Fluctuation results in spin glasses

Matthias Löwe

Mathematisches Kolloquium – Carl von Ossietzky Universität Oldenburg

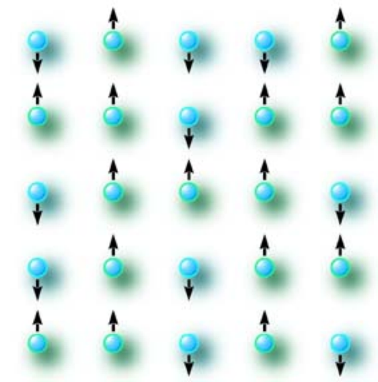
6 June 2012

Today's menu

”Above all, it must be said, that this paper would not have been written without the encouragement of Erwin Bolthausen (the reader, will observe that, as what should have been a three months project, ended only after years of very intense struggle, the word ’grateful’ was omitted in the acknowledgement)” (Michel Talagrand (1996)

The mother of all models?

- Lenz and Ising (~ 1922)
- How do you model a ferromagnet?
- Idea: Locate "atoms" at the sites of a subset of $\Lambda \subset \mathbb{Z}^d$.
- Each of them has a magnetic dipole, a spin $\in \{-1, +1\}$



The mother of all models? (cont.)

- These spins need to cooperate, to create a magnetic behavior.
- Introduce an energy function
- $H_\Lambda(\sigma) = -\sum_{\langle i,j\rangle \in \Lambda} \sigma_i \sigma_j (-h \sum_{j \in \Lambda} \sigma_j), \sigma \in \{\pm 1\}^\Lambda$
- and a probability measure
- $\mu_\Lambda(\sigma) = \frac{e^{-\beta H_\Lambda(\sigma)}}{Z_{\Lambda,\beta}}$ (Gibbs measure)
- with $Z_{\Lambda,\beta} = \sum_{\sigma' \in \{\pm 1\}^\Lambda} e^{-\beta H_\Lambda(\sigma')}$
- and $\beta > 0$.

Ising's predicition

- The model is uninteresting in dimension $d = 1$ (Right!)
- It also does not model magnetic behavior in $d = 2$ Wrong!
- Indeed, the model does show a phase transition from paramagentic to magnetic behavior in higher dimensions.

A related model

- An even simpler model with magnetic behavior:
- Replace interaction of σ_i and other $\sigma_j' s$
- by interaction of σ_i with an average spin: $\frac{1}{|\Lambda|}\sum_{j\in\Lambda_j}\sigma_j$
- The energy function becomes

$$H_N = -\frac{1}{2N}\sum_{i,j=1}^{N}\sigma_i\sigma_j$$

- This is known as the Curie-Weiss model
- Advantage: H_N is a function of the mean magnetization $m_N(\sigma) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i$.

A phase transition

- The Curie-Weiss model exhibits a phase transition at $\beta = 1$
- This can be in various ways, e.g.
- For $\beta \leq 1$: m_N converges to 0 under $\mu_{N,\beta}$
- For $\beta > 1$: The distribution of m_N under $\mu_{N,\beta}$ converges to $\frac{1}{2}\delta_{z(\beta)} + \delta_{-z(\beta)}$
- where $z(\beta)$ is the largest solution of

$$z = \tanh(\beta z).$$

Spin glasses

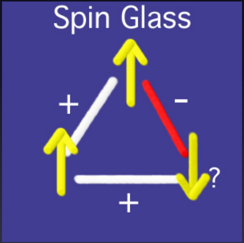
- In the early 1970's:
- Add randomness to the interactions
- i.e., consider the new energy function

$$H(\sigma) = -\sum_{i,j} \sigma_i \sigma_j J_{ij}$$

- where J_{ij} are random variables with
- $\mathbb{E}J_{ij} = 0$
- tossed in advance

Frustration

- Spin glasses are on another level of difficulty
- compared to ferromagnets
- Basic reason: frustration



- Result: Ground states cannot be easily read off
- Many metastable states

Spin glass models

- Edwards-Anderson model
- Disordered Ising model
- $H_\Lambda(\sigma) = -\sum_{\langle i,j\rangle \in \Lambda} \sigma_i \sigma_j J_{ij}$
- with e.g. J_{ij} i.i.d. $\mathcal{N}(0,1)$
- and the corresponding Gibbs measure
- Way to difficult

Spin glass models (cont.)

- Sherrington-Kirkpatrick model (SK model)
- ”Mean-field Spin glass”
- $H_N(\sigma) = -\frac{1}{\sqrt{N}} \sum_{1=i<j\leq N} \sigma_i \sigma_j J_{ij}$
- with again J_{ij} i.i.d. $\mathcal{N}(0,1)$
- and the corresponding Gibbs measure
- still very difficult
- ”Solved” by the physicists by the replica method”

Spin glass models (cont.)

- Hopfield model
- Another "Mean-field Spin glass"
- with a device to tune the difficulty of the model
- $H_N(\sigma) = -\sum_{1=i<j\leq N} \sigma_i \sigma_j J_{ij}$
- with $J_{ij} = \frac{1}{N} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu$
- and ξ_i^μ i.i.d. Bernoulli random variables
- with $\mathbb{P}(\xi_i^\mu = \pm 1) = \frac{1}{2}$
- and M may and will depend on N

The Hopfield model

- For $M \equiv 1 \Rightarrow$ Curie-Weiss model
- For $M = N$, morally

$$
\begin{aligned}
J_{ij} &= \frac{1}{N} \sum_{\mu=1}^{N} \xi_i^\mu \xi_j^\mu \\
&= \frac{1}{\sqrt{N}} \frac{1}{\sqrt{N}} \sum_{\mu=1}^{N} \xi_i^\mu \xi_j^\mu \\
&\sim \frac{1}{\sqrt{N}} Gij
\end{aligned}
$$

- with standard Gaussian random variables G_{ij}
- Hence Hopfield model and SK-model can be expected to behave similarly.

The Hopfield model

- The Hopfield model has various interpretations:
- As a spin glass (Pastur and Figotin, 1976)
- As a neural network (Hopfield 1982)
- Here the vectors ξ^μ are interpreted as information to be stored in a brain
- so-called patterns
- It can be shown, that for
- $\beta > 1$ and $M \ll N$
- the measure $\mu_{N,\beta}$ is concentrated on small balls centered in the $\pm\xi^\mu$ (Bovier/Gayrard)

The Hopfield model

- The Hopfield can also be interpreted as a model of social choice (Cont/L. 2003 and 2009).
- Here the σ are decisions to be made.
- The ξ_i^μ are characteristics of the i'th individual.
- The more two individuals resemble
- the more likely is it, that their decision is the same.

Spin glass models (cont.)

- The Random Energy Model (REM)
- Introduced by Derrida (1984)
- Basic idea: The energy function in the SK-model is a Gaussian process in the σ
- Ignore its covariance structure
- Then we obtain an energy function of the form
- $H_N(\sigma) = \sqrt{N} X_\sigma$
- and the corresponding Gibbs measure
- This is a caricature of a spin glass.

Types of Fluctuation

- Let X_n be a sequence of random variables (possibly in several dimensions)
- Assume there is a Law of Large Numbers on a scale a_n
- i.e. $\frac{X_n}{a_n}$ converges to some point or deterministic vector
- We distinguish the following three types of fluctuations

Central Limit theorem

- There is a scale $b_n \ll a_n$ such that
- X_n/b_n converges to a non-degenerate distribution
- generic example (of course)
- X_n is a sum of n i.i.d. random variables
- and $b_n = \sqrt{n}$

Large deviations

- We say that X_n obey a Large Deviations Principle(LDP), if
- there exists a lower semi-continuous function $I(\cdot)$
- with compact level sets $\{x : I(x) \leq L\}$
- such that for all measurable sets A
- and some scale α_n

$$\begin{aligned} \inf_{x \in A^\circ} I(x) &\leq \liminf \frac{1}{\alpha_n} \log \mathbb{P}(X_n/a_n \in A^\circ) \\ &\leq \limsup \frac{1}{\alpha_n} \log \mathbb{P}(X_n/a_n \in \overline{A}) \leq \inf_{x \in \overline{A}} I(x) \end{aligned}$$

- Example: Cramérs theorem

Moderate deviations

- We say that X_n obey a Moderate Deviations Principle(MDP), if
- there exists a lower semi-continuous function $J(\cdot)$
- with compact level sets $\{x : J(x) \leq L\}$
- such that for all measurable sets A
- and scales c_n and γ_n

$$\begin{aligned} \inf_{x \in A^\circ} I(x) &\leq \liminf \frac{1}{\gamma_n} \log \mathbb{P}((X_n - \mathbb{E}X_n)/c_n \in A^\circ) \\ &\leq \limsup \frac{1}{\gamma_n} \log \mathbb{P}((X_n - \mathbb{E}X_n)/c_n \in \overline{A}) \leq \inf_{x \in \overline{A}} I(x) \end{aligned}$$

Connections

- General folklore (?):
- Often, most often, always (?)
- If a centered sequence X_n obeys an LDP with rate function I
- and there is a CLT on scale b_n with limiting d.f. F, s.t. $1 - F(x) \sim f(x)$ for x large
- Then there is also an MDP with rate function J and
- $J(x) = \log f(x)$
- and similarly for the large deviations: $I(x) \sim J(x)$ for x small.

The free energy in the REM

- Important quantity in statistical physics: the free energy
- $\Phi_{N,\beta} = \frac{1}{N} \log Z_{N,\beta}$
- In disordered systems, this is a random variable in the disorder
- Consider the expected free energy

$$f_N(\beta) = \frac{1}{N} \mathbb{E} \log Z_{N,\beta}$$

- and its thermodynamical limit

$$f(\beta) = \lim_{N \to \infty} f_N$$

A phase transition in the REM

- In the REM $Z_{N,\beta} = \frac{1}{2^N} \sum_{\sigma \in \{\pm 1\}^N} e^{\sqrt{N} X_\sigma}$
- Then one can proof the following result:

Theorem (Derrida?, Talagrand, Bovier)

In the REM:

$$f = \begin{cases} \beta^2/2 & \text{if } \beta \leq \beta_c \\ \beta_c^2/2 + (\beta - \beta_c)\beta_c & \text{if } b > \beta_c \end{cases}$$

where $\beta_c = \sqrt{2 \log 2}$

A phase transition in the REM

- This theorem constitutes a phase transition in the REM
- This phase transition is of third order, as
- the limiting free energy has a jump in its second derivative

Fluctuations in the REM (cont.)

- But even in the high temperature regime there is another surprise

Theorem

The free energy of the REM has the following fluctuations:

- *If* $\beta < \sqrt{\log 2/2}$, *then*

$$e^{\frac{N}{2}(\log 2 - \beta^2)} \log \frac{Z_{N,\beta}}{\mathbb{E}Z_{N,\beta}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

- *If* $\beta = \sqrt{\ln 2/2}$, *then*

$$e^{\frac{N}{2}(\log 2 - \beta^2)} \log \frac{Z_{N,\beta}}{\mathbb{E}Z_{N,\beta}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1/2).$$

Fluctuations in the REM (cont.)

Theorem

- *Let* $\alpha \equiv \beta/\sqrt{2\log 2}$*. If* $\sqrt{\log 2/2} < \beta < \sqrt{2\log 2}$*, then*

$$e^{\frac{N}{2}(\sqrt{2\log 2}-\beta)^2+\frac{\alpha}{2}(\ln(N\log 2)+\log 4\pi)}\log\frac{Z_{N,\beta}}{\mathbb{E}Z_{N,\beta}}$$

$$\stackrel{\mathcal{D}}{\to}\int_{-\infty}^{\infty} e^{\alpha z}(\mathcal{P}(dz)-e^{-z}dz),$$

where $\mathcal{P}$ *denotes the Poisson point process on* $\mathcal{R}$ *with intensity measure* $e^{-x}dx$*.*

Discussion

- This means, there is a second phase transition in the REM
- at $\beta = \sqrt{\log 2/2} = \beta_c/2$.
- For $\beta < \beta_c/2$ the theorem is proved via showing that the Lindeberg condition for triangular arrays holds.
- This condition breaks down at $\beta_c/2$.
- Reason: influence of the extreme values of the summands.
- This is also reflected in the occurrence of the Poisson point processes for $\beta_c/2 < \beta < \beta_c$.

Fluctuations in the Hopfield model

- Which quantity should be considered?
- Observe that in the Hopfield model

$$H_N(\sigma) = -\frac{N}{2}||m_N(\sigma)||^2$$

- where $m_N(\sigma) = (m_N^\mu(\sigma))_{\mu=1}^M$ is the so-called overlap vector
- $m_N^\mu(\sigma) = \frac{1}{N}\sum_{i=1}^N \sigma_i \xi_i^\mu$
- The overlap is an order parameter of the system.

A limit theorem for the overlap

- A phase transition in the Hopfield model can be read off from the behavior of the overlap parameter

Theorem (Bovier, Gayrard, Picco)

For $M \ll N$ and almost all choices of the ξ^μ

- *For $\beta \leq 1$ the distribution of m_N weakly converges to δ_0.*
- *For $\beta > 1$ for any metrics d metrizing weak convergence,*

$$d(\mathbb{P} \circ m_N^{-1}, \frac{1}{2M} \sum_{\mu=1}^{M} \delta_{e^\mu z^+(\beta)} + \delta_{-e_\mu z^+(\beta)}) \to 0$$

with e_μ the $\mu' th$ unit vector and z^+ the largest solution to

$$z = \tanh(\beta z)$$

Fluctuations of the overlap

- The phase transition can also be read off from the fluctuations
- Bovier and Gayrard prove an almost sure LDP for the overlap in the above situation.
- The speed is n
- and the rate function
  - has a unique minimum in 0 for $\beta \leq 1$
  - and several minima in all the $\delta_{\pm e_\mu z^+}$ for $\beta > 1$.

Fluctuations of the overlap (cont.)

- Can we also see something happening of the level of Central Limit Theorem?
- A partial answer is given by the following classical CLT

Theorem (Gentz, Bovier/Gayrard)

If $\beta < 1$and $M \ll N$, the almost surely in the ξ^μ every finite-dimensional projection of the rescaled-overlap vector $\sqrt{N} m_N$ converges in distribution to $\mathcal{N}(0, \Sigma)$, where

$$\Sigma = \frac{1}{1-\beta} \, Id \, .$$

Central idea in the proof (and the following proofs)

- The energy function $H_N(\sigma) = \frac{N}{2}||m_N(\sigma)||^2$ is quadratic (and therefore makes the computation difficult).
- It can linearlised by the so called Hubbard-Stratonovich transformation.

$$\mu_{N,\beta} \circ (\sqrt{a_N N} m_N)^{-1} * \mathcal{N}(0, a_N/\beta) = \chi_{N,\beta,a_N,\xi}$$

where $1/N \leq a_N le 1$

- This new measure possesses a density $f_{N,\beta}$ w.r.t. to Lebesgue measure
- and

$$f_{N,\beta}(x) \sim \exp(-N\beta\Psi_{N,\beta}(x/\sqrt{Na_N}))$$

Central idea in the proof (cont.)

- Here

$$\Psi_{N,\beta}(x) = \frac{1}{2}||x||_2^2 - \frac{1}{\beta N}\sum_{i=1}^{N} \log\cosh(\beta\langle\xi_i, x\rangle).$$

- Expand Ψ
- Analyze $f_{N,\beta}$ (and χ)
- Relate this to the original measure.

What about $\beta = 1$

- At $\beta = 1$ the classical CLT obviously must fail.
- We have to rescale m_N differently.
- This is known form the Curie-Weiss model
- (by $N^{1/4}$)
- However, in the Hopfield model, there is a surprise waiting.

What about $\beta = 1$ (cont).

Theorem (with B. Gentz)

If M is independent of N. Then the distribution of the random probability measure $\mu_{N,\beta} \circ (\sqrt{N^{1/4} N} m_N)^{-1}$ converges weakly to the distribution of a random probability measure Q_η on $\mathbb{R}^M$. Its density is proportional to

$$\exp\left(-\frac{1}{12}|x||_4^4 - \sum_{\mu<\nu} x_\mu^2 x_\nu^2 + \sum_{\mu<\nu} \eta_{\mu,\nu} x_\mu x_\nu\right)$$

where $\eta_{\mu,\nu}$ are independent standard normally distributed r.vs.

What about $\beta = 1$ (cont).

- Note that other than for different values of β
- there is no almost sure result
- The limiting object stays random
- A similar result can also be shown for growing M
- as long as $M \ll N^{1/13}$
- Talagrand analyzes $M \ll N^{1/3}$

Why?

- Taylor-expand $\cosh$ in

$$\Psi_{N,\beta}(x) = \frac{1}{2}||x||_2^2 - \frac{1}{\beta N}\sum_{i=1}^{N} \log\cosh(\beta\langle\xi_i, x\rangle).$$

- At $\beta = 1$ the quadratic term $\frac{1}{2}||x||_2^2$ cancels with part of the second order term of $\log\cosh(\cdot)$
- It remains $\frac{1}{\sqrt{N}}\sum_{i=1}\sum_{\mu<\nu}\xi_i^\mu x^\mu \xi_i^\nu x^\nu$
- plus fourth order terms
- of which $\frac{1}{12}||x||_4^4$ and $\sum_{\mu<\nu} x_\mu^2 x_\nu^2$ survive

What about moderate deviations

- Can the phase transition be detected from the moderate deviations analysis?
- For $\beta < 1$ and $M \ll N$ the moderate deviations of m_N behave as they should
- i.e. for $1 \ll b_N \ll \sqrt{N}$ the rescaled overlap $b_N m_N$
- almost surely obeys an MDP with speed N/b_N^2 and quadratic rate function

What about moderate deviations

- At $\beta = 1$ the situation changes.
- "Away" from the CLT, i.e. for
- $M^6 \ll N$ and $M^2 << N^{1-4\gamma}$
- the rescaled overlap $N^\gamma m_N$, $0 < \gamma < 1/4$
- almost surely obeys an MDP with speed $N^{1-4\gamma}$ and rate function
- $I(x) = \frac{1}{12}|x||_4^4 - \sum_{\mu<\nu} x_\mu^2 x_\nu^2$
- However, even for finite M, for $b_N \ll \sqrt[4]{\log} \log N$
- the sequence $(N^{1/4}/b_N)m_N$ does not obey an a.s. MDP

Conclusion

Today we have

- Met several disordered models from statistical physics
- Seen the influence of a phase transition on fluctuation results
- Seen that in the REM we can detect a second phase transition on the level of a CLT
- Seen that in the Hopfield model the phase transition is characterized by
- a non-standard CLT for the overlap
- and by a breakdown of moderate deviations

Bedankt voor jullie aandacht!