Εθνικο Μετσοβιο Πολυτεχνειο

Μέθοδοι Στατιστικής Φυσικής στη Θεωρητική Πληροφορική

Διαμαντίδης Δημήτρης

Επιβλέπων Καθηγητής: Λουλάκης Μιχαήλ, Αναπληρωτής Καθηγητής ΕΜΠ

Οκτώβριος 2019

Εθνικο Μετσοβιο Πολυτεχνειο

Μέθοδοι Στατιστικής Φυσικής στη Θεωρητική Πληροφορική

Διαμαντίδης Δημήτρης

Τριμελής Επιτροπή

Λουλάκης Μιχαήλ, Αναπληρωρής Καθηγητής ΕΜΠ

Παπανικολάου Βασίλειος, Καθηγητής ΕΜΠ

Φουσκάκης Δημήτριος, Αναπληρωτής Καθηγητής ΕΜΠ

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντά μου καθηγητή κ. Λουλάκη Μιχαήλ, που μου ανέθεσε ένα τόσο ενδιαφέρον και ξεχωριστό θέμα και με βοήθησε στην ολοκλήρωσή του. Επίσης, θα ήθελα να ευχαριστήσω τα μέλη τις τριμελούς μου επιτροπής ξεχωριστά για την συνεισφορά τους στην παρουσίαση της διπλωματικής μου εργασίας. Τέλος, ευχαριστώ την οικογένεια και τους φίλους μου που στάθηκαν στο πλευρό μου καθ όλη την διάρκεια των σπουδών μου.

Περιεχόμενα

Πρόλογος

Σκοπός της παρούσας διπλωματικής είναι η μελέτη συνδυαστικών προβλημάτων με τις πιθανοθεωρητικές μεθόδους της στατιστικής φυσικής. Συγκεκριμένα χρησιμοποιείται η κατανομή Bolzmann που στην φυσική περιγράφει απομονωμένα συστήματα σε θερμική ισορροπεία για την ανάλυση αυτών των προβλημάτων. Αρχικά εισάγουμε θεμελιώδεις έννοιες όπως η εντροπία και αποδεικνύουμε στοιχειώδη αποτελέσματα για αυτήν. Κατόπιν ορίζουμε την κατανομή Bolzmann και παρουσιάζουμε κάποιες σημαντικές ιδιότητές της. Ακολούθως εισάγουμε το μοντέλο Ising και δείχνουμε την άμεση συσχέτισή του με την θεωρητική πληροφορική. Τέλος χρησιμοποιούμε τα πιθανοθεωρήτικα εργαλεία που αναπτύξαμε στην ανάλυση πολλών τυχαίων στιγμιοτύπων δύσκολων υπολογιστικά προβλημάτων.

Η συσχέτιση ενός κλάδου της φυσικής με την πληροφορική δεν πρέπει να προκαλεί έκπληξη καθώς είναι συνυφασμένη από την αρχή της με την φυσική. Ωστόσο η συσχέτιση των δύο κλάδων αφορά κυρίως θέματα υλικού. Συνεπώς η επιρροή που ασκεί στο λογισμικό που αφορά την θεωρητική πληροφορική είναι ως προς το υπολογιστικό μοντέλο που αποτελεί το υλικό. Για παράδειγμα μπορούμε να γράψουμε διαφορετικούς αλγορίθμους σε έναν κβαντικό υπολογιστή από ότι σε έναν κλασσικό. Η ιδιαιτερότητα λοιπόν της παρούσας διπλωματικής είναι η δυνατότητα να μελετήσουμε ένα συνδυαστικό πρόβλημα όπως ο χρωματισμός ενός γράφου με τα θεωρητικά και υπολογιστικά εργαλεία της στατιστικής φυσικής.

Συμβολισμός

1. $\log x$ ο λογάριθμος με βάση e του x

2. $f(x) = O(g(x))$ εάν υπάρχουν $C \in (0,\infty), x_0 \in \mathbb{R}$ τέτοια ώστε για κάθε $x > x_0$ να ισχύει $|f(x)| < Cg(x)$

3. $f(x) = \Omega(g(x))$ εάν υπάρχουν $C \in (0,\infty), x_0 \in \mathbb{R}$ τέτοια ώστε για κάθε $x > x_0$ να ισχύει $|f(x)| > Cg(x)$

4. $f(x) = \Theta(g(x))$ εάν $f(x) = O(g(x))$ και $f(x) = \Omega(g(x))$

5. $S_p = \{x \in \mathbb{X} : p(x) > 0\}$ στήριγμα της κατανομής p

6. p_X είναι η κατανομή της τυχαίας μεταβλητής X

7. $S_X = \{x \in \mathbb{X} : p_X(x) > 0\}$ στήριγμα της κατανομής p_X της τυχαίας μεταβλητής X

8. $a_n \doteq b_n \iff \lim_n \frac{1}{n} \log \frac{a_n}{b_n} = 0$ δηλαδή οι ακολουθίες a_n, b_n έχουν την ίδια εκθετική τάξη

9. $\overline{A}$ η κλειστότητα του A.

10. $R_n = \{j \in \mathbb{N}^* : j \leq n\}$, όπου $n \in \mathbb{N}^*$.

11. C_n η κλίκα με κόμβους R_n.

12. $\mathbb{I}_A$ η δείκτρια του συνόλου A.

13. $sgn(x) = \mathbb{I}_{(0,\infty)}(x) - \mathbb{I}_{(-\infty,0)}(x)$ η συνάρτηση προσήμου.

Σημαντικά Αποτελέσματα

Πρόταση 0.1 (Βελτιωμένη προσέγγιση Striling). *Για* $n \geq 1$ *ισχύει* $n! = e^{r(n)}\sqrt{2\pi n}n^n e^{-n}$ *όπου* $\frac{1}{12n+1} < r(n) < \frac{1}{12n}$

Θεώρημα 0.2 (Εκτίμηση σαγματικού σημείου). *Έστω συναρτήσεις* $f, g \in C^0[a,b] \cap C^2(a,b)$. *Εάν υπάρχει μοναδικό* $x_0 \in (a,b)$ *με τις ιδιότητες* $g(x_0) = \max_{x\in[a,b]} g(x)$ *και* $g''(x_0) < 0$ *και επιπλέον για αυτό το* x_0 *ισχύει* $f(x_0) \neq 0$ *τότε* $\exists \lambda_0 > 0$ *τέτοιο ώστε για* $\lambda \geq \lambda_0$

$$\int_a^b f(x)e^{\lambda g(x)}dx = f(x_0)e^{\lambda g(x_0)}\sqrt{\frac{2\pi}{-\lambda g''(x_0)}}(1 + O(\frac{1}{\lambda}))$$

1 Θεμελιώδεις έννοιες

1.1 Θεωρία Πληροφοριών

Στο παρόν κεφάλαιο θα αναπτύξουμε σημαντικούς ορισμούς για τη μελέτη διακριτών τυχαίων μεταβλητών. Η πιο θεμελιώδης έννοια είναι αυτή της εντροπίας που αναπτύσσουμε παρακάτω.

1.1.1 Εντροπία

Ορισμός 1.1 (Εντροπία). *Η εντροπία μιας διακριτής τυχαίας μεταβλητής X με κατανομή $p(x)$ στον χώρο $\mathbb{X}$ όπου $S = \{x \in \mathbb{X} : p(x) > 0\}$ (στήριγμα της κατανομής p) είναι*

$$H_X = -\sum_{x \in S} p(x) \log_a p(x) \quad a > 1$$

Αρχικά παρατηρούμε ότι $-\log_a p(x) \geq 0 \quad \forall x \in S$. Το οποίο σημαίνει ότι το παραπάνω άθροισμα είναι καλώς ορισμένο. Δηλαδή ισούται με κάποιο μη αρνητικό αριθμό ή είναι άπειρο. Για να εξηγηθεί η διαίσθηση του ορισμού 1.1 θα αποδείξουμε και κατόπιν θα αναλύσουμε τις παρακάτω ιδιότητές του.

Πρόταση 1.2 (Ιδιότητες Εντροπίας). *Θεωρούμε X διακριτή τυχαία μεταβλητή στο $\mathbb{X}$.*

1. $H_X \geq 0$
2. $H_X = 0 \implies \exists! x_0 \in \mathbb{X} : p(x_0) = 1$ *δηλαδή η τυχαία μεταβλητή λαμβάνει με πιθανότητα 1 την x_0.*
3. $|\mathbb{X}| = M \implies H_X \leq \log_a M$ *και η ισότητα ισχύει αν και μόνο αν $X \sim U(\mathbb{X})$*

Απόδειξη. Η ανισότητα 1 έχει ήδη αποδειχθεί από τα προαναφερθέντα ($-\log_a p(x) \geq 0 \quad \forall x \in S$). Για να αποδείξω την ιδιότητα 2 αρχικά χωρίζω το S σε δύο υποσύνολά του A, B ώστε $A = \{x \in S : p(x) = 1\}$ και $B = \{x \in S : 0 < p(x) < 1\}$. Είναι φανερό από τη συνθήκη $\sum_{x \in S} p(x) = 1$ ότι ή το A είναι μονοσύνολο και το B είναι κενό ή το A είναι κενό και το B μη κενό. Επίσης $\forall x \in A - p(x) \log_a p(x) = -1 \log_a 1 = 0$ από το οποίο έπεται $-\sum_{x \in A} p(x) \log_a p(x) = 0$ και συνεπώς

$$-\sum_{x \in S} p(x) \log_a p(x) =$$

$$-\sum_{x \in A} p(x) \log_a p(x) - \sum_{x \in B} p(x) \log_a p(x) =$$

$$-\sum_{x \in B} p(x) \log_a p(x)$$

Άρα αν υποθέσουμε $H_X = 0$ τότε

$$-\sum_{x \in B} p(x) \log_a p(x) = 0$$

Όμως $\forall x \in B\ -p(x)\log_a p(x) > 0$. Το οποίο σημαίνει πως αν το A δεν είναι μονοσύνολο τότε το B είναι μη κενό και έπεται

$$-\sum_{x\in B} p(x)\log_a p(x) > 0$$

Άτοπο. Άρα το A είναι μονοσύνολο το οποίο είναι το ζητούμενο της 2.

Για να αποδείξω την 3 υποθέτω $|\mathbb{X}| = M$ και θεωρώ τη συνάρτηση $f : [0,\infty) \mapsto \mathbb{R}$ με τύπο $f(t) = -t\log_a t$ για $t \in (0,\infty)$ και $f(0) = 0$. Όπως είναι γνωστό από την ανάλυση $\lim_{t\to 0} t\log t = 0$ και $\log_a t = \frac{\log t}{\log a}$ από τα οποία έπεται ότι η f είναι συνεχής. Επίσης η f είναι δύο φορές παραγωγίσημη στο $(0,\infty)$ και έχει $f''(t) = -\frac{1}{t\log a} < 0$ άρα η f είναι κοίλη.

$$\frac{1}{M}\sum_{x\in\mathbb{X}} f(p(x)) = \frac{1}{M}\sum_{x\in\mathbb{X}\setminus S} f(p(x)) + \frac{1}{M}\sum_{x\in S} f(p(x)) =$$

$$\frac{1}{M}\sum_{x\in\mathbb{X}\setminus S} f(0) + \frac{1}{M}\sum_{x\in S} f(p(x)) = \frac{1}{M}\sum_{x\in S} f(p(x)) =$$

$$\frac{1}{M}\sum_{x\in S} -p(x)\log_a p(x) = \frac{1}{M}H_X$$

Από την ανισότητα Jensen έχουμε

$$\frac{1}{M}H_X \le f(\frac{1}{M}\sum_{x\in\mathbb{X}} p(x)) = \frac{1}{M}\log_a M \implies$$

$$H_X \le \log_a M$$

Η ισότητα ισχύει για την ανισότητα Jensen αν και μόνο αν $p(x) = c \quad \forall x \in \mathbb{X}$ και συνεπώς $p(x) = \frac{1}{M} \quad \forall x \in \mathbb{X}$ λόγω της συνθήκης $\sum_{x\in\mathbb{X}} p(x) = 1$.

□

Σε έναν πεπερασμένο χώρο $\mathbb{X}$ η εντροπία της X λαμβάνει ελάχιστο όταν η X λαμβάνει μια τιμή με πιθανότητα 1 και μέγιστο όταν η X κατανέμεται ομοιόμορφα στον $\mathbb{X}$, δηλαδή έχουμε την λιγότερη δυνατή πληροφορία για την τιμή της X. Είναι φανερό πλέον ότι η παραπάνω ορισμένη εντροπία είναι η αβεβαιότητα της X ή η έλλειψη πληροφορίας που έχουμε a priori για την τιμή της X. Για παράδειγμα αν ρίχναμε ένα αμερόληπτο κέρμα n φορές θα είχαμε μια τυχαία αλληλουχία από κορόνες και γράμματα X ομοιόμορφα κατανεμημένη σε έναν χώρο μεγέθους 2^n με εντροπία $n\log_a 2$. Έχει ιδιαίτερο ενδιαφέρον αν παρατηρήσουμε ότι για $a = 2$ η εντροπία είναι n όσα ακριβώς είναι τα bits που δεν γνωρίζουμε για την τιμή της εν λόγο τυχαίας μεταβλητής. Η επιλογή του a καθορίζει την μονάδα που εκφράζεται η εντροπία. Για $a = 2$ η εντροπία εκφράζεται σε bits και αυτή η βάση για την εντροπία χρησιμοποιείται σε προβλήματα ψηφιακών τηλεπικοινωνιών. Αν $a = e$ τότε η εντροπία εκφράζεται σε nats. Φυσικά από τη μαθηματική σκοπιά η επιλογή του a απλά μεταβάλει την εντροπία κατά μια πολλαπλασιαστική σταθερά και δεν παίζει ιδιαίτερο ρόλο. Συνεπώς από εδώ και πέρα θα χρησιμοποιούμε $a = e$.

1.1.2 Σχετική Εντροπία

Ορισμός 1.3 (Σχετική Εντροπία). *Για δύο κατανομές p, q στον διακριτό χώρο $\mathbb{X}$ με την ιδιότητα $(q(x) = 0 \implies p(x) = 0) \quad \forall x \in \mathbb{X}$ ή ισοδύναμα $S_p \subseteq S_q$ ορίζουμε σχετική εντροπία ή απόκλιση KL (Kullback-Leibler) της p ως προς την q*

$$D_{KL}(p\|q) = \sum_{x \in S_p} p(x) \log \frac{p(x)}{q(x)}$$

Αξίζει να σημειωθεί ότι ο παραπάνω ορισμός γενικεύεται για ζεύγος p, q μέτρων πιθανοτήτων σε μετρήσιμο χώρο $(\Omega, \mathbb{F})$. Ωστόσο στην παρούσα μελέτη ασχολούμαστε μόνο με διακριτές τυχαίες μεταβλητές.

Όπως και στην περίπτωση της εντροπίας θα δούμε πρώτα μερικές ιδιότητες του ορισμού 1.3 και έπειτα θα σχολιάσουμε την διαίσθηση πίσω από τον ορισμό.

Πρόταση 1.4 (Ανισότητα Gibbs).

$$D_{KL}(p\|q) \geq 0$$

και η ισότητα ισχύει αν και μόνο αν $p \equiv q$

Απόδειξη. Θεωρούμε p, q κατανομές στον διακριτό χώρο $\mathbb{X}$ με $S_p \subseteq S_q$. Έστω $f : [0, \infty) \mapsto \mathbb{R}$ με τύπο $f(x) = x \log x$ για $x \in (0, \infty)$ και $f(0) = 0$. Στην παράγραφο 1.1 δείξαμε ότι η $-f$ είναι κοίλη άρα η f είναι κυρτή.

$$\sum_{x \in S_q} f\Big(\frac{p(x)}{q(x)}\Big) q(x) = \sum_{x \in S_p} f\Big(\frac{p(x)}{q(x)}\Big) q(x) + \sum_{x \in S_q \setminus S_p} f\Big(\frac{p(x)}{q(x)}\Big) q(x) =$$

$$D_{KL}(p\|q) + \sum_{x \in S_q \setminus S_p} f\Big(\frac{0}{q(x)}\Big) q(x) = D_{KL}(p\|q)$$

Από την ανισότητα Jensen λαμβάνουμε

$$D_{KL}(p\|q) \geq f\Big(\sum_{x \in S_q} \frac{p(x)}{q(x)} q(x)\Big) =$$

$$f\Big(\sum_{x \in S_q} p(x)\Big) = f(1) = 0$$

Η ισότητα ισχύει αν και μόνο αν $p(x) = q(x) \quad \forall x \in S_q$ ή ισοδύναμα $p \equiv q$. $\square$

Η σχετική εντροπία αποτελεί μια έννοια ομοιότητας μεταξύ των κατανομών p, q. Ωστόσο δεν είναι μετρική διότι δεν ισχύει $D_{KL}(p\|q) = D_{KL}(q\|p)$. Πιο συγκεκριμένα μπορεί ενώ ορίζεται η $D_{KL}(p\|q)$ να μην ορίζεται η $D_{KL}(q\|p)$ λόγω των απαιτήσεων $S_p \subseteq S_q$ και $S_q \subseteq S_p$ αντίστοιχα. Από την ανισότητα Gibbs προκύπτει το παρακάτω πολύ ενδιαφέρον αποτέλεσμα.

Πόρισμα 1.4.1 (Υποαθροιστικότητα της εντροπίας)**.** *Έστω X, Y τυχαίες μεταβλητές στα αριθμήσιμα σύνολα $\mathbb{X}, \mathbb{Y}$ αντίστοιχα. τότε*

$$H_{X,Y} \leq H_X + H_Y$$

Η ισότητα ισχύει αν και μόνο αν οι μεταβλητές είναι ανεξάρτητες, δηλαδή $p_{X,Y}(x,y) = p_X(x)p_Y(y) \quad \forall (x,y) \in \mathbb{X} \times \mathbb{Y}$.

Απόδειξη. Στο αριθμήσιμο σύνολο $\mathbb{X} \times \mathbb{Y}$ έστω οι κατανομές $u(x,y) = p_{X,Y}(x,y)$ και $v(x,y) = p_X(x)p_Y(y)$ που ορίζονται από τις κατανομές των τυχαίων μεταβλητών X, Y αντίστοιχα. Για ένα ζευγάρι (x_0, y_0) τέτοιο ώστε $v(x_0, y_0) = 0$ έχουμε

$$v(x_0, y_0) = 0 \implies p_X(x_0) = 0 \vee p_Y(y_0) = 0$$

Αν υποθέσουμε $p_X(x_0) = 0$ τότε

$$0 = p_X(x_0) = \sum_{y \in \mathbb{Y}} p_{X,Y}(x_0, y) \implies p_{X,Y}(x_0, y) = 0 \quad \forall y \in \mathbb{X} \implies$$

$$p_{X,Y}(x_0, y_0) = 0$$

Στην περίπτωση $p_Y(y_0) = 0$ ομοίως έχουμε

$$0 = p_Y(y_0) = \sum_{x \in \mathbb{X}} p_{X,Y}(x, y_0) \implies p_{X,Y}(x_0, y_0) = 0$$

Άρα προκύπτει $v(x_0, y_0) = 0 \implies u(x_0, y_0) = p_{X,Y}(x_0, y_0) = 0$. Από το οποίο έπεται ότι η $D_{KL}(u \| v)$ είναι καλώς ορισμένη.

$$0 \leq D_{KL}(u \| v) = \sum_{(x,y) \in S_u} u(x,y) \log \frac{u(x,y)}{v(x,y)} =$$

$$\sum_{(x,y) \in S_u} u(x,y) \log u(x,y) - \sum_{(x,y) \in S_u} u(x,y) \log v(x,y) =$$

$$-H_{X,Y} - \sum_{(x,y) \in S_u} u(x,y) \log v(x,y)$$

Το δεύτερο άθροισμα όμως είναι

$$- \sum_{(x,y) \in S_u} u(x,y) \log v(x,y) = - \sum_{(x,y) \in S_u} p_{X,Y}(x,y) \log p_X(x) - \sum_{(x,y) \in S_u} p_{X,Y}(x,y) \log p_Y(y)$$

Επειδή $p_{X,Y}(x,y) = 0 \quad \forall (x,y) \in S_v \setminus S_u$

$$- \sum_{(x,y) \in S_u} u(x,y) \log v(x,y) = - \sum_{(x,y) \in S_v} p_{X,Y}(x,y) \log p_X(x) - \sum_{(x,y) \in S_v} p_{X,Y}(x,y) \log p_Y(y)$$

Τέλος παρατηρούμε ότι $v(x,y) > 0 \iff p_X(x) > 0 \wedge p_Y(y) > 0$ και άρα $S_v = S_X \times S_Y$. Επίσης $-p_{X,Y}(x,y) \log p_X(x) > 0$ και $-p_{X,Y}(x,y) \log p_Y(y) > 0$.

Συνεπώς από το Θεώρημα Fubini-Tonelli έπεται ότι

$$-\sum_{(x,y)\in S_u} u(x,y)\log v(x,y) = -\sum_{x\in S_X}(\sum_{y\in S_Y} p_{X,Y}(x,y))\log p_X(x) - \sum_{y\in S_Y}(\sum_{x\in S_X} p_{X,Y}(x,y))\log p_Y(y)$$

$$= -\sum_{x\in S_X} p_X(x)\log p_X(x) - \sum_{y\in S_Y} p_Y(y)\log p_Y(y) = H_X + H_Y$$

Άρα έχουμε $H_{X,Y} \leq H_X + H_Y$ με ισότητα αν και μόνο αν $u \equiv v \iff p_{X,Y}(x,y) = p_X(x)p_Y(y) \quad \forall (x,y) \in \mathbb{X} \times \mathbb{Y}$, δηλαδή οι X, Y είναι ανεξάρτητες. $\square$

Πραγματικά αυτό που μας λέει το παραπάνω πόρισμα είναι η αβεβαιότητα του ζεύγους X, Y τυχαίων μεταβλητών είναι το πολύ το άθροισμα των εντροπιών τους και την λαμβάνει αυτή την τιμή αν και μόνο αν οι μεταβλητές είναι ανεξάρτητες, δηλαδή η γνώση της τιμής μιας εξ αυτών δεν δίνει κάποια πληροφορία για την άλλη. Στην επόμενη παράγραφο θα δούμε έννοιες εντροπίας που αφορούν ζεύγη τυχαίων μεταβλητών και ασχολούνται με την πληροφορία που προκύπτει από τη συσχέτισή τους.

1.1.3 Δεσμευμένη και Αμοιβαία Εντροπία

Ορισμός 1.5 (Δεσμευμένη Εντροπία). *Για το ζεύγος (X, Y) διακριτών τυχαίων μεταβλητών κατανεμημένο στο $\mathbb{X} \times \mathbb{Y}$ ορίζουμε ως δεσμευμένη εντροπία του Y ως προς το X*

$$H_{Y|X} = H_{X,Y} - H_X$$

Η δεσμευμένη εντροπία εκφράζει την αβεβαιότητα της Y γνωρίζοντας την X.

Ορισμός 1.6 (Αμοιβαία Εντροπία). *Ορίζουμε ως αμοιβαία εντροπία του ζεύγους (X, Y) διακριτών τυχαίων μεταβλητών κατανεμημένο στο $\mathbb{X} \times \mathbb{Y}$*

$$I_{X,Y} = H_X - H_{X|Y} = H_Y - H_{Y|X} = H_X + H_Y - H_{X,Y}$$

Η παραπάνω ποσότητα είναι καλώς ορισμένη διότι $H_X - H_{X|Y} = H_X - (H_{X,Y} - H_Y) = H_X + H_Y - H_{X,Y} = H_Y - (H_{X,Y} - H_X) = H_Y - H_{Y|X}$. Από αυτό το γεγονός επίσης έπεται $I_{X,Y} = I_{Y,X}$

Η αμοιβαία εντροπία εκφράζει τη μείωση της αβεβαιότητας της X γνωρίζοντας την Y ή τη μείωση της αβεβαιότητας της Y γνωρίζοντας την X.

Από το πόρισμα 1.4.1 προκύπτει ότι $I_{X,Y} = D_{KL}(p_{X,Y} \| p_X p_Y) \geq 0$ με ισότητα αν και μόνο αν X, Y ανεξάρτητες.

Τέλος παρατηρούμε ότι $H_{X,Y} - H_{X|Y} - H_{Y|X} = H_{X,Y} - (H_{X,Y} - H_Y) - H_{Y|X} = H_Y - H_{Y|X} = I_{X,Y}$

Για λόγους απλότητας από εδώ και πέρα θα γράφουμε $H_X = -\sum_{x \in \mathbb{X}} p(x) \log p(x)$ θεωρώντας $0 \log 0 = 0$. Επίσης θα λέμε H_p την εντροπία της κατανομής p. Τέλος για συναρτήσεις μάζας πιθανότητας p, q που δεν ισχύει $S_p \subseteq S_q$ ορίζουμε $D_{KL}(p\|q) = \infty$.

1.2 Μεγάλες Αποκλίσεις

Η θεωρία μεγάλων αποκλίσεων ασχολείται με την συμπεριφορά των ακραίων φαινομένων όταν $n \to \infty$ σε μια ακολουθία τυχαίων μεταβλητών υπό κατάλληλες προϋποθέσεις. Σε αντίθεση δηλαδή με γνωστά θεωρήματα όπως ο νόμος των μεγάλων αριθμών, το κεντρικό οριακό θεώρημα ή το εργοδικό θεώρημα μαρκοβιανών αλυσίδων που μελετούν τα πιθανότερα φαινόμενα.

Είναι εύλογο λοιπόν να δώσουμε τους παρακάτω ορισμούς

Ορισμός 1.7 (Αρχή μεγάλων αποκλίσεων)**.** *Λέμε ότι η ακολουθία διακριτών κατανομών $(p_n)_{n\in\mathbb{N}}$ στον χώρο $\mathbb{X}$ ακολουθεί μια αρχή μεγάλων αποκλίσεων με συνάρτηση ρυθμού $I : \mathbb{X} \mapsto [0,\infty)$ εάν*

$$p_n(x) \doteq \exp(-nI(x))$$

Ορισμός 1.8. *Για μια πεπερασμένη ακολουθία $x = (x_k)_{k=1}^n \in \mathbb{X}^n$ λέμε τύπο της ακολουθίας*

$$q_x(y) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}(x_k = y)$$

1.2.1 Θεώρημα Sanov

Αρχικά θα εξετασθεί η περίπτωση μιας ακολουθίας ανεξάρτητων ισόνομων τυχαίων μεταβλητών σε ένα πεπερασμένο σύνολο $\mathbb{X}$. Γενικά θεωρώ σε αυτό το κεφάλαιο πως για μια τυχαία μεταβλητή X στο $\mathbb{X}$ ισχύει $S_X = \mathbb{X}$. Σε περίπτωση που έχω $S_X \subset \mathbb{X}$ τότε μπορούμε να αφαιρέσουμε τα στοιχεία του $\mathbb{X}$ που έχουν πιθανότητα 0 διότι τα φαινόμενα που εμπεριέχουν επιλογή με πιθανότητα 0 έχουν πιθανότητα 0 άρα δεν χρειάζεται κάποιο θεώρημα για τη μελέτη τους.

Θεώρημα 1.9 (Θεώρημα Sanov)**.** *Έστω $(X_k)_{k\in\mathbb{N}}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών στο σύνολο $\mathbb{X}$ με $|\mathbb{X}| = m$ και κατανομή $p(x)$. Εάν $\mathbb{M}(\mathbb{X})$ το σύνολο των κατανομών πάνω στο X και $K = \overline{U}$ όπου U ανοικτό και μη κενό υποσύνολο του $\mathbb{M}(\mathbb{X})$ (ως προς οποιαδήποτε p-νόρμα αφού ο χώρος $\mathbb{M}(\mathbb{X})$ είναι υποσύνολο του $\mathbb{R}^m$ και οι p-νόρμες είναι ισοδύναμες. $\mathbb{M}(\mathbb{X}) = \{x = (x_1, x_2, ..., x_m) \in \mathbb{R}^m : \sum_{k=1}^m x_k = 1 \quad 0 \le x_k \le 1\}$). Τέλος θέτουμε q_n τον τύπο των τυχαίων μεταβλητών $X_1, X_2, ..., X_n$. Τότε ισχύει*

$$P(q_n \in K) \doteq \exp(-ND_{KL}(q*||p))$$

όπου $q* = argmin_{q\in K} D_{KL}(q||p)$

Πριν προχωρήσουμε στην απόδειξη του 1.9 θα χρειαστούμε κάποια λήμματα.

Λήμμα 1.10 (Εκτίμηση n επιλογών από m κατηγορίες)**.**

$$\frac{n!}{\prod_{j=1}^m k_j!} = \exp(O(\log n) - n\sum_{j:k_j>0} \frac{k_j}{n}\log\frac{k_j}{n}) = \exp(O(\log n) + nH_p)$$

όπου p η κατανομή $(\frac{k_j}{n})_{j=1}^n$

Απόδειξη. Εφόσον $0! = 1$

$$\frac{n!}{\prod_{j=1}^{m} k_j!} = \frac{n!}{\prod_{j:k_j>0} k_j!}$$

Από την 0.1 λαμβάνουμε

$$\frac{n!}{\prod_{j:k_j>0} k_j!} = \frac{e^{r(n)}(2\pi n)^{\frac{1}{2}} n^n e^{-n}}{e^{\sum_{j:k_j>0} r(k_j)} \prod_{j:k_j>0}(2\pi k_j)^{\frac{1}{2}} k_j^{k_j} e^{-k_j}}$$

Επειδή $\sum_{j:k_j>0} k_j = \sum_{j=1}^{m} k_j = n$ απλοποιούνται οι όροι e^{-n}, e^{-k_j}

$$\frac{n!}{\prod_{j:k_j>0} k_j!} = \frac{e^{r(n)}(2\pi n)^{\frac{1}{2}} n^n}{e^{\sum_{j:k_j>0} r(k_j)} \prod_{j:k_j>0}(2\pi k_j)^{\frac{1}{2}} k_j^{k_j}} =$$

$$\exp(r(n) - \sum_{j:k_j>0} r(k_j) + \frac{1}{2}\log(2\pi n) - \frac{1}{2}\sum_{j:k_j>0}\log(2\pi k_j) + n\log n - \sum_{j:k_j>0} k_j \log k_j) =$$

Θέτω $u(n) = r(n) - \sum_{j:k_j>0} r(k_j) + \frac{1}{2}\log(2\pi n) - \frac{1}{2}\sum_{j:k_j>0}\log(2\pi k_j)$

$$\exp(u(n) + n\log n - \sum_{j:k_j>0} k_j \log k_j) = \exp(u(n) + \sum_{j:k_j>0} k_j \log n - \sum_{j:k_j>0} k_j \log k_j) =$$

$$\exp(u(n) - n\sum_{j:k_j>0} \frac{k_j}{n}\log\frac{k_j}{n})$$

Οπότε αρκεί να δείξω $u(n) = O(\log n)$

$$|u(n)| \leq r(n) + \sum_{j:k_j>0} r(k_j) + \frac{1}{2}\log(2\pi n) + \frac{1}{2}\sum_{j:k_j>0}\log(2\pi k_j)$$

Θέτω $m' = |\{j : k_j > 0\}|$

Από τον ορισμό του $r(n)$ και το γεγονός ότι $k_j \geq 1$ έπεται ότι

$$r(n) + \sum_{j:k_j>0} r(k_j) \leq \frac{1}{12n} + \sum_{j:k_j>0} \frac{1}{k_j} \leq \frac{1}{12n} + m'\frac{1}{12} = O(1)$$

Τέλος από την ανισότητα Jensen λαμβάνουμε

$$\frac{1}{m'}\sum_{j:k_j>0}\log(2\pi k_j) \leq \log(\frac{2\pi n}{m'})$$

Από το οποίο προκύπτει

$$\frac{1}{2}\sum_{j:k_j>0}\log(2\pi k_j) = O(\log n)$$

Και συνεπώς λαμβάνουμε

$$|u(n)| \leq O(1) + \frac{1}{2}\log(2\pi n) + O(\log n) = O(\log n) \implies$$

$$u(n) = O(\log n)$$

$\square$

Βλέπουμε λοιπόν ότι η εκθετική τάξη των n επιλογών από m διακριτές κατηγορίες κυριαρχείται από την εντροπία της κατανομής των επιλογών που έγιναν. Προφανώς αυτό αποτελεί ένα κομβικό λήμμα διότι συνδέει άμεσα την συνδυαστική με τη θεωρία πληροφοριών. Εάν από τις m κατηγορίες επιλέγαμε με πιθανότητα $\frac{1}{m}$ (ομοιόμορφα) τότε η πιθανότητα ενός συγκεκριμένου τύπου $q = (\frac{k_j}{n})_{j=1}^n$ είναι $P(q_n = q) = \frac{n!}{\prod_{j=1}^m k_j!}(\frac{1}{m})^n = \exp(O(\log n) - n\log m + nH_q)$. Δεδομένου ότι $H_q \leq \log m$ με την ισότητα να ισχύει αν και μόνο αν q είναι η ομοιόμορφη έπεται ότι σε όλες τις περιπτώσεις που η q δεν είναι η ομοιόμορφη η πιθανότητα εμφάνισής της φθίνει εκθετικά. Αυτή την παρατήρηση θα γενικεύσουμε με το λήμμα που ακολουθεί για μια οποιαδήποτε κατανομή $p \in \mathbb{M}(\mathbb{X})$.

Λήμμα 1.11.

$$P(q_n = q) = \exp(O(\log n) - nD_{KL}(q\|p))$$

Απόδειξη. Από το λήμμα 1.10

$$P(q_n = q) = \frac{n!}{\prod_{j=1}^m k_j!}\prod_{j=1}^m p_j^{k_j} =$$

$$\exp(O(\log n) - n\sum_{j=1}^m \frac{k_j}{n}\log\frac{k_j}{n} + \sum_{j=1}^k k_j \log p_j)) =$$

$$\exp(O(\log n) - n\sum_{j=1}^m q_j \log\frac{q_j}{p_j}) = \exp(O(\log n) - nD_{KL}(q\|p))$$

$\square$

Στη γενική περίπτωση n τυχαίων επιλογών m αντικειμένων από την κατανομή p η εκθετική τάξη της πιθανότητας εμφάνισης μιας κατανομής q κυριαρχείται από τη σχετική εντροπία της q ως προς την p.

Θεωρούμε Q_n το σύνολο των κατανομών που μπορούμε να λάβουμε ως q_n δηλαδή $Q_n = \{q \in \mathbb{M}(\mathbb{X}) : q = (\frac{k_j}{n}) \quad k_j \in \mathbb{N}\}$. Επίσης είναι σαφές ότι τα Q_n είναι μεγιστικά $\frac{1}{n}$- διαχωρισμένα σύνολα του $\mathbb{M}(\mathbb{X})$. Με αυτή την παρατήρηση μπορούμε να προχωρήσουμε στην απόδειξη του Θεωρήματος Sanov.

Απόδειξη Θεωρήματος Sanov. Εφόσον $q_n \in Q_n$ εκ κατασκευής και Q_n πεπερασμένο

$$P(q_n \in K) = P(q_n \in K \cap Q_n) = \sum_{q \in K \cap Q_n} P(q_n = q)$$

Θέτουμε $q^* = argmin_{q \in K} D_{KL}(q \| p)$ και $q_n^* = argmin_{q \in K \cap Q_n} D_{KL}(q \| p)$

$$P(q_n \in K) \geq P(q_n = q_n^*)$$

Χρησιμοποιώντας το λήμμα 1.11 λαμβάνουμε

$$\frac{1}{n} \log P(q_n \in K) \geq O(\frac{\log n}{n}) - D_{KL}(q_n^* \| p) \tag{1}$$

$$\frac{1}{n} \log P(q_n \in K) = \frac{1}{n} \log(\sum_{q \in K \cap Q_n} \exp(O(\log n) - nD_{KL}(q \| p))) \leq$$

$$\frac{1}{n} \log(|K \cap Q_n| \exp(O(\log n) - nD_{KL}(q_n^* \| p))) \leq$$

$$\frac{1}{n} \log(|Q_n|) + O(\frac{\log n}{n}) - D_{KL}(q_n^* \| p)$$

Το πλήθος των στοιχείων του Q_n είναι ίσο με το πλήθος των τρόπων που γράφεται $n = \sum_{j=1}^{m} x_j$ με $x_j \in \mathbb{N}$. Άρα $|Q_n| = \binom{n+m-1}{m-1} = O(n^{m-1}) \implies \frac{1}{n} \log |Q_n| = O(\frac{\log n}{n})$. Έπεται ότι

$$\frac{1}{n} \log P(q_n \in K) \leq O(\frac{\log n}{n}) - D_{KL}(q_n^* \| p) \tag{2}$$

Έστω $\epsilon > 0$ και $x \in K$. Επειδή $K = \overline{U}$ υπάρχει $y \in U$ τέτοιο ώστε $\|y - x\| < \frac{\epsilon}{2}$. Όμως από το γεγονός ότι το U είναι ανοικτό και μη κενό σύνολο υπάρχει $\delta > 0$ τέτοιο ώστε $B(y, \delta) \subseteq U \subseteq K$. Τέλος επειδή τα Q_n είναι μεγιστκά $\frac{1}{n}$-διαχωρισμένα σύνολα για κάθε $n \geq n_0 = \lfloor \frac{1}{\min(\delta, \frac{\epsilon}{2})} \rfloor + 1$ ισχύει

$$\exists z \in Q_n : \|z - y\| \leq \frac{1}{n} \leq \frac{1}{n_0} < \frac{1}{\frac{1}{\min(\delta, \frac{\epsilon}{2})}} = \min(\delta, \frac{\epsilon}{2})$$

Άρα $z \in Q_n \cap U \subseteq Q_n \cap K$ και $\|z - x\| \leq \|z - y\| + \|y - x\| < \min(\delta, \frac{\epsilon}{2}) + \frac{\epsilon}{2} \leq \epsilon$

Συνεπώς για κάθε $\epsilon > 0$ και για κάθε $x \in K$ υπάρχει n_0 που εξαρτάται από το x και το ϵ τέτοιο ώστε για κάθε $n \geq n_0$ να ισχύει $\exists z_n \in Q_n \cap K : \|z_n - x\| < \epsilon$.

Η συνάρτηση $f : \mathbb{M}(\mathbb{X}) \mapsto \mathbb{R}$ με τύπο $f(q) = D_{KL}(q \| p) = \sum_{x \in \mathbb{X}} q(x) \log \frac{q(x)}{p(x)}$ είναι συνεχής στο $\mathbb{M}(\mathbb{X})$. Με βάση την προηγούμενη πρόταση υπάρχει $n_0 \in N$ και $z_n \in Q_n \cap K$ για κάθε $n \geq n_0$ που ισχύει $\|z_n - q^*\| < \delta$ ώστε $|f(z_n) - f(q^*)| < \epsilon$. Άρα $f(z_n) \leq f(q^*) + \epsilon$ για κάθε $n \geq n_0$ όπου το n_0 εξαρτάται από το ϵ (το q^* δεν μεταβάλλεται). Όμως ισχύουν $f(q^*) \leq f(q_n^*) \leq f(z_n)$ εξ ορισμού των q^*, q_n^*. Συνεπώς έχουμε

$$f(q^*) \leq f(q_n^*) \leq f(q^*) + \epsilon \quad \forall n \geq n_0 = n_0(\epsilon) \implies \lim_{n} f(q_n^*) = f(q^*)$$

Από το οποίο και τις (1), (2) έπεται το ζητούμενο. $\square$

Πόρισμα 1.11.1. *Έστω $(X_k)_{k\in\mathbb{N}}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών στο σύνολο $\mathbb{X}$ με $|\mathbb{X}| = m$ ακολουθούν την κατανομή $p(x)$, $f : \mathbb{X} \mapsto \mathbb{R}$, $a, b \in \mathbb{R} : \min_{x\in\mathbb{X}} f(x) \leq a < b \leq \max_{x\in\mathbb{X}} f(x)$*

$$P(\frac{1}{n}\sum_{j=1}^{n} f(X_j) \in [a,b]) \doteq \exp(-nD_{KL}(q^*\|p))$$

όπου $q^* = argmin_{q\in K} D_{KL}(q\|p)$ και $K = \{q \in \mathbb{M}(\mathbb{X}) : \sum_{x\in\mathbb{X}} f(x)q(x) \in [a,b]\}$

Απόδειξη. Θέτω $U = \{q \in \mathbb{M}(\mathbb{X}) : \sum_{x\in\mathbb{X}} f(x)q(x) \in (a,b)\}$. Θέτουμε $m = argmin_{x\in\mathbb{X}} f(x)$ και $M = argmax_{x\in\mathbb{X}} f(x)$. Για $c \in (a,b)$ ισχύει $f(m) \leq a < c < b \leq f(M)$. Άρα μπορούμε να βρούμε κατάλληλο $\lambda \in [0,1]$ έτσι ώστε $\lambda f(m) + (1-\lambda)f(M) = c$ και συνεπώς η κατανομή που λαμβάνει με πιθανότητα λ την τιμή m και $1-\lambda$ την τιμή M ανήκει στο U. Επειδή η συνάρτηση $g : \mathbb{M}(\mathbb{X}) \mapsto \mathbb{R}$ με τύπο $g(q) = \sum_{x\in\mathbb{X}} f(x)q(x)$ είναι περιορισμός γραμμικής απεικόνισης στο $\mathbb{M}(\mathbb{X})$ έχουμε ότι g συνεχής. Συνεπώς $U = g^{-1}((a,b))$ ανοικτό και $K = g^{-1}([a,b]) = g^{-1}(\overline{(a,b)}) = \overline{U}$. Εφαρμόζοντας το θεώρημα Sanov για το K που είναι η κλειστότητα μη κενού ανοικτού συνόλου λαμβάνουμε

$$P(q_n \in K) \doteq \exp(-nD(q^*\|p))$$

όπου $q* = argmin_{q\in K} D_{KL}(q\|p)$.

Όμως $q_n \in K \iff \sum_{x\in\mathbb{X}} f(x)q_n(x) \in [a,b]$ και $\sum_{x\in\mathbb{X}} f(x)q_n(x) = \frac{1}{n}\sum_{j=1}^{n} f(X_j)$ Άρα

$$P(\frac{1}{n}\sum_{j=1}^{n} f(X_j) \in [a,b]) \doteq \exp(-nD(q^*\|p))$$

□

2 Στατιστική Φυσική

Η στατιστική φυσική είναι ένα τυχαιοκρατικό μοντέλο περιγραφής φυσικών συστημάτων που αποτελούνται από πολλά στοιχειώδη μέλη όπως άτομα ή μόρια. Η κατανομή Bolzmann περιγράφει ένα απομονωμένο σύστημα σε θερμική ισορροπία που η κάθε κατάστασή έχει πιθανότητα εμφάνισης που φθίνει εκθετικά όσο μεγαλύτερη είναι η ενέργειά της και η αντίστροφη θερμοκρασία του συστήματος.

2.1 Κατανομή Bolzmann

Ορισμός 2.1 (Κατανομή Bolzmann)**.** *Έστω ένα πεπερασμένο σύνολο $\mathbb{X}$ και μια συνάρτηση $E : \mathbb{X} \mapsto \mathbb{R}$. Εάν*

$$Z(\beta) = \sum_{x \in \mathbb{X}} \exp(-\beta E(x)) \quad \beta \in (0, \infty)$$

τότε ορίζουμε ως κατανομή Bolzmann πάνω στο $\mathbb{X}$ την συνάρτηση μάζας πιθανότητας

$$p_\beta(x) = \frac{e^{-\beta E(x)}}{Z(\beta)} \quad x \in \mathbb{X} \quad \beta \in (0, \infty)$$

Λέμε το σύνολο $\mathbb{X}$ χώρο καταστάσεων, την συνάρτηση E ενέργεια, την παράμετρο β αντίστροφη θερμοκρασία και την Z συνάρτηση επιμερισμού. Στην φυσική η παράμετρος β είναι $\beta = \frac{1}{k_B T}$ όπου k_B η σταθερά Bolzmann. Επίσης θα μπορούσαμε να ορίσουμε την κατανομή Bolzmann σε έναν οποιοδήποτε χώρο μέτρου (Ω, F, μ) αλλά δεν θα ασχολειθούμε με τέτοια περίπτωση.

Θα συμβολίζουμε $\mathbb{B}(\mathbb{X}, E, \beta)$ την κατανομή Bolzmann στον χώρο καταστάσεων $\mathbb{X}$ με συνάρτηση ενέργειας E και αντίστροφη θερμοκρασία β.

Για κατανομες Bolzmann της μορφής $\mathbb{B}(\mathbb{X}^n, E, \beta)$ είναι σαφές ότι οι n μεταβλητές που αποτελούν μια κατάσταση αναπτύσσουν πολύπλοκες σχέσεις εξάρτησης μεταξύ τους ανάλογα με την παράμετρο της αντίστροφης θερμοκρασίας και την ενέργειά τους.

Θα αποδείξουμε μερικές ιδιότητες της κατανομής Bolzmann.

Πρόταση 2.2. *Οι συναρτήσεις ενέργειας $E, \overline{E} : \mathbb{X} \mapsto R$ όπου $\exists c \in \mathbb{R} : \overline{E}(x) = E(x) + c \quad \forall x \in \mathbb{X}$ δίνουν ακριβώς τις ίδιες κατανομές Bolzmann p_b και $\overline{p}_b$ αντίστοιχα.*

Απόδειξη.

$$\overline{Z}(\beta) = \sum_{x \in \mathbb{X}} \exp(-\beta \overline{E}(x)) = \sum_{x \in \mathbb{X}} \exp(-\beta E(x)) \exp(-\beta c) = \exp(-\beta c) Z(\beta)$$

Άρα έχουμε ότι

$$\overline{p}_\beta(x) = \frac{\exp(-\beta \overline{E}(x))}{\overline{Z}(\beta)} = \frac{\exp(-\beta E(x)) \exp(-\beta c)}{Z(\beta) \exp(-\beta c)} = p_\beta(x)$$

□

Πρόταση 2.3 (Οριακές θερμοκρασίες). *Έστω $\mathbb{X}$ το σύνολο καταστάσεων, E η συνάρτηση ενέργειας και $\mathbb{X}_0 = \{x \in \mathbb{X} : E(x) = \min_{y\in\mathbb{X}} E(y)\}$ τότε ισχύουν*

1. $\lim_{\beta\to\infty} p_\beta(x) = \frac{\mathbb{I}(x\in\mathbb{X}_0)}{|\mathbb{X}_0|}$, *δηλαδή στο όριο χαμηλής θερμοκρασίας η κατανομή Bolzmann συγκλίνει στη $U(\mathbb{X}_0)$*
2. $\lim_{\beta\to 0} p_\beta(x) = \frac{1}{\mathbb{X}}$, *δηλαδή στο όριο υψηλής θερμοκρασίας η κατανομή Bolzmann συγκλίνει στη $U(\mathbb{X})$*

Απόδειξη. Έστω $E_0 = \min_{x\in\mathbb{X}} E(x)$

$$p_\beta(x) = \frac{\exp(-\beta E(x))}{Z(\beta)} = \frac{\exp(-\beta(E(x) - E_0))}{\sum_{y\in\mathbb{X}} \exp(-\beta(E(y) - E_0))} =$$

$$\frac{\exp(-\beta(E(x) - E_0))}{\sum_{y\in\mathbb{X}\setminus\mathbb{X}_0} \exp(-\beta(E(y) - E_0)) + \sum_{y\in\mathbb{X}_0} \exp(-\beta(E(y) - E_0))} =$$

$$\frac{\exp(-\beta(E(x) - E_0))}{\sum_{y\in\mathbb{X}\setminus\mathbb{X}_0} \exp(-\beta(E(y) - E_0)) + \sum_{y\in\mathbb{X}_0} \exp(0)} =$$

$$\frac{\exp(-\beta(E(x) - E_0))}{\sum_{y\in\mathbb{X}\setminus\mathbb{X}_0} \exp(-\beta(E(y) - E_0)) + |\mathbb{X}_0|}$$

Εφόσον $E(x) > E_0 \quad \forall x \in \mathbb{X} \setminus \mathbb{X}_0$ και άρα $\lim_{\beta\to\infty}[\exp(-\beta(E(x) - E_0))] = 0 \quad \forall x \in \mathbb{X} \setminus \mathbb{X}_0$ έχουμε ότι

$$\lim_{\beta\to\infty} p_\beta(x) = \frac{I(x \in \mathbb{X}_0)}{|\mathbb{X}_0|}$$

Από την άλλη

$$\lim_{\beta\to 0}[\exp(-\beta E(x))] = 1 \quad \forall x \in \mathbb{X} \implies$$

$$\lim_{\beta\to 0} p_\beta(x) = \frac{1}{|\mathbb{X}|}$$

$\square$

2.2 Θερμοδυναμικές Συναρτήσεις

Για να μελετήσουμε ένα σύστημα στατιστικής μηχανικής πολλές φορές θα χρειαστούμε συναρτήσεις που εμπεριέχουν πληροφορίες για το σύστημα και είναι ευκολότερος ο χειρισμός τους. Οι σημαντικότερες συναρτήσεις για αυτό το σκοπό δίνονται παρακάτω.

Ορισμός 2.4 (Θερμοδυναμικές Συναρτήσεις)**.** *Για την κατανομή* $\mathbb{B}(\mathbb{X}, E, \beta)$ *ορίζουμε τις συναρτήσεις*

1. *(Ελεύθερη Ενέργεια)* $F(\beta) = -\frac{1}{\beta}\log(Z(\beta))$
2. *(Ελεύθερη Εντροπία)* $\Phi(\beta) = \log(Z(\beta))$
3. *(Εσωτερική Ενέργεια)* $U(\beta) = -\frac{\partial(\Phi(\beta))}{\partial\beta}$
4. *(Κανονική Εντροπία)* $S(\beta) = \beta^2\frac{\partial F}{\partial\beta}$

Για λόγους συντομίας θα γράφουμε Z, F, Φ, U, S αντί $Z(\beta), F(\beta)$ κλπ.

Η μαθηματική όσο και η φυσική σημασία των δυο τελευταίων συναρτήσεων γίνεται αντιληπτή με την επόμενη πρόταση.

Πρόταση 2.5. *Εάν* $X \sim \mathbb{B}(\mathbb{X}, E, \beta)$ *τότε*

1. $U = \mathbb{E}(E(X))$
2. $S = \Phi + \beta U = H_X$
3. $\frac{\partial^2}{\partial\beta^2}\Phi = Var(E(X))$

Απόδειξη.

$$U = -\frac{\partial\Phi}{\partial\beta} = -\frac{\partial\log Z}{\partial\beta} = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} =$$

$$-\frac{1}{Z}\sum_{x\in\mathbb{X}}\frac{\partial(e^{-\beta E(x)})}{\partial\beta} = \frac{1}{Z}\sum_{x\in\mathbb{X}}E(x)e^{-\beta E(x)} = \mathbb{E}(E(X))$$

$$H_X = -\sum_{x\in\mathbb{X}}p_\beta(x)\log p_\beta(x) = -\sum_{x\in\mathbb{X}}p_\beta(x)(-\beta E(x) - \log Z) =$$

$$\log Z + \beta\sum_{x\in\mathbb{X}}E(x)p_\beta(x) = \Phi + \beta U = \beta^2(\frac{1}{\beta^2}\Phi - \frac{1}{\beta}\frac{\partial\Phi}{\partial\beta}) =$$

$$\beta^2\frac{\partial(-\frac{1}{\beta}\Phi)}{\partial\beta} = \beta^2\frac{\partial F}{\partial\beta} = S$$

$$\frac{\partial^2}{\partial\beta^2}\Phi = -\frac{\partial}{\partial\beta}U = -\sum_{x\in\mathbb{X}}E(x)\frac{\partial}{\partial\beta}(p_\beta(x)) =$$

$$-\sum_{x\in\mathbb{X}} E(x)(-E(x)\frac{e^{-\beta E(x)}}{Z} - \frac{1}{Z}\frac{\partial Z}{\partial \beta}\frac{e^{-\beta E(x)}}{Z}) =$$

$$\sum_{x\in\mathbb{X}} E(x)^2 p_\beta(x) - (-\frac{1}{Z}\frac{\partial Z}{\partial \beta})\mathbb{E}(E(X)) =$$

$$\mathbb{E}(E(X)^2) - (\mathbb{E}(E(x)))^2 = Var(E(X))$$

□

Θα μελετήσουμε τη συμπεριφορά αυτών των συναρτήσεων στα όρια χαμηλής και υψηλής θερμοκρασίας. Αφενός για να αναπτύξουμε περισσότερη διαίσθηση για την κατανομή Bolzmann. Αφετέρου, για να αποκτήσουμε σημαντικά εργαλεία για την μελέτη αυτών των ορίων στα συστήματα που θα συναντήσουμε.

Πρόταση 2.6. *Για την κατανομή $\mathbb{B}(\mathbb{X}, E, \beta)$ ισχύουν τα ακόλουθα*

1. $\lim_{\beta\to 0} \Phi = \log|\mathbb{X}|$
2. $\lim_{\beta\to 0} U = \mathbb{E}_{X\sim U(\mathbb{X})}(E(X))$
3. $\lim_{\beta\to 0} S = \log|\mathbb{X}|$
4. $\lim_{\beta\to \infty} F = E_0$
5. $\lim_{\beta\to \infty} U = E_0$
6. $\lim_{\beta\to \infty} S = \log|\mathbb{X}_0|$

Απόδειξη. Από τις προτάσεις 2.3, 2.5

$$\lim_{\beta\to 0} U = \lim_{\beta\to 0}\sum_{x\in\mathbb{X}} E(x)p_\beta(x) = \sum_{x\in\mathbb{X}} E(x)\frac{1}{|\mathbb{X}|} = \mathbb{E}_{X\sim U(\mathbb{X})}(E(X))$$

$$\lim_{\beta\to 0} S = \lim_{\beta\to 0}[-\sum_{x\in\mathbb{X}} p_\beta(x)\log p_\beta(x)] = -\sum_{x\in\mathbb{X}}\frac{1}{|\mathbb{X}|}\log\frac{1}{|\mathbb{X}|} = \log|\mathbb{X}|$$

Ομοίως για την περίπτωση $\beta \to \infty$

$$\lim_{\beta\to \infty} U = \lim_{\beta\to \infty}\sum_{x\in\mathbb{X}} E(x)p_\beta(x) = \sum_{x\in\mathbb{X}} E(x)\frac{\mathbb{I}(x\in\mathbb{X}_0)}{|\mathbb{X}_0|} = E_0$$

$$\lim_{\beta\to \infty} S = \lim_{\beta\to \infty}[-\sum_{x\in\mathbb{X}} p_\beta(x)\log p_\beta(x)] =$$

$$-\sum_{x\in\mathbb{X}}\frac{\mathbb{I}(x\in\mathbb{X}_0)}{|\mathbb{X}_0|}\log\frac{\mathbb{I}(x\in\mathbb{X}_0)}{|\mathbb{X}_0|} = \log|\mathbb{X}_0|$$

Όμως από την πρόταση 2.5 έχουμε

$$S = \Phi + \beta U \implies F = U - \frac{1}{\beta}S$$

Έπεται ότι

$$\lim_{\beta \to 0} \Phi = \lim_{\beta \to 0} (S - \beta U) = \log |\mathbb{X}|$$

$$\lim_{\beta \to \infty} F = \lim_{\beta \to \infty} (U - \frac{1}{\beta} S) = E_0$$

$\square$

2.3 Σχέση Διακύμανσης-Διάλυσης

Σε αυτή την ενότητα θα εξετάσουμε την περίπτωση η ενέργεια να έχει παραμετρική εξάρτηση. Σε φυσικά συστήματα μπορεί να είναι κάποιο βαρυτικό ή ηλεκτρομαγνητικό πεδίο. Σε συνδυαστικά προβλήματα που θα ασχοληθούμε αυτό είναι εξίσου σύνηθες. Σκοπός μας είναι να μετρήσουμε τις μεταβολές στο σύστημα ανάλογα με τη μεταβολή της παραμέτρου.

Θεώρημα 2.7 (Σχέση Διακύμανσης-Διάλυσης)**.** *Έστω $E : \mathbb{X} \times I \mapsto \mathbb{R}$ όπου I διάστημα του $\mathbb{R}$ και X_λ ακολουθεί την κατανομή Bolzmann με συνάρτηση ενέργειας $E(*, \lambda)$ για $\lambda \in I$. Εάν $f : \mathbb{X} \times I \mapsto \mathbb{R}$ και οι E, f είναι παραγωγίσιμες στο I για κάθε $x \in \mathbb{X}$ τότε*

$$\frac{\partial \mathbb{E}(f(X_\lambda, \lambda))}{\partial \lambda} = \mathbb{E}(\frac{\partial f(X_\lambda, \lambda)}{\partial \lambda}) - \beta Cov(f(X_\lambda, \lambda), \frac{\partial E}{\partial \lambda}(X_\lambda, \lambda))$$

Απόδειξη. Αρχικά υπολογίζουμε

$$\frac{\partial Z}{\partial \lambda} = -\sum_{x \in \mathbb{X}} \beta \frac{\partial E}{\partial \lambda} e^{-\beta E(x,\lambda)}$$

$$\frac{\partial}{\partial \lambda}[\mathbb{E}(f(X_\lambda, \lambda)] = \frac{\partial}{\partial \lambda}[\sum_{x \in \mathbb{X}} f(x, \lambda) \frac{e^{-\beta E(x,\lambda)}}{Z}] =$$

$$\sum_{x \in \mathbb{X}} \frac{\partial f(x, \lambda)}{\partial \lambda} \frac{e^{-\beta E(x,\lambda)}}{Z} - \frac{\partial Z}{\partial \lambda} \sum_{x \in \mathbb{X}} f(x, \lambda) \frac{e^{-\beta E(x,\lambda)}}{Z^2} - \beta \sum_{x \in X} f(x, \lambda) \frac{\partial E}{\partial \lambda} \frac{e^{-\beta E(x,\lambda)}}{Z} =$$

$$\mathbb{E}(\frac{\partial f(X_\lambda, \lambda)}{\partial \lambda}) + \beta \sum_{x \in \mathbb{X}} \frac{\partial E}{\partial \lambda} \frac{e^{-\beta E(x,\lambda)}}{Z} \sum_{x \in \mathbb{X}} f(x, \lambda) \frac{e^{-\beta E(x,\lambda)}}{Z} - \beta \mathbb{E}(f(X_\lambda, \lambda) \frac{\partial E}{\partial \lambda}(X_\lambda, \lambda)) =$$

$$\mathbb{E}(\frac{\partial f(X_\lambda, \lambda)}{\partial \lambda}) - \beta[\mathbb{E}(f(X_\lambda, \lambda) \frac{\partial E}{\partial \lambda}(X_\lambda, \lambda)) - \mathbb{E}(f(X_\lambda, \lambda))\mathbb{E}(\frac{\partial E}{\partial \lambda}(X_\lambda, \lambda))] =$$

$$\mathbb{E}(\frac{\partial f(X_\lambda, \lambda)}{\partial \lambda}) - \beta Cov(f(X_\lambda, \lambda), \frac{\partial E}{\partial \lambda}(X_\lambda, \lambda))$$

□

2.4 Θερμοδυναμικό Όριο

Για μια κατανομή Bolzmann όπου ο χώρος καταστάσεων είναι της μορφής $\mathbb{X}^n$ εύλογα θα αναρωτηθεί κάποιος ποια είναι η συμπεριφορά του συστήματος όταν $n \to \infty$. Το όριο αυτό λέγεται θερμοδυναμικό όριο και είναι πάρα πολύ χρήσιμο διότι περιγράφει με έναν αφαιρετικό τρόπο χαρακτηριστικά του συστήματος για αρκετά μεγάλα n.

Ορισμός 2.8. *Για τις κατανομές Bolzmann $\mathbb{B}(\mathbb{X}^n, E_n, \beta)$ με θερμοδυναμικές συναρτήσεις F_n, Φ_n, U_n, S_n αντίστοιχα για $n \in \mathbb{N}^*$ ορίζουμε της επόμενες συναρτήσεις*

1. *(Πυκνότητα ελεύθερης ενέργειας) $f = \lim_n \frac{F_n}{n}$*
2. *(Πυκνότητα ελεύθερης εντροπίας) $\phi = \lim_n \frac{\Phi_n}{n}$*
3. *(Πυκνότητα εσωτερικής ενέργειας) $u = \lim_n \frac{U_n}{n}$*
4. *(Πυκνότητα κανονικής εντροπίας) $s = \lim_n \frac{S_n}{n}$*

για όλα τα $\beta \in (0, \infty)$ που τα αντίστοιχα όρια συγκλίνουν στο $\mathbb{R}$.

Στη φυσική πολλές φορές έχουμε ακούσει τον όρο αλλαγή φάσης να περιγράφει την μετάβαση μιας συλλογής μορίων από στερεή, υγρή ή αέρια κατάσταση σε μια άλλη. Η συμπεριφορά της πυκνότητας ελεύθερης ενέργειας καθορίζει πλήρως τα σημεία που έχουμε αλλαγή φάσης.

Ορισμός 2.9 (Αλλαγή φάσης)**.** *Λέμε ότι στο θερμοδυναμικό όριο των $\mathbb{B}(\mathbb{X}^n, E_n, \beta)$ όπου η συνάρτηση ενέργειας ενδεχομένως εξαρτάται από τις παραμέτρους $(\lambda_1, \lambda_2, ..., \lambda_d) \in \Omega$ όπου Ω χωρίο του $\mathbb{R}^d$ έχουμε αλλαγή φάσης στο σημείο $(\lambda_1, \lambda_2, ..., \lambda_d, \beta_c)$ εάν $\exists k \ \frac{\partial^k}{\partial \beta^k} f$ ασυνεχής ή μη ορισμένη σε αυτό το σημείο και λέμε τάξη της αλλαγής φάσης το μικρότερο δυνατό k που συμβαίνει αυτό.*

Εάν $k = 0$ τότε απλά λέμε ότι έχουμε αλλαγή φάσης. Σε περίπτωση που όλα τα σημεία της ευθείας $\{\lambda = \lambda_c\}$ έχουν αλλαγή φάσης τάξης k τότε απλά λέμε ότι έχουμε αλλαγή φάσης στο λ_c.

Για τις κατανομές Bolzmann $\mathbb{B}(\mathbb{X}^n, E_n, \beta)$ θα ορίσουμε μια ακόμη συνάρτηση η οποία συνδέεται άμεσα με την ασυμπτωτική συμπεριφορά της ενέργειας. Αρχικά ορίζουμε το **ενεργειακό φάσμα**

$$\Omega^n_\Delta(E) = \{x \in \mathbb{X}^n : E \leq E_n(x) < E + \Delta\}$$

και το πλήθος των καταστάσεων του ενεργειακού φάσματος

$$N^n_\Delta(E) = |\Omega^n_\Delta(E)|$$

Ορισμός 2.10 (Μικρο-κανονική πυκνότητα εντροπίας)**.** *Ορίζουμε ως μικροκανονική πυκνότητα εντροπίας την συνάρτηση $\sigma : \mathbb{R} \mapsto \mathbb{R} \cup \{-\infty\}$ που ικανοποιεί την εξίσωση*

$$\sup_{x \in [E, E+\Delta]} \sigma(x) = \lim_n \frac{1}{n} \log N^n_{n\Delta}(nE) \quad \forall \Delta \in \mathbb{R}^+ \quad \forall E \in \mathbb{R}$$

Λήμμα 2.11 (Ασθενής εκτίμηση σαγματικού σημείου)**.** *Έστω* $f, g : [a, b] \mapsto \mathbb{R}$ *όπου* $f \in C^0[a, b]$ *και* $g > 0$*. Τότε ισχύει*

$$\lim_{\lambda\to\infty} \frac{1}{\lambda} \log \int_a^b g(x)e^{\lambda f(x)}dx = \max_{x\in[a,b]} f(x)$$

Απόδειξη. Ως συνεχής η f έχει μέγιστο και ελάχιστο. Ορίζουμε $x_0 = argmax_{x\in[a,b]} f(x)$. Μπορούμε να θεωρήσουμε λόγω συνέχειας της f την περιοχή του x_0, $V(\delta) = (x_0 - \delta, x_0 + \delta) \cap [a, b]$ όπου $f(x) - \epsilon \leq f(x) \quad \forall x \in V(\delta)$ για όποιο $\epsilon > 0$ επιλέξουμε. Ισχύει

$$\int_a^b g(x)e^{\lambda f(x)}dx \geq \int_{V(\delta)} g(x)e^{\lambda f(x)}dx \geq$$

$$e^{\lambda(f(x_0)-\epsilon)} \int_{V(\delta)} g(x)dx \implies$$

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda} \log \int_a^b g(x)e^{\lambda f(x)}dx \geq f(x_0) - \epsilon$$

Η τελευταία ανισότητα ισχύει για όποιο $\epsilon > 0$ και να επιλέξουμε. Έπεται ότι

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda} \log \int_a^b g(x)e^{\lambda f(x)}dx \geq f(x_0)$$

Από την άλλη

$$\int_a^b g(x)e^{\lambda f(x)}dx \leq e^{\lambda f(x_0)} \int_a^b g(x)dx \implies$$

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \int_a^b g(x)e^{\lambda f(x)}dx \leq f(x_0)$$

Από τις 2 τελευταίες ανισότητες προκύπτει το ζητούμενο. □

Θεώρημα 2.12. *Για τις κατανομές Bolzmann* $\mathbb{B}(\mathbb{X}^n, E_n, \beta)$ *που η* σ *ορίζεται στο θερμοδυναμικό όριο και είναι πεπερασμένη και συνεχής στο* $[a, b]$ *και ισχύει* $a \leq \frac{E_n}{n} \leq b$ *τότε*

$$\phi(\beta) = max_{x\in[a,b]}(\sigma(x) - \beta x)$$

Απόδειξη. Θεωρούμε μια κανονική διαμέριση του $[a, b]$ σε N διαστήματα. Θέτουμε $\delta = \frac{1}{N}(b-a)$ και $I_{k,N} = [a+k\delta, a+(k+1)\delta)$. Παρατηρούμε ότι $\Omega^n_{n\delta}(n(a+k\delta)) = \{x \in \mathbb{X}^n : a + k\delta \leq \frac{E_n(x)}{n} < a + (k + 1)\delta\} = \{x \in \mathbb{X}^n : \frac{E_n(x)}{n} \in I_{k,N}\}$. Από τα προηγούμενα και την υπόθεση $a \leq \frac{E_n}{n} \leq b$ έχουμε για την συνάρτηση επιμερισμού

$$Z_n = \sum_{x\in\mathbb{X}^n} e^{-\beta E_n(x)} = \sum_{k=0}^{N-1} \sum_{x\in\Omega^n_{n\delta}(n(a+k\delta))} e^{-\beta E_n(x)} \tag{3}$$

Από την ισότητα (3) και τον ορισμό του ενεργειακού φάσματος προκύπτει άνω φράγμα για την συνάρτηση επιμερισμού

$$Z_n \leq \sum_{k=0}^{N-1} N_{n\delta}^n(n(a+k\delta))e^{-n\beta(a+k\delta)}$$

Για n αρκετά μεγάλο ισχύει από τον ορισμό της σ

$$\exp(n \sup_{x\in[a+k\delta,a+(k+1)\delta]} \sigma(x))e^{-n\epsilon} < N_{n\delta}^n(n(a+k\delta)) < \exp(n \sup_{x\in[a+k\delta,a+(k+1)\delta]} \sigma(x))e^{n\epsilon}$$

Επειδή η σ είναι συνεχής στο συμπαγές σύνολο $[a,b]$ (κλειστό και φραγμένο υποσύνολο του $\mathbb{R}$) έπεται ότι είναι ομοιόμορφα συνεχής. Οπώτε η παραπάνω ανισότητα γίνεται για δ αρκετά μικρό

$$\exp(n\sigma(a+k\delta))e^{-n2\epsilon} < N_{n\delta}^n(n(a+k\delta)) < \exp(n\sigma(a+k\delta))e^{n2\epsilon}$$

Συνεπώς έχουμε

$$Z_n \leq e^{n2\epsilon} \sum_{k=0}^{N-1} e^{n\sigma(a+k\delta)}e^{-n\beta(a+k\delta)}$$

Από το θεώρημα μέσης τιμής ολοκληρωτικού λογισμού

$$\exists c_k \in (a+k\delta, a+(k+1)\delta) : e^{n(\sigma(c_k)-\beta c_k)} = \int_{a+k\delta}^{a+(k+1)\delta} e^{n(\sigma(x)-\beta x)}dx$$

Η συνάρτηση $g(x) - \beta x$ είναι επίσης συνεχής στο συμπαγές $[a,b]$ και συνεπώς έχουμε

$$e^{n(\sigma(c_k)-\beta c_k)}e^{-n\epsilon} < e^{n(\sigma(a+k\delta)-\beta(a+k\delta))} < e^{n(\sigma(c_k)-\beta c_k)}e^{n\epsilon}$$

Έπεται ότι

$$Z_n \leq e^{n3\epsilon} \sum_{k=0}^{N-1} e^{n(\sigma(c_k)-\beta c_k)} = e^{n3\epsilon} \int_a^b e^{n(\sigma(x)-\beta x)}dx \implies$$

$$\limsup_n \frac{1}{n} \log \frac{Z_n}{\int_a^b e^{n(\sigma(x)-\beta x)}dx} \leq 3\epsilon$$

όπου ϵ αυθαίρετος θετικός και συνεπώς

$$\limsup_n \frac{1}{n} \log \frac{Z_n}{\int_a^b e^{n(\sigma(x)-\beta x)}dx} \leq 0$$

Από την (3) και τον ορισμό του ενεργειακού φάσματος επίσης παίρνουμε το κάτω φράγμα

$$Z_n \geq e^{-n\beta\delta} \sum_{k=0}^{N-1} N_{n\delta}^n(n(a+k\delta))e^{-n\beta(a+k\delta)} \geq$$

$$e^{-n(\beta\delta+2\epsilon)}\sum_{k=0}^{N-1} e^{n(\sigma(a+k\delta)-\beta(a+k\delta))} \geq$$

$$e^{-n(\beta\delta+3\epsilon)}\sum_{k=0}^{N-1} e^{n(\sigma(c_k)-\beta c_k)} =$$

$$e^{-n(\beta\delta+3\epsilon)}\int_a^b e^{n(\sigma(x)-\beta x)}dx \implies$$

$$\liminf_n \frac{1}{n}\log\frac{Z_n}{\int_a^b e^{n(\sigma(x)-\beta x)}dx} \geq -\beta\delta - 3\epsilon$$

από την οποία λαμβάνουμε

$$\liminf_n \frac{1}{n}\log\frac{Z_n}{\int_a^b e^{n(\sigma(x)-\beta x)}dx} \geq 0$$

Άρα έχουμε $Z_n \doteq \int_a^b e^{n(\sigma(x)-\beta x)}dx$. Από το λήμμα 2.11 προκύπτει τελικα

$$\phi(\beta) = \lim_n \frac{1}{n}\log Z_n = \lim_n \frac{1}{n}\int_a^b e^{n(\sigma(x)-\beta x)}dx = \max_{x\in[a,b]}(\sigma(x)-\beta x)$$

□

3 Μοντέλο Ising

Το μοντέλο Ising είναι μια μαθηματική περιγραφή στα πλαίσια της στατιστικής φυσικής ενός υλικού με μαγνητικές ιδιότητες. Πρώτα θα ορίσουμε μια γενικευμένη μορφή του εφόσον ο τελικός μας σκοπός είναι να το χρησιμοποιήσουμε σε συνδυαστικά προβλήματα. Κατόπιν θα ασχοληθούμε με κάποιες απλές περιπτώσεις μοντέλων Ising και θα αναφέρουμε αλγορίθμους για προσεγγιστική επίλυση του μοντέλου. Τέλος θα δούμε την πολυπλοκότητα του γενικευμένου μοντέλου Ising και κάποιες παραλλαγές του.

3.1 Γενικευμένο μοντέλο Ising

Για να ορίσουμε ένα **γενικευμένο μοντέλο Ising** χρειαζόμαστε

1. Ένα πεπερασμένο μη κατευθυνόμενο γράφημα $G = (V, E)$
2. Μία συνάρτηση $J : E \mapsto \mathbb{R}$ (μαγνητικές αλληλεπιδράσεις)
3. Μία συνάρτηση $B : V \mapsto \mathbb{R}$ (εξωτερικό μαγνητικό πεδίο)

Ορίζουμε τη συνάρτηση ενέργειας

$$E(\sigma) = -\sum_{\{j,k\}\in E} J_{jk}\sigma_j\sigma_k - \sum_{j\in V} B_j\sigma_j \quad \sigma \in \{-1, 1\}^V$$

όπου $\{-1, 1\}^V$ είναι το σύνολο όλων των συναρτήσεων $V \mapsto \{-1, 1\}$. Η κατανομή $\mathbb{B}(\{-1, 1\}^V, E, \beta)$ που προκύπτει αποτελεί ουσιαστικά ένα γενικευμένο μοντέλο Ising.

Το **μοντέλο Ising** είναι η περίπτωση του γενικευμένου μοντέλου όπου το γράφημα είναι ένα πλέγμα της μορφής $\{1, 2, ..., n\}^d$ που με ακμές συνδέονται τα σημεία που απέχουν 1 και οι μαγνητικές αλληλεπιδράσεις $J_{jk} = 1$.

Όταν έχουμε $J_{jk} > 0$ τότε λέμε ότι η αλληλεπίδραση είναι **φερομαγνητική** δηλαδή οι μαγνήτες j, k τείνουν να προσανατολιστουν. Για να αντιληφθεί κανείς το γεγονός αυτό αρκεί να παρατηρήσει ότι εάν για $\{j, k\} \in E$ ισχύει $\sigma_j\sigma_k > 0$ η συνεισφορά της αλληλεπίδρασης στην ενέργεια γίνεται $-J_{jk}$ που είναι το ελάχιστο δυνατό. Αντίστοιχα όταν έχουμε έχουμε $J_{jk} < 0$ λέμε ότι είναι **αντιφερομαγνητική** η αλληλεπίδραση, δηλαδή οι μαγνήτες τείνουν να έχουν αντίθετα spins. Αυτό μπορεί να υποβάλει την ιδέα ότι μπορούμε για ένα δεδομένο σταθερό εξωτερικό μαγνητικό πεδίο B είναι εύκολος ο υπολογισμός της κατάστασης ελάχιστης ενέργειας. Αυτό όμως γίνεται αντιληπτό ότι δεν ισχύει αν θεωρήσουμε το εξής αντιπαράδειγμα. Έστω $G = (\{1, 2, 3\}, \{\{1, 2\}, \{2, 3\}, \{3, 1\}\})$ με $B = 0$ και $J_{jk} = -1 \quad \forall\{j, k\}$. Βλέπουμε πως και οι 3 μαγνήτες τείνουν να έχουν αντίθετο spin από ότι οι άλλοι, αλλά δεν είναι εφικτό αυτό διότι τουλάχιστον 2 θα έχουν πάντα το ίδιο spin. Υλικά με μαγνητικές αλληλεπιδράσεις όπως του παραπάνω παραδείγματος, δηλαδή απαγορευτικές για να επιτύχουμε την ελάχιστη δυνατή ενέργεια, λέγονται **spin glasses**. Καταλήγοντας, η εύρεση της ελάχιστης δυνατής ενέργειας είναι ένα εξαιρετικά πολύπλοκο πρόβλημα και όπως θα δούμε το αντίστοιχο πρόβλημα απόφασης είναι NP-Complete.

3.2 Επιλύσιμα μοντέλα

Σκοπός μας είναι να μπορέσουμε να μελετήσουμε το θερμοδυναμικό όριο ενός μοντέλου. Έτσι είναι εφικτό να κατανοήσουμε ποιοτικά τη συμπεριφορά του συστήματος. Για να το πετύχουμε αυτό χρειάζεται να υπολογίσουμε την πυκνότητα ελεύθερης ενέργειας. Παρόλο που είναι απλό το μοντέλο Ising δεν έχει επιλυθεί για $d > 2$. Ακόμα και η περίπτωση $d = 2$ έχει αρκετά περίπλοκη λύση και χρειάστηκαν αρκετά χρόνια για να λυθεί. Σε αυτό το κεφάλαιο θα μελετήσουμε το μονοδιάστατου μοντέλο Ising και το μοντέλο Curie-Weiss.

3.2.1 Μονοδιάστατο Μοντέλο Ising

Σε αυτή την ενότητα θα μελετήσουμε το μονοδιάστατο μοντέλο Ising με ομογενές μαγνητικό πεδίο B. Άρα μελετάμε το μοντέλο $G = (V, E)$ όπου $V = \{j \in \mathbb{N} : 1 \leq j \leq n\}$, $E = \{\{j,k\} : k = j+1 \wedge j,k \in V\}$, $J_{jk} = 1 \quad \forall \{j,k\} \in E$ και $B_j = B \quad \forall j \in V$. Για να υπολογίσουμε την πυκνότητα ελεύθερης ενέργειας του θερμοδυναμικού ορίου πρέπει να εκτιμήσουμε την εκθετική τάξη της συνάρτησης επιμερισμού Z. Για λόγους συντομίας θα γράφουμε $\sum_{\sigma_1,\sigma_2,\ldots,\sigma_n}$ ή $\sum_\sigma$ αντί για $\sum_{\sigma\in\{-1,1\}^V}$.

$$Z_n = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_n} \exp(-\beta E(\sigma)) = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_n} \exp(\beta(\sum_{j=1}^{n-1} \sigma_j\sigma_{j+1} + B\sum_{j=1}^{n} \sigma_j))$$

Όπως καταλαβαίνει κανείς ή μέσο του παραπάνω τύπου ή λόγο φυσικής διαίσθησης του προβλήματος μπορούμε να εκφράσουμε αναδρομικά την συνάρτηση επιμερισμού. Το γράφημά μας είναι μια αλυσίδα με n στοιχεία Συνεπώς αναζητούμε έκφραση της συνάρτησης επιμερισμού των $k+1$ στοιχείων από τα k και την τιμή του $k+1$ spin σ_{k+1}. Για αυτό ορίζουμε

$$z_k(\sigma_k) = \sum_{\sigma_1,\sigma_2,\ldots\sigma_{k-1}} \exp(\beta(\sum_{j=1}^{k-1} \sigma_j\sigma_{j+1} + B\sum_{j=1}^{k} \sigma_j)) \quad 2 \leq k \leq n$$

και

$$z_1(\sigma_1) = \exp(\beta B\sigma_1)$$

Τώρα παρατηρούμε για κάθε $\sigma_{k+1} \in \{-1, 1\}$

$$z_{k+1}(\sigma_{k+1}) = \sum_{\sigma_1,\sigma_2,\ldots\sigma_k} \exp(\beta(\sum_{j=1}^{k} \sigma_j\sigma_{j+1} + B\sum_{j=1}^{k+1} \sigma_j)) =$$

$$\sum_{\sigma_k}\sum_{\sigma_1,\sigma_2,\ldots,\sigma_{k-1}} \exp(\beta(\sum_{j=1}^{k-1} \sigma_j\sigma_{j+1} + B\sum_{j=1}^{k} \sigma_j) + \beta\sigma_k\sigma_{k+1} + \beta B\sigma_{k+1}) =$$

$$\sum_{\sigma_k} \exp(\beta\sigma_k\sigma_{k+1} + \beta B\sigma_{k+1}) z_k(\sigma_k)$$

και επίσης για κάθε $\sigma_2 \in \{-1, 1\}$

$$z_2(\sigma_2) = \sum_{\sigma_1} \exp(\beta\sigma_1\sigma_2 + \beta B\sigma_1 + \beta B\sigma_2) = \sum_{\sigma_1} z_1(\sigma_1)\exp(\beta\sigma_1\sigma_2 + \beta B\sigma_2)$$

Σε μορφή πίνακα οι παραπάνω παρατηρήσεις γράφονται

$$\begin{pmatrix} z_{k+1}(1) \\ z_{k+1}(-1) \end{pmatrix} = \begin{pmatrix} e^{\beta+\beta B} & e^{-\beta+\beta B} \\ e^{-\beta-\beta B} & e^{\beta-\beta B} \end{pmatrix} \begin{pmatrix} z_k(1) \\ z_k(-1) \end{pmatrix}$$

Ο πίνακας

$$T = \begin{pmatrix} e^{\beta+\beta B} & e^{-\beta+\beta B} \\ e^{-\beta-\beta B} & e^{\beta-\beta B} \end{pmatrix}$$

λέγεται **πίνακας μεταφοράς**.
Τέλος έχουμε ότι

$$Z_n = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_n} \exp(\beta(\sum_{j=1}^{n-1}\sigma_j\sigma_{j+1} + B\sum_{j=1}^{n}\sigma_j)) =$$

$$\sum_{\sigma_n}\sum_{\sigma_1,\sigma_2,\ldots,\sigma_{n-1}} \exp(\beta(\sum_{j=1}^{n-1}\sigma_j\sigma_{j+1} + B\sum_{j=1}^{n}\sigma_j)) = \sum_{\sigma_n} z_n(\sigma_n) =$$

$$< \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} z_n(1) \\ z_n(-1) \end{pmatrix} > = < \begin{pmatrix} 1 \\ 1 \end{pmatrix}, T^{n-1}\begin{pmatrix} e^{\beta B} \\ e^{-\beta B} \end{pmatrix} >$$

Οι ιδιοτιμές του πίνακα T είναι

$$\lambda_1 = e^{\beta}\cosh(\beta B) + \sqrt{e^{2\beta}\sinh^2(\beta B) + e^{-2\beta}}$$

$$\lambda_2 = e^{\beta}\cosh(\beta B) - \sqrt{e^{2\beta}\sinh^2(\beta B) + e^{-2\beta}}$$

Έστω u_1, u_2 ιδιοδιανύσματα των ιδιοτιμών λ_1, λ_2 αντίστοιχα. Τα u_1, u_2 είναι γραμμικώς ανεξάρτητα και συνεπώς αποτελούν βάση του $\mathbb{R}^2$. Άρα εάν $(e^{\beta B}, e^{-\beta B}) = a_1u_1 + a_2u_2$

$$Z_n = < \begin{pmatrix} 1 \\ 1 \end{pmatrix}, T^{n-1}(a_1u_1 + a_2u_2) > = < \begin{pmatrix} 1 \\ 1 \end{pmatrix}, a_1\lambda_1^{n-1}u_1 + a_2\lambda_2^{n-1}u_2 >$$

και συνεπώς

$$Z_n = c_1\lambda_1^{n-1} + c_2\lambda_2^{n-1}$$

όπου $c_1 = < \begin{pmatrix} 1 \\ 1 \end{pmatrix}, u_1 > a_1$

Για να κυριαρχεί ο όρος λ_1^{n-1} όπου $\lambda_1 > \lambda_2$ αρκεί $c_1 \neq 0$. Από το θεώρημα Perron-Frobenius γνωρίζουμε ότι η λ_1 είναι ιδιοτιμή Perron-Frobenius και μπορούμε να διαλέξουμε το u_1 ώστε όλες οι συντεταγμένες του να είναι θετικές. Συνεπώς αρκεί $a_1 \neq 0$. Εάν $a_1 = 0$ τότε $\begin{pmatrix} e^{\beta B} \\ e^{-\beta B} \end{pmatrix} = a_2u_2$ το οποίο θα σήμαινε ότι είναι ιδιοδιάνυσμα του T το οποίο με πράξεις βλέπουμε ότι είναι άτοπο. Συνεπώς η πυκνότητα ελεύθερης εντροπίας είναι

$$\phi = \lim_n \frac{\log Z_n}{n} = \lim_n \frac{\log(c_1\lambda_1^{n-1} + c_2\lambda_2^{n-1})}{n} =$$

$$\lim_n \frac{\log(c_1 + c_2\frac{\lambda_2}{\lambda_1}^{n-1}) + (n-1)\log\lambda_1}{n} = \log\lambda_1$$

Άρα

$$\phi = \log(e^\beta\cosh(\beta B) + \sqrt{e^{2\beta}\sinh^2(\beta B) + e^{-2\beta}})$$

Εφόσον $f = -\frac{1}{\beta}\phi$ συμπεραίνουμε ότι η f είναι αναλυτική στο $(0,\infty)\times\mathbb{R}$ ως συνάρτηση του (β, B) και άρα δεν έχουμε καμία αλλαγή φάσης.

Μια σημαντική για την μελέτη της επιρροής του μαγνητικού πεδίου B στο σύστημα ποσότητα είναι ο **μέσος μαγνητισμός** ο οποίος είναι

$$M = \frac{1}{|V|}\sum_{j\in V}\mathbb{E}(\sigma_j) = \mathbb{E}(\frac{1}{|V|}\sum_{j\in V}\sigma_j)$$

Επίσης θα λέμε **στιγμιαίο μαγνητισμό** την ποσότητα

$$m(\sigma) = \frac{1}{|V|}\sum_{j\in V}\sigma_j \quad \sigma\in\{-1,1\}^V$$

η οποία στο επόμενο μοντέλο που θα μελετήσουμε παίζει κομβικό ρόλο.

Για τον υπολογισμό του μέσου μαγνητισμού παρατηρούμε ότι

$$\frac{\partial\log Z_n}{\partial B} = \frac{1}{Z_n}\sum_\sigma \frac{\partial\exp(\beta(\sum_{j=1}^{n-1}\sigma_j\sigma_{j+1} + B\sum_{j=1}^n\sigma_j))}{\partial B} =$$

$$\beta\sum_\sigma(\sum_{j=1}^n\sigma_j)\exp(\beta(\sum_{j=1}^{n-1}\sigma_j\sigma_{j+1} + B\sum_{j=1}^n\sigma_j)) =$$

$$\beta\sum_\sigma(\sum_{j=1}^n\sigma_j)p_\beta(\sigma)Z_n = \beta n M_n Z_n \implies$$

$$M_n = \frac{1}{\beta n Z_n}\frac{\partial\log Z_n}{\partial B}$$

όπου M_n ο μέσος μαγνητισμός του μοντέλου με n spins

Στο θερμοδυναμικό όριο η παραπάνω ποσότητα γίνεται

$$M_n = \frac{1}{\beta n}\frac{\frac{\partial\log Z_n}{\partial B}}{Z_n} = \frac{1}{\beta n}\frac{\frac{\partial c_1}{\partial B}\lambda_1^{n-1} + \frac{\partial c_2}{\partial B}\lambda_2^{n-1} + (n-1)[c_1\lambda_1^{n-2}\frac{\partial\lambda_1}{\partial B} + c_2\lambda_2^{n-2}\frac{\partial\lambda_2}{\partial B}]}{c_1\lambda_1^{n-1} + c_2\lambda_2^{n-1}} =$$

$$\frac{1}{\beta}\frac{\frac{1}{n}[\frac{\partial c_1}{\partial B} + \frac{\partial c_2}{\partial B}(\frac{\lambda_2}{\lambda_1})^{n-1}] + (1-\frac{1}{n})\frac{1}{\lambda_1}[c_1\frac{\partial\lambda_1}{\partial B} + c_2(\frac{\lambda_2}{\lambda_1})^{n-2}\frac{\partial\lambda_2}{\partial B}]}{c_1 + c_2\frac{\lambda_2}{\lambda_1}^{n-1}} \implies$$

$$\lim_n M_n = \frac{1}{\beta}\frac{c_1 \frac{1}{\lambda_1}\frac{\partial \lambda_1}{\partial B}}{c_1} = \frac{1}{\beta}\frac{\partial \log \lambda_1}{\partial B} = \frac{1}{\beta}\frac{\partial \phi}{\partial B}$$

Επειδή σκοπός μας είναι να μελετήσουμε συστήματα που προκύπτουν από μετασχηματισμούς συνδυαστικών προβλημάτων από εδώ και πέρα θα λέμε μοντέλο Ising θα εννοούμε το γενικευμένο μοντέλο και όχι το κλασσικό μοντέλο.

3.2.2 Μοντέλο Curie-Weiss

Το μοντέλο Curie-Weiss είναι ένα ακόμη απλό μοντέλο το οποίο ανήκει στην ευρύτερη κατηγορία των **μοντέλων μέσου πεδίου**. Το γράφημα του μοντέλου είναι μια κλίκα με n κόμβους και οι μαγνητικές αλληλεπιδράσεις έχουν ισχύ $\frac{1}{n}$. Πάλι για λόγους απλότητας θα θεωρήσουμε ότι το μαγνητικό πεδίο είναι ομοιογενές με ισχύ B. Άρα έχουμε

$$E_n(\sigma) = -\frac{1}{n}\sum_{j<k}\sigma_j\sigma_k - B\sum_{j=1}^{n}\sigma_j$$

Παρατηρούμε ότι

$$n^2 m_n(\sigma)^2 = \sum_{j=1}^{n}\sigma_j^2 + 2\sum_{j<k}\sigma_j\sigma_k \implies$$

$$\frac{1}{2}(n^2 m_n(\sigma)^2 - n) = \sum_{j<k}\sigma_j\sigma_k \implies$$

$$E_n(\sigma) = \frac{1}{2} - \frac{n}{2}m_n(\sigma)^2 - nBm_n(\sigma)$$

Συνεπώς η ενέργεια εξαρτάται μόνο από τον στιγμιαίο μαγνητισμό. Ένα μοντέλο με αυτή την ιδιότητα λέγεται **μοντέλο μέσου πεδίου**. Άρα λαμβάνουμε την απλοποιημένη συνάρτηση ενέργειας

$$E_n(m) = \frac{1}{2} - \frac{n}{2}m^2 - nBm$$

Η ενέργεια αυτή απλοποιείται περαιτέρω από την πρόταση 2.2 στην

$$E_n(m) = -\frac{n}{2}m^2 - nBm$$

Το πρόβλημα που έχουμε τώρα είναι ότι υπάρχουν πολλές καταστάσεις σ που αντιστοιχούν σε μια τιμή m. Άρα για να εκφράσουμε την συνάρτηση επιμερισμού με την απλοποιημένη ενέργεια πρέπει να βρούμε το πλήθος $N_n(m)$ των καταστάσεων σ που έχουν στιγμιαίο μαγνητισμό m.

Εάν από τα n spins τα k είναι 1 και τα υπόλοιπα -1 έχουμε $m = \frac{k-(n-k)}{n} = -1+2\frac{k}{n}$. Αυτές οι καταστάσεις είναι συνολικά $\binom{n}{k}$ όπου λύνοντας την προηγούμενη εξίσωση ως προς k λαμβάνουμε $k = n\frac{1+m}{2}$. Άρα έχουμε

$$Z_n = \sum_{m\in G_n} N_n(m)\exp(-\beta E_n(m)) = \sum_{m\in G_n}\binom{n}{n\frac{1+m}{2}}\exp(n(\frac{\beta}{2}m^2 + \beta Bm))$$

Όπου G_n το σύνολο όλων των δυνατών τιμών του m για n spins, δηλαδή $G_n = \{-1+2\frac{k}{n} : k \in \mathbb{N} \wedge 0 \leq k \leq n\}$. Παρατηρούμε ότι το G_n είναι μια ομοιόμορφη διαμέριση του $[-1,1]$ σε n διαστήματα. Θέτουμε $x_k = -1+2\frac{k}{n} \quad k \in \mathbb{N} \wedge 0 \leq k \leq n$.

Έστω $h(p) = -p\log p - (1-p)\log(1-p) \quad p \in [0,1]$. Από το λήμμα 1.10 λαμβάνουμε

$$Z_n = \sum_{m \in G_n} \exp(O(\log n + nh(\frac{1+m}{2})) \exp(n(\frac{\beta}{2}m^2 + \beta Bm)) =$$

$$\exp(O(\log n)) \sum_{k=0}^{n} \exp(n[\frac{\beta}{2}x_k^2 + \beta Bx_k + h(\frac{1+x_k}{2})])$$

Έστω $g(x) = \frac{\beta}{2}x^2 + \beta Bx + h(\frac{1+x}{2}) \quad x \in [-1,1]$. Έχουμε g συνεχής στο $[-1,1]$ και το $[-1,1]$ είναι συμπαγές ως κλειστό και φραγμένο υποσύνολο του $\mathbb{R}$. Συνεπώς έχουμε g ομοιόμορφα συνεχής στο $[-1,1]$. Από το θεώρημα μέσης τιμής ολοκληρωτικού λογισμού έχουμε για $0 \leq k \leq n-1$

$$\exists c_k \in (x_k, x_{k+1}) : \frac{2}{n}\exp(ng(c_k)) = \int_{x_k}^{x_{k+1}} \exp(ng(x))dx$$

Για $\epsilon > 0$ από την ομοιόμορφη συνέχεια της g υπάρχει $\delta > 0$ ώστε $|u-v| < \delta \implies |g(u)-g(v)| < \epsilon$. Εφόσον $x_{k+1} - x_k = \frac{2}{n}$ για n αρκετά μεγάλο $\frac{2}{n} < \delta$ και συνεπώς $|g(c_k) - g(x_k)| < \epsilon \implies e^{-n\epsilon} < \exp(n(g(c_k) - g(x_k)) < e^{n\epsilon}$

$$Z_n = e^{O(\log n)} \sum_{k=0}^{n} e^{ng(x_k)} < e^{O(\log n)} e^{n\epsilon} \sum_{k=1}^{n-1} e^{ng(c_k)} + e^{O(\log n)} e^{ng(x_n)} =$$

$$e^{O(\log n)} e^{n\epsilon} \int_{-1}^{1} e^{ng(x)}dx + e^{O(\log n)} e^{ng(1)} \implies$$

$$\frac{Z_n}{\int_{-1}^{1} e^{ng(x)}dx} < e^{O(\log n)}[e^{n\epsilon} + \frac{e^{ng(1)}}{\int_{-1}^{1} e^{ng(x)}dx}]$$

Εφόσον $\lim_{x\to 1} g'(x) = \lim_{x \to 1}(\beta x + \beta B + \frac{1}{2}\log\frac{1-x}{1+x}) = -\infty$ το οποίο σημαίνει ότι η g έχει τοπικό ελάχιστο στο 1. Άρα αν θεωρήσουμε μια περιοχή του 1 $(1-\xi, 1], \xi > 0$ όπου $g(x) > g(1) \quad \forall x \in (1-\xi, 1]$ λαμβάνουμε

$$\int_{-1}^{1} e^{n(g(x)-g(1))}dx > \int_{1-\xi}^{1} e^{n(g(x)-g(1))}dx > \int_{1-\xi}^{1} 1dx = \xi \implies$$

$$\frac{e^{ng(1)}}{\int_{-1}^{1} e^{ng(x)}dx} = O(1) \implies$$

$$\frac{Z_n}{\int_{-1}^{1} e^{ng(x)}dx} < e^{O(\log n)}[e^{n\epsilon} + O(1)] \implies$$

$$\frac{1}{n}\log\frac{Z_n}{\int_{-1}^{1}e^{ng(x)}dx} < O(\frac{\log n}{n}) + \frac{1}{n}\log(e^{n\epsilon}+O(1)) \implies$$

$$\limsup_n \frac{1}{n}\log\frac{Z_n}{\int_{-1}^{1}e^{ng(x)}dx} \leq \epsilon$$

Επίσης έχουμε

$$Z_n = e^{O(\log n)}\sum_{k=0}^{n} e^{ng(x_k)} > e^{O(\log n)}\sum_{k=0}^{n-1} e^{ng(x_k)} >$$

$$e^{O(\log n)}e^{-n\epsilon}\sum_{k=0}^{n-1} e^{ng(c_k)} = e^{O(\log n)}e^{-n\epsilon}\int_{-1}^{1}e^{ng(x)}dx \implies$$

$$\frac{1}{n}\log\frac{Z_n}{\int_{-1}^{1}e^{ng(x)}dx} > O(\frac{\log n}{n}) - \epsilon \implies$$

$$\liminf_n \frac{1}{n}\log\frac{Z_n}{\int_{-1}^{1}e^{ng(x)}dx} \geq -\epsilon$$

Όμως το ϵ το επιλέγουμε αυθαίρετα και συνεπώς έχουμε

$$Z_n \doteq \int_{-1}^{1}e^{ng(x)}dx$$

Αρκεί λοιπόν να υπολογίσουμε την εκθετική τάξη του ολοκληρώματος $\int_{-1}^{1}e^{ng(x)}dx$ για τον υπολογισμό της ϕ. Από το λήμμα 2.11 έχουμε

$$\lim_n \frac{1}{n}\log\int_{-1}^{1}e^{ng(x)}dx = g(x_0)$$

Άρα η ϕ είναι

$$\phi = g(x_0)$$

Επίσης ισχύει $\lim_{x\to -1} g'(x) = \lim_{x\to -1}(\beta x + \beta B + \frac{1}{2}\log\frac{1-x}{1+x}) = \infty$ και άρα το -1 είναι τοπικό ελάχιστο. Εφόσον το σημείο μέγιστου x_0 που ψάχνουμε είναι για όλες τις τιμές των β, B στο $(-1,1)$, λόγω των πλευρικών ορίων που είδαμε προηγούμενος, πρέπει να ικανοποιεί την εξίσωση $g'(x) = 0$. Θα μελετήσουμε την περίπτωση όπου $B = 0$ που είναι η απλούστερη. Εάν $\beta \leq 1$ έχουμε

$$g''(x) = \beta - \frac{1}{1-x^2} \leq 0 \quad \forall x \in [-1,1]$$

όπου η ισότητα ισχύει αν και μόνο αν $x = 0, \beta = 1$. Άρα g' γνησίως φθίνουσα. Επίσης $g'(0) = 0$ άρα είναι η μοναδική ρίζα του 0 και συνεπώς πρέπει να είναι το σημείο μεγίστου που ψάχνουμε. Άρα

$$\phi = g(0) = \log 2 \quad \beta \leq 1$$

Από την άλλη εάν $\beta \geq 1$ θέτουμε $a = \sqrt{1-\frac{1}{\beta}}$ και παρατηρούμε ότι g'' είναι φθίνουσα στο $[0,1)$ και είναι άρτια από αυτά προκύπτουν

$$g''(x) = g''(-x) = \beta - \frac{1}{1-(-x)^2} < \beta - \frac{1}{1-(-a)^2} = 0 \quad \forall x \in (-1,-a)$$

$$g''(x) = g''(|x|) = \beta - \frac{1}{1-|x|^2} > \beta - \frac{1}{1-|a|^2} = 0 \quad \forall x \in (-a,a)$$

$$g''(x) = \beta - \frac{1}{1-x^2} < \beta - \frac{1}{1-a^2} = 0 \quad \forall x \in (a,1)$$

Σε αυτήν την περίπτωση λοιπόν το κρίσιμο σημείο στο 0 είναι τοπικό ελάχιστο αφού $g''(0) > 0$. Η g' είναι αύξουσα στο $(-a,a)$ άρα $g'(a) > g'(0) = 0$. Όμως $\lim_{x\to 1} g'(x) = -\infty$ άρα υπάρχει μοναδικό $\xi \in (a,1)$ τέτοιο ώστε $g'(\xi) = 0$. Άρα στο ξ αφού $g''(\xi) < 0$. Τέλος η g είναι άρτια και συνεπώς στο $-\xi$ έχουμε δεύτερο τοπικό μέγιστο με ίδια τιμή με το ξ. Οπότε προκύπτει

$$\phi = g(\xi) \quad \beta > 1$$

Το ξ μεταβάλλεται με το β οπότε η ϕ είναι μη σταθερή για $\beta > 1$ και σταθερή για $\beta \leq 1$ το οποίο σημαίνει ότι δεν είναι αναλυτική η ϕ στο 1. Άρα για $\beta_c = 1$ έχουμε αλλαγή φάσης στο μοντέλο Curie-Weiss. Η συγκεκριμένη αλλαγή φάσης προκύπτει από την διακλάδωση της g για $B = 0$ στο $\beta = 1$ όπου το σημείο τοπικού μεγίστου γίνεται τοπικού ελαχίστου και προκύπτουν δύο νέα σημεία τοπικού μεγίστου γύρο από αυτό. Η περίπτωση που $B \neq 0$ μελετάται με αντίστοιχο τρόπο δηλαδή βρίσκουμε τα σημεία διακλάδωσης του ενός μεγίστου σε τρία κρίσιμα σημεία ένα εκ των οποίων είναι το μέγιστο.

3.3 Πολυπλοκότητα μοντέλου Ising

Σε αυτή την ενότητα θα μελετήσουμε την υπολογιστική πολυπλοκότητα του υπολογίσμου της ελάχιστης ενέργειας ενός μοντέλου Ising. Όπως όλα τα υπολογιστικά προβλήματα έχει υπολογιστική μορφή και μορφή απόφασης. Για την θεωρία πολυπλοκότητας σημαντικότερη είναι η μορφή απόφασης την οποία και θα χρησιμοποιήσουμε.

Πρόβλημα 3.1 (Πρόβλημα απόφασης άνω φράγματος ελάχιστης ενέργειας μοντέλου Ising)**.** *Να αποφανθούμε εάν η ελάχιστη δυνατή ενέργεια του μοντέλου Ising με γράφημα $G = (V, E)$, μαγνητικές αλληλεπιδράσεις J και μαγνητικό πεδίο B είναι μικρότερη η ίση από E_0. Μας δίνονται σαν είσοδος το γράφημα G, τα βάρη J, το μαγνητικό πεδίο B και το άνω φράγμα της ελάχιστης ενέργειας E_0.*

Έστω μοντέλο Ising με γράφημα $G = (V, E)$, μαγνητικές αλληλεπιδράσεις J, μαγνητικό πεδίο B και άνω φράγμα ελάχιστης ενέργειας E_0. Για μια κατάσταση σ με $E(\sigma) \leq E_0$ υπολογίζουμε $E(\sigma)$ σε $O(|V| + |E|)$ πράξεις το οποίο σημαίνει ότι μπορούμε να ελέγξουμε εάν η απάντηση στο πρόβλημα απόφασης του μοντέλου Ising είναι καταφατική σε πολυωνυμικό χρόνο. Άρα το πρόβλημα απόφασης άνω φράγματος ελάχιστης ενέργειας του μοντέλου Ising είναι NP.

Πρόβλημα 3.2 (Πρόβλημα απόφασης κάτω φράγματος weighted max-cut)**.** *Να αποφανθούμε εάν σε ακατεύθυντο γράφημα $G = (V, E)$ και βάρη w εάν υπάρχει υποσύνολο των κόμβων U με άθροισμα βαρών ακμών από το $V \setminus U$ στο U τουλάχιστον c. Μας δίνονται ως είσοδος το γράφημα G, τα βάρη w και το κάτω φράγμα c.*

Πρόταση 3.3. *Το πρόβλημα απόφασης weighted max-cut ανάγεται σε πολυωνυμικό χρόνο στο πρόβλημα απόφασης ελάχιστης ενέργειας του μοντέλου Ising*

Απόδειξη. Έστω $G = (V, E)$ με βάρη w και το κάτω φράγμα c. Κατασκευάζουμε το μοντέλο Ising με γράφημα G, μαγνητικές αλληλεπιδράσεις $J = -w$ και ομογενές μαγνητικό πεδίο 0. Θέτουμε $s = \sum_{\{j,k\}\in E} w_{jk}$ και $E_0 = s - 2c$. Εάν για κατάσταση σ έχουμε

$$E(\sigma) \leq E_0$$

Θέτουμε $U = \{j \in V : \sigma_j = 1\}$, E_U τις ακμές μεταξύ των κόμβων του U, $E_{V\setminus U}$ τις ακμές μεταξύ των κόμβων του $V \setminus U$ και E^* τις ακμές από κόμβο του U σε κόμβο του $V \setminus U$.

$$E(\sigma) = -\sum_{\{j,k\}\in E} J_{jk}\sigma_j\sigma_k = -\sum_{\{j,k\}\in E_U} J_{jk} - \sum_{\{j,k\}\in E_{V\setminus U}} J_{jk} + \sum_{\{j,k\}\in E^*} J_{jk} =$$

$$-\sum_{\{j,k\}\in E_U} J_{jk} - \sum_{\{j,k\}\in E_{V\setminus U}} J_{jk} - \sum_{\{j,k\}\in E^*} J_{jk} + 2\sum_{\{j,k\}\in E^*} J_{jk} =$$

$$-\sum_{\{j,k\}\in E} J_{jk} + 2\sum_{\{j,k\}\in E^*} J_{jk} =$$

$$s - 2 \sum_{\{j,k\} \in E^*} w_{jk}$$

Άρα

$$E(\sigma) \leq E_0 \implies s - 2 \sum_{\{j,k\} \in E^*} w_{jk} \leq s - 2c \implies$$

$$\sum_{\{j,k\} \in E^*} w_{jk} \geq c$$

Άρα η επιλογή U δίνει άθροισμα βαρών των ακμών μεταξύ U και $V \setminus U$ τουλάχιστον c. Αντίστροφα εάν μια επιλογή U δίνει άθροισμα βαρών των ακμών μεταξύ U και $V \setminus U$ τουλάχιστον c τότε κατασκευάζουμε την κατάσταση σ με $\sigma_j = 1 \quad \forall j \in U$, $\sigma_j = -1 \quad \forall j \in V \setminus U$. Θέτουμε E_U τις ακμές μεταξύ των κόμβων του U, $E_{V \setminus U}$ τις ακμές μεταξύ των κόμβων του $V \setminus U$ και E^* τις ακμές από κόμβο του U σε κόμβο του $V \setminus U$. Προκύπτει πάλι

$$E(\sigma) = s - 2 \sum_{\{j,k\} \in E^*} w_{jk}$$

όμως $\sum_{\{j,k\} \in E^*} w_{jk} \geq c$ άρα

$$E(\sigma) \leq s - 2c = E_0$$

Άρα μετατρέψαμε μια είσοδο του προβλήματος απόφασης weighted max-cut σε μια ισοδύναμη είσοδο του προβλήματος απόφασης του μοντέλου Ising σε πολυωνυμικό χρόνο. □

Από την παραπάνω αναγωγή έπεται ότι το πρόβλημα απόφασης άνω φράγματος ελάχιστης ενέργεια μοντέλου Ising είναι NP-Complete εφόσον είναι γνωστό ότι το πρόβλημα απόφασης weighted max-cut είναι NP-Complete.

3.4 Παρόμοια μοντέλα

Σε αυτό το κεφάλαιο θα συζητήσουμε για παρόμοια μοντέλα με το μοντέλο Ising.

3.4.1 Μοντέλο Potts

Όπως και στην περίπτωση του μοντέλου Ising έχουμε ένα ακατεύθυντο γράφημα $G = (V, E)$ με μαγνητικές αλληλεπιδράσεις $J : E \mapsto \mathbb{R}$. Ορίζουμε ως ενέργεια του μοντέλου Potts με $q \in \mathbb{N} \setminus \{0, 1\}$ spins την συνάρτηση ενέργειας $E : \{n \in \mathbb{N} : 1 \leq n \leq q\}^V \mapsto \mathbb{R}$ με τύπο

$$E(\sigma) = -\sum_{\{j,k\}\in E} J_{jk}\delta(\sigma_j, \sigma_k)$$

όπου δ το δέλτα του Kronecker.

Το μοντέλο Potts με q spins είναι η κατανομή Bolzmann $\mathbb{B}(\mathbb{X}, E, \beta)$, όπου $\mathbb{X} = \{n \in \mathbb{N} : 1 \leq n \leq q\}^V$

Ουσιαστικά το μοντέλο Potts με q spins αποτελεί παραλλαγή του μοντέλου Ising όπου οι μαγνήτες μπορούν να προσανατολιστούν σε q διαφορετικές κατευθύνσεις.

Πρόταση 3.4. *Το μοντέλο Potts με 2 spins, γράφημα $G = (V, E)$ και μαγνητικές αλληλεπιδράσεις J είναι ισοδύναμο με το μοντέλο Ising με γράφημα $G = (V, E)$, μαγνητικές αλληλεπιδράσεις $\frac{1}{2}J$ και ομογενές μαγνητικό πεδίο 0.*

Απόδειξη. Η ενέργεια του μοντέλου Potts είναι

$$E_{potts}(\sigma) = -\sum_{\{j,k\}\in E} J_{jk}\delta(\sigma_j, \sigma_k)$$

μπορούμε να αντικαταστήσουμε το σύνολο τιμών των spins του μοντέλου Potts $\{1, 2\}$ με το $\{-1, 1\}$ αντιστοιχίζοντας $1 \mapsto -1$, $2 \mapsto 1$ εφόσον αυτό δεν επηρεάζει την ενέργεια. Θεωρώ $U = \{j \in V : \sigma_j = 1\}$, E_U τις ακμές μεταξύ κόμβων του U, $E_{V\setminus U}$ τις ακμές μεταξύ κόμβων του $V \setminus U$ και E^* τις ακμές από κόμβο του U σε κόμβο του $V \setminus U$. Τέλος θέτω $s = \sum_{\{j,k\}\in E} J_{jk}$

$$E_{potts}(\sigma) = -\sum_{\{j,k\}\in E} J_{jk}\delta(\sigma_j, \sigma_k) = -\sum_{\{j,k\}\in E_U} J_{jk} - \sum_{\{j,k\}\in E_{V\setminus U}} J_{jk}$$

$$E_{ising}(\sigma) = -\sum_{\{j,k\}\in E} \frac{1}{2}J_{jk}\sigma_j\sigma_k =$$

$$\frac{1}{2}(-\sum_{\{j,k\}\in E_U} J_{j,k} - \sum_{\{j,k\}\in E_{V\setminus U}} J_{j,k} + \sum_{\{j,k\}\in E^*} J_{j,k}) =$$

$$\frac{1}{2}(2E_{potts}(\sigma) + s) = E_{potts}(\sigma) + \frac{s}{2}$$

Από την πρόταση 2.2 και την παραπάνω ισότητα πέρνουμε το ζητούμενο. □

Αρά το μοντέλο Potts είναι μια γενίκευση του μοντέλου Ising χωρίς την επίδραση μαγνητικού πεδίου. Αυτό έχει ως συνέπεια το αντίστοιχο πρόβλημα απόφασης του μοντέλου Potts να είναι επίσης NP-Complete.

Ορισμός 3.5 (Πρόβλημα απόφασης μοντέλου Potts). *Να αποφανθούμε εάν η ελάχιστη δυνατή ενέργεια του μοντέλου Potts με q spins, γράφημα $G = (V, E)$ και μαγνητικές αλληλεπιδράσεις J είναι μικρότερη η ίση από E_0. Μας δίνονται σαν είσοδος το q, το γράφημα G, οι μαγνητικές αλληλεπιδράσεις J και ο αριθμός E_0.*

Πρώτα παρατηρούμε ότι η αναγωγή από το πρόβλημα απόφασης weighted max-cut στο πρόβλημα απόφασης του μοντέλου Ising μεταφέρει στιγμιότυπα του προβλήματος απόφασης weighted max-cut σε στιγμιότυπα του προβλήματος απόφασης του μοντέλου Ising που δεν έχουν μαγνητικό πεδίο. Άρα το πρόβλημα απόφασης του μοντέλου Ising χωρίς μαγνητικό πεδίο είναι επίσης NP-Complete. Από την πρόταση 3.4 προκύπτει άμεσα ο μετασχηματισμός ενός μοντέλου Ising σε ένα ισοδύναμο μοντέλο Potts όπου η ενέργεια του ενός διαφέρει κατά σταθερά από το άλλο. Αυτό σημαίνει ότι το πρόβλημα απόφασης του μοντέλου Ising ανάγεται στο πρόβλημα απόφασης του μοντέλου Potts. Τέλος το γεγονός ότι για κατάσταση σ μπορούμε σε πολυωνυμικό χρόνο να υπολογίσουμε $E(\sigma)$ και να συγκρίνουμε τα $E(\sigma), E_0$ σημαίνει ότι το πρόβλημα απόφασης του μοντέλου Potts είναι NP από το οποίο έπεται ότι τελικά είναι NP-Complete.

3.4.2 Μοντέλο Ising p-spin glasses

Σε αυτό το μοντέλο τα πράγματα είναι λίγο διαφορετικά. Ξεκινάμε πάλι με ένα σύνολο κόμβων V και ένα σύνολο p-άδων E του V, όπου $p \in \mathbb{N} \setminus \{0, 1\}$. Οι μαγνητικές αλληλεπιδράσεις είναι μια συνάρτηση $J : E \mapsto \mathbb{R}$. Σε κάθε κόμβο του V αντιστοιχούμε ένα spin με τιμή -1 ή 1. Η συνάρτηση ενέργειας του μοντέλου Ising p-spin glasses είναι

$$E(\sigma) = -\sum_{\{j_1, j_2, \ldots, j_p\} \in E} J_{j_1 j_2 \ldots j_p} \sigma_{j_1} \sigma_{j_2} \ldots \sigma_{j_p}$$

Το μοντέλο Ising p-spin glasses είναι η κατανομή Bolzmann που προκύπτει από την παραπάνω συνάρτηση ενέργειας στο σύνολο των καταστάσεων σ που περιγράφουν την αντιστοίχηση ενός κόμβου με ένα spin.

Αυτό το μοντέλο είναι άμεση γενίκευση του μοντέλου Ising. Για $p = 2$ είναι το μοντέλο Ising. Πάλι όπως και στην περίπτωση του μοντέλου Potts προκύπτει ότι είναι NP-Complete το αντίστοιχο πρόβλημα απόφασης με όμοιο τρόπο.

3.5 Προσεγγιστικοί Αλγόριθμοι

Σε αυτή την ενότητα θα μελετήσουμε αλγορίθμους τύπου Monte Carlo Markov Chain για να προσεγγίσουμε την κατανομή $\mathbb{B}(\mathbb{X}, E, \beta)$ όταν μας είναι δύσκολο να υπολογίσουμε άμεσα τη συνάρτηση επιμερισμού. Οι αλγόριθμοι που θα συζητήσουμε στηρίζονται στο εργοδικό θεώρημα μαρκοβιανών αλυσίδων.

3.5.1 Αλγόριθμος Metropolis-Hastings

Έστω ένα μοντέλο Ising με γράφημα $G = (V, E)$, μαγνητικές αλληλεπιδράσεις J και μαγνητικό πεδίο B. Η ιδέα του αλγορίθμου είναι να κατασκευάσουμε μια μαρκοβιανή αλυσίδα στον χώρο των καταστάσεων $\mathbb{X} = \{-1, 1\}^V$ η οποία να έχει αναλλοίωτη κατανομή την κατανομή $\mathbb{B}(\mathbb{X}, E, \beta)$, όπου E η συνάρτηση ενέργειας του μοντέλου Ising. Κατόπιν για ένα προκαθορισμένο αριθμό βημάτων N κινούμαστε τυχαία στην μαρκοβιανή αλυσίδα συλλέγοντας δείγματα και έτσι υπολογίζουμε την αναλλοίωτη κατανομή της ή την μέση τιμή κάποιας συνάρτησης ως προς την αναλλοίωτη κατανομή.

Algorithm 1 Αλγόριθμος Metropolis-Hastings

Require: G, J, B, β
  Επιλέγουμε $\sigma \sim U(\{-1, 1\}^V)$
  Υπολογίζουμε $E(\sigma)$
  $N_x \leftarrow 0 \quad x \in V \setminus \{\sigma\}$
  $N_\sigma \leftarrow 1$
  for $m \in \mathbb{Z}_N$ **do**
    Επιλέγουμε $j \sim U(V)$
    $\tau \leftarrow \sigma$
    $\tau_j \leftarrow -\tau_j$
    Υπολογίζουμε $E(\tau)$
    $\Delta E \leftarrow E(\tau) - E(\sigma)$
    Επιλέγουμε $u \sim U(0, 1)$
    if $u \leq e^{-\beta \Delta E}$ **then**
      $\sigma \leftarrow \tau$
    end if
    $N_\sigma \leftarrow N_\sigma + 1$
  end for
  return $\frac{1}{N} N_x \quad x \in V$

Η μαρκοβιανή αλυσίδα που προκύπτει από τον παραπάνω αλγόριθμο έχει πιθανότητα μετάβασης

$$w(\sigma \to \tau) = \frac{1}{|V|} \sum_{j \in V} e^{-\beta \Delta E_j(\sigma)} \delta(\tau, \sigma^{(j)}) + (1 - \frac{1}{|V|} \sum_{j \in V} e^{-\beta \Delta E_j(\sigma)}) \delta(\tau, \sigma)$$

όπου $\sigma^{(j)}$ η κατάσταση που προκύπτει από την αλλαγή του spin j στην σ και $\Delta E_j(\sigma) = \max(E(\sigma^{(j)}) - E(\sigma), 0)$.

Εάν η αλυσίδα είναι μη υποβιβάσιμη τότε γνωρίζουμε ότι $\frac{N_\sigma}{N}$ συγκλίνει σε αναλλοίωτη κατανομή από το εργοδικό θεώρημα μαρκοβιανών αλυσίδων. Εάν επιπλέον είναι απεριοδική και η κατανομή Bolzmann του μοντέλου Ising είναι αναλλοίωτη κατανομή της αλυσίδας, τότε θα έχουμε $\frac{N_\sigma}{N} \to p_\beta(\sigma)$

Αρχικά παρατηρούμε ότι όλες οι καταστάσεις της αλυσίδας είναι προσβάσιμες από μια αρχική κατάσταση σ διότι η πιθανότητα αλλαγής ενός spin είναι πάντα θετική. Άρα η αλυσίδα είναι μη υποβιβάσιμη. Θα δείξουμε ότι η κατανομή Bolzmann είναι αναλλοίωτη της αλυσίδας.

$$\sum_{\tau \in \{-1,1\}^V} w(\tau \to \sigma) p_\beta(\tau) = \sum_{j \in V} w(\sigma^{(j)} \to \sigma) p_\beta(\sigma^{(j)}) + w(\sigma \to \sigma) p_\beta(\sigma) =$$

$$\frac{1}{Z}\Big(\frac{1}{|V|}\sum_{j\in V} e^{-\beta \max(E(\sigma)-E(\sigma^{(j)}),0)} e^{-\beta E(\sigma^{(j)})} + [1 - \frac{1}{|V|}\sum_{j\in V} e^{-\beta \max(E(\sigma^{(j)})-E(\sigma),0)}] e^{-\beta E(\sigma)}\Big) =$$

$$\frac{1}{Z}\Big(\frac{1}{|V|}\sum_{j\in V} e^{-\beta \max(E(\sigma),E(\sigma^{(j)}))} - \frac{1}{|V|}\sum_{j\in V} e^{-\beta \max(E(\sigma^{(j)}),E(\sigma))} + e^{-\beta E(\sigma)}\Big) = p_\beta(\sigma)$$

Όμως η αλυσίδα αυτή δεν είναι πάντα απεριοδική και αυτό φαίνεται στο απλό παράδειγμα ενός μοντέλου Ising με 1 spin χωρίς μαγνητικό πεδίο. Σε αυτό το μοντέλο έχουμε $E(-1) = E(1) = 0$ και άρα $w(-1 \to 1) = w(1 \to -1) = 1$, $w(1 \to 1) = w(-1 \to -1) = 0$. Συνεπώς η αλυσίδα που προκύπτει σε αυτό το μοντέλο Ising έχει περίοδο 2. Μπορούμε να διορθώσουμε αυτή την ατέλεια του αλγόριθμου. Εάν στον κυρίως βρόγχο του αλγορίθμου με πιθανότητα $1 > \epsilon > 0$ δεν τροποποιούμε την κατάσταση σ και με πιθανότητα $1 - \epsilon$ εκτελούμε τα βήματά του, τότε προκύπτουν οι πιθανότητες μετάβασης

$$w'(\sigma \to \tau) = (1-\epsilon) w(\sigma \to \tau) + \epsilon \delta(\sigma, \tau)$$

Πάλι η αλυσίδα που προκύπτει είναι μη υποβιβάσιμη για τον ίδιο ακριβώς λόγο. Επίσης ισχύει

$$\sum_{\tau \in \{-1,1\}^V} w'(\tau \to \sigma) p_\beta(\tau) = (1-\epsilon)[\sum_{\tau \in \{-1,1\}^V} w(\tau \to \sigma) p_\beta(\tau)] +$$

$$\epsilon[\sum_{\tau \in \{-1,1\}^V} \delta(\tau,\sigma) p_\beta(\tau)] = (1-\epsilon) p_\beta(\sigma) + \epsilon p_\beta(\sigma) = p_\beta(\sigma)$$

Άρα πάλι η κατανομή Bolzmann του μοντέλου Ising είναι αναλλοίωτη της αλυσίδας. Τέλος επειδή $w'(\sigma \to \sigma) \geq \epsilon > 0$ η αλυσίδα είναι απεριοδική.

3.5.2 Αλγόριθμος Heat Bath

Ο αλγόριθμος Heat Bath αποτελεί παραλλαγή του αλγορίθμου Metropolis-Hastings.

Algorithm 2 Αλγόριθμος Heat Bath

Require: G, J, B, β
  Επιλέγουμε $\sigma \sim U(\{-1,1\}^V)$
  Υπολογίζουμε $E(\sigma)$
  $N_x \leftarrow 0 \quad x \in V \setminus \{\sigma\}$
  $N_\sigma \leftarrow 1$
  for $m \in \mathbb{Z}_N$ **do**
    Επιλέγουμε $j \sim U(V)$
    $\tau \leftarrow \sigma$
    $\tau_j \leftarrow -\tau_j$
    Υπολογίζουμε $E(\tau)$
    $\Delta E \leftarrow E(\tau) - E(\sigma)$
    Επιλέγουμε $u \sim U(0,1)$
    if $u \leq \frac{1}{2}[1 - \tanh(\frac{\beta\Delta E}{2})]$ **then**
      $\sigma \leftarrow \tau$
    end if
    $N_\sigma \leftarrow N_\sigma + 1$
  end for
  return $\frac{1}{N} N_x \quad x \in V$

Οι πιθανότητες μετάβασης είναι

$$w(\sigma \to \tau) = \frac{1}{|V|}\sum_{j\in V}\frac{1}{2}[1-\tanh(\frac{\beta\Delta E_j(\sigma)}{2})]\delta(\tau,\sigma^{(j)})+[1-\frac{1}{|V|}\sum_{j\in V}\frac{1}{2}[1-\tanh(\frac{\beta\Delta E_(\sigma)}{2})]]\delta(\tau,\sigma)$$

όπου $\sigma^{(j)}$ η κατάσταση που προκύπτει από την αλλαγή του spin j στην σ και $\Delta E_j(\sigma) = E(\sigma^{(j)}) - E(\sigma)$.

Αρχικά παρατηρούμε ότι

$$\frac{1}{2}(1-\tanh x) = \frac{1}{2}(1 - \frac{e^x - e^{-x}}{e^x + e^{-x}}) = \frac{e^{-x}}{e^x + e^{-x}}$$

Η αλυσίδα είναι μη υποβιβάσιμη εφόσον η πιθανότητα αλλαγή ενός spin είναι θετική για κάθε κατάσταση. Παρατηρούμε ότι

$$w(\sigma \to \sigma) = 1 - \frac{1}{|V|}\sum_{j\in V}\frac{e^{-\beta\Delta E_j}}{e^{\beta\Delta E_j} + e^{-\beta\Delta E_j}} > 1 - \frac{1}{|V|}\sum_{j\in V} 1 = 0$$

άρα η αλυσίδα είναι απεριοδική. Τέλος έχουμε

$$\sum_{\tau\in\{-1,1\}^V} w(\tau \to \sigma)p_\beta(\tau) = \sum_{j\in V} w(\sigma^{(j)} \to \sigma)p_\beta(\sigma^{(j)}) + w(\sigma \to \sigma)p_\beta(\sigma) =$$

$$\frac{1}{Z}[\frac{1}{|V|}\sum_{j\in V}\frac{e^{-\beta\Delta E_j(\sigma^{(j)})}}{e^{\beta\Delta E_j(\sigma^{(j)})}+e^{-\beta\Delta E_j(\sigma^{(j)})}}e^{-\beta E(\sigma^{(j)})}+$$

$$(1-\frac{1}{|V|}\sum_{j\in V}\frac{e^{-\beta\Delta E_j(\sigma)}}{e^{\beta\Delta E_j(\sigma)}+e^{-\beta\Delta E_j(\sigma)}})e^{-\beta E(\sigma)}]$$

Όμως ισχύει $\Delta E_j(\sigma^{(j)}) = -\Delta E_j(\sigma)$ και συνεπώς έχουμε

$$\sum_{\tau\in\{-1,1\}^V} w(\tau\to\sigma)p_\beta(\tau) =$$

$$\frac{1}{Z}[\frac{1}{|V|}\sum_{j\in V}\frac{e^{-\frac{\beta}{2}(E(\sigma)+E(\sigma^{(j)}))}}{e^{\beta\Delta E_j(\sigma)}+e^{-\beta\Delta E_j(\sigma)}}+$$

$$e^{-\beta E(\sigma)}-\frac{1}{|V|}\sum_{j\in V}\frac{e^{-\frac{\beta}{2}(E(\sigma)+E(\sigma^{(j)}))}}{e^{\beta\Delta E_j(\sigma)}+e^{-\beta\Delta E_j(\sigma)}}] = p_\beta(\sigma)$$

Άρα η κατανομή Bolzmann του μοντέλου Ising είναι η μοναδική αναλλοίωτη της αλυσίδας στην οποία θα συγκλίνει ο λόγος $\frac{N_\sigma}{N}$

3.5.3 Προσομοιωμένη ανόπτηση (simulated annealing)

Εφόσον με αλγορίθμους Monte Carlo Markov Chain μπορούμε να προσομοιώσουμε την κατανομή Bolzmann ενός μοντέλου Ising προκύπτει το ερώτημα κατά πόσο είναι εφικτό να εκτιμήσουμε την ελάχιστη δυνατή ενέργειά του. Σαν ερώτημα αυτό έχει ιδιαίτερο ενδιαφέρον διότι στο επόμενο κεφάλαιο θα ανάγουμε σημαντικά προβλήματα της συνδυαστικής σε μοντέλα Ising. Μια αφελής προσέγγιση θα ήταν να χρησιμοποιούσαμε αυτούς τους αλγορίθμους με ένα μεγάλο β, βασιζόμενοι στο γεγονός ότι οι πιο ενεργειακά χαμηλές καταστάσεις θα έχουν πολύ μεγαλύτερη πιθανότητα. Δεν γνωρίζουμε ωστόσο πόσο μεγάλο β χρειαζόμαστε. Ούτε γνωρίζουμε τον χρόνο που χρειαζόμαστε για να βρούμε μια τέτοια κατάσταση. Το σοβαρότερο πρόβλημα είναι ότι όσο μεγαλώνουμε το β αυξάνεται εκθετικά ο χρόνος που μένει ο αλγόριθμος σε μια κατάσταση τοπικού ελαχίστου.

Η προσομοιωμένη ανόπτηση έχει ως εξής. Αρχικά επιλέγουμε ένα πρόγραμμα ψύξης $(\beta_j, n_j) \in (0,\infty)\times\mathbb{N}^*$ όπου $j\in\mathbb{N}^*$, $1\le j\le N$, $N\in\mathbb{N}^*$. Για n_1 βήματα με αντίστροφη θερμοκρασία β_1 τρέχουμε τον αλγόριθμο Metropolis-Hastings ή τον Heat Bath. Ο αλγόριθμος επιστρέφει μια προσέγγιση της κατανομής Bolzmann. Έπειτα τρέχουμε πάλι τον ίδιο αλγόριθμο για n_2 βήματα με αντίστροφη θερμοκρασία β_2 επιλέγοντας όμως την αρχική κατάσταση σ στο πρώτο βήμα όχι από την ομοιόμορφη κατανομή αλλά από την προσεγγιστική κατανομή που μόλις αποκτήσαμε. Επαναλαμβάνουμε την ίδια διαδικασία για όλα τα ζεύγη (β_j, n_j) μέχρι να εξαντλήσουμε το πρόγραμμα ψύξης. Έτσι μπορούμε με μεγάλη πιθανότητα να βρεθούμε σε μια κατάσταση με ενέργεια η οποία να μην απέχει περισσότερο από $\epsilon > 0$ από την ελάχιστη δυνατή, χωρίς τον κίνδυνο εγκλωβισμού σε τοπικά ελάχιστα.

4 Αναγωγές σε μοντέλα στατιστικής φυσικής

Σε αυτό το κεφάλαιο θα κατασκευάσουμε κατανομές Bolzmann για στιγμιότυπα συνδυαστικών προβλημάτων ώστε οι αποδοτικότερες ενεργειακά καταστάσεις τους να είναι αυτές που αναζητούμε. Η πρόταση 3.3 είναι μια τέτοια αναγωγή πολυωνυμικού χρόνου η οποία χρησιμοποιήθηκε για να κατατάξουμε το πρόβλημα απόφασης του μοντέλου Ising στην κλάση πολυπλοκότητας NP-Complete. Θέλοντας να μελετήσουμε την συνδυαστική με μεθόδους στατιστικής φυσικής είναι σαφές ότι τέτοιοι μετασχηματισμοί παίζουν πιο κομβικό ρόλο στη μελέτη μας από τη κατάταξη σε κλάσεις πολυπλοκότητας. Για αυτό και θα αναπτύξουμε αρκετές αναγωγές στις επόμενες ενότητες για προβλήματα όλων των ειδών (βελτιστοποίησης, υπολογιστικά, απόφασης).

Φυσικά όπως είναι γνωστό μπορούμε πολυωνυμικά να ανάγουμε οποιοδήποτε πρόβλημα NP σε ένα οποιοδήποτε NP-Complete. Έτσι μπορούμε οποιοδήποτε NP πρόβλημα να το μετατρέψουμε σε ένα ισοδύναμο μοντέλο Ising. Ωστόσο η επίλυση ακόμα και απλών μοντέλων Ising είναι αρκετά περίπλοκη όπως διαπιστώσαμε. Συνεπώς θέλουμε να κατασκευάσουμε την απλούστερη δυνατή αναγωγή σε ισοδύναμο μοντέλο και όχι μια σύνθετη που δεν θα μπορούμε να μελετήσουμε στα πλαίσια της στατιστικής φυσικής.

4.1 Number Partitioning

Πρόβλημα 4.1 (Number Partitioning). *Για δεδομένη είσοδο $(a_j)_{j=1}^n \subseteq \mathbb{N}^*$ να βρεθεί*

$$U = argmin_{A\subseteq R_n} |\sum_{j\in A} a_j - \sum_{j\in R_n\setminus A} a_j|$$

Έστω ακολουθία $(a_j)_{j=1}^n \subseteq \mathbb{N}^*$ ένα στιγμιότυπο Number Partitioning. Θέτουμε για $A \subseteq R_n$

$$C_A = |\sum_{j\in A} a_j - \sum_{j\in R_n\setminus A} a_j|$$

Κάθε υποσύνολο A του R_n αντιστοιχείται αμφιμονοσήμαντα στην δείκτρια συνάρτηση $\mathbb{I}_A : R_n \mapsto \{0,1\}$. Εάν θεωρήσουμε $\sigma^{(A)} : R_n \mapsto \{-1,1\}$ με τύπο $\sigma_j^{(A)} = 2\mathbb{I}_A(j) - 1$ παρατηρούμε ότι έχουμε κατασκευάσει μια ιδιαίτερη δείκτρια η οποία σε όλα τα στοιχεία του A κάνει 1 και στα υπόλοιπα στοιχεία -1. Έτσι τα στοιχεία του R_n λαμβάνουν ένα spin ανάλογα με το εάν ανήκουν ή όχι στο A. Για αυτό τον λόγο θα λέμε την $\sigma^{(A)}$ δείκτρια spin του A. Έπεται ότι

$$C_A = |\sum_{j\in A} a_j - \sum_{j\in R_n\setminus A} a_j| = |\sum_{j\in R_n} a_j \sigma_j^{(A)}|$$

Οπότε μπορούμε να ορίσουμε το κόστος μιας ανάθεσης spin στα $j \in R_n$

$$C(\sigma) = |\sum_{j\in R_n} a_j \sigma_j|$$

Ένα ισοδύναμο μοντέλο για το πρόβλημα Number Partitioning , το οποίο θα μελετήσουμε στο επόμενο κεφάλαιο, προκύπτει από την κατανομή $\mathbb{B}(\{-1,1\}^{R_n}, C, \beta)$, όπου οι αποδοτικότερες ενεργειακά καταστάσεις είναι αυτές που αναζητούμε.

Φυσικά δεν υπάρχει μόνο ένα μοντέλο με αυτή την ιδιότητα. Αν θεωρήσουμε το τετράγωνο του $C(\sigma)$ ως ενέργεια έχουμε

$$E(\sigma) = (\sum_j a_j\sigma_j)^2 = \sum_j a_j^2\sigma_j^2 + 2\sum_{j<k} a_j a_k \sigma_j \sigma_k =$$

$$\sum_j a_j^2 + 2\sum_{j<k} a_j a_k \sigma_j \sigma_k$$

Από το λήμμα 2.2 η παραπάνω συνάρτηση ενέργειας είναι ισοδύναμη με την

$$E(\sigma) = 2\sum_{j<k} a_j a_k \sigma_j \sigma_k$$

Εκτός από ισοδύναμο μοντέλο για το πρόβλημα Number Partitioning η παραπάνω συνάρτηση ενέργειας αντιστοιχεί και σε μοντέλο Ising στο γράφημα C_n με μαγνητικές αλληλεπιδράσεις $J_{jk} = -2a_j a_k$ χωρίς μαγνητικό πεδίο. Οπότε μπορούμε να χρησιμοποιήσουμε τους προσεγγιστικούς αλγορίθμους της ενότητας 3.5 για να εκτιμήσουμε την λύση του Number Partitioning.

4.2 Vertex Cover

Για γράφημα $G = (V, E)$ ορίζουμε

$$E_v = \{\{v, u\} : \{v, u\} \in E\}$$

τις ακμές του v.

Πρόβλημα 4.2 (Vertex Cover). *Για γράφημα $G = (V, E)$ να βρεθεί το ελαχιστικό $A \subseteq V$ τέτοιο ώστε*

$$\bigcup_{v \in A} E_v = E$$

Η έννοια του κόστους μιας λύσης σε αυτό το πρόβλημα δεν είναι τόσο άμεση όσο προηγουμένως. Το κόστος ενός $A \subseteq V$ εξαρτάται κυρίως από πλήθος των ακάλυπτων ακμών και το πλήθος των κόμβων του δευτερευόντως. Αρχικά θα τα ορίσουμε ως ξεχωριστά κόστη και έπειτα θα τα συνδυάσουμε σε μια συνάρτηση κόστους. Όμοια με πριν θα αντικαταστήσουμε το σύνολο A με την δείκτρια spin του, οπότε θα ορίσουμε τα κόστη απευθείας για αναθέσεις spin. Συνεπώς έχουμε

$$P(\sigma) = \sum_{\{j,k\} \in E} \frac{1 - \sigma_j}{2} \frac{1 - \sigma_k}{2}$$

το πλήθος των ακάλυπτων ακμών και

$$Q(\sigma) = \sum_{j \in V} \frac{1 + \sigma_j}{2}$$

το πλήθος των κόμβων που έχουν επιλεγεί. Ορίζουμε

$$C(\sigma) = P(\sigma) + \frac{1}{|V|} Q(\sigma)$$

Επειδή $\frac{1}{|V|} Q(\sigma) \leq 1$ είναι σαφές όλες οι καταστάσεις με κόστος μεγαλύτερο του 1 αντιστοιχούν σε επιλογές κόμβων που δεν καλύπτουν πλήρως τις ακμές. Άρα η εύρεση ελαχιστικού A ανάγεται στην εύρεση ανάθεσης spin στο μοντέλο $\mathbb{B}(\{-1, 1\}^V, C, \beta)$ με ελάχιστη ενέργεια $E(\sigma) = \frac{K}{|V|}$, όπου K ο ελάχιστος αριθμός κόμβων που απαιτούνται για την κάλυψη των ακμών. Τέλος παρατηρούμε ότι

$$C(\sigma) = \sum_{\{j,k\} \in E} \frac{1 - \sigma_j}{2} \frac{1 - \sigma_k}{2} + \frac{1}{|V|} \sum_{j \in V} \frac{1 + \sigma_j}{2} =$$

$$\frac{|E|}{4} - \frac{1}{4} \sum_{\{j,k\} \in E} \sigma_j - \frac{1}{4} \sum_{\{j,k\} \in E} \sigma_k +$$

$$\frac{1}{4} \sum_{\{j,k\} \in E} \sigma_j \sigma_k + \frac{1}{2} + \frac{1}{2|V|} \sum_{j \in V} \sigma_j =$$

$$\frac{|E|}{4}+\frac{1}{2}-\frac{1}{2}\sum_{j\in V}|E_j|\sigma_j+$$

$$\frac{1}{2|V|}\sum_{j\in V}\sigma_j+\frac{1}{4}\sum_{\{j,k\}\in E}\sigma_j\sigma_k$$

Από το λήμμα 2.2 είναι ισοδύναμη η ενέργεια

$$E(\sigma)=\sum_{j\in V}(\frac{1}{2|V|}-\frac{|E_j|}{2})\sigma_j+\sum_{\{j,k\}\in E}\frac{1}{4}\sigma_j\sigma_k$$

Η οποία αντιστοιχεί σε μοντέλο Ising στο γράφημα $G=(V,E)$ με μαγνητικές αλληλεπιδράσεις $J_{jk}=-\frac{1}{4}$ και μαγνητικό πεδίο $B_j=\frac{|E_j|}{2}-\frac{1}{2|V|}$. Άρα αρκεί να υπολογίσουμε μια κατάσταση ελάχιστης ενέργειας του παραπάνω μοντέλου Ising για να βρούμε ένα ελαχιστικό $A\subseteq V$ που να καλύπτει τις ακμές.

4.3 k-SAT

Πρόβλημα 4.3 (Satisfiability). *Μπορεί η λογική πρόταση $A(x_1, x_2, ..., x_n)$ των λογικών μεταβλητών $(x_j)_{j=1}^n$ να είναι αληθής για κάποια ανάθεσή τους.*

Το πρόβλημα της ικανοποιησημότητας (Satisfiability) είναι το πιο θεμελιώδες NP-Complete πρόβλημα. Αυτό αποτυπώνεται στην απόδειξη του θεωρήματος Cook-Levin που μετατρέπεται μια μη ντετερμινιστική μηχανή Turing που επιλύει ένα NP πρόβλημα σε μια λογική παράσταση ικανοποιήσημη αν και μόνο αν το πρόβλημα έχει καταφατική απάντηση. Αυτό κάνει την ανάλυση ενός τέτοιου προβλήματος ιδιαίτερα ελκυστική.

Είναι γνωστό από την λογική ότι μπορούμε να μετατρέψουμε μια οποιαδήποτε λογική πρόταση σε μια της μορφής **k-CNF** (Conjunctive Normal Form) η οποία είναι ικανοποιήσιμη αν και μόνο αν η αρχική είναι. Αυτό μας οδηγεί στο παρακάτω εξίσου στοιχειώδες NP-Complete πρόβλημα.

Πρόβλημα 4.4 (k-SAT). *Μπορεί η λογική πρόταση μορφής k-CNF $A = \bigwedge_{j=1}^m C_j$ με **λογικούς περιορισμούς** $C_j = \bigvee_{l=1}^k z_l^{(j)}$, όπου $z_l^{(j)}$ τα **κατηγορήματα** είναι κάποια από τις μεταβλητές $(x_j)_{j=1}^n$ ή οι αρνήσεις τους, να είναι ικανοποιήσιμη.*

Για την λογική πρόταση της μορφής k-CNF $A = \bigwedge_{j=1}^m C_j$ με μεταβλητές $(x_j)_{j=1}^n$ το πλήθος των ανικανοποίητων περιορισμών είναι η απλούστερη δυνατή έννοια ενέργειας.

$$E(\tau) = \sum_{j=1}^m \mathbb{I}(C_j(\tau) = 0) \quad \tau \in \{0,1\}^n$$

Όπως και πριν προκύπτει ένα σύστημα τύπου spin glasses εάν θεωρήσουμε $\sigma_j = 2\tau_j - 1$. Για τους περιορισμούς C_j που δεν είναι ταυτοτικοί, δηλαδή δεν εμφανίζεται κάποια μεταβλητή με την καταφατική και την αρνητική της μορφή στον περιορισμό, μπορούμε να ορίσουμε τα βάρη W_{jl} ως εξής. $W_{jl} = 1$ εάν εμφανίζεται η καταφατική μορφή της x_l στον C_j, $W_{jl} = -1$ εάν εμφανίζεται η αρνητική μορφή της x_l στον C_j και $W_{jl} = 0$ αλλιώς. Επίσης ορίζουμε $|C_j|$ το πλήθος των διαφορετικών μεταβλητών που εμφανίζονται με οποιαδήποτε μορφή στον C_j. Με τη βοήθεια των παραπάνω μπορούμε να εκφράσουμε την δείκτρια ως προς τα spins

$$\mathbb{I}(C_j(\tau) = 0) = \frac{1}{2^{|C_j|}} \prod_{l=1}^m (1 - W_{jl}\sigma_l)$$

Από αυτό προκύπτει η ενέργεια του συστήματος spin glasses

$$E(\sigma) = \sum_{l=1}^m \frac{1}{2^{|C_j|}} \prod_{j=1}^n (1 - W_{jl}\sigma_j)$$

Παρατηρούμε με απλές πράξεις ότι $\prod_{j=1}^n (1 - W_{jl}\sigma_j)$ περιέχει γινόμενα της μορφής $J_{j_1 j_2 ... j_p}\sigma_{j_1}\sigma_{j_2}...\sigma_{j_p}$ με $p \le k$. Συνεπώς το παραπάνω σύστημα είναι ένα μοντέλο Ising p-spin glasses.

4.4 Graph Coloring

Πρόβλημα 4.5 (Graph Coloring)**.** *Για γράφημα $G = (V, E)$ και $q \in \mathbb{N}^*$ χρώματα. Να αποφανθούμε εάν είναι εφικτός ο χρωματισμός των κόμβων του γραφήματος με αυτά τα χρώματα ώστε να μην υπάρχει ακμή μεταξύ κόμβων ίδιου χρώματος.*

Μπορούμε να παραστήσουμε κάθε κατάσταση σ ως μια απεικόνιση $V \mapsto R_q$. Το κόστος της κάθε κατάστασης είναι απλά το πλήθος των ακμών μεταξύ κόμβων του ίδιου χρώματος. Συνεπώς έχουμε

$$C(\sigma) = \sum_{\{j,k\} \in E} \delta(\sigma_j, \sigma_k)$$

το οποίο είναι η ενέργεια ενός μοντέλου Potts στο γράφημα $G = (V, E)$ με μαγνητικές αλληλεπιδράσεις $J_{jk} = -1$.

5 Συλλογές στιγμιοτύπων

Εκεί που είναι πραγματικά αξιοσημείωτοι οι μέθοδοι της στατιστικής φυσικής ε-ίναι στην μελέτη συλλογών στιγμιοτύπων. Μια συλλογή στιγμιοτύπων είναι απλά μια ακολουθία από τυχαία στιγμιότυπα ενός προβλήματος. Για παράδειγμα στο πρόβλημα Vertex Cover θα μπορούσαμε να θεωρήσουμε μια ακολουθία από τυχαία γραφήματα n κόμβων που παράγονται από την κλίκα C_n αφαιρώντας κάθε κόμβο με πιθανότητα p. Σε αυτό το κεφάλαιο θα χρησιμοποιήσουμε τεχνικές μελέτης αλλαγών φάσεων για να μελετήσουμε την ύπαρξη λύσης και την πολυπλοκότητα υπολογιστικών προβλημάτων.

5.1 Number Partitioning

Έστω $(a_j)_{j=1}^n \subseteq \mathbb{N}^*$. Η συνάρτηση ενέργειας στο $\{-1,1\}^{R_n}$

$$E(\sigma) = |\sum_{j=1}^{n} a_j \sigma_j|$$

περιγράφει κατανομή Bolzmann που οι αποδοτικότερες ενεργειακά καταστάσεις είναι οι ζητούμενες διαμερίσεις.

Η συνάρτηση επιμερισμού είναι

$$Z = \sum_{\sigma} e^{-\beta|\sum_{j=1}^{n} a_j\sigma_j|} = \sum_{\sigma}\int_{-\infty}^{\infty} e^{-|x|}\delta(x - \beta\sum_{j=1}^{n} a_j\sigma_j)dx$$

όπου $\delta(x)$ η συνάρτηση δέλτα του Dirac. Από την θεωρία των γενικευμένων συναρτήσεων και των μετασχηματισμών Fourier είναι γνωστό ότι

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)e^{iyx-iya}dxdy$$

Υπολογίζουμε

$$\int_{-\infty}^{\infty} e^{-|x|+ixy}dx = \int_{0}^{\infty} e^{-x(1+iy)}dx + \int_{0}^{\infty} e^{-x(1-iy)}dx = \frac{2}{1+y^2}$$

Συνεπώς έχουμε

$$Z = \sum_{\sigma}\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+y^2}e^{-i\beta y\sum_{j=1}^{n} a_j\sigma_j}dy = \frac{2^n}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+y^2}\prod_{j=1}^{n}\cos(\beta a_j y)dy$$

Ορίζουμε $p, s : \mathbb{R} \mapsto \mathbb{R}$

$$p(y) = \prod_{j=1}^{n}\cos(\beta a_j y)$$

$$s(y) = sgn(p(y))$$

Ισχύει

$$p(y) = s(y)|p(y)| = s(y)\prod_{j=1}^{n}|\cos(\beta a_j y)| =$$

$$s(y)\exp(\sum_{j=1}^{n}\log|\cos(\beta a_j y)|) = s(y)\exp(n[\frac{1}{n}\sum_{j=1}^{n}\log|\cos(\beta a_j y)|])$$

Εάν τα $(a_j)_{j=1}^{n}$ επιλέγονται ανεξάρτητα και τυχαία από την ομοιόμορφη κατανομή $U(R_{2^b})$, όπου $b \in \mathbb{N}^*$. Από τον νόμο των μεγάλων αριθμών ισχύει

$$\frac{1}{n}\sum_{j=1}^{n}\log|\cos(\beta a_j y)| \to \frac{1}{2^b}\sum_{\kappa=1}^{2^b}\log|\cos(\beta\kappa y)|$$

Το οποίο μας υποδεικνύει να αντικαταστήσουμε το $\frac{1}{n}\sum_{j=1}^{n}\log|\cos(\beta a_j y)|$ με $\frac{1}{2^b}\sum_{\kappa=1}^{2^b}\log|\cos(\beta\kappa y)|$. Επίσης θα αντικαταστήσουμε $s(y)$ με $sgn(\prod_{\kappa=1}^{2^b}\cos(\beta\kappa y))$ για να εκτιμήσουμε ευρετικά την συνάρτηση επιμερισμού. Ορίζουμε

$$r(y) = sgn(\prod_{\kappa=1}^{2^b}\cos(\beta\kappa y))$$

$$g(y) = \frac{1}{2^b}\sum_{\kappa=1}^{2^b}\log|\cos(\beta\kappa y)|$$

Οπότε προκύπτει η εκτίμηση της συνάρτησης επιμερισμού

$$Z = \frac{2^n}{\pi}\int_{-\infty}^{\infty}\frac{r(y)}{1+y^2}\exp(ng(y))dy$$

Παρατηρούμε ότι

$$g(y) = \frac{1}{2^b}\sum_{\kappa=1}^{2^b}\log|\cos(\beta\kappa y)| \leq \frac{1}{2^b}\sum_{\kappa=1}^{2^b}\log 1 = 0$$

Η ισότητα ισχύει αν και μόνο αν υπάρχει $m \in \mathbb{Z}$ τέτοιο ώστε $\beta y = m\pi$. Στα σημεία $y = 2m\frac{\pi}{\beta}$, όπου $m \in \mathbb{Z}$, έχουμε $\cos(\beta\kappa y) = \cos(2m\kappa\pi) = 1$ ενώ για $y = (2m+1)\frac{\pi}{\beta}$, όπου $m \in \mathbb{Z}$, ισχύει $\cos(\beta\kappa y) = \cos((2m+1)\kappa\pi) = (-1)^\kappa$. Από αυτό συμπεραίνουμε ότι

$$r(2m\frac{\pi}{\beta}) = 1$$

και

$$r((2m+1)\frac{\pi}{\beta}) = \prod_{\kappa=1}^{2^b}(-1)^\kappa = 1$$

Παρατηρούμε ότι οποιοδήποτε τοπικό μέγιστο y_0 που δεν είναι ολικό μέγιστο είναι $g(y_0) < 0$, όμως επειδή g περιοδική λαμβάνουμε $g(y_0) \leq a < 0$. Θέτουμε

$$y_m = m\frac{\pi}{\beta}$$

Μπορούμε να θεωρήσουμε περιοχές των y_m λόγω τις περιοδικότητας και της συνέχειας της g

$$V_m = (y_m - \delta, y_m + \delta)$$

όπου $\delta > 0$, $r(y) = 1$ και $g(y) > a$. Επίσης θέτουμε

$$A = \mathbb{R} \setminus \bigcup_{m\in\mathbb{Z}} V_m$$

Οπότε έχουμε

$$Z = \frac{2^n}{\pi}(\sum_{m\in\mathbb{Z}} \int_{V_m} \frac{1}{1+y^2}\exp(ng(y))dy + \int_A \frac{r(y)}{1+y^2}\exp(ng(y))dy)$$

αλλά ισχύει

$$|\int_A \frac{r(y)}{1+y^2}\exp(ng(y))dy| \leq \int_A \frac{1}{1+y^2}\exp(na)dy$$

$$< \int_{-\infty}^{\infty} \frac{1}{1+y^2}\exp(na)dy = \pi\exp(na)$$

Από αυτό και την εκτίμηση σαγματικού σημείου λαμβάνουμε

$$Z = \frac{2^n}{\pi}(\sum_{m\in\mathbb{Z}} \frac{1}{1+(m\frac{\pi}{\beta})^2}\sqrt{\frac{2\pi}{-ng''(m\frac{\pi}{\beta})}}(1+O(\frac{1}{n})) + O(e^{na}))$$

Υπολογίζουμε

$$g''(y) = -\frac{1}{2^b}\beta^2 \sum_{kappa=1}^{2^b} \kappa^2(1+tan(\beta\kappa y)^2)$$

και συνεπώς

$$g''(y_m) = -\beta^2\mathbb{E}(a_1^2)$$

Συνεπώς έχουμε

$$Z = 2^n(\sqrt{\frac{2}{\pi n\beta^2\mathbb{E}(a_1^2)}}\sum_{m\in\mathbb{Z}} \frac{1}{1+(m\frac{\pi}{\beta})^2}(1+O(\frac{1}{n})) + O(e^{na}))$$

Από την

$$\sum_{m\in\mathbb{Z}} \frac{1}{1+(xm)^2} = \frac{\pi}{x}\coth\frac{\pi}{x}$$

έχουμε

$$Z = 2^n(\sqrt{\frac{2}{\pi n \mathbb{E}(a_1^2)}}\coth(\beta)(1+O(\frac{1}{n})) + O(e^{na}))$$

Από το οποίο εκτιμάμε ότι η ελεύθερη ενέργεια είναι

$$F = -\frac{1}{\beta}(n\log 2 + \frac{1}{2}\log\frac{2}{\pi n \mathbb{E}(a_1^2)} + \log\coth\beta)$$

Από την οποία λαμβάνουμε

$$U = \frac{1}{\sinh(\beta)\cosh(\beta)}$$

Παρατηρούμε ότι $\lim_{\beta\to\infty} U = 0$ το οποίο από την πρόταση 2.6 σημαίνει ότι οι τέλειες διαμερίσεις είναι οι καταστάσεις ελάχιστης ενέργειας. Επίσης από την πρόταση 2.6 έπεται ότι το πλήθος των τέλειων διαμερίσεων είναι $e^{\lim_{\beta\to\infty} S}$. Άρα υπολογίζουμε από την πρόταση 2.5

$$S = \beta(U-F) = \beta\frac{1}{\sinh\beta\cosh\beta} + \log\coth\beta + n\log 2 + \frac{1}{2}\log\frac{2}{\pi n\mathbb{E}(a_1^2)}$$

και συνεπώς

$$S_\infty = \lim_{\beta\to\infty} S = n\log 2 + \frac{1}{2}\log\frac{2}{\pi n\mathbb{E}(a_1^2)} = n\log 2 + \frac{1}{2}\log\frac{2}{\pi n} - \frac{1}{2}\log\mathbb{E}(a_1^2) =$$

$$n\log 2(1 + \frac{1}{2n\log 2}\log\frac{2}{\pi n} - \frac{1}{2n\log 2}\log\mathbb{E}(a_1^2))$$

Υπολογίζουμε

$$\mathbb{E}(a_1^2) = \frac{1}{2^b}\sum_{\kappa=1}^{2^b}\kappa^2 = \frac{(2^b+1)(2^{b+1}+1)}{6}$$

και άρα

$$\log\mathbb{E}(a_1^2) = \log(\frac{1}{3}2^{2b} + \frac{1}{2}2^b + \frac{1}{6}) = -\log 3 + 2b\log 2$$

$$+\log(1 + \frac{3}{2}2^{-b} + \frac{1}{2}2^{-2b}) = -\log 3 + 2b\log 2 + O(2^{-b})$$

Άρα έχουμε

$$S_\infty = n\log 2(1 + \frac{1}{2n}\log_2\frac{6}{\pi n} - \frac{b}{n} + O(2^{-b}))$$

Άρα όταν έχουμε b αρκετά μεγάλο και

$$\frac{b}{n} < 1 + \frac{1}{2n}\log_2\frac{6}{\pi n}$$

τότε η εκθετική τάξη του πλήθους των τέλειων διαμερίσεων είναι θετική. Άρα αναμένουμε η απάντηση στο πρόβλημα απόφασης Number Partitioning να είναι καταφατική και μάλιστα να είναι εύκολη η εύρεση της λύσης, λόγο του σχετικά μεγάλου πλήθους των διαθέσιμων λύσεων.

Στην περίπτωση όμως που

$$\frac{b}{n} > 1 + \frac{1}{2n}\log_2\frac{6}{\pi n}$$

αναμένουμε να είναι αρνητική η απάντηση στο πρόβλημα απόφασης Number Partitioning και η εύρεση λύσης είναι πολύ δύσκολη καθώς αν υπάρχει κατάσταση που είναι τέλεια διαμέριση έχει μικρή πιθανότητα να βρεθεί με τους γνωστούς αλγορίθμους που στην χειρότερη περίπτωση κάνουν αναζήτηση σε σχεδόν όλο των χώρο των 2^n καταστάσεων.

Σε αυτό το σημείο οφείλουμε να σημειώσουμε ότι η περίπτωση

$$\frac{b}{n} > 1 + \frac{1}{2n}\log_2\frac{6}{\pi n}$$

είναι προβληματική σαν εκτίμηση διότι αφενός έχουμε χρησιμοποιήσει τον νόμο των μεγάλων αριθμών για n ανεξάρτητα δείγματα από την κατανομή $U(R_{2^b})$ και αφετέρου η προσέγγιση που έχουμε δίνει αρνητική εντροπία σε αυτή την περίπτωση. Εντούτοις το συμπέρασμά της δεν είναι εσφαλμένο και συμφωνεί με την διαίσθηση και εμπειρικά στοιχεία.

Σημείωση Εάν θεωρούσαμε το μοντέλο Ising της παραγράφου 4.1 τότε θα είχαμε

$$Z = \sum_{\sigma}\exp(-\beta(\sum_{j=1}^{n}\sigma_j a_j)^2)$$

όμως γνωρίζουμε για $a \in \mathbb{C}$

$$\sqrt{\pi} = \int_{-\infty}^{\infty}\exp(-(x-a)^2)dx \implies$$

$$\exp(a^2) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2+2xa)dx$$

εάν

$$a = i\sqrt{\beta}\sum_{j=1}^{n}\sigma_j a_j$$

έχουμε

$$\exp(-\beta(\sum_{j=1}^{n}\sigma_j a_j)^2) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2+2xi\sqrt{\beta}\sum_{j=1}^{n}\sigma_j a_j)dx$$

Έπεται ότι

$$Z = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2)\sum_{\sigma}\exp(2xi\sqrt{\beta}\sum_{j=1}^{n}\sigma_j a_j)dx =$$

$$\frac{2^n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2)\prod_{j=1}^{n}\cos(2\sqrt{\beta}a_j x)dx$$

Το οποίο μπορούμε να το εκτιμήσουμε με ακριβώς ανάλογο τρόπο.

5.2 k-SAT

Παρόλο που μπορούμε να αναγάγουμε σε μοντέλο Ising k-spin glasses το k-SAT δεν θα ασχοληθούμε με αυτό το μοντέλο σε αυτή την ενότητα. Θα αναλύσουμε δομικά την πιθανότητα να είναι ικανοποιήσιμη η λογική πρόταση της μορφής k-CNF A με λογικές μεταβλητές $(x_j)_{j=1}^n$ και περιορισμούς $(C_j)_{j=1}^m$, όπου $m = \lfloor an \rfloor$ για $a \in (0, \infty)$. Θα θεωρήσουμε ότι παράγονται ανεξάρτητα οι λογικοί περιορισμοί C_j επιλέγοντας k κατηγορήματα ανεξάρτητα και ομοιόμορφα από τις λογικές μεταβλητές x_j ή τις αρνήσεις τους. Τέλος συμβολίζουμε $P_n(k, a)$ την πιθανότητα μια τέτοια λογική πρόταση να είναι ικανοποιήσιμη.

Διαισθητικά είναι φανερό ότι για a πολύ μεγάλο η πρόταση είναι σχεδόν απίθανο να είναι ικανοποιήσιμη διότι έχουμε πολλούς περιορισμούς. Αντίστοιχα για a αρκετά μικρό μπορούμε εύκολα να βρούμε ένα συνδυασμό τιμών για τις λογικές μεταβλητές που να ικανοποιεί τους λίγους περιορισμούς. Αρκεί να εκτιμήσουμε το άνω όριο που όταν το a το υπερβαίνει μηδενίζεται η πιθανότητα ικανοποιησιμότητας και το κάτω όριο που για a μικρότερα από αυτό είναι σχεδόν βέβαια ικανοποιήσιμη η τυχαία λογική πρόταση. Για να το πετύχουμε αυτό θα χρησιμοποιήσουμε τις μεθόδους των ροπών πρώτης και δευτέρας τάξεως.

Θεώρημα 5.1 (Εκτίμηση άνω ορίου).

$$a > -\frac{\log 2}{\log(1 - 2^{-k})} \implies \lim_n P_n(k, a) = 0$$

Απόδειξη. Έστω A λογική τυχαία λογική πρόταση όπως παραπάνω.

Παρατηρούμε ότι

$$\mathbb{I}(\exists \tau : A(\tau) = 1) \leq N(A)$$

όπου $N(A)$ το πλήθος των αναθέσεων των x_j ώστε η A να είναι ικανοποιήσιμη. Συνεπώς έχουμε

$$P_n(k, a) = \mathbb{P}(\exists \tau : A(\tau) = 1) \leq \mathbb{E}[N(A)]$$

Όμως ισχύει

$$N(A) = \sum_{\tau \in \{0,1\}^n} \mathbb{I}(A(\tau) = 1) = \sum_{\tau \in \{0,1\}^n} \prod_{j=1}^m \mathbb{I}(C_j(\tau) = 1)$$

Και λόγο τις ανεξαρτησίας των λογικών περιορισμών C_j έπεται ότι

$$\mathbb{E}[N(A)] = \sum_{\tau \in \{0,1\}^n} \mathbb{P}(A(\tau) = 1) = \sum_{\tau \in \{0,1\}^n} \prod_{j=1}^m \mathbb{P}(C_j(\tau) = 1)$$

Για να μην ικανοποιείται ο περιορισμός C_j πρέπει όλα τα κατηγορήματά του να μην ικανοποιούνται από την ανάθεση τ. Το κάθε κατηγόρημα έχει πιθανότητα $\frac{1}{2}$ να μην ικανοποιείται από μια ανάθεση και εφόσον είναι ανεξάρτητα τα κατηγορήματα μεταξύ τους προκύπτει ότι

$$\mathbb{P}(C_j(\tau) = 0) = 2^{-k}$$

Από αυτό υπολογίζουμε

$$\mathbb{E}[N(A)] = \sum_{\tau\in\{0,1\}^n}\prod_{j=1}^{m}\mathbb{P}(C_j(\tau)=1) =$$

$$\sum_{\tau\in\{0,1\}^n}\prod_{j=1}^{m}(1-2^{-k}) = 2^n(1-2^{-k})^m =$$

$$\exp(n\log 2 + m\log(1-2^{-k})) = \exp[n(\log 2 + \frac{m}{n}\log(1-2^{-k}))]$$

Επειδή $m = \lfloor an \rfloor \implies \frac{m}{n} = a + O(\frac{1}{n})$ τελικά έχουμε

$$\mathbb{E}[N(A)] = \exp[n(\log 2 + a\log(1-2^{-k}) + O(\frac{1}{n}))]$$

Συνεπώς για $a > -\frac{\log 2}{\log(1-2^{-k})}$ μπορούμε να βρούμε n αρκετά μεγάλο ώστε ο εκθέτης $\log 2 + \frac{m}{n}\log(1-2^{-k})$ να είναι αρνητικός. Από το οποίο προκύπτει

$$\lim_n \mathbb{E}[N(A)] = 0$$

και με βάση την ανισότητα που αποδείξαμε

$$\lim_n P_n(k,a) = 0$$

□

Η μέθοδος των ροπών δευτέρας τάξεως στηρίζεται στο ακόλουθο λήμμα.

Λήμμα 5.2. *Για μη αρνητική τυχαία μεταβλητή X που $\mathbb{P}(X=0)<1$ ισχύει*

$$\mathbb{P}(X>0) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}$$

Απόδειξη. Για μη αρνητική τυχαία μεταβλητή X έχουμε

$$X = \mathbb{I}(X>0)X \implies$$

$$\mathbb{E}(X)^2 = \mathbb{E}(\mathbb{I}(X>0)X)^2$$

Από την ανισότητα Cauchy-Schwarz λαμβάνουμε

$$\mathbb{E}(X)^2 \leq \mathbb{P}(X>0)\mathbb{E}(X^2)$$

Όμως επειδή $\mathbb{P}(X=0)<1 \implies \mathbb{P}(X>0)>0$ και άρα $\mathbb{E}(X^2)>0$ προκύπτει το ζητούμενο. □

Θεώρημα 5.3 (Εκτίμηση κάτω ορίου). *Για $k \geq 3$ ισχύει*

$$a < [2^k - 2(k+1)]\log 2 - 1 \implies \lim_n P(k,a) = 1$$

Απόδειξη. Αναζητούμε μια συνάρτηση από το σύνολο των λογικών προτάσεων στους μη αρνητικούς πραγματικούς αριθμούς η οποία να κάνει μηδέν αν και μόνο αν το όρισμα της είναι μια μη ικανοποιήσιμη λογική πρόταση. Δηλαδή

$$\mathbb{P}(F(A) > 0) = \mathbb{P}(\exists \tau : A(\tau) = 1)$$

Θεωρούμε A τυχαία λογική πρόταση όπως παραπάνω και μια συνάρτηση $\phi : \mathbb{N} \mapsto [0, \infty)$ τέτοια ώστε $\phi(x) = 0 \iff x = 0$. Ορίζουμε

$$W(A) = \sum_{\tau} \prod_{j=1}^{m} \phi(r(\tau, C_j))$$

Όπου $r(\tau, C)$ το πλήθος από τα k κατηγορήματα του C που ικανοποιούνται από την ανάθεση τ. Λόγω τις ανεξαρτησίας των λογικών περιορισμών C_j έχουμε

$$\mathbb{E}[W(A)] = \sum_{\tau} \mathbb{E}[\prod_{j=1}^{m} \phi(r(\tau, C_j)) =$$

$$\sum_{\tau} \prod_{j=1}^{m} \mathbb{E}[\phi(r(\tau, C_j))] = 2^n (\mathbb{E}[\phi(r(\tau, C))])^m$$

Για το $W^2(A)$ έχουμε

$$W^2(A) = [\sum_{\tau} \prod_{j=1}^{m} \phi(r(\tau, C_j))]^2 =$$

$$(\sum_{\tau} \prod_{j=1}^{m} \phi(r(\tau, C_j))(\sum_{\sigma} \prod_{j=1}^{m} \phi(r(\sigma, C_j)) = \sum_{\tau} \sum_{\sigma} \prod_{j=1}^{m} \phi(r(\tau, C_j))\phi(r(\sigma, C_j))$$

Από το οποίο έπεται ότι

$$\mathbb{E}[W^2(A)] = \mathbb{E}[\sum_{\tau} \sum_{\sigma} \prod_{j=1}^{m} \phi(r(\tau, C_j))\phi(r(\sigma, C_j)) =$$

$$\sum_{\tau} \sum_{\sigma} (\mathbb{E}[\phi(r(\tau, C_j))\phi(r(\sigma, C_j))])^m$$

Έστω αναθέσεις τ, σ που ταυτίζονται σε l σημεία και διαφέρουν στα υπόλοιπα $n-l$. Επειδή τα κατηγορήματα ενός λογικού περιορισμού C επιλέγονται ομοιόμορφα και ανεξάρτητα η κατανομή του $(r(\tau, C), r(\sigma, C))$ δεν εξαρτάται από ποιες επιλογές τ, σ θα κάνουμε αλλά μόνο από το πλήθος των σημείων l που ταυτίζονται. Συνεπώς μπορούμε να θέσουμε $g_n(l) = \mathbb{E}[\phi(r(\tau, C_j))\phi(r(\sigma, C_j))]$ για μεταθέσεις τ, σ όπως παραπάνω. Έτσι ομαδοποιώντας τα ζεύγη τ, σ που ταυτίζονται σε ακριβώς l σημεία προκύπτει

$$\mathbb{E}[W^2(A)] = 2^n \sum_{l=0}^{n} \binom{n}{l} (g_n(l))^m$$

Η πιθανότητα να ικανοποιείται ένα κατηγόρημα της C από την ανάθεση τ είναι $\frac{1}{2}$. Συνεπώς το $r(\sigma, C)$ είναι το πλήθος των επιτυχιών k ανεξάρτητων δοκιμών Bernoulli με πιθανότητα $\frac{1}{2}$ και άρα $r(\sigma, C) \sim B(k, \frac{1}{2})$. Θα θεωρήσουμε $\phi(s) = \mathbb{I}(s > 0)\lambda^s$, όπου $\lambda \in (0, \infty)$. Συνεπώς προκύπτει

$$\mathbb{E}[\phi(r(\tau, C))] = \sum_{s=1} 2^{-k}\binom{k}{s}\lambda^s = \frac{(1+\lambda)^k - 1}{2^k} \implies$$

$$\mathbb{E}[W(A)] = \exp(n\log 2 + m\log[(1+\lambda)^k - 1] - mk\log 2) \implies$$

$$\mathbb{E}[W(A)] = \exp(np)$$

όπου $p = \log 2 - ak\log 2 + a\log[(1+\lambda)^k - 1] + O(\frac{1}{n})$. Για να υπολογίσουμε το $g_n(l)$ θα θεωρήσουμε τα $u, v \in \{0,1\}^k$ με $u_s = 1$ εάν ικανοποιείται το s κατηγόρημα της C από την ανάθεση τ αλλιώς $u_s = 0$ και αντίστοιχα ορίζουμε το v για την ανάθεση σ. Για κάθε s που $u_s = v_s$ παρατηρούμε ότι αντιστοιχεί στην επιλογή μιας μεταβλητής σε κατηγόρημα της C που έχουν αναθέσει την ίδια τιμή οι τ, σ. Αντίστοιχα για κάθε s που $u_s \neq v_s$ βλέπουμε ότι αντιστοιχεί στην επιλογή μιας μεταβλητής σε κατηγόρημα της C που διαφέρουν οι αναθέσεις των τ, σ. Επίσης θα συμβολίζουμε για $x \in \{0,1\}^k$ $|x| = |\{j : x_j = 1\}|$. Άρα για τ, σ που ταυτίζονται για l μεταβλητές προκύπτει

$$g_n(l) = \mathbb{E}[\phi(r(\tau, C_j))\phi(r(\sigma, C_j))] =$$

$$\sum_{u,v\in\{0,1\}^k} \phi(|u|)\phi(|v|)2^{-k}\prod_{s=1}^{k}(\frac{l}{n})^{\mathbb{I}(u_s=v_s)}(1-\frac{l}{n})^{\mathbb{I}(u_s\neq v_s)}$$

εφόσον η πιθανότητα επιλογής ενός κατηγορήματος από τις l μεταβλητές που ταυτίζονται οι τ, σ είναι $\frac{l}{n}$. Έστω $d : \{0,1\}^k \times \{0,1\}^k \mapsto \mathbb{N}$ η απόσταση Hamming δηλαδή το πλήθος των σημείων που u, v δεν ταυτίζονται τότε έχουμε

$$g_n(l) = 2^{-k}\sum_{u,v\in\{0,1\}^k}\phi(|u|)\phi(|v|)(\frac{l}{n})^{k-d(u,v)}(1-\frac{l}{n})^{d(u,v)}$$

Από το λήμμα 1.10 προκύπτει

$$\mathbb{E}[W^2(A)] = 2^n\sum_{l=0}^{n}\binom{n}{l}(g_n(l))^m =$$

$$2^n\sum_{l=0}^{n}\exp(O(\log n) + nh(\frac{l}{n}))(g_n(l))^m =$$

$$\sum_{l=0}^{n}\exp(n[O(\frac{\log n}{n}) + \log 2 + h(\frac{l}{n}) + a\log g_n(l)])$$

Θέτουμε $f : [0,1] \mapsto \mathbb{R}$

$$f(z) = \log 2 + h(z) + a\log[2^{-k}\sum_{u,v\in\{0,1\}^k}\phi(|u|)\phi(|v|)z^{k-d(u,v)}(1-z)^{d(u,v)}]$$

Παρατηρούμε ότι

$$f(\frac{1}{2}) = \log 2 + h(\frac{1}{2}) + a\log[2^{-2k}\sum_{u,v\in\{0,1\}^k}\phi(|u|)\phi(|v|)] =$$

$$2\log 2 - 2ak\log 2 + 2a\log[\sum_{u\in\{0,1\}^k}\phi(|u|)] =$$

$$2\log 2 - 2ak\log 2 + 2a\log[(1+\lambda)^k - 1] = 2p \implies$$

$$\mathbb{E}[W(A)]^2 = \exp(nf(\frac{1}{2}))$$

Εάν εκτιμήσουμε με τη μέθοδο του σαγματικού σημείου την ροπή δεύτερης τάξης του $W(A)$

$$\mathbb{E}[W^2(A)] = \exp(n\max_{z\in[0,1]} f(z))(1+o(1))$$

Άρα για να έχουμε το ζητούμενο αρκεί στο $\frac{1}{2}$ να έχει μοναδικό μέγιστο η f. Το οποίο συνεπάγεται

$$\frac{1}{1-\lambda} = (1+\lambda)^{k-1}$$

Λύνοντας την παραπάνω εξίσωση ως προς λ λαμβάνουμε την κατάλληλη συνάρτηση ϕ που μας δίνει το ζητούμενο για το δοσμένο k. □

Τα παραπάνω αποτελέσματα καθώς και το γεγονός ότι το Number Partitioning έχει σημείο αλλαγής φάσης μας υποβάλλουν την ιδέα ότι και το k-SAT έχει ένα τέτοιο σημείο. Έτσι κάνουμε την παρακάτω εικασία.

Εικασία 5.4. *Για κάθε $k \geq 2$ υπάρχει σημείο αλλαγής φάσης για το πρόβλημα k-SAT $a_c(k)$ για το οποίο ισχύει*

$$a > a_c(k) \implies \lim_n P_n(k,a) = 0$$

και

$$a < a_c(k) \implies \lim_n P_n(k,a) = 1$$

Ωστόσο το μόνο που γνωρίζουμε μέχρι τώρα είναι ότι το σημείο αλλαγής φάσης του k-SAT εάν υπάρχει είναι $a_c(k) = \Theta(2^k)$. Αξίζει να σημειωθεί ότι για $k = 2$ μπορούμε να χρησιμοποιήσουμε την εύκολη δομή του προβλήματος για να υπολογίσουμε ότι $a_c(2) = 1$. Γνωρίζουμε ότι $(x \vee y) \iff (\neg x \implies y) \iff (\neg y \implies x)$ το οποίο μας επιτρέπει να παραστήσουμε το πρόβλημα ως ένα γράφημα με κόμβους τα κατηγορήματα και ακμές τους λογικούς περιορισμούς. Αρκεί λοιπόν στο κατευθυντό γράφημα που προκύπτει να μην ανήκουν τα κατηγορήματα μιας μεταβλητής σε έναν κύκλο για να είναι ικανοποιήσιμη η λογική πρόταση. Έτσι εκτός από την εύκολη αλγοριθμική επίλυση που επιδέχεται αυτή η δομή μπορούμε πιθανοθεωρήτικα να αναλύσουμε την περίπτωση να μην υπάρχει κύκλος για $a < 1$. Αυτό μαζί με το πρώτο θεώρημα που αποδείξαμε που για $k = 2$ μας δίνει άνω όριο επίσης το 1 και προκύπτει $a_c(2) = 1$.

Βιβλιογραφία

[1] Marc Mézard, Andrea Montanari. *Information, Physics, and Computation.* Oxford University Press, 2009.

[2] Andrew Lucas. *Ising formulations of many NP problems.* arXiv:1302.5843v3, 2014

[3] Stephan Mertens. *A physicist's approach to number partitioning.* Theoretical Computer Science volume 265 pages 79–108, 2001

[4] Dimitris Achlioptas, Yuval Peres *The Threshold for Random k-SAT is* $2^k \log 2 - O(k)$. arXiv:cs/0305009v2, 2003