An iterative construction of solutions of the TAP equations.

Erwin Bolthausen, University of Zürich

Disordered Media, Warwick, Sept. 7, 2011

Sherrington-Kirkpatrick model:

Random interactions: Independent centered Gaussians $g_{ij},\ i < j$, with variance $1/N$, defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

Inverse temperature $\beta > 0$, strength $h \geq 0$ of the external field.

Hamiltonian:

$$H_{N,\omega}(\boldsymbol{\sigma}) \stackrel{\mathrm{def}}{=} \beta \sum_{1 \leq i < j \leq N} g_{ij}(\omega)\, \sigma_i \sigma_j + h \sum_{i=1}^{N} \sigma_i,\ \ \sigma_i \in \{-1, 1\}$$

Partition function:

$$Z_{N,\omega} \stackrel{\mathrm{def}}{=} \sum_{\boldsymbol{\sigma}} \exp\left[H_{N,\omega}(\boldsymbol{\sigma})\right].$$

Gibbs measure on Σ_N :

$$\frac{\exp\left[H_{N,\omega}(\boldsymbol{\sigma})\right]}{Z_{N,\omega}}.$$

Gibbs-expectations (fixed ω) are written as $\langle \cdot \rangle$.

$$m_i(\omega) \stackrel{\mathrm{def}}{=} \langle \sigma_i \rangle .$$

TAP equations (Thouless-Anderson-Palmer):

$$m_i = \tanh\left(h + \beta \sum\nolimits_{j=1}^{N} g_{ij} m_j - \beta^2 (1-q)\, m_i\right),$$

where $g_{ij} = g_{ji},\ g_{ii} = 0$, and $q = q(\beta, h)$ is the (unique for $h > 0$) solutions of

$$q = \int \tanh^2\left(h + \beta\sqrt{q}x\right)\phi(dx),\ \phi(dx) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx.$$

Has to be understood in limiting $N \to \infty$ sense.
The **Onsager term** $\beta^2 (1-q)\, m_i$ is an Itô type correction: Standard mean-field theory:

$$m_i \simeq \tanh\left(h + \beta \sum\nolimits_j g_{ij} m_j\right).$$

For SK correct by replacing m_j on the rhs by $m_j^{\mathrm{cut}\ i}$ from the system with the connections to i cut. Expanding in first order for the g_{ij}:

$$m_j^{\mathrm{cut}\ i} \simeq m_j - \beta\left(1 - m_j^2\right) m_i g_{ij}.$$

Mathematical proofs of TAP only for small β : Talagrand, Chatterjee.

Free energy:

$$f(\beta,h) \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \log Z_N = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N .$$

The **replica symmetric "solution"**, given by SK:

$$\begin{aligned} \mathrm{RS}(\beta,h) &= \int \log\cosh\left(h + \beta\sqrt{q}z\right)\phi(dz) \\ &+ \frac{\beta^2}{4}(1-q)^2 + \log 2. \end{aligned}$$

Aizenman-Lebowitz-Ruelle ($h = 0$), Talagrand ($h \neq 0$):
Theorem For small enough β :

$$f(\beta,h) = \mathrm{RS}(\beta,h)$$

"Small enough β" is believed to mean that the **de Almayda–Thouless-condition** is satisfied:

$$\textbf{(AT)} \;:\; \beta^2 \int \frac{\phi(dz)}{\cosh^4\left(h + \beta\sqrt{q}z\right)} \leq 1.$$

Proposal for a direct construction of “solutions” of TAP: Assume $h > 0$: Recursive approximations $\left\{m_i^{(k)}\right\}_{1\leq i\leq N}$: Given (g_{ij})

$$m_i^{(0)} \stackrel{\text{def}}{=} 0,\ m_i^{(1)} \stackrel{\text{def}}{=} \sqrt{q},$$

$$m_i^{(k)} \stackrel{\text{def}}{=} \tanh\left(h + \beta \sum\nolimits_j g_{ij} m_j^{(k-1)} - \beta^2 (1-q)\, m_i^{(k-2)}\right),\ k \geq 2.$$

Questions:

- Structure of the dependence of $m^{(k)}$ on (g_{ij})?
- Convergence as $k \to \infty$?
- Relation to SK?

Second point:

Theorem: (AT) is satisfied iff

$$\lim_{k,k'\to\infty} \lim_{N\to\infty} \mathbb{E}\frac{1}{N}\sum_{i=1}^{N}\left(m_i^{(k)} - m_i^{(k')}\right)^2 = 0.$$

Proof by an evaluation of

$$\rho(j,k) \stackrel{\mathrm{def}}{=} \lim_{N\to\infty} \mathbb{E}\frac{1}{N}\sum_{i=1}^{N} m_i^{(k)} m_i^{(j)}.$$

Theorem

$$\rho(k,k) = q, \ \forall k, \ \ \rho(j,k) = \rho_j, \ \forall j < k,$$

where

$$\rho_{j+1} = \psi(\rho_j), \ \ \rho_1 \stackrel{\mathrm{def}}{=} \sqrt{q\psi(0)} < q.$$

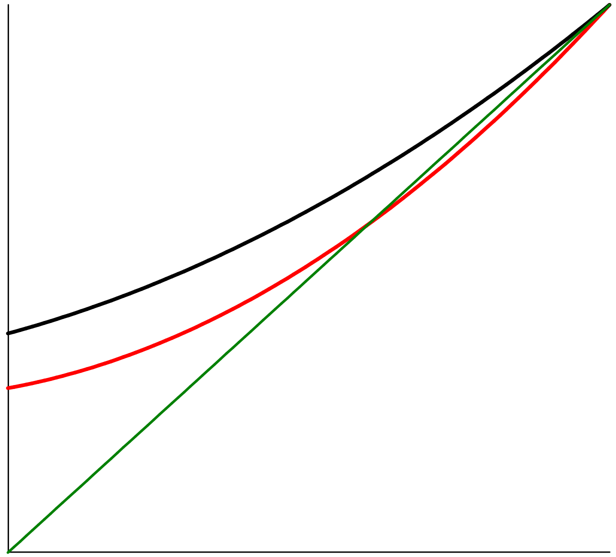
$\psi : [0,q] \to (0,q]$ defined by

$$\psi(t) = \int \tanh\left(h + \beta\sqrt{t}x_1 + \beta\sqrt{q-t}x_2\right)\tanh\left(h + \beta\sqrt{t}x_1 + \beta\sqrt{q-t}x_3\right)\phi^{\otimes 3}(dx).$$

$$\psi(q) = q.$$

Lemma ψ is increasing and convex on $[0, q]$. Furthermore

$$\psi'(q) = \beta^2 \int \frac{\phi(dx)}{\cosh^4\left(h + \beta\sqrt{q}x\right)}.$$



$$\psi^n(\rho_1) \to q \Longleftrightarrow \textbf{(AT)}$$

Structure of the (g_{ij})-dependence of the $m^{(k)}$: Alternative representation:

$$m_i^{(k)} \simeq \tanh\left(h + \beta \sum\nolimits_j g_{ij}^{(k-1)} m_j^{(k-1)} + \beta \sum\nolimits_{r=1}^{k-2} \gamma_r \xi_i^{(r)}\right),$$

$$\gamma_r \stackrel{\text{def}}{=} \frac{\rho_r - \sum_{j=1}^{r-1} \gamma_j^2}{\sqrt{q - \sum_{j=1}^{r-1} \gamma_j^2}},\ \xi_i^{(r)} \stackrel{\text{def}}{=} \sum_j g_{ij}^{(r)} \hat{m}_j^{(r)}.$$

$\hat{m}^{(1)}, \hat{m}^{(2)}, \ldots$ come from $m^{(1)}, m^{(2)}, \ldots$ via Gram-Schmidt in $\mathbb{R}^N$ w.r.t. the inner product

$$\langle x, y \rangle = \frac{1}{N} \sum\nolimits_i x_i y_i.$$

Let $\mathcal{F}_k = \sigma\left(\xi^{(1)}, \ldots, \xi^{(k)}\right)$. $m^{(k)}, \hat{m}^{(k)}$ are $\mathcal{F}_{k-1}$-measurable.
$\mathcal{L}\left(g^{(k)} \middle| \mathcal{F}_{k-2}\right)$ Gaussian, and $g^{(k)}$ is conditionally independent of $\mathcal{F}_{k-1}$.

$$\textbf{(AT)} \iff \sum\nolimits_{r=1}^{\infty} \gamma_r^2 = q$$

$$\textbf{(AT)} \iff \frac{1}{N}\sum_i \left[\sum\nolimits_j g_{ij}^{(k-1)} m_j^{(k-1)}\right]^2 \to_{k\to\infty} 0.$$

Using this representation, one can prove the claims by iteratively applying the LLN, conditionally successively on $\mathcal{F}_{k-2}, \mathcal{F}_{k-3}, \ldots$.
For that one needs the conditional covariance structure of the $\xi^{(k)}$:

$$\mathbb{E}\left(\xi_i^{(k)}\xi_j^{(k)}\middle|\,\mathcal{F}_{k-1}\right) = \begin{cases} 1 + O\left(N^{-1}\right) & \text{for } i = j, \\ \frac{\hat{m}_i^{(k)}\hat{m}_j^{(k)}}{N} - \frac{1}{N}\sum_{r=1}^{k-1}\frac{\hat{m}_i^{(r)}\hat{m}_j^{(r)}}{N} + O\left(N^{-2}\right) & \text{for } i \neq j. \end{cases}$$

Illustration from the first steps:

$$m_i^{(1)} = \sqrt{q},\ m_i^{(2)} = \tanh\left(h + \beta\sqrt{q}\xi_i^{(1)}\right),\ \xi_i^{(1)} \stackrel{\text{def}}{=} \sum\nolimits_j g_{ij},$$
$$\frac{1}{N}\sum\nolimits_{i=1}^N m_i^{(2)} \simeq \int \tanh\left(h + \beta\sqrt{q}x\right)\phi\left(dx\right) = \gamma_1,$$
$$\frac{1}{N}\sum\nolimits_{i=1}^N m_i^{(2)2} \simeq \int \tanh^2\left(h + \beta\sqrt{q}x\right)\phi\left(dx\right) = q,$$

are evident from LLN.

$$m_i^{(3)} = \tanh\left(h + \beta \sum\nolimits_j g_{ij} m_j^{(2)} - \beta^2 (1-q)\sqrt{q}\right).$$

Here one does the shift from g to $g^{(2)}$ which is independent of $\xi^{(1)}$. Essentially

$$g_{ij}^{(2)} \simeq g_{ij} - N^{-1}\left(\xi_i^{(1)} + \xi_j^{(1)}\right).$$

The correction inside the $\tanh$ of $m_i^{(3)}$:

$$\begin{aligned} \beta N^{-1} \sum_j \left(\xi_i^{(1)} + \xi_j^{(1)}\right) m_j^{(2)} &\simeq \beta\gamma_1 \xi_i^{(1)} + \beta \int x \tanh\left(h + \beta\sqrt{q}x\right) \phi(dx) \\ &= \beta\gamma_1 \xi_i^{(1)} + \beta^2 \sqrt{q}\,(1-q). \end{aligned}$$

$$\Longrightarrow m_i^{(3)} \simeq \tanh\left(h + \beta \sum\nolimits_j g_{ij}^{(2)} m_j^{(2)} + \beta \xi_i^{(1)}\right).$$

In the general case: $g \to g^{(2)} \to g^{(3)} \cdots$ successively eats up the Onsager term and produces the $\gamma_r \xi^{(r)}$ terms.

Open problem: Behavior beyond the (AT) line of

$$\sum\nolimits_j g_{ij}^{(k-1)} m_j^{(k-1)}.$$

Original motivation: (Re)prove $f(\beta, h) = \mathrm{RS}(\beta, h)$ hopefully up to **(AT)**, by a change of measure argument:

$$2^{-N} Z_N = \sum_\sigma \exp\left[-H(\sigma)\right] P^{\text{coin}-\text{toss}}(\sigma).$$

$$p_{m_i}^{\text{tilt}}(\sigma_i) = \frac{\exp\left[h_i \sigma_i\right]}{\cosh(h_i)} p^{\text{coin}-\text{toss}}(\sigma_i),\ m_i = \tanh(h_i).$$

The m_i from the TAP approximations.

$$\frac{1}{N} \log Z_N = \frac{1}{N} \log \sum_\sigma \exp\left[-H(\sigma) - \sum\nolimits_i h_i \sigma_i\right] P^{\text{tilt}}(\sigma) + \log 2$$
$$+ \underbrace{\frac{1}{N} \sum_i \log\cosh(h_i)}_{\simeq \int \log\cosh(h + \beta x)\phi(dx)}.$$

After some computations

$$\sum_{\sigma} \exp\left[-H(\sigma) - \sum\nolimits_i h_i \sigma_i\right] P^{\mathrm{tilt}}(\sigma) \simeq \exp\left[\frac{\beta^2 N}{4}(1-q)^2\right]$$
$$\times \sum_{\sigma} \exp\left[\beta \sum_{i<j} g_{ij} \hat{\sigma}_i \hat{\sigma}_j - \frac{\beta^2}{2N} \sum_{i<j} \hat{\sigma}_i^2 \hat{\sigma}_j^2\right] P^{\mathrm{tilt}}(\sigma),$$

where $\hat{\sigma}_i \stackrel{\mathrm{def}}{=} \sigma_i - m_i$. For the latter factor one should get (by second moment)

$$\lim_{N\to\infty} \frac{1}{N} \log[\cdot] = 0.$$

For small β, this is o.k., but not up to **(AT)**.

My favorite spin glass: Perceptron: $g_{ij},\ 1 \leq i, j \leq N$ i.i.d. Gaussian with variance N^{-1}.

$$H(\sigma) \stackrel{\text{def}}{=} \sum_{i=1}^{N} f\Big(\underbrace{\sum\nolimits_{j=1}^{N} g_{ij}\sigma_j}_{y_{\sigma,i}}\Big) = N \int f(x)\, L_{N,\sigma}(dx)$$

$$L_{N,\sigma} \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^{N} \delta_{y_{\sigma,i}}.$$

Question: Is there a quenched LDP in the sense that

$$\#\{\sigma : L_{N,\sigma} \sim \mu\} \sim 2^N \exp[-NJ(\mu)],\ \text{a.s.}$$

A natural question: Given the g_{ij}, is there a typical law of those σ's for which $L_{N,\sigma} \sim \mu$?