}$ τέτοιο ώστε $F_{n_0}(x)\ge F(x)-\varepsilon$. Υπάρχει $n_0\in\mathbb{N}$ έτσι ώστε

$$n_0d(x,y)\ge F(x)-F(y)-\varepsilon,\quad \text{για κάθε } y\in X\setminus\{x\}. \tag{28}$$

Πράγματι, αρκεί να αποδείξουμε ότι η συνάρτηση

$$\frac{F(x)-F(y)-\varepsilon}{d(x,y)}\,,\quad y\in X\setminus\{x\}$$

είναι άνω φραγμένη. Έστω ότι αυτό δε συμβαίνει. Τότε υπάρχει ακολουθία (y_k) στο X με $y_k\neq x$ και

$$\frac{F(x)-F(y_k)-\varepsilon}{d(x,y_k)}\longrightarrow+\infty. \tag{29}$$

Τότε, θα ισχύει επίσης,

$$\frac{F(x)-\varepsilon}{d(x,y_k)}\longrightarrow+\infty\,,$$

αφού η F είναι μη αρνητική. Άρα $d(x,y_k)\to 0$, δηλαδή $y_k\to x$. Τότε, όμως, λόγω κάτω ημισυνέχειας της F, θα έχουμε $F(x)\le\liminf_{k\to+\infty}F(y_k)$, άρα $\limsup_{k\to+\infty}(F(x)-F(y_k)-\varepsilon)\le-\varepsilon<0$ κάτι που έρχεται σε αντίφαση με την (29).

Επιστρέφοντας στη σχέση (28) και προσθέτοντας $F(y)$, έχουμε

$$F(y)+n_0d(x,y)\ge F(x)-\varepsilon,\quad\forall y\in X$$

οπότε και

$$F_{n_0}(x) = \inf_{y \in X}[F(y) + n_0 d(x, y)] \geq F(x) - \varepsilon,$$

όπως επιθυμούσαμε.

∎

ΚΕΦΑΛΑΙΟ 1: Η Δυϊκότητα Kantorovich

Στην 1η ενότητα του κεφαλαίου, θα παρουσιάσουμε το πρόβλημα μεταφοράς μάζας του Kantorovich ενώ στη συνέχεια, θα μελετήσουμε και θα αποδείξουμε το θεώρημα δυϊκότητας, το οποίο αποδείχθηκε για πρώτη φορά από τον Kantorovich. Προκειμένου να τονίσουμε τη γενικότητα αυτής της αρχής, η οποία δεν απαιτεί κάποια γεωμετρική δομή, θα δουλέψουμε σε αφηρημένους Πολωνικούς χώρους (δηλ. πλήρεις και διαχωρίσιμους μετρικούς χώρους). Αργότερα, στη 2η ενότητα του κεφαλαίου, θα ασχοληθούμε με την περίπτωση που η συνάρτηση κόστους είναι μια μετρική. Στην 3η ενότητα, θα παρουσιάσουμε μια ειδική περίπτωση του θεωρήματος δυϊκότητας του Kantorovich δίνοντας μια διαφορετική απόδειξη, ενώ στην 4η και τελευταία ενότητα του κεφαλαίου, θα μελετήσουμε την περίπτωση που η συνάρτηση κόστους παίρνει τιμές στο σύνολο $\{0, 1\}$.

1.1. Γενική Δυϊκότητα

Κεντρικό ρόλο στην πρώτη ενότητα του κεφαλαίου, κατέχει το Θεώρημα Δυϊκότητας του Kantorovich (Θεώρημα 1.3). Θα παρουσιάσουμε, ακόμα, μερικές βασικές προτάσεις που στόχο έχουν να μας οδηγήσουν στην απόδειξη αυτού.

1.1.1. Ορισμοί και προκαταρκτικά.

Έστω $(X, \mathcal{B}_X)$ και $(Y, \mathcal{B}_Y)$ δύο τοπολογικοί χώροι εφοδιασμένοι με την Borel σ −άλγεβρα, μ και ν δύο (Borel) μέτρα πιθανότητας ορισμένα στους X, Y αντίστοιχα, και έστω $c: X \times Y \to [0, +\infty]$ μια μετρήσιμη συνάρτηση. Ορίζουμε το σύνολο $\Pi(\mu, \nu)$ ως το σύνολο των μέτρων πιθανότητας στον $X \times Y$ που έχουν περιθώρια μέτρα μ στον X και ν στον Y. Στον χώρο $X \times Y$ θεωρούμε τη σ −άλγεβρα που γεννάται από τα ορθογώνια $A \times B$, όπου $A \in \mathcal{B}_X$ και $B \in \mathcal{B}_Y$.

Ειδικότερα, $\pi \in \Pi(\mu, \nu)$ αν και μόνο αν το π είναι ένα μη αρνητικό μέτρο που ικανοποιεί τις σχέσεις

$$\pi[A \times Y] = \mu[A], \qquad \pi[X \times B] = \nu[B]. \tag{1.1}$$

για κάθε $A \in \mathcal{B}_X$ και $B \in \mathcal{B}_Y$.

Σημειώνουμε ότι αυτός ο ορισμός εξαναγκάζει το π να είναι ένα μέτρο πιθανότητας, αφού $\pi[X \times Y] = \mu[X] = 1$.

Το πρόβλημα μεταφοράς μάζας του Kantorovich αφορά την ελαχιστοποίηση πάνω στο σύνολο $\Pi(\mu,\nu)$ του γραμμικού συναρτησοειδούς

$$I[\pi] = \int_{X\times Y} c(x,y)d\pi(x,y).$$

Ένα οποιοδήποτε μέτρο $\pi \in \Pi(\mu,\nu)$ λέγεται αποδεκτό πλάνο μεταφοράς για το πρόβλημα Kantorovich. Το σύνολο $\Pi(\mu,\nu)$ είναι μη κενό, αφού το μέτρο γινόμενο $\mu\otimes\nu$ ανήκει στο $\Pi(\mu,\nu)$. Επίσης, το $\Pi(\mu,\nu)$ είναι κυρτό σύνολο.

Πράγματι, έστω $\pi_1,\pi_2 \in \Pi(\mu,\nu)$ και $\lambda\in[0,1]$. Τότε, για κάθε $A\in\mathcal{B}_X$, έχουμε

$$(\lambda\pi_1+(1-\lambda)\pi_2)[A\times Y] = \lambda\pi_1[A\times Y]+(1-\lambda)\pi_2[A\times Y] = \lambda\mu[A]+(1-\lambda)\mu[A] = \mu[A].$$

Εντελώς ανάλογα, για κάθε $B\in\mathcal{B}_Y$, έχουμε $(\lambda\pi_1+(1-\lambda)\pi_2)[X\times B]=\nu[B]$.

Άρα, $\lambda\pi_1+(1-\lambda)\pi_2 \in \Pi(\mu,\nu)$ και το $\Pi(\mu,\nu)$ είναι κυρτό.

Θα αποδείξουμε ότι ο ορισμός του $\Pi(\mu,\nu)$ μέσω της σχέσης (1.1) είναι ισοδύναμος με τον εξής:

$\pi\in\Pi(\mu,\nu)$ αν και μόνο αν το π είναι ένα μη αρνητικό μέτρο στον $X\times Y$ τέτοιο ώστε, για όλα τα ζεύγη μετρήσιμων συναρτήσεων $(\varphi,\psi)\in L^1(d\mu)\times L^1(d\nu)$, ή ισοδύναμα $(\varphi,\psi)\in L^\infty(d\mu)\times L^\infty(d\nu)$, να ισχύει

$$\int_{X\times Y}[\varphi(x)+\psi(y)]d\pi(x,y) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu. \tag{1.2}$$

<u>Απόδειξη:</u>

Για το ευθύ, έστω ότι το π είναι ένα μη αρνητικό μέτρο που ικανοποιεί την (1.1).

Αν $\varphi = 1_A,\ \psi = 1_B$ όπου $A\in\mathcal{B}_X$ και $B\in\mathcal{B}_Y$, τότε:

$$\begin{aligned}\int_{X\times Y}[\varphi(x)+\psi(y)]\,d\pi(x,y) &= \int_{X\times Y}1_A(x)\,d\pi + \int_{X\times Y}1_B(y)\,d\pi\\ &= \int_{X\times Y}1_{A\times Y}(x,y)\,d\pi + \int_{X\times Y}1_{X\times B}(x,y)\,d\pi = \pi[A\times Y]+\pi[X\times B] = \mu[A]+\nu[B]\\ &= \int_X\varphi\,d\mu + \int_Y\psi\,d\nu.\end{aligned}$$

Αν

$$\varphi = \sum_{j=1}^{n} a_j 1_{A_j} \quad \text{και} \quad \psi = \sum_{i=1}^{m} b_i 1_{B_i}$$

είναι δύο απλές συναρτήσεις, τότε:

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x,y) = \int_{X\times Y} \sum_{j=1}^{n} a_j 1_{A_j}(x)\, d\pi + \int_{X\times Y} \sum_{i=1}^{m} b_i 1_{B_i}(y)\, d\pi$$
$$= \sum_{j=1}^{n} a_j \int_{X\times Y} 1_{A_j}(x) d\pi + \sum_{i=1}^{m} b_i \int_{X\times Y} 1_{B_i}(y) d\pi = \sum_{j=1}^{n} a_j \cdot \mu[A_j] + \sum_{i=1}^{m} b_i \cdot \nu[B_i]$$
$$= \int_X \varphi\, d\mu + \int_Y \psi\, d\nu.$$

Αν $\varphi \geq 0,\ \ \psi \geq 0,$ τότε υπάρχουν ακολουθίες απλών συναρτήσεων $0 \leq s_n,\ t_n$ τέτοιες ώστε $s_n(x) \nearrow \varphi(x)$ και $t_n(y) \nearrow \psi(y)$. Τότε, χρησιμοποιώντας το Θεώρημα Μονότονης Σύγκλισης:

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x,y) = \lim_{n\to\infty} \int_{X\times Y} [s_n(x) + t_n(y)] d\pi(x,y) = \lim_{n\to\infty} \int_X s_n d\mu + \lim_{n\to\infty} \int_Y t_n d\nu$$
$$= \int_X \varphi\, d\mu + \int_Y \psi\, d\nu.$$

Αν φ, ψ είναι ολοκληρώσιμες συναρτήσεις, τότε $\varphi = \varphi^+ - \varphi^-$ και $\psi = \psi^+ - \psi^-$, οπότε:

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x,y) = \int_{X\times Y} [\varphi^+(x) + \psi^+(y)] d\pi - \int_{X\times Y} [\varphi^-(x) + \psi^-(y)] d\pi$$
$$= \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu + \int_Y \psi^+ d\nu - \int_Y \psi^- d\nu = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu.$$

Για το αντίστροφο, παίρνοντας $\varphi = 1_A,\ \ \psi = 1_B$ όπου $A \in \mathcal{B}_X,\ \ B \in \mathcal{B}_Y$, έχουμε:

$$\pi[A \times Y] = \int_{X\times Y} [1_A(x) + 0] d\pi(x,y) = \int_X 1_A(x) d\mu(x) + 0 = \mu[A].$$

$$\pi[X \times B] = \int_{X\times Y} [0 + 1_B(y)] d\pi(x,y) = 0 + \int_Y 1_B(y) d\nu(y) = \nu[B].$$

Εφόσον, λοιπόν, η σχέση (1.2) ικανοποιείται για τις συναρτήσεις $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$, θα ικανοποιείται και για τα ζεύγη $(\varphi,\psi) \in L^\infty(d\mu) \times L^\infty(d\nu)$, αφού $L^\infty(d\mu) \times L^\infty(d\nu) \subseteq L^1(d\mu) \times L^1(d\nu)$. Στο αντίστροφο, οι συναρτήσεις $\varphi = 1_A$ και $\psi = 1_B$ που χρησιμοποιήσαμε, προφανώς είναι φραγμένες. Έτσι, λοιπόν, μπορούμε στην (1.2) να θεωρούμε είτε ζεύγη $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$ είτε ζεύγη $(\varphi,\psi) \in L^\infty(d\mu) \times L^\infty(d\nu)$.

■

Στον τελευταίο ορισμό, η κλάση των δοκιμαστικών συναρτήσεων είναι αρκετά ευρεία, αφού $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$. Θα ήταν πιο βολικό να χρησιμοποιήσουμε μικρότερες κλάσεις τέτοιων συναρτήσεων. Αυτό μπορεί να συμβεί εάν οι χώροι $(X, \mathcal{B}_X, \mu)$ και $(Y, \mathcal{B}_Y, \nu)$ έχουν κάποιες επιπλέον ιδιότητες.

Ισχυρισμός 1.

Αν οι X, Y είναι Πολωνικοί χώροι, δηλαδή πλήρεις και διαχωρίσιμοι μετρικοί χώροι, και τα μ, ν είναι Borel μέτρα πιθανότητας, τότε αρκεί να υποθέσουμε τη σχέση (1.2) για ζεύγη συναρτήσεων $(\varphi, \psi) \in C_b(X) \times C_b(Y)$.

Απόδειξη:

Αρκεί να δείξουμε ότι ο $C_b(X) \times C_b(Y)$ είναι πυκνός στον $L^1(d\mu) \times L^1(d\nu)$ ως προς την L^1 −νόρμα. Δηλαδή, αν $\varepsilon > 0$ και $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$, τότε υπάρχει ζεύγος συναρτήσεων $(f, g) \in C_b(X) \times C_b(Y)$ έτσι ώστε $\|\varphi - f\|_{L^1} < \varepsilon$ και $\|\psi - g\|_{L^1} < \varepsilon$.

Έχοντας αποδείξει αυτό και υποθέτοντας ότι το ζεύγος (f, g) ικανοποιεί τη σχέση (1.2), δηλαδή

$$\int_{X\times Y} [f(x) + g(y)] d\pi = \int_X f\, d\mu + \int_Y g\, d\nu,$$

θα ισχύει

$$\left| \int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi - \int_{X\times Y} [f(x) + g(y)]\, d\pi \right| =$$

$$\left| \int_{X\times Y} [\varphi(x) - f(x)]\, d\pi + \int_{X\times Y} [\psi(y) - g(y)]\, d\pi \right| =$$

$$\left| \int_X [\varphi(x) - f(x)]\, d\mu + \int_Y [\psi(y) - g(y)]\, d\nu \right| \le \|\varphi - f\|_{L^1} + \|\psi - g\|_{L^1} < 2\varepsilon.$$

Ομοίως,

$$\left| \int_X f\, d\mu + \int_Y g\, d\nu - \left(\int_X \varphi\, d\mu + \int_Y \psi\, d\nu \right) \right| < 2\varepsilon.$$

Συνεπώς,

$$\left| \int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi - \left(\int_X \varphi\, d\mu + \int_Y \psi\, d\nu \right) \right| < 4\varepsilon,$$

και εφόσον το ε ήταν ένας τυχαίος θετικός αριθμός, έχουμε

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu.$$

Αρκεί, επομένως, να υποθέτουμε τη σχέση (1.2) για ζεύγη συναρτήσεων $(\varphi, \psi) \in C_b(X) \times C_b(Y)$.

Για την απόδειξη της πυκνότητας του $C_b(X)$ στον $L^1(d\mu)$, έστω $\varepsilon > 0$ και $A \in \mathcal{B}_X$. Εφόσον ο X είναι Πολωνικός χώρος, το μ είναι σφικτό (Πρόταση 25, Εισαγωγικό κεφάλαιο), δηλαδή υπάρχουν $K \subseteq X$ συμπαγές και $U \subseteq X$ ανοικτό τέτοια ώστε $K \subseteq A \subseteq U$ και $\mu[U - K] < \varepsilon$.

Από το λήμμα του Urysohn, υπάρχει $f \in C_b(X)$ τέτοια ώστε $1_K \leq f \leq 1_U$. Επειδή για το A ισχύει, επίσης, $1_K \leq 1_A \leq 1_U$ έχουμε

$$\|f - 1_A\|_{L^1} = \int_X |f - 1_A|\, d\mu \leq \mu[U - K] < \varepsilon.$$

Στη συνέχεια, αν

$$\varphi = \sum_{i=1}^{n} a_i\, 1_{A_i}$$

είναι μια απλή συνάρτηση, τότε βρίσκουμε συναρτήσεις $f_i \in C_b(X)$, $i = 1,2,\dots,n$ ώστε $\left\|f_i - 1_{A_i}\right\|_{L^1} < \varepsilon$. Τότε, για τη συνάρτηση

$$f = \sum_{i=1}^{n} a_i f_i$$

ισχύει $f \in C_b(X)$ και

$$\|f - \varphi\|_{L^1} = \left\|\sum_{i=1}^{n} a_i\left(f_i - 1_{A_i}\right)\right\|_{L^1} \leq \sum_{i=1}^{n} |a_i|\, \left\|f_i - 1_{A_i}\right\|_{L^1} \leq \varepsilon\left(\sum_{i=1}^{n} |a_i|\right).$$

Άρα, μπορούμε να προσεγγίσουμε την φ οσοδήποτε θέλουμε με συνάρτηση $f \in C_b(X)$.

Επίσης, εφόσον το σύνολο των απλών συναρτήσεων είναι πυκνό στον L^1, τελικά ο $C_b(X)$ είναι πυκνός στον $L^1(d\mu)$.

Ομοίως, αποδεικνύουμε ότι ο $C_b(Y)$ είναι πυκνός στον $L^1(d\nu)$. Άρα, ο $C_b(X) \times C_b(Y)$ είναι πυκνός στον $L^1(d\mu) \times L^1(d\nu)$.

■

Ισχυρισμός 2.

Αν οι X, Y είναι τοπικά συμπαγείς χώροι Hausdorff, δηλαδή για κάθε σημείο του χώρου υπάρχει περιοχή αυτού με συμπαγή κλειστότητα, τότε αρκεί να χρησιμοποιήσουμε στην (1.2) ζεύγη συναρτήσεων $(\varphi, \psi) \in C_0(X) \times C_0(Y)$.

Απόδειξη:

Η δυνατότητα για τον περιορισμό της κλάσης των δοκιμαστικών συναρτήσεων στην $C_0(X) \times C_0(Y)$, οφείλεται στο θεώρημα του Riesz (Θεώρημα 22, Εισαγωγικό Κεφάλαιο), το οποίο ταυτίζει τον χώρο $M(X)$ των κανονικών προσημασμένων μέτρων Borel που έχουν πεπερασμένη ολική κύμανση με τον τοπολογικό δυϊκό χώρο του $C_0(X)$.

Πράγματι, θέλουμε να δείξουμε ότι αν για κάθε $(f,g) \in C_0(X) \times C_0(Y)$ ισχύει

$$\int_{X\times Y} [f(x) + g(x)]\, d\pi = \int_X f\, d\mu + \int_Y g\, d\nu,$$

τότε τα μ, ν είναι τα περιθώρια μέτρα του π στους χώρους X, Y αντίστοιχα.

Έστω, λοιπόν, π_X και π_Y τα περιθώρια μέτρα του π. Τότε, σύμφωνα με την (1.2), για κάθε $(f,g) \in C_0(X) \times C_0(Y)$ ισχύει

$$\int_{X\times Y} [f(x) + g(x)]\, d\pi = \int_X f\, d\pi_X + \int_Y g\, d\pi_Y.$$

Επιλέγοντας $g \equiv 0$ στις παραπάνω δύο ισότητες, παίρνουμε

$$\int_X f\, d\mu = \int_X f\, d\pi_X, \quad \forall f \in C_0(X).$$

Τότε, σύμφωνα με το Θεώρημα του Riesz, τα μέτρα μ και π_X ορίζουν το ίδιο συναρτησοειδές, και συνεπώς ταυτίζονται. Ομοίως, τα μέτρα ν και π_Y ταυτίζονται.

■

Στη συνέχεια, θα αναφέρουμε δύο βασικά παραδείγματα που θα μας κάνουν να εκτιμήσουμε το περιεχόμενο των παραπάνω ισχυρισμών. Στο πρώτο παράδειγμα αποδεικνύουμε ότι ο $C([0,1];\mathbb{R})$ δεν είναι τοπικά συμπαγής, αφού κάθε συμπαγές υποσύνολο αυτού έχει κενό εσωτερικό, και συνεπώς δεν μπορεί να εφαρμοστεί το θεώρημα του Riesz για τον περιορισμό των δοκιμαστικών συναρτήσεων στο σύνολο C_0. Στο δεύτερο παράδειγμα, δείχνουμε ότι οι χώροι $L^\infty\big((0,1)\big)$ και $C_b(\mathbb{R}^n)$ δεν είναι διαχωρίσιμοι, επομένως ούτε Πολωνικοί.

Παράδειγμα 1.1

Έστω K ένα συμπαγές υποσύνολο του $C([0,1];\mathbb{R})$. Θα δείξουμε ότι $Int(K) = \emptyset$. Θα συμπεράνουμε, στη συνέχεια, ότι $C_0\big(C([01];\mathbb{R})\big) = \{0\}$, και ειδικότερα ότι το συμπέρασμα του θεωρήματος Riesz δεν εφαρμόζεται στον $C([0,1];\mathbb{R})$.

Απόδειξη:

Για να αποδείξουμε ότι $Int(K) = \emptyset$, αρκεί να δείξουμε ότι:

$\forall u \in K$ και $\forall \varepsilon > 0$ μπορούμε να κατασκευάσουμε συνάρτηση $v \in C([0,1];\mathbb{R}) - K$ τέτοια ώστε $\|u - v\|_\infty < \varepsilon$.

Έστω $\eta > 0$. Θα δείξουμε ότι υπάρχει $\delta > 0$ τέτοιο ώστε

$$|x-y| < \delta \implies \sup_{f\in K}|f(x) - f(y)| < \eta.$$

Πράγματι, αφού το K είναι συμπαγές, υπάρχουν μπάλες $B_i = B\left(f_i, \frac{\eta}{3}\right) \subseteq C([0,1];\mathbb{R})$, $i = 1,2,\dots,n$ έτσι ώστε

$$K \subseteq \bigcup_{i=1}^{n} B_i.$$

Εφόσον οι συναρτήσεις f_i είναι συνεχείς και ορισμένες στο συμπαγές σύνολο $[0,1]$, θα είναι ομοιόμορφα συνεχείς. Άρα, για κάθε $i = 1,2,\dots,n$ υπάρχει $\delta_i > 0$ τέτοιο ώστε

$$|x-y| < \delta_i \implies |f_i(x) - f_i(y)| < \frac{\eta}{3}.$$

Ορίζουμε $\delta = \min\{\delta_1, \delta_2, \dots, \delta_n\}$ οπότε έχουμε:

$$|x-y| < \delta \implies |f_i(x) - f_i(y)| < \frac{\eta}{3}, \quad i = 1,2,\dots,n.$$

Έστω, τώρα, $f \in K$. Υπάρχει $i \in \{1,2,\dots,n\}$ έτσι ώστε $f \in B\left(f_i, \frac{\eta}{3}\right)$, οπότε για κάθε $x, y \in [0,1]$ με $|x-y| < \delta$ έχουμε:

$$|f(x) - f(y)| \le |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| < \frac{\eta}{3} + \frac{\eta}{3} + \frac{\eta}{3} = \eta.$$

Έστω, λοιπόν, $u \in K$ και ε τυχαίος θετικός αριθμός. Για τον $\eta = \frac{\varepsilon}{4} > 0$, βρίσκουμε $0 < \delta \le 2$ έτσι ώστε

$$|x-y| < \delta \implies \sup_{f\in K}|f(x) - f(y)| < \frac{\varepsilon}{4}.$$

Ειδικότερα για τη συνάρτηση u έχουμε:

$$|x-y| < \delta \implies |u(x) - u(y)| < \frac{\varepsilon}{4}.$$

Κατασκευάζουμε συνάρτηση $v: [0,1] \longrightarrow \mathbb{R}$ για την οποία ισχύουν τα εξής:

$$v(0) = u(0),$$

στο διάστημα $\left[0, \frac{\delta}{2}\right]$ η v είναι γραμμική

$$v\left(\frac{\delta}{2}\right) = u\left(\frac{\delta}{2}\right) + \frac{\varepsilon}{2},$$

στο διάστημα $\left[\frac{\delta}{2},1\right]$ η v είναι συνεχής και ικανοποιεί τη σχέση

$$|v(x)-u(x)|<\frac{\varepsilon}{2}, \quad x\in\left(\frac{\delta}{2},1\right].$$

Τότε,

$$\|u-v\|_\infty=\sup_{x\in[0,1]}|u(x)-v(x)|<\varepsilon.$$

Πράγματι, για κάθε $x\in\left[0,\frac{\delta}{2}\right)$, $|u(x)-u(0)|<\frac{\varepsilon}{4}$ και επίσης, εφόσον η συνάρτηση v στο διάστημα αυτό είναι γραμμική, ισχύει

$$|v(x)-v(0)|<\left|v(0)-v\left(\frac{\delta}{2}\right)\right|=\left|u(0)-u\left(\frac{\delta}{2}\right)-\frac{\varepsilon}{2}\right|<\frac{\varepsilon}{4}+\frac{\varepsilon}{2}=\frac{3\varepsilon}{4}.$$

Συνεπώς,

$$|u(x)-v(x)|\le|u(x)-u(0)|+|v(0)-v(x)|<\frac{\varepsilon}{4}+\frac{3\varepsilon}{4}=\varepsilon.$$

Επίσης, στο διάστημα $\left[\frac{\delta}{2},1\right]$ εκ κατασκευής της συνάρτησης v ισχύει:

$$|u(x)-v(x)|\le\frac{\varepsilon}{2}.$$

Ακόμα, $v\notin K$, αφού δεν ικανοποιεί την ιδιότητα:

$$|x-y|<\delta\Longrightarrow|v(x)-v(y)|<\frac{\varepsilon}{4}.$$

Πράγματι,

$$\left|v\left(\frac{\delta}{2}\right)-v(0)\right|=\left|v\left(\frac{\delta}{2}\right)-u\left(\frac{\delta}{2}\right)+u\left(\frac{\delta}{2}\right)-v(0)\right|\ge\left|v\left(\frac{\delta}{2}\right)-u\left(\frac{\delta}{2}\right)\right|-\left|u\left(\frac{\delta}{2}\right)-u(0)\right|$$
$$>\frac{\varepsilon}{2}-\frac{\varepsilon}{4}=\frac{\varepsilon}{4}$$

το οποίο αντιφάσκει με την ιδιότητα:

$$\left|\frac{\delta}{2}-0\right|<\delta\Longrightarrow\left|v\left(\frac{\delta}{2}\right)-v(0)\right|<\frac{\varepsilon}{4}.$$

Έτσι, ολοκληρώνεται η απόδειξη του πρώτου σκέλους.

Για το δεύτερο σκέλος του παραδείγματος, έστω $f\in C_0\big(C([0,1];\mathbb{R})\big)$ και $\varepsilon>0$.

Υπάρχει K συμπαγές υποσύνολο του $C([0,1];\mathbb{R})$ έτσι ώστε $|f(u)|\le\varepsilon,\ \ \forall u\in K^c$.

Επειδή $Int(K)=\emptyset$, ισχύει $\overline{K^c}=C([0,1];\mathbb{R})$, δηλαδή κάθε στοιχείο του K μπορεί να προσεγγιστεί από στοιχεία του K^c, με αποτέλεσμα, αφού η f είναι συνεχής:

$$|f(u)| \leq \varepsilon, \quad \forall u \in C([0,1];\mathbb{R})$$

Αφού το ε ήταν τυχαίος θετικός αριθμός, προκύπτει ότι $f \equiv 0$, δηλαδή $C_0\big(C([0,1];\mathbb{R})\big) = \{0\}$.

∎

Παράδειγμα 1.2

Οι χώροι $L^\infty\big((0,1)\big)$, $C_b(R^n)$ δεν είναι διαχωρίσιμοι, συνεπώς ούτε Πολωνικοί.

<u>Απόδειξη:</u>

Η τεχνική που θα ακολουθήσουμε και στις δύο περιπτώσεις είναι η κατασκευή μιας μη-αριθμήσιμης οικογένειας στοιχείων, οποιαδήποτε δύο από τα οποία θα έχουν απόσταση ακριβώς ίση με 1.

$$L^\infty\big((0,1)\big) = \left\{ f:(0,1) \to \mathbb{R} \ : \ f \text{ μετρήσιμη και } \|f\|_\infty = \operatorname*{ess\,sup}_{x\in(0,1)} |f(x)| < +\infty \right\}$$

Τα στοιχεία του συνόλου $L^\infty\big((0,1)\big)$ είναι κλάσεις ισοδυναμίας, όπου δύο συναρτήσεις ανήκουν στην ίδια κλάση αν και μόνο αν ταυτίζονται σχεδόν παντού.

Για να αποδείξουμε ότι ο L^∞ δεν είναι διαχωρίσιμος αρκεί να βρούμε οικογένεια $(f_i)_{i\in I}$ στοιχείων του τέτοια ώστε $\|f_i - f_j\|_\infty = 1$, $\forall i,j \in I$ (με $i \neq j$) και το σύνολο δεικτών I να είναι υπεραριθμήσιμο.

Τότε, παίρνοντας $O_i = B_{\frac{1}{3}}(f_i) = \left\{ g \in L^\infty : \ \|g - f_i\|_\infty < \frac{1}{3} \right\}$ για $i \in I$, έχουμε:

Το O_i είναι μη-κενό ανοικτό υποσύνολο του $L^\infty\big((0,1)\big)$.
$O_i \cap O_j = \emptyset$, για $i \neq j$ δηλαδή ανά δύο είναι ξένα αφού οποιαδήποτε δύο στοιχεία της ακολουθίας $(f_i)_{i\in I}$ απέχουν απόσταση ακριβώς ίση με 1.

Αν, λοιπόν, υποθέσουμε ότι A είναι ένα αριθμήσιμο και πυκνό υποσύνολο του $L^\infty\big((0,1)\big)$, τότε θα πρέπει να τέμνει κάθε ανοικτό υποσύνολο του χώρου, δηλαδή $A \cap O_i \neq \emptyset$, $\forall i \in I$, συνεπώς $\forall i \in I$, $\exists g_i \in A \cap O_i$.

Έτσι, όμως, δημιουργείται μια $1-1$ απεικόνιση $I \to A : i \mapsto g_i$ και συνεπώς το σύνολο A είναι υπεραριθμήσιμο. Αυτό, όμως, αντιφάσκει με την υπόθεση της αριθμησιμότητας του συνόλου A, άρα έχουμε οδηγηθεί σε άτοπο και επομένως ο $L^\infty\big((0,1)\big)$ δεν είναι διαχωρίσιμος.

Μένει, λοιπόν, να κατασκευάσουμε την ακολουθία $(f_i)_{i\in I}$. Ορίζουμε:

$I = \big\{ \varepsilon = (\varepsilon_1, \varepsilon_2, \dots) : \varepsilon_j \in \{0,1\} \big\}$ το σύνολο των ακολουθιών από 0 και 1, το οποίο είναι υπεραριθμήσιμο.

$(A_j)_{j\in\mathbb{N}}$ μια ακολουθία ανοικτών, ξένων ανά δύο διαστημάτων μέσα στο διάστημα $(0,1)$, και

$$f_\varepsilon = \sum_{j=1}^{\infty} \varepsilon_j \cdot 1_{A_j}$$

(που ανήκει στον $L^\infty((0,1))$ αφού $f_\varepsilon \geq 0$ και η μέγιστη τιμή που μπορεί να πάρει είναι ίση με 1.

Ισχύει ότι αν $\varepsilon \neq \varepsilon'$, τότε $\|f_\varepsilon - f_{\varepsilon'}\|_\infty = 1$, αφού οι ακολουθίες ε και ε' θα διαφέρουν τουλάχιστον για κάποιο ε_{j_0}, οπότε για $x \in A_{j_0}$: $|f_\varepsilon(x) - f_{\varepsilon'}(x)| = 1$ και $m(A_{j_0}) > 0$ (m το μέτρο Lebesgue στον $\mathbb{R}$).

$C_b(\mathbb{R}^n) = \{f: \mathbb{R}^n \to \mathbb{R} : f \text{ συνεχής \& φραγμένη}\}$

Θεωρούμε το σύνολο των φυσικών αριθμών $\mathbb{N}$ και για κάθε $A \subseteq \mathbb{N}$ παίρνουμε τη χαρακτηριστική του συνάρτηση 1_A. Το δυναμοσύνολο του $\mathbb{N}$ είναι ένα υπεραριθμήσιμο σύνολο, οπότε έχουμε πάρει μια υπεραριθμήσιμη οικογένεια συναρτήσεων τέτοια ώστε:

αν $A, B \subseteq \mathbb{N}$ με $A \neq B$, τότε υπάρχει τουλάχιστον ένα $x \in \mathbb{N}$ με $x \in A - B$ ή $x \in B - A$, οπότε $|1_A(x) - 1_B(x)| = 1$, δηλαδή $\|1_A - 1_B\|_\infty = 1$.

Στη συνέχεια, επεκτείνουμε την 1_A σε μια συνεχή συνάρτηση από το $\mathbb{N}$ σε όλο το $\mathbb{R}$ με τέτοιο τρόπο ώστε οι τιμές της να παραμένουν μεταξύ 0 και 1.

Στο σημείο αυτό, έχουμε κατασκευάσει μια υπεραριθμήσιμη οικογένεια συνεχών και φραγμένων συναρτήσεων $f: \mathbb{R} \to \mathbb{R}$ ώστε κάθε δύο από αυτές να έχουν απόσταση ίση με 1.

Η κατασκευή αυτή μπορεί να γενικευτεί και στον $\mathbb{R}^n$ ως εξής:

Για κάθε συνάρτηση $f: \mathbb{R} \longrightarrow \mathbb{R}$ από αυτές που κατασκευάσαμε προηγουμένως, ορίζουμε την

$$g: \mathbb{R}^n \longrightarrow \mathbb{R} \text{ με } g(x_1, x_2, \dots, x_n) = f(x_1)$$

Έτσι, λοιπόν, ο $C_b(\mathbb{R}^n)$ δεν είναι διαχωρίσιμος.

■

1.1.2. Δυϊκότητα.

Στην παράγραφο αυτή θα διατυπώσουμε το Θεώρημα δυϊκότητας, το οποίο όπως θα δούμε ισχύει κάτω από αρκετά γενικές υποθέσεις.

Θεώρημα 1.3 (Δυϊκότητα Kantorovich)

Έστω X, Y δύο Πολωνικοί χώροι, $\mu \in P(X)$, $\nu \in P(Y)$, και $c: X \times Y \longrightarrow [0, +\infty]$ μια κάτω ημισυνεχής συνάρτηση. Στο $P(X \times Y)$ και στο $L^1(d\mu) \times L^1(d\nu)$, ορίζουμε, αντίστοιχα, τα συναρτησοειδή

$$I[\pi] = \int_{X\times Y} c(x,y)\, d\pi(x,y) \quad \text{και} \quad J(\varphi,\psi) = \int_X \varphi\ d\mu + \int_Y \psi\ d\nu.$$

Θυμίζουμε ότι $\Pi(\mu,\nu)$ είναι το σύνολο όλων των μέτρων πιθανότητας π στον $X \times Y$ τέτοιων ώστε για κάθε $A \in \mathcal{B}_X$ και $B \in \mathcal{B}_Y$, να ισχύει

$$\pi[A \times Y] = \mu[A], \qquad \pi[X \times B] = \nu[B].$$

Ορίζουμε Φ_c να είναι το σύνολο όλων των ζευγών μετρήσιμων συναρτήσεων $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$ που ικανοποιούν την ανισότητα

$$\varphi(x) + \psi(y) \leq c(x,y) \tag{1.3}$$

για $d\mu$ −σχεδόν όλα τα $x \in X$ και $d\nu$ −σχεδόν όλα τα $y \in Y$.

Τότε, ισχύει:

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_{\Phi_c} J(\varphi,\psi). \tag{1.4}$$

Επιπλέον, το infimum στο αριστερό μέλος της (1.4) λαμβάνεται, δηλαδή υπάρχει $\pi_* \in \Pi(\mu,\nu)$ τέτοιο ώστε $I[\pi_*] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]$.

Επίσης, μπορούμε να περιορίσουμε τον ορισμό του Φ_c σε ζεύγη συναρτήσεων $(\varphi,\psi) \in C_b(X) \times C_b(Y)$ χωρίς να αλλάξει η τιμή του supremum στο δεξί μέλος της (1.4).

Πριν δώσουμε την απόδειξη του θεωρήματος, θεωρούμε σκόπιμο να αναφέρουμε μερικές παρατηρήσεις.

Παρατηρήσεις 1.4

- **i.** Όσον αφορά στον ορισμό του συνόλου Φ_c, επιτρέπουμε στις συναρτήσεις φ και ψ να παίρνουν τιμές στο $\mathbb{R} \cup \{-\infty\}$.
- **ii.** Θυμίζουμε ότι μια συνάρτηση f ορισμένη σ' έναν μετρικό χώρο X λέγεται κάτω ημισυνεχής, αν για κάθε $x \in X$, ισχύει $f(x) \leq \liminf_{y\to x} f(y)$.
- **iii.** Δεν είναι εκ των προτέρων ξεκάθαρο ότι η τιμή του $\sup J$ δεν αλλάζει όταν περιορίσουμε τον ορισμό του συνόλου Φ_c στις συνεχείς και φραγμένες συναρτήσεις. Παρόλα αυτά, μόλις αποδείξουμε το Θεώρημα 1.3, θα δούμε ότι ο ισχυρισμός αυτός όντως αληθεύει. Όταν θα θέλουμε να τονίσουμε τη διάκριση μεταξύ αυτών των ορισμών, θα γράφουμε άτυπα $\Phi_c \cap C_b$ ή $\Phi_c \cap L^1$.
- **iv.** Στο Θεώρημα 1.3, ούτε καν υποθέτουμε ότι η τιμή του infimum είναι πεπερασμένη.
- **v.** Σε αυτό το κεφάλαιο, δεν μας ενδιαφέρει αν το $\sup J$ λαμβάνεται ή όχι. Αυτή η ερώτηση θα απαντηθεί αργότερα στο κεφάλαιο 2.

vi. Η ανισότητα (1.3) ισχύει για $d\mu$ −σχεδόν όλα τα $x \in X$ και $d\nu$ −σχεδόν όλα τα $y \in Y$. Με αυτό εννοούμε ότι υπάρχουν μετρήσιμα σύνολα $N_X \subseteq X$ και $N_Y \subseteq Y$ με $\mu[N_X] = \nu[N_Y] = 0$ και τέτοια ώστε να ισχύει $\varphi(x) + \psi(y) \leq c(x,y), \quad \forall (x,y) \in N_X^c \times N_Y^c$.

1.1.3. Το πρόβλημα του αποστολέα.

Ένα ωραίο παράδειγμα, με το οποίο μπορούμε να δούμε μια ερμηνεία του Θεωρήματος 1.3 είναι το εξής. Υποθέστε ότι ένας βιομήχανος, που ασχολείται με την εξόρυξη κάρβουνου, επεξεργάζεται το πρόβλημα της μεταφοράς αυτού από το ορυχείο στα εργοστάσια επεξεργασίας. Για να κάνει τη μεταφορά, πρέπει να νοικιάσει φορτηγά και να πληρώσει αντίτιμο $c(x,y)$ για τη μεταφορά ενός τόνου κάρβουνου από την τοποθεσία x στην τοποθεσία y. Η ποσότητα του κάρβουνου που εξορύσσεται σε κάθε θέση καθώς και η ποσότητα που θα τοποθετηθεί σε κάθε εργοστάσιο είναι δεδομένες και περιγράφονται, αντίστοιχα, από τα μέτρα μ, ν. Κάθε $\pi \in \Pi(\mu,\nu)$ περιγράφει την ποσότητα κάρβουνου που θα μεταφερθεί από κάθε θέση του ορυχείου σε κάθε εργοστάσιο, άρα περιγράφει ένα πλάνο μεταφοράς. Καθώς ο βιομήχανος επιδιώκει να βρει πλάνο μεταφοράς που ελαχιστοποιεί το κόστος, εμφανίζεται ένας μαθηματικός και του προτείνει την παρακάτω λύση. Προτίθεται να κάνει τη μεταφορά με δικά του φορτηγά χρεώνοντας τον βιομήχανο μια τιμή $\varphi(x)$ για τη φόρτωση ενός τόνου κάρβουνου από την τοποθεσία x και μια τιμή $\psi(y)$ για το ξεφόρτωμά του στην τοποθεσία y. Θα διαμορφώσει τις τιμές $\varphi(x)$ και $\psi(y)$ με τέτοιον τρόπο ώστε να είναι οικονομικά συμφέρον για τον βιομήχανο να επιλέξει την προσφορά του. Ο μαθηματικός είναι, μάλιστα, διατεθειμένος να επιστρέψει χρήματα στον βιομήχανο προκειμένου να πετύχει ότι για κάθε x και y θα ισχύει

$$\varphi(x) + \psi(y) \leq c(x,y).$$

Φυσικά, ο βιομήχανος αποδέχεται την προσφορά. Αυτό που μας λέει η δυϊκότητα Kantorovich, είναι ότι ο μαθηματικός μπορεί να διαμορφώσει τις τιμές με τέτοιο τρόπο ώστε ο βιομήχανος να πληρώσει όσα (σχεδόν) επρόκειτο να πληρώσει και με την άλλη μέθοδο.

1.1.4. Προκαταρκτική παρατήρηση.

Πριν την απόδειξη του Θεωρήματος 1.3, θα αποδείξουμε ένα εύκολο κομμάτι αυτού, την παρακάτω Πρόταση, την οποία μάλιστα θα χρειαστεί να επικαλεστούμε και στην πλήρη απόδειξη του Θεωρήματος λίγο αργότερα.

Πρόταση 1.5 (Εύκολο κομμάτι της δυϊκότητας Kantorovich)

Κάτω από τις ίδιες υποθέσεις του Θεωρήματος 1.3, ισχύει

$$\sup_{\Phi_c \cap C_b} J(\varphi,\psi) \leq \sup_{\Phi_c \cap L^1} J(\varphi,\psi) \leq \inf_{\Pi(\mu,\nu)} I[\pi]. \tag{1.5}$$

Απόδειξη:

Η ανισότητα στα αριστερά της (1.5) είναι τετριμμένη, καθώς $C_b(X) \times C_b(Y) \subseteq L^1(d\mu) \times L^1(d\nu)$. Συνεπώς, έχουμε να αποδείξουμε μόνο την ανισότητα στα δεξιά. Έστω $(\varphi, \psi) \in \Phi_c \cap L^1$, και έστω π ένα οποιοδήποτε στοιχείο του $\Pi(\mu,\nu)$. Τότε, από τον ορισμό του $\Pi(\mu,\nu)$,

$$J(\varphi,\psi) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu = \int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x,y).$$

Όμως, εφόσον η ανισότητα $\varphi(x) + \psi(y) \leq c(x,y)$ ισχύει για $d\mu$ −σχεδόν όλα τα $x \in X$ και $d\nu$ −σχεδόν όλα τα $y \in Y$, θα δείξουμε ότι ισχύει και $d\pi(x,y)$ −σχεδόν παντού. Πράγματι, έστω N_X, N_Y δύο μετρήσιμα σύνολα με $\mu[N_X] = 0$, $\nu[N_Y] = 0$, και τέτοια ώστε η ανισότητα $\varphi(x) + \psi(y) \leq c(x,y)$ να ισχύει για κάθε $(x,y) \in N_X^c \times N_Y^c$. Εφόσον το π έχει περιθώρια μέτρα τα μ και ν, ισχύει $\pi[N_X \times Y] = \mu[N_X] = 0$, $\pi[X \times N_Y] = \nu[N_Y] = 0$, και άρα

$$\pi[(N_X^c \times N_Y^c)^c] \leq \pi[N_X \times Y] + \pi[X \times N_Y] = 0.$$

Συνεπώς,

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x,y) \leq \int_{X\times Y} c(x,y)\, d\pi(x,y) = I[\pi].$$

Έχουμε, λοιπόν, αποδείξει ότι για τυχαίο ζεύγος $(\varphi,\psi) \in \Phi_c \cap L^1$ και τυχαίο $\pi \in \Pi(\mu,\nu)$, ισχύει η ανισότητα $J(\varphi,\psi) \leq I[\pi]$. Άρα,

$$\sup_{\Phi_c \cap L^1} J(\varphi,\psi) \leq \inf_{\Pi(\mu,\nu)} I[\pi]$$

όπως επιθυμούσαμε.

■

Παρατήρηση 1.6

Από την πρόταση 1.5 προκύπτει ότι η δυϊκότητα $\sup_{\Phi_c \cap C_b} J(\varphi,\psi) = \inf_{\Pi(\mu,\nu)} I[\pi]$ συνεπάγεται αυτόματα ότι $\sup_{\Phi_c \cap L_1} J(\varphi,\psi) = \sup_{\Phi_c \cap C_b} J(\varphi,\psi)$.

1.1.5. Μια αρχή ελαχιστομεγίστου.

Το παρακάτω Θεώρημα 1.7, το οποίο θα χρειαστεί να επικαλεστούμε στην απόδειξη του Θεωρήματος 1.3, αποτελεί μια αρχή ελαχιστομεγίστου. Πριν όμως διατυπώσουμε το θεώρημα, θα πρέπει να εισαγάγουμε την έννοια του μετασχηματισμού Legendre-Fenchel.

Έστω E ένας διανυσματικός χώρος με νόρμα και $\Theta: E \to \mathbb{R} \cup \{+\infty\}$ μια κυρτή συνάρτηση. Θυμίζουμε ότι η υπόθεση της κυρτότητας σημαίνει:

$$\forall (z_1, z_2, \lambda) \in E \times E \times [0,1], \quad \Theta(\lambda z_1 + (1-\lambda) z_2) \leq \lambda \Theta(z_1) + (1-\lambda)\Theta(z_2),$$

με τις προφανείς επεκτάσεις των πράξεων του $\mathbb{R}$ στο $\mathbb{R} \cup \{+\infty\}$.

Θυμίζουμε, ακόμα, ότι ο τοπολογικός δυϊκός του E είναι το σύνολο των συνεχών γραμμικών απεικονίσεων από τον E στο $\mathbb{R}$, δηλαδή $E^* = \{f: E \to \mathbb{R} \ : \ f$ γραμμική και συνεχής$\}$.

Τότε, ο μετασχηματισμός Legendre-Fenchel της Θ είναι η συνάρτηση Θ^*, ορισμένη στον τοπολογικό δυϊκό E^* του E μέσω του τύπου:

$$\Theta^*(z^*) = \sup_{z \in E} [\langle z^*, z \rangle - \Theta(z)].$$

Θεώρημα 1.7 (Δυϊκότητα Fenchel – Rockafellar)

Έστω E ένας διανυσματικός χώρος με νόρμα, E^* ο τοπολογικός δυϊκός χώρος αυτού και $\Theta, \Xi : E \longrightarrow \mathbb{R} \cup \{+\infty\}$ δύο κυρτές συναρτήσεις. Έστω, επίσης, Θ^*, Ξ^* οι μετασχηματισμοί Legendre - Fenchel των Θ, Ξ αντίστοιχα. Υποθέτουμε ότι υπάρχει $z_0 \in E$ τέτοιο ώστε

$$\Theta(z_0) < +\infty, \qquad \Xi(z_0) < +\infty, \qquad \text{και η } \Theta \text{ είναι συνεχής στο } z_0.$$

Τότε,

$$\inf_{z \in E}[\Theta(z) + \Xi(z)] = \sup_{z^* \in E^*}[-\Theta^*(-z^*) - \Xi^*(z^*)] = \max_{z^* \in E^*}[-\Theta^*(-z^*) - \Xi^*(z^*)] \tag{1.6}$$

Ισχυρισμός 3. Η πρώτη ισότητα στη σχέση (1.6) γράφεται ως εξής:

$$\sup_{z^* \in E^*} \inf_{z,u \in E} \big(\langle z^*, z-u \rangle + \Theta(z) + \Xi(u)\big) = \inf_{z,u \in E} \sup_{z^* \in E^*} \big(\langle z^*, z-u \rangle + \Theta(z) + \Xi(u)\big)$$

<u>Απόδειξη:</u>

Το δεύτερο μέλος της (1.6)γράφεται ως εξής:

$$\sup_{z^* \in E^*}[-\Theta^*(-z^*) - \Xi^*(z^*)] = \sup_{z^* \in E^*}\left[-\sup_{z \in E}\big(\langle -z^*, z \rangle - \Theta(z)\big) - \sup_{z \in E}\big(\langle z^*, z \rangle - \Xi(z)\big)\right]$$

$$= \sup_{z^* \in E^*}\left[\inf_{z \in E}\big(\langle z^*, z \rangle + \Theta(z)\big) + \inf_{u \in E}\big(\langle -z^*, u \rangle + \Xi(u)\big)\right]$$

$$= \sup_{z^* \in E^*} \inf_{z,u \in E}\big(\langle z^*, z-u \rangle + \Theta(z) + \Xi(u)\big).$$

Επίσης,

$$\inf_{z,u\in E}\sup_{z^*\in E^*}\big(\langle z^*,z-u\rangle+\Theta(z)+\Xi(u)\big)=\inf_{z,u\in E}\left[\sup_{z^*\in E^*}\left(\langle z^*,z-u\rangle\right)+\Theta(z)+\Xi(u)\right]$$

$$=\inf_{z\in E}\big(\Theta(z)+\Xi(z)\big)$$

που είναι το πρώτο μέλος της σχέσης (1.6). Η τελευταία ισότητα ισχύει γιατί, αν $z\neq u$, τότε $\sup_{z^*\in E^*}\langle z^*,z-u\rangle=+\infty$. Συνδυάζοντας, λοιπόν, τα δύο παραπάνω αποτελέσματα, έχουμε ότι η (1.6) γράφεται ως εξής:

$$\sup_{z^*\in E^*}\inf_{z,u\in E}\big(\langle z^*,z-u\rangle+\Theta(z)+\Xi(u)\big)=\inf_{z,u\in E}\sup_{z^*\in E^*}\big(\langle z^*,z-u\rangle+\Theta(z)+\Xi(u)\big).$$

■

Ισχυρισμοί αυτής της μορφής αναφέρονται ως «αρχές ελαχιστομεγίστου».

Για την απόδειξη του Θεωρήματος 1.7 θα χρειαστούμε το παρακάτω λήμμα.

Λήμμα 1.8

Έστω $C\subseteq E$ ένα κυρτό σύνολο. Τότε, το $IntC$ είναι, επίσης, κυρτό. Αν, επιπλέον, $IntC\neq\emptyset$, τότε $\overline{C}=\overline{IntC}$.

Απόδειξη του λήμματος 1.8:

Για την απόδειξη του λήμματος, θα χρειαστούμε τον παρακάτω ισχυρισμό.

Ισχυρισμός 4. Έστω C ένα κυρτό υποσύνολο του E. Έστω, ακόμα, $x\in intC$ και $y\in\overline{C}$. Τότε, για κάθε $\lambda\in(0,1]$ ισχύει: $\lambda x+(1-\lambda)y\in IntC$.

Απόδειξη του Ισχυρισμού 4:

Το εσωτερικό και η κλειστότητα του C μπορούν να εκφραστούν από τους τύπους:

$$IntC=\{x\in E:\ \exists\,\varepsilon>0\ \ \text{τ.ω}\ \ x+\varepsilon B\subseteq C\}$$

$$\overline{C}=\bigcap_{\varepsilon>0}(C+\varepsilon B).$$

όπου $B=B_1(0)=\{x\in E:\ \|x\|<1\}$.

Έστω, λοιπόν, $\lambda\in(0,1]$. Πρέπει να δείξουμε ότι

$$\text{υπάρχει }\ \varepsilon>0\ \text{ έτσι ώστε }\ [\lambda x+(1-\lambda)y]+\varepsilon B\subseteq C.$$

Εφόσον $y\in\overline{C}$, θα ισχύει $y\in C+\varepsilon B$ για κάθε $\varepsilon>0$. Επομένως, έχοντας την κυρτότητα των συνόλων B και C, για κάθε $\varepsilon>0$ έχουμε:

$$[\lambda x + (1-\lambda)y] + \varepsilon B \subseteq \lambda x + (1-\lambda)(C + \varepsilon B) + \varepsilon B = \lambda x + (1-\lambda)C + \varepsilon(2-\lambda)B$$
$$= \lambda\left[x + \varepsilon \cdot \frac{2-\lambda}{\lambda} \cdot B\right] + (1-\lambda)C \subseteq \lambda C + (1-\lambda)C = C$$

όπου ο τελευταίος εγκλεισμός ισχύει αν το $\varepsilon > 0$ επιλεγεί έτσι ώστε να ισχύει

$$x + \varepsilon \cdot \frac{2-\lambda}{\lambda} \cdot B \subseteq C,$$

το οποίο είναι δυνατό, αφού $x \in IntC$.

■

Επιστρέφουμε στην απόδειξη του λήμματος.

Για να δούμε ότι το $IntC$ είναι κυρτό, αρκεί να πάρουμε το y να είναι στοιχείο του $IntC$ στον παραπάνω ισχυρισμό. Έστω, επιπλέον, ότι $IntC \neq \emptyset$. Θα δείξουμε ότι $\overline{IntC} = \overline{C}$.

Για τον πρώτο εγκλεισμό, ισχύει $IntC \subseteq C$, άρα $\overline{IntC} \subseteq \overline{C}$.

Για τον αντίστροφο εγκλεισμό, έστω $y \in \overline{C}$. Έστω, επίσης, $x \in IntC$ (υπάρχει τέτοιο, αφού $IntC \neq \emptyset$). Λόγω του ισχυρισμού 4, $\forall \lambda \in (0,1] : \lambda x + (1-\lambda)y \in IntC$. Άρα, παίρνοντας μια ακολουθία $\lambda_n \in (0,1]$ με $\lambda_n \to 0$, έχουμε ότι: $\lambda_n x + (1-\lambda_n)y \to y$, δηλαδή το y αποτελεί όριο ακολουθίας στοιχείων του $IntC$, άρα $y \in \overline{IntC}$.

■

<u>Απόδειξη του Θεωρήματος 1.7:</u>

Θέτουμε

$$a = \inf_{z\in E}[\Theta(z) + \Xi(z)]$$

$$b = \sup_{z^*\in E^*}[-\Theta^*(-z^*) - \Xi^*(z^*)].$$

Εύκολα μπορούμε να επαληθεύσουμε ότι $b \leq a$. Πράγματι,

$$b = \sup_{z^*\in E^*}\{-\Theta^*(-z^*) - \Xi^*(z^*)\}$$
$$= \sup_{z^*\in E^*}\left\{-\sup_{x\in E}\big(-\langle z^*, x\rangle - \Theta(x)\big) - \sup_{y\in E}\big(\langle z^*, y\rangle - \Xi(y)\big)\right\}$$
$$= \sup_{z^*\in E^*}\left\{\inf_{x\in E}\big(\langle z^*, x\rangle + \Theta(x)\big) + \inf_{y\in E}\big(-\langle z^*, y\rangle + \Xi(y)\big)\right\}$$
$$= \sup_{z^*\in E^*}\left\{\inf_{x,y\in E}(\Theta(x) + \Xi(y) + \langle z^*, x-y\rangle)\right\} \leq \sup_{z^*\in E^*}\{\Theta(x) + \Xi(x)\} = \Theta(x) + \Xi(x),$$

όπου η τελευταία ανισότητα προκύπτει θέτοντας $y = x$.

Δηλαδή, για τυχαίο $x \in E$, ισχύει $b \leq \Theta(x) + \Xi(x)$. Άρα, $b \leq \inf_{x \in E}\big(\Theta(x) + \Xi(x)\big) = a$.

Εξάλλου, έχουμε είτε $a \in \mathbb{R}$ είτε $a = -\infty$ (αφού $\Theta(z_0) < +\infty$, $\Xi(z_0) < +\infty$). Αν το $a = -\infty$, τότε η ανισότητα $a \leq b$ είναι προφανής και η απόδειξη του θεωρήματος πλήρης. Ας υποθέσουμε, λοιπόν, ότι $a \in \mathbb{R}$. Έστω C το επιγράφημα της Θ, δηλαδή το σύνολο

$$C = epi\Theta = \{(z, \lambda) \in E \times \mathbb{R} : \Theta(z) \leq \lambda\}.$$

Εφόσον $\Theta(z_0) < +\infty$ και η Θ είναι συνεχής στο z_0, θα υπάρχει περιοχή του z_0, έστω U_{z_0}, τέτοια ώστε η $\Theta(z)$ να είναι άνω φραγμένη στο U_{z_0}. Επομένως, $IntC \neq \emptyset$.

Πράγματι, έστω $\lambda_0 \in \mathbb{R}$ τέτοιο ώστε $\lambda_0 > \Theta(z) + 1$, για κάθε $z \in U_{z_0}$. Τότε, το σύνολο $U_{z_0} \times (\lambda_0 - 1, \lambda_0 + 1)$ αποτελεί μια ανοιχτή περιοχή του σημείου $(z_0, \lambda_0) \in C$ και εξ ορισμού του C : $U_{z_0} \times (\lambda_0 - 1, \lambda_0 + 1) \subseteq C$. Επομένως, $IntC \neq \emptyset$.

Θα εφαρμόσουμε τώρα το Θεώρημα Hahn-Banach για τα κυρτά σύνολα $A = intC$ και

$$B = \{(z, \lambda) \in E \times \mathbb{R} : \quad \lambda \leq a - \Xi(z)\}.$$

Το σύνολο C είναι κυρτό. Πράγματι, το $C = epi\Theta$ είναι κυρτό σύνολο στο $E \times \mathbb{R}$ αν και μόνο αν η Θ είναι κυρτή συνάρτηση. Ακόμα, σύμφωνα με το λήμμα 1.8, το $A = IntC$ είναι κυρτό, και προφανώς, ανοιχτό σύνολο.

Το σύνολο B είναι μη-κενό, αφού $\Xi(z_0) < +\infty$, και θα επαληθεύσουμε ακόμα ότι είναι κυρτό. Πράγματι, έστω (z_1, λ_1) και (z_2, λ_2) σημεία του συνόλου B. Τότε θα ισχύει $\lambda_1 \leq a - \Xi(z_1)$ και $\lambda_2 \leq a - \Xi(z_2)$. Έστω $t \in (0, 1)$. Τότε,

$$t(z_1, \lambda_1) + (1 - t)(z_2, \lambda_2) = (tz_1 + (1 - t)z_2, t\lambda_1 + (1 - t)\lambda_2)$$

οπότε,

$$\begin{aligned} t\lambda_1 + (1 - t)\lambda_2 &\leq t\big(a - \Xi(z_1)\big) + (1 - t)\big(a - \Xi(z_2)\big) = ta + (1 - t)a - \big(t\Xi(z_1) + (1 - t)\Xi(z_2)\big) \\ &\leq a - \Xi(tz_1 + (1 - t)z_2) \end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει λόγω της κυρτότητας της Ξ. Άρα, $t(z_1, \lambda_1) + (1 - t)(z_2, \lambda_2) \in B$ και συνεπώς το σύνολο B είναι κυρτό.

Τέλος, θα επαληθεύσουμε ότι τα A, B είναι ξένα. Πράγματι, αν $(z, \lambda) \in A = intC$, έχουμε

$$\lambda > \Theta(z) \geq a - \Xi(z)$$

(σύμφωνα με τον ορισμό του a) και άρα $(z, \lambda) \notin B$. Συνεπώς, $A \cap B = \emptyset$.

Οι υποθέσεις του Θεωρήματος Hahn - Banach ικανοποιούνται, επομένως υπάρχει ένα κλειστό υπερεπίπεδο H που διαχωρίζει τα A και B. Δηλαδή, υπάρχει $c \in \mathbb{R}$ και συνεχές γραμμικό, μη-μηδενικό, συναρτησοειδές Φ στο $E \times \mathbb{R}$ τέτοιο ώστε $H = [\Phi = c]$ και

$$\Phi(z,\lambda) \geq c, \quad \forall (z,\lambda) \in A$$

$$\Phi(z,\lambda) \leq c, \quad \forall (z,\lambda) \in B.$$

Λόγω της συνέχειας του Φ, ισχύει $\Phi(z,\lambda) \geq c$ για κάθε $(z,\lambda) \in \overline{A}$. Άρα, το H διαχωρίζει επίσης τα $\overline{A}$ και B.

Επειδή $\overline{A} = \overline{IntC} = \overline{C}$ (λόγω του λήμματος 1.8), το υπερεπίπεδο H διαχωρίζει τα C και B. Όμως κάθε συνεχές γραμμικό συναρτησοειδές στον $E \times \mathbb{R}$ είναι της μορφής:

$$\Phi(z,\lambda) = \langle z^*, z\rangle + k\lambda,$$

όπου $z^* \in E^*$ και $k \in \mathbb{R}$.

Πράγματι, η απεικόνιση $E \longrightarrow \mathbb{R}, \; z \longmapsto \Phi(z,0)$ είναι ένα συνεχές γραμμικό συναρτησοειδές πάνω στον E, δηλαδή

$$\Phi(z,0) = \langle z^*, z\rangle \;\text{ για κάποιο } z^* \in E^*.$$

Αν, λοιπόν, θέσουμε $k = \Phi(0,1)$, τότε για κάθε $(z,\lambda) \in E \times \mathbb{R}$, θα έχουμε:

$$\Phi(z,\lambda) = \Phi[(z,0) + (0,\lambda)] = \Phi(z,0) + \lambda\Phi(0,1) = \langle z^*, z\rangle + k\lambda.$$

Συνεχίζοντας, λοιπόν, έχουμε:

$$\Phi(z,\lambda) = \langle z^*, z\rangle + k\lambda \geq c, \qquad \forall (z,\lambda) \in C \tag{1.7}$$

$$\Phi(z,\lambda) = \langle z^*, z\rangle + k\lambda \leq c, \qquad \forall (z,\lambda) \in B. \tag{1.8}$$

Επιλέγοντας $z = z_0$ στη σχέση (1.7), παίρνουμε

$$\langle z^*, z_0\rangle + k\lambda \geq c,$$

για κάθε $\lambda \geq \Theta(z_0)$, άρα

$$k\lambda \geq c - \langle z^*, z_0\rangle = \Gamma \quad \in \mathbb{R}.$$

Άρα, $\lim\limits_{\lambda\to+\infty}(k\lambda) \geq \Gamma$ και εφόσον το όριο αυτό δεν είναι ίσο με $-\infty$, συμπεραίνουμε ότι $k \geq 0$.

Στη συνέχεια, θα δείξουμε, με απαγωγή σε άτοπο, ότι $k > 0$.

Κατ' αρχήν $\Phi \neq 0$, άρα $\|z^*\| + |k| \neq 0$. Ας υποθέσουμε ότι $k = 0$. Τότε, λόγω των σχέσεων (1.7) και (1.8) ισχύει

$$\langle z^*, z\rangle \geq c, \quad \forall z \in Dom(\Theta)$$

$$\langle z^*, z\rangle \leq c, \quad \forall z \in Dom(\Xi),$$

όπου $Dom(\Theta) = \{z \in E : \; \Theta(z) < +\infty\}$.

Επειδή η Θ είναι συνεχής στο z_0 και $\Theta(z_0) < +\infty$, θα υπάρχει $\varepsilon_0 > 0$ ώστε $B_{\varepsilon_0}(z_0) \subseteq Dom(\Theta)$ και συνεπώς

$$\langle z^*, z_0 + \varepsilon_0 z\rangle \geq c, \quad \forall z \in B_1(0),$$

από την οποία προκύπτει ότι:

$$\langle z^*, z_0\rangle + \varepsilon_0 \langle z^*, z\rangle \geq c, \quad \forall z \in B_1(0)$$

$$\inf_{z \in B_1(0)} \{\langle z^*, z_0\rangle + \varepsilon_0 \langle z^*, z\rangle\} \geq c$$

$$\langle z^*, z_0\rangle + \varepsilon_0 \inf_{\|z\| \leq 1} \{\langle z^*, z\rangle\} \geq c$$

$$\langle z^*, z_0\rangle - \varepsilon_0 \sup_{\|z\| \leq 1} \{-\langle z^*, z\rangle\} \geq c$$

$$\langle z^*, z_0\rangle - \varepsilon_0 \|z^*\| \geq c$$

$$\langle z^*, z_0\rangle \geq c + \varepsilon_0 \|z^*\|$$

Από την άλλη, όμως, $\langle z^*, z_0\rangle \leq c$, αφού $z_0 \in Dom(\Xi)$.

Άρα, θα πρέπει $z^* = 0$, το οποίο όμως είναι άτοπο, αφού $k = 0$ και $\|z^*\| + |k| \neq 0$. Επομένως $k > 0$.

Από τη σχέση (1.7), διαιρώντας με $-k < 0$, έχουμε:

$$\langle -\frac{z^*}{k}, z\rangle - \lambda \leq -\frac{c}{k}, \quad \forall (z, \lambda) \in C.$$

Όμως, $\big(z, \Theta(z)\big) \in C$, οπότε για $\lambda = \Theta(z)$ παίρνουμε

$$\langle -\frac{z^*}{k}, z\rangle - \Theta(z) \leq -\frac{c}{k}, \quad \forall z \in E.$$

Η ανισότητα προκύπτει από τα παραπάνω για $z \in Dom(\Theta)$, αλλά φανερά ισχύει και όταν $\Theta(z) = +\infty$. Επομένως, ισχύει για κάθε $z \in E$.

Άρα,

$$\Theta^*\left(-\frac{z^*}{k}\right) = \sup_{z \in E} \left\{\langle -\frac{z^*}{k}, z\rangle - \Theta(z)\right\} \leq -\frac{c}{k}. \tag{1.9}$$

Αντίστοιχα, από τη σχέση (1.8), διαιρώντας με $k > 0$, έχουμε:

$$\langle \frac{z^*}{k}, z\rangle + \lambda \leq \frac{c}{k}, \quad \forall (z, \lambda) \in B.$$

Αν $\Xi(z) \neq +\infty$, τότε εφόσον $\big(z, a - \Xi(z)\big) \in B$, επιλέγοντας $\lambda = a - \Xi(z)$, θα έχουμε

$$\langle\frac{z^*}{k},z\rangle - \Xi(z) \leq \frac{c}{k} - a,$$

η οποία ανισότητα, όμως, ισχύει και στην περίπτωση που $\Xi(z) = +\infty$.

Άρα,

$$\Xi^*\left(\frac{z^*}{k}\right) = \sup_{z\in E}\left\{\langle\frac{z^*}{k},z\rangle - \Xi(z)\right\} \leq \frac{c}{k} - a. \tag{1.10}$$

Συνεπώς, από τις σχέσεις (1.9) και (1.10):

$$-\Theta^*\left(-\frac{z^*}{k}\right) - \Xi^*\left(\frac{z^*}{k}\right) \geq \frac{c}{k} + a - \frac{c}{k} = a.$$

Επειδή, όμως, από τον ορισμό του b έχουμε

$$-\Theta^*\left(-\frac{z^*}{k}\right) - \Xi^*\left(\frac{z^*}{k}\right) \leq b,$$

και άρα,

$$-\Theta^*\left(-\frac{z^*}{k}\right) - \Xi^*\left(\frac{z^*}{k}\right) \leq b \leq \alpha \leq -\Theta^*\left(-\frac{z^*}{k}\right) - \Xi^*\left(\frac{z^*}{k}\right)$$

συμπεραίνουμε ότι

$$a = b = -\Theta^*\left(-\frac{z^*}{k}\right) - \Xi^*\left(\frac{z^*}{k}\right).$$

■

1.1.6. Απόδειξη της δυϊκότητας Kantorovich.

Στο σημείο αυτό, έχουμε όλα όσα χρειαζόμαστε για να προχωρήσουμε στην απόδειξη του Θεωρήματος 1.3.

<u>Απόδειξη του Θεωρήματος 1.3:</u>

Θα χωρίσουμε την απόδειξη σε τέσσερα σκέλη, ξεκινώντας από την πιο ειδική περίπτωση και πηγαίνοντας προς την γενικότερη.

<u>1ο σκέλος:</u> Υποθέτουμε ότι οι X, Y είναι συμπαγείς Πολωνικοί χώροι και η συνάρτηση c είναι συνεχής στον $X \times Y$ με πραγματικές τιμές.

Έστω $E = C(X \times Y)$ ο χώρος των συνεχών συναρτήσεων στον $X \times Y$, εφοδιασμένος με τη supremum νόρμα $\|\cdot\|_\infty$. Σημειώνουμε ότι, εφόσον οι χώροι X, Y υποτέθηκαν συμπαγείς, το ίδιο ισχύει και για τον χώρο $X \times Y$. Άρα, οι χώροι συναρτήσεων $C(X \times Y)$ και $C_b(X \times Y)$ ταυτίζονται.

Από το Θεώρημα αναπαράστασης του Riesz (Θεώρημα 22, Εισαγωγικό Κεφάλαιο), γνωρίζουμε ότι ο τοπολογικός δυϊκός του E είναι ο χώρος των κανονικών προσημασμένων μέτρων Borel στον $X \times Y$, δηλαδή $E^* = M(X \times Y)$ εφοδιασμένος με τη νόρμα της ολικής κύμανσης. Ειδικότερα, κάθε μη αρνητικό γραμμικό συναρτησοειδές ορισμένο στον E, αναπαρίσταται από ένα μη-αρνητικό μέτρο Borel.
Στη συνέχεια, εισάγουμε τις συναρτήσεις

$$\Theta : u \in C(X \times Y) \longmapsto \begin{cases} 0, & \text{αν } u(x,y) \geq -c(x,y) \\ +\infty, & \text{αλλιώς} \end{cases}$$

$$\Xi : u \in C(X \times Y) \longmapsto \begin{cases} \int_X \varphi \, d\mu + \int_Y \psi \, d\nu, & \text{αν } u(x,y) = \varphi(x) + \psi(y) \\ +\infty, & \text{αλλιώς} \end{cases}$$

Θα πρέπει να ελέγξουμε ότι η Ξ είναι καλώς ορισμένη. Πράγματι, ας υποθέσουμε ότι

$$\forall (x,y) \in X \times Y: \quad \varphi(x) + \psi(y) = \tilde{\varphi}(x) + \tilde{\psi}(y),$$

δηλαδή ότι η $u(x,y)$ μπορεί να γραφεί με δύο τρόπους ως άθροισμα μιας συνάρτησης του x και μιας συνάρτησης του y.
Τότε,

$$\forall (x,y) \in X \times Y: \quad \varphi(x) - \tilde{\varphi}(x) = \tilde{\psi}(y) - \psi(y).$$

Όμως, η μόνη περίπτωση όπου οι τιμές δύο συναρτήσεων διαφορετικών μεταβλητών είναι ίσες για όλες τις δυνατές τιμές των μεταβλητών αυτών, είναι όταν οι δύο αυτές συναρτήσεις είναι σταθερές. Άρα, υπάρχει $s \in \mathbb{R}$ ώστε:

$$\varphi(x) - \tilde{\varphi}(x) = s, \qquad \tilde{\psi}(y) - \psi(y) = s,$$

ή ισοδύναμα,

$$\varphi = \tilde{\varphi} + s, \qquad \psi = \tilde{\psi} - s.$$

Τότε, όμως,

$$\int_X \varphi \, d\mu + \int_Y \psi \, d\nu = \int_X \tilde{\varphi} \, d\mu + \int_Y \tilde{\psi} \, d\nu.$$

Στη συνέχεια, θα θέλαμε να εφαρμόσουμε το Θεώρημα 1.7, οπότε θα ελέγξουμε αν ικανοποιούνται οι υποθέσεις του για τις συναρτήσεις Θ και Ξ.
Έστω $u, v \in C(X \times Y)$ και $\lambda \in (0,1)$. Για την κυρτότητα της Θ, θα πρέπει να ισχύει:

$$\Theta(\lambda u + (1-\lambda)v) \leq \lambda\Theta(u) + (1-\lambda)\Theta(v).$$

Αν $\Theta(u) = +\infty$ ή $\Theta(v) = +\infty$, τότε η ανισότητα είναι, προφανώς, αληθής.

Η μόνη περίπτωση που χρειάζεται να εξεταστεί είναι όταν $u, v \geq -c$. Τότε,

$$\lambda u + (1-\lambda)v \geq -\lambda c - (1-\lambda)c = -c,$$

άρα και τα δύο μέλη της ανισότητας είναι ίσα με 0.

Ομοίως, για την κυρτότητα της Ξ, θα πρέπει να ισχύει:

$$\Xi(\lambda u + (1-\lambda)v) \leq \lambda\Xi(u) + (1-\lambda)\Xi(v).$$

Αν $\Xi(u) = +\infty$ ή $\Xi(v) = +\infty$, τότε η ανισότητα είναι, προφανώς, αληθής.

Η μόνη περίπτωση που χρειάζεται να εξεταστεί είναι όταν $u = \varphi_1 + \psi_1$ και $v = \varphi_2 + \psi_2$. Τότε,

$$\begin{aligned} \Xi(\lambda u + (1-\lambda)v) &= \int_X (\lambda\varphi_1 + (1-\lambda)\varphi_2)\, d\mu + \int_Y (\lambda\psi_1 + (1-\lambda)\psi_2)\, d\nu \\ &= \lambda\left[\int_X \varphi_1\, d\mu + \int_Y \psi_1\, d\nu\right] + (1-\lambda)\left[\int_X \varphi_2\, d\mu + \int_Y \psi_2\, d\nu\right] = \lambda\Xi(u) + (1-\lambda)\Xi(v) \end{aligned}$$

δηλαδή, η ανισότητα είναι αληθής ως ισότητα.
Ακόμα, επιλέγοντας $z_0 \equiv 1$ στο Θεώρημα 1.7, τότε $\Theta(1) = 0$, $\Xi(1) = 1$ και η Θ είναι συνεχής στο z_0. Επομένως, οι υποθέσεις του Θεωρήματος ικανοποιούνται πλήρως, και για το εφαρμόσουμε αρκεί να υπολογίσουμε ξεχωριστά τις τιμές των δύο μελών της σχέσης (1.6).
Το αριστερό μέλος της (1.6) είναι:

$$\begin{aligned} \inf_{u\in E}\{\Theta(u) + \Xi(u)\} &= \inf\left\{\int_X \varphi\, d\mu + \int_Y \psi\, d\nu : \varphi(x) + \psi(y) \geq -c(x,y)\right\} \\ &= \inf\left\{-\int_X \varphi\, d\mu - \int_Y \psi\, d\nu : \varphi(x) + \psi(y) \leq c(x,y)\right\} \\ &= -\sup\left\{\int_X \varphi\, d\mu + \int_Y \psi\, d\nu : \varphi(x) + \psi(y) \leq c(x,y)\right\} = -\sup\{J(\varphi,\psi) : (\varphi,\psi) \in \Phi_c \cap C\}. \end{aligned}$$

Για τον υπολογισμό του δεξιού μέλους της (1.6), θα πρέπει πρώτα να υπολογίσουμε τους μετασχηματισμούς Legendre-Fenchel των Θ, Ξ. Για κάθε $\pi \in E^* = M(X \times Y)$, έχουμε:

$$\begin{aligned} \Theta^*(-\pi) &= \sup_{u\in E}[\langle -\pi, u\rangle - \Theta(u)] \\ &= \sup\left\{-\int u(x,y) d\pi(x,y) - 0 : \ u \in C(X\times Y), \quad u(x,y) \geq -c(x,y)\right\} \\ &= \sup\left\{\int u(x,y) d\pi(x,y) : \ u \in C(X\times Y), \quad u(x,y) \leq c(x,y)\right\} \end{aligned}$$

Αν το π δεν είναι ένα μη-αρνητικό μέτρο, τότε υπάρχει συνάρτηση $v \in C(X \times Y)$ με $v \leq 0$, τέτοια ώστε $\int_{X\times Y} v\, d\pi > 0$. Τότε, μπορούμε να επιλέξουμε $u = \lambda v$, $\lambda > 0$ και να αφήσουμε το $\lambda \to +\infty$, οπότε θα έχουμε

$$\Theta^*(-\pi) \geq \lambda \cdot \int_{X\times Y} v\, d\pi \to +\infty.$$

Στη δεύτερη περίπτωση, όπου το π είναι μη-αρνητικό μέτρο, τότε προφανώς η τιμή του supremum είναι ίση με

$$\int_{X\times Y} c\, d\pi.$$

Τελικά,

$$\Theta^*(-\pi) = \begin{cases} \int_{X\times Y} c(x,y)\, d\pi(x,y), & \text{αν } \pi \in M_+(X\times Y) \\ +\infty, & \text{αλλιώς} \end{cases}$$

Με τον ίδιο τρόπο,

$$\Xi^*(\pi) = \sup_{u\in E}[\langle \pi, u\rangle - \Xi(u)] = \sup_{(\varphi,\psi)\in C(X)\times C(Y)} \left[\int [\varphi(x)+\psi(y)]\, d\pi(x,y) - \int_X \varphi\, d\mu - \int_Y \psi\, d\nu \right]$$

Περιορίζουμε τη διερεύνηση σε μέτρα $\pi \in M_+(X\times Y)$, καθώς σε άλλη περίπτωση ισχύει $\Theta^*(-\pi) = +\infty$ και αυτό δε συνεισφέρει τίποτα στο δεύτερο μέλος της σχέσης (1.6) την οποία θέλουμε να εφαρμόσουμε.

Αν $\pi \in \Pi(\mu,\nu)$, τότε από τον ορισμό του συνόλου $\Pi(\mu,\nu)$ μέσω της σχέσης (27), προκύπτει ότι $\Xi^*(\pi) = 0$. Αν $\pi \notin \Pi(\mu,\nu)$, τότε και πάλι από τον ορισμό του $\Pi(\mu,\nu)$ μέσω της σχέσης (1.2) και τον Ισχυρισμό 1, υπάρχει ζεύγος συναρτήσεων $(\varphi_1,\psi_1) \in C(X)\times C(Y)$ έτσι ώστε

$$\int_{X\times Y} [\varphi_1(x)+\psi_1(y)]\, d\pi(x,y) - \int_X \varphi_1\, d\mu - \int_Y \psi_1\, d\nu = K \neq 0.$$

Αν $K>0$, επιλέγοντας $\varphi = \lambda\varphi_1,\ \psi = \lambda\psi_1$ με $\lambda > 0$ και αφήνοντας το $\lambda \to +\infty$, έχουμε:

$$\Xi^*(\pi) \geq \lambda K \to +\infty.$$

Αν $K<0$, επιλέγοντας $\varphi = \lambda\varphi_1,\ \psi = \lambda\psi_1$ με $\lambda < 0$ και αφήνοντας το $\lambda \to -\infty$, έχουμε:

$$\Xi^*(\pi) \geq \lambda K \to +\infty.$$

Τελικά,

$$\Xi^*(\pi) = \begin{cases} 0, & \text{αν } \pi \in \Pi(\mu,\nu) \\ +\infty, & \text{αλλιώς} \end{cases}$$

Εφαρμόζοντας, λοιπόν, τη σχέση (1.6) έχουμε:

$$\inf_{u\in E}\{\Theta(u)+\Xi(u)\} = \sup_{\pi\in\Pi(\mu,\nu)}\{-\Theta^*(-\pi) - \Xi^*(\pi)\}$$

άρα

$$-\sup_{\Phi_c\cap C} J(\varphi,\psi) = \sup_{\pi\in\Pi(\mu,\nu)} \left\{ -\int c(x,y)\, d\pi(x,y) - 0 \right\}$$

άρα

$$\sup_{\Phi_c\cap C} J(\varphi,\psi) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Συνδυάζοντας το αποτέλεσμα αυτό με τη σχέση (1.5), παίρνουμε

$$\sup_{\Phi_c\cap L^1} J(\varphi,\psi) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]$$

και έτσι τελειώνει η απόδειξη του πρώτου σκέλους.

<u>2ο σκέλος:</u> Υποθέτουμε ότι οι X, Y είναι Πολωνικοί χώροι και η συνάρτηση c είναι φραγμένη και ομοιόμορφα συνεχής.

Ισχυρισμός 5. Το σύνολο $\Pi(\mu,\nu)$ είναι ακολουθιακά συμπαγές στον $P(X\times Y)$ αναφορικά με την ασθενή τοπολογία των μέτρων πιθανότητας, δηλαδή την τοπολογία που επάγεται από τον $C_b(X\times Y)$.

<u>Απόδειξη:</u>

Αρκεί να δείξουμε ότι το $\Pi(\mu,\nu)$ είναι σχετικά ακολουθιακά συμπαγές και κλειστό στο $P(X\times Y)$ αναφορικά με την ασθενή τοπολογία.

Έστω $\varepsilon>0$. Τα μέτρα μ και ν, εφόσον είναι ορισμένα πάνω στους Πολωνικούς χώρους X και Y αντίστοιχα, είναι σφικτά (Πρόταση 25, Εισαγωγικό Κεφάλαιο). Άρα, υπάρχουν συμπαγή σύνολα $K_\varepsilon\subseteq X,\ L_\varepsilon\subseteq Y$ τέτοια ώστε

$$\mu[X-K_\varepsilon]\le\frac{\varepsilon}{2}\qquad\text{και}\qquad\nu[Y-L_\varepsilon]\le\frac{\varepsilon}{2}.$$

Τότε, για κάθε $\pi\in\Pi(\mu,\nu)$, έχουμε:

$$\begin{aligned}\pi[(X\times Y)-(K_\varepsilon\times L_\varepsilon)]&\le\pi[(X-K_\varepsilon)\times Y]+\pi[X\times(Y-L_\varepsilon)]=\mu[X-K_\varepsilon]+\nu[Y-L_\varepsilon]\\&\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon\end{aligned}$$

και το σύνολο $K_\varepsilon\times L_\varepsilon$ είναι συμπαγές ως καρτεσιανό γινόμενο δύο συμπαγών συνόλων.

Αυτό αποδεικνύει ότι το σύνολο $\Pi(\mu,\nu)$ είναι σφικτό, και επομένως σύμφωνα με το θεώρημα του Prokhorov, σχετικά ακολουθιακά συμπαγές στον $P(X\times Y)$, αναφορικά με την ασθενή τοπολογία.

Για να δείξουμε ότι το $\Pi(\mu,\nu)$ είναι (ασθενώς) κλειστό στον $P(X\times Y)$, έστω $(\pi_n)_{n\in\mathbb{N}}$ μια ακολουθία μέτρων στο $\Pi(\mu,\nu)$, η οποία συγκλίνει ασθενώς στο μέτρο $\pi_*\in P(X\times Y)$. Αυτό σημαίνει ότι για κάθε $f\in C_b(X\times Y)$, ισχύει:

$$\lim_{n\to\infty}\int_{X\times Y}f(x,y)\,d\pi_n(x,y)=\int_{X\times Y}f(x,y)\,d\pi_*(x,y).$$

Για να δείξουμε ότι $\pi_*\in\Pi(\mu,\nu)$, θεωρούμε στην παραπάνω σχέση μια οποιαδήποτε συνάρτηση $f(x,y)=g(x)\in C_b(X\times Y)$. Εφόσον, για κάθε $n\in\mathbb{N}$ ισχύει $\pi_n\in\Pi(\mu,\nu)$, θα έχουμε:

$$\int_{X\times Y}g(x)\,d\pi_n(x,y)=\int_X g(x)\,d\mu.$$

Τελικά,

$$\int_{X\times Y}g(x)\,d\pi_*(x,y)=\lim_{n\to\infty}\int_{X\times Y}g(x)\,d\pi_n(x,y)=\int_X g(x)\,d\mu,$$

που σημαίνει ότι το μ είναι το περιθώριο μέτρο του π_* στον X.

Ομοίως, δείχνουμε ότι το ν είναι το περιθώριο μέτρο του π_* στον Y, και συνεπώς $\pi_* \in \Pi(\mu,\nu)$.

■

Ισχυρισμός 6. Υπάρχει βέλτιστο πλάνο μεταφοράς $\pi_* \in \Pi(\mu,\nu)$ για το πρόβλημα Kantorovich, δηλαδή ένα πλάνο μεταφοράς που ικανοποιεί τη σχέση:

$$I[\pi_*] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Απόδειξη:

Έστω μια ακολουθία $(\pi_n)_{n\in\mathbb{N}}$ στο $\Pi(\mu,\nu)$ για την οποία ισχύει

$$\lim_{n\to\infty} I[\pi_n] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Λόγω της ακολουθιακής συμπάγειας του $\Pi(\mu,\nu)$, η ακολουθία αυτή θα έχει σημείο συσσώρευσης $\pi_* \in \Pi(\mu,\nu)$. Υπάρχει, δηλαδή, υπακολουθία $\left(\pi_{n_k}\right)_{k\in\mathbb{N}}$ της $(\pi_n)_{n\in\mathbb{N}}$ η οποία συγκλίνει ασθενώς στο π_*. Αφού η συνάρτηση κόστους $c(x,y)$ είναι συνεχής και φραγμένη, εξ ορισμού της ασθενούς σύγκλισης, θα ισχύει:

$$\lim_{k\to+\infty} \int_{X\times Y} c(x,y)\, d\pi_{n_k}(x,y) = \int_{X\times Y} c(x,y)\, d\pi_*(x,y).$$

Έχουμε, λοιπόν:

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \lim_{k\to+\infty} I\left[\pi_{n_k}\right] = \lim_{k\to+\infty} \int_{X\times Y} c(x,y)\, d\pi_{n_k}(x,y) = \int_{X\times Y} c(x,y)\, d\pi_*(x,y) = I[\pi_*].$$

■

Συνεχίζοντας με το 2ο σκέλος της απόδειξης του θεωρήματος, έστω $0 < \delta < \frac{1}{2}$.

Από την σφικτότητα των μέτρων μ και ν, υπάρχουν συμπαγή σύνολα $X_0 \subseteq X$, $Y_0 \subseteq Y$ τέτοια ώστε $\mu[X - X_0] \leq \delta, \quad \nu[Y - Y_0] \leq \delta$. Επομένως,

$$\begin{aligned}\pi_*[(X\times Y) - (X_0\times Y_0)] &\leq \pi_*[(X - X_0)\times Y] + \pi_*[X\times(Y - Y_0)] = \mu[X - X_0] + \nu[Y - Y_0] \leq \delta + \delta \\ &= 2\delta.\end{aligned}$$

Ακόμα, ισχύει $\pi_*[X_0\times Y_0] > 0$, αφού το δ έχει επιλεγεί μικρότερο του $\frac{1}{2}$.

Ορίζουμε

$$\pi_{*0} = \frac{1_{X_0\times Y_0}}{\pi_*[X_0\times Y_0]}\pi_*,$$

το οποίο είναι ένα μέτρο πιθανότητας πάνω στον $X_0 \times Y_0$, και έστω μ_0, ν_0 τα περιθώρια μέτρα του π_{*0} πάνω στους X_0, Y_0 αντίστοιχα. Ορίζουμε το $\Pi_0(\mu_0, \nu_0)$ ως το σύνολο των μέτρων πιθανότητας π_0 στον $X_0 \times Y_0$ που έχουν περιθώρια μέτρα μ_0, ν_0 και για κάθε $\pi_0 \in \Pi_0(\mu_0, \nu_0)$ ορίζουμε το συναρτησοειδές I_0 ως εξής:

$$I_0[\pi_0] = \int_{X_0 \times Y_0} c(x,y)\, d\pi_0(x,y).$$

Έστω $\tilde{\pi}_0 \in \Pi_0(\mu_0, \nu_0)$ τέτοιο ώστε

$$I_0[\tilde{\pi}_0] = \inf_{\pi_0 \in \Pi_0(\mu_0, \nu_0)} I_0[\pi_0].$$

Θέλουμε, τώρα, από το $\tilde{\pi}_0$ να κατασκευάσουμε ένα $\tilde{\pi} \in \Pi(\mu, \nu)$ συνδυάζοντας το $\tilde{\pi}_0$ με το π_*. Ορίζουμε, λοιπόν,

$$\tilde{\pi} = \pi_*[X_0 \times Y_0]\tilde{\pi}_0 + 1_{(X_0 \times Y_0)^c}\pi_*$$

Ισχυρισμός 7. $\tilde{\pi} \in \Pi(\mu, \nu)$.

Απόδειξη:

Αρκεί να δείξουμε ότι για κάθε $A \in \mathcal{B}_X$, $B \in \mathcal{B}_Y$, ισχύει

$$\tilde{\pi}[A \times Y] = \mu[A], \quad \tilde{\pi}[X \times B] = \nu[B].$$

Έχουμε

$$\tilde{\pi}[A \times Y] = \pi_*[X_0 \times Y_0]\tilde{\pi}_0[(A \times Y) \cap (X_0 \times Y_0)] + \pi_*[(A \times Y) \cap (X_0 \times Y_0)^c],$$

όπου

$$\tilde{\pi}_0[(A \times Y) \cap (X_0 \times Y_0)] = \tilde{\pi}_0[(A \cap X_0) \times Y_0] = \mu_0(A \cap X_0)$$

εφόσον $\tilde{\pi}_0 \in \Pi_0(\mu_0, \nu_0)$, και επίσης,

$$\mu_0(A \cap X_0) = \pi_{*0}[(A \cap X_0) \times Y_0] = \frac{\pi_*[(A \cap X_0) \times Y_0]}{\pi_*[X_0 \times Y_0]}.$$

Άρα, επιστρέφοντας στην πρώτη σχέση, έχουμε:

$$\begin{aligned}
\tilde{\pi}[A \times Y] &= \pi_*[X_0 \times Y_0]\frac{\pi_*[(A \cap X_0) \times Y_0]}{\pi_*[X_0 \times Y_0]} + \pi_*[(A \times Y) \cap (X_0 \times Y_0)^c] \\
&= \pi_*[(A \cap X_0) \times Y_0] + \pi_*[(A \times Y) \cap (X_0 \times Y_0)^c] \\
&= \pi_*\big[\big((A \cap X_0) \times Y_0\big) \cup \big((A \times Y) \cap (X_0 \times Y_0)^c\big)\big] = \pi_*[A \times Y] = \mu[A].
\end{aligned}$$

Ομοίως, δείχνουμε ότι $\tilde{\pi}[X \times B] = \nu[B]$, και έτσι η απόδειξη του ισχυρισμού είναι πλήρης.

∎

Επιστρέφοντας ξανά στην απόδειξη του 2ου σκέλους του θεωρήματος, έχουμε:

$$I[\tilde{\pi}] = \pi_*[X_0 \times Y_0] I_0[\tilde{\pi}_0] + \int_{(X_0\times Y_0)^c} c(x,y)\, d\pi_*(x,y) \leq I_0[\tilde{\pi}_0] + \|c\|_\infty \int_{(X_0\times Y_0)^c} 1\, d\pi_*(x,y)$$
$$= I_0[\tilde{\pi}_0] + \|c\|_\infty \pi_*[(X_0 \times Y_0)^c] \leq\ I_0[\tilde{\pi}_0] + 2\|c\|_\infty \delta = \inf_{\pi_0 \in \Pi_0(\mu_0,\nu_0)} I_0 + 2\|c\|_\infty \delta.$$

Έπεται ότι

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] \leq \inf_{\pi_0\in\Pi_0(\mu_0,\nu_0)} I_0 + 2\|c\|_\infty \delta.$$

Στη συνέχεια, εισάγουμε το συναρτησοειδές

$$J_0(\varphi_0,\psi_0) = \int_{X_0} \varphi_0\, d\mu_0 + \int_{Y_0} \psi_0\, d\nu_0,$$

όπου $(\varphi_0,\psi_0) \in L^1(d\mu_0) \times L^1(d\nu_0)$. Εφόσον X_0, Y_0 είναι συμπαγή σύνολα, από το 1ο σκέλος της απόδειξης, γνωρίζουμε ότι $\inf I_0 = \sup J_0$, όπου το supremum λαμβάνεται πάνω σε όλα τα αποδεκτά ζεύγη συναρτήσεων $(\varphi_0,\psi_0) \in L^1(d\mu_0) \times L^1(d\nu_0)$, δηλαδή πάνω στα ζεύγη που ικανοποιούν την ανισότητα $\varphi_0(x) + \psi_0(y) \leq c(x,y)$ σχεδόν για όλα τα $x \in X_0$ και $y \in Y_0$, με την έννοια που αναφέρεται στην παρατήρηση 1.4.vi.

Σημειώνουμε ότι $\sup J_0 = \inf I_0 < +\infty$, αφού η συνάρτηση c έχει υποτεθεί φραγμένη. Εξ' ορισμού του supremum, υπάρχει αποδεκτό ζεύγος συναρτήσεων $(\tilde{\varphi}_0,\tilde{\psi}_0) \in C(X_0) \times C(Y_0)$ τέτοιο ώστε $J_0(\tilde{\varphi}_0,\tilde{\psi}_0) \geq \sup J_0 - \delta$.

Στη συνέχεια, θέλουμε να κατασκευάσουμε από το $(\tilde{\varphi}_0,\tilde{\psi}_0)$ ένα ζεύγος συναρτήσεων $(\overline{\varphi}_0,\overline{\psi}_0)$ για το οποίο θα αποδείξουμε την ανισότητα $J(\overline{\varphi}_0,\overline{\psi}_0) \geq \inf I$.

Στην προσπάθεια αυτή, θα είναι χρήσιμο να διασφαλίσουμε ότι η ανισότητα $\tilde{\varphi}_0(x) + \tilde{\psi}_0(y) \leq c(x,y)$ ισχύει για όλα τα $x \in X_0$ και $y \in Y_0$, και όχι μόνο για σχεδόν όλα. Θεωρούμε, λοιπόν, σύνολα $N_X \subseteq X_0$ και $N_Y \subseteq Y_0$ για τα οποία $\mu_0[N_X] = 0$, $\nu_0[N_Y] = 0$ και η ανισότητα $\tilde{\varphi}_0(x) + \tilde{\psi}_0(y) \leq c(x,y)$ ισχύει για κάθε $(x,y) \in N_X^c \times N_Y^c$. Έπειτα, αλλάζουμε τις τιμές των $\tilde{\varphi}_0$ και $\tilde{\psi}_0$ ώστε να είναι ίσες με $-\infty$ πάνω στα σύνολα N_X και N_Y αντίστοιχα, και έτσι έχουμε το ζητούμενο αποτέλεσμα.

Σε πρώτη φάση, θα ελέγξουμε από κάτω τις συναρτήσεις $\tilde{\varphi}_0$, $\tilde{\psi}_0$ σε κάποιο σημείο του $X \times Y$. Εφόσον $J_0(0,0) = 0$, θα έχουμε ότι $supJ_0 \geq 0$, άρα και $J_0(\tilde{\varphi}_0,\tilde{\psi}_0) \geq 0 - \delta \geq -1$. Γράφοντας το συναρτησοειδές J_0 στη μορφή

$$J_0(\tilde{\varphi}_0,\tilde{\psi}_0) = \int_{X\times Y} [\tilde{\varphi}_0(x) + \tilde{\psi}_0(y)]\, d\pi_0(x,y) \geq -1,$$

όπου π_0 είναι ένα οποιοδήποτε στοιχείο του $\Pi_0(\mu_0,\nu_0)$, συμπεραίνουμε ότι υπάρχει $(x_0,y_0) \in X_0 \times Y_0$ τέτοιο ώστε $\tilde{\varphi}_0(x_0) + \tilde{\psi}_0(y_0) \geq -1$.

Αν αντικαταστήσουμε το ζεύγος $(\tilde{\varphi}_0, \tilde{\psi}_0)$ με το $(\tilde{\varphi}_0 + s, \tilde{\psi}_0 - s)$, όπου $s \in \mathbb{R}$, τότε η τιμή του συναρτησοειδούς $J_0(\tilde{\varphi}_0, \tilde{\psi}_0)$ δεν αλλάζει και επίσης, το νέο αυτό ζεύγος συναρτήσεων παραμένει αποδεκτό. Μάλιστα, επιλέγοντας κατάλληλα το $s \in \mathbb{R}$, μπορούμε να επιτύχουμε:

$$\tilde{\varphi}_0(x_0) \geq -\frac{1}{2} \quad \text{και} \quad \tilde{\psi}_0(y_0) \geq -\frac{1}{2}.$$

Τότε, λόγω της ανισότητας $\tilde{\varphi}_0(x) + \tilde{\psi}_0(y) \leq c(x,y)$, που ισχύει για κάθε $x \in X_0$ και $y \in Y_0$, θα έχουμε ότι: για κάθε $(x,y) \in X_0 \times Y_0$,

$$\tilde{\varphi}_0(x) \leq c(x, y_0) - \tilde{\psi}_0(y_0) \leq c(x, y_0) + \frac{1}{2}$$

$$\tilde{\psi}_0(y) \leq c(x_0, y) - \tilde{\varphi}_0(x_0) \leq c(x_0, y) + \frac{1}{2}.$$

Για να προχωρήσουμε παραπέρα την απόδειξη, θα χρησιμοποιήσουμε ένα τέχνασμα ώστε να κατασκευάσουμε ζεύγος συναρτήσεων $(\overline{\varphi}_0, \overline{\psi}_0) \in \Phi_c$ τέτοιο ώστε $\overline{\varphi}_0 \geq \tilde{\varphi}_0$ στο X_0 και $\overline{\psi}_0 \geq \tilde{\psi}_0$ στο Y_0.

Για κάθε $x \in X$, ορίζουμε

$$\overline{\varphi}_0(x) = \inf_{y \in Y_0} \left[c(x,y) - \tilde{\psi}_0(y)\right].$$

Η συνάρτηση $\overline{\varphi}_0$ είναι μετρήσιμη, όπως θα αποδείξουμε στην παρατήρηση πιο κάτω.

Από την ανισότητα $\tilde{\varphi}_0(x) \leq c(x,y) - \tilde{\psi}_0(y),\ x \in X_0$, βλέπουμε ότι

$$\tilde{\varphi}_0(x) \leq \inf_{y \in Y_0} \left[c(x,y) - \tilde{\psi}_0(y)\right] = \overline{\varphi}_0(x),$$

δηλαδή $\tilde{\varphi}_0 \leq \overline{\varphi}_0$ πάνω στο X_0.

Ακόμα, από τον ορισμό της $\overline{\varphi}_0$, έχουμε ότι

$$\overline{\varphi}_0(x) \leq c(x, y_0) - \tilde{\psi}_0(y_0) \quad \text{για κάθε } x \in X,$$

οπότε

$$\int_{X_0} \left|\overline{\varphi}_0\right| d\mu_0 \leq \|c\|_\infty + \left|\tilde{\psi}_0(y_0)\right| < +\infty$$

δηλαδή $\overline{\varphi}_0 \in L^1(d\mu_0)$.

Άρα, ορίζεται το $J_0(\overline{\varphi}_0, \tilde{\psi}_0)$ και $J_0(\overline{\varphi}_0, \tilde{\psi}_0) \geq J_0(\tilde{\varphi}_0, \tilde{\psi}_0)$.

Επιπλέον, για κάθε $x \in X$ έχουμε έναν έλεγχο της $\overline{\varphi}_0(x)$ από πάνω και κάτω, αναφορικά με τη συνάρτηση κόστους:

$$\overline{\varphi}_0(x) = \inf_{y\in Y_0}\left[c(x,y) - \tilde{\psi}_0(y)\right] \geq \inf_{y\in Y_0}[c(x,y) - c(x_0,y)] - \frac{1}{2}$$

$$\overline{\varphi}_0(x) = \inf_{y\in Y_0}\left[c(x,y) - \tilde{\psi}_0(y)\right] \leq c(x,y_0) - \tilde{\psi}_0(y_0) \leq c(x,y_0) + \frac{1}{2}.$$

Τέλος, για κάθε $y \in Y$, ορίζουμε

$$\overline{\psi}_0(y) = \inf_{x\in X}\left[c(x,y) - \overline{\varphi}_0(x)\right].$$

Η συνάρτηση $\overline{\psi}_0$ είναι μετρήσιμη, όπως θα αποδείξουμε στην παρατήρηση πιο κάτω.

Επίσης, συνεχίζει να ισχύει $\left(\overline{\varphi}_0, \overline{\psi}_0\right) \in \Phi_c$. Πράγματι, λόγω των παραπάνω εκτιμήσεων για την $\overline{\varphi}_0$ και εφόσον η c είναι φραγμένη, παρατηρούμε ότι η $\overline{\varphi}_0$ παίρνει τιμές στο $\mathbb{R}$. Επομένως, για κάθε $(x_1, y_1) \in X \times Y$, έχουμε:

$$\overline{\varphi}_0(x_1) + \overline{\psi}_0(y_1) = \overline{\varphi}_0(x_1) + \inf_{x\in X}\left[c(x,y_1) - \overline{\varphi}_0(x)\right] \leq \overline{\varphi}_0(x_1) + c(x_1,y_1) - \overline{\varphi}_0(x_1) = c(x_1,y_1).$$

Επίσης, για κάθε $(x_1, y_1) \in X_0 \times Y_0$,

$$\overline{\varphi}_0(x_1) = \inf_{y\in Y_0}\left[c(x_1,y) - \tilde{\psi}_0(y)\right] \leq c(x_1,y_1) - \tilde{\psi}_0(y_1)$$

άρα,

$$\tilde{\psi}_0(y_1) \leq c(x_1,y_1) - \overline{\varphi}_0(x_1)$$

και

$$\tilde{\psi}_0(y_1) \leq \inf_{x_1\in X_0}\left[c(x_1,y_1) - \overline{\varphi}_0(x_1)\right] = \overline{\psi}_0(y_1),$$

δηλαδή $\overline{\psi}_0 \geq \tilde{\psi}_0$ στο Y_0.

Ακόμα, εφόσον εξ ορισμού της $\overline{\psi}_0$ ισχύει

$$\overline{\psi}_0(y) \leq c(x_0,y) - \overline{\varphi}_0(x_0), \quad \text{για κάθε } y \in Y,$$

θα έχουμε

$$\int_{Y_0} \left|\overline{\psi}_0\right| d\nu_0 \leq \|c\|_\infty + \left|\overline{\varphi}_0(x_0)\right| < +\infty,$$

δηλαδή $\overline{\psi}_0 \in L^1(d\nu_0)$.

Συνεπώς, ορίζεται το $J_0\left(\overline{\varphi}_0, \overline{\psi}_0\right)$ και παίρνουμε ότι $J_0\left(\overline{\varphi}_0, \overline{\psi}_0\right) \geq J_0\left(\overline{\varphi}_0, \tilde{\psi}_0\right) \geq J_0(\tilde{\varphi}_0, \tilde{\psi}_0)$.

Επιπλέον, για κάθε $y \in Y$,

$$\overline{\psi}_0(y) = \inf_{x \in X}\left[c(x,y) - \overline{\varphi}_0(x)\right] \geq \inf_{x \in X}[c(x,y) - c(x,y_0)] - \frac{1}{2},$$

$$\overline{\psi}_0(y) \leq c(x_0,y) - \overline{\varphi}_0(x_0) \leq c(x_0,y) - \tilde{\varphi}_0(x_0) \leq c(x_0,y) + \frac{1}{2}.$$

Ειδικότερα,

$$\overline{\varphi}_0(x) \geq -\|c\|_\infty - \frac{1}{2},$$

$$\overline{\psi}_0(y) \geq -\|c\|_\infty - \frac{1}{2}.$$

Εφόσον έχουμε διασφαλίσει αυτά τα φράγματα, ισχύει:

$$\begin{aligned}
J(\overline{\varphi}_0, \overline{\psi}_0) &= \int_X \overline{\varphi}_0\, d\mu + \int_Y \overline{\psi}_0\, d\nu = \int_{X \times Y} \left[\overline{\varphi}_0(x) + \overline{\psi}_0(y)\right] d\pi_*(x,y) \\
&= \pi_*[X_0 \times Y_0] \int_{X_0 \times Y_0} \left[\overline{\varphi}_0(x) + \overline{\psi}_0(y)\right] d\pi_{*0}(x,y) + \int_{(X_0 \times Y_0)^c} \left[\overline{\varphi}_0(x) + \overline{\psi}_0(y)\right] d\pi_*(x,y) \\
&\geq (1 - 2\delta)\left(\int_{X_0} \overline{\varphi}_0\, d\mu_0 + \int_{Y_0} \overline{\psi}_0\, d\nu_0\right) - (2\|c\|_\infty + 1)\pi_*[(X_0 \times Y_0)^c] \\
&\geq (1 - 2\delta)J_0(\overline{\varphi}_0, \overline{\psi}_0) - 2(2\|c\|_\infty + 1)\delta \geq (1 - 2\delta)J_0(\tilde{\varphi}_0, \tilde{\psi}_0) - 2(2\|c\|_\infty + 1)\delta \\
&\geq (1 - 2\delta)(inf I_0 - \delta) - 2(2\|c\|_\infty + 1)\delta \\
&\geq (1 - 2\delta)(inf I - (2\|c\|_\infty + 1)\delta) - 2(2\|c\|_\infty + 1)\delta.
\end{aligned}$$

Αφού το δ ήταν ένας τυχαίος θετικός αριθμός μικρότερος του $\frac{1}{2}$, συμπεραίνουμε ότι $\sup J(\varphi,\psi) \geq \inf I$. Έτσι, λοιπόν, αποδεικνύεται η ισότητα $\sup J(\varphi,\psi) = \inf I$, αφού έχουμε ήδη αποδείξει την αντίστροφη ανισότητα στην Πρόταση 1.5.

Παρατήρηση: Εφόσον η c υποτέθηκε ομοιόμορφα συνεχής συνάρτηση, οι συναρτήσεις $\overline{\varphi}_0, \overline{\psi}_0$ είναι ομοιόμορφα συνεχείς στους χώρους X, Y αντίστοιχα. Επιπλέον, αφού οι $c, \tilde{\varphi}_0, \tilde{\psi}_0$ είναι φραγμένες, το ίδιο ισχύει για τις $\overline{\varphi}_0, \overline{\psi}_0$.

Πράγματι, έστω $\varepsilon > 0$. Από την ομοιόμορφη συνέχεια της c, υπάρχει $\delta > 0$ τέτοιο ώστε για κάθε $x_1, x_2 \in X$ με $d(x_1,x_2) < \delta$ να ισχύει $|c(x_1,y) - c(x_2,y)| < \varepsilon, \;\; y \in Y_0$.

Άρα, και

$$\left|\left(c(x_1,y) - \tilde{\psi}_0(y)\right) - \left(c(x_2,y) - \tilde{\psi}_0(y)\right)\right| < \varepsilon, \qquad y \in Y_0.$$

Εύκολα μπορούμε να δούμε ότι ισχύει και η παρακάτω ανισότητα

$$\left|\inf_{y \in Y_0}\left(c(x_1,y) - \tilde{\psi}_0(y)\right) - \inf_{y \in Y_0}\left(c(x_2,y) - \tilde{\psi}_0(y)\right)\right| < \varepsilon,$$

δηλαδή $\left|\overline{\varphi}_0(x_1) - \overline{\varphi}_0(x_2)\right| < \varepsilon$ και έτσι η $\overline{\varphi}_0$ προκύπτει ομοιόμορφα συνεχής στον X.

Ομοίως δείχνουμε ότι η $\overline{\psi}_0$ είναι ομοιόμορφα συνεχής στον Y.

Έτσι, δεν υπάρχει κάποια διαφορά αν θεωρούμε το supremum του συναρτησοειδούς J πάνω στις ολοκληρώσιμες συναρτήσεις (δηλ. στο $\Phi_c \cap L^1$) ή πάνω στις συνεχείς και φραγμένες συναρτήσεις (δηλ. στο $\Phi_c \cap C_b$).

Ακόμα, η παρατήρηση αυτή αποδεικνύει ότι οι συναρτήσεις $\overline{\varphi}_0, \overline{\psi}_0$, εφόσον είναι συνεχείς, είναι και μετρήσιμες.

<u>3ο σκέλος:</u> Τέλος, θα ασχοληθούμε με τη γενική περίπτωση όπου οι X, Y είναι Πολωνικοί χώροι και η συνάρτηση κόστους c είναι μια κάτω ημισυνεχής συνάρτηση.

Σύμφωνα με την Πρόταση 36 του εισαγωγικού κεφαλαίου, μπορούμε να γράψουμε $c = \sup_{n\in\mathbb{N}} c_n$, όπου $(c_n)_{n=1}^{\infty}$ είναι μια αύξουσα ακολουθία μη-αρνητικών και ομοιόμορφα συνεχών συναρτήσεων. Μπορούμε να υποθέσουμε ότι κάθε c_n είναι μια φραγμένη συνάρτηση, αφού σε διαφορετική περίπτωση μπορούμε να αντικαταστήσουμε την c_n με την $\inf\{c_n, n\}$.

Στη συνέχεια, στο $\Pi(\mu,\nu)$, ορίζουμε το συναρτησοειδές I_n ως εξής:

$$I_n[\pi] = \int_{X\times Y} c_n \, d\pi.$$

Εφόσον έχουμε υποθέσει ότι οι συναρτήσεις c_n είναι ομοιόμορφα συνεχείς και φραγμένες, από το 2ο σκέλος της απόδειξης γνωρίζουμε ότι

$$\inf_{\pi\in\Pi(\mu,\nu)} I_n[\pi] = \sup_{(\varphi,\psi)\in\Phi_{c_n}} J(\varphi,\psi). \tag{1.11}$$

Θα ολοκληρώσουμε την απόδειξη δείχνοντας ότι

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_n \inf_{\pi\in\Pi(\mu,\nu)} I_n[\pi] \tag{1.12}$$

και επίσης ότι, για κάθε $n \in \mathbb{N}$,

$$\sup_{(\varphi,\psi)\in\Phi_{c_n}} J(\varphi,\psi) \le \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi). \tag{1.13}$$

Πράγματι, ο συνδυασμός των σχέσεων (1.11), (1.12) και (1.13) συνεπάγεται ότι

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] \le \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi),$$

ενώ ήδη γνωρίζουμε, από την Πρόταση 1.5, ότι η αντίστροφη ανισότητα είναι αληθής.

Εφόσον, εκ κατασκευής, ισχύει $c_n \leq c$, έπεται ότι $\Phi_{c_n} \subseteq \Phi_c$ και συνεπώς η σχέση (1.13) είναι αληθής. Επιπλέον, δεν έχει καμία σημασία αν ορίζουμε τα σύνολα Φ_{c_n} και Φ_c ως υποσύνολα του $C_b(X) \times C_b(Y)$ ή του $L^1(d\mu) \times L^1(d\nu)$.

Εφόσον η $(I_n)_{n=1}^{\infty}$ είναι μια αύξουσα ακολουθία συναρτησοειδών, θα ισχύει ότι και η $(\inf I_n)_{n=1}^{\infty}$ θα είναι μια αύξουσα ακολουθία αριθμών, και μάλιστα φραγμένη από πάνω από το $\inf I$. Δηλαδή, για κάθε $n \in \mathbb{N}$,

$$\inf_{\pi \in \Pi(\mu,\nu)} I_n[\pi] \leq \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

Άρα, αρκεί να αποδείξουμε ότι

$$\lim_{n\to\infty} \inf_{\pi \in \Pi(\mu,\nu)} I_n[\pi] \geq \inf_{\pi \in \Pi(\mu,\nu)} I[\pi]. \tag{1.14}$$

Εφόσον οι c_n είναι φραγμένες και ομοιόμορφα συνεχείς, από τον Ισχυρισμό 5 γνωρίζουμε ότι υπάρχει βέλτιστο πλάνο μεταφοράς $\pi_n \in \Pi(\mu,\nu)$, δηλαδή πλάνο μεταφοράς που ικανοποιεί τη σχέση $I_n[\pi_n] = \inf_{\pi \in \Pi(\mu,\nu)} I_n[\pi]$.

Λόγω της ακολουθιακής συμπάγειας του $\Pi(\mu,\nu)$, η ακολουθία $(\pi_n)_{n\in\mathbb{N}}$ έχει σημείο συσσώρευσης $\pi_* \in \Pi(\mu,\nu)$. Αυτό σημαίνει ότι υπάρχει υπακολουθία της $(\pi_n)_{n\in\mathbb{N}}$, έστω $\left(\pi_{n_k}\right)_{k\in\mathbb{N}}$, που συγκλίνει ασθενώς στο π_*.

Αν $n \geq m$, τότε έχουμε $I_n[\pi_n] \geq I_m[\pi_n]$.

Έτσι, λόγω συνέχειας του συναρτησοειδούς I_m,

$$\lim_{n\to\infty} I_n[\pi_n] \geq \limsup_{n\to\infty} I_m[\pi_n] \geq \lim_{k\to\infty} I_m[\pi_{n_k}] = I_m[\pi_*].$$

Λόγω μονότονης σύγκλισης, $I_m[\pi_*] \to I[\pi_*]$ καθώς $m \to \infty$, και επομένως

$$\lim_{n\to\infty} I_n[\pi_n] \geq \lim_{m\to\infty} I_m[\pi_*] = I[\pi_*] \geq \inf_{\pi \in \Pi(\mu,\nu)} I[\pi],$$

το οποίο αποδεικνύει την (1.14), και έτσι ολοκληρώνεται η απόδειξη της σχέσης (1.4).

<u>4ο σκέλος:</u> Για να ολοκληρώσουμε την απόδειξη του θεωρήματος 1.3, μένει μόνο να δείξουμε ότι το infimum λαμβάνεται, δηλαδή ότι υπάρχει $\pi_* \in \Pi(\mu,\nu)$ ώστε $I[\pi_*] = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi]$.

Αυτό θα είναι, επίσης, συνέπεια της ακολουθιακής συμπάγειας του συνόλου $\Pi(\mu,\nu)$.

Πράγματι, έστω $(\pi_k)_{k=1}^{\infty}$ μια ακολουθία στο $\Pi(\mu,\nu)$ τέτοια ώστε $\lim_{k\to\infty} I[\pi_k] = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi]$.

Η ακολουθία $(\pi_k)_{k=1}^{\infty}$, λόγω της ακολουθιακής συμπάγειας του $\Pi(\mu,\nu)$, θα έχει σημείο συσσώρευσης $\pi_* \in \Pi(\mu,\nu)$, δηλαδή υπάρχει υπακολουθία της $(\pi_k)_{k=1}^{\infty}$, έστω $\left(\pi_{k_l}\right)_{l=1}^{\infty}$, που συγκλίνει ασθενώς στο π_*. Τότε, χρησιμοποιώντας το Θεώρημα Μονότονης Σύγκλισης για την αύξουσα ακολουθία (c_n) που θεωρήσαμε στο 3ο σκέλος της απόδειξης, θα έχουμε:

$$I[\pi_*] = \lim_{n\to\infty} I_n[\pi_*] = \lim_{n\to\infty} (\lim_{l\to\infty} I_n[\pi_{k_l}]) \leq \lim_{n\to\infty} \left(\limsup_{k\to\infty} I_n[\pi_k]\right) \leq \limsup_{k\to\infty} I[\pi_k] = \inf I.$$

■

Παρατήρηση 1.9 (c −κοίλες συναρτήσεις)

Θεωρούμε την περίπτωση στην οποία η c είναι φραγμένη και, επιπλέον, ισχύει η εξής ιδιότητα: για κάθε μετρήσιμη φραγμένη $\varphi : X \longrightarrow \mathbb{R}$ και κάθε μετρήσιμη φραγμένη $\psi : Y \longrightarrow \mathbb{R}$, οι φ^c, ψ^c είναι μετρήσιμες, όπου οι συναρτήσεις φ^c και φ^{cc} ορίζονται από τις σχέσεις:

$$\varphi^c(y) = \inf_{x\in X}[c(x,y) - \varphi(x)] \quad \text{και} \quad \varphi^{cc}(x) = \inf_{y\in Y}[c(x,y) - \varphi^c(y)].$$

Τότε, το supremum στο δεξί μέλος της σχέσης (1.4) μπορεί να περιοριστεί στα ζεύγη συναρτήσεων $(\varphi^{cc}, \varphi^c)$, όπου η συνάρτηση φ είναι φραγμένη ή ακόμη φραγμένη και συνεχής.

Πράγματι, από τον ορισμό της φ^{cc} προκύπτει ότι $\varphi^{cc}(x) + \varphi^c(y) \leq c(x,y)$, άρα $(\varphi^{cc}, \varphi^c) \in \Phi_c$. Επιπλέον, αν $(\varphi, \psi) \in \Phi_c$, από τον ορισμό της φ^c, παίρνουμε $\varphi^c \geq \psi$ και από τον ορισμό της φ^{cc}, παίρνουμε $\varphi^{cc} \geq \varphi$, άρα $J(\varphi^{cc}, \varphi^c) \geq J(\varphi, \psi)$.

Το ζεύγος $(\varphi^{cc}, \varphi^c)$ ονομάζεται **ζεύγος συζυγών c −κοίλων συναρτήσεων.**

Ισχυρισμός 8. $(\varphi^{cc})^c = \varphi^c$

<u>Απόδειξη:</u>

Καταρχήν, θα δείξουμε ότι για κάθε $(x,y) \in X \times Y$, ισχύει $\varphi(x) + \varphi^c(y) \leq c(x,y)$.

Έστω $(x_0, y_0) \in X \times Y$. Τότε,

$$\varphi(x_0) + \varphi^c(y_0) = \varphi(x_0) + \inf_{x\in X}[c(x,y_0) - \varphi(x)] \leq \varphi(x_0) + c(x_0,y_0) - \varphi(x_0) = c(x_0,y_0).$$

Επομένως, για κάθε $(x,y) \in X \times Y$, έχουμε $c(x,y) - \varphi^c(y) \geq \varphi(x)$, οπότε και για κάθε $x \in X$, έχουμε:

$$\varphi^{cc}(x) = \inf_{y\in Y}[c(x,y) - \varphi^c(y)] \geq \varphi(x).$$

Άρα,

$$(\varphi^{cc})^c(y) = \inf_{x\in X}[c(x,y) - \varphi^{cc}(x)] \leq \inf_{x\in X}[c(x,y) - \varphi(x)] = \varphi^c(y). \tag{1.15}$$

Αν στην ανισότητα που προηγήθηκε της σχέσης (1.15) βάλουμε τη συνάρτηση φ^c στη θέση της φ, τότε παίρνουμε:

$$(\varphi^{cc})^c(y) \geq \varphi^c(y). \tag{1.16}$$

Από τις σχέσεις (1.15) και (1.16), η απόδειξη του ισχυρισμού είναι πλήρης.

■

Αν η c είναι φραγμένη και συνεχής, η φ^c είναι το infimum μιας οικογένειας συνεχών συναρτήσεων, άρα είναι άνω ημισυνεχής. Συνεπώς, η φ^c είναι μετρήσιμη. Ομοίως, η φ^{cc} είναι μετρήσιμη.

Παρατήρηση 1.10 (Εκτιμήσεις για φραγμένες συναρτήσεις κόστους)

Θα αποδείξουμε κάποιες εκτιμήσεις για τις παραπάνω συναρτήσεις φ^c και φ^{cc}, στην περίπτωση που ικανοποιούνται οι υποθέσεις της Παρατήρησης 1.9 και $\varphi = \varphi^{cc}$.

$$\varphi^c(y) = \inf_{x\in X}[c(x,y) - \varphi(x)] \leq \inf_{x\in X}[\|c\|_\infty - \varphi(x)] = \|c\|_\infty - \sup_{x\in X}\varphi(x).$$

Από την άλλη, $0 \leq c(x,y)$ και $\varphi(x) \leq \sup_{x\in X}\varphi(x)$, οπότε έχουμε $c(x,y) - \varphi(x) \geq 0 - \sup_{x\in X}\varphi(x)$.

Άρα,

$$\varphi^c(y) = \inf_{x\in X}[c(x,y) - \varphi(x)] \geq -\sup_{x\in X}\varphi(x).$$

Τελικά, έχουμε την ακόλουθη εκτίμηση για την φ^c:

$$-\sup\varphi \leq \varphi^c \leq \|c\|_\infty - \sup\varphi. \tag{1.17}$$

Ομοίως για την φ^{cc}:

$$-\sup\varphi^c \leq \varphi^{cc} \leq \|c\|_\infty - \sup\varphi^c. \tag{1.18}$$

Επίσης, αφού για κάθε $s \in \mathbb{R}$: $J(\varphi + s, \psi - s) = J(\varphi,\psi)$, και

$$(\varphi + s)^c(y) = \inf_{x\in X}[c(x,y) - \varphi(x) - s] = \inf_{x\in X}[c(x,y) - \varphi(x)] - s = \varphi^c(y) - s,$$

μπορούμε να υποθέσουμε, χωρίς βλάβη της γενικότητας, ότι $\sup\varphi = \|c\|_\infty$. Τότε, από τη σχέση (1.17) προκύπτει ότι: $-\|c\|_\infty \leq \varphi^c \leq 0$.

Από τη σχέση (1.18) έπεται, αφού $\varphi = \varphi^{cc}$, ότι $\sup\varphi - \inf\varphi \leq \|c\|_\infty$, άρα, $\inf\varphi \geq 0$.

Με βάση τα αποτελέσματα αυτά καθώς και την Παρατήρηση 1.9, στην περίπτωση που εξετάζουμε, το supremum στο δεξί μέλος της σχέσης (1.4) μπορεί να περιοριστεί και ως εξής:

$$\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = \sup\{J(\varphi,\psi) : (\varphi,\psi) \in \Phi_c, 0 \leq \varphi \leq \|c\|_\infty, -\|c\|_\infty \leq \psi \leq 0\}.$$

1.2. Μετρικές συναρτήσεις κόστους.

Στη δεύτερη ενότητα του κεφαλαίου, θα ασχοληθούμε με την περίπτωση όπου οι χώροι X, Y ταυτίζονται και η συνάρτηση κόστους είναι μια μετρική $c(x,y) = d(x,y)$ στον $X = Y$.

1.2.1. Το Θεώρημα Kantorovich – Rubinstein

Θεώρημα 1.11 (Θεώρημα Kantorovich – Rubinstein)

Έστω $X = Y$ ένας Πολωνικός χώρος, $\mu, \nu \in P(X)$ και έστω $d: X \times X \to [0, +\infty)$ μια κάτω ημι-συνεχής μετρική στον X. Υποθέτουμε ότι για κάθε μετρήσιμη φραγμένη συνάρτηση $\varphi: X \longrightarrow \mathbb{R}$, η φ^d είναι μετρήσιμη. Έστω, επίσης, T_d το κόστος της βέλτιστης μεταφοράς αναφορικά με τη συνάρτηση κόστους $c(x,y) = d(x,y)$, δηλαδή

$$T_d(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times X} d(x,y) d\pi(x,y).$$

Συμβολίζουμε με $Lip(X)$ τον χώρο όλων των d —Lipschitz συναρτήσεων στον X, και ορίζουμε

$$\|\varphi\|_{Lip} = \sup_{x\neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)}.$$

Τότε,

$$T_d(\mu,\nu) = \sup\left\{\int_X \varphi\, d(\mu-\nu) :\ \varphi \in L^1(d|\mu-\nu|), \qquad \|\varphi\|_{Lip} \leq 1\right\}.$$

Επιπλέον, μπορούμε να περιορίσουμε το παραπάνω supremum στις φραγμένες συναρτήσεις φ χωρίς να μεταβληθεί η τιμή του.

Η ποσότητα $T_d(\mu,\nu)$ λέγεται και απόσταση Kantorovich – Rubinstein ανάμεσα στα μέτρα πιθανότητας μ, ν.

<u>Απόδειξη του Θεωρήματος 1.11:</u>

Για κάθε $n \in \mathbb{N}$ ορίζουμε:

$$d_n = \frac{d}{1 + \frac{d}{n}}.$$

Η d_n είναι μια κάτω ημισυνεχής μετρική που, για κάθε $n \in \mathbb{N}$, ικανοποιεί την ανισότητα $d_n \leq d$, και επίσης, για κάθε $x, y \in X$ ισχύει $d_n(x,y) \nearrow d(x,y)$ καθώς το $n \to \infty$. Ειδικότερα, αν

$$\sup_{x\neq y}\frac{|\varphi(x)-\varphi(y)|}{d_n(x,y)}\leq 1,$$

τότε, αφού $d_n \leq d$,

$$\sup_{x\neq y}\frac{|\varphi(x)-\varphi(y)|}{d(x,y)}\leq 1,$$

δηλαδή το σύνολο των 1-Lipschitz συναρτήσεων για τη μετρική d_n περιλαμβάνεται στο σύνολο των 1-Lipschitz συναρτήσεων για τη μετρική d. Λαμβάνοντας υπ' όψιν ότι, λόγω του Θεωρήματος 1.3, το $T_d(\mu,\nu)$ λαμβάνεται ως ολοκλήρωμα για κάποιο $\pi \in \Pi(\mu,\nu)$ μιμούμενοι το 3ο σκέλος της απόδειξης του Θεωρήματος 1.3, βλέπουμε ότι αρκεί να αποδείξουμε το Θεώρημα 1.11 αντικαθιστώντας την d με την d_n. Συνεπώς, μπορούμε στη συνέχεια, να υποθέσουμε, χωρίς βλάβη της γενικότητας, ότι η μετρική d είναι φραγμένη, δηλαδή

$$\exists\, M>0 \ \text{τ.ω}\ \forall x,y\in X:\ \ d(x,y)\leq M.$$

Σ' αυτή την περίπτωση, όλες οι Lipschitz συναρτήσεις είναι φραγμένες, αφού αν επιλέξουμε κάποιο $x_0 \in X$, τότε

$$\forall x\in X,\quad |\varphi(x)-\varphi(x_0)|\leq d(x,x_0)\leq M$$

και επομένως,

$$\forall x\in X,\quad |\varphi(x)|\leq M+|\varphi(x_0)|.$$

Τότε, επίσης, όλες οι Lipschitz συναρτήσεις είναι μετρήσιμες (αφού είναι συνεχείς) και ολοκληρώσιμες αναφορικά με τα μέτρα πιθανότητας μ,ν. Σύμφωνα με το Θεώρημα 1.3,

$$T_d(\mu,\nu)=\sup_{(\varphi,\psi)\in\Phi_d} J(\varphi,\psi),$$

και συνεπώς, για να αποδείξουμε το Θεώρημα 1.11, αρκεί να δείξουμε ότι

$$\sup_{(\varphi,\psi)\in\Phi_d} J(\varphi,\psi)=\sup\left\{\int_X \varphi\, d(\mu-\nu)\ :\ \|\varphi\|_{Lip}\leq 1\right\}$$

όπου J είναι το γνωστό συναρτησοειδές

$$J(\varphi,\psi)=\int_X \varphi\, d\mu+\int_Y \psi\, d\nu.$$

Από την παρατήρηση 1.9, γνωρίζουμε ότι

$$\sup_{(\varphi,\psi)\in\Phi_d} J(\varphi,\psi)=\sup_{\varphi\in C_b(X)} J(\varphi^{dd},\varphi^{d})$$

όπου

$$\varphi^d(y)\coloneqq\inf_{x\in X}[d(x,y)-\varphi(x)]\quad\text{και}\quad \varphi^{dd}(x)\coloneqq\inf_{y\in X}[d(x,y)-\varphi^d(y)].$$

Στη συνέχεια, θα αποδείξουμε ότι για κάθε φραγμένη φ, η συνάρτηση φ^d είναι 1-Lipschitz. Πράγματι, έστω $\varepsilon > 0$ και $y_2 \in X$. Από τον ορισμό του infimum, θα υπάρχει $x_0 \in X$ τέτοιο ώστε

$$d(x_0, y_2) - \varphi(x_0) < \inf_{x \in X}[d(x, y_2) - \varphi(x)] + \varepsilon$$

οπότε,

$$-\inf_{x \in X}[d(x, y_2) - \varphi(x)] < -\big(d(x_0, y_2) - \varphi(x_0)\big) + \varepsilon.$$

Επομένως, για κάθε $y_1, y_2 \in X$ έχουμε:

$$\begin{aligned}\varphi^d(y_1) - \varphi^d(y_2) &= \inf_{x \in X}[d(x, y_1) - \varphi(x)] - \inf_{x \in X}[d(x, y_2) - \varphi(x)] \\ &< \inf_{x \in X}[d(x, y_1) - \varphi(x)] - \big(d(x_0, y_2) - \varphi(x_0)\big) + \varepsilon \\ &\leq d(x_0, y_1) - \varphi(x_0) - d(x_0, y_2) + \varphi(x_0) + \varepsilon \leq d(y_1, y_2) + \varepsilon,\end{aligned}$$

αφού, λόγω τριγωνικής ανισότητας, $d(x_0, y_1) - d(x_0, y_2) \leq d(y_1, y_2)$. Προφανώς, ισχύει η ίδια ανισότητα αν εναλλάξουμε τους ρόλους των y_1 και y_2.

Εφόσον, το ε ήταν ένας τυχαίος θετικός αριθμός, συμπεραίνουμε ότι

$$\forall y_1, y_2 \in X, \quad |\varphi^d(y_1) - \varphi^d(y_2)| \leq d(y_1, y_2)$$

δηλαδή η συνάρτηση φ^d είναι 1-Lipschitz.

Στη συνέχεια, λόγω αυτής της ιδιότητας της φ^d, για κάθε $x, y \in X$, έχουμε $|\varphi^d(x) - \varphi^d(y)| \leq d(x, y)$, άρα $-d(x, y) \leq \varphi^d(x) - \varphi^d(y) \leq d(x, y)$.

Οπότε, από την πρώτη ανισότητα, έχουμε $-\varphi^d(x) \leq d(x, y) - \varphi^d(y)$, άρα και

$$-\varphi^d(x) \leq \inf_{y \in X}[d(x, y) - \varphi^d(y)] = \varphi^{dd}(x).$$

Επίσης,

$$\varphi^{dd}(x) = \inf_{x \in X}[d(x, y) - \varphi^d(x)] \leq -\varphi^d(x),$$

το οποίο προκύπτει επιλέγοντας $x = y$ στο infimum.

Τελικά, $-\varphi^d(x) \leq \varphi^{dd}(x) \leq -\varphi^d(x)$, από όπου προκύπτει ότι $\varphi^{dd} = -\varphi^d$.

Έτσι, έχουμε λοιπόν,

$$\sup_{\Phi_d} J(\varphi, \psi) = \sup_{\varphi \in C_b(X)} J(\varphi^{dd}, \varphi^d) = \sup_{\varphi \in C_b(X)} J(-\varphi^d, \varphi^d) \leq \sup_{\|\varphi\|_{Lip} \leq 1} J(\varphi, -\varphi) \leq \sup_{\Phi_d} J(\varphi, \psi),$$

όπου η προτελευταία ανισότητα ισχύει αφού η φ^d, όπως δείξαμε, είναι 1-Lipschitz. Λαμβάνοντας υπόψη ότι

$$J(\varphi, -\varphi) = \int_X \varphi d\mu + \int_X (-\varphi) d\nu = \int_X \varphi d(\mu - \nu).$$

έχουμε αποδείξει ότι

$$\sup_{\Phi_d} J(\varphi, \psi) = \sup\left\{\int_X \varphi d(\mu - \nu) \,:\, \|\varphi\|_{Lip} \leq 1\right\}$$

όπως και επιθυμούσαμε.

■

Παρατήρηση 1.12

Στην περίπτωση όπου η d είναι φραγμένη, τότε λόγω της Παρατήρησης 1.10, μπορούμε να περιορίσουμε το supremum σε φραγμένες συναρτήσεις φ που ικανοποιούν, μάλιστα, την ανισότητα: $0 \leq \varphi \leq \|d\|_\infty$.

Πόρισμα 1.13 (Αναλλοίωτο της απόστασης Kantorovich – Rubinstein ως προς αφαίρεση μάζας)

Έστω $X = Y$ ένας Πολωνικός χώρος και έστω d μια κάτω ημισυνεχής μετρική στον X. Έστω, ακόμα, μ, ν και σ τρία μη αρνητικά μέτρα Borel στον X, τέτοια ώστε $\mu[X] = \nu[X] < +\infty$, $\sigma[X] < +\infty$. Τότε,

$$T_d(\mu + \sigma, \nu + \sigma) = T_d(\mu, \nu).$$

<u>Απόδειξη:</u>

Σύμφωνα με το Θεώρημα 1.11, έχουμε:

$$\begin{aligned}
&T_d(\mu + \sigma, \nu + \sigma) \\
&= \sup\left\{\int_X \varphi\, d(\mu + \sigma - \nu - \sigma) \,:\, \varphi \in L^1(d|\mu + \sigma - \nu - \sigma|), \qquad \|\varphi\|_{Lip} \leq 1\right\} \\
&= \sup\left\{\int_X \varphi\, d(\mu - \nu) \,:\, \varphi \in L^1(d|\mu - \nu|), \qquad \|\varphi\|_{Lip} \leq 1\right\} \\
&= T_d(\mu, \nu).
\end{aligned}$$

■

Παράδειγμα 1.14 (Τύπος ολικής κύμανσης)

Έστω X ένας Πολωνικός χώρος. Τότε, οι υποθέσεις του Θεωρήματος 1.11 ικανοποιούνται από τη συνάρτηση κόστους $c(x,y) = 1_{x\neq y}$. Ακόμα, αν μ, ν είναι Borel μέτρα πιθανότητας στον X, τότε ισχύουν τα εξής:

1)

$$\inf_{\pi\in\Pi(\mu,\nu)} \pi[\{x\neq y\}] = \sup\left\{\int_X f\, d(\mu-\nu) : \ 0\leq f\leq 1, \qquad \|f\|_{Lip}\leq 1\right\}.$$

2)

$$\sup\left\{\int_X f\, d(\mu-\nu) : \ 0\leq f\leq 1, \qquad \|f\|_{Lip}\leq 1\right\} = (\mu-\nu)_+[X] = (\mu-\nu)_-[X]$$
$$= \frac{1}{2}\|\mu-\nu\|_{TV}$$

όπου $\|\cdot\|_{TV}$ είναι η νόρμα της ολικής κύμανσης, δηλαδή $\|\mu\|_{TV} = \inf\{\mu_+(X)+\mu_-(X)\}$ και το infimum λαμβάνεται πάνω σε όλα τα μη αρνητικά μέτρα μ_+, μ_- τέτοια ώστε $\mu = \mu_+ - \mu_-$.

3)

$$(\mu-\nu)_+[X] = \sup_{\substack{K\subseteq X\\ K \text{ συμπαγές}}} (\mu-\nu)_+[K] = \sup_{\substack{K\subseteq X\\ K \text{ συμπαγές}}} (\mu[K]-\nu[K]).$$

<u>Απόδειξη:</u>

Θα πρέπει να δείξουμε ότι η $c(x,y) = 1_{x\neq y}$ είναι μια κάτω ημι-συνεχής μετρική, δηλαδή

$$\forall (x,y)\in X\times X, \quad c(x,y)\leq \liminf_{(z,w)\to(x,y)} c(z,w).$$

Έστω, λοιπόν, $(x,y)\in X\times X$.

Αν $x=y$, τότε $c(x,y) = 0 \leq c(z,w),\ \ \forall (z,w)\in X\times X$. Άρα, $c(x,y)\leq \liminf_{(z,w)\to(x,y)} c(z,w)$.

Αν $x\neq y$, τότε $c(x,y)=1$ και υπάρχει περιοχή $N_\varepsilon(x,y)$ του σημείου (x,y) ώστε $c(z,w) = 1,\ \ \forall (z,w)\in N_\varepsilon(x,y)$. Άρα, $\liminf_{(z,w)\to(x,y)} c(z,w) = 1 = c(x,y)$.

Επιπλέον, έστω $\varphi : X\longrightarrow \mathbb{R}$ μετρήσιμη και φραγμένη συνάρτηση. Τότε,

$$\varphi^c(y) = \inf(\{1-\varphi(x) : \ \ x\neq y\}\cup\{-\varphi(y)\}).$$

Άρα,

$$\varphi^c(y) = \begin{cases} -\varphi(y), & -\varphi(y)\leq \inf(1-\varphi) \\ \inf(1-\varphi), & \text{αλλιώς} \end{cases}$$

και συνεπώς, η φ^c είναι μετρήσιμη.

Εφόσον, λοιπόν, οι υποθέσεις του Θεωρήματος 1.11 ικανοποιούνται, από την εφαρμογή του έχουμε:

$$T_d(\mu,\nu) = \sup\left\{\int_X \varphi\, d(\mu-\nu) :\ \varphi \in L^1(d|\mu-\nu|), \|\varphi\|_{Lip} \leq 1\right\}$$

άρα

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times X} 1_{x\neq y}\, d\pi = \sup\left\{\int_X \varphi\, d(\mu-\nu) :\ 0 \leq \varphi \leq 1, \|\varphi\|_{Lip} \leq 1\right\}$$

άρα

$$\inf_{\pi\in\Pi(\mu,\nu)} \pi[\{x \neq y\}] = \sup\left\{\int_X \varphi\, d(\mu-\nu) :\ 0 \leq \varphi \leq 1, \|\varphi\|_{Lip} \leq 1\right\}.$$

Ο περιορισμός του supremum στις συναρτήσεις που ικανοποιούν την ανισότητα $0 \leq \varphi \leq 1$, είναι δυνατός λόγω της Παρατήρησης 1.12, αφού η μετρική $1_{x\neq y}$ είναι φραγμένη.

Στη συνέχεια, θα δείξουμε ότι

$$\sup\left\{\int_X f\, d(\mu-\nu) :\ \ 0 \leq f \leq 1, \qquad \|f\|_{Lip} \leq 1\right\} = (\mu-\nu)_+[X].$$

Χρησιμοποιώντας την ανάλυση Hahn, παίρνουμε $X = A \cup B$, όπου A, B είναι σύνολα Borel με $A \cap B = \emptyset$ και τέτοια ώστε σε κάθε υποσύνολο του A, $\mu - \nu \geq 0$, ενώ σε κάθε υποσύνολο του B, $\mu - \nu \leq 0$. Όπως είναι γνωστό, ισχύει

$$(\mu-\nu)_+[X] = (\mu-\nu)[A].$$

Τότε, για οποιαδήποτε συνάρτηση f, με $0 \leq f \leq 1$, θα έχουμε

$$\int_X f\, d(\mu-\nu) = \int_A f\, d(\mu-\nu) + \int_B f\, d(\mu-\nu)$$

$$\leq \int_A f\, d(\mu-\nu) \leq \int_A 1\, d(\mu-\nu) = (\mu-\nu)[A] = (\mu-\nu)_+[X].$$

Για την αντίστροφη ανισότητα, έχουμε

$$(\mu-\nu)_+[X] = (\mu-\nu)[A] = \int_X 1_A\, d(\mu-\nu)$$

και επίσης, $\|1_A\|_{Lip} \leq 1$, αφού

$$\|1_A\|_{Lip} = \sup_{x\neq y} \frac{|1_A(x) - 1_A(y)|}{1} = |1_A(x) - 1_A(y)| \leq 1.$$

Άρα,

$$(\mu-\nu)_+[X] \leq \sup\left\{\int_X f\,d(\mu-\nu):\ \ 0\leq f\leq 1, \qquad \|f\|_{Lip}\leq 1\right\}.$$

Έπειτα,

$$0=(\mu-\nu)[X]=(\mu-\nu)_+[X]-(\mu-\nu)_-[X]$$

άρα

$$(\mu-\nu)_+[X]=(\mu-\nu)_-[X]=\frac{1}{2}\|\mu-\nu\|_{TV}$$

αφού $\|\mu-\nu\|_{TV}=(\mu-\nu)_+[X]+(\mu-\nu)_-[X]$.

Θα αποδείξουμε, ακόμα, την ισότητα

$$(\mu-\nu)_+[X]=\sup\{(\mu-\nu)_+[K]:K\subseteq X \text{ συμπαγές}\}$$

$$=\sup\{\mu[K]-\nu[K]:\ K\subseteq X \text{ συμπαγές}\}.$$

Η πρώτη ισότητα προκύπτει από την κανονικότητα του μέτρου $(\mu-\nu)_+$, οπότε συνεχίζουμε στην απόδειξη της δεύτερης.

Θεωρούμε, πάλι, τα σύνολα A, B όπως προηγουμένως. Τότε,

$$(\mu-\nu)_+[X]=(\mu-\nu)[A]=\sup\{(\mu-\nu)[K]:K\subseteq A \text{ συμπαγές}\}$$
$$\leq \sup\{(\mu-\nu)[K]:K\subseteq X \text{ συμπαγές}\}=\sup\{\mu[K]-\nu[K]:\ K\subseteq X \text{ συμπαγές}\}$$

Για την αντίστροφη ανισότητα, αν $K\subseteq X$ συμπαγές, τότε

$$(\mu-\nu)_+[X]\geq(\mu-\nu)_+[K]\geq(\mu-\nu)[K]=\mu[K]-\nu[K].$$

Άρα,

$$(\mu-\nu)_+[X]\geq \sup\{\mu[K]-\nu[K]:\ K\subseteq X \text{ συμπαγές}\}.$$

■

Παράδειγμα 1.15 (Ο δυϊκός του $C^{0,a}$)

Έστω Ω ένα φραγμένο ανοικτό υποσύνολο του $\mathbb{R}^n$, $a\in(0,1)$, και έστω $E=C^{0,a}(\Omega)$ ο χώρος Banach των a −Hölder συνεχών πραγματικών συναρτήσεων στο Ω, εφοδιασμένος με τη νόρμα

$$\|f\|_E=\sup_{x\in\Omega}|f(x)|+\sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|^a}.$$

Φυσικά, μπορούμε να θεωρήσουμε το χώρο $M(\Omega)$ ως ένα υποσύνολο του E^*, εφοδιασμένο με τη νόρμα $\|l\|_{E^*} = \sup_{\|f\|_E \le 1} \langle l, f \rangle$.

Έστω μ, ν δύο μέτρα πιθανότητας στο E^* και $c(x,y) = |x-y|^a$. Θα αποδείξουμε ότι

$$\|\mu - \nu\|_{E^*} \le \min\{\|\mu - \nu\|_{TV}, T_c(\mu, \nu)\}.$$

<u>Απόδειξη:</u>

Αρκεί να δείξουμε ότι $\|\mu - \nu\|_{E^*} \le \|\mu - \nu\|_{TV}$ και $\|\mu - \nu\|_{E^*} \le T_c(\mu, \nu)$.

Για την απόδειξη της πρώτης ανισότητας, έχουμε:

αν

$$\|f\|_E = \sup_{x \in \Omega} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^a} \le 1,$$

τότε και

$$\sup_{x \in \Omega} |f(x)| \le 1.$$

Άρα,

$$\begin{aligned}
\|\mu - \nu\|_{E^*} &= \sup_{\|f\|_E \le 1} \langle \mu - \nu, f \rangle = \sup_{\|f\|_E \le 1} \int_\Omega f \, d(\mu - \nu) \le \sup_{\|f\|_E \le 1} \left| \int_\Omega f d(\mu - \nu) \right| \\
&\le \sup_{\|f\|_E \le 1} \left\{ \left| \int_\Omega f d(\mu - \nu)_+ - \int_\Omega f d(\mu - \nu)_- \right| \right\} \le \sup_{\|f\|_E \le 1} \left\{ \left| \int_\Omega f d(\mu - \nu)_+ \right| + \left| \int_\Omega f d(\mu - \nu)_- \right| \right\} \\
&\le \sup_{\|f\|_E \le 1} \{(\mu - \nu)_+[\Omega] + (\mu - \nu)_-[\Omega]\} = \sup_{\|f\|_E \le 1} \{\|\mu - \nu\|_{TV}\} = \|\mu - \nu\|_{TV}.
\end{aligned}$$

Προχωράμε στην απόδειξη της δεύτερης ανισότητας.

Ισχυρισμός 9. Η συνάρτηση κόστους $c(x,y) = |x-y|^a$ είναι μια κάτω ημισυνεχής μετρική στο $\Omega \subseteq \mathbb{R}^n$.

<u>Απόδειξη του Ισχυρισμού 9:</u>

Πράγματι, η c είναι συνεχής, άρα και κάτω ημισυνεχής. Για να δούμε ότι είναι μετρική, αρκεί να επαληθεύσουμε την τριγωνική ανισότητα. Για κάθε $x, y, z \in \Omega$ ισχύει:

$$|x-y|^a \le (|x-z| + |z-y|)^a \le |x-z|^a + |z-y|^a,$$

όπου για να επαληθεύσουμε την τελευταία ανισότητα, θα αποδείξουμε ότι

$$\forall x, y \ge 0, \quad (x+y)^a \le x^a + y^a.$$

Αν $x = y = 0$, η ανισότητα είναι προφανής. Υποθέτουμε ότι $y \neq 0$. Διαιρώντας με y^a, έχουμε:

$$\left(\frac{x}{y}+1\right)^a \leq \left(\frac{x}{y}\right)^a + 1.$$

Θέτουμε $z = \frac{x}{y}$, οπότε έχουμε να αποδείξουμε την εξής ανισότητα:

$$\forall z \geq 0 : \ (z+1)^a \leq z^a + 1.$$

Θα μελετήσουμε ως προς τη μονοτονία τη συνάρτηση

$$f(z) = z^a + 1 - (z+1)^a, \quad z \geq 0, \qquad a \in (0,1).$$

Για κάθε $z > 0$, η συνάρτηση f είναι παραγωγίσιμη με παράγωγο

$$f'(z) = az^{a-1} - a(z+1)^{a-1} = a(z^{a-1} - (z+1)^{a-1}).$$

Για κάθε $z > 0$, έχουμε $f'(z) > 0$, γιατί εφόσον $a \in (0,1)$ ισχύει η ανισότητα $z^{a-1} - (z+1)^{a-1} > 0$. Επομένως, η συνάρτηση f είναι γνησίως αύξουσα στο $[0, +\infty)$, οπότε

$$\forall z \geq 0, \quad f(z) \geq f(0) = 0$$

άρα

$$\forall z \geq 0, \quad z^a + 1 - (z+1)^a \geq 0.$$

■

Επομένως, μπορούμε να εφαρμόσουμε το Θεώρημα 1.11, γιατί το Ω, ως ανοικτό υποσύνολο Πολωνικού χώρου, είναι επίσης, Πολωνικός χώρος. Έχουμε, λοιπόν:

$$\|\mu - \nu\|_{E^*} = \sup_{\|f\|_E \leq 1} \langle \mu - \nu, f \rangle = \sup_{\|f\|_E \leq 1} \int_\Omega f \, d(\mu - \nu) \leq \sup_{\|f\|_{Lip} \leq 1} \int_\Omega f \, d(\mu - \nu) = T_c(\mu, \nu)$$

όπου η ανισότητα ισχύει αφού:

$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{c(x,y)} \leq \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^a} + \sup_{x \in \Omega} |f(x)| = \|f\|_E.$$

Άρα, αν $\|f\|_E \leq 1$, τότε ισχύει $\|f\|_{Lip} \leq 1$, δηλαδή

$$\{f : \Omega \to \mathbb{R} : \ \|f\|_E \leq 1\} \subseteq \left\{f : \Omega \to \mathbb{R} : \ \|f\|_{Lip} \leq 1\right\}.$$

■

1.2.2. Μεταφόρτωση.

Σύμφωνα με το Θεώρημα Kantorovich – Rubinstein, το οποίο διατυπώσαμε και αποδείξαμε παραπάνω, όταν $X = Y$ είναι ένας Πολωνικός χώρος, το βέλτιστο κόστος μεταφοράς εξαρτάται μόνο από τη διαφορά των μέτρων $\mu - \nu$. Συνεπώς, αν $\mu, \nu \in P(X)$ και η συνάρτηση κόστους είναι μια μετρική στον X, το πρόβλημα βέλτιστης μεταφοράς του Kantorovich είναι ισοδύναμο με το **πρόβλημα μεταφόρτωσης των Kantorovich – Rubinstein**, που ορίζεται ως εξής:

Ελαχιστοποιήστε το συναρτησοειδές

$$I[\pi] = \int_{X\times X} d(x,y)d\pi(x,y)$$

πάνω σε όλα τα μέτρα π που, για κάθε $A \in \mathcal{B}_X$, ικανοποιούν τη σχέση

$$\pi[A\times X] - \pi[X\times A] = (\mu - \nu)[A].$$

Ας θυμηθούμε τη συνθήκη που απαιτήσαμε ώστε $\pi \in \Pi(\mu,\nu)$, η οποία ήταν $\pi[A\times X] = \mu[A]$ και $\pi[X\times A] = \nu[A]$ για κάθε $A \in \mathcal{B}_X$. Συγκρίνοντας τις δύο συνθήκες για τα προβλήματα μεταφοράς και μεταφόρτωσης, βλέπουμε ότι, για μια γενική συνάρτηση κόστους, το πρόβλημα μεταφόρτωσης είναι μια αρκετά πιο χαλαρή εκδοχή του προβλήματος μεταφοράς.

Για να γίνει αυτό πιο κατανοητό, θα αποδείξουμε ότι στην περίπτωση που η συνάρτηση κόστους είναι η $c(x,y) = |x-y|^2$ στον $\mathbb{R}^n$, η οποία βέβαια δεν είναι μετρική, το βέλτιστο κόστος μεταφόρτωσης είναι ίσο με 0. Το ίδιο συμβαίνει και όταν $c(x,y) = |x-y|^p$, με $p > 1$.

Παράδειγμα 1.16 (Η μεταφόρτωση κάποιες φορές δεν κοστίζει σχεδόν τίποτα)

Θεωρούμε τη συνάρτηση κόστους $c(x,y) = |x-y|^2$ στον $\mathbb{R}^n$ και μ,ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ τέτοια ώστε $T_c(\mu,\nu) < +\infty$. Έστω, ακόμα, $\pi \in \Pi(\mu,\nu)$ ένα πλάνο μεταφοράς τέτοιο ώστε $I[\pi] < +\infty$. Μπορούμε, βέβαια, να θεωρήσουμε το π και ως ένα πλάνο μεταφόρτωσης με το αντίστοιχο κόστος μεταφόρτωσης $I[\pi]$. Προκειμένου να ελαττώσουμε το κόστος μεταφοράς, εφαρμόζουμε την ακόλουθη στρατηγική:

Αν x και y είναι ένα αρχικό και ένα τελικό σημείο, αντίστοιχα, τότε αντί να μεταφέρουμε το x στο y, μεταφέρουμε το x στο $\frac{x+y}{2}$ και ταυτόχρονα το $\frac{x+y}{2}$ στο y.

Ισχυρισμός 10. Η παραπάνω στρατηγική εφαρμόζεται με ένα αποδεκτό πλάνο μεταφόρτωσης και το κόστος μεταφόρτωσης υποδιπλασιάζεται. Η διαδοχική εφαρμογή αυτής της στρατηγικής οδηγεί στο συμπέρασμα ότι το βέλτιστο κόστος μεταφόρτωσης είναι ίσο με 0.

<u>Απόδειξη:</u>

Θεωρούμε τις συναρτήσεις

$$\varphi: \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \times \mathbb{R}^n \ \text{ με } \ \varphi(x,y) = \left(x, \frac{x+y}{2}\right)$$

$$\psi: \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \times \mathbb{R}^n \ \text{ με } \ \psi(x,y) = \left(\frac{x+y}{2}, y\right)$$

και ορίζουμε το μέτρο $\pi_1 \coloneqq \varphi(\pi) + \psi(\pi)$.

Θα δείξουμε ότι το π_1 είναι ένα αποδεκτό πλάνο μεταφόρτωσης, δηλαδή ότι για κάθε $A \subseteq \mathbb{R}^n$ μετρήσιμο, ισχύει:

$$\pi_1[A \times \mathbb{R}^n] - \pi_1[\mathbb{R}^n \times A] = (\mu - \nu)[A].$$

Πράγματι, ισχύει

$$\varphi^{-1}(A \times \mathbb{R}^n) = \{(x, y): \ x \in A\} = A \times \mathbb{R}^n$$

$$\varphi^{-1}(\mathbb{R}^n \times A) = \left\{(x, y): \frac{x + y}{2} \in A\right\} = \psi^{-1}(A \times \mathbb{R}^n)$$

$$\psi^{-1}(A \times \mathbb{R}^n) = \{(x, y): \ y \in A\} = \mathbb{R}^n \times A$$

οπότε έχουμε:

$$\begin{aligned} &\pi_1[A \times \mathbb{R}^n] - \pi_1[\mathbb{R}^n \times A] \\ &\quad = \pi[\varphi^{-1}(A \times \mathbb{R}^n)] + \pi[\psi^{-1}(A \times \mathbb{R}^n)] - \pi[\varphi^{-1}(\mathbb{R}^n \times A)] - \pi[\psi^{-1}(\mathbb{R}^n \times A)] \\ &\quad = \pi[A \times \mathbb{R}^n] - \pi[\mathbb{R}^n \times A] = \mu[A] - \nu[A] = (\mu - \nu)[A]. \end{aligned}$$

Δείχνουμε στη συνέχεια, ότι με την εφαρμογή της συγκεκριμένης στρατηγικής το κόστος μεταφόρτωσης υποδιπλασιάζεται.

$$\begin{aligned} I[\pi_1] &= \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\pi_1 = \int_{\mathbb{R}^n \times \mathbb{R}^n} c\big(\varphi(x, y)\big) d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} c\big(\psi(x, y)\big) d\pi \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} c\left(x, \frac{x + y}{2}\right) d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} c\left(\frac{x + y}{2}, y\right) d\pi \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|x - \frac{x + y}{2}\right|^2 d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x + y}{2} - y\right|^2 d\pi = 2 \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x - y}{2}\right|^2 d\pi \\ &= \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\pi \\ &= \frac{1}{2} I[\pi]. \end{aligned}$$

Στη συνέχεια, ορίζουμε το μέτρο π_2 ως εξής: $\pi_2 \coloneqq \varphi(\pi_1) + \psi(\pi_1)$.

Για κάθε $A \subseteq \mathbb{R}^n$ μετρήσιμο, ισχύει:

$$\begin{aligned} &\pi_2[A \times \mathbb{R}^n] - \pi_2[\mathbb{R}^n \times A] \\ &\quad = \pi_1[\varphi^{-1}(A \times \mathbb{R}^n)] + \pi_1[\psi^{-1}(A \times \mathbb{R}^n)] - \pi_1[\varphi^{-1}(\mathbb{R}^n \times A)] - \pi_1[\psi^{-1}(\mathbb{R}^n \times A)] \\ &\quad = \pi_1[A \times \mathbb{R}^n] - \pi_1[\mathbb{R}^n \times A] = (\mu - \nu)[A]. \end{aligned}$$

Δηλαδή, το π_2 είναι και αυτό ένα αποδεκτό πλάνο μεταφόρτωσης. Ισχύει, όπως και προηγουμένως, ότι

$$I[\pi_2] = \frac{1}{2} I[\pi_1].$$

Με τον τρόπο αυτόν, κατασκευάζουμε μια ακολουθία πλάνων μεταφόρτωσης $(\pi_n)_{n=1}^{\infty}$ έτσι ώστε

$$\forall n \in \mathbb{N}, \quad I[\pi_{n+1}] = \frac{1}{2} I[\pi_n] = \frac{1}{2^n} I[\pi].$$

Με τη διαδοχική εφαρμογή αυτής της στρατηγικής, συμπεραίνουμε ότι $\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = 0$, δηλαδή ότι το βέλτιστο κόστος μεταφόρτωσης είναι ίσο με 0, το οποίο βέβαια δεν μπορεί να επιτευχθεί εκτός αν $\mu = \nu$.

■

Ισχυρισμός 11. Για οποιαδήποτε συνάρτηση κόστους $c(x,y) = |x-y|^p$, με $p > 1$ το βέλτιστο κόστος μεταφόρτωσης είναι ίσο με 0.

Απόδειξη:

Το κόστος για το πλάνο μεταφόρτωσης π_1 θα είναι

$$I[\pi_1] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y) d\pi_1 = \int_{\mathbb{R}^n \times \mathbb{R}^n} c\big(\varphi(x,y)\big) d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} c\big(\psi(x,y)\big) d\pi$$

$$= \int_{\mathbb{R}^n \times \mathbb{R}^n} c\left(x, \frac{x+y}{2}\right) d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} c\left(\frac{x+y}{2}, y\right) d\pi$$

$$= \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|x - \frac{x+y}{2}\right|^p d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x+y}{2} - y\right|^p d\pi = \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x-y}{2}\right|^p d\pi + \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x-y}{2}\right|^p d\pi$$

$$= 2 \int_{\mathbb{R}^n \times \mathbb{R}^n} \left|\frac{x-y}{2}\right|^p d\pi = 2^{1-p} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^p \, d\pi = 2^{1-p} I[\pi],$$

και γενικότερα, για κάθε $n \in \mathbb{N}$, το κόστος μεταφόρτωσης αναφορικά με το π_n είναι $I[\pi_n] = 2^{n(1-p)} I[\pi]$. Εφόσον $p > 1$, ισχύει $2^{n(1-p)} \to 0$, καθώς $n \to +\infty$, άρα $\lim_{n \to +\infty} I[\pi_n] = 0$. Συμπεραίνουμε ότι $\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = 0$.

■

Ισχυρισμός 12. Το συμπέρασμα του ισχυρισμού 10 παραμένει έγκυρο για οποιαδήποτε συνάρτηση $c(x,y) = f(|x-y|)$, όπου $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ είναι μια κυρτή συνάρτηση με $f(0) = 0$, $f'(0) = 0$.

<u>Απόδειξη:</u>

Το κόστος για το πλάνο μεταφόρτωσης π_1 θα είναι

$$\begin{aligned}
I[\pi_1] &= \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)d\pi_1 = \int_{\mathbb{R}^n\times\mathbb{R}^n} c\big(\varphi(x,y)\big)d\pi + \int_{\mathbb{R}^n\times\mathbb{R}^n} c\big(\psi(x,y)\big)d\pi \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} c\left(x,\frac{x+y}{2}\right)d\pi + \int_{\mathbb{R}^n\times\mathbb{R}^n} c\left(\frac{x+y}{2},y\right)d\pi \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\left|x-\frac{x+y}{2}\right|\right)d\pi + \int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\left|\frac{x+y}{2}-y\right|\right)d\pi \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\frac{|x-y|}{2}\right)d\pi + \int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\frac{|x-y|}{2}\right)d\pi = 2\int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\frac{|x-y|}{2}\right)d\pi
\end{aligned}$$

και γενικότερα, για κάθε $n \in \mathbb{N}$, το κόστος μεταφόρτωσης αναφορικά με το π_n είναι

$$I[\pi_n] = 2^n \int_{\mathbb{R}^n\times\mathbb{R}^n} f\left(\frac{|x-y|}{2^n}\right)d\pi = \int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{f\left(\frac{|x-y|}{2^n}\right) - f(0)}{\frac{1}{2^n}}d\pi$$

αφού $f(0) = 0$.

Στη συνέχεια, για κάθε $x, y \in \mathbb{R}^n$ έχουμε:

$$\lim_{n\to+\infty} \frac{f\left(\frac{|x-y|}{2^n}\right) - f(0)}{\frac{1}{2^n}} = |x-y| \lim_{n\to+\infty} \frac{f\left(\frac{|x-y|}{2^n}\right) - f(0)}{\frac{|x-y|}{2^n}} = |x-y|f'(0) = 0.$$

Ακόμα, εφόσον η f είναι κυρτή, έχουμε:

$$f\left(\frac{|x-y|}{2^{n+1}}\right) = f\left(\frac{1}{2}\frac{|x-y|}{2^n} + \frac{1}{2}0\right) \le \frac{1}{2}f\left(\frac{|x-y|}{2^n}\right) + \frac{1}{2}f(0) = \frac{1}{2}f\left(\frac{|x-y|}{2^n}\right).$$

Άρα,

$$f\left(\frac{|x-y|}{2^n}\right) \ge 2f\left(\frac{|x-y|}{2^{n+1}}\right)$$

και πολλαπλασιάζοντας με 2^n, προκύπτει

$$\frac{f\left(\frac{|x-y|}{2^n}\right) - f(0)}{\frac{1}{2^n}} \ge \frac{f\left(\frac{|x-y|}{2^{n+1}}\right) - f(0)}{\frac{1}{2^{n+1}}}.$$

Επομένως, από το Θεώρημα Μονότονης Σύγκλισης, επειδή $I[\pi] < +\infty$, έχουμε

$$\lim_{n\to+\infty} I[\pi_n] = \lim_{n\to+\infty} \int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{f\left(\frac{|x-y|}{2^n}\right) - f(0)}{\frac{1}{2^n}} \, d\pi = \int_{\mathbb{R}^n\times\mathbb{R}^n} 0 \, d\pi = 0.$$

Τελικά, συμπεραίνουμε ότι $\inf\limits_{\pi\in\Pi(\mu,\nu)} I[\pi] = 0$.

■

1.3. Το επιχείρημα δυϊκότητας στον $C_b(X \times Y)$.

Στην πρώτη ενότητα και ειδικότερα στο 1ο σκέλος της απόδειξης του Θεωρήματος 1.3, χρησιμοποιήσαμε το Θεώρημα δυϊκότητας των Fenchel – Rockafellar στην περίπτωση που οι χώροι X, Y ήταν συμπαγείς. Στην παρούσα ενότητα, θα ασχοληθούμε με την εφαρμογή του Θεωρήματος δυϊκότητας στην περίπτωση που οι X, Y είναι τοπικά συμπαγείς χώροι και όχι συμπαγείς. Αυτό θα είναι το αντικείμενο της Πρότασης 1.17, η οποία αποτελεί ειδική περίπτωση του Θεωρήματος 1.3. Εδώ θα δώσουμε μια διαφορετική απόδειξη, για την οποία θα χρειαστούμε τα λήμματα 1.20 και 1.21.

Πρόταση 1.17 (Δυϊκότητα Kantorovich)

Έστω X, Y δύο τοπικά συμπαγείς Πολωνικοί χώροι, $c: X \times Y \longrightarrow [0, +\infty)$ μια κάτω ημισυνεχής συνάρτηση, και μ, ν δύο Borel μέτρα πιθανότητας στους X, Y αντίστοιχα. Τότε, ισχύει

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi)\in\Phi_c\cap C_b} J(\varphi, \psi). \tag{1.19}$$

Για την απόδειξη, θα στηριχθούμε στο 1ο σκέλος της απόδειξης του Θεωρήματος 1.3. Θα επιλέξουμε $E = C_b(X \times Y)$.

Αν $l \in E^*$ είναι ένα συνεχές γραμμικό συναρτησοειδές, τότε και ο περιορισμός του πάνω στον $C_0(X \times Y)$ $\big($συμβ. $l|_{C_0(X\times Y)}\big)$ είναι, επίσης, συνεχής. Επομένως, από το Θεώρημα Αναπαράστασης του Riesz (Θεώρημα 22, Εισαγωγικό κεφάλαιο), υπάρχει μοναδικό κανονικό προσημασμένο μέτρο $\pi \in M(X \times Y)$ τέτοιο ώστε:

$$\langle l, u\rangle = \int_{X\times Y} u \, d\pi, \qquad u \in C_0(X \times Y).$$

Έστω $k \in E^*$ το συνεχές γραμμικό συναρτησοειδές που ορίζει το μέτρο π. Το k ορίζεται από τη σχέση

$$\langle k, u\rangle = \int_{X\times Y} u\, d\pi, \qquad u \in C_b(X\times Y).$$

Τότε, οι περιορισμοί των l και k πάνω στον $C_0(X\times Y)$ ταυτίζονται, δηλαδή $l|_{C_0(X\times Y)} = k|_{C_0(X\times Y)}$.

Αν, λοιπόν, ορίσουμε $R = l - k$, τότε αυτό είναι ένα συνεχές γραμμικό συναρτησοειδές, για το οποίο ισχύει $R|_{C_0(X\times Y)} = 0$.

Άρα, μπορούμε να γράψουμε $l = k + R$, ή αλλιώς $l = \pi + R$, όπου το R είναι ένα συνεχές γραμμικό συναρτησοειδές συγκεντρωμένο στο άπειρο, το οποίο σημαίνει ότι

$$u \in C_0(X\times Y) \Longrightarrow \langle R, u\rangle = 0.$$

Το παράδειγμα που ακολουθεί σκοπό έχει να μας δείξει ότι μερικά από αυτά τα συναρτησοειδή R μπορεί να είναι αρκετά περίεργα και συνεπώς θα πρέπει να τα αντιμετωπίζουμε με προσοχή.

Ορισμός 1.18

Μια συνάρτηση $u \in C_b(X)$ δέχεται ένα όριο $u(\infty)$ στο άπειρο, αν για κάθε $\varepsilon > 0$, υπάρχει συμπαγές $K_\varepsilon \subseteq X$ τέτοιο ώστε $x \notin K_\varepsilon \Longrightarrow |u(x) - u(\infty)| \le \varepsilon$.

Παράδειγμα 1.19

Έστω X, Y τοπικά συμπαγείς Πολωνικοί χώροι, $\mu \in P(X)$, $\nu \in P(Y)$. Τότε, υπάρχει συνεχές γραμμικό συναρτησοειδές l στο $C_b(X\times Y)$ συγκεντρωμένο στο άπειρο έτσι ώστε

$$\forall(\varphi,\psi) \in C_0(X)\times C_0(Y), \quad \langle l, \varphi+\psi\rangle = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu. \tag{1.20}$$

Σημειώνουμε ότι η παραπάνω ισότητα δεν εγγυάται ότι $l \in \Pi(\mu,\nu)$, καθώς το l δεν είναι κατ' ανάγκην μη αρνητικό.

Απόδειξη:

Έστω

$$E = \left\{ \begin{array}{c} u \in C_b(X\times Y) : \\ \forall x \in X \text{ υπάρχει το } u(x,\infty) \text{ και η συνάρτηση } x \longmapsto u(x,\infty) \text{ είναι μετρήσιμη} \\ \text{και } \forall y \in Y \text{ υπάρχει το } u(\infty,y) \text{ και η συνάρτηση } y \longmapsto u(\infty,y) \text{ είναι μετρήσιμη} \end{array} \right\}$$

Τότε, ορίζεται το συναρτησοειδές $f: E \longrightarrow \mathbb{R}$ με

$$f(u) = \int_X u(x,\infty)\, d\mu(x) + \int_Y u(\infty,y)\, d\nu(y).$$

Το f είναι συνεχές γραμμικό συναρτησοειδές γιατί $|f(u)| \leq 2\|u\|_\infty$. Επομένως, υπάρχει επέκταση l του f σε συνεχές γραμμικό συναρτησοειδές στο $C_b(X \times Y)$.

Το l είναι συγκεντρωμένο στο άπειρο, αφού αν $u \in C_0(X \times Y)$, τότε $u(x,\infty) = u(\infty,y) = 0$ για κάθε $x \in X,\ y \in Y$.

Επιπλέον, έστω $(\varphi,\psi) \in C_0(X) \times C_0(Y)$. Τότε, αν $u(x,y) = \varphi(x) + \psi(y)$, για κάθε $x \in X$ ισχύει $u(x,\infty) = \varphi(x)$ και για κάθε $y \in Y$ ισχύει $u(\infty,y) = \psi(y)$. Άρα, $u \in E$ και

$$l(u) = f(u) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu.$$

■

Λήμμα 1.20 (Ανάλυση μη αρνητικών στοιχείων του $(C_b)^*$)

Έστω X, Y τοπικά συμπαγείς, σ −συμπαγείς Πολωνικοί χώροι. Έστω l ένα μη αρνητικό γραμμικό συναρτησιακό ορισμένο πάνω στον $C_b(X \times Y)$. Τότε, αυτό μπορεί να γραφεί στη μορφή $l = \pi + R$, όπου π είναι ένα μη αρνητικό μέτρο και R ένα μη αρνητικό συνεχές γραμμικό συναρτησοειδές συγκεντρωμένο στο άπειρο.

Θυμίζουμε ότι ένας χώρος X λέγεται σ −συμπαγής, όταν μπορεί να γραφεί στη μορφή

$$X = \bigcup_{n=1}^{\infty} K_n,$$

όπου $(K_n)_{n=1}^{\infty}$ είναι μια ακολουθία συμπαγών υποσυνόλων του X.

<u>Απόδειξη:</u>

Σε συνέχεια όσων αναφέραμε πριν το παράδειγμα, γράφουμε $l = \pi + R$, όπου $\pi \in M(X \times Y)$ και R είναι ένα συνεχές γραμμικό συναρτησοειδές συγκεντρωμένο στο άπειρο. Εφόσον το l είναι μη αρνητικό, ο περιορισμός αυτού πάνω στον $C_0(X \times Y)$, δηλαδή το μέτρο π, θα είναι επίσης μη αρνητικό. Συνεπώς, αυτό που μένει να αποδείξουμε είναι ότι το R είναι μη αρνητικό. Αρκεί να δείξουμε ότι $l \geq \pi$ στον $C_b(X \times Y)$, δηλαδή ότι για κάθε μη αρνητική συνάρτηση $u \in C_b(X \times Y)$, ισχύει

$$\langle l, u \rangle \geq \int_{X \times Y} u\, d\pi.$$

Έστω, λοιπόν, $u(x,y)$ μια οποιαδήποτε συνεχής, φραγμένη και μη αρνητική συνάρτηση.

Εφόσον έχουμε υποθέσει ότι οι χώροι X και Y είναι σ −συμπαγείς, το ίδιο θα συμβαίνει και με τον $X \times Y$, άρα μπορούμε να γράψουμε

$$X \times Y = \bigcup_{n=1}^{\infty} C_n,$$

όπου $(C_n)_{n=1}^{\infty}$ είναι μια ακολουθία συμπαγών υποσυνόλων του $X \times Y$. Εφόσον έχουμε, ακόμα, υποθέσει ότι οι χώροι X, Y είναι τοπικά συμπαγείς, το ίδιο θα συμβαίνει και με τον $X \times Y$. Θα εφαρμόσουμε την παρακάτω διαδικασία ώστε να αποδείξουμε ότι

$$X \times Y = \bigcup_{n=1}^{\infty} K_n,$$

όπου τα K_n είναι συμπαγή και για κάθε $n \in \mathbb{N}$ ισχύει $K_n \subseteq Int(K_{n+1})$.

Θέτουμε $K_1 = C_1$. Λόγω της τοπικής συμπάγειας, για κάθε $x \in K_1$ υπάρχει ανοιχτή περιοχή U_x του x έτσι ώστε το $\overline{U_x}$ να είναι συμπαγές. Λόγω της συμπάγειας, υπάρχουν πεπερασμένα το πλήθος σημεία $x_1, x_2, \dots, x_\rho \in K_1$ τέτοια ώστε

$$K_1 \subseteq \bigcup_{i=1}^{\rho} U_{x_i} \subseteq \bigcup_{i=1}^{\rho} \overline{U_{x_i}}.$$

Ονομάζουμε

$$A_2 = \bigcup_{i=1}^{\rho} U_{x_i}$$

το οποίο είναι ένα ανοιχτό σύνολο και

$$B_2 = \bigcup_{i=1}^{\rho} \overline{U_{x_i}}$$

το οποίο είναι ένα συμπαγές σύνολο.

Στη συνέχεια, θέτουμε $K_2 = B_2 \cup C_2$, το οποίο είναι συμπαγές, $K_2 \supseteq C_2$ και παρατηρούμε ότι

$$K_1 \subseteq A_2 \subseteq Int(B_2) \subseteq Int(B_2 \cup C_2) = Int(K_2).$$

Συνεχίζοντας, κατασκευάζουμε αύξουσα ακολουθία συμπαγών συνόλων $(K_n)_{n=1}^{\infty}$ που έχει τις ζητούμενες ιδιότητες.

Επομένως, για κάθε $n \in \mathbb{N}$, ικανοποιούνται οι υποθέσεις του λήμματος του Urysohn για τα σύνολα K_n, $Int(K_{n+1})$. Έτσι, για κάθε $n \in \mathbb{N}$, υπάρχει συνάρτηση $\chi_n \in C_0(X \times Y)$, με $0 \leq \chi_n \leq 1$ στον $X \times Y$, για την οποία ισχύει $\chi_n = 1$ στο K_n και $\chi_n = 0$ στο K_{n+1}^c. Ειδικότερα, αφού η $(K_n)_{n=1}^{\infty}$ είναι μια αύξουσα ακολουθία συμπαγών υποσυνόλων του $X \times Y$, η $\{\chi_n\}_{n=1}^{\infty}$ είναι και αυτή μια αύξουσα ακολουθία συνεχών συναρτήσεων με συμπαγή φορέα, για την οποία μάλιστα ισχύει ότι, για κάθε $(x, y) \in X \times Y$, $\chi_n(x, y) \to 1$ καθώς $n \to \infty$.

Άρα,

$$\int_{X\times Y} u\, d\pi = \lim_{n\to\infty}\int_{X\times Y} u\chi_n\, d\pi = \lim_{n\to\infty}\langle l, u\chi_n\rangle$$

όπου η πρώτη ισότητα προκύπτει από το Θεώρημα Μονότονης Σύγκλισης. Στη συνέχεια, αφού $u - u\chi_n \geq 0$, ισχύει $\langle l, u - u\chi_n\rangle \geq 0$, οπότε λόγω γραμμικότητας, $\langle l, u\rangle \geq \langle l, u\chi_n\rangle$. Έτσι, λοιπόν, προκύπτει

$$\int_{X\times Y} u\, d\pi = \lim_{n\to\infty}\langle l, u\chi_n\rangle \leq \langle l, u\rangle$$

όπως επιθυμούσαμε.

■

Λήμμα 1.21 (Περιθώρια συνθήκη στον $(C_b)^*$)

Έστω X, Y δύο τοπικά συμπαγείς, σ −συμπαγείς Πολωνικοί χώροι, και έστω $\mu \in P(X)$, $\nu \in P(Y)$ δύο Borel μέτρα πιθανότητας. Θεωρούμε l ένα μη αρνητικό γραμμικό συναρτησοειδές στον $C_b(X\times Y)$ τέτοιο ώστε, για κάθε ζεύγος συναρτήσεων $(\varphi, \psi) \in C_b(X)\times C_b(Y)$ να ισχύει

$$\langle l, \varphi + \psi\rangle = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu. \tag{1.21}$$

Τότε, το l είναι ένα μη αρνητικό μέτρο, και μάλιστα $l \in \Pi(\mu, \nu)$.

<u>Απόδειξη:</u>

Από τη σχέση (1.21), για $\varphi \equiv 1, \psi \equiv 0$, προκύπτει $\langle l, 1\rangle = 1$. Την ισότητα αυτή θα χρειαστεί να την επικαλεστούμε αργότερα στην παρούσα απόδειξη.

Όπως στην απόδειξη του προηγούμενου λήμματος, μπορούμε να γράψουμε

$$X = \bigcup_{n=1}^{\infty} K_n \quad \text{και} \quad Y = \bigcup_{n=1}^{\infty} L_n$$

όπου $(K_n)_{n=1}^{\infty}$ και $(L_n)_{n=1}^{\infty}$ είναι δύο αύξουσες ακολουθίες συμπαγών υποσυνόλων των X και Y αντίστοιχα, τέτοιες ώστε $K_n \subseteq Int(K_{n+1})$ και $L_n \subseteq Int(L_{n+1})$.

Από το λήμμα του Urysohn, βρίσκουμε μια αύξουσα ακολουθία συναρτήσεων $(\varphi_n)_{n=1}^{\infty} \in C_0(X)$ τέτοια ώστε για κάθε $n \in \mathbb{N}$, $0 \leq \varphi_n \leq 1$ στον X, $\varphi_n = 1$ στο K_n και $\varphi_n = 0$ στο K_{n+1}^c. Ομοίως, βρίσκουμε μια αύξουσα ακολουθία συναρτήσεων $(\psi_n)_{n=1}^{\infty} \in C_0(Y)$ τέτοια ώστε για κάθε $n \in \mathbb{N}$, $0 \leq \psi_n \leq 1$ στον Y, $\psi_n = 1$ στο L_n και $\psi_n = 0$ στο L_{n+1}^c.

Στη συνέχεια, με φ_n και ψ_n συμβολίζουμε αντίστοιχα τις συναρτήσεις $(x,y) \longmapsto \varphi_n(x)$ και $(x,y) \longmapsto \psi_n(y)$.

Ειδικότερα, για κάθε $n \in \mathbb{N}$, έχουμε $(1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n \leq 1$, πάνω σε όλο τον $X \times Y$ και μάλιστα $(1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n \longrightarrow 0$, κατά σημείο στον $X \times Y$ καθώς το $n \to +\infty$.

Πράγματι, έστω $(x,y) \in X \times Y$. Διακρίνουμε τρεις περιπτώσεις:

Αν $(x,y) \in K_{n+1} \times L_{n+1}$, τότε $\varphi_{n+1}(x) = 1$ και $\psi_{n+1}(y) = 1$, οπότε

$$\big(1-\varphi_{n+1}(x)\big)\psi_n(y) + \big(1-\psi_{n+1}(y)\big)\varphi_n(x) = 0.$$

Αν $x \notin K_{n+1}$, τότε $\varphi_n(x) = 0$, οπότε

$$\big(1-\varphi_{n+1}(x)\big)\psi_n(y) + \big(1-\psi_{n+1}(y)\big)\varphi_n(x) = \big(1-\varphi_{n+1}(x)\big)\psi_n(y) \leq 1.$$

Αν $y \notin L_{n+1}$, τότε $\psi_n(y) = 0$, οπότε

$$\big(1-\varphi_{n+1}(x)\big)\psi_n(y) + \big(1-\psi_{n+1}(y)\big)\varphi_n(x) = \big(1-\psi_{n+1}(y)\big)\varphi_n(x) \leq 1.$$

Ακόμα, για να αποδείξουμε την κατά σημείο σύγκλιση, έστω $(x,y) \in X \times Y$. Τότε, $x \in K_{n_1}$ για κάποιο $n_1 \in \mathbb{N}$ και $y \in L_{n_2}$ για κάποιο $n_2 \in \mathbb{N}$. Έστω $n_0 = \max\{n_1, n_2\}$. Τότε, $(x,y) \in K_{n_0} \times L_{n_0}$ και επομένως

$$\varphi_{n_0}(x) = \varphi_{n_0+1}(x) = 1$$

$$\psi_{n_0}(y) = \psi_{n_0+1}(y) = 1$$

και εφόσον οι ακολουθίες $(\varphi_n)_{n=1}^{\infty}$ και $(\psi_n)_{n=1}^{\infty}$ είναι αύξουσες, η ίδια ισότητα θα ισχύει και για κάθε $n \geq n_0$.

Συνεπώς, για κάθε $n \geq n_0$, έχουμε

$$\big(1-\varphi_{n+1}(x)\big)\psi_n(y) + \big(1-\psi_{n+1}(y)\big)\varphi_n(x) = 0.$$

Στη συνέχεια, έστω

$$A_n = \langle l, (1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n \rangle.$$

Σύμφωνα με το λήμμα 1.20, $l = \pi + R$, όπου το π είναι ένα μη αρνητικό μέτρο και το R ένα μη αρνητικό συνεχές γραμμικό συναρτησοειδές συγκεντρωμένο στο άπειρο. Επομένως, για την ποσότητα A_n θα ισχύει:

$$\begin{aligned} A_n &= \int_{X\times Y} [(1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n]\, d\pi(x,y) + \langle R, (1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n \rangle \\ &\leq \int_{X\times Y} [(1-\varphi_{n+1})\psi_n + (1-\psi_{n+1})\varphi_n]\, d\pi(x,y) + \langle R, 1 \rangle \end{aligned}$$

όπου η ανισότητα ισχύει αφού το R είναι μη αρνητικό.

Λόγω του Θεωρήματος Κυριαρχημένης Σύγκλισης, η τελευταία ποσότητα συγκλίνει καθώς το $n \to +\infty$ στο $\langle R, 1\rangle$ και

$$\langle R, 1\rangle = \langle l, 1\rangle - \langle \pi, 1\rangle = 1 - \int_{X\times Y} d\pi.$$

Επομένως,

$$\liminf_{n\to+\infty} A_n \leq 1 - \pi[X \times Y]. \qquad (1.22)$$

Από την άλλη, λόγω γραμμικότητας του l, η ποσότητα A_n είναι ίση με

$$A_n = \langle l, \varphi_n + \psi_n\rangle - \langle l, \varphi_{n+1}\psi_n + \varphi_n\psi_{n+1}\rangle.$$

Παρατηρώντας ότι $\varphi_{n+1}\psi_n + \varphi_n\psi_{n+1} \in C_0(X \times Y)$, από τη σχέση (1.21) και το γεγονός ότι το R είναι συγκεντρωμένο στο άπειρο, προκύπτει ότι

$$A_n = \int_X \varphi_n \, d\mu + \int_Y \psi_n \, d\nu - \int_{X\times Y} (\varphi_{n+1}\psi_n + \varphi_n\psi_{n+1}) \, d\pi.$$

Λόγω του Θεωρήματος Μονότονης Σύγκλισης,

$$\lim_{n\to\infty} A_n = 1 + 1 - 2\int_{X\times Y} d\pi = 2 - 2\pi[X \times Y]. \qquad (1.23)$$

Από τις σχέσεις (1.22) και (1.23) έχουμε: $2 - 2\pi[X \times Y] \leq 1 - \pi[X \times Y]$, άρα $\pi[X \times Y] \geq 1$.

Ειδικότερα, $\langle R, 1\rangle = \langle l, 1\rangle - \langle \pi, 1\rangle \leq 0$ και εφόσον το R είναι μη αρνητικό, έπεται ότι $\langle R, 1\rangle = 0$. Τότε, όπως θα δείξουμε, $R = 0$, οπότε και $l = \pi$.

Για να δούμε ότι, όντως $R = 0$, έστω $u \in C_b(X \times Y)$. Έχουμε:

$$\langle R, u\rangle \leq \langle R, \|u\|_\infty\rangle = \|u\|_\infty \langle R, 1\rangle = 0$$

και

$$\langle R, u\rangle \geq \langle R, -\|u\|_\infty\rangle = -\|u\|_\infty \langle R, 1\rangle = 0,$$

άρα $\langle R, u\rangle = 0$.

Επομένως, το l είναι ένα μη αρνητικό μέτρο, για κάθε $(\varphi, \psi) \in C_b(X) \times C_b(Y)$ ισχύει

$$\int_{X\times Y} [\varphi(x) + \psi(y)] \, d\pi(x, y) = \langle l, \varphi + \psi\rangle = \int_X \varphi \, d\mu + \int_Y \psi \, d\nu,$$

οι χώροι X, Y είναι Πολωνικοί και τα μ, ν είναι Borel μέτρα πιθανότητας.

Άρα, λόγω του ορισμού του $\Pi(\mu,\nu)$ μέσω της σχέσης (1.2) καθώς και του Ισχυρισμού 1, το l είναι στοιχείο του $\Pi(\mu,\nu)$.

■

Απόδειξη της Πρότασης 1.17:

Ειδική περίπτωση: X, Y σ −συμπαγείς χώροι.

Στο $C_b(X\times Y)$ ορίζουμε, όπως και στο 1ο σκέλος της απόδειξης του Θεωρήματος 1.3, τις συναρτήσεις Θ και Ξ. Ειδικότερα,

$$\Theta : u \in C_b(X\times Y) \longmapsto \begin{cases} 0, & \text{αν } u(x,y) \geq -c(x,y) \\ +\infty, & \text{αλλιώς} \end{cases}$$

$$\Xi : u \in C_b(X\times Y) \longmapsto \begin{cases} \int_X \varphi\, d\mu + \int_Y \psi\, d\nu, & \text{αν } u(x,y) = \varphi(x) + \psi(y) \\ +\infty, & \text{αλλιώς} \end{cases}$$

Στη συνέχεια, θα θέλαμε να εφαρμόσουμε το Θεώρημα 1.7 επιλέγοντας $z_0 = 0$. Παρατηρούμε, όμως, ότι αν $\inf c(x,y) = 0$ (το οποίο είναι κάτι που μπορεί να συμβαίνει, εάν π.χ. έχουμε επιλέξει τους χώρους X, Y να ταυτίζονται και η c είναι μια απόσταση, οπότε $c(x,x) = 0$), τότε η Θ δεν είναι συνεχής στο 0. Για να ξεπεράσουμε αυτό το πρόβλημα, θεωρούμε $\varepsilon > 0$ και παίρνουμε τη συνάρτηση $c_\varepsilon = c + \varepsilon$ στη θέση της c. Ορίζουμε, λοιπόν, τη συνάρτηση Θ ως εξής:

$$\Theta : u \in C_b(X\times Y) \longmapsto \begin{cases} 0, & \text{αν } u(x,y) \geq -c(x,y) - \varepsilon \\ +\infty, & \text{αλλιώς} \end{cases}$$

Τότε, αν $u \in C_b(X\times Y)$ με $\|u\|_\infty \leq \varepsilon$, ισχύει $\Theta(u) = 0$, δηλαδή η Θ είναι συνεχής στο 0. Εφόσον, επίσης, $\Theta(0) = \Xi(0) = 0 < +\infty$, οι υποθέσεις του Θεωρήματος 1.7 ικανοποιούνται.

Παρατηρούμε, τώρα, ότι για κάθε $\pi \in \Pi(\mu,\nu)$,

$$I_\varepsilon[\pi] = \int_{X\times Y} c_\varepsilon(x,y)\, d\pi(x,y) = \int_{X\times Y} (c(x,y) + \varepsilon)\, d\pi(x,y) = I[\pi] + \varepsilon$$

οπότε

$$\inf_{\pi\in\Pi(\mu,\nu)} I_\varepsilon[\pi] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi] + \varepsilon.$$

Ακόμα,

$$\begin{aligned} \sup_{(\varphi,\psi)\in\Phi_{c_\varepsilon}} J(\varphi,\psi) &= \sup\left\{\int_X \varphi\, d\mu + \int_Y \psi\, d\nu :\ \varphi(x) + \psi(y) \leq c(x,y) + \varepsilon\right\} \\ &= \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) + \varepsilon. \end{aligned}$$

Αυτό σημαίνει ότι η ισότητα

$$\inf_{\pi\in\Pi(\mu,\nu)} I_\varepsilon[\pi] = \sup_{(\varphi,\psi)\in\Phi_{c_\varepsilon}} J(\varphi,\psi)$$

για τη συνάρτηση κόστους c_ε συνεπάγεται την ισότητα

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi)$$

για την αρχική συνάρτηση κόστους c.

Επομένως, αρκεί να αποδείξουμε το ζητούμενο με την επιπλέον υπόθεση ότι $c \geq \varepsilon$.

Έτσι, λοιπόν, από το Θεώρημα 1.7 έχουμε

$$\inf_{u\in C_b(X\times Y)}[\Theta(u)+\Xi(u)] = \sup_{l\in C_b(X\times Y)^*}[-\Theta^*(-l)-\Xi^*(l)],$$

την οποία ισότητα μπορούμε να γράψουμε και ως εξής:

$$\sup_{u\in C_b(X\times Y)}[-\Theta(u)-\Xi(u)] = \inf_{l\in C_b(X\times Y)^*}[\Theta^*(-l)+\Xi^*(l)]. \tag{1.24}$$

Όπως και στο 1ο σκέλος της απόδειξης του Θεωρήματος 1.3, το αριστερό μέλος της σχέσης (1.24) είναι:

$$\sup_{u\in C_b(X\times Y)}[-\Theta(u)-\Xi(u)] = \sup_{\Phi_c\cap C_b} J(\varphi,\psi).$$

Για να υπολογίσουμε την τιμή του δεξιού μέλους της σχέσης (1.24), ας δούμε πρώτα τον μετασχηματισμό Legendre-Fenchel της Θ. Για κάθε $u \in C_b(X\times Y)$ έχουμε

$$\Theta^*(-l) = \sup_{u\in C_b(X\times Y)}[\langle -l,u\rangle - \Theta(u)] = \sup_{u\geq -c}\langle -l,u\rangle = \sup_{u\leq c}\langle l,u\rangle.$$

Αν υποθέσουμε ότι το l δεν είναι ένα μη αρνητικό γραμμικό συναρτησοειδές στο $C_b(X\times Y)$, τότε, θα υπάρχει συνάρτηση $v\in C_b(X\times Y)$ τέτοια ώστε $v\leq 0$ αλλά $\langle l,v\rangle > 0$. Τότε, για κάθε $\lambda>0$, θα ισχύει $\lambda v\leq c$, συνεπώς

$$\sup_{u\leq c}\langle l,u\rangle \geq \sup_{\lambda>0}\langle l,\lambda v\rangle = \sup_{\lambda>0}\lambda\langle l,v\rangle.$$

Αφήνοντας το $\lambda\to+\infty$, βλέπουμε ότι $\Theta^*(-l)=\sup_{u\leq c}\langle l,u\rangle = +\infty$, άρα δεν συνεισφέρει τίποτα στο infimum του δεξιού μέλους της σχέσης (1.24). Θα λάβουμε, επομένως, υπόψη μας, μόνο τα μη αρνητικά συνεχή γραμμικά συναρτησοειδή στο $C_b(X\times Y)$, οπότε έχουμε

$$\Theta^*(-l) = \sup_{u\leq c}\langle l,u\rangle = \langle l,c\rangle.$$

Για την Ξ^*, έχουμε

$$\Xi^*(l) = \sup_{u \in C_b(X\times Y)} [\langle l, u\rangle - \Xi(u)] = \sup\left[\langle l, \varphi + \psi\rangle - \int_X \varphi\, d\mu - \int_Y \psi\, d\nu : \quad (\varphi,\psi) \in C_b(X) \times C_b(Y)\right].$$

Αν για κάθε $(\varphi,\psi) \in C_b(X) \times C_b(Y)$, ισχύει

$$\langle l, \varphi + \psi\rangle = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu,$$

τότε $\Xi^*(l) = 0$.

Αν υπάρχει ζεύγος συναρτήσεων $(\varphi_1,\psi_1) \in C_b(X) \times C_b(Y)$ τέτοιο ώστε

$$\langle l, \varphi_1 + \psi_1\rangle - \int_X \varphi_1\, d\mu - \int_Y \psi_1\, d\nu = K \neq 0,$$

τότε:
αν $K > 0$, επιλέγοντας $\varphi = \lambda\varphi_1,\ \psi = \lambda\psi_1,\ \ \lambda > 0$ και αφήνοντας το $\lambda \to +\infty$ θα έχουμε

$$\Xi^*(l) \geq \lambda K \to +\infty$$

ενώ, αν $K < 0$, επιλέγοντας $\varphi = \lambda\varphi_1,\ \psi = \lambda\psi_1,\ \ \lambda < 0$ και αφήνοντας το $\lambda \to -\infty$ θα έχουμε και πάλι:

$$\Xi^*(l) \geq \lambda K \to +\infty.$$

Τελικά,

$$\Xi^*(l) = \begin{cases} 0, & \text{αν} \begin{bmatrix} \forall(\varphi,\psi) \in C_b(X)\times C_b(Y), \\ \langle l, \varphi+\psi\rangle = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu \end{bmatrix} \\ +\infty, & \text{αλλιώς} \end{cases}$$

Στην περίπτωση που το l είναι μη αρνητικό γραμμικό συναρτησοειδές, τότε σύμφωνα με το λήμμα 1.21, $l \in \Pi(\mu,\nu)$, και έχουμε:

$$\Xi^*(l) = \begin{cases} 0, & \text{αν } l \in \Pi(\mu,\nu) \\ +\infty, & \text{αλλιώς} \end{cases}$$

Συνοψίζοντας, το δεξί μέλος της σχέσης (1.24) γίνεται:

$$\inf_{l \in C_b(X\times Y)^*} [\Theta^*(-l) + \Xi^*(l)] = \inf_{\pi \in \Pi(\mu,\nu)} \left[\int_{X\times Y} c\, d\pi + 0\right],$$

δηλαδή

$$\inf_{l \in C_b(X\times Y)^*} [\Theta^*(-l) + \Xi^*(l)] = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

Επιστρέφοντας, λοιπόν, στην ισότητα

$$\sup_{u \in C_b(X\times Y)} [-\Theta(u) - \Xi(u)] = \inf_{l \in C_b(X\times Y)^*} [\Theta^*(-l) + \Xi^*(l)],$$

και αντικαθιστώντας τα δύο μέλη, έχουμε:

$$\sup_{\Phi_c \cap C_b} J(\varphi, \psi) = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

όπως και επιθυμούσαμε.

<u>Γενική περίπτωση:</u>

Από το λήμμα του Ulam (Πρόταση 26, Εισαγωγικό κεφάλαιο), υπάρχουν σ −συμπαγή σύνολα $S \subseteq X$, $T \subseteq Y$ έτσι ώστε $\mu[S] = 1$, $\nu[T] = 1$. Τα S, T μπορούν να επιλεγούν ώστε να είναι ανοικτά σύνολα. Πράγματι, έστω

$$S = \bigcup_{n=1}^{\infty} K_n$$

όπου τα K_n είναι συμπαγή. Επειδή ο X είναι τοπικά συμπαγής χώρος, κάθε συμπαγές σύνολο περιέχεται σε ένα ανοικτό, σχετικά συμπαγές σύνολο. Άρα, υπάρχουν ανοικτά σύνολα U_n, $n = 1, 2, \dots$ και συμπαγή σύνολα F_n, $n = 1, 2, \dots$ έτσι ώστε

$$K_1 \subseteq U_1 \subseteq F_1 \cup K_2 \subseteq U_2 \subseteq F_2 \cup K_3 \subseteq \cdots \subseteq U_n \subseteq F_n \cup K_{n+1} \subseteq \cdots$$

Τότε, το σύνολο

$$S = \bigcup_{n=1}^{\infty} U_n = \bigcup_{n=1}^{\infty} (F_n \cup K_{n+1})$$

είναι σ −συμπαγές και ανοικτό, και ως ανοικτό υποσύνολο τοπικά συμπαγούς Πολωνικού χώρου είναι, επίσης, τοπικά συμπαγής Πολωνικός χώρος.

Θέτουμε

$$I_{X,Y} = \sup\{I[\pi] : \ \pi \in P(X \times Y), \pi \in \Pi(\mu,\nu)\}$$

$$I_{S,T} = \sup\{I[\pi] : \ \pi \in P(S \times T), \pi \in \Pi(\mu,\nu)\}$$

$$J_{X,Y} = \sup\left\{\int_X \varphi \, d\mu + \int_Y \psi \, d\nu : \ (\varphi,\psi) \in \big(C_b(X) \times C_b(Y)\big) \cap \Phi_c\right\}$$

$$J_{S,T} = \sup\left\{\int_T \varphi \, d\mu + \int_S \psi \, d\nu : \ (\varphi,\psi) \in \big(C_b(S) \times C_b(T)\big) \cap \Phi_c\right\}$$

Τότε, $I_{X,Y} = I_{S,T} = J_{S,T} = J_{X,Y}$.

Η πρώτη ισότητα είναι προφανής, η δεύτερη έπεται από την ειδική περίπτωση και η τρίτη προκύπτει από τον τελευταίο ισχυρισμό στη διατύπωση του Θεωρήματος 1.3, αφού κάθε $\varphi \in C_b(S)$ επεκτείνεται σε μια συνάρτηση στο $L^1(d\mu)$ και κάθε $\psi \in C_b(T)$ σε μια συνάρτηση στο $L^1(d\nu)$.

■

1.4. Συναρτήσεις κόστους με τιμές στο σύνολο $\{0,1\}$.

Στην τελευταία ενότητα του 1ου κεφαλαίου, θα δούμε τη μορφή που παίρνει το Θεώρημα δυϊκότητας του Kantorovich στην περίπτωση που η συνάρτηση κόστους παίρνει τιμές στο σύνολο $\{0,1\}$, δηλαδή όταν είναι της μορφής $c(x,y) = 1_C(x,y)$, όπου C είναι ένα υποσύνολο του χώρου $X \times Y$. Τότε, το συναρτησοειδές $I[\pi]$ παίρνει τη μορφή

$$I[\pi] = \int_{X\times Y} 1_C(x,y)\, d\pi(x,y) = \pi[C].$$

Θεώρημα 1.22 (Δυϊκότητα Kantorovich για τη συνάρτηση κόστους $\mathbf{1}_C(x,y)$)

Έστω X, Y δύο Πολωνικοί χώροι, $\mu \in P(X)$, $\nu \in P(Y)$ και έστω C ένα μη κενό ανοικτό υποσύνολο του $X \times Y$. Ορίζουμε $A^C \coloneqq \{y \in Y :\ \exists x \in A\ \text{τ.ω}\ (x,y) \notin C\}$. Τότε,

$$\inf_{\pi\in\Pi(\mu,\nu)} \pi[C] = \sup_{\substack{A\subseteq X \\ A\ \text{κλειστό}}} \{\mu[A] - \nu[A^C]\}.$$

Συνέπεια του Θεωρήματος 1.22 είναι το παρακάτω πόρισμα, το οποίο αναφέρεται και ως θεώρημα του Strassen.

Πόρισμα 1.23 (Θεώρημα Strassen)

Έστω (X,d) ένας Πολωνικός χώρος, $\mu,\nu \in P(X)$ και $\varepsilon \geq 0$. Ορίζουμε $A^\varepsilon \coloneqq \{y \in X : d(y,A) \leq \varepsilon\}$. Τότε:

$$\inf_{\pi\in\Pi(\mu,\nu)} \pi[\{d(x,y) > \varepsilon\}] = \sup_{\substack{A\subseteq X \\ A\ \text{κλειστό}}} \{\mu[A] - \nu[A^\varepsilon]\}.$$

Στο παραπάνω πόρισμα χρησιμοποιούμε το συνήθη συμβολισμό για την απόσταση σημείου από σύνολο $d(y,A) = \inf_{x\in A} d(x,y)$.

<u>Απόδειξη του πορίσματος 1.23:</u>

Έστω $\varepsilon \geq 0$. Στο Θεώρημα 1.22, επιλέγουμε $X = Y$ και $C = \{(x,y) \in X \times X :\ \ d(x,y) > \varepsilon\}$ το οποίο είναι ένα ανοικτό υποσύνολο του $X \times X$. Θα δείξουμε ότι αν $A \subseteq X$ συμπαγές, τότε $A^C = A^\varepsilon$.

Για την απόδειξη του εγκλεισμού $A^C \subseteq A^\varepsilon$, έστω $y \in A^C$. Υπάρχει $x \in A$ τέτοιο ώστε $(x,y) \notin C$. Αυτό σημαίνει ότι $d(x,y) \leq \varepsilon$, άρα $d(y,A) \leq \varepsilon$, δηλαδή $y \in A^\varepsilon$.

Για τον αντίστροφο εγκλεισμό $A^{\varepsilon} \subseteq A^{C}$, έστω $y \in A^{\varepsilon}$. Τότε, $d(y, A) \leq \varepsilon$ και επειδή το A είναι συμπαγές, υπάρχει $x \in A$ τέτοιο ώστε $d(x, y) \leq \varepsilon$. Αυτό σημαίνει ότι $(x, y) \notin C$ και επομένως $y \in A^{C}$.

Έχοντας διασφαλίσει την ισότητα $A^{C} = A^{\varepsilon}$ για τα συμπαγή υποσύνολα του X, αρκεί να αποδείξουμε ότι τα supremum που εμφανίζονται στο Θεώρημα 1.22 και στο πόρισμα δεν αλλάζουν αν περιοριστούμε σε συμπαγή υποσύνολα A του X.

Πράγματι, έστω $A \subseteq X$ κλειστό και $\delta > 0$. Το μ, ως μέτρο ορισμένο πάνω σε Πολωνικό χώρο, είναι κανονικό, δηλαδή υπάρχει συμπαγές $K \subseteq A$ τέτοιο ώστε $\mu[A] \leq \mu[K] + \delta$. Επίσης, εύκολα βλέπουμε ότι $K^{C} \subseteq A^{C}$ και $K^{\varepsilon} \subseteq A^{\varepsilon}$, οπότε $\nu[K^{C}] \leq \nu[A^{C}]$ και $\nu[K^{\varepsilon}] \leq \nu[A^{\varepsilon}]$ αντίστοιχα. Συνεπώς,

$$\mu[A] - \nu[A^{C}] \leq \mu[K] - \nu[K^{C}] + \delta$$

$$\mu[A] - \nu[A^{\varepsilon}] \leq \mu[K] - \nu[K^{\varepsilon}] + \delta,$$

δηλαδή

$$\sup_{\substack{A \subseteq X \\ A \text{ κλειστό}}} \{\mu[A] - \nu[A^{C}]\} \leq \sup_{\substack{K \subseteq X \\ K \text{ συμπαγές}}} \{\mu[K] - \nu[K^{C}]\} + \delta$$

και

$$\sup_{\substack{A \subseteq X \\ A \text{ κλειστό}}} \{\mu[A] - \nu[A^{\varepsilon}]\} \leq \sup_{\substack{K \subseteq X \\ K \text{ συμπαγές}}} \{\mu[K] - \nu[K^{\varepsilon}]\} + \delta.$$

Εφόσον το δ ήταν ένας αυθαίρετος θετικός αριθμός, προκύπτει ότι οι τιμές των supremum δεν αλλάζουν αν περιοριστούμε στα συμπαγή υποσύνολα του X.

Τέλος, με την εφαρμογή του Θεωρήματος 1.22, προκύπτει το ζητούμενο.

■

Απόδειξη του Θεωρήματος 1.22:

Εφόσον για το σύνολο C έχουμε υποθέσει ότι είναι ανοικτό στον $X \times Y$, η συνάρτηση κόστους $c(x, y) = 1_{C}(x, y)$ θα είναι μια κάτω ημι-συνεχής συνάρτηση στον $X \times Y$. Πράγματι, αρκεί για κάθε $(x, y) \in X \times Y$ να ισχύει $1_{C}(x, y) \leq \liminf\limits_{(z,w) \to (x,y)} 1_{C}(z, w)$.

Αν $(x, y) \in C$, τότε $1_{C}(x, y) = 1$, ενώ αφού το C είναι ανοικτό, υπάρχει περιοχή του σημείου (x, y), $N_{\varepsilon}(x, y) \subseteq C$, δηλαδή $1_{C}(z, w) = 1$, για κάθε $(z, w) \in N_{\varepsilon}(x, y)$. Επομένως,

$$\liminf_{(z,w) \to (x,y)} 1_{C}(z, w) = 1 = 1_{C}(x, y).$$

Αν $(x, y) \notin C$, τότε $1_{C}(x, y) = 0 \leq 1_{C}(z, w)$, για κάθε $(z, w) \in X \times Y$. Επομένως,

$$1_C(x,y) \le \liminf_{(z,w)\to(x,y)} 1_C(z,w).$$

Στη συνέχεια, όπως αποδείξαμε στο εισαγωγικό κεφάλαιο (Πρόταση 36), η 1_C, ως μια μη αρνητική κάτω ημισυνεχής συνάρτηση στον μετρικό χώρο $X \times Y$, μπορεί να γραφεί στη μορφή $1_C = \sup_{k\in\mathbb{N}} c_k$, όπου $(c_k)_{k=1}^{\infty}$ είναι μια αύξουσα ακολουθία μη αρνητικών και ομοιόμορφα συνεχών συναρτήσεων οι οποίες για κάθε $k \in \mathbb{N}$, ικανοποιούν την ανισότητα $0 \le c_k \le 1_C$.

Θα ισχύει τότε:

$$\begin{aligned}\inf_{\pi\in\Pi(\mu,\nu)} \pi[C] &= \inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} 1_C \, d\pi = \lim_{k\to\infty} \inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c_k \, d\pi = \lim_{k\to\infty} \inf_{\pi\in\Pi(\mu,\nu)} I_k[\pi] \\ &= \lim_{k\to\infty} \sup_{(\varphi,\psi)\in\Phi_{c_k}} J(\varphi,\psi) = \lim_{k\to\infty} \sup\left\{\int_X \varphi \, d\mu + \int_Y \psi \, d\nu : \ (\varphi,\psi) \in \Phi_{c_k}\right\}.\end{aligned}$$

Σημειώνουμε ότι η δεύτερη ισότητα ισχύει λόγω της σχέσης (1.12).

Από τις Παρατηρήσεις 1.9 και 1.10, εφόσον προφανώς $\|c_k\|_\infty \le \|1_C\|_\infty = 1$, βλέπουμε ότι για κάθε $k \in \mathbb{N}$, μπορούμε να περιορίσουμε το supremum σε εκείνα τα ζεύγη συναρτήσεων $(\varphi,\psi) \in \Phi_{c_k}$ για τις οποίες $0 \le \varphi \le 1$, $-1 \le \psi \le 0$, όπου μάλιστα η ψ είναι συνεχής συνάρτηση και η φ άνω ημισυνεχής που ικανοποιούν την ανισότητα $\varphi(x) + \psi(y) \le c_k(x,y) \le c(x,y) = 1_C(x,y)$, για κάθε $(x,y) \in X \times Y$. Άρα, αν ονομάσουμε $\widetilde{\Phi_c}$ το σύνολο όλων των ζευγών $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$ τέτοιες ώστε

$$(1.25) \qquad \begin{cases} \varphi(x) + \psi(y) \le c(x,y) = 1_C(x,y), \quad \text{για κάθε } (x,y) \in X \times Y \\ 0 \le \varphi \le 1, \quad -1 \le \psi \le 0 \\ \varphi \quad \text{άνω ημι} - \text{συνεχής} \end{cases}$$

συμπεραίνουμε ότι

$$(1.26) \qquad \inf_{\pi\in\Pi(\mu,\nu)} \pi[C] = \sup\left\{\int_X \varphi \, d\mu + \int_Y \psi \, d\nu : \ (\varphi,\psi) \in \widetilde{\Phi_c}\right\}.$$

Παρατηρούμε ότι το σύνολο $\widetilde{\Phi_c}$ είναι κυρτό.

Πράγματι, έστω $(\varphi_1,\psi_1), (\varphi_2,\psi_2) \in \widetilde{\Phi_c}$ και $\lambda \in [0,1]$. Θα πρέπει να ελέγξουμε ότι

$$\lambda(\varphi_1,\psi_1) + (1-\lambda)(\varphi_2,\psi_2) = (\lambda\varphi_1 + (1-\lambda)\varphi_2, \lambda\psi_1 + (1-\lambda)\psi_2) \in \widetilde{\Phi_c}.$$

Για κάθε $(x,y) \in X \times Y$,

$$\begin{aligned}&\lambda\varphi_1(x) + (1-\lambda)\varphi_2(x) + \lambda\psi_1(y) + (1-\lambda)\psi_2(y) \\ &\quad = \lambda\big(\varphi_1(x) + \psi_1(y)\big) + (1-\lambda)\big(\varphi_2(x) + \psi_2(y)\big) \le \lambda c(x,y) + (1-\lambda)c(x,y) = c(x,y).\end{aligned}$$

Για κάθε $x \in X$, $0 \le \lambda\varphi_1(x) + (1-\lambda)\varphi_2(x) \le 1$, αφού $0 \le \varphi_1(x) \le 1$ και $0 \le \varphi_2(x) \le 1$. Ομοίως, για κάθε $y \in Y$, $-1 \le \lambda\psi_1(y) + (1-\lambda)\psi_2(y) \le 0$.

Η $\lambda\varphi_1 + (1-\lambda)\varphi_2$ είναι άνω ημι-συνεχής, αφού οι φ_1 και φ_2 είναι, οπότε για κάθε $x \in X$,

$$\lambda\varphi_1(x) + (1-\lambda)\varphi_2(x) \geq \lambda \cdot \limsup_{z\to x} \varphi_1(z) + (1-\lambda) \cdot \limsup_{z\to x} \varphi_2(z)$$
$$\geq \limsup_{z\to x} \,[\lambda\varphi_1(z) + (1-\lambda)\varphi_2(z)].$$

Τελικά, $\lambda(\varphi_1,\psi_1) + (1-\lambda)(\varphi_2,\psi_2) \in \widetilde{\Phi_c}$ και το $\widetilde{\Phi_c}$ είναι κυρτό σύνολο.

Ισχυρισμός 13. Κάθε ζεύγος $(\varphi,\psi) \in \widetilde{\Phi_c}$ μπορεί να αναπαρασταθεί ως ένας κυρτός συνδυασμός ζευγών της μορφής $(1_A, -1_B) \in \widetilde{\Phi_c}$, όπου με κυρτό συνδυασμό εννοούμε ένα ολοκλήρωμα της μορφής

$$\int_0^1 (1_{A_s}, -1_{B_s})ds$$

και τα A_s, B_s είναι κλειστά υποσύνολα του X.

<u>Απόδειξη του Ισχυρισμού 13:</u>

Καταρχήν, κάθε μετρήσιμη απεικόνιση $u: X \longrightarrow [0,1]$ μπορεί να γραφεί ως

$$u = \int_0^1 1_{\{u\geq s\}}\, ds$$

και αντίστοιχα κάθε μετρήσιμη απεικόνιση $v: X \longrightarrow [-1,0]$ μπορεί να γραφεί ως

$$v = \int_0^1 -1_{\{v\leq -s\}}\, ds.$$

Πράγματι, για κάθε $x \in X$, έχουμε

$$\int_0^1 1_{\{u\geq s\}}(x)\, ds = \int_0^{u(x)} ds = u(x)$$

και

$$\int_0^1 -1_{\{v\leq -s\}}(x)\, ds = -\int_0^{-v(x)} ds = v(x).$$

Επομένως, αν $(\varphi,\psi) \in \widetilde{\Phi_c}$, (οπότε $0 \leq \varphi \leq 1$ και $-1 \leq \psi \leq 0$) μπορούμε να γράψουμε

$$(\varphi,\psi) = \int_0^1 (1_{\{\varphi\geq s\}}, -1_{\{\psi\leq -s\}})ds.$$

Αφού η φ είναι άνω ημισυνεχής, για κάθε $s \in [0,1]$ το σύνολο $\{x \in X\colon \varphi(x) \geq s\}$ είναι κλειστό στον X. Επομένως, για να ολοκληρώσουμε την απόδειξη του ισχυρισμού, αρκεί να ελέγξουμε ότι $\left(1_{\{\varphi\geq s\}}, -1_{\{\psi\leq -s\}}\right) \in \widetilde{\Phi_c}$, δηλαδή ότι για κάθε $s \in [0,1]$ και για κάθε $(x,y) \in X \times Y$, ισχύει

$$1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) \leq c(x,y) = 1_C(x,y).$$

Έστω, λοιπόν, $(x,y) \in X \times Y$. Διακρίνουμε τις περιπτώσεις:

Αν $\varphi(x) < s,\ \psi(y) \leq -s$, τότε $1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) = 0 - 1 = -1 \leq 1_C(x,y)$.

Αν $\varphi(x) < s,\ \psi(y) > -s$, τότε $1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) = 0 - 0 = 0 \leq 1_C(x,y)$.

Αν $\varphi(x) \geq s,\ \psi(y) \leq -s$, τότε $1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) = 1 - 1 = 0 \leq 1_C(x,y)$.

Στην τελευταία περίπτωση, όπου $\varphi(x) \geq s,\ \psi(y) > -s$, έχουμε

$$1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) = 1 - 0 = 1.$$

Αφού $(\varphi, \psi) \in \Phi_c$,

$$c(x,y) \geq \varphi(x) + \psi(y) > 0.$$

Τώρα, αφού η $c(x,y) = 1_C(x,y)$ παίρνει τιμές στο σύνολο $\{0,1\}$, θα ισχύει $1_C(x,y) = 1$, άρα και πάλι

$$1_{\{\varphi \geq s\}}(x) - 1_{\{\psi \leq -s\}}(y) \leq 1_C(x,y).$$

Έτσι, λοιπόν, η απόδειξη του ισχυρισμού είναι πλήρης.

■

Επιστρέφουμε, τώρα, στην απόδειξη του Θεωρήματος.

Για κάθε ζεύγος $(\varphi, \psi) \in \widetilde{\Phi_c}$, υπάρχει ζεύγος της μορφής $(1_A, -1_B) \in \widetilde{\Phi_c}$ με A κλειστό και τέτοιο ώστε $J(1_A, -1_B) \geq J(\varphi, \psi)$.

Πράγματι, έστω $(\varphi, \psi) \in \widetilde{\Phi_c}$. Λόγω του ισχυρισμού 12, μπορούμε να γράψουμε

$$\varphi(x) = \int_0^1 1_{A_s}(x)\, ds = \int_0^1 1_{\{\varphi \geq s\}}(x)\, ds$$

με τα A_s κλειστά και

$$\psi(y) = \int_0^1 -1_{B_s}(y)\, ds = \int_0^1 -1_{\{\psi \leq -s\}}(y)\, ds.$$

Τότε,

$$\begin{aligned}
J(\varphi, \psi) &= \int_X \varphi\, d\mu + \int_Y \psi\, d\nu = \int_X \left(\int_0^1 1_{A_s}(x)\, ds \right) d\mu + \int_Y \left(\int_0^1 -1_{B_s}(y)\, ds \right) d\nu \\
&= \int_0^1 \left(\int_X 1_{A_s}(x)\, d\mu \right) ds + \int_0^1 \left(\int_Y -1_{B_s}(y)\, d\nu \right) ds \\
&= \int_0^1 \left(\int_X 1_{A_s}(x)\, d\mu + \int_Y -1_{B_s}(y)\, d\nu \right) ds.
\end{aligned}$$

Από την παραπάνω ισότητα, συμπεραίνουμε ότι υπάρχει $s \in (0,1)$ έτσι ώστε

$$J(\varphi,\psi) \leq \int_X 1_{A_s}(x)\,d\mu + \int_Y -1_{B_s}(y)\,d\nu = J\big(1_{A_s}, -1_{B_s}\big).$$

Ισχυρισμός 14. Το supremum στο δεξί μέλος της σχέσης (1.26) μπορεί να περιοριστεί πάνω στα ζεύγη της μορφής $(1_A, -1_B)$, όπου A είναι κλειστό, χωρίς να υπάρχει αλλαγή στην τιμή του. Ειδικότερα, ισχύει:

$$\begin{aligned}&\sup\left\{\int_X \varphi\,d\mu + \int_Y \varphi\,d\nu \;:\; (\varphi,\psi) \in \widetilde{\Phi_c}\right\}\\ &\quad= \sup\left\{\int_X 1_A\,d\mu - \int_Y 1_B\,d\nu \;:\; (1_A,-1_B) \in \widetilde{\Phi_c}\,, A \subseteq X \text{ κλειστό}\right\}\\ &\quad= \sup\{\mu[A] - \nu[A^C] :\; A \subseteq X \text{ κλειστό}\}.\end{aligned} \tag{1.27}$$

<u>Απόδειξη του Ισχυρισμού 14:</u>

Η πρώτη ισότητα προκύπτει από τον Ισχυρισμό 13. Για τη δεύτερη ισότητα, παρατηρούμε ότι αν $(1_A,-1_B) \in \widetilde{\Phi_c}$, τότε για κάθε $(x,y) \in X \times Y$ ισχύει $1_A(x) - 1_B(y) \leq 1_C(x,y)$, δηλαδή $1_B(y) \geq 1_A(x) - 1_C(x,y)$, οπότε για κάθε $y \in Y$, $1_B(y) \geq \sup_{x\in X}[1_A(x) - 1_C(x,y)] = 1_{A^C}(y)$.

Η τελευταία ισότητα είναι αληθής, αφού:

Αν $y \in A^C$, τότε υπάρχει $x_0 \in A$ τέτοιο ώστε $(x_0,y) \notin C$. Επομένως, $\sup_{x\in X}[1_A(x) - 1_C(x,y)] \geq 1_A(x_0) - 1_C(x_0,y) = 1 - 0 = 1$ και εφόσον από την άλλη $\sup_{x\in X}[1_A(x) - 1_C(x,y)] \leq 1$, ισχύει η ισότητα $\sup_{x\in X}[1_A(x) - 1_C(x,y)] = 1 = 1_{A^C}(y)$.

Αν $y \notin A^C$, τότε για κάθε $x \in A$, $(x,y) \in C$. Επομένως, για κάθε $x \in X$, $1_A(x) \leq 1_C(x,y)$, άρα και $\sup_{x\in X}[1_A(x) - 1_C(x,y)] \leq 0$. Από την άλλη, αν $x_0 \in A$, έχουμε $\sup_{x\in X}[1_A(x) - 1_C(x,y)] \geq 1_A(x_0) - 1_C(x_0,y) = 1 - 1 = 0$. Τελικά, $\sup_{x\in X}[1_A(x) - 1_C(x,y)] = 0 = 1_{A^C}(y)$.

Έχοντας, λοιπόν, διασφαλίσει ότι για κάθε $y \in Y$, ισχύει $1_B(y) \geq 1_{A^C}(y)$ και συνεπώς $A^C \subseteq B$, έχουμε $\nu[A^C] \leq \nu[B]$ και άρα $\mu[A] - \nu[B] \leq \mu[A] - \nu[A^C]$.

Έτσι, λοιπόν, έχουμε

$$\begin{aligned}&\sup\big\{\mu[A] - \nu[B] :\; (1_A,-1_B) \in \widetilde{\Phi_c}, \quad A \subseteq X \text{ κλειστό}\big\}\\ &\quad\leq \sup\{\mu[A] - \nu[A^C] :\; A \subseteq X \text{ κλειστό}\}.\end{aligned}$$

Η αντίστροφη ανισότητα είναι, επίσης, αληθής αφού για κάθε $A \subseteq X$ κλειστό ισχύει $\left(1_A, -1_{A^c}\right) \in \widetilde{\Phi_c}$.

Πράγματι, για κάθε $(x, y) \in X \times Y$ ισχύει $1_A(x) - 1_{A^c}(y) \leq 1_C(x, y)$. Αν $x \in A$, $y \notin A^C$, τότε $(x, y) \in C$, άρα έχουμε $1_A(x) - 1_{A^c}(y) = 1 - 0 \leq 1 = 1_C(x, y)$. Στις άλλες περιπτώσεις η ανισότητα είναι προφανής. Άρα, ο Ισχυρισμός 14 έχει αποδειχθεί.

■

Επιστρέφοντας στην απόδειξη του Θεωρήματος, σύμφωνα με τη σχέση (1.26), τον Ισχυρισμό 14 και τα αναφερόμενα πριν από αυτόν, έχουμε αποδείξει ότι

$$\inf_{\pi \in \Pi(\mu,\nu)} \pi[C] = \sup\{\mu[A] - \nu[A^C] : \ A \subseteq X \text{ κλειστό}\}$$

όπως επιθυμούσαμε. Η απόδειξη του Θεωρήματος 1.22 είναι πλήρης.

■

ΚΕΦΑΛΑΙΟ 2: Η γεωμετρία του προβλήματος βέλτιστης μεταφοράς

Στο δεύτερο κεφάλαιο, θα ασχοληθούμε με την αναγνώριση των βέλτιστων πλάνων μεταφοράς. Λόγω απλότητας των αποτελεσμάτων αλλά και χρησιμότητας στις διάφορες εφαρμογές, θα δουλέψουμε στην περίπτωση που η συνάρτηση κόστους είναι μια τετραγωνική συνάρτηση στον $\mathbb{R}^n$. Το βασικό αποτέλεσμα θα είναι **το κριτήριο Knott-Smith**, το οποίο προϋποθέτει ότι τα μέτρα πιθανότητας μ και ν με τα οποία δουλεύουμε έχουν πεπερασμένες ροπές 2ης τάξης. Σύμφωνα με το κριτήριο αυτό, ένα πλάνο μεταφοράς είναι βέλτιστο αν και μόνο αν συγκεντρώνεται στο υποδιαφορικό μιας κυρτής συνάρτησης. Ένα ακόμα αποτέλεσμα, το οποίο θα αναφέρουμε ως το **Θεώρημα Brenier**, ισχυρίζεται ότι το βέλτιστο πλάνο μεταφοράς είναι μοναδικό, υπό την επιπλέον υπόθεση ότι το μέτρο μ δεν δίνει μάζα στα «μικρά» σύνολα. Αναφέρουμε ως «μικρό» σύνολο κάθε υποσύνολο του $\mathbb{R}^n$ που έχει μέτρο Lebesgue ίσο με 0.

2.1. Μια απόδειξη για τετραγωνική συνάρτηση κόστους βασισμένη στη δυϊκότητα.

Θεωρούμε τους τοπολογικούς χώρους $X = Y = \mathbb{R}^n$ εφοδιασμένους με την Borel σ-άλγεβρα. Θεωρούμε, ακόμα, μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$, τα οποία έχουν πεπερασμένες ροπές 2ης τάξης, κάτι που μπορεί να εκφραστεί μέσω της σχέσης:

$$M_2 := \int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y) < +\infty. \tag{2.1}$$

Η συνάρτηση κόστους είναι το τετράγωνο της Ευκλείδειας νόρμας, πολλαπλασιασμένη με τον παράγοντα $\frac{1}{2}$ μόνο για λόγους απλότητας των αποδείξεων. Δηλαδή, $c(x,y) = \frac{|x-y|^2}{2}$.

Το $\Pi(\mu,\nu)$ είναι το σύνολο των μέτρων πιθανότητας στον $\mathbb{R}^n \times \mathbb{R}^n$, που έχουν περιθώρια μέτρα μ και ν.

Έτσι, το συναρτησοειδές I είναι:

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|x-y|^2}{2} d\pi(x,y). \tag{2.2}$$

Η συνθήκη στη σχέση (2.1) εξασφαλίζει ότι το συναρτησοειδές I είναι πάντοτε πεπερασμένο στο $\Pi(\mu,\nu)$. Πράγματι, αν $\pi \in \Pi(\mu,\nu)$, έχουμε:

$$I[\pi] = \int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|x-y|^2}{2}\,d\pi(x,y) \leq \int_{\mathbb{R}^n\times\mathbb{R}^n} (|x|^2+|y|^2)\,d\pi(x,y)$$
$$= \int_{\mathbb{R}^n} |x|^2\,d\mu(x) + \int_{\mathbb{R}^n} |y|^2\,d\nu(y) = 2M_2 < +\infty.$$

Θυμίζουμε ότι το πρόβλημα μεταφοράς μάζας του Kantorovich συνίσταται στην ελαχιστοποίηση του γραμμικού συναρτησοειδούς I πάνω στο σύνολο $\Pi(\mu,\nu)$.

Η αρχή δυϊκότητας του Kantorovich, την οποία αποδείξαμε στο Θεώρημα 1.3 του 1[ου] κεφαλαίου, εκφράζεται από τη σχέση

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi). \tag{2.3}$$

Θυμίζουμε ότι το συναρτησοειδές J ορίζεται πάνω στο σύνολο $L^1(d\mu)\times L^1(d\nu)$ μέσω του τύπου

$$J(\varphi,\psi) = \int_X \varphi\,d\mu + \int_Y \psi\,d\nu$$

και Φ_c ορίστηκε να είναι το σύνολο όλων των ζευγών μετρήσιμων συναρτήσεων $(\varphi,\psi)\in L^1(d\mu)\times L^1(d\nu)$ που ικανοποιούν την ανισότητα $\varphi(x)+\psi(y)\leq c(x,y)$ για $d\mu$ −σχεδόν όλα τα $x\in\mathbb{R}^n$ και $d\nu$ −σχεδόν όλα τα $y\in\mathbb{R}^n$, υπό την έννοια που αναφέρεται στην παρατήρηση 1.4.(vi) του 1[ου] κεφαλαίου.

2.1.1. Το πρωταρχικό πρόβλημα.

Πρόταση 2.1 (Ύπαρξη ενός βέλτιστου πλάνου μεταφοράς)

Υπάρχει βέλτιστο πλάνο μεταφοράς για το πρόβλημα ελαχιστοποίησης του Kantorovich, δηλαδή υπάρχει $\pi_*\in\Pi(\mu,\nu)$ τέτοιο ώστε

$$I[\pi_*] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Παρόλο που έχουμε ήδη αποδείξει την πρόταση αυτή στο προηγούμενο κεφάλαιο (στην απόδειξη του θεωρήματος 1.3), θα επαναλάβουμε την απόδειξη χρησιμοποιώντας μόνο την ιδιότητα της κάτω ημισυνέχειας της συνάρτησης κόστους $c(x,y)$.

<u>Απόδειξη:</u>

Όπως έχουμε ήδη αποδείξει στο 1ο κεφάλαιο (Ισχυρισμός 5 στην απόδειξη του θεωρήματος 1.3), το σύνολο $\Pi(\mu,\nu)$ είναι ακολουθιακά συμπαγές στον $P(X\times Y)$ αναφορικά με την ασθενή τοπολογία των μέτρων πιθανότητας, δηλαδή την τοπολογία που επάγεται από τον $C_b(X\times Y)$. Η ακολουθιακή συμπάγεια του $\Pi(\mu,\nu)$ εξασφαλίζει την υπάρξη του βέλτιστου πλάνου μεταφοράς για το πρόβλημα Kantorovich.

Πράγματι, έστω μια ακολουθία $(\pi_n)_{n\in\mathbb{N}}$ στο $\Pi(\mu,\nu)$ για την οποία ισχύει

$$\lim_{n\to\infty} I[\pi_n] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Λόγω της ακολουθιακής συμπάγειας του $\Pi(\mu,\nu)$, η ακολουθία αυτή θα έχει σημείο συσσώρευσης $\pi_* \in \Pi(\mu,\nu)$. Υπάρχει, δηλαδή, υπακολουθία $\left(\pi_{n_k}\right)_{k\in\mathbb{N}}$ της $(\pi_n)_{n\in\mathbb{N}}$ η οποία συγκλίνει ασθενώς στο π_*. Αυτό σημαίνει ότι για κάθε $f\in C_b(\mathbb{R}^n\times\mathbb{R}^n)$, ισχύει:

$$\lim_{k\to\infty}\int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,y)\,d\pi_{n_k}(x,y) = \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,y)\,d\pi_*(x,y).$$

Θεωρούμε τις συναρτήσεις $c_l = \inf\{c,l\}$, $l=1,2,\dots$ Τότε, οι c_l είναι συνεχείς, φραγμένες και

$$c(x,y) = \sup_{l\in\mathbb{N}} c_l(x,y),$$

Έχουμε, λοιπόν:

$$\begin{aligned} I[\pi_*] &= \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,d\pi_*(x,y) = \lim_{l\to\infty}\int_{\mathbb{R}^n\times\mathbb{R}^n} c_l(x,y)\,d\pi_*(x,y) \\ &= \lim_{l\to\infty}\left(\lim_{k\to\infty}\int_{\mathbb{R}^n\times\mathbb{R}^n} c_l(x,y)\,d\pi_{n_k}(x,y)\right) \le \lim_{k\to\infty}\int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,d\pi_{n_k}(x,y) = \lim_{k\to\infty} I\left[\pi_{n_k}\right] \\ &= \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]. \end{aligned}$$

Σημειώνουμε ότι η δεύτερη ισότητα ισχύει λόγω του Θεωρήματος Μονότονης Σύγκλισης, η τρίτη ισότητα λόγω του ότι οι c_l είναι συνεχείς και φραγμένες, ενώ η ανισότητα ισχύει επειδή $c_l \le c$. Έπεται, λοιπόν, ότι

$$I[\pi_*] = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]$$

δηλαδή το πλάνο μεταφοράς π_* είναι βέλτιστο όσον αφορά την ελαχιστοποίηση του συναρτησοειδούς I.

∎

2.1.2. Το δυϊκό πρόβλημα.

Η συνθήκη ώστε το ζεύγος (φ,ψ) να ανήκει στο σύνολο Φ_c είναι $\varphi(x)+\psi(y)\leq\frac{|x-y|^2}{2}$ η οποία, υπό την έννοια της παρατήρησης 1.4(vi), ισχύει για $d\mu$ −σχεδόν όλα τα $x\in\mathbb{R}^n$ και $d\nu$ −σχεδόν όλα τα $y\in\mathbb{R}^n$. Αναπτύσσοντας το δεύτερο μέλος της ανισότητας, έχουμε

$$\varphi(x)+\psi(y)\leq\frac{|x|^2}{2}+\frac{|y|^2}{2}-x\cdot y$$

και αναδιατάσσοντας τους όρους, παίρνουμε

$$x\cdot y\leq\left[\frac{|x|^2}{2}-\varphi(x)\right]+\left[\frac{|y|^2}{2}-\psi(y)\right].$$

Η μορφή αυτή μας υποδεικνύει να θεωρήσουμε στο εξής ως νέους αγνώστους τις συναρτήσεις

$$\tilde{\varphi}(x)=\frac{|x|^2}{2}-\varphi(x),\qquad\tilde{\psi}(y)=\frac{|y|^2}{2}-\psi(y).$$

Έχουμε, λοιπόν:

$$\inf_{\pi\in\Pi(\mu,\nu)}I[\pi]=\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}\frac{|x-y|^2}{2}d\pi(x,y)=\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}\left(\frac{|x|^2}{2}+\frac{|y|^2}{2}-x\cdot y\right)d\pi(x,y)$$

και χρησιμοποιώντας τη σχέση (2.1), έχουμε:

$$\inf_{\pi\in\Pi(\mu,\nu)}I[\pi]=M_2-\sup_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi(x,y).\tag{2.4}$$

Ακόμα,

$$\begin{aligned}\sup_{(\varphi,\psi)\in\Phi_c}J(\varphi,\psi)&=\sup_{(\varphi,\psi)\in\Phi_c}\left\{\int_{\mathbb{R}^n}\varphi(x)\,d\mu(x)+\int_{\mathbb{R}^n}\psi(y)\,d\nu(y)\right\}\\&=\sup_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}}\left\{\int_{\mathbb{R}^n}\left(\frac{|x|^2}{2}-\tilde{\varphi}(x)\right)d\mu(x)+\int_{\mathbb{R}^n}\left(\frac{|y|^2}{2}-\tilde{\psi}(y)\right)d\nu(y)\right\}\end{aligned}$$

όπου $\tilde{\Phi}$ είναι το σύνολο των ζευγών συναρτήσεων $(\tilde{\varphi},\tilde{\psi})\in L^1(d\mu)\times L^1(d\nu)$ που παίρνουν τιμές στο σύνολο $\mathbb{R}\cup\{+\infty\}$ και ικανοποιούν την ανισότητα

$$x\cdot y\leq\tilde{\varphi}(x)+\tilde{\psi}(y)\tag{2.5}$$

για $d\mu$ −σχεδόν όλα τα $x\in\mathbb{R}^n$ και $d\nu$ −σχεδόν όλα τα $y\in\mathbb{R}^n$.

Χρησιμοποιώντας και πάλι τη σχέση (2.1), έχουμε

$$\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = M_2 - \inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}} J(\tilde{\varphi},\tilde{\psi}). \tag{2.6}$$

Εφαρμόζοντας τη σχέση (2.3) και αξιοποιώντας τις σχέσεις (2.4) και (2.6), η δυϊκότητα Kantorovich μετασχηματίζεται στη σχέση:

$$\sup_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} (x\cdot y)\,d\pi(x,y) = \inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}} J(\tilde{\varphi},\tilde{\psi}). \tag{2.7}$$

Στη συνέχεια, θα περιγράψουμε τη μέθοδο της διπλής κυρτοποίησης, την οποία θα χρησιμοποιήσουμε προκειμένου να βελτιώσουμε τα αποδεκτά ζεύγη συναρτήσεων που χρησιμοποιούνται στο δυϊκό πρόβλημα.

Θεωρούμε ένα ζεύγος $(\varphi,\psi)\in\tilde{\Phi}$. Από την παρατήρηση 1.4(vi) και τον ορισμό του συνόλου $\tilde{\Phi}$, γνωρίζουμε ότι υπάρχουν μετρήσιμα σύνολα $N_1, N_2 \subseteq \mathbb{R}^n$ για τα οποία ισχύει $\mu[N_1]=\nu[N_2]=0$ και η ανισότητα (2.5) ισχύει για κάθε $(x,y)\in N_1^c\times N_2^c$. Αλλάζουμε τις τιμές των φ και ψ, έτσι ώστε να είναι ίσες με $+\infty$ πάνω στα σύνολα N_1 και N_2 αντίστοιχα. Τότε, το ζεύγος (φ,ψ) συνεχίζει να ανήκει στο $\tilde{\Phi}$, η ανισότητα (2.5) ισχύει για κάθε $x,y\in\mathbb{R}^n$, ενώ ακόμα, η τιμή του συναρτησοειδούς $J(\varphi,\psi)$ παραμένει αμετάβλητη.

Έτσι, λοιπόν, για κάθε $y\in\mathbb{R}^n$ ορίζουμε $\varphi^*(y) \coloneqq \sup_{x\in\mathbb{R}^n}[x\cdot y-\varphi(x)]$.

Από την ανισότητα (2.5) προκύπτει ότι για κάθε $y\in\mathbb{R}^n$ ισχύει:

$$\varphi^*(y)\le\psi(y) \tag{2.8}$$

οπότε αν, επιπλέον, δεχτούμε ότι $\varphi^*\in L^1(d\nu)$ τότε έχουμε

$$J(\varphi,\varphi^*)\le J(\varphi,\psi). \tag{2.9}$$

Στη συνέχεια, για κάθε $x\in\mathbb{R}^n$ ορίζουμε $\varphi^{**}(x)\coloneqq\sup_{y\in\mathbb{R}^n}[x\cdot y-\varphi^*(y)]$.

Από τον ορισμό της φ^*, έχουμε ότι για κάθε $y\in\mathbb{R}^n$ ισχύει $\varphi^*(y)\ge x\cdot y-\varphi(x)$, $x\in\mathbb{R}^n$ και άρα, για κάθε $x\in\mathbb{R}^n$, μπορούμε να γράψουμε $\varphi(x)\ge x\cdot y-\varphi^*(y)$, $y\in\mathbb{R}^n$.

Επομένως, προκύπτει ότι

$$\varphi^{**}(x)\le\varphi(x). \tag{2.10}$$

Αν και πάλι δεχτούμε ότι $\varphi^{**}\in L^1(d\mu)$, τότε έχουμε

$$J(\varphi^{**},\varphi^*)\le J(\varphi,\varphi^*). \tag{2.11}$$

Συνδυάζοντας τις ανισότητες (2.9) και (2.11), παίρνουμε $J(\varphi^{**},\varphi^*)\le J(\varphi,\psi)$, άρα και

$$\inf_{\varphi\in L^1(d\mu)} J(\varphi^{**},\varphi^*) \le \inf_{(\varphi,\psi)\in\widetilde{\Phi}} J(\varphi,\psi). \tag{2.12}$$

Για κάθε $x_0, y_0 \in \mathbb{R}^n$, έχουμε

$$\varphi^{**}(x_0)+\varphi^*(y_0) = \sup_{y\in\mathbb{R}^n}[x_0\cdot y-\varphi^*(y)]+\varphi^*(y_0) \ge x_0\cdot y_0-\varphi^*(y_0)+\varphi^*(y_0) = x_0\cdot y_0$$

δηλαδή το ζεύγος (φ^{**},φ^*) ικανοποιεί την ανισότητα (2.5) (αν $\varphi^*(y_0)=+\infty$, η ανισότητα είναι προφανής). Ακόμα, έχοντας δεχτεί ότι $(\varphi^{**},\varphi^*)\in L^1(d\mu)\times L^1(d\nu)$, προκύπτει ότι $(\varphi^{**},\varphi^*)\in\widetilde{\Phi}$. Συμπεραίνουμε ότι μπορούμε να περιορίσουμε το συναρτησοειδές J πάνω στο υποσύνολο του $\widetilde{\Phi}$ που αποτελείται από τα ζεύγη συναρτήσεων (φ^{**},φ^*) χωρίς να μεταβληθεί η τιμή του $\inf_{\widetilde{\Phi}} J$.

Πριν συνεχίσουμε τη μελέτη του δυϊκού προβλήματος, θα αφιερώσουμε την επόμενη υποενότητα στην καταγραφή και απόδειξη μερικών βασικών ιδιοτήτων των κυρτών συναρτήσεων.

2.1.3. Ορισμοί και προτάσεις από την κυρτή ανάλυση.

1. Ορισμοί.

Μια συνάρτηση $\varphi:\mathbb{R}^n\to[-\infty,+\infty]$ λέγεται κανονική αν και μόνο αν $\varphi(x)<+\infty$ για τουλάχιστον ένα x και $\varphi(x)>-\infty$ για κάθε x.

Μια συνάρτηση φ λέγεται κυρτή αν για κάθε $x,y\in\mathbb{R}^n$ και κάθε $t\in(0,1)$ ισχύει:

$$\varphi(tx+(1-t)y)\le t\varphi(x)+(1-t)\varphi(y). \tag{2.13}$$

Η φ ονομάζεται αυστηρά κυρτή αν για κάθε $x,y\in\mathbb{R}^n$ με $x\neq y$ και κάθε $t\in(0,1)$ ισχύει:

$$\varphi(tx+(1-t)y)< t\varphi(x)+(1-t)\varphi(y). \tag{2.14}$$

Ορίζουμε

$$Dom(\varphi) \coloneqq \{x\in\mathbb{R}^n \ : \ \varphi(x)<+\infty\},$$

δηλαδή $Dom(\varphi)$ είναι το (κυρτό) σύνολο των σημείων του $\mathbb{R}^n$ στα οποία η φ λαμβάνει πεπερασμένες τιμές, και

$$epi\varphi \coloneqq \{(x,\mu)\in\mathbb{R}^n\times\mathbb{R}: \ \ \mu\ge\varphi(x)\}$$

το οποίο ονομάζεται επιγράφημα της φ.

Ένας ισοδύναμος ορισμός της κυρτότητας είναι ο εξής:

Μια συνάρτηση φ λέγεται κυρτή αν το σύνολο $epi\varphi$ είναι κυρτό ως υποσύνολο του $\mathbb{R}^{n+1}$.

Απόδειξη της ισοδυναμίας των δύο ορισμών:

Υποθέτουμε ότι για κάθε $x, y \in \mathbb{R}^n$ ισχύει

$$\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y)$$

και θα δείξουμε ότι το σύνολο $epi\varphi$ είναι κυρτό.

Έστω δύο σημεία $(x_1, \mu_1), (x_2, \mu_2) \in epi\varphi$ και $t \in (0,1)$.

Αρκεί να δείξουμε ότι $t(x_1, \mu_1) + (1-t)(x_2, \mu_2) = (tx_1 + (1-t)x_2, t\mu_1 + (1-t)\mu_2) \in epi\varphi$.

Πράγματι, $t\mu_1 + (1-t)\mu_2 \geq t\varphi(x_1) + (1-t)\varphi(x_2) \geq \varphi(tx_1 + (1-t)x_2)$ το οποίο εξ ορισμού σημαίνει ότι $t(x_1, \mu_1) + (1-t)(x_2, \mu_2) \in epi\varphi$ και άρα το $epi\varphi$ είναι κυρτό.

Αντίστροφα, έστω ότι το σύνολο $epi\varphi$ είναι κυρτό. Αν $\mu_1, \mu_2 \in \mathbb{R}$ με $\mu_1 \geq \varphi(x_1)$, $\mu_2 \geq \varphi(x_2)$ και $t \in (0,1)$, τότε $t(x_1, \mu_1) + (1-t)(x_2, \mu_2) = (tx_1 + (1-t)x_2, t\mu_1 + (1-t)\mu_2) \in epi\varphi$.
Άρα, $\varphi(tx_1 + (1-t)x_2) \leq t\mu_1 + (1-t)\mu_2$.
Ειδικότερα, η ανισότητα ισχύει για $\mu_1 = \varphi(x_1)$ και $\mu_2 = \varphi(x_2)$, επομένως $\varphi(tx_1 + (1-t)x_2) \leq t\varphi(x_1) + (1-t)\varphi(x_2)$, δηλαδή η φ είναι κυρτή.

■

Λήμμα 2.2

Έστω $C \subseteq \mathbb{R}^n$ κυρτό, $x \in Int(C)$ και $y \in \overline{C}$. Τότε, $(1-\lambda)x + \lambda y \in Int(C)$ για $0 \leq \lambda < 1$.

Απόδειξη:

Έστω $\lambda \in [0,1)$. Θα δείξουμε ότι υπάρχει $\varepsilon > 0$ τέτοιο ώστε $(1-\lambda)x + \lambda y + \varepsilon B \subseteq C$, όπου $B = B_1(0)$.

Αφού $y \in \overline{C}$, $y \in C + \varepsilon B$ για κάθε $\varepsilon > 0$. Άρα, για κάθε $\varepsilon > 0$

$$(1-\lambda)x + \lambda y + \varepsilon B \subseteq (1-\lambda)x + \lambda(C + \varepsilon B) + \varepsilon B = (1-\lambda)\left(x + \varepsilon \frac{1+\lambda}{1-\lambda} B\right) + \lambda C.$$

Αφού $x \in Int(C)$, το τελευταίο σύνολο περιέχεται στο $(1-\lambda)C + \lambda C = C$ αν πάρουμε το ε αρκετά μικρό.

■

Λήμμα 2.3

Αν η φ είναι κυρτή, τότε

$$Int(epi\varphi) = \{(x,\mu) \in \mathbb{R}^n \times \mathbb{R} : \quad x \in Int(Dom(\varphi)) \ \text{ και } \ \mu > \varphi(x)\}.$$

<u>Απόδειξη:</u>

Έστω $(x_0, \mu_0) \in Int(epi\varphi)$. Βρίσκουμε μπάλα κέντρου (x_0, μ_0) και ακτίνας ε στον $\mathbb{R}^{n+1}$, την οποία συμβολίζουμε $B_\varepsilon(x_0, \mu_0)$, τέτοια ώστε $B_\varepsilon(x_0, \mu_0) \subseteq epi\varphi$.

Παίρνοντας την προβολή της $B_\varepsilon(x_0, \mu_0)$ στον $\mathbb{R}^n$ προκύπτει η μπάλα $B_\varepsilon(x_0) \subseteq \mathbb{R}^n$ για την οποία ισχύει $\varphi(x) < +\infty, \ \ x \in B_\varepsilon(x_0)$. Άρα, $x_0 \in Int(Dom(\varphi))$ και $\mu_0 > \mu_0 - \varepsilon \geq \varphi(x_0)$.

Για την απόδειξη του αντίστροφου εγκλεισμού, έστω $x_0 \in Int(Dom(\varphi))$ και $\mu_0 \in \mathbb{R}$ με $\mu_0 > \varphi(x_0)$. Έστω $a_1, \dots, a_r \in Dom(\varphi)$ τέτοια ώστε $x_0 \in Int(P)$, όπου $P = Conv\{a_1, \dots, a_r\}$ είναι το σύνολο των κυρτών συνδυασμών των σημείων $a_1, \dots, a_r$. Έστω, ακόμα, $a = \max_{1 \leq i \leq r}\{\varphi(a_i)\}$. Για κάθε $x \in P$, μπορούμε να εκφράσουμε το x ως κυρτό συνδυασμό

$$x = \lambda_1 a_1 + \cdots + \lambda_r a_r, \qquad \lambda_i \geq 0, \ \ \lambda_1 + \cdots + \lambda_r = 1$$

και άρα

$$\varphi(x) \leq \lambda_1 \varphi(a_1) + \cdots + \lambda_r \varphi(a_r) \leq (\lambda_1 + \cdots + \lambda_r)a = a$$

Άρα, το ανοικτό σύνολο $\{(x,\mu) : \ x \in Int(P), \ \mu > a\}$ περιέχεται στο $epi\varphi$. Ειδικότερα, για κάθε $\mu > a$, έχουμε $(x_0, \mu) \in Int(epi\varphi)$ και προφανώς $(x_0, \varphi(x_0)) \in epi\varphi$. Άρα, το (x_0, μ_0) είναι εσωτερικό σημείο ενός κατακόρυφου ευθύγραμμου τμήματος στο $epi\varphi$ το οποίο συναντά το $Int(epi\varphi)$. Από το λήμμα 2.2, $(x_0, \mu_0) \in Int(epi\varphi)$.

■

Λήμμα 2.4

Αν η φ είναι μη κανονική κυρτή συνάρτηση, τότε $\varphi(x) = -\infty$ για κάθε $x \in Int(Dom(\varphi))$.

<u>Απόδειξη:</u>

Αν $Dom(\varphi) \neq \emptyset$, τότε το $Dom(\varphi)$ περιέχει σημεία όπου η φ έχει την τιμή $-\infty$. Έστω, λοιπόν, $u \in Dom(\varphi)$ με $\varphi(u) = -\infty$ και $x \in Int(Dom(\varphi))$. Υπάρχει $\mu > 1$ τέτοιο ώστε $y = (1-\mu)u + \mu x \in Dom(\varphi)$. Έχουμε, τότε $x = (1-\lambda)u + \lambda y$ με $0 < \lambda = \frac{1}{\mu} < 1$. Από την κυρτότητα της φ, έχουμε $\varphi(x) = \varphi((1-\lambda)u + \lambda y) < (1-\lambda)\alpha + \lambda\beta$ για οποιαδήποτε $\alpha > \varphi(u)$ και $\beta > \varphi(y)$. Αφού $\varphi(u) = -\infty$ και $\varphi(y) < +\infty$, προκύπτει ότι $\varphi(x) = -\infty$.

■

Λήμμα 2.5

Έστω μια συνάρτηση $\varphi\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Οι ακόλουθες προτάσεις είναι ισοδύναμες:

a) Η φ είναι κάτω ημισυνεχής στον $\mathbb{R}^n$.
b) Για κάθε $a \in \mathbb{R}$, το σύνολο $A = \{x \in \mathbb{R}^n : \varphi(x) \leq a\}$ είναι κλειστό.
c) Το $epi\varphi$ είναι ένα κλειστό υποσύνολο του $\mathbb{R}^{n+1}$.

<u>Απόδειξη:</u>

Καταρχήν θα αποδείξουμε την ισοδυναμία των προτάσεων (a) και (c).

Έστω ότι η φ είναι κάτω ημισυνεχής στον $\mathbb{R}^n$. Έστω, άκομα, $\{(x_k, \mu_k)\}_{k=1}^{\infty}$ μια ακολουθία στοιχείων του $epi\varphi$ που συγκλίνει στο $(x, \mu) \in \mathbb{R}^{n+1}$. Τότε, η σύγκλιση είναι κατά συντεταγμένη, δηλαδή $x_k \to x$ και $\mu_k \to \mu$ καθώς $k \to \infty$, και για κάθε $k \in \mathbb{N}$, έχουμε $\mu_k \geq \varphi(x_k)$.

Άρα, $\mu = \lim_{k\to\infty} \mu_k \geq \liminf_{k\to\infty} \varphi(x_k)$.
Αφού η φ είναι κάτω ημισυνεχής, δηλαδή $\liminf_{k\to\infty} \varphi(x_k) \geq \varphi(x)$, προκύπτει ότι $\mu \geq \varphi(x)$, δηλαδή $(x, \mu) \in epi\varphi$. Συνεπώς, το $epi\varphi$ είναι κλειστό στον $\mathbb{R}^{n+1}$.

Αντίστροφα, έστω ότι το $epi\varphi$ είναι κλειστό. Υποθέτουμε, προς απαγωγή εις άτοπο, ότι η φ δεν είναι κάτω ημισυνεχής, δηλαδή υπάρχει $x_0 \in \mathbb{R}^n$ τέτοιο ώστε $\liminf_{x\to x_0} \varphi(x) < \varphi(x_0)$.

Τότε, υπάρχει $\delta > 0$ τέτοιο ώστε $\liminf_{x\to x_0} \varphi(x) < \varphi(x_0) - \delta$.

Εξ ορισμού $\liminf_{x\to x_0} \varphi(x) = \lim_{\varepsilon\downarrow 0} \inf_{x\in B_\varepsilon(x_0)} \varphi(x)$, και συνεπώς για κάθε $\varepsilon > 0$, ισχύει $\inf_{x\in B_\varepsilon(x_0)} \varphi(x) < \varphi(x_0) - \delta$.
Άρα, για κάθε $k \in \mathbb{N}$, υπάρχει $x_k \in B_{\frac{1}{k}}(x_0)$ τέτοιο ώστε $\varphi(x_k) < \varphi(x_0) - \delta$. Παρατηρούμε ότι $x_k \to x_0$ καθώς $k \to +\infty$. Επίσης, παίρνουμε τη σταθερή ακολουθία $\mu_k = \varphi(x_0) - \delta,\quad k \in \mathbb{N}$.

Τότε, για κάθε $k \in \mathbb{N}$, ισχύει $(x_k, \mu_k) \in epi\varphi$ και $(x_k, \mu_k) \to (x_0, \varphi(x_0) - \delta)$. Λόγω κλειστότητας του $epi\varphi$, ισχύει $(x_0, \varphi(x_0) - \delta) \in epi\varphi$, δηλαδή $\varphi(x_0) - \delta \geq \varphi(x_0)$.

Έχουμε καταλήξει σε αντίφαση, άρα η φ είναι κάτω ημισυνεχής.

Θα αποδείξουμε τώρα την ισοδυναμία των προτάσεων (a) και (b).

Έστω ότι η φ είναι κάτω ημισυνεχής στον $\mathbb{R}^n$ και έστω $a \in \mathbb{R}$. Παίρνουμε μια ακολουθία $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^n$ με $\varphi(x_k) \leq a$ και $x_k \to x$. Τότε $\varphi(x) \leq \liminf_{k\to\infty} \varphi(x_k) \leq a$, δηλαδή $x \in A$ και το σύνολο A είναι κλειστό.

Αντίστροφα, παίρνουμε μια ακολουθία $\{x_k\}_{k=1}^{\infty}$ με $x_k \to x$ και έστω $\mu = \liminf_{k\to\infty} \varphi(x_k)$.

Υπάρχει υπακολουθία x_{k_n} της x_k τέτοια ώστε $\varphi(x_{k_n}) \to \mu$. Αν $\alpha > \mu$, τότε υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $\varphi(x_{k_n}) < \alpha, \quad n \geq n_0$.

Επομένως, για το x που είναι το όριο της x_{k_n}, ισχύει: $x \in \overline{\{y \in \mathbb{R}^n : \varphi(y) \leq a\}}$.

Άρα, αν υποθέσουμε ότι το σύνολο A είναι κλειστό, τότε καταλήγουμε στο εξής: $x \in \{y \in \mathbb{R}^n : \varphi(y) \leq a\}$, δηλαδή $\varphi(x) \leq \alpha$.

Επειδή το α ήταν ένας αυθαίρετος αριθμός μεγαλύτερος του μ, συμπεραίνουμε ότι $\varphi(x) \leq \mu = \liminf_{k\to\infty} \varphi(x_k)$, δηλαδή η συνάρτηση φ είναι κάτω ημισυνεχής.

■

Παρατήρηση: Αντίστοιχα, για τις άνω ημισυνεχείς συναρτήσεις ισχύει:

Μια συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι άνω ημισυνεχής στον $\mathbb{R}^n$ αν και μόνο αν για κάθε $a \in \mathbb{R}$, το σύνολο $A = \{x \in \mathbb{R}^n : \varphi(x) \geq a\}$ είναι κλειστό.

Πρόταση 2.6

Αν $\varphi, \psi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι δύο κυρτές κάτω ημισυνεχείς συναρτήσεις τέτοιες ώστε $Int(Dom(\varphi)) = Int(Dom(\psi))$ και επίσης συμπίπτουν στο σύνολο αυτό, τότε είναι ίσες.

<u>Απόδειξη:</u>

Από την υπόθεση $\varphi = \psi$ στο σύνολο $Int(Dom(\varphi)) = Int(Dom(\psi))$ και το λήμμα 2.3, προκύπτει ότι $Int(epi\varphi) = Int(epi\psi)$. Άρα, $\overline{Int(epi\varphi)} = \overline{Int(epi\psi)}$.

Από το λήμμα 1.8 του 1ου Κεφαλαίου και εφόσον το $epi\varphi$ είναι κυρτό σύνολο, ισχύει $\overline{epi\varphi} = \overline{Int(epi\varphi)}$. Άρα, $\overline{epi\varphi} = \overline{epi\psi}$.

Λόγω της κάτω ημισυνέχειας των φ και ψ, από το λήμμα 2.5 τα σύνολα $epi\varphi$ και $epi\psi$ είναι κλειστά, άρα $epi\varphi = epi\psi$. Έπεται ότι $\varphi(x) = \psi(x), \ \ x \in \mathbb{R}^n$.

■

Ορισμός 2.7

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική συνάρτηση. Ορίζουμε ως κλειστότητα της φ, $cl\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, τη μεγαλύτερη κάτω ημισυνεχή συνάρτηση που κυριαρχείται από την φ.

Ένας ισοδύναμος ορισμός της συνάρτησης $cl\varphi$ δίνεται μέσω της σχέσης $epi(cl\varphi) \coloneqq \overline{epi\varphi}$.

<u>Απόδειξη της ισοδυναμίας των δύο ορισμών:</u>

Καταρχήν, πρέπει να δείξουμε ότι το σύνολο $K = \overline{epi\varphi}$ είναι όντως το επιγράφημα κάποιας συνάρτησης. Προς την απόδειξη αυτού, έστω ένα σημείο $(x, \mu) \in K$ και $k > \mu$. Υπάρχει ακολουθία $(x_n, \mu_n) \in epi\varphi$ τέτοια ώστε $(x_n, \mu_n) \longrightarrow (x, \mu)$, δηλαδή $x_n \to x$ και $\mu_n \to \mu$. Φτιάχνουμε την ακολουθία $k_n = \mu_n + (k - \mu)$ για την οποία ισχύει $k_n \to k$ και $k_n > \mu_n \geq \varphi(x_n)$. Άρα, $(x_n, k_n) \in epi\varphi$ και το όριο αυτής, $(x, k) \in K$. Αποδείξαμε, λοιπόν, ότι το σύνολο K περιέχει επίσης την κατακόρυφη ημιευθεία $\{(x, k) \in \mathbb{R}^n \times \mathbb{R} : \ k > \mu\}$ και συνεπώς, αφού είναι κλειστό, όντως αποτελεί το επιγράφημα κάποιας συνάρτησης.

Ονομάζουμε ψ τη συνάρτηση αυτή. Δηλαδή $K = epi\psi$ και θα δείξουμε ότι η ψ ικανοποιεί τον πρώτο ορισμό. Έχουμε, λοιπόν:

Το σύνολο K είναι κλειστό, άρα η ψ είναι κάτω ημισυνεχής. Ισχύει $epi\varphi \subseteq epi\psi$, άρα $\psi \leq \varphi$. Ακόμα, αν g είναι μια κάτω ημισυνεχής συνάρτηση με $g \leq \varphi$, τότε το $epig$ είναι κλειστό και $epig \supseteq epi\varphi$. Άρα, ισχύει: $epig \supseteq \overline{epi\varphi} = epi\psi$, δηλαδή $g \leq \psi$.

Τελικά, η συνάρτηση ψ είναι η μεγαλύτερη κάτω ημισυνεχής συνάρτηση που κυριαρχείται από την φ.

■

Μια κανονική συνάρτηση φ λέγεται κλειστή, αν $cl\varphi = \varphi$. Άρα, για μια κανονική συνάρτηση, η κλειστότητα ισοδυναμεί με την κάτω ημισυνέχεια.

Αν φ είναι μια κανονική συνάρτηση , τότε εξ ορισμού η κλειστότητά της μπορεί να εκφραστεί μέσω του τύπου $(cl\varphi)(x) = \liminf_{y \to x} \varphi(y)$.

Πράγματι, εφόσον η $cl\varphi$ είναι μια κάτω ημισυνεχής συνάρτηση που κυριαρχείται από την φ, έχουμε $cl\varphi(x) \leq \liminf_{y \to x} cl\varphi(y) \leq \liminf_{y \to x} \varphi(y)$.

Για την απόδειξη της αντίστροφης ανισότητας, ας θεωρήσουμε ένα σημείο $\big(x, cl\varphi(x)\big) \in epi(cl\varphi) = \overline{epi\varphi}$. Τότε, υπάρχει ακολουθία $(x_n, \mu_n) \in epi\varphi$ τέτοια ώστε $(x_n, \mu_n) \longrightarrow \big(x, cl\varphi(x)\big)$, δηλαδή $x_n \to x$ και $\mu_n \to cl\varphi(x)$. Τότε ισχύει $\mu_n \geq \varphi(x_n)$ και έτσι προκύπτει ότι

$$cl\varphi(x) \geq \liminf_{n \to \infty} \varphi(x_n) \geq \liminf_{y \to x} \varphi(y).$$

Λήμμα 2.8

Αν φ είναι μια κανονική κυρτή συνάρτηση στον $\mathbb{R}^n$, τότε

$$(cl\varphi)(x) = \varphi(x), \qquad x \in \mathbb{R}^n - \partial Dom(\varphi).$$

Απόδειξη:

Έστω $x_0 \in Int\big(Dom(\varphi)\big)$. Θεωρούμε την κατακόρυφη ευθεία $M \coloneqq \{(x_0, \mu) : \ \mu \in \mathbb{R}\}$.

Τότε $M \cap epi\varphi = \{(x_0, \mu) : \ \mu \geq \varphi(x_0)\}$, το οποίο είναι ένα κλειστό υποσύνολο του $\mathbb{R}^{n+1}$. Επίσης, εξ ορισμού ισχύει $epi(cl\varphi) = \overline{epi\varphi}$, οπότε έχουμε $M \cap epi(cl\varphi) = M \cap \overline{epi\varphi}$.

Τέλος, εφόσον το σύνολο $epi\varphi$ είναι κυρτό και $M \cap Int(epi\varphi) \neq \emptyset$, σύμφωνα με την παρακάτω παρατήρηση, έχουμε ότι $\overline{M \cap epi\varphi} = M \cap \overline{epi\varphi}$. Επομένως,

$$M \cap epi(cl\varphi) = M \cap \overline{epi\varphi} = \overline{M \cap epi\varphi} = M \cap epi\varphi.$$

Άρα, $(cl\varphi)(x_0) = \varphi(x_0)$.

Έστω τώρα $x_0 \in \mathbb{R}^n - \overline{Dom(\varphi)}$. Τότε, επειδή $\overline{Dom(\varphi)} \supseteq Dom(cl\varphi) \supseteq Dom(\varphi)$, έπεται ότι $(cl\varphi)(x_0) = \varphi(x_0) = +\infty$.

Ο εγκλεισμός $Dom(cl\varphi) \subseteq \overline{Dom(\varphi)}$ είναι, πράγματι, σωστός αφού αν $x_0 \in Dom(cl\varphi)$, τότε $+\infty > cl\varphi(x_0) = \liminf_{x \to x_0} \varphi(x)$ και άρα, υπάρχει ακολουθία $x_n \to x_0$ τέτοια ώστε $\varphi(x_n) \to cl\varphi(x_0)$. Τελικά $\varphi(x_n) < +\infty$, άρα $x_n \in Dom(\varphi)$ και $x_0 \in \overline{Dom(\varphi)}$.

■

Παρατήρηση ([13], Πόρισμα 6.5.1)

Έστω C ένα κυρτό σύνολο και M ένα αφινικό σύνολο τέτοιο ώστε $M \cap Int(C) \neq \emptyset$. Τότε,

$$Int(M \cap C) = M \cap Int(C) \qquad \text{και} \qquad \overline{M \cap C} = M \cap \bar{C}$$

Πρόταση 2.9

Αν η φ είναι κυρτή συνάρτηση, τότε είναι συνεχής σε κάθε ανοικτό κυρτό σύνολο $C \subseteq \mathbb{R}^n$ στο οποίο δεν παίρνει την τιμή $+\infty$. Ειδικότερα, η φ είναι συνεχής στο $Int\big(Dom(\varphi)\big)$.

Απόδειξη:

Η συνάρτηση

$$g(x) = \begin{cases} \varphi(x), & x \in C \\ +\infty, & \text{αλλού} \end{cases}$$

έχει $Dom(g) = C$. Άρα, αντικαθιστώντας την φ με την g αν χρειαστεί, ανάγουμε το θεώρημα στην περίπτωση που $C = Int\big(Dom(\varphi)\big)$.

Αν η φ είναι μη κανονική, τότε είναι ταυτοτικά ίση με $-\infty$ (λήμμα 2.4), οπότε η συνέχεια είναι τετριμμένη. Υποθέτουμε, λοιπόν, ότι η φ είναι κανονική.

Από το λήμμα 2.8, $(cl\varphi)(x) = \varphi(x), \ \ x \in C$, οπότε η φ είναι κάτω ημισυνεχής στο C.

Αφού το $C = Int\big(Dom(\varphi)\big)$ είναι ανοικτό, από το λήμμα 2.3, έχουμε

$$Int(epi\varphi) = \{(x, \mu) \in C \times \mathbb{R} : \quad \mu > \varphi(x)\}.$$

Επομένως, για κάθε $a \in \mathbb{R}$, το σύνολο $\{x : \varphi(x) < a\}$ είναι η προβολή στον $\mathbb{R}^n$ του συνόλου $Int(epi\varphi) \cap \{(x, \mu) \in \mathbb{R}^n \times \mathbb{R} : \quad \mu < a\}$, το οποίο είναι ανοιχτό ως τομή δύο ανοιχτών συνόλων.

Έπεται ότι το $\{x : \varphi(x) < a\}$ είναι ανοιχτό, και άρα το συμπλήρωμα αυτού, το σύνολο $\{x : \varphi(x) \geq a\}$ είναι κλειστό για κάθε $a \in \mathbb{R}$. Σύμφωνα με την παρατήρηση που έπεται του λήμματος 2.5, η φ είναι άνω ημισυνεχής στο C.

Αφού, λοιπόν, η φ είναι άνω και κάτω ημισυνεχής στο C, είναι συνεχής στο C.

■

Πρόταση 2.10

Αν η φ είναι μια κανονική κυρτή συνάρτηση, τότε είναι τοπικά Lipschitz στο $Int\big(Dom(\varphi)\big)$.

Θυμίζουμε ότι η φ λέγεται τοπικά Lipschitz στο σύνολο S, αν για κάθε συμπαγές $K \subseteq S$ υπάρχει σταθερά C_K τέτοια ώστε $|\varphi(x) - \varphi(y)| \leq C_K|x - y|, \ \ x, y \in K$.

<u>Απόδειξη:</u>

Έστω, $K \subseteq Int\big(Dom(\varphi)\big)$ συμπαγές. Με B συμβολίζουμε την κλειστή μοναδιαία μπάλα στον $\mathbb{R}^n$. Για κάθε $\varepsilon > 0$, το σύνολο $K + \varepsilon B \coloneqq \{y \in \mathbb{R}^n : \quad d(y, K) \leq \varepsilon\}$ είναι επίσης συμπαγές. Θέτουμε

$$S_\varepsilon \coloneqq (K + \varepsilon B) \cap \Big(\mathbb{R}^n - Int\big(Dom(\varphi)\big)\Big), \qquad \varepsilon > 0.$$

Αν υποθέσουμε ότι

$$\bigcap_{\varepsilon > 0} S_\varepsilon \neq \emptyset,$$

τότε θα υπάρχει $x \in \mathbb{R}^n$ τέτοιο ώστε $x \in K + \varepsilon B$, για κάθε $\varepsilon > 0$ και $x \notin Int\big(Dom(\varphi)\big)$. Δηλαδή $x \in K$ και $x \notin Int\big(Dom(\varphi)\big)$, το οποίο αντιφάσκει με την επιλογή του K ως υποσυνόλου του $Int\big(Dom(\varphi)\big)$. Άρα,

$$\bigcap_{\varepsilon>0} S_\varepsilon = \emptyset,$$

και επομένως υπάρχει $\varepsilon_1 > 0$ τέτοιο ώστε $(K + \varepsilon_1 B) \cap \left(\mathbb{R}^n - Int\big(Dom(\varphi)\big)\right) = \emptyset$, δηλαδή

$$K + \varepsilon_1 B \subseteq Int\big(Dom(\varphi)\big).$$

Εφόσον η φ είναι συνεχής στο $K + \varepsilon_1 B$ και το σύνολο αυτό είναι συμπαγές, η φ είναι φραγμένη στο $K + \varepsilon_1 B$. Έστω $a_1, a_2 \in \mathbb{R}$ τέτοια ώστε $a_1 \leq \varphi(x) \leq a_2,\;\; x \in\; K + \varepsilon_1 B$.

Έστω $x, y \in K$ με $x \neq y$. Θεωρούμε το σημείο

$$z = y + \frac{\varepsilon_1}{|y - x|}(y - x).$$

Τότε $z \in K + \varepsilon_1 B$ και $y = \lambda z + (1 - \lambda)x$, όπου

$$\lambda \coloneqq \frac{|y - x|}{\varepsilon_1 + |y - x|} > 0.$$

Λόγω της κυρτότητας της φ, έχουμε $\varphi(y) \leq \lambda\varphi(z) + (1 - \lambda)\varphi(x) = \varphi(x) + \lambda\big(\varphi(z) - \varphi(x)\big)$.

Συνεπώς, $\varphi(y) - \varphi(x) = \lambda\big(\varphi(z) - \varphi(x)\big) \leq \lambda(a_2 - a_1) \leq a|y - x|$, όπου

$$a = \frac{a_2 - a_1}{\varepsilon_1}.$$

Η παραπάνω ανισότητα ισχύει για κάθε $x, y \in K$, άρα η φ είναι Lipschitz στο K με $C_K = a$.

■

Πρόταση 2.11

Το σύνορο ενός κυρτού συνόλου στον $\mathbb{R}^n$ έχει μέτρο Lebesgue ίσο με 0, δηλαδή αν $K \subseteq \mathbb{R}^n$ κυρτό, τότε $m(\partial K) = 0$.

Απόδειξη:

Η απόδειξη θα γίνει με επαγωγή ως προς το n. Για $n = 1$, το ∂K περιέχει το πολύ δύο στοιχεία.

Έστω ότι το συμπέρασμα ισχύει στον $\mathbb{R}^n$. Θεωρούμε $K \subseteq \mathbb{R}^{n+1}$ κυρτό. Έστω η προβολή $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ με $\pi(x_1, \dots, x_n, x_{n+1}) = (x_1, \dots, x_n)$ και $C = \pi(K)$. Το C είναι κυρτό. Θεωρούμε τις συναρτήσεις

$$f : C \to \mathbb{R} \cup \{-\infty\}, \quad f(x) = \inf\{t : (x, t) \in K\}$$

$$g : C \to \mathbb{R} \cup \{+\infty\}, \quad f(x) = \sup\{t : (x, t) \in K\}.$$

Λόγω της κυρτότητας του K, η f είναι κυρτή και η g κοίλη. Αν η f δεν είναι κανονική, τότε σύμφωνα με το λήμμα 2.4, $f(x) = -\infty$ για κάθε $x \in Int(C)$.

Άρα,

$$f(x) > -\infty \text{ για κάθε } x \in Int(C) \quad \text{ή} \quad f(x) = -\infty \text{ για κάθε } x \in Int(C).$$

Ομοίως,

$$g(x) < +\infty \text{ για κάθε } x \in Int(C) \quad \text{ή} \quad g(x) = +\infty \text{ για κάθε } x \in Int(C).$$

Σύμφωνα με την πρόταση 2.9, οι f, g είναι συνεχείς στο $Int(C)$. Άρα, αν $x \in Int(C)$ με $f(x) < t < g(x)$, τότε $(x, t) \in Int(K)$. Άρα,

$$\partial K \subseteq (\partial C \times \mathbb{R}) \cup \{(x, f(x)) : x \in Int(C), f(x) > -\infty\}$$
$$\cup \{(x, g(x)) : x \in Int(C), g(x) < +\infty\}.$$

Από την επαγωγική υπόθεση αλλά και λόγω της συνέχειας (άρα μετρησιμότητας) των f, g στο $Int(C)$, και τα τρία σύνολα στο δεξιό μέλος έχουν μέτρο Lebesgue ίσο με 0.

■

2. Υποδιαφορισιμότητα.

Το υποδιαφορικό $\partial\varphi$ μιας κυρτής συνάρτησης φ είναι μια απεικόνιση, η οποία σε κάθε $x \in \mathbb{R}^n$ αντιστοιχεί ένα σύνολο στον $\mathbb{R}^n$ ως εξής:

$$y \in \partial\varphi(x) \Leftrightarrow [\,\forall z \in \mathbb{R}^n, \quad \varphi(z) \geq \varphi(x) + y \cdot (z - x)] \tag{2.15}$$

Τα στοιχεία του $\partial\varphi(x)$ λέγονται υποπαράγωγοι της φ στο σημείο x.

Στη συνέχεια, θα ταυτίζουμε την απεικόνιση του υποδιαφορικού με το γράφημά της, το οποίο θα συμβολίζουμε με $Graph(\partial\varphi)$. Θεωρούμε, δηλαδή, το υποδιαφορικό ως ένα υποσύνολο του $\mathbb{R}^n \times \mathbb{R}^n$ για το οποίο ισχύει: $(x, y) \in Graph(\partial\varphi) \Leftrightarrow y \in \partial\varphi(x)$.

Ορίζουμε $Dom(\partial\varphi) \coloneqq \{x \in \mathbb{R}^n : \quad \partial\varphi(x) \neq \emptyset\}$.

Πρόταση 2.12

Έστω φ μια κανονική κυρτή συνάρτηση. Για κάθε $x \in Int(Dom(\varphi))$, $\partial\varphi(x) \neq \emptyset$ ενώ για κάθε $x \notin Dom(\varphi)$, $\partial\varphi(x) = \emptyset$. Με άλλα λόγια,

$$Int(Dom(\varphi)) \subseteq Dom(\partial\varphi) \subseteq Dom(\varphi).$$

Απόδειξη:

Έστω $x_0 \notin Dom(\varphi)$. Υποθέτουμε ότι υπάρχει $x^* \in \partial\varphi(x_0)$. Τότε, εξ ορισμού

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x_0) + x^* \cdot (z - x).$$

Παίρνοντας το z να είναι στοιχείο του $Dom(\varphi)$, βλέπουμε ότι η παραπάνω ανισότητα δεν μπορεί να ισχύει. Άρα, $\partial\varphi(x_0) = \emptyset$.

Έστω, τώρα, $x_0 \in Int\big(Dom(\varphi)\big)$. Από το λήμμα 2.3, για κάθε $\mu \in \mathbb{R}$ με $\mu > \varphi(x_0)$ ισχύει $(x_0, \mu) \in Int(epi\varphi)$, ενώ το σημείο $\big(x_0, \varphi(x_0)\big)$ ανήκει στο σύνορο $\partial(epi\varphi)$.

Θα εφαρμόσουμε την πρώτη γεωμετρική μορφή του Θεωρήματος Hahn-Banach για τα κυρτά σύνολα $A = Int(epi\varphi)$ και $B = \big\{\big(x_0, \varphi(x_0)\big)\big\}$. Προφανώς, το A είναι ανοιχτό και $A \cap B = \emptyset$.

Σύμφωνα, λοιπόν, με το Θεώρημα Hahn-Banach, υπάρχει κλειστό υπερεπίπεδο H στον $\mathbb{R}^{n+1}$ και συνεχές γραμμικό συναρτησοειδές Φ στον $\mathbb{R}^{n+1}$ τέτοιο ώστε $H = [\Phi = a]$, για κάποιο $a \in \mathbb{R}$ και το H διαχωρίζει τα A, B υπό την εξής έννοια: $\Phi(x, \mu) \geq \alpha$, $(x, \mu) \in Int(epi\varphi)$ και $\Phi\big(x_0, \varphi(x_0)\big) \leq \alpha$.

Λόγω συνέχειας του Φ, η ανισότητα $\Phi(x, \mu) \geq \alpha$ ισχύει και για κάθε $(x, \mu) \in \overline{Int(epi\varphi)} = \overline{epi\varphi}$, όπου η ισότητα ισχύει λόγω της κυρτότητας του συνόλου $epi\varphi$ (λήμμα 1.8). Εφόσον $\big(x_0, \varphi(x_0)\big) \in \overline{epi\varphi}$, προκύπτει, τελικά, ότι $\Phi\big(x_0, \varphi(x_0)\big) = \alpha$.

Υπάρχουν $y \in \mathbb{R}^n$ και $k \in \mathbb{R}$ με $(y, k) \neq (0,0) \in \mathbb{R}^n \times \mathbb{R}$ τέτοια ώστε το συναρτησοειδές Φ να γράφεται στη μορφή $\Phi(x, \mu) = y \cdot x + k\mu$, $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}$.

Τότε, $a = \Phi\big(x_0, \varphi(x_0)\big) = y \cdot x_0 + k\varphi(x_0)$.

Αφού $\Phi \geq a$ στο $\overline{A}$, έχουμε

$$y \cdot x + k\mu \geq a, \qquad (x, \mu) \in \overline{Int(epi\varphi)}. \tag{2.16}$$

Αφού η ανισότητα (2.16) ισχύει για κάθε $\mu > \varphi(x)$, αφήνοντας το $\mu \to +\infty$, προκύπτει ότι $k \geq 0$. Θα αποκλείσουμε την περίπτωση $k = 0$.

Πράγματι, αν υποθέσουμε ότι $k = 0$, τότε αναγκαστικά $y \neq 0$.

Ακόμα, $a = y \cdot x_0 + k\varphi(x_0) = y \cdot x_0$ και σύμφωνα με την ανισότητα (2.16) έχουμε

$$y \cdot (x - x_0) \geq 0, \quad x \in Dom(\varphi).$$

Εφόσον $x_0 \in Int\big(Dom(\varphi)\big)$, μπορούμε να βρούμε $\delta > 0$ τέτοιο ώστε

$$x_1 = x_0 - \frac{\delta}{\|y\|} y \ \in Dom(\varphi).$$

Θα έχουμε, λοιπόν:

$$0 \leq y \cdot (x_1 - x_0) = y \cdot \left(-\frac{\delta}{\|y\|} y\right) = -\delta\|y\|.$$

Συνεπάγεται ότι $\|y\| = 0$, δηλαδή $y = 0$. Καταλήξαμε σε αντίφαση, άρα $k \neq 0$.

Ειδικότερα, η ανισότητα (2.16) ισχύει για $\mu = \varphi(x)$, δηλαδή

$$y \cdot x + k\varphi(x) \geq a = y \cdot x_0 + k\varphi(x_0), \qquad x \in Dom(\varphi).$$

Διαιρώντας την τελευταία ανισότητα με $k > 0$, παίρνουμε

$$\varphi(x) \geq \varphi(x_0) - \frac{1}{k} y \cdot (x - x_0), \qquad x \in Dom(\varphi).$$

Εφόσον $\varphi(x) = +\infty$ για κάθε $x \in \mathbb{R}^n - Dom(\varphi)$, η παραπάνω ανισότητα ισχύει για κάθε $x \in \mathbb{R}^n$. Αυτό σημαίνει ότι

$$-\frac{1}{k} y \in \partial\varphi(x_0)$$

και άρα το $\partial\varphi(x_0)$ είναι μη κενό.

■

3. Διαφορισιμότητα.

Ορισμός 2.13

Μια συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι διαφορίσιμη στο σημείο $x \in \mathbb{R}^n$ αν $\varphi(x) \in \mathbb{R}$ και υπάρχει διάνυσμα, το οποίο συμβολίζουμε με $\nabla\varphi(x)$, τέτοιο ώστε

$$\lim_{y \to x} \frac{|\varphi(y) - \varphi(x) - \nabla\varphi(x) \cdot (x - y)|}{|x - y|} = 0.$$

Αν υπάρχει τέτοιο διάνυσμα, τότε είναι μοναδικό και ορίζουμε να είναι η παράγωγος της φ στο σημείο $x \in \mathbb{R}^n$, την οποία συμβολίζουμε με $D\varphi(x)$.

Συνέπεια της παρακάτω πρότασης είναι ότι μια κανονική κυρτή συνάρτηση είναι διαφορίσιμη σχεδόν παντού στο $Int\big(Dom(\varphi)\big)$, αναφορικά με το μέτρο Lebesgue, το οποίο συμβολίζουμε $\mathcal{L}^n$.

Πρόταση 2.14 (Θεώρημα Rademacher)

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση τοπικά Lipschitz. Τότε, η φ είναι διαφορίσιμη $\mathcal{L}^n$ −σχεδόν παντού στον $\mathbb{R}^n$.

Απόδειξη:

Εφόσον η διαφορισιμότητα είναι τοπική έννοια, μπορούμε να υποθέσουμε ότι η φ είναι Lipschitz και συμβολίζουμε με $Lip(\varphi)$ τη σταθερά Lipschitz.

Σταθεροποιούμε ένα $v \in \mathbb{R}^n$ με $|v| = 1$, και ορίζουμε

$$D_v\varphi(x) \coloneqq \lim_{t\to 0}\frac{\varphi(x+tv)-\varphi(x)}{t}, \qquad x \in \mathbb{R}^n$$

υπό την προϋπόθεση ότι το όριο υπάρχει.

Ισχυρισμός 1. Το $D_v\varphi(x)$ υπάρχει $\mathcal{L}^n$ −σχεδόν παντού στον $\mathbb{R}^n$.

Απόδειξη του Ισχυρισμού 1:

Εφόσον η φ είναι τοπικά Lipschitz, είναι και συνεχής, οπότε οι συναρτήσεις

$$\overline{D_v}\varphi(x) \coloneqq \limsup_{t\to 0}\frac{\varphi(x+tv)-\varphi(x)}{t} = \lim_{k\to\infty}\sup_{\substack{0<|t|<\frac{1}{k}\\ t\ \rho\eta\tau\acute{o}\varsigma}}\frac{\varphi(x+tv)-\varphi(x)}{t}$$

και

$$\underline{D_v}\varphi(x) \coloneqq \liminf_{t\to 0}\frac{\varphi(x+tv)-\varphi(x)}{t} = \lim_{k\to\infty}\inf_{\substack{0<|t|<\frac{1}{k}\\ t\ \rho\eta\tau\acute{o}\varsigma}}\frac{\varphi(x+tv)-\varphi(x)}{t}$$

είναι Borel μετρήσιμες.

Άρα και το σύνολο

$$A_v \coloneqq \{x \in \mathbb{R}^n : \quad D_v\varphi(x) \ \text{δεν υπάρχει}\} = \left\{x \in \mathbb{R}^n : \quad \underline{D_v}\varphi(x) < \overline{D_v}\varphi(x)\right\}$$

είναι Borel μετρήσιμο.

Για κάθε $x, v \in \mathbb{R}^n$ με $|v| = 1$, ορίζουμε τη συνάρτηση $f_x : \mathbb{R} \to \mathbb{R}$ με

$$f_x(t) \coloneqq \varphi(x+tv), \qquad t \in \mathbb{R}.$$

Τότε, για κάθε $t_1, t_2 \in \mathbb{R}$ και εφόσον η φ έχει υποτεθεί Lipschitz, έχουμε

$$|f_x(t_1) - f_x(t_2)| = |\varphi(x+t_1v) - \varphi(x+t_2v)| \leq Lip(\varphi)|(t_1-t_2)v| = Lip(\varphi)|t_1-t_2|,$$

αφού $|v| = 1$. Άρα, η f_x είναι Lipschitz στο $\mathbb{R}$.

Έστω $\varepsilon > 0$. Επιλέγουμε θετικό αριθμό $\delta < \frac{\varepsilon}{Lip(\varphi)}$ και ακολουθία ξένων ανά δύο διαστημάτων (a_k, b_k), $k = 1,2,\dots,N$ τέτοια ώστε

$$\sum_{k=1}^{N} |b_k - a_k| < \delta.$$

Τότε

$$\sum_{k=1}^{N} |\varphi(b_k) - \varphi(a_k)| \leq Lip(\varphi) \sum_{k=1}^{N} |b_k - a_k| < Lip(\varphi)\delta < Lip(\varphi) \frac{\varepsilon}{Lip(\varphi)} = \varepsilon.$$

Επομένως η f_x είναι απόλυτα συνεχής και άρα διαφορίσιμη $\mathcal{L}^1$ −σχεδόν παντού στο $\mathbb{R}$. Δηλαδή, αν $A_x = \{t \in \mathbb{R} : \ f_x \text{ όχι διαφορίσιμη στο } t\}$, τότε $\mathcal{L}^1(A_x) = 0$.

Έστω $E = v^{\perp}$. Τότε ο E μπορεί να ταυτιστεί με τον $\mathbb{R}^{n-1}$ και ο $\mathbb{R}^n$ με τον $E \times \mathbb{R}$. Ισχύει

$$\begin{aligned} A_v &= \{(x,t) \in E \times \mathbb{R} : \ D_v\varphi \text{ δεν υπάρχει στο } (x,t)\} = \{(x,t) \in E \times \mathbb{R} : \ f_x \text{ όχι διαφορίσιμη στο } t\} \\ &= \{(x,t) \in E \times \mathbb{R} : \ t \in A_x\}. \end{aligned}$$

Άρα,

$$\mathcal{L}^n(A_v) = \int_E \mathcal{L}^1(A_x)\, d\mathcal{L}^{n-1}(x) = 0$$

επειδή το A_v είναι μετρήσιμο.

■

Ως συνέπεια του Ισχυρισμού 1, βλέπουμε ότι το

$$\nabla\varphi(x) \coloneqq \left(\frac{\partial\varphi}{\partial x_1}(x), \dots, \frac{\partial\varphi}{\partial x_n}(x) \right)$$

υπάρχει $\mathcal{L}^n$ −σχεδόν παντού στον $\mathbb{R}^n$.

Ισχυρισμός 2. $D_v\varphi(x) = v \cdot \nabla\varphi(x), \quad \mathcal{L}^n$ −σχεδόν παντού στον $\mathbb{R}^n$.

Απόδειξη του Ισχυρισμού 2:

Επειδή η ιδιότητα που θέλουμε να αποδείξουμε είναι τοπική, μπορούμε, χωρίς βλάβη της γενικότητας, να υποθέσουμε ότι η φ έχει συμπαγή φορέα.

Έστω μια συνάρτηση $\zeta \in C_c^{\infty}(\mathbb{R}^n)$, δηλαδή άπειρες φορές συνεχώς παραγωγισιμη με συμπαγή φορέα. Τότε, ισχύει:

$$\int_{\mathbb{R}^n} \varphi(x + tv)\zeta(x)\, dx = \int_{\mathbb{R}^n} \varphi(x)\zeta(x - tv)\, dx$$

το οποίο φαίνεται ξεκάθαρα αν κάνουμε την αλλαγή μεταβλητής $y = x + tv$ στο πρώτο ολοκλήρωμα. Αφαιρώντας, στη συνέχεια, και από τα δύο μέλη της ισότητας το

$$\int_{\mathbb{R}^n} \varphi(x)\zeta(x)\,dx$$

και διαιρώντας με το t, παίρνουμε τη σχέση:

$$\int_{\mathbb{R}^n} \left[\frac{\varphi(x+tv)-\varphi(x)}{t}\right]\zeta(x)\,dx = -\int_{\mathbb{R}^n} \varphi(x)\left[\frac{\zeta(x)-\zeta(x-tv)}{t}\right]dx. \tag{2.17}$$

Θέτουμε $t = \frac{1}{k},\;\; k \in \mathbb{N}$ στην παραπάνω ισότητα και παρατηρούμε ότι

$$\left|\frac{\varphi\left(x+\frac{1}{k}v\right)-\varphi(x)}{\frac{1}{k}}\right| = k\left|\varphi\left(x+\frac{1}{k}v\right)-\varphi(x)\right| \le kLip(\varphi)\left|\frac{1}{k}v\right| = Lip(\varphi),$$

αφού $|v| = 1$. Ομοίως, εφόσον η $\zeta \in C_c^\infty(\mathbb{R}^n)$ είναι και αυτή Lipschitz, άρα

$$\left|\frac{\zeta(x)-\zeta\left(x-\frac{1}{k}v\right)}{\frac{1}{k}}\right| = k\left|\zeta(x)-\zeta\left(x-\frac{1}{k}v\right)\right| \le Lip(\zeta).$$

Εφόσον $\zeta \in C_c^\infty(\mathbb{R}^n)$, υπάρχει $M_1 \in \mathbb{R}$ ώστε $|\zeta(x)| \le M_1,\;\; x \in \mathbb{R}^n$, και υπάρχει συμπαγής σφαίρα $K \subseteq \mathbb{R}^n$ώστε $\zeta(x) = 0,\;\; x \notin K$. Παίρνοντας τη συμπαγή σφαίρα $B = K + B_1(0)$, όπου $B_1(0)$ η μοναδιαία σφαίρα με κέντρο το 0, πετυχαίνουμε ότι για κάθε $k \in \mathbb{N}$ ισχύει $\zeta\left(x-\frac{1}{k}v\right) = 0,\;\; x \notin B$. Αφού η συνάρτηση φ είναι Lipschitz, είναι και συνεχής, άρα υπάρχει $M_2 \in \mathbb{R}$ ώστε $|\varphi(x)| \le M_2,\;\; x \in B$. Επομένως, για κάθε $k \in \mathbb{N}$, έχουμε:

$$\left|\left[\frac{\varphi\left(x+\frac{1}{k}v\right)-\varphi(x)}{\frac{1}{k}}\right]\zeta(x)\right| \le f(x) = \begin{cases} Lip(\varphi)M_1, & x \in B \\ 0, & x \notin B \end{cases}$$

και

$$\left|\varphi(x)\left[\frac{\zeta(x)-\zeta\left(x-\frac{1}{k}v\right)}{\frac{1}{k}}\right]\right| \le g(x) = \begin{cases} M_2 Lip(\zeta), & x \in B \\ 0, & x \notin B \end{cases}$$

και οι συναρτήσεις f και g είναι ολοκληρώσιμες.

Άρα, χρησιμοποιώντας το Θεώρημα Κυριαρχημένης Σύγκλισης, προκύπτει ότι

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \left[\frac{\varphi\left(x+\frac{1}{k}v\right)-\varphi(x)}{\frac{1}{k}}\right]\zeta(x)\,dx = \int_{\mathbb{R}^n} D_v\varphi(x)\zeta(x)\,dx$$

και

$$\lim_{k\to\infty}\int_{\mathbb{R}^n}\varphi(x)\left[\frac{\zeta(x)-\zeta(x-tv)}{t}\right]dx=\int_{\mathbb{R}^n}\varphi(x)D_v\zeta(x)\,dx.$$

Χρησιμοποιώντας τη σχέση (2.17), το Θεώρημα Fubini και την απόλυτη συνέχεια της φ πάνω σε ευθείες, έχουμε:

$$\int_{\mathbb{R}^n}D_v\varphi(x)\zeta(x)\,dx=-\int_{\mathbb{R}^n}\varphi(x)D_v\zeta(x)\,dx=-\sum_{i=1}^{n}v_i\int_{\mathbb{R}^n}\varphi(x)\frac{\partial\zeta}{\partial x_i}(x)\,dx$$
$$=\sum_{i=1}^{n}v_i\int_{\mathbb{R}^n}\frac{\partial\varphi}{\partial x_i}(x)\zeta(x)\,dx=\int_{\mathbb{R}^n}\big(v\cdot\nabla\varphi(x)\big)\zeta(x)\,dx.$$

Θα εξηγήσουμε την τρίτη ισότητα δείχνοντας ότι

$$\int_{\mathbb{R}^n}\varphi(x)\frac{\partial\zeta}{\partial x_1}(x)\,dx=-\int_{\mathbb{R}^n}\frac{\partial\varphi}{\partial x_1}(x)\zeta(x)\,dx.$$

Πράγματι, αρκεί να δείξουμε ότι

$$\int_{\mathbb{R}^n}\left[\varphi(x)\frac{\partial\zeta}{\partial x_1}(x)+\frac{\partial\varphi}{\partial x_1}(x)\zeta(x)\right]dx=0.$$

Έχουμε, λοιπόν

$$\int_{\mathbb{R}^n}\left[\varphi(x)\frac{\partial\zeta}{\partial x_1}(x)+\frac{\partial\varphi}{\partial x_1}(x)\zeta(x)\right]dx$$
$$=\int_{\mathbb{R}^{n-1}}\int_{-\infty}^{+\infty}\left[\varphi(x_1,\dots,x_n)\frac{\partial\zeta}{\partial x_1}(x_1,\dots,x_n)+\frac{\partial\varphi}{\partial x_1}(x_1,\dots,x_n)\zeta(x_1,\dots,x_n)\right]dx_1\dots dx_n$$
$$=\int_{\mathbb{R}^{n-1}}\int_{-\infty}^{+\infty}\frac{\partial}{\partial x_1}\big(\varphi(x)\zeta(x)\big)\,dx_1\dots dx_n=0$$

γιατί η συνάρτηση ζ έχει συμπαγή φορέα.

Επιστρέφουμε στην ισότητα

$$\int_{\mathbb{R}^n}D_v\varphi(x)\zeta(x)\,dx=\int_{\mathbb{R}^n}\big(v\cdot\nabla\varphi(x)\big)\zeta(x)\,dx$$

η οποία μπορεί να γραφεί ως εξής

$$\int_{\mathbb{R}^n}\big(D_v\varphi(x)-v\cdot\nabla\varphi(x)\big)\zeta(x)\,dx=0$$

και ισχύει για κάθε συνάρτηση $\zeta\in C_c^\infty(\mathbb{R}^n)$. Μπορούμε να βρούμε ακολουθία συναρτήσεων $\zeta_n\in C_c^\infty(\mathbb{R}^n)$ τέτοια ώστε

$$\zeta_n\longrightarrow\big(D_v\varphi(x)-v\cdot\nabla\varphi(x)\big),\quad \text{ως προς την } L^2-\text{νόρμα}.$$

Τότε, από το Θεώρημα Κυριαρχημένης Σύγκλισης έχουμε:

$$0 = \lim_{n\to\infty} \int_{\mathbb{R}^n} \big(D_v\varphi(x) - v \cdot \nabla\varphi(x)\big)\zeta_n(x)\, dx = \int_{\mathbb{R}^n} \big(D_v\varphi(x) - v \cdot \nabla\varphi(x)\big)^2\, dx.$$

Προκύπτει ότι

$$D_v\, \varphi(x) = v \cdot \nabla\varphi(x), \qquad \mathcal{L}^n - \text{σχεδόν παντού στον } \mathbb{R}^n.$$

■

Επιλέγουμε ένα αριθμήσιμο πυκνό υποσύνολο $\{v_k\}_{k=1}^{\infty}$ του $\partial B_1(0)$, όπου $\partial B_1(0)$ το σύνορο της μπάλας κέντρου 0 και ακτίνας 1.

Για $k \in \mathbb{N}$, θέτουμε

$$A_k \coloneqq \left\{x \in \mathbb{R}^n : \quad D_{v_k}\varphi(x), \nabla\varphi(x) \text{ υπάρχουν και } D_{v_k}\varphi(x) = v_k \cdot \nabla\varphi(x)\right\}$$

και ορίζουμε

$$A \coloneqq \bigcap_{k=1}^{\infty} A_k.$$

Από τους ισχυρισμούς 1 και 2, βλέπουμε ότι $\mathcal{L}^n(\mathbb{R}^n - A) = 0$.

Ισχυρισμός 3. Η φ είναι διαφορίσιμη σε κάθε σημείο $x \in A$.

<u>Απόδειξη του Ισχυρισμού 3:</u>

Σταθεροποιούμε ένα $x \in A$. Επιλέγουμε $v \in \partial B_1(0)$, $t \in \mathbb{R} - \{0\}$ και ορίζουμε

$$Q(x, v, t) \coloneqq \frac{\varphi(x + tv) - \varphi(x)}{t} - v \cdot \nabla\varphi(x).$$

Αν $v' \in \partial B_1(0)$, έχουμε

$$\begin{aligned} |Q(x, v, t) - Q(x, v', t)| &\le \left|\frac{\varphi(x + tv) - \varphi(x + tv')}{t}\right| + |(v - v') \cdot \nabla\varphi(x)| \\ &\le \frac{1}{|t|} Lip(\varphi)|t(v - v')| + |\nabla\varphi(x)||v - v'| = (Lip(\varphi) + |\nabla\varphi(x)|)|v - v'| \\ &\le \big(\sqrt{n} + 1\big) Lip(\varphi)|v - v'|. \end{aligned}$$

Έστω $\varepsilon > 0$. Επιλέγουμε N αρκούντως μεγάλο ώστε αν $v' \in \partial B_1(0)$, τότε

$$|v' - v_k| \le \frac{\varepsilon}{2\big(\sqrt{n} + 1\big) Lip(\varphi)} \tag{2.18}$$

για κάποιο $k \in \{1, 2, \dots, N\}$. Σημειώνουμε ότι υπάρχει τέτοιο N λόγω της πυκνότητας του συνόλου $\{v_k\}_{k=1}^{\infty}$ μέσα στο συμπαγές σύνολο $\partial B_1(0)$.

Τώρα, για κάθε $k \in \{1, 2, \dots, N\}$, έχουμε

$$\lim_{t \to 0} Q(x, v_k, t) = \lim_{t \to 0} \left[\frac{\varphi(x + tv_k) - \varphi(x)}{t}\right] - v_k \cdot \nabla\varphi(x) = D_{v_k}\varphi(x) - v_k \cdot \nabla\varphi(x) = 0, \qquad \text{εφόσον } x \in A.$$

Άρα, υπάρχει $\delta > 0$ τέτοιο ώστε για κάθε $k \in \{1, 2, \dots, N\}$ να έχουμε

$$|Q(x, v_k, t)| < \frac{\varepsilon}{2}, \qquad 0 < |t| < \delta. \tag{2.19}$$

Συνεπώς, σύμφωνα με τις σχέσεις (2.18) και (2.19), για κάθε $v \in \partial B_1(0)$, υπάρχει $k \in \{1, 2, \dots, N\}$ τέτοιο ώστε

$$\begin{aligned}|Q(x, v, t)| &\le |Q(x, v_k, t)| + |Q(x, v, t) - Q(x, v_k, t)| \le |Q(x, v_k, t)| + \left(\sqrt{n} + 1\right) Lip(\varphi)|v - v_k| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\end{aligned}$$

για $0 < |t| < \delta$. Παρατηρούμε ότι το ίδιο $\delta > 0$ δουλεύει για όλα τα $v \in \partial B_1(0)$. Άρα, $\lim_{t \to 0} Q(x, v, t) = 0$ ομοιόμορφα για όλα τα $v \in \partial B_1(0)$.

Τέλος, επιλέγουμε ένα οποιοδήποτε $y \in \mathbb{R}^n - \{x\}$. Παίρνουμε το μοναδιαίο διάνυσμα

$$v_y \coloneqq \frac{y - x}{|y - x|}$$

έτσι ώστε $y = x + tv_y$, όπου $t \coloneqq |y - x|$. Τότε,

$$\begin{aligned}\lim_{y \to x} \frac{|\varphi(y) - \varphi(x) - \nabla\varphi(x) \cdot (y - x)|}{|y - x|} &= \lim_{t \to 0} \frac{\left|\varphi\left(x + tv_y\right) - \varphi(x) - tv_y \cdot \nabla\varphi(x)\right|}{t} \\ &= \lim_{t \to 0} \frac{\left|tQ\left(x, v_y, t\right)\right|}{t} = \lim_{t \to 0} Q\left(x, v_y, t\right) = 0,\end{aligned}$$

λόγω της ομοιόμορφης σύγκλισης. Άρα, η φ είναι διαφορίσιμη στο $x \in A$ με $D\varphi(x) = \nabla\varphi(x)$.

■

Εφόσον $\mathcal{L}^n(\mathbb{R}^n - A) = 0$, η φ είναι διαφορίσιμη $\mathcal{L}^n$ −σχεδόν παντού στον $\mathbb{R}^n$ και η απόδειξη της πρότασης είναι πλήρης.

■

Ορισμός 2.15

Έστω $C \subseteq \mathbb{R}^n$ κυρτό. Η συνάρτηση $\delta^*(x^* \mid C) \coloneqq \sup_{x \in C}[x^* \cdot x]$, $x^* \in \mathbb{R}^n$ ονομάζεται συνάρτηση στήριξης του C.

Ορισμός 2.16

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ και $x \in Dom(\varphi)$. Ορίζουμε τη μονόπλευρη κατά κατεύθυνση παράγωγο της φ στο x αναφορικά με το διάνυσμα $v \in \mathbb{R}^n$, να είναι το όριο:

$$\varphi'(x\,;v) \coloneqq \lim_{\lambda \downarrow 0} \frac{\varphi(x+\lambda v) - \varphi(x)}{\lambda}. \tag{2.20}$$

αν, βέβαια, αυτό υπάρχει.

Πρόταση 2.17

Έστω φ μια κυρτή συνάρτηση και $x \in Dom(\varphi)$. Τότε, για κάθε $v \in \mathbb{R}^n$ υπάρχει το $\varphi'(x\,;v) \in \mathbb{R} \cup \{+\infty\}$ και

$$x^* \in \partial\varphi(x) \iff \forall v \in \mathbb{R}^n, \qquad \varphi'(x\,;v) \geq x^* \cdot v\,. \tag{2.21}$$

Στην πραγματικότητα, η κλειστότητα της $\varphi'(x\,;v)$, θεωρούμενη ως μια κυρτή συνάρτηση του $v \in \mathbb{R}^n$, είναι η συνάρτηση στήριξης του κλειστού κυρτού συνόλου $\partial\varphi(x)$.

<u>Απόδειξη:</u>

Εξ ορισμού ισχύει

$$x^* \in \partial\varphi(x) \iff \forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x) + x^* \cdot (z-x)\,.$$

Θέτοντας $z = x + \lambda v$, $v \in \mathbb{R}^n$, $\lambda > 0$, η παραπάνω ανισότητα γίνεται

$$\frac{\varphi(x+\lambda v) - \varphi(x)}{\lambda} \geq x^* \cdot v, \qquad v \in \mathbb{R}^n, \lambda > 0.$$

Σύμφωνα με την παρακάτω παρατήρηση 2, το πηλίκο

$$\frac{\varphi(x+\lambda v) - \varphi(x)}{\lambda}$$

φθίνει προς το $\varphi'(x\,;v)$ καθώς το $\lambda \downarrow 0$, και άρα η παραπάνω ανισότητα ισοδυναμεί με την

$$\varphi'(x\,;v) = \lim_{\lambda \downarrow 0} \frac{\varphi(x+\lambda v) - \varphi(x)}{\lambda} \geq x^* \cdot v, \qquad v \in \mathbb{R}^n.$$

Εύκολα μπορούμε να ελέγξουμε ότι η συνάρτηση $\varphi'(x\,;v)$, θεωρούμενη ως συνάρτηση του $v \in \mathbb{R}^n$, είναι κυρτή και θετικά ομογενής, δηλαδή ικανοποιεί τη σχέση

$$\varphi'(x\,;\kappa v) = \kappa\varphi'(x\,;v), \;\; \kappa > 0.$$

Επομένως, εφαρμόζοντας το πόρισμα που διατυπώνουμε στην παρατήρηση 3 πιο κάτω, προκύπτει ότι η κλειστότητα της $\varphi'(x\,;v)$ είναι η συνάρτηση στήριξης του κλειστού κυρτού συνόλου $\{x^* \in \mathbb{R}^n : \;\; x^* \cdot v \le \varphi'(x\,;v), \;\; v \in \mathbb{R}^n\}$ δηλαδή, σύμφωνα με τα παραπάνω, του $\partial\varphi(x)$.

Σημειώνουμε ότι η $\varphi'(x\,;v)$ είναι κανονική, αφού $\varphi'(x\,;0) = 0$.

■

Παρατηρήσεις.

1. Μια συνάρτηση φ στον $\mathbb{R}^n$ λέγεται θετικά ομογενής αν για κάθε $x \in \mathbb{R}^n$ ισχύει
$$\varphi(\lambda x) = \lambda\varphi(x), \quad \lambda > 0.$$

2. **Θεώρημα 23.1 [13]**
Έστω $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ μια κυρτή συνάρτηση και $x \in \mathbb{R}^n$ τέτοιο ώστε $\varphi(x) < +\infty$. Τότε, για κάθε $y \in \mathbb{R}^n$, το πηλίκο
$$\frac{\varphi(x+\lambda y) - \varphi(x)}{\lambda}$$
είναι μια αύξουσα συνάρτηση ως προς $\lambda > 0$, δηλαδή το $\varphi'(x\,;y)$ υπάρχει και
$$\varphi'(x\,;y) = \inf_{\lambda>0} \frac{\varphi(x+\lambda y) - \varphi(x)}{\lambda}.$$
Επιπλέον, η $\varphi'(x\,;y)$ είναι μια θετικά ομογενής κυρτή συνάρτηση του y, με $\varphi'(x\,;0) = 0$ και $-\varphi'(x\,;-y) \le \varphi'(x\,;y), \;\; \forall y \in \mathbb{R}^n$.

3. **Πόρισμα 13.2.1 [13]**
Έστω $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ μια θετικά ομογενής κυρτή συνάρτηση, $\varphi \not\equiv +\infty$. Τότε, η $cl\varphi$ είναι η συνάρτηση στήριξης του κλειστού κυρτού συνόλου
$$C = \{x^* \in \mathbb{R}^n : \;\; x^* \cdot x \le \varphi(x), \;\; x \in \mathbb{R}^n\}$$

Πρόταση 2.18

Έστω φ μια κυρτή συνάρτηση στον $\mathbb{R}^n$ και $x \in Dom(\varphi)$. Αν η φ είναι διαφορίσιμη στο x, τότε το διάνυσμα $\nabla\varphi(x)$ είναι η μοναδική υποπαράγωγος της φ στο x, δηλαδή $\partial\varphi(x) = \{\nabla\varphi(x)\}$. Ειδικότερα, ισχύει η σχέση

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \ge \varphi(x) + \nabla\varphi(x)\cdot(z - x). \tag{2.22}$$

Αντίστροφα, αν η φ έχει μοναδική υποπαράγωγο στο x, τότε η φ είναι διαφορίσιμη στο x.

Απόδειξη:

Έστω ότι η φ είναι διαφορίσιμη στο x. Τότε, $\varphi'(x\,;v) = v \cdot \nabla\varphi(x), \;\; v \in \mathbb{R}^n$.

Από την Πρόταση 2.17, προκύπτει ότι $x^* \in \partial\varphi(x) \iff \forall v \in \mathbb{R}^n,\ \ v \cdot \nabla\varphi(x) \geq v \cdot x^*$.

Ισοδύναμα, $\forall v \in \mathbb{R}^n,\ \ v \cdot (\nabla\varphi(x) - x^*) \geq 0$. Παίρνοντας $v = x^* - \nabla\varphi(x)$, προκύπτει:

$$0 \leq \big(x^* - \nabla\varphi(x)\big) \cdot (\nabla\varphi(x) - x^*) = -\|x^* - \nabla\varphi(x)\|^2.$$

Η ανισότητα ικανοποιείται αν και μόνο αν $x^* = \nabla\varphi(x)$. Άρα, $\partial\varphi(x) = \{\nabla\varphi(x)\}$.

Αντίστροφα, υποθέτουμε ότι η φ έχει μοναδική υποπαράγωγο στο x το διάνυσμα x^*. Θα δείξουμε ότι η φ είναι διαφορίσιμη στο x, ακριβέστερα ότι

$$\lim_{z \to x} \frac{\varphi(z) - \varphi(x) - x^* \cdot (z - x)}{|z - x|} = 0.$$

Θέτοντας $y = z - x$, ισοδύναμα θα δείξουμε ότι

$$\lim_{y \to 0} \frac{\varphi(x + y) - \varphi(x) - x^* \cdot y}{|y|} = 0.$$

Ορίζουμε την κυρτή συνάρτηση

$$g(y) \coloneqq \varphi(x + y) - \varphi(x) - x^* \cdot y, \qquad y \in \mathbb{R}^n.$$

Ισχυρισμός 4. $\partial g(0) = \{0\}$.

<u>Απόδειξη του Ισχυρισμού 4:</u>

Έστω $y^* \in \partial g(0)$. Αυτό ισοδυναμεί με το ότι $\forall y \in \mathbb{R}^n,\ \ g(y) \geq g(0) + y^* \cdot (y - 0)$, δηλαδή

$$\forall y \in \mathbb{R}^n, \qquad \varphi(x + y) - \varphi(x) - x^* \cdot y \geq y^* \cdot y,$$

δηλαδή

$$\forall y \in \mathbb{R}^n, \qquad \varphi(x + y) \geq \varphi(x) + (x^* + y^*) \cdot y.$$

Θέτοντας $y = z - x,\ \ z \in \mathbb{R}^n$ παίρνουμε την ανισότητα

$$\forall z \in \mathbb{R}^n,\ \ \varphi(z) \geq \varphi(x) + (x^* + y^*) \cdot (z - x),$$

που εξ ορισμού σημαίνει ότι $x^* + y^* \in \partial\varphi(x)$. Εφόσον η φ έχει μοναδική υποπαράγωγο στο x το διάνυσμα x^*, έπεται ότι η παραπάνω ανισότητα ισχύει για κάθε $z \in \mathbb{R}^n$ αν και μόνο αν $y^* = 0$. Άρα $\partial g(0) = \{0\}$.

■

Ισχυρισμός 5.

$$\lim_{y \to 0} \frac{g(y)}{|y|} = 0.$$

<u>Απόδειξη του Ισχυρισμού 5:</u>

Από την Πρόταση 2.17, γνωρίζουμε ότι η κλειστότητα της $g'(0\,;v)$ είναι η συνάρτηση στήριξης του $\partial g(0)$, δηλαδή $clg'(0\,;v) = \delta^*\big(v \mid \partial g(0)\big) = 0,\ \ v \in \mathbb{R}^n$.

Από το λήμμα 2.8, έχουμε $g'(0\,;v) = clg'(0\,;v) = 0,\ \ v \in \mathbb{R}^n - \partial Dom\big(g'(0\,;v)\big)$.

Αν υποθέσουμε ότι υπάρχει $v_0 \in \mathbb{R}^n - \overline{Dom\big(g'(0\,;v)\big)}$, τότε

$$g'(0\,;v_0) = +\infty \quad \text{και} \quad g'(0\,;v_0) = clg'(0\,;v_0) = 0.$$

Άρα, $\mathbb{R}^n - \overline{Dom\big(g'(0\,;v)\big)} = \emptyset$, δηλαδή $\overline{Dom\big(g'(0\,;v)\big)} = \mathbb{R}^n$ και επειδή το σύνολο $Dom\big(g'(0\,;v)\big)$ είναι κυρτό, σύμφωνα με την παρακάτω παρατήρηση, έχουμε $Int\left(Dom\big(g'(0\,;v)\big)\right) = \mathbb{R}^n$.

Τελικά, $\partial Dom\big(g'(0\,;v)\big) = \emptyset$ και συνεπώς $g'(0\,;v) = 0$ για κάθε $v \in \mathbb{R}^n$.

Για κάθε $v \in \mathbb{R}^n$, έχουμε, λοιπόν:

$$0 = g'(0\,;v) = \lim_{\lambda \downarrow 0} \frac{g(0+\lambda v)}{\lambda} = \lim_{\lambda \downarrow 0} \frac{g(\lambda v)}{\lambda}$$

και το πηλίκο $\frac{g(\lambda v)}{\lambda}$ φθίνει καθώς $\lambda \downarrow 0$.

Άρα, οι συναρτήσεις

$$h_\lambda(v) \coloneqq \frac{g(\lambda v)}{\lambda}, \qquad \lambda > 0$$

είναι κυρτές και φθίνουν κατά σημείο στη σταθερή συνάρτηση 0, δηλαδή

$$\forall v \in \mathbb{R}^n, \qquad \lim_{\lambda \downarrow 0} h_\lambda(v) = 0.$$

Έστω $B \coloneqq \{v \in \mathbb{R}^n : \ \ |v| = 1\}$ η Ευκλείδεια μοναδιαία σφαίρα και έστω ακόμα $\{a_1, \dots, a_m\} \subseteq \mathbb{R}^n$ ένα οποιοδήποτε πεπερασμένο σύνολο τέτοιο ώστε $B \subseteq Conv(\{a_1, \dots, a_m\})$, όπου $Conv$ συμβολίζει την κυρτή θήκη. Ένα τέτοιο σύνολο θα μπορούσε να είναι το $\{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \in \mathbb{R}^n : \ \ \varepsilon_i = \pm 1\}$ που είναι ένα γενικευμένο «ορθογώνιο» στον $\mathbb{R}^n$.

Κάθε $v \in B$ μπορεί να εκφραστεί ως ένας κυρτός συνδυασμός

$$v = \lambda_1 a_1 + \cdots + \lambda_m a_m, \qquad \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1.$$

Τότε, έχουμε

$$0 \leq h_\lambda(v) \leq \sum_{i=1}^{m} \lambda_i\, h_\lambda(a_i) \leq \max_{1 \leq i \leq m} \{h_\lambda(a_i)\}.$$

Εφόσον για κάθε $i = 1,2, \dots, m$ ισχύει $\lim_{\lambda \downarrow 0} h_\lambda(a_i) = 0$, συμπεραίνουμε ότι $\lim_{\lambda \downarrow 0} h_\lambda(v) = 0$ ομοιόμορφα ως προς $v \in B$.

Έστω $\varepsilon > 0$. Τότε, υπάρχει $\delta > 0$ τέτοιο ώστε για κάθε $\lambda \in (0, \delta]$ και $v \in B$ να ισχύει

$$\frac{g(\lambda v)}{\lambda} \leq \varepsilon.$$

Εφόσον οποιοδήποτε διάνυσμα $y \in \mathbb{R}^n$ με $0 < |y| \leq \delta$ μπορούμε να το γράψουμε ως $y = \lambda v$, με $\lambda = |y|$ και $v \in B$, έχουμε τη συνεπαγωγή

$$0 < |y| \leq \delta \implies \frac{g(y)}{|y|} \leq \varepsilon.$$

Αυτό σημαίνει ότι

$$\lim_{y \to 0} \frac{g(y)}{|y|} = 0$$

όπως επιθυμούσαμε.

■

Παρατήρηση: Αν $C \subseteq \mathbb{R}^n$ κυρτό, τότε $Int(C) = Int(\bar{C})$.

Απόδειξη:

Ο εγκλεισμός $Int(C) \subseteq Int(\bar{C})$ είναι προφανής, οπότε για την απόδειξη του αντίστροφου εγκλεισμού, έστω $z \in Int(\bar{C})$ και $x \in Int(C)$. Υποθέτουμε ότι $x \neq z$, αλλιώς $z \in Int(C)$ τετριμμένα. Θεωρούμε την ευθεία που ορίζουν τα x και z. Για $\mu > 1$ με το $(\mu - 1)$ να είναι αρκετά μικρό, το σημείο $y = (1 - \mu)x + \mu z = z - (\mu - 1)(x - z)$ ανήκει στο $Int(\bar{C})$. Για ένα τέτοιο y, μπορούμε να εκφράσουμε το z στη μορφή $z = (1 - \lambda)x + \lambda y$ με $0 < \lambda = \frac{1}{\mu} < 1$. Από το λήμμα 2.2, έχουμε $z \in Int(C)$.

■

4. Η γραφική παράσταση βρίσκεται πάνω από την εφαπτομένη.

Όπως είδαμε στην Πρόταση 2.18, αν φ είναι μια κυρτή συνάρτηση, τότε για κάθε σημείο $x \in \mathbb{R}^n$ όπου η φ είναι διαφορίσιμη, ισχύει η σχέση:

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x) + \nabla\varphi(x) \cdot (z - x).$$

Η γεωμετρική σημασία της σχέσης αυτής είναι ότι η γραφική παράσταση της φ βρίσκεται πάνω από το εφαπτόμενο υπερεπίπεδό της στο σημείο x.

5. Μονοτονία.

Πρόταση 2.19

Για οποιαδήποτε κυρτή συνάρτηση $\varphi\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, η $\nabla\varphi$ είναι μονότονη συνάρτηση. Δηλαδή, αν η φ είναι διαφορίσιμη στα σημεία x_1 και x_2, τότε ισχύει $\left(\nabla\varphi(x_1) - \nabla\varphi(x_2)\right) \cdot (x_1 - x_2) \geq 0$.

Απόδειξη:

Αν η φ είναι διαφορίσιμη στα σημεία x_1 και x_2, τότε ισχύουν οι σχέσεις:

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x_1) + \nabla\varphi(x_1) \cdot (z - x_1)$$

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x_2) + \nabla\varphi(x_2) \cdot (z - x_2).$$

Θέτουμε $z = x_2$ στην 1η σχέση και $z = x_1$ στη 2η σχέση, οπότε προκύπτει:

$$\varphi(x_2) \geq \varphi(x_1) + \nabla\varphi(x_1) \cdot (x_2 - x_1)$$

$$\varphi(x_1) \geq \varphi(x_2) + \nabla\varphi(x_2) \cdot (x_1 - x_2).$$

Προσθέτοντας κατά μέλη τις δύο ανισότητες, προκύπτει:

$$\left(\nabla\varphi(x_1) - \nabla\varphi(x_2)\right) \cdot (x_1 - x_2) \geq 0$$

όπως επιθυμούσαμε.

■

Πρόταση 2.20

Το υποδιαφορικό μιας κυρτής συνάρτησης είναι μια μονότονη απεικόνιση. Δηλαδή, αν $y_1 \in \partial\varphi(x_1)$ και $y_2 \in \partial\varphi(x_2)$, τότε ισχύει $(y_2 - y_1) \cdot (x_2 - x_1) \geq 0$.

Απόδειξη:

Εξ ορισμού

$$y_1 \in \partial\varphi(x_1) \Leftrightarrow [\ \forall z \in \mathbb{R}^n, \quad \varphi(z) \geq \varphi(x_1) + y_1 \cdot (z - x_1)]$$

$$y_2 \in \partial\varphi(x_2) \Leftrightarrow [\ \forall z \in \mathbb{R}^n, \quad \varphi(z) \geq \varphi(x_2) + y_2 \cdot (z - x_2)].$$

Θέτουμε $z = x_2$ στην 1η σχέση και $z = x_1$ στη 2η σχέση, οπότε προκύπτει:

$$\varphi(x_2) \geq \varphi(x_1) + y_1 \cdot (x_2 - x_1)$$

$$\varphi(x_1) \geq \varphi(x_2) + y_2 \cdot (x_1 - x_2).$$

Προσθέτοντας κατά μέλη τις δύο ανισότητες, προκύπτει $y_1 \cdot (x_2 - x_1) + y_2 \cdot (x_1 - x_2) \leq 0$ ή αλλιώς $(y_2 - y_1) \cdot (x_2 - x_1) \geq 0$ όπως επιθυμούσαμε.

■

6. Συζυγείς συναρτήσεις.

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική συνάρτηση. Για κάθε $y \in \mathbb{R}^n$, ορίζουμε τη συζυγή κυρτή συνάρτηση της φ, μέσω της σχέσης

$$\varphi^*(y) \coloneqq \sup_{x \in \mathbb{R}^n} [x \cdot y - \varphi(x)]. \tag{2.23}$$

Η φ^* λέγεται, επίσης, μετασχηματισμός Legendre της φ και είναι μια ειδική περίπτωση (για $E = \mathbb{R}^n$) του μετασχηματισμού Legendre – Fenchel, που ορίσθηκε στην παράγραφο 1.1.5.

Πρόταση 2.21

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική συνάρτηση. Τότε, η φ^* είναι μια κυρτή κάτω ημισυνεχής συνάρτηση με τιμές στο $\mathbb{R} \cup \{+\infty\}$. Αν, επιπλέον, υποθέσουμε ότι η φ είναι κυρτή και κάτω ημισυνεχής, τότε η φ^* είναι κανονική.

Απόδειξη:

Έστω $y_1, y_2 \in \mathbb{R}^n$ και $t \in (0,1)$.

$$\begin{aligned}
\varphi^*(ty_1 + (1-t)y_2) &= \sup_{x \in \mathbb{R}^n} [x \cdot (ty_1 + (1-t)y_2) - \varphi(x)] \\
&= \sup_{x \in \mathbb{R}^n} [tx \cdot y_1 + (1-t)x \cdot y_2 - t\varphi(x) - (1-t)\varphi(x)] \\
&= \sup_{x \in \mathbb{R}^n} \left[t\big(x \cdot y_1 - \varphi(x)\big) + (1-t)\big(x \cdot y_2 - \varphi(x)\big)\right] \\
&\leq t \sup_{x \in \mathbb{R}^n} [x \cdot y_1 - \varphi(x)] + (1-t) \sup_{x \in \mathbb{R}^n} [x \cdot y_2 - \varphi(x)] = t\varphi^*(y_1) + (1-t)\varphi^*(y_2).
\end{aligned}$$

Άρα, η φ^* είναι κυρτή.

Για την απόδειξη της κάτω ημισυνέχειας της φ^*, έστω $a \in \mathbb{R}$. Έχουμε, τότε:

$$\begin{aligned}
\{y \in \mathbb{R}^n : \ \varphi^*(y) > a\} &= \left\{y \in \mathbb{R}^n : \ \sup_{x \in \mathbb{R}^n} [x \cdot y - \varphi(x)] > a\right\} \\
&= \{y \in \mathbb{R}^n : \ \exists x \in \mathbb{R}^n \ \text{τ.ω} \ x \cdot y - \varphi(x) > a\} = \bigcup_{x \in \mathbb{R}^n} \{y \in \mathbb{R}^n : \ x \cdot y - \varphi(x) > a\}.
\end{aligned}$$

Για κάθε $x \in \mathbb{R}^n$, η συνάρτηση $x \cdot y - \varphi(x)$ είναι συνεχής ως συνάρτηση του y και συνεπώς, το αντίστοιχο σύνολο είναι ανοιχτό. Άρα, και το σύνολο $\{y \in \mathbb{R}^n : \ \varphi^*(y) > a\}$ είναι ανοιχτό, ως ένωση ανοιχτών συνόλων. Άρα, το συμπλήρωμα αυτού, το σύνολο $\{y \in \mathbb{R}^n : \ \varphi^*(y) \leq a\}$ είναι κλειστό, το οποίο σύμφωνα με την Πρόταση 2.5, σημαίνει ότι η φ^* είναι κάτω ημισυνεχής.

Επιπλέον, επειδή η φ είναι κανονική, υπάρχει $x \in \mathbb{R}^n$ με $x \cdot y - \varphi(x) \in \mathbb{R}$, άρα $\varphi^*(x) > -\infty$.

Υποθέτουμε, ακόμα, ότι η φ είναι κυρτή και κάτω ημι-συνεχής. Τότε, το $epi\varphi$ είναι κυρτό και κλειστό σύνολο. Έστω $x_0 \in Dom(\varphi)$ και $\mu_0 < \varphi(x_0)$.

Θα εφαρμόσουμε τη δεύτερη γεωμετρική μορφή του Θεωρήματος Hahn-Banach για τα κυρτά σύνολα $A = epi\varphi$ και $B = \{(x_0, \mu_0)\}$. Το A είναι κλειστό, το B είναι συμπαγές και προφανώς $A \cap B = \emptyset$. Σύμφωνα, λοιπόν, με το Θεώρημα Hahn-Banach, υπάρχει κλειστό υπερεπίπεδο H στον $\mathbb{R}^{n+1}$ και συνεχές γραμμικό συναρτησοειδές Φ στον $\mathbb{R}^{n+1}$ τέτοιο ώστε $H = [\Phi = a]$ για κάποιο $a \in \mathbb{R}$, και το H διαχωρίζει τα A, B υπό την εξής έννοια:

υπάρχει $\varepsilon > 0$ τέτοιο ώστε $\Phi(x, \mu) \geq a + \varepsilon$, $(x, \mu) \in epi\varphi$ και $\Phi(x_0, \mu_0) \leq a - \varepsilon$.

Υπάρχουν $y \in \mathbb{R}^n$ και $k \in \mathbb{R}$ έτσι ώστε το γραμμικό συναρτησοειδές Φ να γράφεται στη μορφή

$$\Phi(x, \mu) = y \cdot x + k\mu, \qquad (x, \mu) \in \mathbb{R}^n \times \mathbb{R}.$$

Επειδή ισχύει $\Phi > a$ στο A και $\Phi < a$ στο B, παίρνουμε $y \cdot x + k\mu > a$, $(x, \mu) \in epi\varphi$ και $y \cdot x_0 + k\mu_0 < a$.

Ειδικότερα, η ανισότητα $y \cdot x + k\mu > a$ ισχύει για $x \in Dom(\varphi)$ και $\mu = \varphi(x)$, οπότε

$$y \cdot x + k\varphi(x) > a, \qquad x \in Dom(\varphi). \tag{2.24}$$

Επομένως, εφόσον $x_0 \in Dom(\varphi)$, έχουμε $y \cdot x_0 + k\varphi(x_0) > a > y \cdot x_0 + k\mu_0$, άρα, $k(\varphi(x_0) - \mu_0) > 0$ και συνεπώς $k > 0$.

Διαιρώντας με $-k < 0$ στη σχέση (2.24), παίρνουμε

$$-\frac{1}{k} y \cdot x - \varphi(x) < -\frac{a}{k}, \qquad x \in Dom(\varphi).$$

Συμπεραίνουμε ότι η φ^* είναι κανονική, αφού

$$\varphi^*\left(-\frac{1}{k} y\right) = \sup_{x \in \mathbb{R}^n} \left[-\frac{1}{k} y \cdot x - \varphi(x)\right] \leq -\frac{a}{k} < +\infty.$$

∎

Από τον ορισμό της φ^* μέσω της σχέσης (2.23), εύκολα προκύπτει ότι:

$$\forall x, y \in \mathbb{R}^n, \quad x \cdot y \leq \varphi(x) + \varphi^*(y). \tag{2.25}$$

Επίσης, για κάθε $x \in \mathbb{R}^n$, ορίζουμε τη συνάρτηση

$$\varphi^{**}(x) \coloneqq \sup_{y \in \mathbb{R}^n} \big(x \cdot y - \varphi^*(y)\big), \tag{2.26}$$

η οποία, επίσης, είναι κανονική, κυρτή και κάτω ημι-συνεχής. Εύκολα βλέπουμε ότι ικανοποιείται η ανισότητα

$$\forall x, y \in \mathbb{R}^n, \quad x \cdot y \leq \varphi^{**}(x) + \varphi^*(y). \tag{2.27}$$

Αρκετά σημαντικός είναι ο χαρακτηρισμός των περιπτώσεων στις οποίες ισχύει η ισότητα στη σχέση (2.25).

Πρόταση 2.22 (Χαρακτηρισμός του υποδιαφορικού)

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική κάτω ημισυνεχής κυρτή συνάρτηση. Τότε, για κάθε $x, y \in \mathbb{R}^n$, ισχύει $x \cdot y = \varphi(x) + \varphi^*(y) \iff y \in \partial\varphi(x)$.

<u>Απόδειξη:</u>

$$x \cdot y = \varphi(x) + \varphi^*(y) \iff x \cdot y \geq \varphi(x) + \varphi^*(y), \quad \text{λόγω της (2.25)}$$

$$\iff \varphi(x) \in \mathbb{R} \quad \text{και} \quad \forall z \in \mathbb{R}^n, \quad x \cdot y \geq \varphi(x) + y \cdot z - \varphi(z), \quad \text{λόγω της (2.23)}$$

$$\iff \forall z \in \mathbb{R}^n, \quad \varphi(z) \geq \varphi(x) + y \cdot (z - x)$$

$$\iff y \in \partial\varphi(x), \quad \text{λόγω της (2.15).}$$

■

Πρόταση 2.23 (Δυικότητα Legendre για κάτω ημι-συνεχείς κυρτές συναρτήσεις)

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική συνάρτηση τέτοια ώστε και η φ^* να είναι κανονική. Οι ακόλουθες τρεις προτάσεις είναι ισοδύναμες:

(i). Η φ είναι κυρτή κάτω ημισυνεχής.
(ii). $\varphi = \psi^*$ για κάποια κανονική συνάρτηση ψ.
(iii). $\varphi^{**} = \varphi$.

<u>Απόδειξη:</u>

Η πρόταση (iii) συνεπάγεται τη (ii), αφού αν $\varphi^{**} = \varphi$, τότε $\varphi = \psi^*$ με $\psi = \varphi^*$ και η ψ είναι κανονική συνάρτηση λόγω της υπόθεσης. Ακόμα, η (ii) συνεπάγεται την (i) λόγω της Πρότασης

2.21. Μένει, λοιπόν, να δείξουμε ότι η (i) συνεπάγεται την (iii). Η απόδειξη θα χωριστεί σε τρία βήματα.

Βήμα 1ο: Από τη σχέση (2.25), έχουμε: $\forall x, y \in \mathbb{R}^n, \ \ x \cdot y \le \varphi(x) + \varphi^*(y)$ οπότε

$$\forall x \in \mathbb{R}^n, \quad \varphi(x) \ge \sup_{y \in \mathbb{R}^n} [x \cdot y - \varphi^*(y)]$$

δηλαδή $\forall x \in \mathbb{R}^n, \ \ \varphi(x) \ge \varphi^{**}(x)$.

Βήμα 2ο: Έστω $x \in Int\big(Dom(\varphi)\big)$. Εφόσον, από την Πρόταση 2.12, ισχύει $\partial\varphi(x) \neq \emptyset$, επιλέγουμε $y_0 \in \partial\varphi(x)$. Λόγω της Πρότασης 2.22, ισχύει $\varphi(x) + \varphi^*(y_0) = x \cdot y_0$ άρα $\varphi(x) \le \sup_{y \in \mathbb{R}^n} [x \cdot y - \varphi^*(y)]$ δηλαδή $\varphi(x) \le \varphi^{**}(x)$.

Έχουμε δείξει, λοιπόν, ότι $\varphi = \varphi^{**}$ στο $Int\big(Dom(\varphi)\big)$. Ειδικότερα, αν $Dom(\varphi) = \mathbb{R}^n$, τότε $\varphi = \varphi^{**}$ σε όλο τον $\mathbb{R}^n$ και η πρόταση έχει αποδειχθεί.

Βήμα 3ο: Προκειμένου να αποδείξουμε την πρόταση και στη γενική περίπτωση, θα εφαρμόσουμε infimal συνέλιξη στην φ. Έστω $\varepsilon > 0$. Θέτουμε

$$\psi_\varepsilon(x) = \frac{|x|^2}{2\varepsilon}, \qquad x \in \mathbb{R}^n$$

και $\varphi_\varepsilon = \varphi *_{inf} \psi_\varepsilon$.

Για κάθε $x \in \mathbb{R}^n$, η συνάρτηση φ_ε γράφεται ως εξής:

$$\varphi_\varepsilon(x) = \inf_{z_1 + z_2 = x} \{\varphi(z_1) + \psi_\varepsilon(z_2)\} = \inf_{z \in \mathbb{R}^n} \{\varphi(z) + \psi_\varepsilon(x - z)\} = \inf_{z \in \mathbb{R}^n} \left\{\varphi(z) + \frac{|x - z|^2}{2\varepsilon}\right\}.$$

Σταθεροποιούμε ένα σημείο $x \in \mathbb{R}^n$. Θέτοντας $z = x$ στην παράσταση που βρίσκεται εντός των αγκυλών, προκύπτει η ανισότητα $\varphi_\varepsilon(x) \le \varphi(x)$.

Ισχυρισμός 6.

$$\forall x \in \mathbb{R}^n, \qquad \varphi(x) = \lim_{\varepsilon \to 0} \varphi_\varepsilon(x).$$

Απόδειξη του Ισχυρισμού 6:

Έστω ένας πραγματικός αριθμός $a < \varphi(x)$. Ζητάμε να βρούμε $\varepsilon > 0$ έτσι ώστε

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) + \frac{|x - z|^2}{2\varepsilon} > a.$$

Τότε θα ισχύει επίσης:

$$\varphi_\varepsilon(x) = \inf_{z \in \mathbb{R}^n} \left\{\varphi(z) + \frac{|x - z|^2}{2\varepsilon}\right\} \ge a,$$

και η ανισότητα αυτή εξασφαλίζει το ζητούμενο.

Εφόσον η φ είναι κάτω ημισυνεχής, ισχύει $\liminf_{z\to x} \varphi(z) \geq \varphi(x)$.

Άρα, υπάρχει $\delta > 0$ τέτοιο ώστε $|z-x| < \delta \Longrightarrow \varphi(z) > a$ και τότε επίσης

$$|z-x| < \delta \Longrightarrow \varphi(z) + \frac{|x-z|^2}{2\varepsilon} > a.$$

Η φ είναι κάτω φραγμένη από μια αφινική συνάρτηση. Πράγματι, λόγω της σχέσης (2.25), έχουμε: $\forall z, y \in \mathbb{R}^n$, $\varphi(z) \geq z \cdot y - \varphi^*(y)$. Άρα, επιλέγοντας $y_0 \in \mathbb{R}^n - \{0\}$ τέτοιο ώστε $\varphi^*(y_0) < +\infty$, έχουμε: $\forall z \in \mathbb{R}^n$, $\varphi(z) \geq z \cdot y_0 - \varphi^*(y_0)$.

Επομένως, $\varphi(z) \geq \gamma + \beta \cdot z$ ή αλλιώς $\varphi(z) \geq \gamma + \beta \cdot (z-x)$ για κάποια $\beta \in \mathbb{R}^n$ και $\gamma \in \mathbb{R}$.

Αν, τώρα, $|z-x| \geq \delta$, θα έχουμε:

$$\begin{aligned}\varphi(z) + \frac{|x-z|^2}{2\varepsilon} &\geq \gamma + \beta \cdot (z-x) + \frac{|x-z|^2}{2\varepsilon} \geq \gamma - |\beta||z-x| + \frac{|x-z|^2}{2\varepsilon} \\ &= |z-x|\left(\frac{|x-z|}{2\varepsilon} - |\beta|\right) + \gamma \geq \delta\left(\frac{\delta}{2\varepsilon} - |\beta|\right) + \gamma > \alpha,\end{aligned}$$

όπου

$$\frac{|x-z|}{2\varepsilon} - |\beta| > 0$$

καθώς οι δυο τελευταίες ανισότητες ισχύουν, αρκεί να επιλέξουμε το ε αρκούντως μικρό.

Συνδυάζοντας τα παραπάνω, βλέπουμε ότι μπορούμε να πετύχουμε την ανισότητα

$$\varphi(z) + \frac{|x-z|^2}{2\varepsilon} > a,$$

για κάθε $z \in \mathbb{R}^n$, αρκεί να επιλέξουμε το ε αρκούντως μικρό. Τότε,

$$\varphi_\varepsilon(x) = \inf_{z\in\mathbb{R}^n}\left\{\varphi(z) + \frac{|x-z|^2}{2\varepsilon}\right\} \geq a$$

που σημαίνει ότι $\lim_{\varepsilon\to 0} \varphi_\varepsilon(x) = \varphi(x)$.

■

Η φ_ε είναι κυρτή (σύμφωνα με την παρακάτω παρατήρηση) και $Dom(\varphi_\varepsilon) = \mathbb{R}^n$, οπότε η φ_ε είναι συνεχής. Άρα, από το 2ο βήμα της απόδειξης γνωρίζουμε ότι $(\varphi_\varepsilon)^{**} = \varphi_\varepsilon$. Από την άλλη, εφόσον $\varphi_\varepsilon(x) \leq \varphi(x)$, θα ισχύει $\varphi_\varepsilon^*(y) \geq \varphi^*(y)$ και $\varphi_\varepsilon^{**}(x) \leq \varphi^{**}(x)$. Συμπεραίνουμε ότι για κάθε $x \in \mathbb{R}^n$ ισχύει: $\varphi^{**}(x) \geq \liminf_{\varepsilon\to 0} \varphi_\varepsilon^{**}(x) = \liminf_{\varepsilon\to 0} \varphi_\varepsilon(x) = \varphi(x)$.

Συνδυάζοντας την ανισότητα αυτή με το 1ο βήμα της απόδειξης, προκύπτει ότι $\varphi^{**} = \varphi$, όπως επιθυμούσαμε.

■

Παρατήρηση (Θεώρημα 5.4, [13])

Έστω $\varphi_1, \dots, \varphi_m$ κανονικές κυρτές συναρτήσεις στον $\mathbb{R}^n$ και

$$\varphi(x) = \inf\{\varphi_1(x_1) + \cdots + \varphi_m(x_m) \;:\; x_i \in \mathbb{R}^n, x_1 + \cdots + x_m = x\}.$$

Τότε, η φ είναι κυρτή στον $\mathbb{R}^n$.

Πόρισμα 2.24

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική κάτω ημισυνεχής κυρτή συνάρτηση. Τότε, για κάθε $x, y \in \mathbb{R}^n$, ισχύει $y \in \partial\varphi(x) \Leftrightarrow x \in \partial\varphi^*(y)$.

Απόδειξη:

Σύμφωνα με τη Πρόταση 2.23, εφόσον η φ είναι μια κανονική κάτω ημι-συνεχής κυρτή συνάρτηση, ισχύει $\varphi^{**} = \varphi$. Επομένως:

$$y \in \partial\varphi(x) \Leftrightarrow x \cdot y = \varphi(x) + \varphi^*(y) \Leftrightarrow x \cdot y = \varphi^{**}(x) + \varphi^*(y) \Leftrightarrow x \in \partial\varphi^*(y)$$

Η πρώτη και η τελευταία ισοδυναμία ισχύουν λόγω της Πρότασης 2.22.

■

Πρόταση 2.25

Αν η φ είναι μια κανονική κάτω ημισυνεχής κυρτή συνάρτηση, τότε η απεικόνιση του υποδιαφορικού είναι συνεχής σε όλο τον $\mathbb{R}^n$, υπό την εξής έννοια:

$$\text{αν } x_k \to x \text{ και } \partial\varphi(x_k) \ni y_k \to y, \text{ τότε } y \in \partial\varphi(x).$$

Απόδειξη:

Για κάθε $k \in \mathbb{N}$, έχουμε: $y_k \in \partial\varphi(x_k) \Leftrightarrow x_k \cdot y_k = \varphi(x_k) + \varphi^*(y_k)$ και λόγω της σχέσης (2.25): $y_k \in \partial\varphi(x_k) \Leftrightarrow x_k \cdot y_k \geq \varphi(x_k) + \varphi^*(y_k)$.

Άρα, $\liminf_{k\to\infty}(x_k \cdot y_k) \geq \liminf_{k\to\infty}\big(\varphi(x_k) + \varphi^*(y_k)\big)$, δηλαδή

$$x \cdot y \geq \liminf_{k\to\infty} \varphi(x_k) + \liminf_{k\to\infty} \varphi^*(y_k) \geq \varphi(x) + \varphi^*(y).$$

Και πάλι λόγω της (2.25), έχουμε $x \cdot y = \varphi(x) + \varphi^*(y)$ η οποία ισοδυναμεί με $y \in \partial\varphi(x)$.

■

Πόρισμα 2.26

Αν η φ είναι μια κανονική κάτω ημισυνεχής κυρτή συνάρτηση, τότε το $Graph(\partial\varphi)$ είναι κλειστό ως υποσύνολο του $\mathbb{R}^{n+1}$.

Αν $(x_k)_{k=1}^{\infty}$ είναι μια ακολουθία που συγκλίνει στο x και $(y_k)_{k=1}^{\infty}$ μια ακολουθία τέτοια ώστε $y_k \in \partial\varphi(x_k)$, τότε η y_k δεν είναι αναγκαστικά συγκλίνουσα. Αν, όμως το σύνολο $\partial\varphi(x)$ αποτελείται από ένα και μοναδικό στοιχείο, τότε η παρακάτω πρόταση δείχνει ότι η ακολουθία y_k συγκλίνει και έχει όριο το $\nabla\varphi(x)$.

Πρόταση 2.27 (Συνέχεια της παραγώγου)

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κάτω ημι-συνεχής κανονική κυρτή συνάρτηση και έστω x ένα σημείο στο οποίο η φ είναι διαφορίσιμη. Συμβολίζουμε $y = \nabla\varphi(x)$ και με $B_r(x_0)$ σημειώνουμε την Ευκλείδεια μπάλα με κέντρο το x_0 και ακτίνα r. Τότε, για κάθε $\varepsilon > 0$, υπάρχει $\delta > 0$ τέτοιο ώστε

$$\partial\varphi\big(B_\delta(x)\big) \subseteq B_\varepsilon(y). \tag{2.28}$$

Απόδειξη:

Θεωρούμε ακολουθία $(x_k)_{k=1}^{\infty}$ με $x_k \to x$ και ακολουθία $(y_k)_{k=1}^{\infty}$ με $y_k \in \partial\varphi(x_k)$. Αρκεί να αποδείξουμε ότι η $(y_k)_{k=1}^{\infty}$ είναι φραγμένη.

Πράγματι, αν η $(y_k)_{k=1}^{\infty}$ είναι φραγμένη, τότε θα υπάρχει συγκλίνουσα υπακολουθία $\left(y_{k_n}\right)_{n=1}^{\infty}$, η οποία σύμφωνα με την Πρόταση 2.25, θα συγκλίνει στο $y = \nabla\varphi(x)$. Αν υποθέσουμε ότι υπάρχει $\varepsilon > 0$ τέτοιο ώστε για κάθε $\delta > 0$, να ισχύει $\partial\varphi\big(B_\delta(x)\big) \nsubseteq B_\varepsilon(y)$, τότε, για κάθε $k \in \mathbb{N}$, επιλέγουμε $x_k \in B_{\frac{1}{k}}(x)$ και $y_k \in \partial\varphi(x_k)$, με $y_k \notin B_\varepsilon(y)$. Από την άλλη όμως, υπάρχει συγκλίνουσα υπακολουθία $y_{k_n} \to y$. Έχουμε καταλήξει σε αντίφαση.

Θα τελειώσουμε την απόδειξη δείχνοντας ότι η $(y_k)_{k=1}^{\infty}$ είναι φραγμένη. Λόγω της συνέχειας της φ στο x, υπάρχουν $r > 0$, $M > 0$ έτσι ώστε $|\varphi(z)| \leq M$, για κάθε $z \in B_r(x)$.

Για $\kappa = 1,2,\dots$ επιλέγουμε

$$z_k = x_k + \frac{r}{2|y_k|} y_k.$$

Τότε υπάρχει $k_0 \in \mathbb{N}$ τέτοιο ώστε $z_k \in B_r(x)$, $k \geq k_0$. Οπότε, για $k \geq k_0$ ισχύει

$$M \geq \varphi(z_k) \geq \varphi(x_k) + y_k \cdot (z_k - x_k) = \varphi(x_k) + y_k \cdot \frac{r}{2|y_k|} y_k = \varphi(x_k) + \frac{r}{2}|y_k|.$$

Άρα, για $k \geq k_0$ ισχύει

$$|y_k| \leq \frac{2}{r}\big(M - \varphi(x_k)\big)$$

και η ακολουθία $\varphi(x_k)$ είναι συγκλίνουσα με όριο το $\varphi(x)$, άρα φραγμένη.

■

7. Κανονικοποίηση.

Ορισμός 2.28

Αν φ, ψ είναι δύο κανονικές κυρτές συναρτήσεις στο $\mathbb{R}^n$, ορίζουμε την infimal συνέλιξη των δύο συναρτήσεων μέσω της σχέσης:

$$\big(\varphi *_{inf} \psi\big)(z) = \inf_{x_1+x_2=z} [\varphi(x_1) + \psi(x_2)],$$

την οποία μπορούμε να γράψουμε και ως εξής:

$$\big(\varphi *_{inf} \psi\big)(z) = \inf_{x\in\mathbb{R}^n} [\varphi(x) + \psi(z - x)].$$

Πρόταση 2.29

Αν οι φ, ψ είναι κανονικές κυρτές συναρτήσεις, τότε $\big(\varphi *_{inf} \psi\big)^* = \varphi^* + \psi^*$.

<u>Απόδειξη:</u>

$$\begin{aligned}\big(\varphi *_{inf} \psi\big)^*(y) &= \sup_{z\in\mathbb{R}^n}\big(z \cdot y - \big(\varphi *_{inf} \psi\big)(z)\big) = \sup_{z\in\mathbb{R}^n}\left(z \cdot y - \inf_{x_1+x_2=z}[\varphi(x_1) + \psi(x_2)]\right) \\ &= \sup_{z\in\mathbb{R}^n} \sup_{x_1+x_2=z} \Big(z \cdot y - \big(\varphi(x_1) + \psi(x_2)\big)\Big) = \sup_{x_1,x_2\in\mathbb{R}^n}\big(x_1 \cdot y - \varphi(x_1) + x_2 \cdot y - \psi(x_2)\big) \\ &= \sup_{x_1\in\mathbb{R}^n}\big(x_1 \cdot y - \varphi(x_1)\big) + \sup_{x_2\in\mathbb{R}^n}\big(x_2 \cdot y - \psi(x_2)\big) = \varphi^*(y) + \psi^*(y).\end{aligned}$$

■

Παρατηρήσεις 2.30

(i). Υπάρχουν κυρτές συναρτήσεις φ, οι οποίες δεν είναι κάτω ημισυνεχείς, και άρα $\varphi^{**} \neq \varphi$. Θα κατασκευάσουμε μια τέτοια συνάρτηση και θα δούμε ότι η ιδιομορφία αυτή μπορεί να συμβαίνει μόνο στο $\partial Dom(\varphi)$.

Απόδειξη:

Ορίζουμε τη συνάρτηση $\varphi: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ με

$$\varphi(x) = \begin{cases} 0, & x \in (0,1) \\ +\infty, & x \notin (0,1) \end{cases}$$

Η φ έχει επιγράφημα $epi\varphi = (0,1) \times [0,+\infty)$, το οποίο δεν είναι κλειστό σύνολο και επομένως, σύμφωνα με το λήμμα 2.5, η φ δεν είναι κάτω ημι-συνεχής. Ακόμα,

$$\varphi^*(y) = \sup_{x \in \mathbb{R}}[xy - \varphi(x)] = \begin{cases} y, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

και

$$\varphi^{**}(x) = \sup_{y \in \mathbb{R}}[xy - \varphi^*(y)] = \begin{cases} 0, & x \in [0,1] \\ +\infty, & x \notin [0,1] \end{cases}$$

Παρατηρούμε, λοιπόν, ότι $\varphi^{**} \neq \varphi$, αλλά η ιδιομορφία αυτή συμβαίνει μόνο στο $\partial Dom(\varphi) = \{0, 1\}$.

■

(ii). Αν η φ είναι κάτω ημισυνεχής συνάρτηση και αυστηρά κυρτή στην περιοχή κάποιου $x \in \mathbb{R}^n$, τότε η φ^* είναι διαφορίσιμη στο $\partial\varphi(x)$, και για κάθε $y \in \partial\varphi(x)$ ισχύει $\nabla\varphi^*(y) = x$.

Απόδειξη:

Καταρχήν, παρατηρούμε ότι αν υπάρχει $z \in \partial\varphi(x) \cap \partial\varphi(u)$, τότε $x = u$. Πράγματι, ας υποθέσουμε ότι $x \neq u$. Τότε, για κάθε $t \in [0, 1]$, από τον ορισμό του υποδιαφορικού μέσω της σχέσης (2.15), έχουμε $\varphi(tx + (1-t)u) \geq \varphi(x) + (1-t)z \cdot (u - x)$ και $\varphi(tx + (1-t)u) \geq \varphi(u) - tz \cdot (u - x)$.

Άρα, σύμφωνα με τις παραπάνω σχέσεις και την κυρτότητα της φ, για $t \in [0,1]$ έχουμε:

$$\begin{aligned} \varphi(tx + (1-t)u) &= t\varphi(tx + (1-t)u) + (1-t)\varphi(tx + (1-t)u) \\ &\geq t\big(\varphi(x) + (1-t)z \cdot (u-x)\big) + (1-t)\big(\varphi(u) - tz \cdot (u-x)\big) = t\varphi(x) + (1-t)\varphi(u) \\ &\geq \varphi(tx + (1-t)u). \end{aligned}$$

Άρα, παντού ισχύει η ισότητα, δηλαδή για κάθε $t \in [0,1]$ ισχύει $\varphi(tx + (1-t)u) = t\varphi(x) + (1-t)\varphi(u)$ το οποίο, όμως, αντιφάσκει με την υπόθεση ότι η φ είναι αυστηρά κυρτή σε μια περιοχή του x.

Από την Πρόταση 2.18, για να αποδείξουμε το ζητούμενο αρκεί να δείξουμε ότι $\partial\varphi^*(y) = \{x\}$. Από το πόρισμα 2.24, έχουμε ότι $x \in \partial\varphi^*(y)$, αφού $y \in \partial\varphi(x)$. Αν, επιπλέον, υποθέσουμε ότι $u \in \partial\varphi^*(y)$, τότε $y \in \partial\varphi(u)$ και άρα αναγκαστικά $x = u$.

■

(iii). Αν η φ είναι διαφορίσιμη και αυστηρά κυρτή στο $\mathbb{R}^n$, τότε και η φ^* είναι, επίσης, διαφορίσιμη και αυστηρά κυρτή στο $Int\big(Dom(\varphi^*)\big)$ και η $\nabla\varphi$ είναι 1-1. Μάλιστα, ισχύει:

$$(\nabla\varphi)^{-1} = \nabla\varphi^*. \tag{2.29}$$

<u>Απόδειξη:</u>

Έστω $y \in Int\big(Dom(\varphi^*)\big)$. Από την Πρόταση 2.12, υπάρχει $x \in \partial\varphi^*(y)$. Τότε, από το πόρισμα 2.24, $y \in \partial\varphi(x)$. Άρα, από την προηγούμενη παρατήρηση, η φ^* είναι διαφορίσιμη στο y και $\partial\varphi^*(y) = \{\nabla\varphi^*(y)\} = \{x\}$. Αφού η φ είναι διαφορίσιμη, $\nabla\varphi(x) = y$.

Αντίστροφα, για κάθε $x \in \mathbb{R}^n$, η φ είναι διαφορίσιμη στο x. Αν $y = \nabla\varphi(x)$, από την προηγούμενη παρατήρηση παίρνουμε ότι η φ^* είναι διαφορίσιμη στο y (οπότε $y \in Int\big(Dom(\varphi^*)\big)$, αφού η διαφορισιμότητα της φ^* στο y συνεπάγεται ότι η φ^* έχει πεπερασμένες τιμές σε μια περιοχή του y) και $\nabla\varphi^*(y) = x$.

Από τα παραπάνω προκύπτει ότι: αν $x \in \mathbb{R}^n$, $y \in Int\big(Dom(\varphi^*)\big)$, τότε

$$y = \nabla\varphi(x) \quad \text{αν και μόνο αν} \quad x = \nabla\varphi^*(y).$$

Άρα, οι απεικονίσεις

$$\begin{aligned} Int\big(Dom(\varphi^*)\big) &\longrightarrow \mathbb{R}^n \\ y &\longmapsto \nabla\varphi^*(y) \end{aligned}$$

και

$$\begin{aligned} \mathbb{R}^n &\longrightarrow Int\big(Dom(\varphi^*)\big) \\ x &\longmapsto \nabla\varphi(x) \end{aligned}$$

είναι αντίστροφες η μία της άλλης. Άρα, η $\nabla\varphi$ είναι 1-1 και ισχύει η σχέση (2.29). Όμοια, και η $\nabla\varphi^*$ είναι 1-1.

Μένει να αποδείξουμε ότι η φ^* είναι αυστηρά κυρτή. Πριν από αυτό, παρατηρούμε το εξής:

Αν $U \subseteq \mathbb{R}^n$ είναι ένα κυρτό ανοιχτό σύνολο, $\varphi : U \longrightarrow \mathbb{R}$ μια κυρτή συνάρτηση, $h : U \longrightarrow \mathbb{R}$ μια αφινική συνάρτηση, δηλαδή $h(z) = a \cdot z + \beta$, $a \in \mathbb{R}^n, \beta \in \mathbb{R}$, $\varphi(z) \geq h(z)$, $z \in U$ και $\varphi(x_0) = h(x_0)$, τότε $a \in \partial\varphi(x_0)$.

Πράγματι, από τις σχέσεις $\varphi(z) \geq a \cdot z + \beta$ και $\varphi(x_0) = a \cdot x_0 + \beta$ παίρνουμε $\varphi(z) - \varphi(x_0) \geq a \cdot (z - x_0)$, που εξ ορισμού σημαίνει ότι $a \in \partial\varphi(x_0)$.

Υποθέτουμε, λοιπόν, προς απαγωγή σε άτοπο, ότι η φ^* δεν είναι αυστηρά κυρτή. Τότε, υπάρχουν $x_1, x_2 \in Int\big(Dom(\varphi^*)\big)$ και $t \in (0,1)$ έτσι ώστε $x_1 \neq x_2$ και

$$\varphi^*(tx_1 + (1-t)x_2) = t\varphi^*(x_1) + (1-t)\varphi^*(x_2).$$

Έστω $x = tx_1 + (1-t)x_2$ και $h(z) \coloneqq \varphi^*(x) + \nabla\varphi^*(x) \cdot (z - x),\ \ z \in \mathbb{R}^n$.

Τότε, αφού $\nabla\varphi^*(x) \in \partial\varphi^*(x)$, έχουμε $\varphi^*(z) \geq h(z),\ \ z \in \mathbb{R}^n$ και $\varphi^*(x) = h(x)$. Έπεται ότι

$$\varphi^*(x) = t\varphi^*(x_1) + (1-t)\varphi^*(x_2) \geq th(x_1) + (1-t)h(x_2) = h(x) = \varphi^*(x).$$

Άρα, υπάρχει παντού ισότητα και $\varphi^*(x_1) = h(x_1)$ και $\varphi^*(x_2) = h(x_2)$.

Από την παραπάνω παρατήρηση, έπεται ότι $\nabla\varphi^*(x) \in \partial\varphi^*(x_1)$ και $\nabla\varphi^*(x) \in \partial\varphi^*(x_2)$.

Η φ^* είναι διαφορίσιμη, άρα $\{\nabla\varphi^*(x_1)\} = \partial\varphi^*(x_1) = \{\nabla\varphi^*(x)\} = \partial\varphi^*(x_2) = \{\nabla\varphi^*(x_2)\}$.

Έπεται ότι $\nabla\varphi^*(x_1) = \nabla\varphi^*(x_2)$, το οποίο είναι άτοπο αφού η $\nabla\varphi^*$ είναι 1-1.

■

(iv). Αν, επιπλέον, υποθέσουμε ότι η φ είναι υπεργραμμική, δηλαδή

$$\lim_{|x|\to\infty} \frac{\varphi(x)}{|x|} = +\infty,$$

τότε $\nabla\varphi(\mathbb{R}^n) = \mathbb{R}^n$, και στην περίπτωση αυτή η $\nabla\varphi : \mathbb{R}^n \to \mathbb{R}^n$ είναι μια 1-1 και επί συνάρτηση και $(\nabla\varphi)^{-1} = \nabla\varphi^*$.

<u>Απόδειξη:</u>

Σύμφωνα με την παρατήρηση (iii), αρκεί να αποδείξουμε ότι $Int\big(Dom(\varphi^*)\big) = \mathbb{R}^n$, δηλαδή ότι $Dom(\varphi^*) = \mathbb{R}^n$, δηλαδή ότι $\varphi^*(y) < +\infty$, για κάθε $y \in \mathbb{R}^n$.

Προς απαγωγή σε άτοπο, θεωρούμε $y \in \mathbb{R}^n$ με $\varphi^*(y) = +\infty$, δηλαδή $\sup\limits_{x\in\mathbb{R}^n}\big(x \cdot y - \varphi(x)\big) = +\infty$.

Αφού η συνάρτηση $x \cdot y - \varphi(x)$ (ως συνάρτηση του x) είναι συνεχής, είναι φραγμένη στα φραγμένα υποσύνολα του $\mathbb{R}^n$. Άρα, υπάρχει ακολουθία $(x_k)_{k=1}^{\infty}$ με $|x_k| \to +\infty$ και $x_k \cdot y - \varphi(x_k) \to +\infty$. Ειδικότερα, $x_k \cdot y - \varphi(x_k) > 0$ τελικά. Όμως, από την υπόθεση της υπεργραμμικότητας για τη φ παίρνουμε ότι

$$\frac{x_k \cdot y - \varphi(x_k)}{|x_k|} \longrightarrow -\infty$$

αφού

$$\frac{|x_k \cdot y|}{|x_k|} \leq |y|.$$

Άρα, $x_k \cdot y - \varphi(x_k) < 0$ τελικά. Άτοπο!

■

(v). Αν οι συναρτήσεις φ και φ^* είναι δύο φορές διαφορίσιμες, τότε

$$D^2\varphi^*\big(\nabla\varphi(x)\big) = [D^2(\varphi(x)]^{-1}. \tag{2.30}$$

Απόδειξη:

Αν στην παρατήρηση (iii), $y = \nabla\varphi(x)$, είδαμε ότι $x = \nabla\varphi^*(y)$, άρα $\nabla\varphi^*\big(\nabla\varphi(x)\big) = x$. Παραγωγίζουμε τη σχέση αυτή και επειδή έχουμε διανυσματικές συναρτήσεις, παίρνουμε τους Ιακωβιανούς πίνακες D. Επίσης, D^2 είναι το σύμβολο για το $D(\nabla)$ και είναι ο Εσσιανός πίνακας. Στο πρώτο μέλος εφαρμόζουμε τον κανόνα της αλυσίδας.

$$I = D(\nabla\varphi^*)\big(\nabla\varphi(x)\big)\cdot D(\nabla\varphi)(x) = D^2\varphi^*\big(\nabla\varphi(x)\big)\cdot D^2\varphi(x).$$

■

8. Δεύτερη παράγωγος.

Θεώρημα 2.31 (Θεώρημα του Aleksandrov)

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κυρτή συνάρτηση. Τότε, η φ είναι δύο φορές διαφορίσιμη $\mathcal{L}^n$ −σχεδόν παντού στο $Int\big(Dom(\varphi)\big)$.

Ειδικότερα, $\mathcal{L}^n$ −σχεδόν για όλα τα $x \in \mathbb{R}^n$, ισχύει:

$$\lim_{y\to x}\frac{\left|\varphi(y)-\varphi(x)-D\varphi(x)\cdot(y-x)-\frac{1}{2}(y-x)^T\cdot D^2\varphi(x)\cdot(y-x)\right|}{|y-x|^2} = 0.$$

Για την απόδειξη του θεωρήματος παραπέμπουμε στο [7].

Ορισμός 2.32 (Aleksandrov παράγωγος)

Μια συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι δύο φορές παραγωγίσιμη στο σημείο $x_0 \in \mathbb{R}^n$ με Aleksandrov παράγωγο τον πίνακα $\Lambda \coloneqq \nabla^2\varphi(x_0)$, αν υπάρχει το $\nabla\varphi(x_0)$ και για κάθε $\varepsilon > 0$, υπάρχει $\delta > 0$ τέτοιο ώστε: αν $|x - x_0| < \delta$, τότε

$$\sup_{y\in\partial\varphi(x)}|y-\nabla\varphi(x_0)-\Lambda(x-x_0)| < \varepsilon|x-x_0| \tag{2.31}$$

Η Aleksandrov παράγωγος είναι ένας μη αρνητικός (θετικά ημιορισμένος) $n \times n$ πίνακας.

Παρατήρηση:

Αν $\psi(x) = \varphi(x + x_0) - y_0 \cdot x$, τότε $\nabla\psi(0) = \nabla\varphi(x_0) - y_0$ και

$$y \in \partial\psi(0) \text{ αν και μόνο αν } y + y_0 \in \partial\varphi(x_0).$$

Από αυτά έπεται ότι η φ έχει Aleksandrov παράγωγο Λ στο x_0 αν και μόνο αν η ψ έχει Aleksandrov παράγωγο Λ στο 0.

Πρόταση 2.33 (Θεώρημα αντίστροφης συνάρτησης για μονότονες απεικονίσεις)

Υποθέτουμε ότι η συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι κυρτή κάτω ημισυνεχής και έχει Aleksandrov παράγωγο $\Lambda \coloneqq \nabla^2\varphi(x_0)$ στο σημείο $x_0 \in \mathbb{R}^n$, συνεπώς το $\nabla\varphi(x_0)$ υπάρχει και $\varphi < +\infty$ σε μια περιοχή του x_0.

Αν ο Λ είναι αντιστρέψιμος, τότε η φ^* έχει τον Λ^{-1} ως Aleksandrov παράγωγο στο $\nabla\varphi(x_0)$.

Αν ο Λ δεν είναι αντιστρέψιμος, τότε η φ^* αποτυγχάνει να είναι δύο φορές παραγωγίσιμη στο $\nabla\varphi(x_0)$ υπό την έννοια του Aleksandrov.

Απόδειξη:

Θέτουμε $\nabla\varphi(x_0) = y_0$. Αν αντικαταστήσουμε τις συναρτήσεις $\varphi(x)$ και $\varphi^*(y)$ με τις $\psi(x) \coloneqq \varphi(x + x_0) - y_0 \cdot x$ και τον μετασχηματισμό Legendre αυτής, που είναι η συνάρτηση

$$\psi^*(y) = \sup_{x\in\mathbb{R}^n}[y \cdot x - \psi(x)] = \sup_{x\in\mathbb{R}^n}[y \cdot x - \varphi(x + x_0) + y_0 \cdot x]$$

$$= \sup_{x\in\mathbb{R}^n}[(y + y_0) \cdot x - \varphi(x + x_0)] = \sup_{x\in\mathbb{R}^n}[(y + y_0) \cdot (x + x_0) - \varphi(x + x_0) - (y + y_0) \cdot x_0]$$

$$= \sup_{x\in\mathbb{R}^n}[(y + y_0) \cdot (x + x_0) - \varphi(x + x_0)] - (y + y_0) \cdot x_0 = \varphi^*(y + y_0) - (y + y_0) \cdot x_0,$$

τότε η ψ είναι κυρτή και έχει Aleksadrov παράγωγο στο σημείο $x = 0$. Ακόμα, ισχύει $\nabla\psi(x) = \nabla\varphi(x + x_0) - y_0$, οπότε $\nabla\psi(0) = 0$.

Από την παραπάνω παρατήρηση και το γεγονός ότι η προσθήκη μιας σταθεράς δεν επηρεάζει την Aleksandrov παράγωγο, αρκεί να αποδείξουμε την πρόταση στην περίπτωση $x_0 = y_0 = 0$.

Ισχυρισμός 7. Αν ο Λ είναι αντιστρέψιμος, τότε η φ^* είναι διαφορίσιμη στο 0 με $\nabla\varphi^*(0) = 0$.

Απόδειξη του Ισχυρισμού 7:

Ο ισχυρισμός έπεται αν ισχύει η ισοδυναμία: $x \in \partial\varphi^*(0) \iff x = 0$.

Καταρχήν, θα δείξουμε ότι $0 \in \partial\varphi^*(0)$, το οποίο ισοδύναμα σημαίνει ότι $\forall z \in \mathbb{R}^n$, $\varphi^*(z) \geq \varphi^*(0) + 0 \cdot (z - 0) = \varphi^*(0)$.

Πράγματι, αφού $\nabla\varphi(0) = 0$, έπεται ότι $\varphi(0) < +\infty$ και $\forall x \in \mathbb{R}^n$, $\varphi(x) \geq \varphi(0) + 0 \cdot (x - 0) = \varphi(0)$. Άρα, $\varphi^*(0) = \sup_{x\in\mathbb{R}^n}[0 \cdot x - \varphi(x)] = \sup_{x\in\mathbb{R}^n}[-\varphi(x)] = -\varphi(0)$.

Απ' την άλλη, για κάθε $z \in \mathbb{R}^n$, έχουμε $\varphi^*(z) = \sup_{x\in\mathbb{R}^n}[z \cdot x - \varphi(x)] \geq z \cdot 0 - \varphi(0) = -\varphi(0)$, δηλαδή $\forall z \in \mathbb{R}^n$, $\varphi^*(z) \geq -\varphi(0) = \varphi^*(0)$.

Στη συνέχεια, έστω $x \in \partial\varphi^*(0)$, $x \neq 0$. Εφόσον το σύνολο $\partial\varphi^*(0)$ είναι κυρτό και $0 \in \partial\varphi^*(0)$, για κάθε $t \in (0,1]$ θα ισχύει $tx \in \partial\varphi^*(0)$. Σύμφωνα με το Πόρισμα 2.24, αυτό ισοδυναμεί με το ότι $0 \in \partial\varphi(tx)$, ή αλλιώς $(tx, 0) \in \partial\varphi$.

Για κάθε $\varepsilon > 0$, παίρνοντας $t < \frac{\delta}{|x|}$ στην (2.31) έχουμε $|tx| < \delta$ και $\sup_{y\in\partial\varphi(tx)} |y - \Lambda(tx)| < \varepsilon|tx|$ άρα $\sup_{y\in\partial\varphi(tx)} |y - t\Lambda x| < t\varepsilon|x|$ και αφού $0 \in \partial\varphi(tx)$, παίρνουμε

$$|0 - t\Lambda x| \leq \sup_{y\in\partial\varphi(tx)} |y - t\Lambda x| < t\varepsilon|x|$$

δηλαδή $t|\Lambda x| < t\varepsilon|x|$ και άρα $|\Lambda x| < \varepsilon|x|$.

Αφού το ε είναι ένας αυθαίρετος θετικός αριθμός, προκύπτει ότι $|\Lambda x| = 0$, άρα $\Lambda x = 0$, και αφού ο Λ είναι αντιστρέψιμος, $x = 0$.

Τελικά, η φ^* έχει μοναδική υποπαράγωγο στο σημείο 0, το διάνυσμα 0. Άρα, είναι παραγωγίσιμη στο 0, και $\nabla\varphi^*(0) = 0$.

■

Συνεχίζουμε την απόδειξη της πρότασης, δείχνοντας ότι η φ^* είναι δύο φορές διαφορίσιμη στο 0, στην περίπτωση που ο Λ είναι αντιστρέψιμος.

Αρκεί να αποδείξουμε ότι αν $y_k \to 0$ και $x_k \in \partial\varphi^*(y_k)$, τότε

$$\frac{|x_k - \Lambda^{-1}y_k|}{|y_k|} \to 0.$$

Έστω $\varepsilon > 0$. Αφού η φ έχει Aleksandrov παράγωγο στο 0, βρίσκουμε $\delta > 0$ έτσι ώστε, αν $|x| < \delta$, να ισχύει η σχέση (2.31).

Λόγω συνέχειας της απεικόνισης $\partial\varphi^*$ στο 0, σύμφωνα με την Πρόταση 2.25, $x_k \to 0$, καθώς $k \to +\infty$. Άρα, υπάρχει $k_0 \in \mathbb{N}$ τέτοιο ώστε $|x_k| < \delta$ για κάθε $k \geq k_0$.

Άρα για $k \geq k_0$, $x_k \in \partial\varphi^*(y_k)$, το οποίο ισοδυναμεί με $y_k \in \partial\varphi(x_k)$, και επίσης, ισχύει η σχέση (2.31) από την οποία παίρνουμε: $|y_k - \Lambda x_k| < \varepsilon|x_k|$.

Τότε,

$$\frac{\|\Lambda^{-1}\|}{\|\Lambda^{-1}\|}|y_k - \Lambda x_k| < \varepsilon|x_k|$$

άρα

$$\frac{|\Lambda^{-1}(y_k - \Lambda x_k)|}{\|\Lambda^{-1}\|} < \varepsilon|x_k - \Lambda^{-1}y_k + \Lambda^{-1}y_k|$$

άρα

$$\frac{|\Lambda^{-1}y_k - x_k|}{\|\Lambda^{-1}\|} < \varepsilon|x_k - \Lambda^{-1}y_k| + \varepsilon|\Lambda^{-1}y_k|$$

άρα

$$|x_k - \Lambda^{-1}y_k|\left(\frac{1}{\|\Lambda^{-1}\|} - \varepsilon\right) < \varepsilon|\Lambda^{-1}y_k|.$$

Για $\varepsilon < \frac{1}{2\|\Lambda^{-1}\|}$ παίρνουμε:

$$|x_k - \Lambda^{-1}y_k|\frac{1}{2\|\Lambda^{-1}\|} < \varepsilon|\Lambda^{-1}y_k| \leq \varepsilon\|\Lambda^{-1}\||y_k|$$

και συνεπώς $|x_k - \Lambda^{-1}y_k| < 2\varepsilon\|\Lambda^{-1}\|^2|y_k|$.

Άρα, για κάθε $0 < \varepsilon < \frac{1}{2\|\Lambda^{-1}\|}$, υπάρχει $k_0 \in \mathbb{N}$ έτσι ώστε για $k \geq k_0$

$$\frac{|x_k - \Lambda^{-1}y_k|}{|y_k|} < 2\varepsilon\|\Lambda^{-1}\|^2.$$

Η σχέση αυτή εκφράζει το γεγονός ότι η φ^* είναι δύο φορές παραγωγίσιμη στο 0 και μάλιστα ότι $\Lambda^{-1} = \nabla^2\varphi^*(0)$.

Μένει να ασχοληθούμε με την περίπτωση που ο Λ δεν είναι αντιστρέψιμος. Τότε, θα υπάρχει $x_0 \in \mathbb{R}^n - \{0\}$ τέτοιο ώστε $\Lambda x_0 = 0$.

Ισχυρισμός 8. Υπάρχει ακολουθία (x_k) πολλαπλασίων του x_0 με $x_k \to 0$ και ακολουθία $y_k \in \partial\varphi(x_k)$ τέτοια ώστε

$$|y_k| \leq \frac{1}{k}|x_k|.$$

<u>Απόδειξη του Ισχυρισμού 8:</u>

Από τη σχέση (2.31), υπάρχουν $\delta_k > 0$, $k \in \mathbb{N}$ με $\delta_k \to 0$ έτσι ώστε

$$\Lambda x = 0 \ \text{ και } \ |x| < \delta_k \implies \sup_{y \in \partial\varphi(x)} |y| < \frac{1}{k}|x|.$$

Επιλέγουμε ακολουθία αριθμών $(t_k)_{k=1}^{\infty}$ τέτοια ώστε

$$0 < t_k < \frac{\delta_k}{|x_0|}.$$

Για κάθε $k \in \mathbb{N}$, ορίζουμε $x_k := t_k x_0$. Τότε, $x_k \to 0$, καθώς $k \to +\infty$ και $\Lambda x_k = t_k \Lambda x_0 = 0$. Εφόσον $|x_k| = t_k|x_0| < \frac{\delta_k}{|x_0|}|x_0| = \delta_k$, από την επιλογή των δ_k έχουμε ότι

$$\forall k \in \mathbb{N}, \qquad \sup_{y\in\partial\varphi(x_k)} |y| < \frac{1}{k}|x_k|.$$

Αφού $0 \in Int(Dom\varphi)$, τα x_k μπορούν να επιλεγούν στο $Dom\varphi$, οπότε $\partial\varphi(x_k) \neq \emptyset$. Επιλέγοντας $y_k \in \partial\varphi(x_k)$, παίρνουμε τη ζητούμενη ακολουθία.

■

Για κάθε πίνακα Λ' και $\varepsilon > 0$ έχουμε:

$$|x_k - \Lambda' y_k| \geq |x_k| - |\Lambda' y_k| \geq |x_k| - \|\Lambda'\||y_k| \geq k|y_k| - \|\Lambda'\||y_k| = (k - \|\Lambda'\|)|y_k|.$$

Παίρνοντας το k αρκετά μεγάλο ώστε $k - \|\Lambda'\| > \varepsilon$, έχουμε $|x_k - \Lambda' y_k| \geq \varepsilon|y_k|$. Η σχέση αυτή εκφράζει το γεγονός ότι η φ^* αποτυγχάνει να είναι δύο φορές παραγωγίσιμη στο 0.

■

9. Παραμόρφωση του όγκου.

Πρόταση 2.34

Υποθέτουμε ότι η συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι κυρτή, κάτω ημισυνεχής και έχει Aleksandrov παράγωγο $\Lambda := \nabla^2\varphi(x_0)$ στο σημείο $x_0 \in \mathbb{R}^n$. *Αν* με $B_r(x_0)$ συμβολίσουμε τη μπάλα κέντρου x_0 και ακτίνας r, τότε:

$$\text{(2.32)} \qquad \lim_{r\to 0} \frac{vol\left[\partial\varphi\big(B_r(x_0)\big)\right]}{vol[B_r(x_0)]} = \det[\nabla^2\varphi(x_0)\,].$$

Όταν ο Λ είναι αντιστρέψιμος, το σύνολο $\partial\varphi\big(B_r(x_0)\big)$ συρρικνώνεται όμορφα στο σημείο $\nabla\varphi(x_0)$ υπό την εξής έννοια:

το σύνολο $\partial\varphi\big(B_r(x_0)\big)$ περιέχεται σε μια οικογένεια μπαλών $B_{R(r)}$ με $R(r) \to 0$ καθώς $r \to 0$, ενώ ταυτόχρονα περιέχει ένα ποσοστό της μπάλας $B_{R(r)}$ που είναι κάτω φραγμένο από έναν θετικό αριθμό.

<u>Απόδειξη:</u>

Όπως και στην προηγούμενη απόδειξη, αρκεί να αποδείξουμε την πρόταση στην περίπτωση που $x_0 = \nabla\varphi(x_0) = 0$.

Περίπτωση 1η: Ο πίνακας Λ είναι αντιστρέψιμος.

Έστω $\varepsilon > 0$. Σημειώνουμε με B_r την $B_r(0)$ και με ΛB_r την εικόνα της μέσω του πίνακα Λ.

Ισχυρισμός 9. Υπάρχει $\delta > 0$ τέτοιο ώστε για $r < \delta$ να έχουμε

$$\partial\varphi(B_r) \subseteq (1 + \varepsilon\|\Lambda^{-1}\|)\Lambda B_r. \tag{2.33}$$

Απόδειξη του Ισχυρισμού 9:

Έστω

$$y \in \partial\varphi(B_r) = \bigcup_{x\in B_r} \partial\varphi(x).$$

Τότε, $y \in \partial\varphi(x)$ για κάποιο $x \in B_r$. Εφαρμόζοντας τη σχέση (2.31), προκύπτει ότι υπάρχει $\delta > 0$ έτσι ώστε για $r < \delta$ να ισχύει $|y - \Lambda x| < \varepsilon|x|$.

Εφόσον μπορούμε να γράψουμε $y = \Lambda x + (y - \Lambda x)$, έχουμε: $\Lambda^{-1}y = x + \Lambda^{-1}(y - \Lambda x)$ άρα

$$\begin{aligned}|\Lambda^{-1}y| = |x + \Lambda^{-1}(y - \Lambda x)| \leq |x| + \|\Lambda^{-1}\||y - \Lambda x| < |x| + \|\Lambda^{-1}\|\varepsilon|x| = (1 + \varepsilon\|\Lambda^{-1}\|)|x| \\ < (1 + \varepsilon\|\Lambda^{-1}\|)r.\end{aligned}$$

Δηλαδή, $\Lambda^{-1}y \in (1 + \varepsilon\|\Lambda^{-1}\|)B_r$ και επομένως $y \in (1 + \varepsilon\|\Lambda^{-1}\|)\Lambda B_r$.

■

Από την Πρόταση 2.33, γνωρίζουμε ότι η φ^* έχει Aleksandrov παράγωγο Λ^{-1} στο σημείο $\nabla\varphi(0) = 0$.

Ισχυρισμός 10. Υπάρχει $\delta' > 0$ τέτοιο ώστε για $r < \delta'$ να έχουμε $\partial\varphi^*(\Lambda B_r) \subseteq (1 + \varepsilon\|\Lambda\|)B_r$.

Απόδειξη του Ισχυρισμού 10:

Έστω

$$x \in \partial\varphi^*(\Lambda B_r) = \bigcup_{y\in\Lambda B_r} \partial\varphi^*(y).$$

Τότε, $x \in \partial\varphi^*(y)$ για κάποιο $y \in \Lambda B_r$. Εφαρμόζοντας τη σχέση (2.31), προκύπτει ότι υπάρχει $\delta' > 0$ έτσι ώστε για $r < \delta'$ να ισχύει $|x - \Lambda^{-1}y| < \varepsilon|y|$, άρα $x - \Lambda^{-1}y \in \varepsilon\|\Lambda\|B_r$.

Εφόσον μπορούμε να γράψουμε $x = \Lambda^{-1}y + (x - \Lambda^{-1}y)$, έχουμε:

$$x \in \Lambda^{-1}(\Lambda B_r) + \varepsilon\|\Lambda\|B_r = (1 + \varepsilon\|\Lambda\|)B_r$$

■

Από την Πρόταση 2.33, είναι γνωστό ότι η φ^* έχει Aleksandrov παράγωγο στο 0, άρα $0 \in Int\big(Dom(\varphi^*)\big)$. Άρα, παίρνοντας το r μικρότερο αν χρειαστεί έχουμε

$$(1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r \subseteq Int\big(Dom(\varphi^*)\big).$$

Εφαρμόζοντας το προηγούμενο συμπέρασμα με το $(1+\varepsilon\|\Lambda\|)^{-1}B_r$ στη θέση του B_r και παίρνοντας το r μικρότερο αν χρειάζεται, παίρνουμε λόγω δυϊκότητας

$$(1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r \subseteq \partial\varphi\big(\partial\varphi^*((1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r)\big) \subseteq \partial\varphi(B_r),$$

άρα,

$$(1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r \subseteq \partial\varphi(B_r). \tag{2.34}$$

Συνδυάζοντας τις σχέσεις (2.33) και (2.34), παίρνουμε:

$$(1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r \subseteq \partial\varphi(B_r) \subseteq (1+\varepsilon\|\Lambda^{-1}\|)\Lambda B_r.$$

Ακόμα,

$$\frac{vol[(1+\varepsilon\|\Lambda\|)^{-1}\Lambda B_r]}{vol[B_r]} \leq \frac{vol[\partial\varphi(B_r)]}{vol[B_r]} \leq \frac{vol[(1+\varepsilon\|\Lambda^{-1}\|)\Lambda B_r]}{vol[B_r]},$$

ή ισοδύναμα

$$(1+\varepsilon\|\Lambda\|)^{-n}\frac{vol[\Lambda B_r]}{vol[B_r]} \leq \frac{vol[\partial\varphi(B_r)]}{vol[B_r]} \leq \frac{vol[\Lambda B_r]}{vol[B_r]}(1+\varepsilon\|\Lambda^{-1}\|)^n.$$

Χρησιμοποιώντας την ταυτότητα

$$\det(\Lambda) = \frac{vol[\Lambda B_r]}{vol[B_r]},$$

παίρνουμε

$$(1+\varepsilon\|\Lambda\|)^{-n}\det(\Lambda) \leq \frac{vol[\partial\varphi(B_r)]}{vol[B_r]} \leq \det(\Lambda)\,(1+\varepsilon\|\Lambda^{-1}\|)^n.$$

Άρα, για κάθε $\varepsilon > 0$

$$(1+\varepsilon\|\Lambda\|)^{-n}\det(\Lambda) \leq \lim_{r\to 0}\frac{vol[\partial\varphi(B_r)]}{vol[B_r]} \leq \det(\Lambda)\,(1+\varepsilon\|\Lambda^{-1}\|)^n.$$

Άρα,

$$\lim_{r\to 0}\frac{vol[\partial\varphi(B_r)]}{vol[B_r]} = \det(\Lambda).$$

Στη συνέχεια, για να αποδείξουμε ότι το σύνολο $\partial\varphi(B_r)$ συρρικνώνεται όμορφα στο σημείο $\nabla\varphi(0)$, θα χρησιμοποιήσουμε τον εγκλεισμό:

$$\frac{1}{\|\Lambda^{-1}\|}B_r \subseteq \Lambda B_r \subseteq \|\Lambda\|B_r.$$

Από τις σχέσεις (2.33), (2.34) και τους παραπάνω εγκλεισμούς, αν το r είναι αρκούντως μικρό, έχουμε:

$$[(1+\varepsilon\|\Lambda\|)\|\Lambda^{-1}\|]^{-1}B_r \subseteq \partial\varphi(B_r) \subseteq [(1+\varepsilon\|\Lambda^{-1}\|)\|\Lambda\|]B_r$$

δηλαδή

$$\frac{1}{(1+\varepsilon\|\Lambda\|)(1+\varepsilon\|\Lambda^{-1}\|)\|\Lambda\|\|\Lambda^{-1}\|}B_{\|\Lambda\|(1+\varepsilon\|\Lambda^{-1}\|)r} \subseteq \partial\varphi(B_r) \subseteq B_{\|\Lambda\|(1+\varepsilon\|\Lambda^{-1}\|)r}.$$

Επομένως, το σύνολο $\partial\varphi(B_r)$ περιέχεται σε μια οικογένεια μπαλών $B_{R(r)}$ με $R(r)\to 0$ καθώς το $r\to 0$, ενώ ταυτόχρονα περιέχει ένα «ποσοστό» της μπάλας $B_{R(r)}$ που είναι κάτω φραγμένο μακριά από το μηδέν. Αυτό εννοούμε με το ότι το σύνολο $\partial\varphi(B_r)$ συρρικνώνεται όμορφα προς το σημείο $\nabla\varphi(0)=0$.

Περίπτωση 2η: Ο πίνακας Λ είναι μη-αντιστρέψιμος.

Τότε, η ΛB_r περιέχεται σε έναν υπόχωρο E του $\mathbb{R}^n$ διάστασης $n-1$. Έστω $\varepsilon>0$.

Εφαρμόζοντας τη σχέση (2.31), υπάρχει $\delta>0$ έτσι ώστε για $|x|<r\le\delta$ και $y\in\partial\varphi(x)$ να έχουμε: $|y-\Lambda x|<\varepsilon|x|$.

Τότε, $y=\Lambda x+(y-\Lambda x)$, όπου $\Lambda x\in E$, $|\Lambda x|\le\|\Lambda\|r$ και $|y-\Lambda x|<\varepsilon r$. Δηλαδή, το y είναι άθροισμα ενός διανύσματος του συνόλου $\Lambda B_r\subseteq B_{\|\Lambda\|r}$ και ενός διανύσματος που έχει μήκος το πολύ εr. Άρα, το σύνολο $\partial\varphi(B_r)$ περικλείεται σε έναν κύλινδρο H_ε με $(n-1)$ −διάστατη βάση ακτίνας ίσης με $\|\Lambda\|r+\varepsilon r=(\|\Lambda\|+\varepsilon)r$ και ύψος $2\varepsilon r$.

Άρα, $vol[\partial\varphi(B_r)]\le vol[H_\varepsilon]=c[(\|\Lambda\|+\varepsilon)r\ ]^{n-1}2\varepsilon r=2\varepsilon(\|\Lambda\|+\varepsilon)^{n-1}cr^n$, όπου με c_{n-1} συμβολίζουμε τον όγκο της μοναδιαίας μπάλας στον $\mathbb{R}^{n-1}$. Για να ολοκληρώσουμε την απόδειξη, γράφουμε

$$\frac{vol[\partial\varphi(B_r)]}{vol[B_r]}\le\frac{2\varepsilon(\|\Lambda\|+\varepsilon)^{n-1}c_{n-1}r^n}{c_nr^n}=2\varepsilon(\|\Lambda\|+\varepsilon)^{n-1},$$

όπου c_n ο όγκος της μοναδιαίας μπάλας στον $\mathbb{R}^n$.

Για κάθε $\varepsilon>0$ προκύπτει ότι

$$\lim_{r\to 0}\frac{vol[\partial\varphi(B_r)]}{vol[B_r]}\le 2\varepsilon(\|\Lambda\|+\varepsilon)^{n-1}\cdot\frac{c_{n-1}}{c_n},$$

άρα

$$\lim_{r\to 0}\frac{vol[\partial\varphi(B_r)]}{vol[B_r]} = 0 = \det(\Lambda).$$

■

2.1.4. Επιστροφή στη μελέτη του δυϊκού προβλήματος Kantorovich.

Θεώρημα 2.35 (Ύπαρξη ενός βέλτιστου ζεύγους συζυγών κυρτών συναρτήσεων)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ με πεπερασμένες ροπές 2ης τάξης. Θεωρούμε, ακόμα, το σύνολο $\widetilde{\Phi}$, δηλαδή το σύνολο όλων των ζευγών μετρήσιμων συναρτήσεων $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$ που παίρνουν τιμές στο $\mathbb{R} \cup \{+\infty\}$ και ικανοποιούν την ανισότητα $x \cdot y \leq \varphi(x) + \psi(y)$, για $d\mu$ −σχεδόν όλα τα $x \in \mathbb{R}^n$ και $d\nu$ −σχεδόν όλα τα $y \in \mathbb{R}^n$.

Τότε, υπάρχει ζεύγος $(\varphi, \varphi^*) \in L^1(d\mu) \times L^1(d\nu)$ κάτω ημισυνεχών κανονικών συζυγών κυρτών συναρτήσεων στον $\mathbb{R}^n$, τέτοιων ώστε $\inf_{\widetilde{\Phi}} J = J(\varphi, \varphi^*)$.

Λήμμα 2.36 (Λήμμα διπλής κυρτοποίησης)

Έστω μ, ν δύο μέτρα πιθανότητας τα οποία συγκεντρώνονται αντίστοιχα στα υποσύνολα X, Y του $\mathbb{R}^n$ και ικανοποιούν τη σχέση

$$M_2 := \int_X \frac{|x|^2}{2}\, d\mu(x) + \int_Y \frac{|y|^2}{2}\, d\nu(y) < +\infty.$$

Αν $\varphi, \psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ είναι μετρήσιμες συναρτήσεις, ορίζουμε

$$\varphi^*(y) = \sup_{x\in X}[x \cdot y - \varphi(x)], \qquad y \in \mathbb{R}^n. \tag{2.35}$$

$$\psi^*(x) = \sup_{y\in Y}[x \cdot y - \psi(y)], \qquad x \in \mathbb{R}^n. \tag{2.36}$$

Θεωρούμε, ακόμα, το σύνολο $\widetilde{\Phi}$ όπως προηγουμένως και $(\varphi_k, \psi_k)_{k\in\mathbb{N}}$ μια ακολουθία στο $\widetilde{\Phi}$ τέτοια ώστε $\lim_{k\to\infty} J(\varphi_k, \psi_k) = \inf_{\widetilde{\Phi}} J$.

Τότε,

(i). Μπορούμε να τροποποιήσουμε τα ζεύγη (φ_k, ψ_k) πάνω σε σύνολα μέτρου 0 ως προς τα μέτρα μ και ν αντίστοιχα, με τέτοιο τρόπο ώστε η ανισότητα (2.5) να ισχύει για κάθε $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, και χωρίς να αλλάξουν οι τιμές του συναρτησοειδούς $J(\varphi_k, \psi_k)$.

(ii). Αν τα ζεύγη (φ_k, ψ_k) είναι τα τροποποιημένα σύμφωνα με το (i), τότε υπάρχει ακολουθία πραγματικών αριθμών $(a_k)_{k\in\mathbb{N}}$ ώστε η ακολουθία $\left(\overline{\varphi}_k, \overline{\psi}_k\right) = (\varphi_k^{**} + a_k, \varphi_k^* - a_k)$ να ανήκει στο $\widetilde{\Phi}$, να έχει την ιδιότητα $\lim_{k\to\infty} J\left(\overline{\varphi}_k, \overline{\psi}_k\right) = \inf_{\widetilde{\Phi}} J$ και να ισχύουν οι εκτιμήσεις:

$$\forall x \in X, \forall y \in Y\,, \qquad \overline{\varphi}_k(x) \geq -\frac{|x|^2}{2}, \quad \overline{\psi}_k(y) \geq -\frac{|y|^2}{2} \tag{2.37}$$

$$\liminf_{k\to\infty} \inf_{x\in X} \left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) \leq \inf_{\widetilde{\Phi}} J + M_2 \tag{2.38}$$

$$\liminf_{k\to\infty} \inf_{y\in Y} \left(\overline{\psi}_k(y) + \frac{|y|^2}{2}\right) \leq \inf_{\widetilde{\Phi}} J + M_2. \tag{2.39}$$

(iii). Αν $X = Y = \mathbb{R}^n$, τότε οι συναρτήσεις που ορίσαμε μέσω των σχέσεων (2.35) και (2.36) συμπίπτουν με τους μετασχηματισμούς Legendre των φ και ψ αντίστοιχα. Ακόμα, μπορούμε να περιορίσουμε το infimum του J πάνω στα ζεύγη των συζυγών κανονικών κυρτών συναρτήσεων χωρίς να μεταβληθεί η τιμή του, δηλαδή $\inf_{\widetilde{\Phi}} J = \inf_{\varphi\in L^1(d\mu)} J(\varphi^{**}, \varphi^*)$.

<u>Απόδειξη:</u>

Για την απόδειξη του (i), θεωρούμε ένα ζεύγος $(\varphi_k, \psi_k) \in \widetilde{\Phi}$. Γνωρίζουμε, τότε, ότι υπάρχουν μετρήσιμα σύνολα $N_1, N_2 \subseteq \mathbb{R}^n$ για τα οποία ισχύει $\mu[N_1] = \nu[N_2] = 0$ και η ανισότητα (2.5) ισχύει για κάθε $(x,y) \in N_1^c \times N_2^c$. Αλλάζουμε τις τιμές των φ_k και ψ_k, έτσι ώστε να είναι ίσες με $+\infty$ πάνω στα σύνολα N_1 και N_2 αντίστοιχα. Τότε, το ζεύγος (φ_k, ψ_k) συνεχίζει να ανήκει στο $\widetilde{\Phi}$, η ανισότητα (2.5) ισχύει για κάθε $x, y \in \mathbb{R}^n$, ενώ ακόμα, η τιμή του συναρτησοειδούς $J(\varphi_k, \psi_k)$ παραμένει αμετάβλητη.

Συνεχίζουμε με την απόδειξη του (ii).

<u>Βήμα 1ο</u>: Ορισμός της ακολουθίας $(a_k)_{k\in\mathbb{N}}$ και απόδειξη της σχέσης (2.37).

Επικαλούμενοι το μέρος (i) του λήμματος, έχουμε ότι για κάθε $k \in \mathbb{N}$ και για κάθε $x, y \in \mathbb{R}^n$ ισχύει $\varphi_k(x) \geq x\cdot y - \psi_k(y)$. Αφού $\varphi_k \in L^1(d\mu)$ και $\psi_k \in L^1(d\nu)$, υπάρχουν $x_0 \in X$ και $y_0 \in Y$ τέτοια ώστε $\varphi_k(x_0) < +\infty$, $\psi_k(y_0) < +\infty$ και $\varphi_k(x) \geq x\cdot y_0 - \psi_k(y_0)$, $x \in \mathbb{R}^n$. Τότε,

$$\varphi_k^*(y_0) = \sup_{x\in X}[x\cdot y_0 - \varphi_k(x)] \leq \sup_{x\in X}[x\cdot y_0 - x\cdot y_0 + \psi_k(y_0)] = \psi_k(y_0) < +\infty.$$

Άρα, η φ_k^* δεν είναι ταυτοτικά ίση με $+\infty$. Ακόμα, εξ ορισμού ισχύει

$$\varphi_k^*(y) \geq x_0\cdot y - \varphi_k(x_0), \qquad y \in \mathbb{R}^n.$$

Ορίζουμε την ακολουθία των πραγματικών αριθμών

$$a_k := \inf_{y\in Y}\left(\varphi_k^*(y) + \frac{|y|^2}{2}\right), \qquad k \in \mathbb{N}.$$

Για κάθε $k \in \mathbb{N}$, έχουμε:

$$\inf_{y\in Y}\left(\varphi_k^*(y)+\frac{|y|^2}{2}\right)\le \varphi_k^*(y_0)+\frac{|y_0|^2}{2}<+\infty$$

και

$$\inf_{y\in Y}\left(\varphi_k^*(y)+\frac{|y|^2}{2}\right)\ge \inf_{y\in Y}\left(x_0\cdot y+\frac{|y|^2}{2}\right)-\varphi_k(x_0)\ge \inf_{y\in\mathbb{R}^n}\left(x_0\cdot y+\frac{|y|^2}{2}\right)-\varphi_k(x_0)$$
$$\ge \inf_{y\in\mathbb{R}^n}\left(\frac{|y|^2}{2}-|x_0\cdot y|\right)-\varphi_k(x_0)\ge \inf_{y\in\mathbb{R}^n}\left(\frac{|y|^2}{2}-|x_0||y|\right)-\varphi_k(x_0)$$
$$\ge \inf_{y\in\mathbb{R}^n}\left(|y|\left(\frac{|y|}{2}-|x_0|\right)\right)-\varphi_k(x_0)>-\infty,$$

αφού έχουμε συνεχή συνάρτηση με όριο $+\infty$ για $|y|\to+\infty$.

Δηλαδή για κάθε $k \in \mathbb{N}$, ο a_k είναι, όντως, πραγματικός αριθμός. Για κάθε $k \in \mathbb{N}$, ορίζουμε

$$(\overline{\varphi}_k,\overline{\psi}_k)=(\varphi_k^{**}+a_k,\varphi_k^*-a_k).$$

Ισχύει

$$(\overline{\psi}_k)^*(x)=(\varphi_k^*-a_k)^*(x)=\sup_{y\in Y}[x\cdot y-\varphi_k^*(y)]+a_k=\varphi_k^{**}(x)+a_k=\overline{\varphi}_k(x).$$

Εκ κατασκευής της ακολουθίας a_k έχουμε

$$\inf_{y\in Y}\left(\overline{\psi}_k(y)+\frac{|y|^2}{2}\right)=\inf_{y\in Y}\left(\varphi_k^*(y)+\frac{|y|^2}{2}-a_k\right)=\inf_{y\in Y}\left(\varphi_k^*(y)+\frac{|y|^2}{2}\right)-a_k=0,$$

άρα για κάθε $y\in Y$,

$$\overline{\psi}_k(y)\ge-\frac{|y|^2}{2}.$$

Επιπλέον,

$$\overline{\varphi}_k(x)+\frac{|x|^2}{2}=(\overline{\psi}_k)^*(x)+\frac{|x|^2}{2}=\sup_{y\in Y}\left[x\cdot y+\frac{|x|^2}{2}-\overline{\psi}_k(y)\right]\ge\sup_{y\in Y}\left[-\frac{|y|^2}{2}-\overline{\psi}_k(y)\right]$$
$$=-\inf_{y\in Y}\left[\overline{\psi}_k(y)+\frac{|y|^2}{2}\right]=0$$

(η τελευταία ανισότητα ισχύει, αφού $|x|^2+|y|^2+2x\cdot y\ge0$) δηλαδή

$$\overline{\varphi}_k(x)\ge-\frac{|x|^2}{2}.$$

Αποδείξαμε, λοιπόν, ότι η ακολουθία $\left(\overline{\varphi}_k, \overline{\psi}_k\right)$ ικανοποιεί τη σχέση (2.37).

<u>Βήμα 2ο</u>: Θα αποδείξουμε ότι $\left(\overline{\varphi}_k, \overline{\psi}_k\right) \in \widetilde{\Phi}$ και $\lim\limits_{k\to\infty} J\left(\overline{\varphi}_k, \overline{\psi}_k\right) = \inf\limits_{\widetilde{\Phi}} J$.

Όπως περιγράψαμε και στη μέθοδο της διπλής κυρτοποίησης της υποενότητας 2.1.2,

$$J\left(\overline{\varphi}_k, \overline{\psi}_k\right) = J(\varphi_k^{**} + a_k, \varphi_k^{*} - a_k) = J(\varphi_k^{**}, \varphi_k^{*}) \leq J(\varphi_k, \psi_k) < +\infty. \tag{2.40}$$

Ειδικότερα, $J\left(\overline{\varphi}_k, \overline{\psi}_k\right) + M_2 < +\infty$, το οποίο μπορούμε να γράψουμε ως

$$\int_X \left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) d\mu(x) + \int_Y \left(\overline{\psi}_k(y) + \frac{|y|^2}{2}\right) d\nu(y) < +\infty.$$

Εφόσον οι δυο υπό ολοκλήρωση συναρτήσεις είναι μη αρνητικές, συμπεραίνουμε ότι είναι και ολοκληρώσιμες αναφορικά με τα μέτρα πιθανότητας μ, ν αντίστοιχα. Άρα, και οι $\overline{\varphi}_k$, $\overline{\psi}_k$ είναι ολοκληρώσιμες (ως προς τα μέτρα μ και ν αντίστοιχα) ως διαφορές ολοκληρώσιμων συναρτήσεων.

Εξ ορισμού, για κάθε $k \in \mathbb{N}$ και $x, y \in \mathbb{R}^n$ έχουμε $\overline{\varphi}_k(x) + \overline{\psi}_k(y) = \varphi_k^{**}(x) + \varphi_k^{*}(y) \geq x \cdot y$,

άρα $\left(\overline{\varphi}_k, \overline{\psi}_k\right) \in \widetilde{\Phi}$ και επίσης, σύμφωνα με τη σχέση (2.40), ισχύει $\lim\limits_{k\to\infty} J\left(\overline{\varphi}_k, \overline{\psi}_k\right) = \inf\limits_{\widetilde{\Phi}} J$.

<u>Βήμα 3ο</u>: Απόδειξη των σχέσεων (2.38) και (2.39).

Παρατηρούμε ότι

$$\begin{aligned} J\left(\overline{\varphi}_k, \overline{\psi}_k\right) + M_2 &= \int_X \left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) d\mu(x) + \int_Y \left(\overline{\psi}_k(y) + \frac{|y|^2}{2}\right) d\nu(y) \\ &\geq \inf_{x\in X}\left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) + \inf_{y\in Y}\left(\overline{\psi}_k(y) + \frac{|y|^2}{2}\right). \end{aligned}$$

Εφόσον οι συναρτήσεις που βρίσκονται μέσα στις παρενθέσεις είναι μη αρνητικές, προκύπτει ότι

$$\liminf_{k\to\infty} \inf_{x\in X}\left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) \leq \liminf_{k\to\infty} J\left(\overline{\varphi}_k, \overline{\psi}_k\right) + M_2 = \inf_{\widetilde{\Phi}} J + M_2.$$

Ομοίως, δείχνουμε ότι

$$\liminf_{k\to\infty} \inf_{y\in Y}\left(\overline{\psi}_k(y) + \frac{|y|^2}{2}\right) \leq \inf_{\widetilde{\Phi}} J + M_2.$$

Έχοντας ολοκληρώσει την απόδειξη του μέρους (ii) του λήμματος, το μέρος (iii) αποτελεί άμεση συνέπεια αυτού, αφού $\overline{\varphi}_k = (\varphi_k + \alpha_k)^{**}$, $\overline{\psi}_k = (\varphi_k + \alpha_k)^{*}$.

■

Απόδειξη του Θεωρήματος 2.35 στην περίπτωση στην οποία $Supp(\mu) \subseteq X$ και $Supp(\nu) \subseteq Y$, όπου X, Y είναι συμπαγή υποσύνολα του $\mathbb{R}^n$:

Μπορούμε να υποθέσουμε ότι τα X, Y είναι κυρτά, π.χ μπάλες. Θεωρούμε $\{(\varphi_k, \psi_k)\}_{k\in\mathbb{N}}$ μια ακολουθία στο $\widetilde{\Phi}$ με την ιδιότητα $\lim_{k\to\infty} J(\varphi_k, \psi_k) = \inf_{\widetilde{\Phi}} J$.

Θεωρούμε, ακόμα, $\{(\overline{\varphi}_k, \overline{\psi}_k)\}_{k\in\mathbb{N}}$ την ακολουθία που κατασκευάσαμε στο λήμμα 2.36, η οποία επίσης έχει την ιδιότητα $\lim_{k\to\infty} J(\overline{\varphi}_k, \overline{\psi}_k) = \inf_{\widetilde{\Phi}} J$.

Για κάθε $y \in \mathbb{R}^n$, έχουμε $\overline{\psi}_k(y) = \varphi_k^*(y) - a_k = \sup_{x\in X}[x \cdot y - \varphi_k(x)] - a_k$, οπότε για κάθε $y_1, y_2 \in \mathbb{R}^n$, έχουμε

$$\left|\overline{\psi}_k(y_1) - \overline{\psi}_k(y_2)\right| = \left|\sup_{x\in X}[x \cdot y_1 - \varphi_k(x)] - \sup_{x\in X}[x \cdot y_2 - \varphi_k(x)]\right| \leq \sup_{x\in X}|x \cdot (y_1 - y_2)|$$
$$\leq \sup_{x\in X}[|x||y_1 - y_2|] = \left(\sup_{x\in X}|x|\right)|y_1 - y_2|.$$

Η ανισότητα ισχύει επειδή

$$\sup_{x\in X}[x \cdot (y_1 - y_2)] + \sup_{x\in X}[x \cdot y_2 - \varphi_k(x)] \geq \sup_{x\in X}[x \cdot (y_1 - y_2) + x \cdot y_2 - \varphi_k(x)]$$
$$= \sup_{x\in X}[x \cdot y_1 - \varphi_k(x)].$$

Εφόσον το σύνολο X είναι συμπαγές, είναι και φραγμένο οπότε $\sup_{x\in X}|x| < +\infty$. Επομένως, η συνάρτηση $\overline{\psi}_k$ είναι Lipschitz και $\left\|\overline{\psi}_k\right\|_{Lip(Y)} \leq \sup_{x\in X}|x|$. Ομοίως, δείχνουμε ότι η συνάρτηση $\overline{\varphi}_k$ είναι Lipschitz και $\left\|\overline{\varphi}_k\right\|_{Lip(X)} \leq \sup_{y\in Y}|y|$. Συμπεραίνουμε ότι οι συναρτήσεις $\overline{\varphi}_k$ και $\overline{\psi}_k$ είναι ομοιόμορφα (ως προς k) Lipschitz.

Από την άλλη, όπως αποδείξαμε στο λήμμα 2.36,

$$\liminf_{k\to\infty} \inf_{x\in X}\left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) \leq \inf_{\widetilde{\Phi}} J + M_2.$$

Άρα, υπάρχουν οσοδήποτε μεγάλα k έτσι ώστε

$$\inf_{x\in X}\left(\overline{\varphi}_k(x) + \frac{|x|^2}{2}\right) < \inf_{\widetilde{\Phi}} J + M_2 + 1,$$

δηλαδή, για άπειρα k υπάρχουν $x_k \in X$ τέτοια ώστε

$$0 \leq \overline{\varphi}_k(x_k) + \frac{|x_k|^2}{2} \leq \inf_{\widetilde{\Phi}} J + M_2 + 1.$$

Από την ανισότητα στα αριστερά, παίρνουμε

$$\overline{\varphi}_k(x_k) \geq -\frac{|x_k|^2}{2} \geq -\sup_{x\in X}\frac{|x|^2}{2},$$

ενώ από την ανισότητα στα δεξιά, προκύπτει

$$\overline{\varphi}_k(x_k) \leq -\frac{|x_k|^2}{2} + \inf_{\widetilde{\Phi}} J + M_2 + 1 \leq \sup_{x\in X}\frac{|x|^2}{2} + \inf_{\widetilde{\Phi}} J + M_2 + 1.$$

Περιοριζόμαστε σε υπακολουθία της $(\overline{\varphi}_k)$ στην οποία ισχύουν οι παραπάνω ανισότητες. Με περαιτέρω περιορισμό σε υπακολουθία επιτυγχάνουμε ανάλογες ανισότητες για τις $\overline{\psi}_k$. Στη συνέχεια, στη θέση της ακολουθίας $(\overline{\varphi}_k, \overline{\psi}_k)$ θεωρούμε την υπακολουθία στην οποία έχουμε περιορισθεί.

Ισχυρισμός 11. Οι συναρτήσεις $\overline{\varphi}_k$ και $\overline{\psi}_k$ είναι ομοιόμορφα (ως προς k) φραγμένες.

Απόδειξη του Ισχυρισμού 11:

Για κάθε $x \in X$ ισχύει

$$\begin{aligned}\left|\overline{\varphi}_k(x)\right| &\leq \left|\overline{\varphi}_k(x_k) - \overline{\varphi}_k(x)\right| + \left|\overline{\varphi}_k(x_k)\right| \\ &\leq \sup_{y\in Y}|y| \cdot diam(X) + \max\left\{\sup_{x\in X}\frac{|x|^2}{2}, \sup_{x\in X}\frac{|x|^2}{2} + \inf_{\widetilde{\Phi}} J + M_2 + 1\right\}.\end{aligned}$$

■

Στη συνέχεια, εφόσον έχουμε αποδείξει ότι οι συναρτήσεις $\overline{\varphi}_k$ και $\overline{\psi}_k$ είναι ομοιόμορφα φραγμένες και ομοιόμορφα Lipschitz, βλέπουμε ότι ικανοποιούνται οι υποθέσεις του Θεωρήματος Ascoli-Arzela. Σύμφωνα με το θεώρημα αυτό, υπάρχουν υπακολουθίες $\left(\overline{\varphi}_{k_l}\right)_{l\in\mathbb{N}}$ της $\left(\overline{\varphi}_k\right)_{k\in\mathbb{N}}$ και $\left(\overline{\psi}_{k_l}\right)_{l\in\mathbb{N}}$ της $\left(\overline{\psi}_k\right)_{k\in\mathbb{N}}$ και συναρτήσεις $\overline{\varphi} \in C_b(X)$, $\overline{\psi} \in C_b(Y)$ έτσι ώστε

$$\sup_{x\in X}\left|\overline{\varphi}_{k_l}(x) - \overline{\varphi}(x)\right| \longrightarrow 0, \qquad \text{καθώς } l \to +\infty$$

και

$$\sup_{y\in Y}\left|\overline{\psi}_{k_l}(y) - \overline{\psi}(y)\right| \longrightarrow 0, \qquad \text{καθώς } l \to +\infty.$$

Έχουμε τότε

$$\left|J\left(\overline{\varphi}_{k_l},\overline{\psi}_{k_l}\right)-J(\overline{\varphi},\overline{\psi})\right| \leq \int_X \left|\overline{\varphi}_{k_l}(x)-\overline{\varphi}(x)\right| d\mu(x)+\int_Y \left|\overline{\psi}_{k_l}(y)-\overline{\psi}(y)\right| d\nu(y)$$
$$\leq \sup_{x\in X}\left|\overline{\varphi}_{k_l}(x)-\overline{\varphi}(x)\right|+\sup_{y\in Y}\left|\overline{\psi}_{k_l}(y)-\overline{\psi}(y)\right| \longrightarrow 0$$

δηλαδή $\lim_{l\to+\infty} J\left(\overline{\varphi}_{k_l},\overline{\psi}_{k_l}\right)=J(\overline{\varphi},\overline{\psi})$.

Με άλλα λόγια, το ζεύγος $(\overline{\varphi},\overline{\psi}) \in C_b(X)\times C_b(Y)$ (το οποίο προς το παρόν είναι ορισμένο μόνο πάνω στο σύνολο $X\times Y$) είναι βέλτιστο αναφορικά με την ελαχιστοποίηση πάνω στο $\widetilde{\Phi}$ του συναρτησοειδούς J. Επιπλέον, οι $\overline{\varphi},\overline{\psi}$ είναι κυρτές ως όρια κυρτών συναρτήσεων.

Στη συνέχεια, επεκτείνουμε τις συναρτήσεις $\overline{\varphi}$ και $\overline{\psi}$ έξω από τα σύνολα X και Y ορίζοντας

$$\overline{\varphi}(x)=+\infty,\ \ x\in\mathbb{R}^n-X \quad \text{και} \quad \overline{\psi}(y)=+\infty,\ \ y\in\mathbb{R}^n-Y.$$

Οι $\overline{\varphi},\overline{\psi}$ είναι, τότε, κυρτές και κάτω ημισυνεχείς συναρτήσεις. Τέλος, εφαρμόζουμε διπλή κυρτοποίηση στο ζεύγος $(\overline{\varphi},\overline{\psi})$. Ορίζουμε τις συναρτήσεις

$$\overline{\varphi}^*(y) \coloneqq \sup_{x\in\mathbb{R}^n}[x\cdot y-\overline{\varphi}(x)],\ \ y\in\mathbb{R}^n \quad \text{και} \quad \overline{\varphi}^{**}(x)\coloneqq \sup_{y\in\mathbb{R}^n}\left[x\cdot y-\overline{\varphi}^*(y)\right],\ \ x\in\mathbb{R}^n.$$

Τότε, σύμφωνα με την Πρόταση 2.21, οι $\overline{\varphi}^*$ και $\overline{\varphi}^{**}$ είναι κάτω ημισυνεχείς, κανονικές, συζυγείς κυρτές συναρτήσεις ορισμένες στον $\mathbb{R}^n$ και $J\left(\overline{\varphi}^{**},\overline{\varphi}^*\right)\leq J(\overline{\varphi},\overline{\psi})=\inf_{\widetilde{\Phi}} J$, άρα $J\left(\overline{\varphi}^{**},\overline{\varphi}^*\right)=\inf_{\widetilde{\Phi}} J$.

■

2.1.5. Το Θεώρημα Βέλτιστης Μεταφοράς.

Θεώρημα 2.37 (Το Θεώρημα βέλτιστης μεταφοράς για τετραγωνική συνάρτηση κόστους)

Έστω μ,ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$, με πεπερασμένες ροπές 2ης τάξης, κάτι που εκφράστηκε μέσω της σχέσης (2.1) ως εξής:

$$M_2 \coloneqq \int_{\mathbb{R}^n}\frac{|x|^2}{2}d\mu(x)+\int_{\mathbb{R}^n}\frac{|y|^2}{2}d\nu(y)<+\infty.$$

Θεωρούμε το πρόβλημα μεταφοράς μάζας του Kantorovich αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x,y)=|x-y|^2$.

Τότε, ισχύουν τα εξής:

(i). Το κριτήριο βελτιστοποίησης Knott – Smith.

Το $\pi \in \Pi(\mu,\nu)$ είναι βέλτιστο πλάνο μεταφοράς αν και μόνο αν υπάρχει μια κάτω ημισυνεχής κυρτή συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R}$ τέτοια ώστε:

$$Supp(\pi) \subseteq Graph(\partial\varphi), \tag{2.41}$$

ή ισοδύναμα, (επειδή το $Graph(\partial\varphi)$ είναι κλειστό σύνολο)

$$y \in \partial\varphi(x), \quad d\pi - \text{σχεδόν για όλα τα } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n. \tag{2.42}$$

Επιπλέον, το ζεύγος (φ,φ^*) πετυχαίνει την ελαχιστοποίηση πάνω στο $\widetilde{\Phi}$ του συναρτησοειδούς $J(\varphi,\psi)$, δηλαδή $J(\varphi,\varphi^*) = \inf_{(\varphi,\psi)\in\widetilde{\Phi}} J(\varphi,\psi)$.

(ii). Το Θεώρημα Brenier.

Αν υποθέσουμε ότι το μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue, τότε το βέλτιστο μέτρο π είναι μοναδικό. Μάλιστα κατασκευάζεται ως εξής:

$$d\pi(x,y) = d\mu(x)\delta[y = \nabla\varphi(x)], \tag{2.43}$$

ή ισοδύναμα

$$\pi = (Id \times \nabla\varphi)\#\mu, \tag{2.44}$$

όπου $\nabla\varphi$ είναι η μοναδική (δηλαδή με μοναδικό τρόπο καθορισμένη $d\mu$ −σχεδόν παντού) παράγωγος μιας κυρτής συνάρτησης, η οποία μεταφέρει το μ πάνω στο ν, δηλαδή $\nabla\varphi\#\mu = \nu$.

Επιπλέον, $Supp(\nu) = \overline{\nabla\varphi\big(Supp(\mu)\big)}$.

(iii). Αν, επίσης, το μέτρο ν δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue, τότε για $d\mu$ −σχεδόν όλα τα $x \in \mathbb{R}^n$ και $d\nu$ −σχεδόν όλα τα $y \in \mathbb{R}^n$, ισχύει:

$$(\nabla\varphi^* \circ \nabla\varphi)(x) = x, \qquad (\nabla\varphi \circ \nabla\varphi^*)(y) = y,$$

και η $\nabla\varphi^*$ είναι η ($d\nu$ −σχεδόν παντού) μοναδική παράγωγος μιας κυρτής συνάρτησης η οποία μεταφέρει το ν πάνω στο μ.

<u>Απόδειξη:</u>

Εφόσον η συνάρτηση $c(x,y)$ είναι μια τετραγωνική συνάρτηση κόστους, σύμφωνα με όσα αναφέραμε στην υποενότητα 2.1.2, το θεώρημα δυϊκότητας του Kantorovich μετατρέπεται στη σχέση (2.7), δηλαδή

$$\sup_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi(x,y)=\inf_{(\varphi,\psi)\in\widetilde{\Phi}}J(\varphi,\psi).$$

Βήμα 1ο: Απόδειξη του (i).

Από την Πρόταση 2.1 γνωρίζουμε ότι υπάρχει βέλτιστο πλάνο μεταφοράς $\pi_*\in\Pi(\mu,\nu)$. Από το Θεώρημα 2.35, γνωρίζουμε ότι υπάρχει ζεύγος συζυγών κυρτών συναρτήσεων (φ,φ^*), το οποίο είναι βέλτιστο για το δυϊκό πρόβλημα της ελαχιστοποίησης του συναρτησοειδούς J πάνω στο $\widetilde{\Phi}$.

Από τη σχέση (2.7) και τον ορισμό του συνόλου $\Pi(\mu,\nu)$ έχουμε:

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi_*(x,y)=J(\varphi,\varphi^*)=\int_{\mathbb{R}^n}\varphi\,d\mu+\int_{\mathbb{R}^n}\varphi^*\,d\nu=\int_{\mathbb{R}^n\times\mathbb{R}^n}[\varphi(x)+\varphi^*(y)]\,d\pi_*(x,y).$$

Ισοδύναμα,

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}[\varphi(x)+\varphi^*(y)-x\cdot y]\,d\pi_*(x,y)=0.$$

Εφόσον από τη σχέση (2.25) γνωρίζουμε ότι η υπό ολοκλήρωση συνάρτηση είναι μη αρνητική και το π^* είναι ένα μη αρνητικό μέτρο, συμπεραίνουμε ότι

$$\varphi(x)+\varphi^*(y)=x\cdot y,\qquad d\pi_*-\text{σχεδόν παντού στον }\mathbb{R}^n\times\mathbb{R}^n.$$

Από την Πρόταση 2.22, αυτό ισοδυναμεί με τη σχέση (2.42):

$$y\in\partial\varphi(x),\qquad d\pi_*-\text{σχεδόν για όλα τα }(x,y)\in\mathbb{R}^n\times\mathbb{R}^n.$$

Αντίστροφα, έστω $\pi_*\in\Pi(\mu,\nu)$ και κυρτή συνάρτηση φ έτσι ώστε να ικανοποιείται η σχέση

$$y\in\partial\varphi(x),\qquad d\pi_*-\text{σχεδόν για όλα τα }(x,y)\in\mathbb{R}^n\times\mathbb{R}^n.$$

Από την Πρόταση 2.22 και πάλι, αυτό ισοδυναμεί με τη σχέση

$$d\pi_*-\text{σχεδόν για κάθε }x,y\in\mathbb{R}^n,\qquad \varphi(x)+\varphi^*(y)=x\cdot y,$$

όπου φ^* είναι ο μετασχηματισμός Legendre της φ. Άρα, όπως και πριν

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi_*(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}[\varphi(x)+\varphi^*(y)]\,d\pi_*(x,y)=\int_{\mathbb{R}^n}\varphi\,d\mu+\int_{\mathbb{R}^n}\varphi^*\,d\nu.$$

Λόγω της σχέσης (2.7), η τελευταία ισότητα σημαίνει αφ' ενός ότι το π_* είναι βέλτιστο όσον αφορά το πρόβλημα της μεγιστοποίησης πάνω στο $\Pi(\mu,\nu)$ του συναρτησοειδούς

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi(x,y)$$

και αφ' ετέρου ότι το ζεύγος (φ,φ^*) είναι βέλτιστο αναφορικά με την ελαχιστοποίηση πάνω στο $\widetilde{\Phi}$ του συναρτησοειδούς $J(\varphi,\psi)$. Έτσι ολοκληρώνεται η απόδειξη του (i).

Βήμα 2ο: Απόδειξη του (ii).

Τώρα, υποθέτουμε ότι το μ δεν δίνει μάζα στα μικρά σύνολα. Θεωρούμε ένα οποιοδήποτε βέλτιστο πλάνο μεταφοράς $\pi_* \in \Pi(\mu,\nu)$ καθώς και τη συνάρτηση φ που προκύπτει από το (i).

Εφόσον $\varphi \in L^1(d\mu)$, θα ισχύει $\varphi(x) < +\infty$, $d\mu -$ σχεδόν παντού στον $\mathbb{R}^n$, άρα $\mu[Dom(\varphi)] = 1$.

Από την άλλη, σύμφωνα με την Πρόταση 2.11, το σύνορο $\partial Dom(\varphi)$ του κυρτού συνόλου $Dom(\varphi)$ έχει μέτρο Lebesgue ίσο με 0, οπότε $\mu\big[Int\big(Dom(\varphi)\big)\big] = 1$.

Στη συνέχεια, εφόσον η φ είναι κυρτή, σύμφωνα με το Θεώρημα 2.32 το σύνολο των σημείων του $Int\big(Dom(\varphi)\big)$ στα οποία δεν είναι διαφορίσιμη είναι ένα μικρό σύνολο. Άρα, η φ είναι διαφορίσιμη $d\mu$ −σχεδόν παντού στον $\mathbb{R}^n$. Δηλαδή, σύμφωνα με την Πρόταση 2.18

$$\partial\varphi(x) = \{\nabla\varphi(x)\}, \qquad d\mu - \text{σχεδόν για όλα τα } x \in \mathbb{R}^n.$$

Υπάρχει, επομένως, μετρήσιμο σύνολο $A \subseteq \mathbb{R}^n$ με $\mu[\mathbb{R}^n - A] = 0$ τέτοιο ώστε

$$\partial\varphi(x) = \{\nabla\varphi(x)\}, \qquad \text{για κάθε } x \in A.$$

Επίσης, από το (i) υπάρχει μετρήσιμο σύνολο $B \subseteq \mathbb{R}^n \times \mathbb{R}^n$ με $\pi[\mathbb{R}^n \times \mathbb{R}^n - B] = 0$ τέτοιο ώστε

$$y \in \partial\varphi(x), \qquad \text{για κάθε } (x,y) \in B.$$

Τότε, ισχύει $\pi[B \cap (A \times \mathbb{R}^n)] = 1$ και για κάθε $(x,y) \in B \cap (A \times \mathbb{R}^n)$

$$y \in \partial\varphi(x) \quad \text{και} \quad \partial\varphi(x) = \{\nabla\varphi(x)\}.$$

Άρα, $y = \nabla\varphi(x)$, $d\pi -$ σχεδόν παντού στον $\mathbb{R}^n \times \mathbb{R}^n$.

Έως τώρα έχουμε αποδείξει ότι υπάρχει τουλάχιστον ένα βέλτιστο πλάνο μεταφοράς $\pi \in \Pi(\mu,\nu)$ και ότι κάθε τέτοιο π γράφεται στη μορφή $\pi = (Id \times \nabla\varphi)\#\mu$ για κάποια κυρτή συνάρτηση φ τέτοια ώστε $\nabla\varphi\#\mu = \nu$. Θα δείξουμε τώρα τη μοναδικότητα αυτής της συνάρτησης.

Έστω, λοιπόν, μια οποιαδήποτε άλλη κυρτή συνάρτηση $\overline{\varphi}$ τέτοια ώστε $\nabla\overline{\varphi}\#\mu = \nu$. Θα δείξουμε ότι $\nabla\varphi = \nabla\overline{\varphi}$, $d\mu -$ σχεδόν παντού στον $\mathbb{R}^n$. Από την υπόθεση για την $\overline{\varphi}$ έπεται ότι η $\nabla\overline{\varphi}$ ορίζεται $d\mu$ −σχεδόν παντού, άρα το $(Id \times \nabla\overline{\varphi}) \# \mu$ ανήκει στο $\Pi(\mu,\nu)$. Από το (i), γνωρίζουμε ότι το $(Id \times \nabla\overline{\varphi})\#\mu$ είναι ένα βέλτιστο πλάνο μεταφοράς και το ζεύγος $\big(\overline{\varphi},\overline{\varphi}^*\big)$ είναι βέλτιστο για το δυϊκό πρόβλημα, όπως ακριβώς και το (φ,φ^*). Άρα, $J\big(\overline{\varphi},\overline{\varphi}^*\big) = J(\varphi,\varphi^*) = \inf_{\Phi} J$, οπότε

$$\int_{\mathbb{R}^n} \overline{\varphi}\, d\mu + \int_{\mathbb{R}^n} \overline{\varphi}^*\, d\nu = \int_{\mathbb{R}^n} \varphi\, d\mu + \int_{\mathbb{R}^n} \varphi^*\, d\nu.$$

Έστω $\pi_* \in \Pi(\mu,\nu)$ το βέλτιστο πλάνο μεταφοράς που αντιστοιχεί στην φ, δηλαδή $\pi_* = (Id \times \nabla\varphi)\#\mu$. Τότε, από την παραπάνω ισότητα, τον ορισμό του $\Pi(\mu,\nu)$ και την Πρόταση 2.18, έχουμε:

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}[\overline{\varphi}(x)+\overline{\varphi}^*(y)]\,d\pi_*(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}[\varphi(x)+\varphi^*(y)]\,d\pi_*(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,d\pi_*(x,y).$$

Εφόσον $\pi_* = (Id\times\nabla\varphi)\#\mu$, η παραπάνω ισότητα ξαναγράφεται ως εξής:

$$\int_{\mathbb{R}^n}[\overline{\varphi}(x)+\overline{\varphi}^*(y)]\circ(Id\times\nabla\varphi)(x)\,d\mu(x)=\int_{\mathbb{R}^n}(x\cdot y)\circ(Id\times\nabla\varphi)(x)\,d\mu(x),$$

δηλαδή

$$\int_{\mathbb{R}^n}\left[\overline{\varphi}(x)+\overline{\varphi}^*\big(\nabla\varphi(x)\big)\right]d\mu(x)=\int_{\mathbb{R}^n}\big(x\cdot\nabla\varphi(x)\big)\,d\mu(x).$$

Άρα,

$$\int_{\mathbb{R}^n}\left[\overline{\varphi}(x)+\overline{\varphi}^*\big(\nabla\varphi(x)\big)-x\cdot\nabla\varphi(x)\right]d\mu(x)=0.$$

Αφού και πάλι, σύμφωνα με τη σχέση (2.25), η υπό ολοκλήρωση συνάρτηση είναι μη αρνητική και το μ είναι ένα μη αρνητικό μέτρο, θα ισχύει

$$\overline{\varphi}(x)+\overline{\varphi}^*\big(\nabla\varphi(x)\big)=x\cdot\nabla\varphi(x),\qquad d\mu-\text{σχεδόν παντού στον }\mathbb{R}^n.$$

Από την Πρόταση 2.22, παίρνουμε $\nabla\varphi(x)=\nabla\overline{\varphi}(x)$, $d\mu -$ σχεδόν για όλα τα $x\in\mathbb{R}^n$.

Σημειώνουμε ότι εκτός από τη μοναδικότητα του βέλτιστου πλάνου μεταφοράς για το πρόβλημα Monge – Kantorovich, αποδείξαμε και τη μοναδικότητα ($d\mu$ –σχεδόν παντού) της παραγώγου μιας κυρτής συνάρτησης που μεταφέρει το μέτρο μ πάνω στο ν.

Έτσι ολοκληρώνεται η απόδειξη του (ii), εκτός από τον τελευταίο ισχυρισμό.

<u>Βήμα 3ο</u>: Απόδειξη της ισότητας $Supp(\nu)=\overline{\nabla\varphi\big(Supp(\mu)\big)}$.

Έστω $x\in Supp(\mu)$ ένα σημείο στο οποίο η φ είναι διαφορίσιμη. Θέτουμε $y=\nabla\varphi(x)$. Από την Πρόταση 2.27, γνωρίζουμε ότι $\forall\varepsilon>0,\ \exists\delta>0\quad$ τ.ω $\quad\nabla\varphi\big(B_\delta(x)\big)\subseteq B_\varepsilon(y)$.

Ειδικότερα, $\nu[B_\varepsilon(y)]\geq\nu\big[\nabla\varphi\big(B_\delta(x)\big)\big]=\mu\left[(\nabla\varphi)^{-1}\left(\nabla\varphi\big(B_\delta(x)\big)\right)\right]\geq\mu[B_\delta(x)]$.

Εφόσον $x\in Supp(\mu)$, έχουμε $\mu[B_\delta(x)]>0$, άρα και $\nu[B_\varepsilon(y)]>0$. Το ε ήταν ένας αυθαίρετος θετικός αριθμός, άρα συμπεραίνουμε ότι $y\in Supp(\nu)$. Επομένως, $\nabla\varphi\big(Supp(\mu)\big)\subseteq Supp(\nu)$ και εφόσον εξ ορισμού ο φορέας ενός μέτρου είναι ένα κλειστό σύνολο, έχουμε $\overline{\nabla\varphi\big(Supp(\mu)\big)}\subseteq Supp(\nu)$.

Για την απόδειξη του αντίστροφου εγκλεισμού, έχουμε

$$\nu\big[\nabla\varphi\big(Supp(\mu)\big)\big]=\mu\left[(\nabla\varphi)^{-1}\left(\nabla\varphi\big(Supp(\mu)\big)\right)\right]\geq\mu[Supp(\mu)]=1.$$

Άρα, το μέτρο ν είναι συγκεντρωμένο στο σύνολο $\nabla\varphi\big(Supp(\mu)\big)$, οπότε εξ ορισμού και πάλι του φορέα ενός μέτρου, έχουμε $Supp(\nu) \subseteq \overline{\nabla\varphi\big(Supp(\mu)\big)}$.

Βήμα 4ο: Απόδειξη του (iii).

Αν το $\pi_* \in \Pi(\mu,\nu)$ είναι βέλτιστο, γνωρίζουμε ότι

$$y = \nabla\varphi(x), \;\; d\pi - \text{σχεδόν παντού στον } \mathbb{R}^n \times \mathbb{R}^n,$$

το οποίο, σύμφωνα με το Πόρισμα 2.24, ισοδυναμεί με

$$x \in \partial\varphi^*(y), \qquad d\pi - \text{σχεδόν παντού στον } \mathbb{R}^n \times \mathbb{R}^n.$$

Όμως, εφόσον η φ^* είναι πεπερασμένη $d\nu -$ σχεδόν παντού στον $\mathbb{R}^n$, είναι επίσης διαφορίσιμη $d\nu -$ σχεδόν παντού στον $\mathbb{R}^n$. Άρα,

$$x = \nabla\varphi^*(y) = \nabla\varphi^*\big(\nabla\varphi(x)\big), \qquad d\pi - \text{σχεδόν παντού στον } \mathbb{R}^n \times \mathbb{R}^n.$$

Περνώντας στο περιθώριο μέτρο, βλέπουμε ότι ισχύει η σχέση

$$(\nabla\varphi^* \circ \nabla\varphi)(x) = x, \qquad d\mu - \text{σχεδόν παντού στον } \mathbb{R}^n.$$

Λόγω συμμετρίας, και η άλλη σχέση προκύπτει ομοίως.

■

Ορισμός 2.38

Υπό τις υποθέσεις του Θεωρήματος Brenier, η απεικόνιση $\nabla\varphi$ (μοναδικά ορισμένη $d\mu -$σχεδόν παντού) θα λέγεται απεικόνιση Brenier που μεταφέρει το μ πάνω στο ν.

Θεώρημα 2.39 (Ποσοτική εκδοχή του κριτηρίου βελτιστοποίησης Knott-Smith)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ με πεπερασμένες ροπές 2ης τάξης και έστω $\varepsilon > 0$. Θεωρούμε $\pi_* \in \Pi(\mu,\nu)$ και $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ μια κανονική κάτω ημισυνεχή κυρτή συνάρτηση τέτοια ώστε $(\varphi, \varphi^*) \in L^1(d\mu) \times L^1(d\nu)$ και

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} [\varphi(x) + \varphi^*(y) - x \cdot y]\, d\pi_*(x,y) \leq \varepsilon.$$

Τότε,

$$I[\pi_*] \leq \left(\inf_{\pi \in \Pi(\mu,\nu)} I[\pi]\right) + \varepsilon.$$

Απόδειξη:

Από τη δοσμένη σχέση, έχουμε

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} [\varphi(x) + \varphi^*(y)]\, d\pi_*(x,y) \leq \int_{\mathbb{R}^n\times\mathbb{R}^n} (x\cdot y)\, d\pi_*(x,y) + \varepsilon,$$

οπότε, αφού $\pi_* \in \Pi(\mu,\nu)$ και $(\varphi,\varphi^*) \in L^1(d\mu)\times L^1(d\nu)$, έχουμε

$$\int_{\mathbb{R}^n} \varphi(x)\, d\mu(x) + \int_{\mathbb{R}^n} \varphi^*(y)\, d\nu(y) \leq \int_{\mathbb{R}^n\times\mathbb{R}^n} (x\cdot y)\, d\pi_*(x,y) + \varepsilon,$$

δηλαδή

$$J(\varphi,\varphi^*) \leq \int_{\mathbb{R}^n\times\mathbb{R}^n} (x\cdot y)\, d\pi_*(x,y) + \varepsilon.$$

Στη συνέχεια, λόγω της δυϊκότητας Kantorovich και της σχέσης (2.6), για $c(x,y) = \frac{|x-y|^2}{2}$ έχουμε:

$$\begin{aligned}
\inf_{\Pi(\mu,\nu)} I &= \sup_{\Phi_c} J = M_2 - \inf_{\tilde{\Phi}} J \geq M_2 - J(\varphi,\varphi^*)\\
&\geq \int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y) - \int_{\mathbb{R}^n\times\mathbb{R}^n} (x\cdot y)\, d\pi_*(x,y) - \varepsilon\\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} \left(\frac{|x|^2}{2} + \frac{|y|^2}{2} - x\cdot y\right) d\pi_*(x,y) - \varepsilon = \int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|x-y|^2}{2} d\pi_*(x,y) - \varepsilon\\
&= I[\pi_*] - \varepsilon.
\end{aligned}$$

Ισοδύναμα, $I[\pi_*] \leq \left(\inf_{\pi\in\Pi(\mu,\nu)} I[\pi]\right) + \varepsilon.$

■

2.2. Η πραγματική ευθεία.

Σε αυτή την ενότητα, θα μελετήσουμε την ιδιαίτερη περίπτωση του $\mathbb{R}$ εφοδιασμένο με την Borel σ −άλγεβρα.

Ορισμός 2.40

Ένα υποσύνολο Γ του $\mathbb{R}^2$ λέγεται μονότονο, αν ισχύει η συνεπαγωγή

(2.45) $$(x_1,y_1),(x_2,y_2) \in \Gamma \implies (x_1-x_2)(y_1-y_2) \geq 0.$$

Στο $\mathbb{R}$, οι παράγωγοι των κυρτών συναρτήσεων συμπίπτουν με τις αύξουσες συναρτήσεις. Ακόμα, το υποδιαφορικό μιας κυρτής συνάρτησης είναι ένα μέγιστο μονότονο υποσύνολο του $\mathbb{R}^2$.

Πράγματι, αν $\varphi: \mathbb{R} \longrightarrow \mathbb{R}$ είναι μια κυρτή συνάρτηση, το υποδιαφορικό στη θέση $x \in \mathbb{R}$ είναι

$$\partial\varphi(x) = \begin{cases} \{(x, \varphi'(x))\}, & \varphi \text{ παραγωγίσιμη στο } x \\ \{(x, a): \ \varphi'_-(x) \leq a \leq \varphi'_+(x)\}, & \varphi \text{ όχι παραγωγίσιμη στο } x \end{cases}$$

Έστω ότι το $Graph(\partial\varphi)$ είναι γνήσιο υποσύνολο του $A \subseteq \mathbb{R}^2$, όπου A μονότονο σύνολο. Τότε, υπάρχει $(x_0, a_0) \in A$ με $(x_0, a_0) \notin Graph(\partial\varphi)$, δηλαδή $a_0 \notin \partial\varphi(x_0)$. Μια κυρτή συνάρτηση στο $\mathbb{R}$ είναι παραγωγίσιμη σχεδόν παντού, άρα υπάρχουν ακολουθίες $x_n \uparrow x_0$ και $y_n \downarrow x_0$ με την φ να είναι παραγωγίσιμη στα σημεία x_n και y_n. Τότε $(x_n, \varphi'(x_n)) \in A$, $(y_n, \varphi'(y_n)) \in A$. Από την ανισότητα $x_n \leq x_0 \leq y_n$, λόγω μονοτονίας του συνόλου A, έχουμε $\varphi'(x_n) \leq a_0 \leq \varphi'(y_n)$. Επίσης, ισχύει $\varphi'(x_n) \rightarrow \varphi'_-(x_0)$, $\varphi'(y_n) \rightarrow \varphi'_+(x_0)$ ([13], Θεώρημα 24.1). Άρα, $\varphi'_-(x_0) \leq \alpha_0 \leq \varphi'_+(x_0)$, δηλαδή $a_0 \in \partial\varphi(x_0)$. Έχουμε καταλήξει σε αντίφαση, οπότε το $Graph(\partial\varphi)$ είναι μέγιστο.

Γεωμετρικά, ένα μονότονο υποσύνολο του $\mathbb{R}^2$ περιέχεται στο σύνολο που προκύπτει αν στο γράφημα μιας αύξουσας συνάρτησης προστεθούν κάποιες κάθετες γραμμές στα πιθανά σημεία ασυνέχειας, έτσι ώστε να κάνουν το γράφημα «συνεχές». Αν $\varphi: \mathbb{R} \longrightarrow \mathbb{R}$ είναι μια κυρτή συνάρτηση, οι γραμμές αυτές αντιστοιχούν στα σημεία $x \in \mathbb{R}$ για τα οποία η αριστερή και η δεξιά παράγωγος διαφέρουν. Μάλιστα, για κάθε $x \in \mathbb{R}$ ισχύει: $\partial\varphi(x) = [\varphi'_-(x), \varphi'_+(x)]$.

Πρόταση 2.41

Κάθε μονότονο σύνολο $A \subseteq \mathbb{R}^2$ περιέχεται στο γράφημα του υποδιαφορικού μιας κάτω ημισυνεχούς κυρτής συνάρτησης.

Απόδειξη:

1η περίπτωση: $A \subseteq \{x_0\} \times \mathbb{R}$.

Το $\{x_0\} \times \mathbb{R}$ είναι το γράφημα του υποδιαφορικού της συνάρτησης

$$\varphi(x) = \begin{cases} 0, & x = x_0 \\ +\infty, & x \neq 0 \end{cases}$$

2η περίπτωση: Το A περιέχει τουλάχιστον δύο σημεία με διαφορετική πρώτη συντεταγμένη.

Στην περίπτωση αυτή το A περιέχεται σε ένα σύνολο G, που προκύπτει με την εξής διαδικασία:

Θεωρούμε ένα κλειστό διάστημα I και μια αύξουσα συνάρτηση $f: I \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\}$, όπου:

Α) ως κλειστά διαστήματα θεωρούμε τα διαστήματα της μορφής $[a, b]$, $[a, +\infty)$, $(-\infty, a]$ ή $\mathbb{R}$,

Β) η f παίρνει πεπερασμένες τιμές στο $Int(I)$,

Γ) η f είναι συνεχής στα άκρα του I.

Θεωρούμε το σύνολο G που αποτελείται από α) τα σημεία του $Graph(f)$ που έχουν πεπερασμένες συντεταγμένες,

β) τα κατακόρυφα ευθύγραμμα τμήματα που αντιστοιχούν στα σημεία ασυνέχειας της f και έχουν άκρα τις αντίστοιχες πλευρικές τιμές της,

γ) αν η f έχει πεπερασμένη τιμή στο δεξιό άκρο του I, την κατακόρυφη ημιευθεία με άκρο την τιμή αυτή, και αντίστοιχα για το αριστερό άκρο του I.

Έστω $x_0 \in Int(I)$. Θεωρούμε τη συνάρτηση

$$\begin{cases} \int_{x_0}^{x} f(t)dt\,, & \text{αν η } f \text{ είναι ολοκληρώσιμη στο διάστημα με άκρα τα } x_0, x \\ +\infty, & \text{αλλού} \end{cases}$$

Η φ είναι κάτω ημισυνεχής κυρτή συνάρτηση και για κάθε $x \in Dom(\varphi)$ με $f(x) \in \mathbb{R}$ ισχύει $\partial\varphi(x) = \{y \in \mathbb{R} : \quad (x, y) \in G\}$. Έπεται ότι $A \subseteq Graph(\partial\varphi)$.

■

2.2.1. Αθροιστικές συναρτήσεις κατανομής.

Έστω μ ένα μέτρο πιθανότητας ορισμένο στο $\mathbb{R}$. Το μ αναπαρίσταται από την αθροιστική συνάρτηση κατανομής $F: \mathbb{R} \longrightarrow [0,1]$, που ορίζεται ως εξής:

$$F(x) = \int_{-\infty}^{x} d\mu = \mu\big[(-\infty, x]\big].$$

Σημειώνουμε ότι η F είναι αύξουσα, συνεχής από τα δεξιά και έχει όρια στο $-\infty$ και στο $+\infty$ ίσα με $F(-\infty) = 0$ και $F(+\infty) = 1$.
Η F δεν είναι κατά ανάγκην αντιστρέψιμη συνάρτηση, όμως μπορούμε να ορίσουμε τη γενικευμένη αντίστροφο της F, τη συνάρτηση $F^{-1}: [0,1] \longrightarrow \mathbb{R}$, ως εξής:

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : \quad F(x) > t\}.$$

Η F^{-1} είναι, επίσης, αύξουσα, συνεχής από δεξιά και ικανοποιεί τις παρακάτω ανισότητες:

$$\forall x \in \mathbb{R}, \qquad F^{-1}\big(F(x)\big) \geq x$$

$$\forall t \in [0,1], \qquad F\big(\,F^{-1}(t)\big) \geq t.$$

Ένα μέτρο πιθανότητας π ορισμένο πάνω στον χώρο γινόμενο $\mathbb{R} \times \mathbb{R}$ αναπαρίσταται με την από κοινού διδιάστατη αθροιστική συναρτήση κατανομής, $H: \mathbb{R}^2 \longrightarrow [0,1]$, η οποία για κάθε σημείο $(x_0, y_0) \in \mathbb{R}^2$ ορίζεται ως εξής:

$$H(x_0, y_0) = \int_{R(x_0,y_0)} d\pi = \pi[R(x_0, y_0)],$$

όπου $R(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 : \quad x \leq x_0 \quad \text{και} \quad y \leq y_0\}$.

Όταν το μέτρο π αναπαρίσταται από την αθροιστική συνάρτηση κατανομής H, χρησιμοποιούμε το συμβολισμό $d\pi = dH$.

Πρόταση 2.42

Έστω μια συνάρτηση $H: \mathbb{R}^2 \to [0, 1]$ που έχει τις ακόλουθες ιδιότητες:

Α) Είναι αύξουσα και συνεχής από δεξιά και ως προς τις δύο μεταβλητές x, y.

Β) Έχει όριο στο $(+\infty, +\infty)$ ίσο με $H(+\infty, +\infty) = 1$.

Γ) Ισχύει $\lim\limits_{y\to-\infty} H(x, y) = 0,\ \ x \in \mathbb{R}$ και $\lim\limits_{x\to-\infty} H(x, y) = 0,\ \ y \in \mathbb{R}$.

Δ) Για κάθε ορθογώνιο της μορφής $[x_1, x_2] \times [y_1, y_2] \subseteq \mathbb{R}^2$ ισχύει

$$H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \geq 0.$$

Τότε, η H αναπαριστά ένα μοναδικό μέτρο πιθανότητας στο $\mathbb{R}^2$.

Για την απόδειξη παραπέμπουμε στο ([19], Θεώρημα 1.1.6).

2.2.2. Το Θεώρημα βέλτιστης μεταφοράς στο $\mathbb{R}$.

Στην παρούσα υποενότητα, θα διατυπώσουμε τη λύση του προβλήματος βέλτιστης μεταφοράς με τη βοήθεια των αθροιστικών συναρτήσεων κατανομής.

Αν F είναι μια αθροιστική συνάρτηση κατανομής, τότε $F(x^-) \coloneqq \lim\limits_{z\to x^-} F(z)$, είναι το όριο από αριστερά της F στο x, το οποίο λόγω της μονοτονίας της F, πάντοτε υπάρχει. Ακόμα, ο φορέας ενός μέτρου μ ορισμένου στον χώρο X συμβολίζεται με $Supp(\mu)$ και είναι το μικρότερο κλειστό σύνολο τέτοιο ώστε $\mu[X - Supp(\mu)] = 0$. Αποδεικνύεται εύκολα ότι ένα τέτοιο σύνολο υπάρχει.

Θεώρημα 2.43 (Θεώρημα βέλτιστης μεταφοράς για τετραγωνική συνάρτηση κόστους στο $\mathbb{R}$)

Θεωρούμε μ, ν δύο μέτρα πιθανότητας στο $\mathbb{R}$, με πεπερασμένες ροπές 2ης τάξης. Συμβολίζουμε με F, G τις αντίστοιχες αθροιστικές συναρτήσεις κατανομής και έστω π το μέτρο πιθανότητας στον $\mathbb{R}^2$ που ορίζεται με την από κοινού διδιάστατη αθροιστική συνάρτηση κατανομής

(2.46) $$H(x, y) = \min\big(F(x), G(y)\big).$$

Θεωρούμε, ακόμα, το πρόβλημα μεταφοράς μάζας του Kantorovich αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x, y) = |x - y|^2$.

Τότε, $\pi \in \Pi(\mu,\nu)$ και μάλιστα είναι ένα βέλτιστο πλάνο μεταφοράς. Επιπλέον, το βέλτιστο κόστος μεταφοράς δίνεται από τη σχέση

$$T_2(\mu,\nu) = \int_0^1 |F^{-1}(t) - G^{-1}(t)|^2 \, dt.$$

Παρατήρηση: Η συνάρτηση H ικανοποιεί τις υποθέσεις της Πρότασης 2.42 και επομένως, ορίζεται μοναδικό μέτρο π, όπως στην Πρόταση 2.42.

Πράγματι, εύκολα επαληθεύουμε τις ιδιότητες Α, Β, Γ, οπότε χρειάζεται να ελέγξουμε μόνο την ιδιότητα Δ. Θεωρούμε ένα ορθογώνιο της μορφής $[x_1,x_2]\times[y_1,y_2]$ στον $\mathbb{R}^2$. Διακρίνουμε δύο περιπτώσεις.

Αν $F(x_1) \le G(y_1)$, τότε $H(x_2,y_2) - H(x_2,y_1) - H(x_1,y_2) + H(x_1,y_1) = H(x_2,y_2) - H(x_2,y_1) - F(x_1) + F(x_1) = H(x_2,y_2) - H(x_2,y_1) \ge 0$, αφού η συνάρτηση H είναι αύξουσα ως προς τη μεταβλητή y.

Αν $F(x_1) > G(y_1)$, τότε $H(x_2,y_2) - H(x_2,y_1) - H(x_1,y_2) + H(x_1,y_1) = H(x_2,y_2) - G(y_1) - H(x_1,y_2) + G(y_1) = H(x_2,y_2) - H(x_1,y_2) \ge 0$, αφού η συνάρτηση H είναι αύξουσα ως προς τη μεταβλητή x.

Απόδειξη του Θεωρήματος 2.43:

Καταρχήν, θα αποδείξουμε ότι το μέτρο π που ορίζεται από τη συνάρτηση H ανήκει στο $\Pi(\mu,\nu)$. Πράγματι, έστω ένα σύνολο της μορφής $A = (-\infty, x]$. Τότε,

$$\pi[A\times\mathbb{R}] = \pi[(-\infty,x]\times\mathbb{R}] = H(x,+\infty) = \min\{F(x), G(+\infty)\} = F(x) = \mu\big[(-\infty,x]\big] = \mu[A].$$

Ομοίως δείχνουμε ότι $\pi[\mathbb{R}\times A] = \nu[A]$ για ένα οποιοδήποτε σύνολο της μορφής $A = (-\infty, y]$.

Θα ολοκληρώσουμε την απόδειξη του Θεωρήματος σε τρία βήματα.

Βήμα 1ο:

Ισχυρισμός 12.

$$\text{(2.47)} \qquad Supp(\pi) \subseteq \{(x,y)\in\mathbb{R}^2 : \quad F(x^-) \le G(y) \;\text{ και }\; G(y^-) \le F(x)\}$$

Απόδειξη του Ισχυρισμού 12:

Έστω ένα σημείο $(x,y)\in\mathbb{R}^2$ για το οποίο υποθέτουμε ότι $F(x^-) > G(y)$. Όμοια γίνεται η απόδειξη αν $G(y^-) > F(x)$. Θέτουμε $\varepsilon = F(x^-) - G(y) > 0$. Εφόσον $F(x^-)$ είναι το αριστερό όριο της F στο x, εξ ορισμού του ορίου υπάρχει $\delta_1 > 0$ τέτοιο ώστε

$$x' \in [x-\delta_1, x) \implies F(x^-) - F(x') < \frac{\varepsilon}{3}.$$

Εφόσον η G είναι συνεχής από δεξιά στο y, υπάρχει $\delta_2 > 0$ τέτοιο ώστε

$$y' \in (y, y+\delta_2] \implies G(y') - G(y) < \frac{\varepsilon}{3}.$$

Θέτουμε $\delta = \min\{\delta_1, \delta_2\}$ και διακρίνουμε δύο περιπτώσεις.

Αν $x' \in [x-\delta, x)$ και $y' \in (y, y+\delta]$, τότε:

$$\begin{aligned} F(x') - G(y') &= \big(F(x^-) - G(y)\big) - \big(F(x^-) - F(x')\big) - \big(G(y') - G(y)\big) \\ &> \varepsilon - \frac{\varepsilon}{3} - \frac{\varepsilon}{3} = \frac{\varepsilon}{3} > 0. \end{aligned}$$

Αν $x' \in [x, x+\delta]$ ή $y' \in [y-\delta, y]$, τότε λόγω της μονοτονίας των F και G, έχουμε:

$$F(x') - G(y') \geq F(x^-) - G(y) = \varepsilon > 0.$$

Βλέπουμε, λοιπόν, ότι σε κάθε περίπτωση υπάρχει περιοχή του x, $U_x = [x-\delta, x+\delta]$ και περιοχή του y, $U_y = [y-\delta, y+\delta]$ έτσι ώστε

$$F(x') > G(y'), \qquad x' \in U_x, y' \in U_y.$$

Άρα,

$$H(x', y') = \min\{F(x'), G(y')\} = G(y'), \qquad x' \in U_x, y' \in U_y.$$

Θα δείξουμε ότι $\pi\big[U_x \times U_y\big] = 0$. Πράγματι,

$$\begin{aligned} \pi\big[U_x \times U_y\big] &= H(x+\delta, y+\delta) + H(x-\delta, y-\delta) - H(x+\delta, y-\delta) - H(x-\delta, y+\delta) \\ &= G(y+\delta) + G(y-\delta) - G(y-\delta) - G(y+\delta) = 0. \end{aligned}$$

Εφόσον, λοιπόν, το π δίνει μηδενική μάζα στο ορθογώνιο $U_x \times U_y$, έπεται ότι $(x, y) \notin Supp(\pi)$.

■

Ισχυρισμός 13. Το σύνολο $Supp(\pi)$ είναι ένα μονότονο υποσύνολο του $\mathbb{R}^2$.

<u>Απόδειξη του Ισχυρισμού 13:</u>

Έστω δύο σημεία $(x_1, y_1), (x_2, y_2) \in Supp(\pi)$. Υποθέτουμε ότι $x_1 > x_2$ και χρειάζεται να δείξουμε ότι $y_1 \geq y_2$. Εφαρμόζοντας τη σχέση (2.47) και εφόσον η F είναι αύξουσα, προκύπτει

$$G(y_1) \geq F(x_1^-) \geq F(x_2) \geq G(y_2^-).$$

Διακρίνουμε δύο περιπτώσεις:

Αν $G(y_1) > G(y_2^-)$, τότε λόγω μονοτονίας της G έχουμε $y_1 \geq y_2$ όπως επιθυμούσαμε.

Αν $G(y_1) = G(y_2^-)$, τότε αναγκαστικά θα ισχύει

$$G(y_1) = F(x_1^-) = F(x_2) = G(y_2^-).$$

Εάν υποθέσουμε ότι $y_2 > y_1$, τότε η F θα είναι σταθερή στο διάστημα $[x_2, x_1)$ και η G σταθερή στο $[y_1, y_2)$. Θα δείξουμε, με απαγωγή σε άτοπο, ότι αυτό είναι αδύνατο.

Θεωρούμε το σημείο (x_2, y_2) και έστω $\varepsilon > 0$. Θα δείξουμε ότι για το ορθογώνιο

$$R_\varepsilon = [x_2 - \varepsilon, x_2 + \varepsilon] \times [y_2 - \varepsilon, y_2 + \varepsilon]$$

ισχύει $\pi[R_\varepsilon] = 0$, αρκεί το ε να επιλεγεί αρκούντως μικρό. Πράγματι,

$$\pi[R_\varepsilon] = H(x_2 + \varepsilon, y_2 + \varepsilon) + H(x_2 - \varepsilon, y_2 - \varepsilon) - H(x_2 - \varepsilon, y_2 + \varepsilon) - H(x_2 + \varepsilon, y_2 - \varepsilon)$$

Εκμεταλλευόμενοι τον ορισμό της H και τη μονοτονία των F, G, έχουμε:

$H(x_2 + \varepsilon, y_2 + \varepsilon) = F(x_2 + \varepsilon)$, αφού $F(x_2 + \varepsilon) = F(x_2) = G(y_2^-) \leq G(y_2 + \varepsilon)$ αρκεί το ε να επιλεγεί αρκούντως μικρό ώστε $x_2 + \varepsilon \in [x_2, x_1)$.

$H(x_2 - \varepsilon, y_2 - \varepsilon) = F(x_2 - \varepsilon)$, αφού $F(x_2 - \varepsilon) \leq F(x_2) = G(y_2^-) = G(y_2 - \varepsilon)$ αρκεί το ε να επιλεγεί αρκούντως μικρό ώστε $y_2 - \varepsilon \in [y_1, y_2)$.

$H(x_2 - \varepsilon, y_2 + \varepsilon) = F(x_2 - \varepsilon)$, αφού $F(x_2 - \varepsilon) \leq F(x_2) = G(y_2^-) \leq G(y_2 + \varepsilon)$.

$H(x_2 + \varepsilon, y_2 - \varepsilon) = F(x_2 + \varepsilon)$, αφού $F(x_2 + \varepsilon) = F(x_2) = G(y_2^-) = G(y_2 - \varepsilon)$ αρκεί το ε να επιλεγεί αρκούντως μικρό ώστε $x_2 + \varepsilon \in [x_2, x_1)$, $y_2 - \varepsilon \in [y_1, y_2)$.

Αν, λοιπόν, το ε επιλεγεί αρκούντως μικρό έτσι ώστε $x_2 + \varepsilon \in [x_2, x_1)$ και $y_2 - \varepsilon \in [y_1, y_2)$, τότε $\pi[R_\varepsilon] = 0$. Έπεται ότι το σημείο $(x_2, y_2) \notin Supp(\pi)$.

Έχουμε φτάσει σε αντίφαση, οπότε ισχύει $y_1 \geq y_2$. Προκύπτει ότι το $Supp(\pi)$ είναι ένα μονότονο υποσύνολο του $\mathbb{R}^2$.

■

Βήμα 2ο:

Το $Supp(\pi)$ είναι ένα μονότονο υποσύνολο του $\mathbb{R}^2$, άρα περιέχεται στο υποδιαφορικό μιας κάτω ημισυνεχούς κυρτής συνάρτησης (Πρόταση 2.41). Από το Κριτήριο Βελτιστοποίησης Knott – Smith, το π είναι ένα βέλτιστο πλάνο μεταφοράς για το πρόβλημα Kantorovich αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x, y) = |x - y|^2$.

Ισχυρισμός 14. Αν με $\mathcal{L}$ συμβολίσουμε το μέτρο Lebesgue στο διάστημα $[0, 1]$, τότε το μέτρο π ,που ορίζει η συνάρτηση H μπορεί να γραφεί και ως

$$\pi = (F^{-1} \times G^{-1})\#\mathcal{L}. \tag{2.48}$$

Απόδειξη του Ισχυρισμού 14:

Αρκεί να επαληθεύσουμε την ισότητα για ένα τυχαίο ορθογώνιο της μορφής

$$R(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 : \quad x \le x_0 \quad \text{και} \quad y \le y_0\},$$

αφού η κλάση των ορθογωνίων αυτών είναι κλειστή ως προς τις τομές και γεννάει όλα τα Borel σύνολα στον $\mathbb{R}^2$. Για ένα τέτοιο ορθογώνιο, σύμφωνα με τη σχέση (2.48) έχουμε:

$$\begin{aligned} \pi[R(x_0, y_0)] &= \mathcal{L}\big[\{t \in [0,1] : \quad \big(F^{-1}(t), G^{-1}(t)\big) \in R(x_0, y_0)\}\big] \\ &= \mathcal{L}[\{t \in [0,1] : \quad F^{-1}(t) \le x_0 \quad \text{και} \quad G^{-1}(t) \le y_0\}] \\ &= \mathcal{L}[\{t \in [0,1] : \quad F^{-1}(t) \le x_0\} \cap \{t \in [0,1] : \quad G^{-1}(t) \le y_0\}]. \end{aligned}$$

Από τον ορισμό των γενικευμένων αντίστροφων συναρτήσεων F^{-1} και G^{-1}, βλέπουμε ότι το σύνολο $\{t \in [0,1] : \quad F^{-1}(t) \le x_0\}$ θα είναι ίσο με το διάστημα $[0, F(x_0)]$ ή $\big[0, F(x_0)\big)$ ενώ το σύνολο $\{t \in [0,1] : \quad G^{-1}(t) \le y_0\}$ θα είναι ίσο με το διάστημα $[0, G(y_0)]$ ή $\big[0, G(y_0)\big)$. Σε κάθε περίπτωση, το σύνολο του οποίου ζητάμε να υπολογίσουμε το μέτρο Lebesgue είναι ένα διάστημα με άκρα 0 και $\min\{F(x_0), G(y_0)\}$. Άρα, έχει μέτρο Lebesgue ίσο με

$$\min\{F(x_0), G(y_0)\} = H(x_0, y_0).$$

Τελικά, τα μέτρα που ορίζονται από τη σχέση (2.48) και τη συνάρτηση H συμπίπτουν πάνω στα ορθογώνια $R(x_0, y_0)$, άρα και σε όλα τα Borel υποσύνολα του $\mathbb{R}^2$.

■

Βήμα 3ο:

Από τη σχέση (2.48) προκύπτει ότι για κάθε μη αρνητική μετρήσιμη συνάρτηση $u: \mathbb{R}^2 \to \mathbb{R}$ ισχύει:

$$\int_{\mathbb{R}^2} u(x, y)\, d\pi(x, y) = \int_0^1 u\big(F^{-1}(t), G^{-1}(t)\big)\, dt.$$

Για να ολοκληρωθεί η απόδειξη του Θεωρήματος, έχουμε

$$T_2(\mu, \nu) = \int_{\mathbb{R}^2} c(x, y)\, d\pi(x, y) = \int_0^1 c\big(F^{-1}(t), G^{-1}(t)\big)\, dt = \int_0^1 |F^{-1}(t) - G^{-1}(t)|^2\, dt.$$

■

Παρατηρήσεις 2.44

(i). Παραθέτουμε το Θεώρημα των Hoeffding – Frechet, για την απόδειξη του οποίου παραπέμπουμε στο [12].

Θεωρούμε μ, ν δύο μέτρα πιθανότητας στο $\mathbb{R}$ με αντίστοιχες αθροιστικές συναρτήσεις κατανομής F και G. Αν $H: \mathbb{R}^2 \longrightarrow \mathbb{R}$ είναι μια μη αρνητική αύξουσα συνάρτηση, συνεχής από δεξιά και ως προς τις δύο μεταβλητές, τότε αυτή ορίζει ένα μέτρο πιθανότητας $\pi \in \Pi(\mu, \nu)$ αν και μόνο αν για κάθε $(x, y) \in \mathbb{R}^2$ ισχύει

$$F(x) + G(y) - 1 \leq H(x,y) \leq min\{F(x), G(y)\}. \tag{2.49}$$

(ii). Αν τα μέτρα μ, ν δεν δίνουν μάζα στα σημεία, δηλαδή $\mu[\{x\}] = \nu[\{y\}] = 0$ για κάθε $x, y \in \mathbb{R}$, τότε η συνάρτηση $T = G^{-1} \circ F$ μεταφέρει το μ πάνω στο ν, συμβολικά $\nu = T\#\mu$, και ισχύει

$$\int_{-\infty}^{x} d\mu = \int_{-\infty}^{T(x)} d\nu. \tag{2.50}$$

<u>Απόδειξη:</u>

Για κάθε $x \in \mathbb{R}$, ισχύει $F(x) - F(x^-) = \mu(\{x\}) = 0$, δηλαδή η F είναι συνεχής στο $\mathbb{R}$ και για κάθε $t \in [0,1]$, $F\big(F^{-1}(t)\big) = t$. Ομοίως για την G.

Εφόσον τα σύνολα Borel παράγονται από διαστήματα της μορφής $(-\infty, x]$ και το σύνολο αυτών των διαστημάτων είναι κλειστό ως προς τις τομές, αρκεί να δείξουμε ότι για οποιοδήποτε σύνολο της μορφής $A = (-\infty, x]$ ισχύει: $\nu[A] = \mu[T^{-1}(A)]$.

Έχουμε, λοιπόν:

$$\begin{aligned} T^{-1}(A) &= T^{-1}\big((-\infty, x]\big) = \{t \in \mathbb{R} : \ T(t) \leq x\} = \big\{t \in \mathbb{R} : \ G^{-1}\big(F(t)\big) \leq x\big\} \\ &= \{t \in \mathbb{R} : \ F(t) \leq G(x)\} = \big\{t \in \mathbb{R} : \ t \leq F^{-1}\big(G(x)\big)\big\} = \big(-\infty, F^{-1}\big(G(x)\big)\big]. \end{aligned}$$

Συνεπώς,

$$\mu[T^{-1}(A)] = F\left(F^{-1}\big(G(x)\big)\right) = G(x) = \nu[A].$$

Στη συνέχεια, εφόσον για κάθε $x \in \mathbb{R}$ ισχύει $F(x) = G\big(T(x)\big)$, εξ ορισμού των αθροιστικών συναρτήσεων κατανομής παίρνουμε

$$\int_{-\infty}^{x} d\mu = \int_{-\infty}^{T(x)} d\nu.$$

■

(iii). Υποθέτουμε ότι τα μέτρα μ, ν έχουν αντίστοιχα πυκνότητες f, g αναφορικά με το μέτρο Lebesgue, οι οποίες είναι συνεχείς και η g είναι αυστηρά θετική. Τότε, $T \in C^1(\mathbb{R})$ και επίσης

$$f(x) = g\big(T(x)\big)T'(x).$$

Απόδειξη:

Από την υπόθεση έχουμε ότι $F, G \in C^1(\mathbb{R})$ και επίσης, η G είναι γνησίως αύξουσα στο $\mathbb{R}$. Επομένως, η G^{-1} είναι η κανονική αντίστροφη της συνάρτησης G και φυσικά $G^{-1} \in C^1(\mathbb{R})$. Προκύπτει ότι $T \in C^1(\mathbb{R})$ ως σύνθεση δύο C^1 συναρτήσεων. Ακόμα, από τη σχέση (2.50) προκύπτει ότι

$$\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{T(x)} g(t)dt.$$

Παραγωγίζοντας την παραπάνω σχέση, παίρνουμε: $f(x) = g\big(T(x)\big)T'(x)$.

■

Παράδειγμα 2.45

Θεωρούμε μ, ν δύο μέτρα πιθανότητας στο $\mathbb{R}$ με πεπερασμένες ροπές 1ης τάξης για τα οποία υποθέτουμε ότι υπάρχουν ξένα διαστήματα $I, J \subseteq \mathbb{R}$ τέτοια ώστε $Supp(\mu) \subseteq I$, $Supp(\nu) \subseteq J$. Θεωρούμε ακόμα το πρόβλημα μεταφοράς μάζας του Kantorovich με αντίστοιχη συνάρτηση κόστους $c(x,y) = |x-y|$.

Τότε, οποιοδήποτε πλάνο μεταφοράς $\pi \in \Pi(\mu,\nu)$ είναι βέλτιστο.

Απόδειξη:

Αρκεί να δείξουμε ότι για οποιοδήποτε $\pi_0 \in \Pi(\mu,\nu)$, υπάρχει ζεύγος συναρτήσεων $(\varphi_0, \psi_0) \in \Phi_c$ έτσι ώστε $I[\pi_0] = J(\varphi_0, \psi_0)$. Τότε, θα έχουμε

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] \leq I[\pi_0] = J(\varphi_0, \psi_0) \leq \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi),$$

ενώ σύμφωνα με τη δυϊκότητα Kantorovich

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi).$$

Τελικά, στην παραπάνω σχέση θα ισχύει παντού η ισότητα, και άρα το π θα είναι βέλτιστο πλάνο μεταφοράς.

Έστω, λοιπόν, ένα $\pi \in \Pi(\mu,\nu)$. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι το διάστημα I, μέσα στο οποίο «ζει» το μέτρο μ, βρίσκεται πιο δεξιά πάνω στον άξονα των πραγματικών αριθμών από το διάστημα J, μέσα στο οποίο «ζει» το μέτρο ν. Έχουμε, λοιπόν:

$$I[\pi] = \int_{\mathbb{R}^2} |x-y|\,d\pi(x,y) = \int_{I\times J} |x-y|\,d\pi(x,y) = \int_{I\times J} (x-y)\,d\pi(x,y)$$
$$= \int_{I\times J} x\,d\pi(x,y) - \int_{I\times J} y\,d\pi(x,y) = \int_I x\,d\mu + \int_J (-y)\,d\nu = \int_{\mathbb{R}} x\,d\mu + \int_{\mathbb{R}} (-y)\,d\nu = J(\varphi,\psi),$$

με $\varphi(x) = x$ και $\psi(y) = -y$. Προφανώς, $(\varphi,\psi) \in \Phi_c$ και η απόδειξη είναι πλήρης.

■

2.3. Εναλλακτικά επιχειρήματα.

Στην παρούσα ενότητα, θα επιστρέψουμε στην περίπτωση του προβλήματος Kantorovich στον $\mathbb{R}^n$ αναφορικά με μια τετραγωνική συνάρτηση κόστους και θα αναφερθούμε σε κάποιες ιδέες γεωμετρικής φύσης που στηρίζονται στην έννοια της κυκλικής μονοτονίας. Μέσω αυτών θα μπορέσουμε να ανακτήσουμε τα αποτελέσματα της ενότητας 2.1, σε κάποιες περιπτώσεις μάλιστα σε μεγαλύτερη γενικότητα.

2.3.1. Κυκλική μονοτονία.

Παράδειγμα 2.46 (Η διακριτή περίπτωση)

Έστω $x_1,\dots,x_N$ και $y_1,\dots,y_N$ σημεία στον $\mathbb{R}^n$ (όχι απαραίτητα διακεκριμένα). Ορίζουμε τα μέτρα πιθανότητας

$$\mu \coloneqq \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}\ , \quad \nu \coloneqq \frac{1}{N}\sum_{j=1}^{N} \delta_{y_j} \quad \text{και} \quad \pi \coloneqq \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i,y_i)}.$$

Συμβολίζουμε με S_N το σύνολο όλων των μεταθέσεων του συνόλου $\{1,\dots,N\}$. Τότε:

Α) Το π είναι βέλτιστο για το πρόβλημα Monge – Kantorovich της μεταφοράς του μ πάνω στο ν αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x,y) = |x-y|^2$ αν και μόνο αν για όλες τις μεταθέσεις $\sigma \in S_N$, ισχύει

$$\sum_{i=1}^{n} |x_i - y_i|^2 \le \sum_{i=1}^{n} \left|x_i - y_{\sigma(i)}\right|^2 \tag{2.51}$$

Β) Η ισχύς της σχέσης (2.51) για κάθε $\sigma \in S_N$, είναι ισοδύναμη με τον ακόλουθο ισχυρισμό:

Για κάθε $m \leq N$ και για κάθε $i_1, \dots, i_m \in \{1, \dots, N\}$, ισχύει

$$(2.52) \qquad \sum_{k=1}^{m} \left|x_{i_k} - y_{i_k}\right|^2 \leq \sum_{k=1}^{m} \left|x_{i_k} - y_{i_{k-1}}\right|^2,$$

όπου, βάσει σύμβασης, $i_0 = i_m$.

Γ) Η σχέση (2.52) είναι ισοδύναμη με την

$$(2.53) \qquad \sum_{k=1}^{m} y_{i_k} \cdot \left(x_{i_{k+1}} - x_{i_k}\right) \leq 0,$$

όπου, βάσει σύμβασης, $i_{m+1} = i_1$.

Απόδειξη:

Α) Έστω ότι το π είναι βέλτιστο. Θεωρούμε μια μετάθεση $\sigma \in S_N$ και ορίζουμε το μέτρο

$$\tilde{\pi} \coloneqq \frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i, y_{\sigma(i)})}.$$

Ισχύει $\tilde{\pi} \in \Pi(\mu, \nu)$. Πράγματι, έστω $A \subseteq \mathbb{R}^n$. Τότε

$$\tilde{\pi}[A \times \mathbb{R}^n] = \frac{1}{N} \sum_{x_i \in A} \delta_{(x_i, y_{\sigma(i)})}[A \times \mathbb{R}^n] = \frac{1}{N} \sum_{x_i \in A} \delta_{x_i}[A] = \mu[A].$$

Ομοίως, δείχνουμε ότι $\tilde{\pi}[\mathbb{R}^n \times A] = \nu[A]$. Στη συνέχεια, εφόσον το μέτρο π είναι βέλτιστο, έχουμε $I[\pi] \leq I[\tilde{\pi}]$, το οποίο ισοδυναμεί με το ότι

$$\int_{\mathbb{R}^n} |x - y|^2 \, d\pi(x, y) \leq \int_{\mathbb{R}^n} |x - y|^2 \, d\tilde{\pi}(x, y)$$

ισοδύναμα

$$\frac{1}{N} \sum_{i=1}^{n} |x_i - y_i|^2 \leq \frac{1}{N} \sum_{i=1}^{n} \left|x_i - y_{\sigma(i)}\right|^2$$

ισοδύναμα

$$\sum_{i=1}^{n} |x_i - y_i|^2 \leq \sum_{i=1}^{n} \left|x_i - y_{\sigma(i)}\right|^2.$$

Αντίστροφα, έστω ότι ισχύει η σχέση (2.51). Θέλουμε να δείξουμε ότι το μέτρο π είναι βέλτιστο. Κατ' αρχήν, παρατηρούμε ότι κάθε μέτρο $\tilde{\pi} \in \Pi(\mu, \nu)$μπορεί να γραφεί στη μορφή

$$\tilde{\pi} = \frac{1}{N} \sum_{i,j=1}^{N} \pi_{i,j} \delta_{(x_i, y_j)},$$

όπου $(\pi_{i,j})$ είναι ένας διπλά στοχαστικός πίνακας. Αυτό σημαίνει ότι όλα τα στοιχεία $\pi_{i,j}$ είναι μη αρνητικοί αριθμοί και το άθροισμα κάθε γραμμής και στήλης του πίνακα ισούται με 1.

Πράγματι, αφού $\tilde{\pi} \in \Pi(\mu, \nu)$ έχουμε $\tilde{\pi}(\{x_1, \dots, x_N\} \times \mathbb{R}^n) = \mu(\{x_1, \dots, x_N\}) = 1$. Ομοίως, $\tilde{\pi}(\mathbb{R}^n \times \{y_1, \dots, y_N\}) = 1$. Άρα, $\tilde{\pi}(\{x_1, \dots, x_N\} \times \{y_1, \dots, y_N\}) = \tilde{\pi}(\{(x_i, y_i) : \ i, j = 1, \dots, N\}) = 1$. Συνεπώς,

$$\tilde{\pi} = \sum_{i,j=1}^{N} \tilde{\pi}_{i,j} \delta_{(x_i, y_i)} \quad \text{με } \tilde{\pi}_{i,j} \geq 0.$$

Τώρα, αν $A = \{x_i\}$, τότε

$$1 = \mu(A) = \sum_{j=1}^{n} \tilde{\pi}_{i,j} \ , \quad i = 1, \dots, N.$$

Ομοίως,

$$\sum_{i=1}^{n} \tilde{\pi}_{i,j} = 1 \ , \quad j = 1, \dots, N.$$

Συμβολίζουμε με $\mathcal{D}_N$ το σύνολο των $N \times N$ διπλά στοχαστικών πινάκων, το οποίο είναι ένα κυρτό και συμπαγές υποσύνολο του $\mathbb{R}^{N^2}$. Σύμφωνα με το Θεώρημα Krein - Milman ([17], Θεώρημα 5.5), κάθε διπλά στοχαστικός πίνακας γράφεται ως κυρτός συνδυασμός ακραίων σημείων του $\mathcal{D}_N$.

Αν K είναι ένα κυρτό σύνολο, το $x \in K$ λέγεται ακραίο σημείο του K αν δεν υπάρχουν $y, z \in K$, $y \neq z$, και $t \in (0, 1)$ έτσι ώστε $x = ty + (1 - t)z$.

Σύμφωνα με το Θεώρημα Birkhoff ([17], Θεώρημα 5.7), τα ακραία σημεία του $\mathcal{D}_N$ είναι ακριβώς οι πίνακες μετάθεσης.

Επομένως, κάθε $\tilde{\pi} \in \Pi(\mu, \nu)$ γράφεται ως

$$\tilde{\pi} = \frac{1}{N} \sum_{i,j=1}^{N} \tilde{\pi}_{i,j} \delta_{(x_i, y_j)},$$

με

$$(\tilde{\pi}_{i,j}) = \sum_{k=1}^{m} t_k P_k,$$

όπου κάθε P_k είναι ένας πίνακας μετάθεσης, $t_k \geq 0$ και

$$\sum_{k=1}^{m} t_k = 1.$$

Αν, τώρα, P είναι ένας πίνακας μετάθεσης, αυτός αντιστοιχεί σε μια μετάθεση $\sigma \in S_N$ και το αντίστοιχο μέτρο γράφεται

$$\tilde{\pi}_P = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_{\sigma(i)})}.$$

Σύμφωνα με τη σχέση (2.51), για οποιονδήποτε πίνακα μετάθεσης P, έχουμε

$$\frac{1}{N}\sum_{i=1}^{n} |x_i - y_i|^2 \leq \frac{1}{N}\sum_{i=1}^{n} \left|x_i - y_{\sigma(i)}\right|^2 = \int_{\mathbb{R}^n} |x-y|^2 \, d\tilde{\pi}_P.$$

Τελικά, για να δείξουμε ότι το μέτρο

$$\pi = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_i)}$$

είναι βέλτιστο, επιλέγουμε οποιοδήποτε $\tilde{\pi} \in \Pi(\mu, \nu)$ και έχουμε:

$$\begin{aligned} I[\tilde{\pi}] &= \int_{\mathbb{R}^n} |x-y|^2 \, d\tilde{\pi} = \int_{\mathbb{R}^n} |x-y|^2 \, d\left(\sum_{k=1}^{m} t_k \tilde{\pi}_{P_k}\right) = \sum_{k=1}^{m} t_k \int_{\mathbb{R}^n} |x-y|^2 \, d\tilde{\pi}_{P_k} \\ &\geq \sum_{k=1}^{m} t_k \left(\frac{1}{N}\sum_{i=1}^{N} |x_i - y_i|^2\right) = \left(\frac{1}{N}\sum_{i=1}^{N} |x_i - y_i|^2\right)\left(\sum_{k=1}^{m} t_k\right) = \frac{1}{N}\sum_{i=1}^{N} |x_i - y_i|^2 = I[\pi] \end{aligned}$$

αφού

$$\sum_{k=1}^{m} t_k = 1.$$

B) Υποθέτουμε ότι ισχύει η σχέση (2.51) για κάθε $\sigma \in S_N$. Επιλέγουμε $m \in \mathbb{N}$, $m \leq N$ και $i_1, \dots i_m \in \{1,2,\dots,N\}$. Ορίζουμε τη μετάθεση $\sigma \in S_N$ του συνόλου $\{1,2,\dots,N\}$ ως εξής:

$$\begin{cases} \sigma(i_k) = i_{k-1}\,, & k = 1,2,\dots,m \ \text{ με } i_0 = i_m \\ \sigma(j) = j\,, & j \neq i_1, i_2, \dots, i_m \end{cases}$$

Για αυτή τη μετάθεση, εφαρμόζουμε τη σχέση (2.51):

$$\sum_{i=1}^{n} |x_i - y_i|^2 \leq \sum_{i=1}^{n} \left|x_i - y_{\sigma(i)}\right|^2$$

$$\sum_{k=1}^{m}\left|x_{i_k}-y_{i_k}\right|^2+\sum_{j\neq i_1,i_2,\dots,i_m}\left|x_j-y_j\right|^2\leq\sum_{k=1}^{m}\left|x_{i_k}-y_{\sigma(i_k)}\right|^2+\sum_{j\neq i_1,i_2,\dots,i_m}\left|x_j-y_{\sigma(j)}\right|^2$$

Επειδή, όμως, $\sigma(j)=j$ για $j\neq i_1,i_2,\dots,i_m$, διαγράφουμε τους δύο ίδιους όρους και προκύπτει το ζητούμενο

$$\sum_{k=1}^{m}\left|x_{i_k}-y_{i_k}\right|^2\leq\sum_{k=1}^{m}\left|x_{i_k}-y_{i_{k-1}}\right|^2.$$

Αντίστροφα, υποθέτουμε ότι ισχύει η σχέση (2.52) για κάθε $m\leq N$ και για κάθε $i_1,\dots,i_m\in\{1,\dots,N\}$. Θεωρούμε μια μετάθεση $\sigma\in S_N$. Η μετάθεση αυτή μπορεί να γραφεί ως γινόμενο ξένων κύκλων, δηλαδή

$$\sigma=s_1s_2\dots s_m$$

όπου s_i είναι κύκλοι με ξένους φορείς. Σύμφωνα με την υπόθεσή μας, μπορούμε να γράψουμε τη σχέση (2.52) για κάθε έναν κύκλο ξεχωριστά. Άρα,

$$\sum_{i_k\in s_1}\left|x_{i_k}-y_{i_k}\right|^2\leq\sum_{i_k\in s_1}\left|x_{i_k}-y_{i_{k-1}}\right|^2$$

$$\vdots$$

$$\sum_{i_k\in s_m}\left|x_{i_k}-y_{i_k}\right|^2\leq\sum_{i_k\in s_m}\left|x_{i_k}-y_{i_{k-1}}\right|^2.$$

Αθροίζοντας κατά μέλη τις παραπάνω ανισότητες και εφόσον οι κύκλοι είναι ξένοι μεταξύ τους, παίρνουμε τη ζητούμενη σχέση (2.51).

Γ) Για την απόδειξη της ισοδυναμίας των σχέσεων (2.52) και (2.53), έχουμε:

$$\sum_{k=1}^{m}\left|x_{i_k}-y_{i_k}\right|^2\leq\sum_{k=1}^{m}\left|x_{i_k}-y_{i_{k-1}}\right|^2\Leftrightarrow$$

$$\sum_{k=1}^{m}\left(\left|x_{i_k}\right|^2-2x_{i_k}\cdot y_{i_k}+\left|y_{i_k}\right|^2\right)-\sum_{k=1}^{m}\left(\left|x_{i_k}\right|^2-2x_{i_k}\cdot y_{i_{k-1}}+\left|y_{i_{k-1}}\right|^2\right)\leq 0\Leftrightarrow$$

$$\sum_{k=1}^{m}\left|x_{i_k}\right|^2-2\sum_{k=1}^{m}x_{i_k}\cdot y_{i_k}+\sum_{k=1}^{m}\left|y_{i_k}\right|^2-\sum_{k=1}^{m}\left|x_{i_k}\right|^2+2\sum_{k=1}^{m}x_{i_k}\cdot y_{i_{k-1}}-\sum_{k=1}^{m}\left|y_{i_{k-1}}\right|^2\leq 0$$

Διαγράφοντας τους ίδιους όρους και διαιρώντας με το 2, ισοδύναμα προκύπτει:

$$\sum_{k=1}^{m} x_{i_k} \cdot y_{i_{k-1}} - \sum_{k=1}^{m} x_{i_k} \cdot y_{i_k} \leq 0.$$

Αναπτύσσοντας τα δύο αθροίσματα, ισοδύναμα έχουμε

$$x_{i_1} \cdot y_{i_0} + x_{i_2} \cdot y_{i_1} + \cdots + x_{i_m} \cdot y_{i_{m-1}} - x_{i_1} \cdot y_{i_1} - x_{i_2} \cdot y_{i_2} - \cdots - x_{i_m} \cdot y_{i_m} \leq 0$$

ή ισοδύναμα

$$y_{i_1} \cdot \left(x_{i_2} - x_{i_1}\right) + y_{i_2} \cdot \left(x_{i_3} - x_{i_2}\right) + \cdots + y_{i_m} \cdot \left(x_{i_1} - x_{i_m}\right) \leq 0$$

το οποίο πιο σύντομα μπορούμε να γράψουμε

$$\sum_{k=1}^{m} y_{i_k} \cdot \left(x_{i_{k+1}} - x_{i_k}\right) \leq 0.$$

■

Ορισμός 2.47 (Κυκλική μονοτονία)

Ένα σύνολο $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^n$ λέγεται κυκλικά μονότονο αν ικανοποιεί την ακόλουθη συνθήκη:

Για κάθε $m \geq 1$ και για οποιαδήποτε συλλογή σημείων $(x_1, y_1), \ldots, (x_m, y_m) \in \Gamma$,

$$\sum_{i=1}^{m} |x_i - y_i|^2 \leq \sum_{i=1}^{m} |x_i - y_{i-1}|^2,$$

με τη σύμβαση $y_0 = y_m$, ή ισοδύναμα:

$$\sum_{i=1}^{m} y_i \cdot (x_{i+1} - x_i) \leq 0,$$

με τη σύμβαση $x_{m+1} = x_1$.

Παρατήρηση 2.48

Αν υποθέσουμε ότι $\pi_k \longrightarrow \pi$ ασθενώς, τότε για κάθε σημείο $(x, y) \in Supp(\pi)$ υπάρχει ακολουθία σημείων $(x_k, y_k) \in Supp(\pi_k)$ έτσι ώστε $(x_k, y_k) \longrightarrow (x, y)$.

Απόδειξη:

Έστω σημείο $(x, y) \in Supp(\pi)$. Υποθέτουμε ότι δεν υπάρχει ακολουθία σημείων $(x_k, y_k) \in Supp(\pi_k)$ έτσι ώστε $(x_k, y_k) \longrightarrow (x, y)$. Τότε, θα αποδείξουμε τον ακόλουθο ισχυρισμό.

Ισχυρισμός 15. Υπάρχει περιοχή V του (x,y) έτσι ώστε $Supp(\pi_k) \cap V = \emptyset$, για άπειρα k.

<u>Απόδειξη του Ισχυρισμού 15:</u>

Υποθέτουμε ότι αυτό δεν ισχύει. Τότε, για κάθε περιοχή V του (x,y) υπάρχει $k_0 \in \mathbb{N}$ τέτοιο ώστε $Supp(\pi_k) \cap V \neq \emptyset$, $k \geq k_0$.

Για τη μπάλα $B\big((x,y),1\big)$ βρίσκουμε k_1 ώστε $Supp(\pi_k) \cap B\big((x,y),1\big) \neq \emptyset$, $k \geq k_1$.

Για τη μπάλα $B\left((x,y),\frac{1}{2}\right)$ βρίσκουμε $k_2 > k_1$ ώστε $Supp(\pi_k) \cap B\left((x,y),\frac{1}{2}\right) \neq \emptyset$, $k \geq k_2$.

Συνεχίζοντας με τον ίδιο τρόπο, κατασκευάζουμε ακολουθία σημείων

$$\big(x_{k_1}, y_{k_1}\big) \in Supp\big(\pi_{k_1}\big) \cap B\big((x,y),1\big)$$

$$\big(x_{k_1+1}, y_{k_1+1}\big) \in Supp\big(\pi_{k_1+1}\big) \cap B\big((x,y),1\big)$$

$$\vdots$$

$$\big(x_{k_2}, y_{k_2}\big) \in Supp\big(\pi_{k_2}\big) \cap B\left((x,y),\frac{1}{2}\right)$$

$$\vdots$$

για την οποία ισχύει

$$(x_k, y_k) \to (x,y) \qquad \text{και} \qquad (x_k, y_k) \in Supp(\pi_k), \quad k \geq k_1.$$

Αυτό, όμως, έρχεται σε αντίφαση με την αρχική μας υπόθεση ότι δεν υπάρχει τέτοια ακολουθία.

■

Με βάση τον Ισχυρισμό 15, βρίσκουμε μπάλα με κέντρο (x,y) και ακτίνα ρ έτσι ώστε

$$Supp(\pi_k) \cap B\big((x,y),\rho\big) = \emptyset, \qquad \text{για άπειρα } k.$$

Παίρνουμε, στη συνέχεια, συνάρτηση $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow [0,+\infty)$ συνεχή και φραγμένη και η οποία ικανοποιεί τα εξής:

$$\varphi|_{B\left((x,y),\frac{\rho}{2}\right)} = 1 \qquad \text{και} \qquad \varphi|_{\mathbb{R}^n - B((x,y),\rho)} = 0.$$

Τότε,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \varphi \, d\pi_k = 0, \qquad \text{για άπειρα } k$$

και

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi\, d\pi \geq \pi\left[B\left((x,y),\frac{\rho}{2}\right)\right] > 0.$$

Αυτό έρχεται σε αντίφαση με την υπόθεση $\pi_k \to \pi$ ασθενώς, σύμφωνα με την οποία θα έπρεπε

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi\, d\pi_k \to \int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi\, d\pi, \qquad \text{καθώς } k \to \infty.$$

■

Έπεται ότι αν έχουμε μια ακολουθία μέτρων πιθανότητας με κυκλικά μονότονους φορείς που συγκλίνει ασθενώς, τότε και το ασθενές όριο αυτής θα είναι επίσης ένα μέτρο πιθανότητας με κυκλικά μονότονο φορέα. Με άλλα λόγια, η κυκλική μονοτονία είναι κλειστή κάτω από την ασθενή σύγκλιση.

Πρόταση 2.49 (Τα βέλτιστα πλάνα μεταφοράς έχουν κυκλικά μονότονους φορείς)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$. Θεωρούμε το πρόβλημα Kantorovich της μεταφοράς μάζας του μ πάνω στο ν αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x,y) = |x-y|^2$. Έστω, ακόμα, ότι το $\pi \in \Pi(\mu,\nu)$ είναι ένα βέλτιστο πλάνο μεταφοράς.

Τότε, ο φορέας του π είναι ένα κυκλικά μονότονο υποσύνολο του $\mathbb{R}^n \times \mathbb{R}^n$.

Πριν την απόδειξη της πρότασης θα είναι χρήσιμο να θυμηθούμε κάποια πράγματα από τη θεωρία πιθανοτήτων.

Παρατήρηση 2.50

Έστω X ένας τοπολογικός χώρος και μια συλλογή Borel μέτρων πιθανότητας $\pi_1, \dots, \pi_m$ στον X. Τότε, υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{B}, \eta)$, Ω τοπολογικός χώρος και $\mathcal{B}$ η $\sigma-$άλγεβρα των Borel συνόλων, και Borel μετρήσιμες απεικονίσεις $\gamma_1, \dots, \gamma_m : \Omega \longrightarrow X$ έτσι ώστε για κάθε $j = 1,2,\dots,m$ να έχουμε $\pi_j = \gamma_j \# \eta$. Πράγματι, παίρνουμε $\Omega \coloneqq X^m$, $\eta = \pi_1 \times \dots \times \pi_m$ να είναι το μέτρο γινόμενο ορισμένο πάνω στα Borel υποσύνολα του Ω και $\gamma_j(x_1, \dots, x_m) \coloneqq x_j$ να είναι η $j-$οστή προβολή στον X.

Επίσης, αν π είναι ένα μέτρο πιθανότητας στον X και $U \in \mathcal{B}_X$ είναι ένα Borel υποσύνολο του X με μέτρο $\pi[U] > 0$, τότε ορίζουμε τον κανονικοποιημένο περιορισμό του π στο U μέσω της σχέσης

$$\pi[A] \coloneqq \frac{\pi[A \cap U]}{\pi[U]}, \qquad A \in \mathcal{B}_X.$$

Απόδειξη της Πρότασης 2.49:

Έστω ότι το μέτρο π είναι βέλτιστο, δηλαδή ελαχιστοποιεί το συνολικό κόστος μεταφοράς $I[\pi]$ ανάμεσα σε όλα τα μέτρα που ανήκουν στο $\Pi(\mu,\nu)$. Θέλουμε να δείξουμε ότι το $Supp(\pi)$ είναι κυκλικά μονότονο σύνολο.

Υποθέτουμε, προς απαγωγή σε άτοπο, ότι το $Supp(\pi)$ δεν είναι κυκλικά μονότονο. Θεωρούμε τη συνάρτηση

$$f(x_1,\dots,x_m,y_1,\dots,y_m)=\sum_{i=1}^{m}[|x_i-y_{i-1}|^2-|x_i-y_i|^2],$$

όπου $y_0=y_m$. Τότε, θα υπάρχει ακέραιος $m\geq 1$ και σημεία $(x_1,y_1),\dots,(x_m,y_m)\in Supp(\pi)$ έτσι ώστε $f(x_1,\dots,x_m,y_1,\dots,y_m)<0$.

Αφού η f είναι συνεχής συνάρτηση, για κάθε $i=1,2,\dots,m$ υπάρχουν συμπαγείς περιοχές U_i του x_i και V_i του y_i έτσι ώστε

$$f(u_1,\dots,u_m,v_1,\dots,v_m)<0,\qquad u_i\in U_i, v_i\in V_i.$$

Επιπλέον, εφόσον για κάθε $i=1,2,\dots,m$ έχουμε $(x_i,y_i)\in Supp(\pi)$, θα ισχύει

$$\pi[U_i\times V_i]>0,\qquad i=1,2,\dots,m$$

και άρα

$$\lambda\coloneqq\inf_i \pi[U_i\times V_i]>0.$$

Για κάθε $i=1,2,\dots,m$ ορίζουμε τον κανονικοποιημένο περιορισμό του μέτρου $\pi_i\in\Pi(\mu,\nu)$ στο σύνολο $U_i\times V_i$, δηλαδή

$$\pi_i[A]=\frac{\pi[A\cap(U_i\times V_i)]}{\pi[U_i\times V_i]},\qquad A\subseteq\mathbb{R}^n\times\mathbb{R}^n.$$

Έπειτα, το μέτρο

$$\pi'\coloneqq\pi-\sum_{i=1}^{m}\frac{\lambda\pi_i}{m}$$

είναι ένα θετικό μέτρο. Πράγματι, για κάθε $A\subseteq\mathbb{R}^n\times\mathbb{R}^n$ έχουμε:

$$\pi'[A]=\pi[A]-\sum_{i=1}^{m}\frac{\lambda}{m}\pi_i[A]=\pi[A]-\sum_{i=1}^{m}\frac{\lambda}{m}\frac{\pi[A\cap(U_i\times V_i)]}{\pi[U_i\times V_i]}\geq\pi[A]-\sum_{i=1}^{m}\frac{1}{m}\pi[A\cap(U_i\times V_i)]$$
$$\geq\pi[A]-\frac{1}{m}\sum_{i=1}^{m}\pi[A]=0.$$

Σημειώνουμε ότι η πρώτη ανισότητα ισχύει αφού εξ ορισμού $\lambda \coloneqq \inf_i \pi[U_i \times V_i]$ και ως εκ τούτου

$$\frac{\lambda}{\pi[U_i \times V_i]} \leq 1, \quad \text{για κάθε } i = 1,2, \dots, m.$$

Στη συνέχεια, δεδομένης της συλλογής των μέτρων πιθανότητας $\pi_1, \dots, \pi_m$ στον $\mathbb{R}^n \times \mathbb{R}^n$, υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{B}, \eta)$ και μετρήσιμες απεικονίσεις $\gamma_1, \dots, \gamma_m : \Omega \longrightarrow \mathbb{R}^n \times \mathbb{R}^n$ ώστε για κάθε $i = 1,2, \dots, m$ να έχουμε $\pi_i = \gamma_i \# \eta$. Θέτοντας $\gamma_i(\omega) = \big(u_i(\omega), v_i(\omega)\big)$, παίρνουμε $\pi_i = (u_i \times v_i) \# \eta$. Σημειώνουμε ότι κάθε απεικόνιση γ_i παίρνει σχεδόν παντού τιμές στο συμπαγές σύνολο $U_i \times V_i$.

Ορίζουμε το θετικό μέτρο

$$\tilde{\pi} = \pi + \frac{\lambda}{m} \sum_{i=1}^{m} [(u_i \times v_{i-1}) \# \eta - (u_i \times v_i) \# \eta].$$

Τότε, $\tilde{\pi} \in \Pi(\mu, \nu)$.

Αυτό είναι, πράγματι, αληθές αφού για κάθε Borel σύνολο $A \subseteq \mathbb{R}^n$ ισχύει

$$\sum_{i=1}^{m} \big((u_i \times v_{i-1}) \# \eta\big)[\mathbb{R}^n \times A] = \sum_{i=1}^{m} \eta[(u_i \times v_{i-1})^{-1}(\mathbb{R}^n \times A)] = \sum_{i=1}^{m} \eta[\mathbb{R}^n \times v_{i-1}^{-1}(A)]$$
$$= \sum_{i=1}^{m} \eta[\mathbb{R}^n \times v_i^{-1}(A)] = \sum_{i=1}^{m} \big((u_i \times v_i) \# \eta\big)[\mathbb{R}^n \times A].$$

Τέλος, υπολογίζοντας το κόστος μεταφοράς για το $\tilde{\pi}$ φτάνουμε σε αντίφαση με το γεγονός ότι το π είναι βέλτιστο. Παίρνοντας υπόψη ότι το ολοκλήρωμα της συνάρτησης f είναι αρνητικό, έχουμε:

$$I[\tilde{\pi}] - I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\, d\tilde{\pi} - \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\, d\pi = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\, d(\tilde{\pi} - \pi)$$
$$= \frac{\lambda}{m} \int_{\Omega} c(x,y)\, d\left(\sum_{i=1}^{m} [(u_i \times v_{i-1}) \# \eta - (u_i \times v_i) \# \eta] \right) = \frac{\lambda}{m} \int_{\Omega} \sum_{i=1}^{m} [c(u_i, v_{i-1}) - c(u_i, v_i)]\, d\eta$$
$$= \frac{\lambda}{m} \int_{\Omega} \sum_{i=1}^{m} [|u_i - v_{i-1}|^2 - |u_i - v_i|^2]\, d\eta = \frac{\lambda}{m} \int_{\Omega} f(u_1, \dots, u_m, v_1, \dots, v_m)\, d\eta < 0.$$

Τελικά, το μέτρο π έχει κυκλικά μονότονο φορέα και η απόδειξη της πρότασης είναι πλήρης.

■

Η πρόταση που μόλις αποδείξαμε δεν αποτελεί στην πραγματικότητα έναν «χαρακτηρισμό» των βέλτιστων πλάνων μεταφοράς, καθώς δεν έχει αποδειχθεί μέχρι και σήμερα η αντίστροφη πρόταση.

Ανοικτό πρόβλημα 2.51 (Κριτήριο βελτιστοποίησης για τετραγωνική συνάρτηση κόστους)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ με πεπερασμένες ροπές 2ης τάξης. Θεωρούμε το πρόβλημα Kantorovich της μεταφοράς μάζας του μ πάνω στο ν αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x,y) = |x-y|^2$.

Αν $\pi \in \Pi(\mu,\nu)$ είναι ένα πλάνο μεταφοράς με κυκλικά μονότονο φορέα, τότε είναι βέλτιστο;

Όταν το μέτρο μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue, το παραπάνω ερώτημα επιδέχεται θετική απάντηση όπως έπεται από τον ισχυρισμό της μοναδικότητας του Θεωρήματος 2.59 που θα αποδείξουμε παρακάτω. Πράγματι, αφού το π έχει κυκλικά μονότονο φορέα, σύμφωνα με το Θεώρημα 2.52 παρακάτω, αυτός περιέχεται στο υποδιαφορικό μιας κάτω ημισυνεχούς κυρτής συνάρτησης στον $\mathbb{R}^n$. Αφού, σύμφωνα με το Θεώρημα 2.59, υπάρχει μοναδική κάτω ημισυνεχής κυρτή συνάρτηση φ τέτοια ώστε $\nabla\varphi \# \mu = \nu$, από το θεώρημα του Brenier (2.37, (ii)), προκύπτει ότι το π είναι βέλτιστο.

2.3.2. Το Θεώρημα του Rockafellar.

Θεώρημα 2.52 (Το Θεώρημα του Rockafellar)

Ένα μη κενό σύνολο $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^n$ είναι κυκλικά μονότονο αν και μόνο αν περιέχεται στο υποδιαφορικό μιας κανονικής κάτω ημισυνεχούς κυρτής συνάρτησης $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Επιπλέον, τα μέγιστα κυκλικά μονότονα υποσύνολα είναι ακριβώς τα υποδιαφορικά των κανονικών κάτω ημισυνεχών κυρτών συναρτήσεων.

<u>Απόδειξη:</u>

<u>1ο βήμα:</u>

Έστω $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ κανονική κάτω ημισυνεχής κυρτή συνάρτηση. Θεωρούμε $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^n \times \mathbb{R}^n$ τέτοια ώστε για κάθε $i = 1, 2, \dots, m$ να ισχύει $y_i \in \partial\varphi(x_i)$.

Εξ ορισμού, αυτό σημαίνει ότι για κάθε $z \in \mathbb{R}^n$ και για κάθε $i = 1, 2, \dots, m$ ισχύει

$$\varphi(z) \geq \varphi(x_i) + y_i \cdot (z - x_i).$$

Ειδικότερα,

$$\begin{cases} \varphi(x_2) \geq \varphi(x_1) + y_1 \cdot (x_2 - x_1) \\ \varphi(x_3) \geq \varphi(x_2) + y_2 \cdot (x_3 - x_2) \\ \quad\vdots \\ \varphi(x_1) \geq \varphi(x_m) + y_m \cdot (x_1 - x_m) \end{cases}$$

Προσθέτοντας κατά μέλη τις ανισότητες αυτές, παίρνουμε

$$\sum_{i=1}^{m} y_i \cdot (x_{i+1} - x_i) \leq 0,$$

που σημαίνει ότι το υποδιαφορικό της φ είναι ένα κυκλικά μονότονο υποσύνολο του $\mathbb{R}^n \times \mathbb{R}^n$.

<u>2ο βήμα:</u>

Θεωρούμε ένα κυκλικά μονότονο σύνολο $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^n$. Επιλέγουμε $(x_0, y_0) \in \Gamma$ και ορίζουμε τη συνάρτηση

$$\varphi(x) = \sup_{\substack{m \in \mathbb{N} \\ (x_1,y_1),\dots,(x_m,y_m) \in \Gamma}} \{y_m \cdot (x - x_m) + y_{m-1} \cdot (x_m - x_{m-1}) + \cdots + y_0 \cdot (x_1 - x_0)\}.$$

Εφόσον η φ ορίζεται ως το supremum αφινικών συναρτήσεων, προκύπτει ότι είναι μια κάτω ημισυνεχής κυρτή συνάρτηση. Ακόμα, από τον ορισμό της φ εύκολα βλέπουμε ότι $\varphi(x_0) \leq 0$ και έτσι η φ είναι, επίσης, κανονική.

<u>3ο βήμα:</u> Θα δείξουμε ότι $\Gamma \subseteq Graph(\partial\varphi)$.

Έστω, λοιπόν, $(x, y) \in \Gamma$. Πρέπει να ελέγξουμε ότι $y \in \partial\varphi(x)$, δηλαδή ότι

$$\forall z \in \mathbb{R}^n, \qquad \varphi(z) \geq \varphi(x) + y \cdot (z - x).$$

Επιλέγουμε τυχαίο $z \in \mathbb{R}^n$. Αρκεί να δείξουμε ότι για όλα τα $\alpha < \varphi(x)$, ισχύει

$$\varphi(z) \geq \alpha + y \cdot (z - x).$$

Αν, λοιπόν, $\alpha < \varphi(x)$, τότε από τον ορισμό της φ υπάρχουν $m \in \mathbb{N}$ και $(x_1, y_1), \dots, (x_m, y_m) \in \Gamma$ έτσι ώστε

$$\alpha \leq y_m \cdot (x - x_m) + \cdots + y_0 \cdot (x_1 - x_0).$$

Άρα,

$$\alpha + y \cdot (z - x) \leq y \cdot (z - x) + y_m \cdot (x - x_m) + \cdots + y_0 \cdot (x_1 - x_0).$$

Συμβολίζουμε $x = x_{m+1}$, $y = y_{m+1}$, οπότε προκύπτει

$$\alpha + y \cdot (z - x) \leq y_{m+1} \cdot (z - x_{m+1}) + y_m \cdot (x_{m+1} - x_m) + \cdots + y_0 \cdot (x_1 - x_0) \leq \varphi(z).$$

Συνοψίζοντας το 2ο και 3ο βήμα, για ένα οποιοδήποτε κυκλικά μονότονο σύνολο $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^n$, μπορούμε να κατασκευάσουμε μια κανονική κάτω ημισυνεχή κυρτή συνάρτηση φ τέτοια ώστε $\Gamma \subseteq Graph(\partial\varphi)$. Αυτό, σε συνδυασμό με το 1ο βήμα, ολοκληρώνει την απόδειξη του Θεωρήματος.

■

Από την Πρόταση 2.49 και το Θεώρημα 2.52, μπορούμε άμεσα να εξάγουμε το ακόλουθο θεώρημα, σύμφωνα με το οποίο ο φορέας κάθε βέλτιστου πλάνου μεταφοράς συγκεντρώνεται στο υποδιαφορικό μιας κανονικής κάτω ημισυνεχούς κυρτής συνάρτησης.

Θεώρημα 2.53

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ με πεπερασμένες ροπές 2ης τάξης. Θεωρούμε το πρόβλημα Kantorovich αναφορικά με την τετραγωνική συνάρτηση κόστους $c(x,y) = |x-y|^2$ και έστω $\pi \in \Pi(\mu,\nu)$ ένα βέλτιστο πλάνο μεταφοράς.

Τότε, υπάρχει κανονική κάτω ημισυνεχής κυρτή συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ τέτοια ώστε $Supp(\pi) \subseteq Graph(\partial\varphi)$.

Επιπλέον, με τη χρήση των ιδιοτήτων διαφορισιμότητας των κυρτών συναρτήσεων στον $\mathbb{R}^n$, αποδεικνύεται το παρακάτω πόρισμα, χωρίς οποιαδήποτε υπόθεση σχετικά με τις ροπές 2ης τάξης των μ και ν.

Πόρισμα 2.54 (Βελτιωμένη εκδοχή του Θεωρήματος Brenier, υπαρξιακό μέρος)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$. Υποθέτουμε ότι το μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue. Τότε υπάρχει κάτω ημισυνεχής κυρτή συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ τέτοια ώστε $\nabla\varphi \# \mu = \nu$.

Για την απόδειξη του πορίσματος παραπέμπουμε στο ([10], Main Theorem, σελ. 310).

2.3.3. Το Λήμμα του Aleksadrov.

Αντικείμενο αυτής της παραγράφου είναι η απόδειξη της μοναδικότητας του βέλτιστου πλάνου μεταφοράς υπό την υπόθεση ότι το μέτρο μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue, χωρίς να εμπλέξουμε καθόλου τη θεωρία της δυϊκότητας.

Λήμμα 2.55 (Το λήμμα του Aleksandrov)

Έστω $\varphi, \psi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ δύο κυρτές κάτω ημισυνεχείς συναρτήσεις, διαφορίσιμες στο σημείο x_0 με $\varphi(x_0) = \psi(x_0) = 0$ αλλά $\nabla\varphi(x_0) \neq \nabla\psi(x_0)$. Ορίζουμε

$$V \coloneqq \{x \in \mathbb{R}^n : \ \varphi(x) > \psi(x)\} \subseteq Dom(\psi) \quad \text{και} \quad Z \coloneqq \nabla\psi^{-1}\big(\partial\varphi(V)\big).$$

Τότε $Z \subseteq V$, ενώ το σημείο x_0 βρίσκεται σε θετική απόσταση από το σύνολο Z.

Απόδειξη:

Αρχικά, θα αποδείξουμε το λήμμα στην ειδική περίπτωση που $\nabla\psi(x_0) = 0$.

Για να αποδείξουμε τον εγκλεισμό $Z \subseteq V$, θεωρούμε $x \in Z$ και θέτουμε $y \coloneqq \nabla\psi(x)$. Τότε,

$$y \in \partial\varphi(V) = \bigcup_{m \in V} \partial\varphi(m),$$

άρα, υπάρχει $m \in V$ τέτοιο ώστε $y \in \partial\varphi(m)$.

Άρα, για κάθε $z \in \mathbb{R}^n$, ισχύει $\varphi(z) \geq \varphi(m) + y \cdot (z - m)$ και $\psi(z) \geq \psi(x) + y \cdot (z - x)$.

Ειδικότερα, για $z = m$ η δεύτερη ανισότητα γίνεται $\psi(m) \geq \psi(x) + y \cdot (m - x)$.

Όμως, $m \in V$, δηλαδή $\varphi(m) > \psi(m)$. Συνδυάζοντας την 1η, την 3η και την 4η ανισότητα, παίρνουμε:

$$\forall\, z \in \mathbb{R}^n, \qquad \varphi(z) > \psi(x) + y \cdot (z - x). \tag{2.54}$$

Θέτοντας $z = x$, παίρνουμε $\varphi(x) > \psi(x)$, δηλαδή $x \in V$.

Στη συνέχεια, υποθέτουμε ότι $x_0 \in \overline{Z}$. Αυτό σημαίνει ότι υπάρχει ακολουθία $x_\kappa \in Z$ με $x_k \to x_0$. Τότε, υπάρχει ακολουθία $m_k \in V$ τέτοια ώστε $\nabla\psi(x_k) \in \partial\varphi(m_k)$.

Αφού $\nabla\psi(x_0) = 0$, από την κυρτότητα της ψ έπεται ότι

$$\forall\, z \in \mathbb{R}^n, \qquad \psi(z) \geq \psi(x_0) + 0 \cdot (z - x_0) = \psi(x_0) = 0$$

ενώ από τη συνέχεια του υποδιαφορικού $\partial\psi$ προκύπτει ότι $\nabla\psi(x_k) \longrightarrow 0$.

Απ' την άλλη, εφόσον $\nabla\varphi(x_0) \neq 0$, συμπεραίνουμε ότι υπάρχει σημείο $q \in \mathbb{R}^n$, τέτοιο ώστε $\varphi(q) < \varphi(x_0) = 0$. Πράγματι, αν υποθέσουμε ότι για κάθε $z \in \mathbb{R}^n$ ισχύει $\varphi(z) \geq \varphi(x_0)$, τότε η συνάρτηση φ θα παρουσιάζει ελάχιστο στο σημείο x_0, με αποτέλεσμα $\nabla\varphi(x_0) = 0$.

Χρησιμοποιώντας τη σχέση (2.54) με $y_k = \nabla\psi(x_k)$ έχουμε

$$\forall\, z \in \mathbb{R}^n, \qquad \varphi(z) \geq \psi(x_k) + \nabla\psi(x_k) \cdot (z - x_k).$$

και αν θέσουμε $z = q$, παίρνουμε:

$$\varphi(q) > \psi(x_k) + \nabla\psi(x_k) \cdot (q - x_k) \geq 0 - |\nabla\psi(x_k)||q - x_k| = -|\nabla\psi(x_k)||q - x_k|.$$

Εφόσον $x_k \to x_0$ και $\nabla\psi(x_k) \to 0$, προκύπτει ότι $\varphi(q) \geq 0$, αποτέλεσμα το οποίο αντιφάσκει με την ανισότητα $\varphi(q) < \varphi(x_0) = 0$.

Συμπεραίνουμε ότι το $x_0 \notin \overline{Z}$, και συνεπώς βρίσκεται σε θετική απόσταση από το σύνολο Z.

Στη γενική περίπτωση που $\nabla\psi(x_0) \neq 0$, ανάγουμε την απόδειξη στην ειδική περίπτωση θεωρώντας τις συναρτήσεις $\varphi_1(x) = \varphi(x) + a \cdot x$ και $\psi_1(x) = \psi(x) + a \cdot x$ με $a = -\nabla\psi(x_0)$.

Τότε $\nabla\psi_1(x_0) = 0$ και $Z = \nabla\psi_1^{-1}\big(\partial\varphi_1(V)\big) = \{x \in \mathbb{R}^n :\ \nabla\psi_1(x) \in \partial\varphi_1(V)\}$.

Πράγματι, ισχύει $\partial\varphi_1(x) = \partial\varphi(x) + a$, γιατί

$$y \in \partial\varphi_1(x) \Leftrightarrow \forall u \in \mathbb{R}^n,\ \ \varphi_1(u) \geq \varphi_1(x) + y \cdot (u - x)$$

$$\Leftrightarrow \forall u \in \mathbb{R}^n,\ \ \varphi(u) + a \cdot u \geq \varphi(x) + a \cdot x + y \cdot (u - x).$$

Από την άλλη,

$$y \in \partial\varphi(x) + a \Leftrightarrow \forall u \in \mathbb{R}^n,\ \ \varphi(u) \geq \varphi(x) + (y - a) \cdot (u - x).$$

Άρα, $y \in \partial\varphi_1(x) \Leftrightarrow y \in \partial\varphi(x) + a$. Άρα, η φ είναι διαφορίσιμη στο x αν και μόνο αν αυτό ισχύει για την φ_1 και

$$y \in Z \Leftrightarrow \text{υπάρχει } x \in V \text{ με } \nabla\psi(x) \in \partial\varphi(x) \Leftrightarrow \text{υπάρχει } x \in V \text{ με } \nabla\psi_1(x) \in \partial\varphi(x) + a$$

$$\Leftrightarrow \text{υπάρχει } x \in V \text{ με } \nabla\psi(x) \in \partial\varphi_1(x).$$

■

Θεώρημα 2.56

Θεωρούμε $\varphi, \psi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ δύο κυρτές κάτω ημισυνεχείς συναρτήσεις, διαφορίσιμες στο σημείο p με $\varphi(p) = \psi(p)$, αλλά $\nabla\varphi(p) \neq \nabla\psi(p)$. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι το διάνυσμα $\nabla\varphi(p) - \nabla\psi(p)$ έχει διεύθυνση κατά μήκος του άξονα x_1. Τότε υπάρχει μια Lipschitz συνάρτηση $f: \mathbb{R}^{n-1} \to \mathbb{R}$ με $\|f\|_{Lip} = 1$ και περιοχή U του σημείου p έτσι ώστε για κάθε $x \in U$ ισχύει

$$\varphi(x) = \psi(x) \quad \Leftrightarrow \quad x_1 = f(x_2, \dots, x_n).$$

Για την απόδειξη του θεωρήματος αυτού θα χρειαστούμε τα παρακάτω δύο λήμματα.

Λήμμα 2.57

Έστω $f: \mathbb{R} \to \mathbb{R}$ μια συνάρτηση, η οποία παρουσιάζει τοπικό ελάχιστο στο σημείο t. Αν η f έχει πλευρικές παραγώγους f'_- από αριστερά και f'_+ από δεξιά στο σημείο t, τότε $f'_-(t) \leq 0 \leq f'_+(t)$.

Λήμμα 2.58

Θεωρούμε $\varphi, \psi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ δύο κάτω ημισυνεχείς κυρτές συναρτήσεις. Έστω, ακόμα, $p, q \in \mathbb{R}^n$ τέτοια ώστε $p, q \in \mathrm{Int}\big(Dom(\varphi)\big) \cap Int\big(Dom(\psi)\big)$ και $\psi(p) - \varphi(p) = \psi(q) - \varphi(q)$.

Τότε, για κάποιο σημείο $x = (1 - t)p + tq$ με $t \in (0, 1)$, υπάρχουν $y \in \partial\psi(x)$ και $v \in \partial\varphi(x)$ τέτοια ώστε $(y - v) \cdot (q - p) = 0$.

Απόδειξη:

Λόγω κυρτότητας των συναρτήσεων φ και ψ, για κάθε $t \in [0,1]$ ισχύει $(1-t)p + tq \in Int\big(Dom(\varphi)\big) \cap Int\big(Dom(\psi)\big)$. Άρα, η συνάρτηση $f(t) \coloneqq (\psi - \varphi)\big((1-t)p + tq\big)$ είναι συνεχής στο διάστημα $[0,1]$. Εφόσον, εξ υποθέσεως, ισχύει $\psi(p) - \varphi(p) = \psi(q) - \varphi(q)$, έχουμε $f(0) = f(1)$ και επομένως, για κάποιο $t_0 \in (0,1)$ η f παρουσιάζει ακρότατο. Θέτουμε $x_0 = (1-t_0)p + t_0 q$.

Αφού $x_0 \in Int\big(Dom(\psi)\big)$, το υποδιαφορικό $\partial\psi(x_0)$ είναι συμπαγές σύνολο ([13], 23.2 – 23.4). Άρα, η συνάρτηση $y \cdot (q-p)$, $y \in \partial\psi(x_0)$ παρουσιάζει μέγιστο σε κάποιο σημείο $y_1 \in \partial\psi(x_0)$ και ελάχιστο σε κάποιο σημείο $y_2 \in \partial\psi(x_0)$. Ομοίως, το σύνολο $\partial\varphi(x_0)$ είναι συμπαγές, οπότε η συνάρτηση $v \cdot (q-p)$, $v \in \partial\varphi(x_0)$ παρουσιάζει μέγιστο και ελάχιστο, αντίστοιχα, στα σημεία $v_1 \in \partial\varphi(x_0)$ και $v_2 \in \partial\varphi(x_0)$. Άρα,

$$f'_+(t) \leq (y_1 - v_2) \cdot (q-p) \quad \text{και} \quad f'_-(t) \geq (y_2 - v_1) \cdot (q-p).$$

Αν υποθέσουμε ότι η f παρουσιάζει ελάχιστο στο t_0, τότε από το προηγούμενο λήμμα, έχουμε:

$$(y_1 - v_2) \cdot (q-p) \leq 0 \leq (y_2 - v_1) \cdot (q-p).$$

Επιλέγουμε $\lambda \in [0,1]$ τέτοιο ώστε $[(1-\lambda)(y_1 - v_2) + \lambda(y_2 - v_1)] \cdot (q-p) = 0$ και θέτουμε

$$y \coloneqq (1-\lambda)y_1 + \lambda y_2, \qquad v \coloneqq \lambda v_1 + (1-\lambda)v_2.$$

Εφόσον τα σύνολα $\partial\psi(x_0)$ και $\partial\varphi(x_0)$ είναι κυρτά, έχουμε $y \in \partial\psi(x_0)$, $v \in \partial\varphi(x_0)$ και από την παραπάνω σχέση παίρνουμε $(y-v) \cdot (q-p) = 0$ όπως και επιθυμούσαμε.

■

Παρατήρηση: Στην παραπάνω απόδειξη αναφέραμε ότι αν $x_0 \in Int\big(Dom(\psi)\big)$, τότε το $\partial\psi(x_0)$ είναι συμπαγές σύνολο. Για την απόδειξη παραπέμπουμε στο ([13], 23.2 – 4).

Απόδειξη του Θεωρήματος 2.56:

Χωρίς βλάβη της γενικότητας, θεωρούμε ότι $p = 0$.

Θεωρούμε ακολουθία $x_k \to p$ και ακολουθίες $y_k \in \partial\psi(x_k)$ και $v_k \in \partial\varphi(x_k)$. Αφού οι συναρτήσεις φ, ψ είναι διαφορίσιμες στο σημείο p, λόγω της Πρότασης 2.25, θα έχουμε

$$y_k \to \nabla\psi(p) \quad \text{και} \quad v_k \to \nabla\varphi(p).$$

Εφόσον $\nabla\psi(p) - \nabla\varphi(p) = \lambda\hat{x}_1$, όπου $\lambda > 0$ και $\hat{x}_1$ το μοναδιαίο διάνυσμα στην κατεύθυνση του άξονα x_1, από την Πρόταση 2.27 προκύπτει ότι υπάρχει περιοχή W του σημείου p τέτοια ώστε για κάθε $x \in W$, $y \in \partial\psi(x)$, $v \in \partial\varphi(x)$ να ισχύει $y_1 - v_1 > |\pi_{n-1}(y) - \pi_{n-1}(v)|$. Σημειώνουμε ότι y_1 συμβολίζει την πρώτη συντεταγμένη του διανύσματος $y \in \mathbb{R}^n$ και $\pi_{n-1}(y) = (y_2, \dots, y_n)$.

Η περιοχή W μπορεί να επιλεγεί κυρτή με τις φ, ψ να παίρνουν πεπερασμένες τιμές στο W.

Ειδικότερα, $y_1 - v_1 > 0$, οπότε από το λήμμα 2.58 η συνάρτηση $h \coloneqq \psi - \varphi$ είναι γνησίως μονότονη κατά μήκος ευθυγράμμων τμημάτων στο W παράλληλων με τον άξονα x_1. Πράγματι, έστω a, b δύο σημεία σε ένα ευθύγραμμο τμήμα στο W παράλληλο με τον άξονα x_1, με $h(a) = h(b)$ δηλαδή $\psi(a) - \varphi(a) = \psi(b) - \varphi(b)$. Από το λήμμα 2.58, για κάποιο σημείο $x = (1-t)a + tb$, με $t \in (0,1)$, υπάρχουν $y \in \partial\psi(x), v \in \varphi(x)$ τέτοια ώστε $\langle y - v, b - a\rangle = 0$. Όμως, $\langle y - v, b - a\rangle = (y_1 - v_1)(b_1 - a_1)$ και αφού $b_1 - a_1 \neq 0$, έπεται ότι $y_1 - v_1 = 0$. Έχουμε καταλήξει σε αντίφαση. Άρα, η h είναι 1-1 πάνω στο ευθύγραμμο τμήμα με άκρα τα a, b, άρα αφού είναι συνεχής (ως κυρτή) είναι γνησίως μονότονη.

Παίρνουμε σημεία $p_+ \in W$ και $p_- \in W$ πάνω στον άξονα x_1 και εκατέρωθεν του σημείου $p = 0$. Λόγω της συνέχειας της συνάρτησης h, υπάρχει $r > 0$ τέτοιο ώστε $B_r(p_+) \subseteq W$, $B_r(p_-) \subseteq W$ και

$$h(x) > h(p) = 0, \quad x \in B_r(p_+)$$

$$h(x) < h(p) = 0, \quad x \in B_r(p_-).$$

Στη συνέχεια, η περιοχή W μπορεί να αντικατασταθεί με την $U = W \cap \{x \in \mathbb{R}^n : \ |\pi_{n-1}(x)| < r\}$.

Σταθεροποιούμε ένα σημείο $w \in B_r(0) \subseteq \mathbb{R}^{n-1}$. Τότε, η συνάρτηση h λαμβάνει θετικές και αρνητικές τιμές πάνω στην ευθεία $\{x \in U : \ \pi_{n-1}(x) = w\}$. Εφόσον, ακόμα, η h είναι συνεχής και γνήσια μονότονη, υπάρχει μοναδικό σημείο x πάνω σ' αυτή την ευθεία τέτοιο ώστε $h(x) = 0$. Ορίζουμε $f(w) \coloneqq x_1$ (x_1 η πρώτη συντεταγμένη του σημείου x).
Μένει να δείξουμε ότι η $f(w)$ ικανοποιεί το φράγμα Lipschitz στην $B_r(0)$. Εκτός του ότι είναι Lipschitz, δεν κάνουμε οποιαδήποτε άλλη υπόθεση για την f έξω από τη μπάλα $B_r(0)$. Η f βεβαίως μπορεί να επεκταθεί στον $\mathbb{R}^{n-1}$ χωρίς να αλλάξει η σταθερά Lipschitz.

Άρα, παίρνουμε $w, z \in B_r(0) \in \mathbb{R}^{n-1}$. Αφού $h\big((f(w), w)\big) = h\big((f(z), z)\big) = 0$, από το λήμμα 2.58 γνωρίζουμε ότι υπάρχει $x \in \mathbb{R}^n$ πάνω στο ευθύγραμμο τμήμα που ενώνει τα σημεία $(f(w), w)$ και $(f(z), z)$ καθώς και $y \in \partial\psi(x)$, $v \in \partial\varphi(x)$ έτσι ώστε

$$|y_1 - v_1||f(w) - f(z)| = \left|\big(\pi_{n-1}(y) - \pi_{n-1}(v)\big) \cdot (z - w)\right| \leq |\pi_{n-1}(y) - \pi_{n-1}(v)||z - w|.$$

Εφόσον η περιοχή U είναι κυρτή, έχουμε $x \in U$ και από τη σχέση

$$y_1 - v_1 > |\pi_{n-1}(y) - \pi_{n-1}(v)|$$

προκύπτει

$$|f(w) - f(z)| \leq \frac{|\pi_{n-1}(y) - \pi_{n-1}(v)|}{|y_1 - v_1|}|z - w| < \frac{|\pi_{n-1}(y) - \pi_{n-1}(v)|}{|\pi_{n-1}(y) - \pi_{n-1}(v)|}|z - w| = |z - w|.$$

■

Είμαστε, πλέον, έτοιμοι να διατυπώσουμε το κεντρικό θεώρημα αυτής της παραγράφου.

Θεώρημα 2.59 (Εκλεπτυσμένη εκδοχή του Θεωρήματος Brenier)

Έστω μ, ν δύο μέτρα πιθανότητας στον $\mathbb{R}^n$ και έστω ότι το μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue. Τότε υπάρχει μοναδική ($d\mu$ −σχεδόν παντού) κάτω ημισυνεχής κυρτή συνάρτηση $\varphi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ τέτοια ώστε $\nabla\varphi \# \mu = \nu$.
(Θεωρούμε δύο συναρτήσεις $\varphi, \overline{\varphi}$ ίσες όταν $\nabla\varphi = \nabla\overline{\varphi}$).

Απόδειξη:

Το Πόρισμα 2.54 εγγυάται την ύπαρξη της φ, οπότε μένει να αποδείξουμε τη μοναδικότητα αυτής. Υποθέτουμε, προς απαγωγή σε άτοπο, ότι υπάρχουν δύο κάτω ημισυνεχείς κυρτές συναρτήσεις φ και $\overline{\varphi}$ τέτοιες ώστε $\nabla\varphi \# \mu = \nabla\overline{\varphi} \# \mu = \nu$, και οι $\nabla\varphi, \nabla\overline{\varphi}$ δεν είναι ταυτοτικά ίσες στο $Supp(\mu)$. Κάνοντας χρήση του Θεωρήματος 2.56, μπορούμε να δείξουμε ότι το σύνολο $A = \{x \in \mathbb{R}^n : \varphi(x) = \overline{\varphi}(x) \text{ και } \nabla\varphi(x) \neq \nabla\overline{\varphi}(x)\}$ έχει μέτρο Lebesgue ίσο με 0.

Πράγματι: Οι $\varphi, \overline{\varphi}$ είναι διαφορίσιμες $\mathcal{L}^n$ −σχεδόν παντού στα $Int\big(Dom(\varphi)\big)$ και $Int\big(Dom(\overline{\varphi})\big)$ αντίστοιχα ([13], Θεώρημα 25.5). Άρα, αρκεί να δείξουμε ότι το σύνολο

$$B = \{x \in \mathbb{R}^n \ : \ \varphi(x) = \overline{\varphi}(x), \qquad \varphi, \overline{\varphi} \text{ διαφορίσιμες στο } x, \qquad \nabla\varphi(x) \neq \nabla\overline{\varphi}(x)\}$$

έχει μηδενικό μέτρο Lebesgue. Από το Θεώρημα 2.56, για κάθε $x \in B$ υπάρχει περιοχή U_x του x και Lipschitz συνάρτηση f έτσι ώστε $\{y \in U_x : \ \varphi(y) = \overline{\varphi}(y)\} = \{y \in U_x : \ y_1 = f(y_2, \dots, y_n)\}$ μετά από κατάλληλη επιλογή του συστήματος συντεταγμένων.

Άρα, για κάθε $x \in B$ υπάρχει περιοχή U_x του x για την οποία, σύμφωνα με την παρατήρηση μετά το τέλος της απόδειξης, ισχύει $\mathcal{L}^n(B \cap U_x) = 0$. Τα σύνολα U_x είναι ανοικτά και ισχύει

$$\bigcup U_x \supseteq B.$$

Για κάθε περιοχή U_x υπάρχει μπάλα B_x με κέντρο κάποιο σημείο με ρητές συντεταγμένες και ρητή ακτίνα έτσι ώστε $x \in B_x \subseteq U_x$. Περιοριζόμαστε στις μπάλες B_x και τότε βρίσκουμε αριθμήσιμο πλήθος αυτών των μπαλών B_i , $i = 1,2, \dots$ έτσι ώστε

$$\bigcup_{i \in \mathbb{N}} B_i \supseteq B.$$

Άρα,

$$\mathcal{L}^n(B) = \mathcal{L}^n\left(\left(\bigcup_{i \in \mathbb{N}} B_i\right) \cap B\right) = \mathcal{L}^n\left(\bigcup_{i \in \mathbb{N}} (B_i \cap B)\right) = 0.$$

Έστω $x_0 \in Supp(\mu)$ με $\varphi(x_0) = \overline{\varphi}(x_0) = 0$ και $\nabla\varphi(x_0) \neq \nabla\overline{\varphi}(x_0)$. Αφού το μ δεν δίνει μάζα στα σύνολα μηδενικού μέτρου Lebesgue και οι $\nabla\varphi, \nabla\overline{\varphi}$ είναι συνεχείς (σύμφωνα με το [13], Θεώρημα 25.5), ισχύει:

$$\mu[B_\varepsilon(x_0) \cap \{x \in \mathbb{R}^n : \ \varphi(x) > \overline{\varphi}(x)\}] > 0 \quad \text{για κάθε } \varepsilon > 0$$

ή

$$\mu[B_\varepsilon(x_0) \cap \{x \in \mathbb{R}^n : \varphi(x) < \overline{\varphi}(x)\}] > 0 \quad \text{για κάθε } \varepsilon > 0.$$

Υποθέτουμε ότι ισχύει το πρώτο, και τότε μπορούμε να εφαρμόσουμε το λήμμα 2.55. Κάνοντας χρήση του συμβολισμού που είχαμε στο παραπάνω λήμμα, έχουμε:

Εφόσον το x_0 βρίσκεται σε θετική απόσταση από το σύνολο Z, έπεται ότι το Z αποφεύγει ένα σύνολο θετικού μέτρου για το μ, με αποτέλεσμα $\mu[Z] < \mu[V]$.

Αυτό συνεπάγεται ότι $\nabla\varphi \# \mu \neq \nabla\overline{\varphi} \# \mu$, όπως και επιθυμούσαμε. Πράγματι,

$$\nabla\overline{\varphi} \# \mu[\nabla\varphi(V)] = \mu\left[\nabla\overline{\varphi}^{-1}\big(\nabla\varphi(V)\big)\right] = \mu[Z] < \mu[V] \leq \mu\left[\nabla\varphi^{-1}\big(\nabla\varphi(V)\big)\right] = \nabla\varphi \# \mu[\nabla\varphi(V)].$$

■

Παρατήρηση:

Αν $U \subseteq \mathbb{R}^{n-1}$ ανοιχτό και $f: U \longrightarrow \mathbb{R} \cup \{+\infty\}$ συνεχής, τότε για το σύνολο $\Gamma = \big\{\big(x, f(x)\big) : \ x \in U\big\}$ ισχύει $\mathcal{L}^n(\Gamma) = 0$, όπου $\mathcal{L}^n$ το μέτρο Lebesgue στον $\mathbb{R}^n$.

<u>Απόδειξη:</u>

Αρχικά, έστω ότι η f ορίζεται σε ένα συμπαγές σύνολο $K \subseteq \mathbb{R}^{n-1}$. Για κάθε $\varepsilon > 0$, υπάρχει $1 > \delta > 0$ τέτοιο ώστε: $|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$. Υπάρχουν σημεία $x_1, \dots, x_\rho$ ώστε

$$K \subseteq \bigcup_{i=1}^{\rho} B_\delta(x_i).$$

Ισχύει ότι $\big\{\big(x, f(x)\big) : \ x \in B_\delta(x_1)\big\} \subseteq B_\delta(x_1) \times [f(x_1) - \varepsilon, f(x_1) + \varepsilon]$ και γενικότερα

$$\big\{\big(x, f(x)\big) : \ x \in K\big\} \subseteq \bigcup_{i=1}^{\rho} B_\delta(x_i) \times [f(x_i) - \varepsilon, f(x_i) + \varepsilon].$$

Θεωρούμε τα σύνολα

$$A_1 = B_\delta(x_1)$$

$$A_2 = B_\delta(x_2) \setminus A_1$$

$$A_3 = B_\delta(x_3) \setminus (A_1 \cup A_2)$$

$$\vdots$$

Τα σύνολα A_i είναι ξένα μεταξύ τους, οπότε:

$$\mathcal{L}^n\big(\{(x,f(x)) : \ x \in K\}\big) \leq \mathcal{L}^n\left(\bigcup_{i=1}^{\rho} A_i \times [f(x_i)-\varepsilon, f(x_i)+\varepsilon]\right) \leq \sum_{i=1}^{\rho} \mathcal{L}^n(A_i)\cdot 2\varepsilon$$
$$= \mathcal{L}^n\left(\bigcup_{i=1}^{\rho} A_i\right)\cdot 2\varepsilon.$$

Τότε:

$$\bigcup_{i=1}^{\rho} A_i \subseteq K + B_\delta(0) \quad \text{και} \quad \mathcal{L}^n\left(\bigcup_{i=1}^{\rho} A_i\right) \leq \mathcal{L}^n\big(K + B_\delta(0)\big) \leq \mathcal{L}^n\big(K + B_1(0)\big) < +\infty.$$

Αφού το ε ήταν τυχαίος θετικός αριθμός, προκύπτει ότι $\mathcal{L}^n\big(\{(x,f(x)) : \ x \in K\}\big) = 0$.

Όταν η f ορίζεται σε ένα ανοιχτό $U \subseteq \mathbb{R}^{n-1}$, αυτό γράφεται ως αριθμήσιμη ένωση συμπαγών συνόλων K_i με $\mathcal{L}^n\big(\{(x,f(x)) : \ x \in K_i\}\big) = 0$ για κάθε i και έτσι προκύπτει, επίσης, ότι $\mathcal{L}^n(\Gamma) = 0$.

■

2.4. Γενίκευση σε άλλες συναρτήσεις κόστους.

Μπορούμε να γενικεύσουμε το Θεώρημα Βέλτιστης Μεταφοράς και για άλλες συναρτήσεις κόστους βασιζόμενοι είτε στη δυϊκότητα είτε και σε γεωμετρικά επιχειρήματα. Η βασική ιδέα είναι να χρησιμοποιήσουμε τις γενικευμένες έννοιες της κυρτότητας, της συζυγίας, του υποδιαφορικού, της κυκλικής μονοτονίας, κ.τ.λ ώστε να καταγράψουμε και να αποδείξουμε κάποιες βασικές προτάσεις.

2.4.1. Γενικές έννοιες

Ορισμός 2.60 (Γενικευμένες έννοιες της κυρτότητας)

Θεωρούμε δύο μη κενά σύνολα X και Y και μια συνάρτηση $c: X \times Y \to \mathbb{R}$.

Μια συνάρτηση $\varphi: X \to \mathbb{R} \cup \{-\infty\}$ λέγεται c −κοίλη αν υπάρχει συνάρτηση $\psi: Y \longrightarrow \mathbb{R} \cup \{-\infty\}$, όχι ταυτοτικά ίση με $-\infty$, τέτοια ώστε

$$\varphi(x) = \inf_{y\in Y}[c(x,y) - \psi(y)], \ \ x \in X.$$

Ορίζουμε το πεδίο ορισμού της φ να είναι το σύνολο

$$Dom(\varphi) \coloneqq \{x \in X : \ \ \varphi(x) \neq -\infty\}.$$

Ένα υποσύνολο $\Gamma \subseteq X \times Y$ λέγεται $c-$ κυκλικά μονότονο αν για κάθε πεπερασμένο αριθμό σημείων $(x_i, y_i) \in \Gamma,\ i = 1,2, \dots m$ και οποιαδήποτε μετάθεση σ του συνόλου $\{1,2, \dots m\}$, ισχύει

$$\sum_{i=1}^{m} c(x_i, y_i) \leq \sum_{i=1}^{m} c\big(x_{\sigma(i)}, y_i\big).$$

Το $c-$υπερδιαφορικό $\partial^c \varphi$ μιας $c-$ κοίλης συνάρτησης φ ορίζεται να είναι το σύνολο όλων των ζευγών $(x, y) \in X \times Y$ τέτοια ώστε

$$\forall z \in X, \qquad \varphi(z) \leq \varphi(x) + [c(z, y) - c(x, y)].$$

Για οποιαδήποτε συνάρτηση $\varphi: X \to \mathbb{R} \cup \{-\infty\}$, όχι ταυτοτικά ίση με $-\infty$, ορίζουμε τον $c-$ μετασχηματισμό αυτής μέσω της σχέσης

$$\varphi^c(y) = \inf_{x \in X} [c(x, y) - \varphi(x)].$$

Εναλλάσσοντας τους ρόλους των X και Y και κρατώντας τη συνάρτηση $c(x, y)$ σταθερή, παίρνουμε τους ανάλογους ορισμούς της $c-$κοιλότητας, του $c-$υπερδιαφορικού και του $c-$μετασχηματισμού μιας συνάρτησης $\psi: Y \to \mathbb{R} \cup \{-\infty\}$.

Λήμμα 2.61

Έστω $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ κοίλη συνάρτηση, $g, h :\ \mathbb{R}^n \longrightarrow \mathbb{R}$ συναρτήσεις της μορφής $g(x) = \alpha \cdot x + \beta$, $h(x) = \gamma \cdot x + \delta$, τα γραφήματα των οποίων είναι εφαπτόμενα υπερεπίπεδα του γραφήματος της φ, δηλαδή: $g(x) \geq \varphi(x)$, $h(x) \geq \varphi(x)$ για κάθε $x \in \mathbb{R}^n$ και υπάρχουν y, z με $g(y) = \varphi(y)$, $h(z) = \varphi(z)$.

Αν $\alpha = \gamma$, τότε $\beta = \delta$.

<u>Απόδειξη:</u>

Έστω ότι ισχύει $\beta > \delta$. Τότε, $h(x) < g(x),\ x \in \mathbb{R}^n$. Το γράφημα της φ περιέχεται στον ημιχώρο $H = \{(x, t) \in \mathbb{R}^n : t \leq h(x)\}$. Αφού $g(y) > h(y)$, έπεται ότι $\big(y, \varphi(y)\big) = \big(y, g(y)\big) \notin H$, το οποίο είναι άτοπο.

■

Πρόταση 2.62

Έστω $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ κοίλη συνάρτηση. Τότε υπάρχει συνάρτηση $\psi :\ \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{-\infty\}$ τέτοια ώστε:

$$\varphi(x) = \inf_{y \in \mathbb{R}^n} \big(x \cdot y - \psi(y)\big),\ \ x \in \mathbb{R}^n.$$

Απόδειξη:

Έστω $z \in \mathbb{R}^n$. Από την Πρόταση 2.12, το $\partial(-\varphi)(z)$ δεν είναι κενό, άρα υπάρχει $u \in \mathbb{R}^n$ έτσι ώστε $\varphi(x) \leq \varphi(z) + u \cdot (x - z),\ \ x \in \mathbb{R}^n$. Ορίζουμε μια συνάρτηση $h : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ απεικονίζοντας κάθε $z \in \mathbb{R}^n$ σε κάποιο u με την παραπάνω ιδιότητα. Τότε, για κάθε $x, z \in \mathbb{R}^n$ ισχύει

$$\varphi(x) \leq x \cdot h(z) - \big(z \cdot h(z) - \varphi(z)\big).$$

Αφού για $x = z$ η παραπάνω σχέση ισχύει ως ισότητα, έπεται ότι για κάθε $x \in \mathbb{R}^n$

$$\varphi(x) = \inf_{z\in\mathbb{R}^n}\Big(x \cdot h(z) - \big(z \cdot h(z) - \varphi(z)\big)\Big).$$

Από το λήμμα 2.61 έπεται ότι: αν $h(z_1) = h(z_2)$, τότε $z_1 \cdot h(z_1) - \varphi(z_1) = z_2 \cdot h(z_2) - \varphi(z_2)$. Άρα, υπάρχει συνάρτηση $\psi : Im(h) \longrightarrow \mathbb{R}^n$ έτσι ώστε $\psi\big(h(z)\big) = z \cdot h(z) - \varphi(z),\ \ z \in \mathbb{R}^n$, όπου $Im(h) = \{h(z) : \ \ z \in \mathbb{R}^n\}$.

Παίρνουμε ότι για κάθε $x \in \mathbb{R}^n$ ισχύει $\varphi(x) = \inf_{y\in Im(h)}\big(x \cdot y - \psi(y)\big)$. Επεκτείνουμε την ψ στο $\mathbb{R}^n$ θέτοντας $\psi(y) = -\infty,$ αν $y \notin Im(h)$, οπότε παίρνουμε το ζητούμενο.

■

Παρατηρήσεις 2.63

Α) Στην περίπτωση που $X = Y = \mathbb{R}^n$ και $c(x, y) = x \cdot y$, μια συνάρτηση φ είναι c −κοίλη, αν υπάρχει συνάρτηση $\psi: \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ τέτοια ώστε

$$\forall x \in \mathbb{R}^n, \qquad \varphi(x) = \inf_{y\in\mathbb{R}^n}[x \cdot y - \psi(y)].$$

Επομένως, αν η φ παίρνει πραγματικές τιμές, η συνάρτηση φ είναι c −κοίλη αν και μόνο αν η φ είναι κοίλη.

Β) Στην περίπτωση που $X = Y = \mathbb{R}^n$, η φ παίρνει πραγματικές τιμές και $c(x, y) = |x - y|^2$, η συνάρτηση φ είναι c −κοίλη αν και μόνο αν η συνάρτηση $\varphi(x) - |x|^2$ είναι κοίλη.

Απόδειξη του (Β):

Για το ευθύ, υποθέτουμε ότι η φ είναι c −κοίλη. Τότε, υπάρχει συνάρτηση $\psi: \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ έτσι ώστε $\varphi(x) = \inf_{y\in\mathbb{R}^n}[|x - y|^2 - \psi(y)]$. Τότε,

$$\begin{aligned}\varphi(x) - |x|^2 &= \inf_{y\in\mathbb{R}^n}[|x - y|^2 - |x|^2 - \psi(y)] = \inf_{y\in\mathbb{R}^n}[|x|^2 - 2x \cdot y + |y|^2 - |x|^2 - \psi(y)] \\ &= \inf_{y\in\mathbb{R}^n}[x \cdot (-2y) + |y|^2 - \psi(y)]\end{aligned}$$

και άρα η συνάρτηση $\varphi(x) - |x|^2$ είναι κοίλη.

Αντίστροφα, υποθέτουμε ότι η συνάρτηση $\varphi(x) - |x|^2$ είναι κοίλη. Τότε, από την Πρόταση 2.62, υπάρχει συνάρτηση $\psi: \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ έτσι ώστε $\varphi(x) - |x|^2 = \inf_{y \in \mathbb{R}^n} [x \cdot y - \psi(y)]$. Επομένως,

$$\varphi(x) = \inf_{y \in \mathbb{R}^n} [|x|^2 + x \cdot y - \psi(y)] = \inf_{y \in \mathbb{R}^n} \left[|x|^2 - 2x \cdot \left(-\frac{1}{2} y \right) + \frac{1}{4} |y|^2 - \left(\frac{1}{4} |y|^2 + \psi(y) \right) \right]$$

$$= \inf_{y \in \mathbb{R}^n} \left[\left| x - \left(-\frac{y}{2} \right) \right|^2 - h(y) \right]$$

Κάνοντας την αντικατάσταση $t = -\frac{y}{2}$, παίρνουμε $\varphi(x) = \inf_{t \in \mathbb{R}^n} [|x - t|^2 - h(-2t)]$ και συνεπώς η συνάρτηση φ είναι c −κοίλη.

■

Έχοντας ολοκληρώσει τους παραπάνω ορισμούς, είμαστε έτοιμοι να επεκτείνουμε ένα μέρος του Θεωρήματος 2.37 σε ένα αρκετά γενικό σκηνικό.

Θεώρημα 2.64 (Αφηρημένη δυϊκότητα Legendre)

Θεωρούμε δύο μη κενά σύνολα X, Y, μια συνάρτηση $c: X \times Y \to \mathbb{R}$ και συνάρτηση $\varphi: X \to \mathbb{R} \cup \{-\infty\}$ όχι ταυτοτικά ίση με $-\infty$. Ισχύουν τα παρακάτω:

a) Για κάθε $(x, y) \in X \times Y$, $\varphi(x) + \varphi^c(y) \leq c(x, y)$.

b) Έστω ότι η φ^c δεν είναι ταυτοτικά ίση με $-\infty$. Τότε ορίζεται η φ^{cc} και $\varphi^{cc} \geq \varphi$.

c) Έστω ότι οι φ^c και φ^{cc} δεν είναι ταυτοτικά ίσες με $-\infty$. Τότε, $\varphi^{cc} = \varphi$ αν και μόνο αν η φ είναι c −κοίλη.

d) Για κάθε $y \in Y$, η συνάρτηση $h(x) = c(x, y)$ είναι c −κοίλη.

e) Το infimum μιας οικογένειας c −κοίλων συναρτήσεων, αν δεν είναι ταυτοτικά ίσο με $-\infty$, είναι μια c −κοίλη συνάρτηση.

<u>Απόδειξη:</u>

a) Έστω $(x_0, y_0) \in X \times Y$. Τότε,

$$\varphi(x_0) + \varphi^c(y_0) = \varphi(x_0) + \inf_{x \in X} [c(x, y_0) - \varphi(x)] \leq \varphi(x_0) + c(x_0, y_0) - \varphi(x_0) = c(x_0, y_0).$$

b) Από το (a), για κάθε $(x, y) \in X \times Y$ ισχύει η ανισότητα $c(x, y) - \varphi^c(y) \geq \varphi(x)$.
Άρα, για κάθε $x \in X$, $\inf_{y \in Y} [c(x, y) - \varphi^c(y)] \geq \varphi(x)$ δηλαδή για κάθε $x \in X$, $\varphi^{cc}(x) \geq \varphi(x)$.

c) Καταρχήν, θα ελέγξουμε ότι $\varphi^{ccc} = \varphi^c$.
Πράγματι, έχοντας υπόψη την ανισότητα $\varphi^{cc} \geq \varphi$, έχουμε

$$\varphi^{ccc}(y) = \inf_{x\in X}[c(x,y) - \varphi^{cc}(x)] \leq \inf_{x\in X}[c(x,y) - \varphi(x)] = \varphi^c(y).$$

Για την απόδειξη της αντίστροφης ανισότητας, αρκεί στην ανισότητα $\varphi^{cc} \geq \varphi$ να βάλουμε τη συνάρτηση φ^c στη θέση της φ και τότε παίρνουμε $(\varphi^c)^{cc} \geq \varphi^c$.

Αν η φ είναι c −κοίλη και ψ όπως στον ορισμό των c −κοίλων συναρτήσεων, τότε $\varphi = \psi^c$, άρα $\varphi^{cc} = \psi^{ccc} = \psi^c = \varphi$. Επίσης, για κάθε ψ η ψ^c είναι c −κοίλη, άρα αν $\varphi = \varphi^{cc}$, η φ είναι $c -$ κοίλη.

d) Σταθεροποιούμε ένα $y_0 \in Y$. Για να δείξουμε ότι η $h(x) = c(x, y_0)$ είναι c −κοίλη, αρκεί να υπάρχει $\psi: Y \to \mathbb{R} \cup \{-\infty\}$, όχι ταυτοτικά ίση με $-\infty$, τέτοια ώστε για κάθε $x \in X$, $c(x, y_0) = \inf_{y\in Y}[c(x,y) - \psi(y)]$.

Αυτό πράγματι συμβαίνει αν πάρουμε ως συνάρτηση ψ την $\psi(y) = \inf_{z\in X}[c(z,y) - c(z,y_0)]$, για την οποία ισχύει $\psi(y_0) = 0$.

e) Θεωρούμε $(\varphi_i)_{i\in I}$ μια οικογένεια c −κοίλων συναρτήσεων. Τότε υπάρχουν συναρτήσεις ψ_i : $Y \longrightarrow \mathbb{R} \cup \{-\infty\}$, όχι ταυτοτικά ίσες με $-\infty$, έτσι ώστε

$$\varphi_i(x) = \inf_{y\in Y}[c(x,y) - \psi_i(y)], \quad x \in X.$$

Ορίζουμε $\varphi_i(x) = \inf_{i\in I} \varphi_i(x)$. Τότε,

$$\varphi(x) = \inf_{i\in I}\left(\inf_{y\in Y}[c(x,y) - \psi_i(y)]\right) = \inf_{y\in \mathrm{Y}}\left(\inf_{i\in I}[c(x,y) - \psi_i(y)]\right) = \inf_{y\in \mathrm{Y}}\left(c(x,y) + \inf_{i\in I}[-\psi_i(y)]\right)$$
$$= \inf_{y\in \mathrm{Y}}\left(c(x,y) - \sup_{i\in I}[\psi_i(y)]\right).$$

Άρα, η φ είναι μια c −κοίλη συνάρτηση.

■

Στο σημείο αυτό θα είναι χρήσιμο να κάνουμε μερικές υπενθυμίσεις από το 1ο κεφάλαιο σχετικά με το αρχικό πρόβλημα Kantorovich και το δυϊκό πρόβλημα αυτού.

Το αρχικό πρόβλημα Kantorovich συνίσταται στην ελαχιστοποίηση του συναρτησοειδούς

$$I[\pi] = \int_{X\times Y} c(x,y)\, d\pi(x,y)$$

πάνω στο σύνολο $\Pi(\mu,\nu)$ των μέτρων πιθανότητας στον $X \times Y$ που έχουν περιθώρια μέτρα μ στον X και ν στον Y. Το δυϊκό πρόβλημα συνίσταται στη μεγιστοποίηση του συναρτησοειδούς

$$J(\varphi,\psi) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu$$

πάνω στο σύνολο Φ_c όλων των ζευγών μετρήσιμων συναρτήσεων $(\varphi,\psi)\in L^1(d\mu)\times L^1(d\nu)$ που παίρνουν τιμές στο $\mathbb{R}\cup\{-\infty\}$ και ικανοποιούν την ανισότητα $\varphi(x)+\psi(y)\le c(x,y)$ για σχεδόν όλα τα $x\in X$ και σχεδόν όλα τα $y\in Y$. Αν αλλάξουμε τις τιμές των συναρτήσεων φ και ψ πάνω σε σύνολα μηδενικού μέτρου Lebesgue και τις θέσουμε ίσες με $-\infty$, τότε η ανισότητα $\varphi(x)+\psi(y)\le c(x,y)$ αληθεύει για κάθε x,y.

Πρόταση 2.65 (Τα βέλτιστα πλάνα μεταφοράς χαρακτηρίζονται από τον φορέα τους)

Έστω X,Y δύο Πολωνικοί χώροι, $\mu\in P(X)$, $\nu\in P(Y)$ δύο Borel μέτρα πιθανότητας και $c: X\times Y\to\mathbb{R}_+$ μια κάτω ημισυνεχής συνάρτηση. Έστω, ακόμα, ότι το ζεύγος $(\varphi_0,\psi_0)\in\Phi_c$ μεγιστοποιεί το συναρτησοειδές J στο δυϊκό πρόβλημα Kantorovich.

Τότε, το $\pi_0\in\Pi(\mu,\nu)$ είναι ένα βέλτιστο πλάνο μεταφοράς για το αρχικό πρόβλημα Kantorovich αν και μόνο αν συγκεντρώνεται στο σύνολο $\{(x,y)\in X\times Y:\ \ \varphi_0(x)+\psi_0(y)=c(x,y)\}$.

<u>Απόδειξη:</u>

Έστω π_0 είναι βέλτιστο πλάνο μεταφοράς. Τότε, σύμφωνα με τη δυϊκότητα Kantorovich,

$$I[\pi_0]=\inf_{\pi\in\Pi(\mu,\nu)} I[\pi]=\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi)=J(\varphi_0,\psi_0).$$

Άρα,

$$\int_{X\times Y} c(x,y)\,d\pi_0(x,y)=\int_X \varphi_0\,d\mu+\int_Y \psi_0\,d\nu=\int_{X\times Y}[\varphi_0(x)+\psi_0(y)]\,d\pi_0(x,y)$$

ή ισοδύναμα,

$$\int_{X\times Y}[c(x,y)-\varphi_0(x)-\psi_0(y)]\,d\pi_0(x,y)=0.$$

Εξ ορισμού του συνόλου Φ_c, έχουμε $\varphi_0(x)+\psi_0(y)\le c(x,y)$, οπότε η υπό ολοκλήρωση συνάρτηση είναι μεγαλύτερη ή ίση από το 0. Προκύπτει ότι

$$\varphi_0(x)+\psi_0(y)=c(x,y),\qquad d\pi_0-\text{σχεδόν παντού στον } X\times Y,$$

δηλαδή το π_0 συγκεντρώνεται στο σύνολο $\{(x,y)\in X\times Y:\ \ \varphi_0(x)+\psi_0(y)=c(x,y)\}$.

Αντίστροφα, έστω ότι το π_0 συγκεντρώνεται στο σύνολο $\{(x,y)\in X\times Y:\ \ \varphi_0(x)+\psi_0(y)=c(x,y)\}$. Εύκολα βλέπουμε ότι

$$\int_{X\times Y} c(x,y)\,d\pi_0(x,y)=\int_X \varphi_0\,d\mu+\int_Y \psi_0\,d\nu,$$

δηλαδή $I[\pi_0]=J(\varphi_0,\psi_0)$.

Αφού $\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) \le \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]$, έπεται ότι το π_0 είναι βέλτιστο πλάνο μεταφοράς.

■

Στη συνέχεια, αποδεικνύουμε ότι τα βέλτιστα πλάνα μεταφοράς έχουν c −κυκλικά μονότονο φορέα βασιζόμενοι μόνο στη δυϊκότητα Kantorovich.

Πρόταση 2.66

Έστω X, Y δύο Πολωνικοί χώροι, $\mu \in P(X)$, $\nu \in P(Y)$ δύο Borel μέτρα πιθανότητας και $c: X \times Y \to \mathbb{R}_+$ μια κάτω ημισυνεχής συνάρτηση. Υποθέτουμε ότι $\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] < +\infty$ και συμβολίζουμε με π_* ένα βέλτιστο πλάνο μεταφοράς για το αρχικό πρόβλημα Kantorovich.

Α. Υποθέτουμε ότι υπάρχει βέλτιστο ζεύγος (φ_*, ψ_*) συζυγών c −κοίλων συναρτήσεων για το δυϊκό πρόβλημα Kantorovich. Έστω $n \ge 1$ και σ μια μετάθεση του συνόλου $\{1,2,\dots,n\}$. Τότε:

i. Για οποιαδήποτε σημεία $(x_1, y_1), \dots, (x_n, y_n) \in X \times Y$, ισχύει

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \ge \sum_{i=1}^{n} [\varphi_*(x_i) + \psi_*(y_i)].$$

ii. Για $d\pi_*$ −σχεδόν όλα τα $(x,y) \in X \times Y$, ισχύει $\varphi_*(x) + \psi_*(y) = c(x,y)$.

iii. Το π_* συγκεντρώνεται σε ένα c − κυκλικά μονότονο υποσύνολο του $X \times Y$.

Β. Αν δεν ικανοποιείται η υπόθεση του σκέλους (Α), επιλέγουμε μια ακολουθία $(\varphi_n, \psi_n) \in \Phi_c$ τέτοια ώστε $J(\varphi_n, \psi_n) \longrightarrow \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi)$. Τότε:

i. Η ακολουθία των μη αρνητικών συναρτήσεων $c_n(x,y) = c(x,y) - \varphi_n(x) - \psi_n(y)$ συγκλίνει στο 0 στον $L^1(d\pi_*)$.

ii. Υπάρχει υπακολουθία n_k και σύνολο $Z \subseteq X \times Y$ έτσι ώστε $\pi_*[Z] = 1$ και η ακολουθία $c_{n_k}(x,y) = c(x,y) - \varphi_{n_k}(x) - \psi_{n_k}(y)$ να συγκλίνει στο 0 για κάθε $(x,y) \in Z$.

iii. Το π_* συγκεντρώνεται σε ένα c −κυκλικά μονότονο υποσύνολο του $X \times Y$.

<u>Απόδειξη:</u>

Α. Θεωρούμε $n \ge 1$, σημεία $(x_1, y_1), \dots, (x_n, y_n) \in X \times Y$ και σ μια οποιαδήποτε μετάθεση του συνόλου $\{1,2,\dots,n\}$. Τότε, θα έχουμε:

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \geq \sum_{i=1}^{n} \big[\varphi_*(x_i) + \psi_*\big(y_{\sigma(i)}\big)\big] = \sum_{i=1}^{n} [\varphi_*(x_i) + \psi_*(y_i)],$$

όπου η ανισότητα ισχύει αφού $(\varphi_*, \psi_*) \in \Phi_c$.

Από τη δυϊκότητα Kantorovich, έχουμε $I[\pi_*] = J(\varphi_*, \psi_*)$, δηλαδή

$$\int_{X\times Y} c(x,y)\, d\pi_*(x,y) = \int_X \varphi_*\, d\mu + \int_Y \psi_*\, d\nu = \int_{X\times Y} [\varphi_*(x) + \psi_*(y)]\, d\pi_*(x,y),$$

και επομένως,

$$\int_{X\times Y} [c(x,y) - \varphi_*(x) - \psi_*(y)]\, d\pi_*(x,y) = 0.$$

Επειδή, $(\varphi_*, \psi_*) \in \Phi_c$, η υπό ολοκλήρωση συνάρτηση είναι μη αρνητική και συνεπώς

$$\varphi_*(x) + \psi_*(y) = c(x,y), \qquad d\pi_* - \text{σχεδόν παντού στον } X \times Y.$$

Συνδυάζοντας τα δύο προηγούμενα αποτελέσματα, παίρνουμε ότι το π_* συγκεντρώνεται στο σύνολο $\{(x,y) \in X \times Y : \ \varphi_*(x) + \psi_*(y) = c(x,y)\}$, για τα σημεία του οποίου, ισχύει

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \geq \sum_{i=1}^{n} c(x_i, y_i),$$

είναι, δηλαδή, ένα $c -$ κυκλικά μονότονο υποσύνολο του $X \times Y$.

B. Από τη δυϊκότητα Kantorovich και την υπόθεση, έχουμε

$$\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = I[\pi_*] = \int_{X\times Y} c(x,y)\, d\pi_*(x,y).$$

Άρα,

$$J(\varphi_n, \psi_n) \longrightarrow \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = \int_{X\times Y} c(x,y)\, d\pi_*(x,y)$$

οπότε

$$\int_{X\times Y} c(x,y)\, d\pi_*(x,y) - J(\varphi_n, \psi_n) \longrightarrow 0$$

δηλαδή

$$\int_{X\times Y} c_n(x,y)\, d\pi_*(x,y) = \int_{X\times Y} c(x,y)\, d\pi_*(x,y) - \int_X \varphi_n\, d\mu - \int_Y \psi_n\, d\nu \longrightarrow 0.$$

Συνεπώς, αφού $\pi_* \in \Pi(\mu,\nu)$ και οι συναρτήσεις $c_n(x,y) = c(x,y) - \varphi_n(x) - \psi_n(y)$ είναι μη αρνητικές εξ' ορισμού του συνόλου Φ_c, έχουμε:

$$\int_{X\times Y} |c(x,y) - \varphi_n(x) - \psi_n(y)|\, d\pi_*(x,y) \longrightarrow 0.$$

Στη συνέχεια, η σύγκλιση κατά L^1 συνεπάγεται την κατά σημείο σύγκλιση μιας υπακολουθίας της $c_n(x,y) = c(x,y) - \varphi_n(x) - \psi_n(y)$ $d\pi_*$ −σχεδόν παντού στον $X \times Y$. Δηλαδή, υπάρχει υπακολουθία n_k και σύνολο $Z \subseteq X \times Y$ με $\pi_*[Z] = 1$ έτσι ώστε

$$c_{n_k}(x,y) = c(x,y) - \varphi_{n_k}(x) - \psi_{n_k}(y) \longrightarrow 0, \qquad \text{για κάθε } (x,y) \in Z.$$

Τέλος, θεωρούμε σημεία $(x_1,y_1),\dots,(x_n,y_n) \in Z$, $n \geq 1$ και σ μια οποιαδήποτε μετάθεση του συνόλου $\{1,2,\dots,n\}$. Τότε, για κάθε $k \in \mathbb{N}$, έχουμε:

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \geq \sum_{i=1}^{n}\big[\varphi_{n_k}(x_i) + \psi_{n_k}\big(y_{\sigma(i)}\big)\big] = \sum_{i=1}^{n}\big[\varphi_{n_k}(x_i) + \psi_{n_k}(y_i)\big] = \sum_{i=1}^{n}\big[c(x_i,y_i) - c_{n_k}(x_i,y_i)\big].$$

Παίρνοντας το όριο καθώς το $k \to +\infty$, προκύπτει η σχέση

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \geq \sum_{i=1}^{n} c(x_i, y_i),$$

δηλαδή το π_* συγκεντρώνεται στο σύνολο Z, το οποίο είναι ένα c −κυκλικά μονότονο υποσύνολο του $X \times Y$.

∎

Πόρισμα 2.67

Έστω X, Y δύο Πολωνικοί χώροι, $\mu \in P(X)$ και $\nu \in P(Y)$ δύο Borel μέτρα πιθανότητας και $c: X \times Y \to \mathbb{R}_+$ μια συνεχής συνάρτηση. Υποθέτουμε ότι $\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] < +\infty$.

Τότε υπάρχει c −κυκλικά μονότονο σύνολο $S \subseteq X \times Y$ στο οποίο συγκεντρώνονται όλα τα βέλτιστα πλάνα μεταφοράς $\pi \in \Pi(\mu,\nu)$.

<u>Απόδειξη:</u>

Θεωρούμε το σύνολο

$$S = \bigcup_{\substack{\pi\in\Pi(\mu,\nu) \\ \pi \ \text{βέλτιστο}}} Supp(\pi).$$

Θα δείξουμε ότι το S είναι c −κυκλικά μονότονο. Θεωρούμε $n \geq 1$, σημεία $(x_i, y_i) \in S,\ \ 1 \leq i \leq n$ και μια μετάθεση σ του συνόλου $\{1,2,\dots,n\}$. Για κάθε i, από τον ορισμό του συνόλου S, υπάρχει βέλτιστο πλάνο μεταφοράς $\pi_i \in \Pi(\mu,\nu)$ με $(x_i, y_i) \in Supp(\pi_i)$. Ορίζουμε ένα νέο πλάνο μεταφοράς μέσω του κυρτού συνδυασμού

$$\pi_* = \frac{1}{n}\sum_{i=1}^{n} \pi_i.$$

Εφόσον το $\Pi(\mu,\nu)$ είναι κυρτό σύνολο και το I γραμμικό συναρτησοειδές, έχουμε $\pi_* \in \Pi(\mu,\nu)$ και $I[\pi_*] = I[\pi_i]$ για κάθε $i = 1, 2, \dots, n$. Άρα, το π_* είναι, επίσης, ένα βέλτιστο πλάνο μεταφοράς. Άρα, το $Supp(\pi_*)$ είναι c −κυκλικά μονότονο σύνολο ([9], Θεώρημα 2.3). Το $Supp(\pi_*)$ περιέχει τα σύνολα $Supp(\pi_i)$ για κάθε $i = 1,2,\dots,n$ και ειδικότερα περιέχει τα σημεία $(x_i, y_i),\ \ i = 1,2,\dots,n$.

Άρα, επαληθεύται η ανισότητα

$$\sum_{i=1}^{n} c\big(x_i, y_{\sigma(i)}\big) \geq \sum_{i=1}^{n} c(x_i, y_i)$$

και, τελικά, το σύνολο S είναι c −κυκλικά μονότονο.

■

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] VILLANI C., *Topics in Optimal Transportation*. vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, 2003.

[2] ALBERTI G. & AMBROSIO L., *A geometrical approach to monotone functions in* $\mathbb{R}^n$, Math. Z. 230, 2 (1999), 259-316.

[3] AMBROSIO L. & PRATELLI A., *Existence and stability results in the* L^1 *theory of optimal transportation*. To appear in the proceedings of the CIME course held in Martina Franca (Italy), "Optimal transportation and application", 2 – 9 September 2001, Lecture Notes in Mathematics.

[4] BILLINGSLEY P., *Convergence of probability measures*, second ed. John Wiley & Sons Inc., New York, 1999. A Wiley – Interscience Publication.

[5] DUDLEY R. M., *Real analysis and probability*, vol. 74 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[6] EVANS L. C., *Partial differential equations and Monge-Kantorovich mass transfer*. In Current developments in mathematics, 1997 (Cambridge, MA). Int. Press, Boston, MA, 1999, pp. 65 – 126.

[7] EVANS L. C. & GARIEPY R. F., *Measure theory and fine properties of functions*. CRC Press, Boca Raton, FL, 1992.

[8] GANGBO W. & McCANN R. J., *Optimal maps in Monge's mass transportation problem.* C. R. Acad. SCI. Paris Ser. I Math. 321, 12 (1995), 1653 – 1658.

[9] GANGBO W. & McCANN R. J., *The geometry of optimal transportation*. Acta Math. 177, 2 (1996), 113 – 161.

[10] McCANN R. J., *Existence and uniqueness of monotone measure–preserving maps.* Duke Math. J. 80, 2 (1995), 309 – 323.

[11] McCANN R. J., *A convexity principle for interacting gases*. Adv. Math. 128, 1 (1997), 153 – 179.

[12] RACHEV S. & RUSCHENDORF L., *Mass Transportation Problems. Vol. I: Theory, Vol. II: Applications. Probability and its applications*. Springer – Verlag, New York, 1998.

[13] ROCKAFELLAR R. T., *Convex Analysis*. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[14] RUDIN W., *Real and complex analysis*, third ed. McGraw-Hill Book Co., New York, 1987.

[15] RUSCHENDORF L., *On c–optimal random variables*. Statist. Probab. Lett. 27, 3 (1996), 267 – 270.

[16] BREZIS H., *Συναρτησιακή Ανάλυση, Θεωρία και Εφαρμογές*. Μετάφραση – Επιμέλεια: Δημήτρης Κραββαρίτης, καθηγητής Ε.Μ.Π & Ίων Χρυσοβέργης, επίκ. καθηγητής Ε.Μ.Π, Πανεπιστημιακές Εκδόσεις Ε.Μ.Π, Αθήνα 1997.

[17] GRUBER M. PETER, *Convex and Discrete Geometry*, Volume 336, A series of Comprehensive Studies in Mathematics, Springer.

[18] ONNO VAN GAANS, *Probability measures on metric spaces*.

[19] DURRETT RICHARD, *Probability: Theory and Examples,* Fourth Edition, Cambridge University Press.