

Cristalls de *spin*

DAVID MÁRQUEZ-CARRERAS I CARLES ROVIRA

Resum

L'estudi dels cristalls de *spin* iniciat a partir dels anys setanta ha tingut darre-rament grans avenços en termes matemàtics. En aquest article presentem, de manera divulgativa, els conceptes bàsics d'aquesta àrea, detallem el seu marc matemàtic, descrivim els models clàssics que considerem més importants i expliquem algunes de les tècniques que s'utilitzen per a estudiar-los.

Paraules clau: cristalls de *spin*, mesura de Gibbs

Classificació AMS: 82B44.

1 Introducció

Els *cristalls de spin* van ser originàriament desenvolupats per físics amb l'objectiu de modelitzar estats d'agregació de la matèria on es volia estudiar la relació entre l'estructura i el comportament magnètic, aplicant-se, per exemple, a certs materials ferromagnètics o conductors. En general, aquests sistemes, inclosos els més senzills, són força complexos ja que el comportament de les interaccions entre els àtoms és molt complicat.

Aquests models van ser tractats inicialment per la física teòrica dins del que es coneix com la mecànica estadística de sistemes desordenats (podeu veure, per exemple, el llibre de Mezard, Parisi i Vilasoro [10] o l'article dels dos darrers [11]). L'estudi com a objectes purament matemàtics és més recent i té un gran creixement a partir de mitjan anys noranta. Apareixen així una sèrie de problemes matemàtics dins del camp de les probabilitats, generalment difícils

de resoldre. Més concretament, aquests problemes es situen en la teoria de processos estocàstics amb dimensió gran.

Algunes de les àrees on s'apliquen aquests models són: les xarxes neuronals, l'optimització, la intel·ligència artificial o la computació.

Començarem amb un petit exemple clàssic que podeu trobar al llibre tot just acabat d'esmentar [10]. Aquest exemple no està lligat a problemes de magnetisme, però ens permet mostrar quins tipus de models apareixen i introduir alhora dos conceptes importants d'aquests models físics: la frustració i el desordre.

Considerem que tenim N individus que volem classificar en dos grups segons la seva afinitat. Suposem que entre cada dos individus hi ha un sentiment de simpatia o d'antipatia i que aquest sentiment és recíproc. Difícilment podrem separar els individus en dos grups de manera que en cada grup tots es caiguin simpàtics. Aquesta situació la podem formalitzar de la manera següent: a cada parella d'individus i i j , hi associem un nombre $J_{i,j}$ que val 1 si són amics i -1 si no ho són (vaja, si no es cauen bé). A més, a cada individu i , li assignem una variable s_i que pren els valors 1 o -1 segons en quin dels dos grups situem aquest individu. Aleshores, donat el conjunt de relacions $J := \{J_{i,j}, 1 \leq i < j \leq N\}$, l'objectiu és minimitzar el que seria la quantitat de desencís generat per la classificació $s := \{s_i, 1 \leq i \leq N\}$ en dos grups, és a dir, minimitzar

$$H(J, s) = - \sum_{1 \leq i < j \leq N} J_{i,j} s_i s_j.$$

Observem que com més gran és el valor de $H(J, s)$, hi ha més parelles d'individus amb sentiment d'antipatia mutu que estan en el mateix grup, i parelles que es cauen bé en grups diferents; hi ha, per tant, un grau més gran de frustració. Aquesta frustració genera inestabilitat ja que els individus voldran classificar-se d'una altra manera. Es parla de frustració, ja que hi haurà individus que no es sentiran còmodes on els hem classificat.

El càlcul dels s_i que minimitzen el desencís és un problema matemàtic difícil i molt complex. Una manera alternativa de plantejar aquesta situació, és assignar a cada classificació s una certa probabilitat

$$G_N(s) = \frac{\exp(-\beta H(J, s))}{Z_N},$$

on Z_N és la funció de partició o normalitzadora

$$Z_N = \sum_s \exp(-\beta H(J, s))$$

i β un nou paràmetre del model. Aleshores, el problema de determinar la classificació òptima es converteix en un problema equivalent que consisteix a estudiar la mesura de probabilitat G_N . A la mecànica estadística, $H(J, s)$ s'interpreta com l'energia associada a una configuració, a G_N l'anomenen la mesura de Gibbs i β és l'invers de la temperatura.

Ens queda encara l'altre concepte important, el desordre. En aquest cas, l'anomenat desordre ve del fet que el conjunt de relacions J s'escullen a l'atzar segons una determinada llei de probabilitat. Si apliquem aquest model als problemes de magnetisme ens trobem amb materials magnètics que poden ser ferromagnètics ($J_{i,j} > 0$, tendeixen a alinear els *spins*) o antiferromagnètic ($J_{i,j} < 0$).

Resumint, tenim un sistema físic que es pot trobar en un nombre finit de configuracions diferents. Per a cada configuració, el sistema té associada una determinada energia, que depèn de la temperatura del sistema. Llavors, ens preguntem quina és la probabilitat de trobar-nos en una configuració concreta després de deixar el sistema a una temperatura determinada durant prou temps. El desordre (donat pel caràcter aleatori de les J) no depèn de la temperatura. Per això, es parla de desordre congelat —*frozen disorder*.

Destaquem també que d'aquests sistemes de partícules n'hi ha de dos tipus: els models reals i els models en mitjana —*mean-fields*. Els primers són els que més es donen a la naturalesa i és on el model té en compte la localització geomètrica de les partícules (per exemple, els àtoms haurien d'interaccionar bàsicament amb els que tenen al seu voltant); mentre que en els segons s'oblida aquesta localització perquè s'accepta que tots els àtoms interaccionen d'alguna manera entre ells. El recull de models que donarem en aquest treball pertanyen als models en mitjana, ja que els models reals són, avui en dia, força difícils de tractar.

En aquesta nota volem fer una presentació divulgativa, però rigurosa, de l'estudi dels models anomenats *cristalls de spin*. El nostre coneixement d'aquests models l'hem obtingut fonamentalment a partir dels treballs de M. Talagrand. És, per tant, aquesta aproximació als models de *cristalls de spin*, la que presentem aquí. El treball de Talagrand té, a més, una característica excitant: tot i la seva complexitat, les eines que utilitza són senzilles, de manera que no es requereixen coneixements més enllà de la teoria bàsica de la probabilitat i algunes propietats de les lleis gaussianes.

Hi ha altres matemàtics i físics importants que utilitzen altres mètodes per a tractar aquests models. Podem citar entre d'altres: Aizenman, Ben Arous, Bovier, Comets, Ghirlanda, Guerra, Guionnet, Lebowitz, Neveu, Parisi, Pastur, Shcherbina, Stein i Tonelli.

Finalment, l'estructura d'aquest treball és la següent: en la secció 2 formalitzem el marc matemàtic presentat a la introducció. La secció 3 està dedicada a presentar diferents models de *cristalls de spin*, mentre que en la secció 4 es mostren quatre dels aspectes de la mesura de Gibbs que s'estudien. Per acabar, en la secció 5, demostrem un petit resultat amb l'objectiu de presentar dues de les metodologies bàsiques per a estudiar aquests models: el mètode de la cavitat —*cavity method*— i el mètode del camí bo —*smart path method*.

2 Marc matemàtic

Els *crystals de spin* són sistemes de N partícules on cadascuna pot prendre dos valors: $+1$ i -1 . El conjunt dels possibles valors de les N partícules s'anomena espai de configuracions i el designarem per $\Sigma_N = \{-1, 1\}^N$. Així, una configuració $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$ és un element de Σ_N i cada component d'aquesta configuració és un *spin* ($\sigma_i = \pm 1$). En general podem considerar que N serà sempre un valor gran.

Cada configuració $\boldsymbol{\sigma}$ té associada una energia que anomenarem el hamiltonià del sistema i designarem per $H_N(\boldsymbol{\sigma})$. És a dir, H_N és una funció real definida a Σ_N . Serà el tipus de funció d'energia el que determinarà el tipus de model. Com ja hem comentat, a causa del desordre que volem modelitzar, l'energia serà una variable aleatòria. A partir d'aquesta estructura aleatòria es reflectirà la interacció existent entre les diferents configuracions.

Assignarem a cada configuració un pes proporcional al seu factor de Boltzmann $\exp(-\beta H_N(\boldsymbol{\sigma}))$, on β és un paràmetre que representa l'invers de la temperatura. Això ens permetrà definir una probabilitat sobre Σ_N , anomenada *mesura de Gibbs*,

$$G_N(\boldsymbol{\sigma}) = \frac{1}{Z_N} \exp(-\beta H_N(\boldsymbol{\sigma})),$$

on Z_N és el factor de normalització

$$Z_N = Z_N(\beta) = \sum_{\boldsymbol{\sigma} \in \Sigma_N} \exp(-\beta H_N(\boldsymbol{\sigma})).$$

La mesura de Gibbs és un concepte que concerneix la mecànica estadística i físicament la podem interpretar com la probabilitat de trobar-nos en una configuració $\boldsymbol{\sigma}$ quan s'ha arribat a l'equilibri després d'un bany tèrmic prou llarg de temperatura $T = \frac{1}{\beta}$. Parlarem només de models a temperatura alta, és a dir, quan β és petit. L'estudi d'aquests models per β gran acostuma a ser força més complicat.

Aquesta mesura de Gibbs és dues vegades aleatòria, és a dir, tenim dos nivells d'aleatorietat i podem parlar, per tant, de dos tipus d'esperança o mitjana: respecte a l'espai de configuracions i respecte a l'aleatorietat associada al hamiltonià. Per a una funció $f: \Sigma_N \rightarrow \mathbb{R}$ designarem per $\langle \cdot \rangle$ l'esperança respecte a la mesura de Gibbs G_N , i. e.

$$\langle f \rangle = \frac{1}{Z_N} \sum_{\boldsymbol{\sigma} \in \Sigma_N} f(\boldsymbol{\sigma}) \exp(-\beta H_N(\boldsymbol{\sigma})) = \int_{\Sigma_N} f dG_N, \quad (1)$$

i després escriurem

$$\nu(f) = E\langle f \rangle,$$

on E denota l'esperança respecte a l'aleatorietat de l'energia. La fórmula (1) pot generalitzar-se de manera natural per funcions $f: \Sigma_N^n \rightarrow \mathbb{R}$ com

$$\langle f \rangle = \frac{1}{Z_N^n} \sum_{(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in \Sigma_N^n} f(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \exp\left(-\sum_{1 \leq l \leq n} \beta H_N(\boldsymbol{\sigma}^l)\right).$$

3 Diferents models

En aquest apartat presentarem breument diversos models de *cristalls de spin*, començant pels models més senzills.

3.1 El model REM

El model d'*energia aleatòria* de Derrida, que denotarem amb l'acrònim REM, de l'anglès Random Energy Model, va ser introduït per Derrida el 1980 —vegeu, per exemple, els articles [4] i [5]. En aquest cas, la família $\{H_N(\boldsymbol{\sigma})\}_{\boldsymbol{\sigma} \in \Sigma_N}$ està formada per variables aleatòries gaussianes centrades i independents amb variància $\frac{N}{2}$, és a dir, $H_N(\boldsymbol{\sigma}) \sim \mathcal{N}(0, \frac{N}{2})$. Aquest model no té gaire interès real, ja que la no-correlació entre les configuracions indica una situació sense interaccions entre les partícules. No obstant això, resulta molt útil matemàticament perquè, gràcies a la seva simplicitat, permet un estudi complet del seu comportament. A més, ens ajuda a entendre altres models força més complicats.

3.2 El model Sherrington-Kirkpatrick (SK)

El hamiltonià que defineix aquest model és:

$$-H_N(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{i,j} \sigma_i \sigma_j,$$

on $\{g_{i,j}, 1 \leq i < j \leq N\}$ és una família de variables aleatòries independents idènticament distribuïdes amb llei $\mathcal{N}(0, 1)$. Podríem dir que aquest és el model original ja que fou el primer que s'estudià —vegeu l'article de Sherrington i Kirkpatrick [12].

A diferència del model REM —tot i que en aquest cas els hamiltonians també són variables aleatòries gaussianes centrades— si tenim dues configuracions diferents $\boldsymbol{\sigma}^1$ i $\boldsymbol{\sigma}^2$, les seves energies no són independents. Un petit càlcul ens dona

$$\begin{aligned} \mathbb{E} \left(H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) \right) &= \frac{1}{N} \sum_{1 \leq i < j \leq N} \sigma_i^1 \sigma_j^1 \sigma_i^2 \sigma_j^2 = \\ &= \frac{N}{2} \left(\frac{1}{N} \sum_{1 \leq i \leq N} \sigma_i^1 \sigma_i^2 \right)^2 - \frac{1}{2}. \end{aligned}$$

Ens apareix aquí un concepte fonamental, la *superposició (overlap)* entre dues configuracions $\boldsymbol{\sigma}^1$ i $\boldsymbol{\sigma}^2$ que es defineix com

$$R_{1,2} = \frac{1}{N} \sum_{1 \leq i \leq N} \sigma_i^1 \sigma_i^2.$$

Observem que com més semblants són les dues configuracions, més propera a 1 estarà la superposició. A més, es té

$$d(\boldsymbol{\sigma}, \boldsymbol{\sigma}') = \text{card} \{i \leq N : \sigma_i \neq \sigma'_i\} = \frac{N}{2} (1 - R_{1,2}),$$

on $d(\cdot, \cdot)$ és la distància de Hamming que mesura el nombre de coordenades en les quals dues configuracions difereixen.

3.3 El model d'interacció de p -spins amb camp extern

La generalització natural del model acabat de presentar seria un sistema on, d'una banda, la interacció tingués lloc entre p -spins i, d'altra banda, es tingués en compte la influència d'un camp extern. Tenim d'aquesta manera un hamiltonià del tipus següent:

$$-H_N(\boldsymbol{\sigma}) = u_N \sum_{(i_1, \dots, i_p) \in A_N^p} g_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p} + h \sum_{1 \leq i \leq N} \sigma_i,$$

on

$$u_N = \left(\frac{p!}{2N^{p-1}} \right)^{\frac{1}{2}},$$

$$A_N^p = \{(i_1, \dots, i_p) \in \mathbb{N}^p; 1 \leq i_1 < \dots < i_p \leq N\},$$

$h > 0$ i $\{g_{i_1, \dots, i_p}; (i_1, \dots, i_p) \in A_N^p\}$ és una família de variables aleatòries gaussianes independents amb llei $\mathcal{N}(0, 1)$. La part del hamiltonià $h \sum_{1 \leq i \leq N} \sigma_i$ representa el camp magnètic extern que influeix sobre els spins i que com $h > 0$ està agafat de manera que afavoreix els spins positius per damunt dels negatius.

Quan $p = 2$ i $h = 0$ (és a dir, sense camp extern) aquest model coincideix amb el de Sherrington-Kirkpatrick i quan $p \rightarrow \infty$ i $h = 0$ molts comportaments corresponen als del model REM. Per $p = 2$ i $h > 0$ tenim el model SK amb camp extern que serà el que utilitzarem en la propera secció.

El model d'interacció de p -spins sense camp extern fou introduït a la física per tal d'estudiar la distribució dels àtoms sotmesos a interaccions més complexes que les del cas SK —vegeu, per exemple, el treball de Gardner [7].

3.4 El model del perceptró

Considerem el hamiltonià següent:

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{1 \leq k \leq M} u \left(\frac{1}{\sqrt{N}} \sum_{1 \leq i \leq N} g_{i,k} \sigma_i \right),$$

on M és una petita proporció de N (és a dir, $M = \alpha N$, $\alpha > 0$), $u: \mathbb{R} \rightarrow \mathbb{R}$ és una funció prou regular i $\{g_{i,k}, 1 \leq i \leq N, 1 \leq k \leq M\}$ és una família de variables aleatòries independents idènticament distribuïdes amb llei $\mathcal{N}(0, 1)$.

Aquest model té nombroses aplicacions a les xarxes neuronals, a la computació i a la intel·ligència artificial. S'aplica especialment al que s'anomena aprenentatge supervisat de reconeixement de formes —*supervised learning for pattern recognition*—; vegeu per exemple [9]. Per a una presentació més general d'aquest model podeu llegir els articles [6] i [8].

4 Què s'estudia d'aquests models?

Aquesta secció està dedicada a presentar algunes de les maneres de tractar els sistemes de *cristalls de spin*. Com ja hem comentat a la introducció, la manera de fer-ho és estudiant la mesura de Gibbs. Aquest estudi el podem fer, però, des de diversos punts de vista. Aquí n'explicarem quatre.

Per fer-ho tot més comprensible, ho portarem a terme en el cas particular del model de Sherrington-Kirkpatrick amb camp extern que té per hamiltonià

$$-\beta H_N(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{i,j} \sigma_i \sigma_j + h \sum_{1 \leq i \leq N} \sigma_i, \quad (2)$$

on $h > 0$ és el camp extern que actua sobre el sistema tal com s'ha explicat als apartats 3.2 i 3.3. També adrecem el lector a aquestes seccions per a la resta de notació d'aquest hamiltonià.

Aquests estudis es faran tots a temperatura alta (β petita). Alguns dels aspectes que observarem poden tenir comportaments diferents (o encara desconeguts) a temperatura baixa o en altres models.

Finalment, hem d'indicar que el lector interessat pot trobar una bona mostra d'aquestes propietats en les dues referències següents: el capítol 2 del llibre [13], dedicat al model Sherrington-Kirkpatrick, i l'estudi del model *p-spin* fet a [2].

4.1 La superposició (*overlap*)

En la secció 3.2 hem definit la superposició entre dues configuracions $\boldsymbol{\sigma}^1$ i $\boldsymbol{\sigma}^2$ com

$$R_{1,2} = \frac{1}{N} \sum_{1 \leq i \leq N} \sigma_i^1 \sigma_i^2.$$

Hem vist també la relació entre $R_{1,2}$ i la correlació de la família de variables aleatòries $(H_N(\boldsymbol{\sigma}))_{\boldsymbol{\sigma} \in \Sigma_N}$. Com que són aquestes correlacions les que ens defineixen les propietats dels diferents models, una bona manera de conèixer un model és estudiant la llei de la superposició. Una de les primeres maneres d'estudiar la llei de la superposició és amb l'estudi dels seus moments i en concret de la seva convergència en L^{2k} , $k \geq 0$.

Sigui q l'única solució que pertany a $[0, 1]$ de l'equació

$$q = \mathbf{E} \tanh^2(\beta z \sqrt{q} + h), \quad (3)$$

on $z \sim \mathcal{N}(0, 1)$ és independent de la resta de variables. Observeu que si no hi ha camp extern, és a dir, si $h = 0$, es té $q = 0$.

Aleshores, es compleix que per β prou petit

$$\nu \left((R_{1,2} - q)^{2k} \right) = \mathbf{E} \left(\left\langle (R_{1,2} - q)^{2k} \right\rangle \right) \leq \left(\frac{Lk}{N} \right)^k,$$

on L és una constant que no depèn ni de k ni de N . A partir d'aquest resultat (teorema 2.5.1, [13]) s'obtenen també moments exponencials de $R_{1,2} - q$:

$$\nu \left(\exp \frac{N}{L} (R_{1,2} - q)^2 \right) \leq L.$$

Emprant aquestes desigualtats, es demostra l'anomenat teorema del límit central per la superposició: s'obté que la variable aleatòria $\sqrt{N}(\langle R_{1,2} \rangle - q)$ és asimptòticament gaussiana. Més concretament, es demostra que

$$\nu \left(\phi \left(\sqrt{N} (R_{1,2} - q) \right) \right) \xrightarrow{N \rightarrow \infty} \mathbf{E}(\phi(z)),$$

on ϕ és una funció contínua i afitada i $z \sim \mathcal{N}(0, \nu^2)$ amb variància ν^2 especificada.

Finalment, també es fan estudis més fins de les potències de la superposició, buscant els seus desenvolupaments de Taylor. Per il·lustrar-ho donarem a tall d'exemple un resultat que pot trobar-se a [1]: per a qualsevol $n \in \mathbb{N}$ i agafant $h = 0$, s'ha demostrat l'existència de constants $C(\beta, j)$ tals que

$$\nu \left(R_{1,2}^2 \right) = \sum_{j=1}^n \frac{C(\beta, j)}{N^j} + \mathcal{O} \left(\frac{1}{N^{n+1}} \right).$$

4.2 L'energia lliure

Un altre aspecte important per a conèixer la mesura de Gibbs és l'estudi del factor de normalització Z_N definit en la secció 2, anomenat també *funció de partició*. Amb l'objectiu d'estudiar les fluctuacions de la funció de partició, i seguint la denominació utilitzada pels físics, es defineix l'energia lliure associada al model com

$$p_N(\beta) := \frac{1}{N} \mathbf{E}(\log Z_N).$$

El resultat fonamental és el següent: per β prou petit i per $h > 0$ es compleix

$$\lim_{N \rightarrow \infty} p_N(\beta) = \frac{\beta^2}{4} (1 - q)^2 + \log 2 + \mathbf{E}(\log \cosh(\beta z \sqrt{q} + h)),$$

on $z \sim \mathcal{N}(0, 1)$ és independent de la resta de variables i q és l'única solució pertanyent a $[0, 1]$ de l'equació (3) definida a l'apartat anterior. Aquesta igualtat es coneix amb la terminologia anglesa de *replica-symmetric formula*. A més, per $h > 0$ es té

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbf{E}(Z_N) = \frac{\beta^2}{4} + \log 2 + \log \cosh h,$$

de manera que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbf{E}(\log Z_N) \neq \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbf{E}(Z_N).$$

En canvi quan $h = 0$ i per $\beta < 1$ tenim

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbf{E}(\log Z_N) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbf{E}(Z_N) = \frac{\beta^2}{4} + \log 2.$$

Obtenim així comportaments diferents per als casos $h = 0$ i $h > 0$. Mentre que en el primer cas, en escala logarítmica, la mitjana i la mediana de Z_N són la mateixa (asimptòticament), en el cas $h > 0$ ja no coincideixen.

Darrerament, Talagrand —vegeu [15]— ha descrit el comportament asimptòtic de l'energia lliure a qualsevol temperatura per a una àmplia classe de hamiltonians. Ha resolt així un dels problemes fonamentals d'aquests models.

4.3 La mesura de Gibbs com a mesura producte

Una altra manera d'estudiar la mesura de Gibbs és comparant-la amb una mesura producte, tant localment com global. En el model SK sense camp extern o amb un camp extern petit (h petit), la mesura de Gibbs té un comportament diferent en els casos local i global tal com tot seguit exposarem.

Fixat un nombre natural $n < N$, considerem el conjunt de *spins* $\{-1, +1\}^n$. Sigui $G_{N,n}$ la projecció de la mesura de Gibbs sobre aquest conjunt. Definim també una mesura producte $\mu_{h,n}$ sobre el mateix conjunt, tal que

$$\int_{\{-1,+1\}^n} x_i d\mu_{h,n}(\mathbf{x}) = \langle x_i \rangle, \quad \forall i \leq n.$$

És fàcil veure que quan no hi ha camp extern, és a dir, quan $h = 0$, la mesura producte $\mu_{h,n}$ és la mesura uniforme sobre el conjunt $\{-1, +1\}^n$.

D'una banda, per al cas local s'obté

$$\mathbf{E}(|G_{N,n} - \mu_{h,n}|^2) := \mathbf{E}\left(\left[\sum_{\mathbf{x} \in \{-1,+1\}^n} |G_{N,n}(\mathbf{x}) - \mu_{h,n}(\mathbf{x})|\right]^2\right) \leq \frac{K(n)}{N},$$

per una constant positiva $K(n)$. Això vol dir que la mesura de Gibbs vista localment es comportarà asimptòticament com una mesura producte. Més particularment, en el cas de mancança de camp extern aquesta mesura producte és la uniforme.

D'altra banda, quan $h = 0$ o h és prou petit, aquest comportament local de la mesura de Gibbs no es compleix de manera global. En efecte, es poden trobar dos subconjunts $A, B \subset \Sigma_N$ i $K > 0$ que satisfan

$$G_N(A) \geq \frac{2}{3}, \quad \mu_{h,N}(B) \geq \frac{2}{3} \quad \text{i} \quad d(A, B) \geq K,$$

per a tot $h \leq h_0$ (h_0 donat) i N suficientment gran (recordem que d és la distància de Hamming definida a l'apartat 3.2). És a dir, que per a N prou gran, la mesura producte i la mesura de Gibbs són molt diferents.

Aquests resultats es poden trobar a [13] —vegeu el teorema 2.4.10— per al cas local i a [3] per al cas global.

4.4 La magnetització

El comportament local de la mesura de Gibbs semblant al de la mesura producte $\mu_{h,n}$, fa que una manera d'estudiar el comportament de la mesura de Gibbs sigui entendre la família de variables aleatòries $\{\langle\sigma_i\rangle\}_{i\leq n}$.

S'obté el resultat següent: per a β prou petit i per a un nombre natural n fixat, podem trobar n variables aleatòries gaussianes estàndards independents z_1, \dots, z_n tals que

$$\mathbf{E} \left(\sum_{1 \leq i \leq n} [\langle\sigma_i\rangle - \tanh(\beta z_i \sqrt{q} + h)]^2 \right) \leq \frac{K(n, h)}{N},$$

on recordem que q l'hem definit a l'apartat 4.1 com l'única solució que pertany a $[0, 1]$ de l'equació (3).

Aquest resultat el podem interpretar de la manera següent: la mitjana d'un *spin* respecte a l'aleatorietat a l'espai de configuracions convergeix en llei cap a una variable aleatòria identificada, és a dir,

$$\langle\sigma_i\rangle \xrightarrow[N \rightarrow \infty]{} \tanh(\beta z_i \sqrt{q} + h).$$

A més a més, aquestes mitjanes $\langle\sigma_1\rangle, \dots, \langle\sigma_n\rangle$ són asimptòticament independents.

Aquests resultats es poden trobar a [14] —vegeu el teorema 5.14. Relacionades amb la magnetització, hi ha també les equacions de Thouless, Anderson i Palmer —conegudes com a equacions TAP.

5 Un exemple: el comportament asimptòtic de la *superposició*

En les seccions anteriors ens hem dedicat a fer una introducció general al món dels *spins*; ara, en canvi, el nostre objectiu és la demostració d'un resultat concret per a mostrar al lector quines són les tècniques i eines més habituals que s'empren. Presentarem així dos mètodes fonamentals i clàssics en la recerca sobre *spins*: el mètode de la cavitat (introduït per Talagrand) i el mètode del camí bo.

El resultat principal que provarem és el teorema 2.4.2 de [13], que dona la convergència quadràtica de la superposició cap a l'única solució de l'equació (3) en el model de SK amb camp extern. Els arguments usats en aquest apartat es poden comprendre amb coneixements bàsics de probabilitats.

1 TEOREMA Existeix β_0 tal que, per a tot $\beta \leq \beta_0$, tenim

$$v \left((R_{1,2} - q)^2 \right) = \mathbf{E} \left\langle (R_{1,2} - q)^2 \right\rangle \leq \frac{K}{N}, \quad (4)$$

on q és l'única solució de l'equació (3) i $K > 0$.

A continuació donarem tot un seguit de conceptes, nocions, notacions i resultats previs que necessitem per a abordar la prova d'aquest teorema.

El mètode de la cavitat consisteix en un argument per inducció sobre N que redueix el sistema inicial de N -spins a un nou sistema amb un *spin* menys (d'aquí la paraula cavitat). Per això, la idea clau és reagrupar en el hamiltonià (2) els termes que depenen de σ_N :

$$\begin{aligned} -H_{N,\beta}(\boldsymbol{\sigma}) := -\beta H_N(\boldsymbol{\sigma}) &= \left[\frac{\beta}{\sqrt{N}} \sum_{1 \leq i < j \leq N-1} g_{i,j} \sigma_i \sigma_j + h \sum_{1 \leq i \leq N-1} \sigma_i \right] + \\ &+ \sigma_N \left[\frac{\beta}{\sqrt{N}} \sum_{1 \leq i \leq N-1} g_{i,N} \sigma_i + h \right]. \end{aligned} \quad (5)$$

Si designem per $\boldsymbol{\rho} = (\sigma_1, \dots, \sigma_{N-1})$ un element de l'espai de configuracions Σ_{N-1} i reemplacem β per $\beta^- = \beta \sqrt{\frac{N-1}{N}}$, podem veure el primer sumand de (5) com un hamiltonià de $(N-1)$ -spins amb paràmetre β^- :

$$-H_{N-1,\beta^-}(\boldsymbol{\rho}) = \frac{\beta^-}{\sqrt{N-1}} \sum_{1 \leq i < j \leq N-1} g_{i,j} \sigma_i \sigma_j + h \sum_{1 \leq i \leq N-1} \sigma_i.$$

Si, a més, escrivim

$$g(\boldsymbol{\rho}) = \frac{\beta}{\sqrt{N}} \sum_{1 \leq i \leq N-1} g_{i,N} \sigma_i,$$

tindrem

$$-H_{N,\beta}(\boldsymbol{\sigma}) = -H_{N-1,\beta^-}(\boldsymbol{\rho}) + \sigma_N (g(\boldsymbol{\rho}) + h).$$

Per $\langle \cdot \rangle_-$ entendrem la mitjana respecte a la mesura de Gibbs sobre Σ_{N-1} amb hamiltonià H_{N-1,β^-} . Aleshores, per a una funció $f: \Sigma_N \rightarrow \mathbb{R}$, un simple càlcul algebraic ens permet escriure la identitat següent:

$$\langle f \rangle = \frac{\langle \mathbf{A} \mathbf{v} f \exp [\sigma_N (g(\boldsymbol{\rho}) + h)] \rangle_-}{Z}, \quad (6)$$

on $\mathbf{A} \mathbf{v}$ significa la mitjana respecte a l'últim *spin* $\sigma_N = \pm 1$ i

$$Z = \langle \mathbf{A} \mathbf{v} \exp [\sigma_N (g(\boldsymbol{\rho}) + h)] \rangle_- = \langle \cosh (g(\boldsymbol{\rho}) + h) \rangle_-.$$

La fórmula (6) té una generalització immediata per a una funció $f: \Sigma_N^n \rightarrow \mathbb{R}$, $n \geq 1$,

$$\langle f \rangle = \frac{\langle \mathbf{A} \mathbf{v} f \exp [\sum_{l \leq n} \sigma_N^l (g(\boldsymbol{\rho}^l) + h)] \rangle_-}{Z^n},$$

on aquest $\mathbf{A} \mathbf{v}$ vol dir la mitjana respecte a $\sigma_N^1 = \pm 1, \dots, \sigma_N^n = \pm 1$.

Una altra eina molt utilitzada és, tal com hem dit abans, el mètode del camí bo. Notem

$$g_t(\boldsymbol{\rho}) = \sqrt{t} g(\boldsymbol{\rho}) + \beta z \sqrt{1-t} \sqrt{q},$$

on $z \sim \mathcal{N}(0, 1)$ és independent de tots els $g_{i,j}$ i el paràmetre $q \in (0, 1)$ és l'única solució de l'equació (3). Aleshores, per una funció $f : \Sigma_N^n \rightarrow \mathbb{R}$, $n \geq 1$, definim:

$$\langle f \rangle_t = \frac{\langle \mathbf{A} \mathbf{v} f \mathcal{E}_{n,t} \rangle_-}{Z_t^n},$$

amb

$$\mathcal{E}_{n,t} = \mathcal{E}_{n,t}(\boldsymbol{\rho}^1, \dots, \boldsymbol{\rho}^n) = \exp \left[\sum_{l \leq n} \sigma_N^l (g_t(\boldsymbol{\rho}^l) + h) \right],$$

$$Z_t = \langle \mathbf{A} \mathbf{v} \mathcal{E}_{1,t} \rangle = \langle \mathbf{A} \mathbf{v} \exp [\sigma_N (g_t(\boldsymbol{\rho}) + h)] \rangle_- = \langle \cosh (g_t(\boldsymbol{\rho}) + h) \rangle_- ,$$

i també

$$v_t(f) = \mathbf{E} \langle f \rangle_t .$$

Observem ara que

$$v_1(f) = v(f) = \mathbf{E} \langle f \rangle \quad \text{i} \quad v_0(f) = \mathbf{E} \langle \mathbf{A} \mathbf{v} f \rangle_- .$$

Tota aquesta notació que sembla complicar les coses en realitat les simplificarà: d'una banda, habitualment $v_0(f)$ té una estructura molt més simple i és més fàcil de calcular que $v(f)$; d'altra banda, aquestes dues quantitats estan relacionades mitjançant l'expressió

$$v(f) = v_0(f) + \int_0^1 v'_t(f) dt . \quad (7)$$

Per tant, en lloc d'estudiar $v(f)$, tractarem $v_0(f)$ que serà més senzill i provarem de controlar o fitar la derivada $v'_t = \frac{d}{dt} v_t$.

Per a demostrar el teorema, necessitem tres resultats auxiliars. Els dos primers es demostren utilitzant el mètode de la cavitat, mentre que el tercer és un simple càlcul algebraic a partir de la definició de v_0 . Aquests resultats són els següents:

1. Per a $t \in (0, 1)$ i per a tota funció f definida a Σ_N^n tenim

$$\begin{aligned} v'_t(f) &= \beta^2 \sum_{1 \leq l \leq k \leq n} v_t(f \sigma_N^l \sigma_N^k (R_{l,k} - q)) - \\ &\quad - \beta^2 n \sum_{1 \leq l \leq n} v_t(f \sigma_N^l \sigma_N^{n+1} (R_{l,n+1} - q)) + \\ &\quad + \beta^2 \frac{n(n+1)}{2} v_t(f \sigma_N^{n+1} \sigma_N^{n+2} (R_{n+1,n+2} - q)) . \end{aligned} \quad (8)$$

2. Si f és una funció no negativa definida a Σ_N^n , tenim

$$v_t(f) \leq \exp(4n^2 \beta^2) v(f) . \quad (9)$$

3. Si f^- és una funció definida a Σ_{N-1}^n aleshores

$$\nu_0 \left(f^- \sigma_N^1 \sigma_N^2 \right) = E(\tanh^2 Y) \nu_0 (f^-), \quad (10)$$

on $Y = \beta z \sqrt{q} + h$ i z és una variable aleatòria gaussianiana estàndard.

Podem centrar-nos ja en la demostració del teorema.

PROVA DEL TEOREMA 1 Definim una modificació de la superposició quan no considerem el darrer *spin*

$$R_{1,2}^- = \frac{1}{N} \sum_{1 \leq i \leq N-1} \sigma_i^1 \sigma_i^2.$$

Sigui

$$f = (\sigma_N^1 \sigma_N^2 - q)(R_{1,2} - q) = \frac{1}{N} (\sigma_N^1 \sigma_N^2 - q)^2 + (\sigma_N^1 \sigma_N^2 - q)(R_{1,2}^- - \frac{N-1}{N}q).$$

Per un argument de simetria entre els N spins s'obté fàcilment

$$\nu \left((R_{1,2} - q)^2 \right) = \nu \left((\sigma_N^1 \sigma_N^2 - q)(R_{1,2} - q) \right) = \nu(f). \quad (11)$$

Utilitzant ara el mètode del camí bo, en tenim prou afitant $\nu_0(f)$ i $\nu_t'(f)$.

Estudiem en primer lloc $\nu_0(f)$. De la definició de q donada a (3) s'obté $E(\tanh^2 Y) = q$. Aleshores, a partir del tercer resultat auxiliar (10) tenim

$$\nu_0 \left((\sigma_N^1 \sigma_N^2 - q)(R_{1,2}^- - \frac{N-1}{N}q) \right) = [E(\tanh^2 Y) - q] \nu_0 \left(R_{1,2}^- - \frac{N-1}{N}q \right) = 0,$$

i

$$\nu_0 \left((\sigma_N^1 \sigma_N^2 - q)^2 \right) = \nu_0(1 - 2\sigma_N^1 \sigma_N^2 q + q^2) = 1 - 2q E(\tanh^2 Y) + q^2 = 1 - q^2.$$

Per tant

$$\nu_0(f) = \frac{1}{N}(1 - q^2). \quad (12)$$

Tractem ara l'altre terme $\nu_t'(f)$. Emprant el fet que $|\sigma_N^1 \sigma_N^2| \leq 1$ i el primer resultat auxiliar (8) obtenim

$$\begin{aligned} \nu_t'(f) &\leq \beta^2 \nu_t(|f||R_{1,2} - q|) + \\ &\quad + 2\beta^2 [\nu_t(|f||R_{1,3} - q|) + \nu_t(|f||R_{2,3} - q|)] + \\ &\quad + 3\beta^2 \nu_t(|f||R_{3,4} - q|). \end{aligned}$$

A partir de la definició de f , el fet que $|\sigma_N^1 \sigma_N^2 - q| \leq 2$ i la desigualtat de Hölder s'obté

$$\nu_t'(f) \leq 16\beta^2 \nu_t \left((R_{1,2} - q)^2 \right).$$

Al combinar aquesta desigualtat amb el segon resultat auxiliar (9) podem escriure

$$\nu'_i(f) \leq 16\beta^2 \exp(16\beta^2) \nu \left((R_{1,2} - q)^2 \right). \quad (13)$$

Ara, només cal ajuntar (7), (11), (12) i (13) per a obtenir

$$\nu \left((R_{1,2} - q)^2 \right) \leq \frac{1}{N} + 16\beta^2 \exp(16\beta^2) \nu \left((R_{1,2} - q)^2 \right).$$

Si escollim un β prou petit tal que

$$16\beta^2 \exp(16\beta^2) \leq \frac{1}{2},$$

s'acaba la demostració. □

Referències

- [1] BARDINA, X.; MÁRQUEZ-CARRERAS, D.; ROVIRA, C.; TINDEL, S. «Higher order expansions for the overlap of the SK model». Seminar on Stochastic Analysis, Random Fields and Applications IV. *Progr. in Probab.* 58. [Birkhäuser], (2004), 21–43.
- [2] BARDINA, X.; MÁRQUEZ-CARRERAS, D.; ROVIRA, C.; TINDEL, S. «The p -spin interaction model with external field». *Potential Anal.* 21 (4), (2004), 311–362.
- [3] CADEL, A.; MÁRQUEZ-CARRERAS, D.; ROVIRA, C. «The Gibbs' measure in the p -spin interaction model with external field». *Math. Preprint Series*, 356 (2004).
- [4] DERRIDA, B. «Random energy model: limit of a family of disordered systems». *Phys. Rev. Lett.*, 45 (2) (1980), 79–82.
- [5] DERRIDA, B. «Random energy model: an exactly solvable of disordered models». *Phys. Rev. B*, 24 (5) (1981), 2613–2626.
- [6] DERRIDA, B.; GARDNER, E. «Optimal storage properties of neural networks models». *J. Phys. A*, 21 (1988), 271–284.
- [7] GARDNER, E. Spin glasses with p -spin interactions. *Nuclear Phys. B* 257 (6) (1985), 747–765.
- [8] GARDNER, E. «The space of interactions in neural networks models». *J. Phys. A*, 21 (1988), 257–270.
- [9] HERTZ, J.; KROGH, A.; PALMER, R. *Introduction to the Theory of Neural Computation*. Addison-Wesley Publishing Company, 1991.
- [10] MEZARD, M.; PARISI, G.; VIRASORO, M. A. *Spin glass theory and beyond*. World Scientific, 1987.
- [11] PARISI, G; VIRASORO, M. A. «On a mechanism for explicit replica symmetry breaking». *J. Physique*, 50 (22) (1989), 3317–3329.

- [12] SHERRINGTON, D.; KIRKPATRICK, S. «Solvable model of a spin glass». *Phys. Rev. Lett.*, 35 (26) (1975), 1792-1796.
- [13] TALAGRAND, M. *Spin glasses: a challenge for mathematicians*. Springer, 2000.
- [14] TALAGRAND, M. «Mean field models for spin glasses: a first course». *Lecture Notes in Math.*, [Springer, 2003], 1816, 180-285.
- [15] TALAGRAND, M. «The generalized Parisi formula». *C. R. Math. Acad. Sci. Paris*, 337 (2003), 111-114.

FACULTAT DE MATEMÀTIQUES
UNIVERSITAT DE BARCELONA
GRAN VIA DE LES CORTS CATALANES, 585
08007 BARCELONA
davidmarquez@ub.edu
carles.rovira@ub.edu