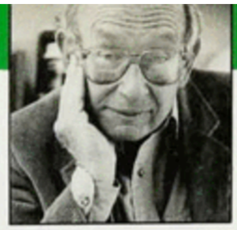
Machine Learning for Physical Scientists

Lecture 9
Intro to Statistical Physics of Inference and Spin Glasses

Statistical Physics and Computer Science: What's the connection?

"Ultrametricity"

REFERENCE FRAME



SPIN GLASS VI: SPIN GLASS AS CORNUCOPIA

Philip W. Anderson

Some attentive readers will recall a remark I made in my fourth column (September 1988, page 9), to the effect that in the difficulties and annoying features encountered in the study of spin glasses, we were beginning to have an inkling of results that would turn out to be among the most important of modern theoretical physics. I shall now try to make that clear to you. I explained one of the key results last time (July, page 9): the discovery by Gérard Toulouse and his collaborators that there are many inequivalent solutions of the TAP theory of the SK long-range spin glass and that those solutions can be arranged in an "ultrametric tree" whose branches already begin dividing as T is lowered below T_c. To remind you what this jargon means: The TAP theory is the mean-field theory David Thouless, Richard Palmer and I constructed. That theory, we thought, would in principle be exact because fluctuations about it should be negligible in view of the many long-range interactions each spin has in the SK spin glass. "Ultrametric" is an ant's-eye view of a tree, in which the only way to get to another leaf is to climb all the way down to the common branch point and back up (see the illustration in my last column).

Scott Kirkpatrick made a second important connection. Scott observed that finding the lowest-energy state of the SK spin glass—in fact, of almost any spin glass—is a complex optimization problem equivalent to one of the classic examples of what computer theorists call the NP-complete problems. This mysterious class of problems includes a great many mathematical "toys," such as bisecting random graphs, setting up mixed-doubles tournaments and inventing tours of length N for traveling salesmen or Chinese postmen; but it also contains many highly practical problems, such as routing telephone networks to N cities, designing chips with N transistors, connecting N chips together, evolving the fittest animal with N genes and doing almost anything useful with N neurons. Large complex optimization problems are everywhere around us, and almost anything that can be learned about them is of immense importance.

An important branch of computer science is complexity theory, which classifies such large problems according to their "size" N. The size of a complex problem may be thought of as the number of bits necessary to state that problem. For instance, the size of the SK spin glass problem is $N(N-1)/2$, the number of J_{ij}'s. It is strongly conjectured that the number of steps it takes a computer to solve an NP-complete problem cannot be less than a number proportional to an exponential of a positive power of the size. For large N, then, it could take forever. This is clearly the reason why Scott, Richard and others had been unable to find a unique lowest-energy state.

Each instance of the dozens of known NP-complete problems can be converted to an instance of any of the other problems by an algorithm taking only N^p time steps—that is, the number of time steps is a polynomial function of the size of the problem. This suggests that a statistical mechanical "solution" of the spin glass problem might be of general interest for all NP-complete problems. But that is not the case, even if one assumes that the "polynomial" algorithm that maps other problems to the spin glass is not more trouble than it is worth. Our statistical mechanical solution gives *average* answers for an ensemble of examples of the given problem. Such an answer is valid for a generic, or typical, instance of the problem. In the case of the spin glass, the average number describes the generic instance of the problem involving the given distribution of J_{ij}'s. But the mapping algorithm might transform that generic instance into a special case or vice versa. This issue was perhaps somewhat clarified in an exchange between Eric Baum (Princeton), on the one hand, and Daniel Stein (University of Arizona), G. Baskaran (MATSCIENCE, Madras, India) and myself, on the other, about NP-complete problems with "golf course" energy landscapes—landscapes that are flat everywhere except one point! Furthermore, proofs of NP completeness in computer science often refer only to the worst possible case, and some NP-complete problems do not look very hard in generic terms. Finally, the computer scientist discusses—for obvious reasons—the problem of finding the *exact* answer for a particular case, not the average answer correct to order N for the generic case.

Nonetheless, specifying exactly the structure of the landscape of energy values as a function in the 2^N-dimensional space of spins tells us a very great deal about such problems. For instance, the existence of a transition temperature T_c tells us that below some value of energy per site E_c the space bifurcates into regions corresponding to different "solutions," and that as we go lower and lower in energy (or temperature) the space breaks up more and more. This gives us a clear reason why such a problem is "exponentially" hard: If we are in the wrong region, we have to cross an energy and entropy barrier of order N to get a better solution. This kind of "freezing" phenomenon had been conjectured by computer scientists but never rigorously proved. To counter it, they had evolved a number of heuristic techniques for getting approximate solutions. We now know why this was necessary—namely, to get over the high barriers and sample the entire space of solutions.

Almost the first effect of the kind of thinking developed to understand spin glasses was to provide a *new* heuristic algorithm for the solution of complex optimization problems. That algorithm is called simulated annealing, and it was introduced by Scott and his colleague C. Daniel Gelatt Jr. Kirkpatrick and Gelatt proposed that one imitate the procedure the spin glassers had already been using, of "warming up" the problem above T_c and slowly cooling it back down, or "annealing" it. This could be done by regarding the "cost" for a given problem—say, the cost of connections on a chip—as a "Hamiltonian" function C of the positions to be varied. One plugs this Hamiltonian into a statistical mechanics simulator program, such as the well-known Metropolis algorithm. Then one chooses an appropriately scaled "temperature" T and minimizes $<e^{-C/T}>_{\text{ave}}$ for increasingly low temperatures. Simulated annealing, it turns out, is the most effective algorithm only for certain problems, but where it works it is very good indeed, and it is already in regular, profitable commercial use. The question of *why* simulated annealing works as well as it does was approached theoretically by Miguel Virasoro, who showed that, at least for the SK model, the lower the energy of a solution is, the larger is the entropy associated with it near T_c. That is, deeper valleys have bigger basins of attraction near T_c, and so one is more likely to start out in such a valley at T_c.

To me the key result here is the beautiful revelation of the structure of the randomly "rugged landscape" that underlies many complex optimization problems. Physics, however, has its own "nattering nabobs of negativism" (in the immortal phrase of William Safire), and they recently have been decrying the importance of the ultrametric structure, saying that it is a property of the SK model, not of physical spin glasses. Such criticism misses the point: Physical spin glasses and the SK model are only a jumping-off point for an amazing cornucopia of wide-ranging applications of the same kind of thinking. I will write about this in the next—and I hope the last—of these columns. ■

Philip Anderson is a condensed matter theorist whose work has also had impact on field theory, astrophysics, computer science and biology. He is Joseph Henry Professor of Physics at Princeton University.

Physics Today **42**, 9 (1989)

Statistical Physics and Inference: What's the connection?

Advances in Physics, 2016
Vol. 65, No. 5, 453–552, http://dx.doi.org/10.1080/00018732.2016.1211393

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Statistical physics of inference: thresholds and algorithms

Lenka Zdeborová[a*] and Florent Krzakala[b]

[a]*Institut de Physique Théorique, CNRS, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France;*
[b]*Laboratoire de Physique Statistique, Sorbonne Universités, Université Pierre et Marie Curie Paris 6, Ecole Normale Superieure, 75005 Paris, France*

(*Received 28 September 2015; accepted 30 June 2016*)

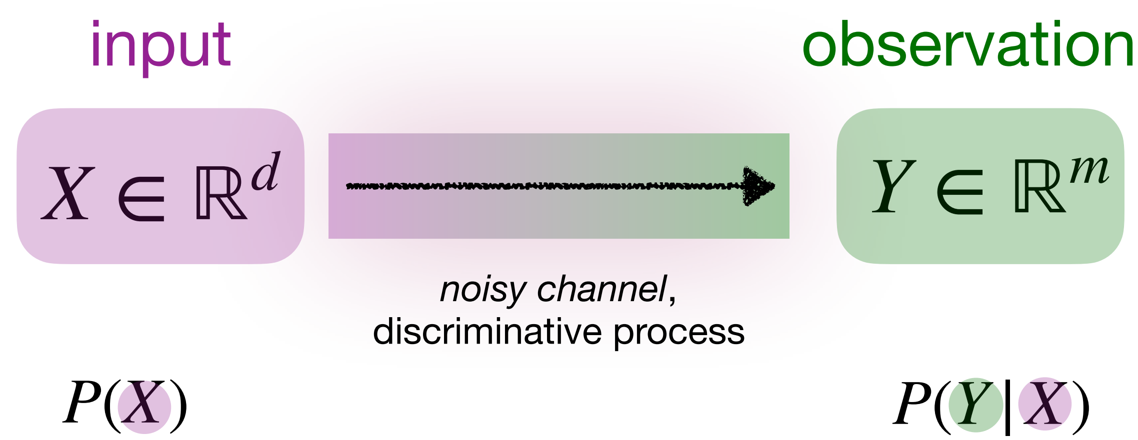
Many questions of fundamental interest in today's science can be formulated as inference problems: some partial, or noisy, observations are performed over a set of variables and the goal is to recover, or infer, the values of the variables based on the indirect information contained in the measurements. For such problems, the central scientific questions are: Under what conditions is the information contained in the measurements sufficient for a satisfactory inference to be possible? What are the most efficient algorithms for this task? A growing body of work has shown that often we can understand and locate these fundamental barriers by thinking of them as phase transitions in the sense of statistical physics. Moreover, it turned out that we can use the gained physical insight to develop new promising algorithms. The connection between inference and statistical physics is currently witnessing an impressive renaissance and we review here the current state-of-the-art, with a pedagogical focus on the Ising model which, formulated as an inference problem, we call the planted spin glass. In terms of applications we review two classes of problems: (i) inference of clusters on graphs and networks, with community detection as a special case and (ii) estimating a signal from its noisy linear measurements, with compressed sensing as a case of sparse estimation. Our goal is to provide a pedagogical review for researchers in physics and other fields interested in this fascinating topic.

PACS: 89.70.Eg interdisciplinary: computational complexity; 75.50.Lk magneticproperties of materials: spin glasses and other random magnets; 64.60.aq networks; 89.75.-k complex systems

Keywords: Bayesian inference; spin glass theory; compressed sensing; stochastic block model; belief propagation; phase transitions in computer science

link to the article

Inference in (equilibrium) Statistical Physics Framework



Given the observation, and known *noisy channel* models, can we *infer* the input?

Inference in (equilibrium) Statistical Physics Framework

input $\quad\quad\quad\quad\quad\quad\quad\quad$ observation

$$X \in \mathbb{R}^d \longrightarrow Y \in \mathbb{R}^m$$

noisy channel,
discriminative process

$P(X)$ $\quad\quad\quad\quad\quad\quad\quad\quad$ $P(Y|X)$

Given the observation, and known *noisy channel* models, can we *infer* the input?

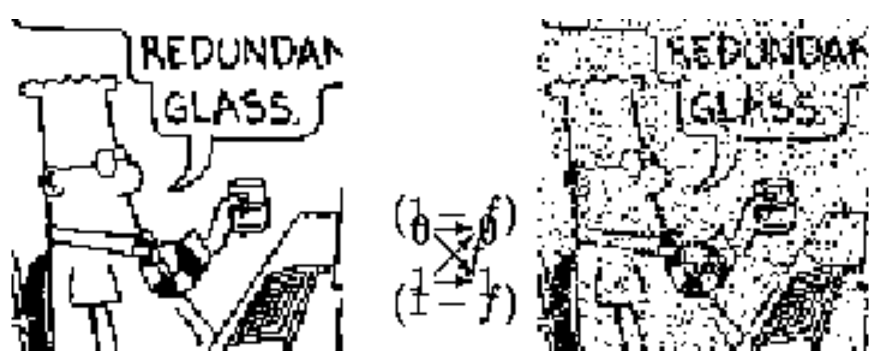
noisy
communication channel

Bob

Alice

compressed sensing/
error-correcting code

True Signal



Noisy Observation

Inference in (equilibrium) Statistical Physics Framework

input $\qquad\qquad$ observation

$$X \in \mathbb{R}^d \longrightarrow Y \in \mathbb{R}^m$$

noisy channel,
discriminative process

$P(X) \qquad\qquad P(Y|X)$

Given the observation, and known *noisy channel* models, can we *infer* the input?

$$\boxed{P(X|Y) = \frac{P(Y|X)P(X)}{Z} = \frac{1}{Z}e^{\log[P(Y|X)P(X)]}}$$

Inference in (equilibrium) Statistical Physics Framework

input $\qquad\qquad\qquad\qquad$ observation

$X \in \mathbb{R}^d \longrightarrow Y \in \mathbb{R}^m$

noisy channel,
discriminative process

$P(X) \qquad\qquad\qquad\qquad P(Y|X)$

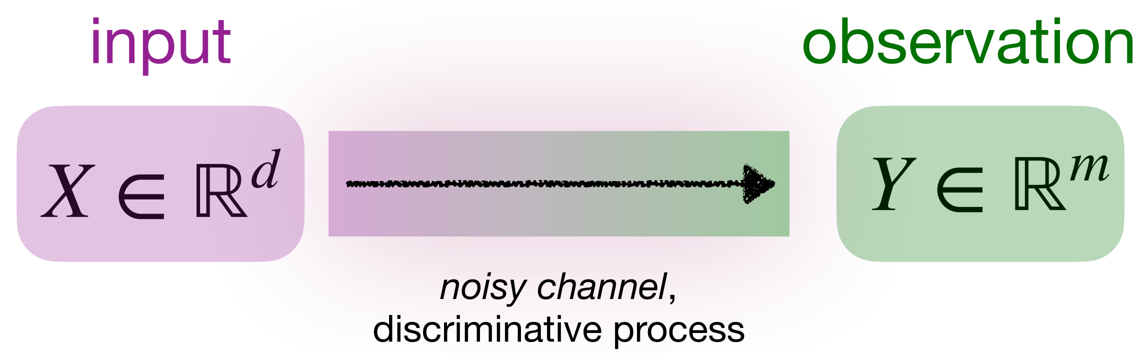
Given the observation, and known *noisy channel* models, can we *infer* the input?

$$\boxed{P(X|Y) = \frac{P(Y|X)P(X)}{Z} = \frac{1}{Z}e^{\log[P(Y|X)P(X)]}}$$

In this case, the **Hamiltonian** is $\mathscr{H}(\mathbf{X})_Y = -\log[P(Y|X)P(X)]$,the inverse temperature $\beta = 1$ and the **partition function** is just the marginal probability

$$Z(\beta) \equiv \sum_{\{X\}} e^{-\beta \mathscr{H}(\mathbf{X})}$$

Inference in (equilibrium) Statistical Physics Framework



$P(X)$ $\qquad\qquad$ $P(Y|X)$

Given the observation, and known *noisy channel* models, can we *infer* the input?

$$\boxed{P(X|Y) = \frac{P(Y|X)P(X)}{Z} = \frac{1}{Z}e^{\log[P(Y|X)P(X)]}}$$

This is also our *cost function/negative log-likelihood* objective function

In this case, the **Hamiltonian** is $\mathscr{H}(\mathbf{X})_Y = -\log[P(Y|X)P(X)]$,the inverse temperature $\beta = 1$ and the **partition function** is just the marginal probability

$$Z(\beta) \equiv \sum_{\{X\}} e^{-\beta \mathscr{H}(\mathbf{X})}$$

If we can compute the Boltzmann factor, then we can obtain a good estimator $\hat{X}$ of X, given the observation Y.

In fact, the *MAP estimator* can be obtained by minimising the cost function, or equivalently finding the ground state energy.

$$\boxed{\lim_{\beta\to\infty} -\partial_\beta \log Z_N(\beta) = \lim_{\beta\to\infty} \langle \mathscr{H}(\mathbf{X}) \rangle_\beta = \min \mathscr{H}(\mathbf{X})}$$

So, the left hand-side (derivative of the Free Energy) is what we have calculated in various models of equilibrium statistical mechanics.

However, the useful ones for inference turns out to be the model of spin glasses, which we haven't studied much.

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathscr{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathcal{N}(0,\Delta)$$

$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathcal{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathcal{N}(0,\Delta)$$

$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Question: what is a ***typical value*** of the minimal cost $e^{\min} = \mathcal{H}^{\min}/N$ as $N \to \infty$

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathcal{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathcal{N}(0,\Delta)$$

$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Question: what is a <u>*typical value*</u> of the minimal cost $e^{\min} = \mathcal{H}^{\min}/N$ as $N \to \infty$

$$e_{\min} = -\mathbb{E}\left[\lim_{\beta\to\infty} \frac{1}{N}\partial_\beta \log Z_N(\beta)\right]$$

$$= -\lim_{\beta\to\infty} \partial_\beta \boxed{\mathbb{E}\left[\frac{1}{N}\log Z_N(\beta)\right]}$$

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathcal{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathcal{N}(0,\Delta)$$
$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Question: what is a *typical value* of the minimal cost $e^{\min} = \mathcal{H}^{\min}/N$ as $N \to \infty$

$$e_{\min} = -\mathbb{E}\left[\lim_{\beta\to\infty} \frac{1}{N}\partial_\beta \log Z_N(\beta)\right]$$
$$= -\lim_{\beta\to\infty} \partial_\beta \mathbb{E}\left[\frac{1}{N}\log Z_N(\beta)\right]$$

average free energy $\Phi_N(\beta)$

Known fact, in $N \to \infty$ limit, typical value of the free energy is close to the expectation

$$\frac{1}{N}\log Z_N(\beta) = \Phi_N(\beta) + O\left(\frac{1}{\sqrt{N}}\right)$$

Physicists' jargon: "The free-energy is self-averaging"
Mathematicians' jargon: "The free-energy concentrates near the mean"

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathscr{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathscr{N}(0,\Delta)$$
$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Question: what is a *typical value* of the minimal cost $e^{\min} = \mathscr{H}^{\min}/N$ as $N \to \infty$

$$e_{\min} = -\mathbb{E}\left[\lim_{\beta\to\infty}\frac{1}{N}\partial_\beta \log Z_N(\beta)\right]$$
$$= -\lim_{\beta\to\infty}\partial_\beta \boxed{\mathbb{E}\left[\frac{1}{N}\log Z_N(\beta)\right]}$$

Thus, e_{min} is a good proxy for typical value of minimal cost.

average free energy $\Phi_N(\beta)$

Known fact, in $N \to \infty$ limit, typical value of the free energy is close to the expectation

$$\frac{1}{N}\log Z_N(\beta) = \Phi_N(\beta) + O\left(\frac{1}{\sqrt{N}}\right)$$

Physicists' jargon: "The free-energy is self-averaging"
Mathematicians' jargon: "The free-energy concentrates near the mean"

Spin Glass and the Replica Trick in a Nutshell

Toy problem: the random-field Ising model

$$\mathscr{H}(\mathbf{S}) \equiv -\sum_i h_i S_i - \frac{1}{2}\left(\frac{\sum_i S_i}{N}\right)^2 \quad \textbf{with} \quad h_i \sim \mathcal{N}(0,\Delta)$$
$$S_i = \pm 1, \quad i = 1,\ldots,N$$

Question: what is a <u>*typical value*</u> of the minimal cost $e^{\min} = \mathscr{H}^{\min}/N$ as $N \to \infty$

$$e_{\min} = -\mathbb{E}\left[\lim_{\beta\to\infty} \frac{1}{N}\partial_\beta \log Z_N(\beta)\right]$$
$$= -\lim_{\beta\to\infty} \partial_\beta \underbrace{\mathbb{E}\left[\frac{1}{N}\log Z_N(\beta)\right]}_{\textit{average free energy } \Phi_N(\beta)}$$

Thus, e_{min} is a good proxy for typical value of minimal cost.

How to compute the expectation of the logarithm???

Known fact, in $N \to \infty$ limit, typical value of the free energy is close to the expectation

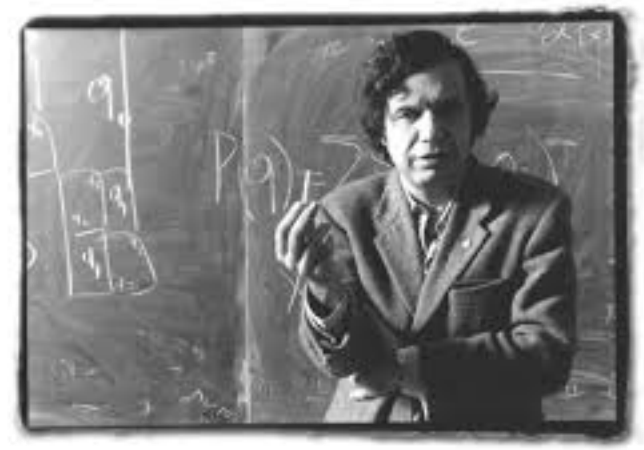
$$\frac{1}{N}\log Z_N(\beta) = \Phi_N(\beta) + O\left(\frac{1}{\sqrt{N}}\right)$$

Physicists' jargon: "The free-energy is self-averaging"
Mathematicians' jargon: "The free-energy concentrates near the mean"

The Replica “Trick”

Mark Kac

Sam Edwards



Giorgio Parisi

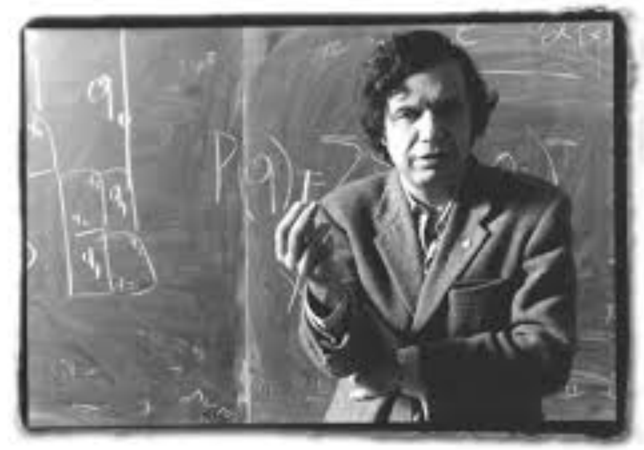
Mark Mezard

For $n \ll 1$, $\mathbb{E}\left[Z^n\right] = \mathbb{E}\left[e^{n \log Z}\right] \approx \mathbb{E}[1 + n \log Z] \approx 1 + n\mathbb{E}[\log Z]$

The Replica "Trick"

Mark Kac

Sam Edwards



Giorgio Parisi

Mark Mezard

For $n \ll 1$, $\quad \mathbb{E}\left[Z^n\right] = \mathbb{E}\left[e^{n \log Z}\right] \approx \mathbb{E}[1 + n \log Z] \approx 1 + n\mathbb{E}[\log Z]$

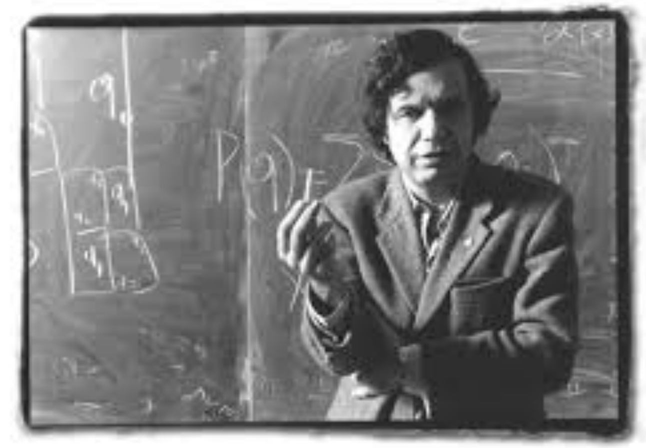
So

$$\boxed{\mathbb{E}[\log Z] = \lim_{n \to 0} \frac{\mathbb{E}\left[Z^n\right] - 1}{n}}$$

The Replica “Trick”

Mark Kac

Sam Edwards



Giorgio Parisi

Mark Mezard

For $n \ll 1$, $\quad \mathbb{E}\left[Z^n\right] = \mathbb{E}\left[e^{n\log Z}\right] \approx \mathbb{E}[1 + n\log Z] \approx 1 + n\mathbb{E}[\log Z]$

So

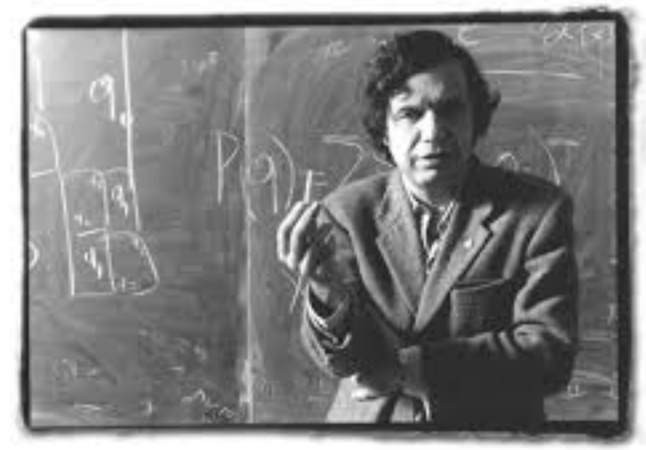
$$\boxed{\mathbb{E}[\log Z] = \lim_{n\to 0} \frac{\mathbb{E}\left[Z^n\right] - 1}{n}}$$

Recall from statistical physics that for *n* independent systems, the total partition function is the product of individual partition function.

The Replica "Trick"

Mark Kac

Sam Edwards



Giorgio Parisi

Mark Mezard

For $n \ll 1$, $\quad \mathbb{E}\left[Z^n\right] = \mathbb{E}\left[e^{n \log Z}\right] \approx \mathbb{E}[1 + n \log Z] \approx 1 + n\mathbb{E}[\log Z]$

So

$$\boxed{\mathbb{E}[\log Z] = \lim_{n \to 0} \frac{\mathbb{E}\left[Z^n\right] - 1}{n}}$$

Recall from statistical physics that for n independent systems, the total partition function is the product of individual partition function.

This heuristic trick is to evaluate the expectation of *n* replicas, then eventually treats *n* as a real number, and assume analytic continuation then take the limit *n.*

We'll take limit $n \to 0$, $N \to \infty$ without worrying about the order of limit…

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k}{N}\right)^2 + \beta \sum_k h_k S_k} \right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k^\alpha}{N}\right)^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

creates *n* replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0,\Delta)$

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2} \left(\frac{\sum_k S_k}{N} \right)^2 + \beta \sum_k h_k S_k} \right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2} \left(\frac{\sum_k S_k^\alpha}{N} \right)^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

creates n replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0,\Delta)$

$$= \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} N \int dm_\alpha \delta \left(N m_\alpha - \sum_k S_k^\alpha \right) e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Introduce auxiliary variables (fields) to disentangle the coupling between random field and the state variable S.

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2} \left(\frac{\sum_k S_k}{N} \right)^2 + \beta \sum_k h_k S_k} \right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2} \left(\frac{\sum_k S_k^\alpha}{N} \right)^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

creates n replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0, \Delta)$

$$= \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} N \int dm_\alpha \delta \left(N m_\alpha - \sum_k S_k^\alpha \right) e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Introduce auxiliary variables (fields) to disentangle the coupling between random field and the state variable *S*.

$$\propto \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} \int dm_\alpha \int d\hat{m}_\alpha e^{i \hat{m}_\alpha \left(N m_\alpha - \sum_k S_k^\alpha \right)} e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Fourier representation of delta function.

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k}{N}\right)^2 + \beta \sum_k h_k S_k} \right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k^\alpha}{N}\right)^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

creates n replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0,\Delta)$

$$= \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} N \int dm_\alpha \delta\left(N m_\alpha - \sum_k S_k^\alpha \right) e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Introduce auxiliary variables (fields) to disentangle the coupling between random field and the state variable *S*.

$$\propto \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} \int dm_\alpha \int d\hat{m}_\alpha e^{i\hat{m}_\alpha\left(N m_\alpha - \sum_k S_k^\alpha\right)} e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Fourier representation of delta function.

decoupled field theory!

$$\propto \prod_{\alpha=1}^{n} \left(\int dm_\alpha \int d\hat{m}_\alpha e^{iN\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \sum_{\{S^\alpha\}} e^{\beta \sum_k h_k S_k^\alpha - i\hat{m}_\alpha \sum_k S_k^\alpha} \right)$$

Swapping the order of sum and integration.

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k}{N}\right)^2 + \beta \sum_k h_k S_k} \right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2}\left(\frac{\sum_k S_k^\alpha}{N}\right)^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

creates n replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0,\Delta)$

$$= \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} N \int dm_\alpha \delta\left(N m_\alpha - \sum_k S_k^\alpha \right) e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Introduce auxiliary variables (fields) to disentangle the coupling between random field and the state variable *S*.

$$\propto \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} \int dm_\alpha \int d\hat{m}_\alpha e^{i\hat{m}_\alpha\left(N m_\alpha - \sum_k S_k^\alpha\right)} e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha} \right)$$

Fourier representation of delta function.

decoupled field theory!

$$\propto \prod_{\alpha=1}^{n} \left(\int dm_\alpha \int d\hat{m}_\alpha e^{iN\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \sum_{\{S^\alpha\}} e^{\beta \sum_k h_k S_k^\alpha - i\hat{m}_\alpha \sum_k S_k^\alpha} \right)$$

Swapping the order of sum and integration.

$$\propto \prod_{\alpha=1}^{n} \left(\int dm_\alpha \int d\hat{m}_\alpha e^{iN\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \prod_k 2\cosh\left(\beta h_k - i\hat{m}_\alpha\right) \right)$$

Evaluate decoupled theory

Replica Trick in Random Field Ising Model

$$Z_N^n = \left(\sum_{\{S\}} e^{\beta \frac{N}{2}\left(\frac{\Sigma_k S_k}{N}\right)^2 + \beta \sum_k h_k S_k}\right)^n = \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} e^{\beta \frac{N}{2}\left(\frac{\Sigma_k S_k^\alpha}{N}\right)^2 + \beta \sum_k h_k S_k^\alpha}\right)$$

creates n replicates/independent systems

remember h_k is i.i.d. $\mathcal{N}(0,\Delta)$

$$= \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} N \int dm_\alpha \delta\left(Nm_\alpha - \sum_k S_k^\alpha\right) e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha}\right)$$

Introduce auxiliary variables (fields) to disentangle the coupling between random field and the state variable S.

$$\propto \prod_{\alpha=1}^{n} \left(\sum_{\{S^\alpha\}} \int dm_\alpha \int d\hat{m}_\alpha e^{i\hat{m}_\alpha \left(Nm_\alpha - \sum_k S_k^\alpha\right)} e^{\beta \frac{N}{2} m_\alpha^2 + \beta \sum_k h_k S_k^\alpha}\right)$$

Fourier representation of delta function.

decoupled field theory!

$$\propto \prod_{\alpha=1}^{n} \left(\int dm_\alpha \int d\hat{m}_\alpha e^{iN\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \sum_{\{S^\alpha\}} e^{\beta \sum_k h_k S_k^\alpha - i\hat{m}_\alpha \sum_k S_k^\alpha}\right)$$

Swapping the order of sum and integration.

$$\propto \prod_{\alpha=1}^{n} \left(\int dm_\alpha \int d\hat{m}_\alpha e^{iN\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \prod_k 2\cosh\left(\beta h_k - i\hat{m}_\alpha\right)\right)$$

Evaluate decoupled theory

So far so good, we haven't done any fancy math. These are standard field-theoretic tricks to decouple the non-linear interaction by introducing auxiliary fields variables.

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha\right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right)\right]^N$$

Remember, we'll take limit $n \to 0$, $N \to \infty$

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha\right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right)\right]^N$$

This limit allows us to guess the saddle point solution of the integral

Remember, we'll take limit $n \to 0$, $N \to \infty$

This limit seems strange and unphysical

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha \right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right) \right]^N$$

This limit allows us to guess the saddle point solution of the integral

Remember, we'll take limit $n \to 0$, $N \to \infty$

This limit seems strange and unphysical

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha\right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta\frac{N}{2}m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right)\right]^N$$

"Replica Symmetric Assumption"

The integral is exponentially dominated by values of $m_\alpha = m$, $\hat{m}_\alpha = \hat{m}\ \forall \alpha$

We won't talk about replica symmetry breaking (where different replicates are dominated by different *m* values), introduced by Giorgio Parisi.

This limit allows us to guess the saddle point solution of the integral

Remember, we'll take limit $n \to 0$, $N \to \infty$

This limit seems strange and unphysical

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha \right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right)\right]^N$$

"Replica Symmetric Assumption"

The integral is exponentially dominated by values of $m_\alpha = m$, $\hat{m}_\alpha = \hat{m}\ \forall \alpha$

We won't talk about replica symmetry breaking (where different replicates are dominated by different *m* values), introduced by Giorgio Parisi.

$$\mathbb{E}\left[Z_N^n\right] \propto \int dm d\hat{m} e^{inN\left(\hat{m}m + \beta\frac{1}{2}m^2\right)} \mathbb{E}\left[(2\cosh(\beta h - i\hat{m}))^n\right]^N$$

Since we'll take the limit of infinitesimal *n,* we now perform another replica trick inside the replica trick =)

$$\mathbb{E}\left[X^n\right] = \mathbb{E}\left[e^{n\log X}\right] \approx \mathbb{E}[1 + n\log X] \approx 1 + n\mathbb{E}[\log X] \approx e^{n\mathbb{E}[\log X]}$$

This limit allows us to guess the saddle point solution of the integral

Remember, we'll take limit $n \to 0$, $N \to \infty$

This limit seems strange and unphysical

$$\mathbb{E}\left[Z_N^n\right] \propto \int \left(\prod_{\alpha=1}^{n} dm_\alpha d\hat{m}_\alpha \right) e^{i\sum_\alpha N\hat{m}_\alpha m_\alpha + \beta \frac{N}{2} m_\alpha^2} \mathbb{E}\left[\prod_\alpha 2\cosh\left(\beta h - i\hat{m}_\alpha\right)\right]^N$$

“Replica Symmetric Assumption”

The integral is exponentially dominated by values of $m_\alpha = m$, $\hat{m}_\alpha = \hat{m}\ \forall \alpha$

We won't talk about replica symmetry breaking (where different replicates are dominated by different *m* values), introduced by Giorgio Parisi.

$$\mathbb{E}\left[Z_N^n\right] \propto \int dm d\hat{m} e^{inN\left(\hat{m}m + \beta\frac{1}{2}m^2\right)} \mathbb{E}\left[(2\cosh(\beta h - i\hat{m}))^n\right]^N$$

Since we'll take the limit of infinitesimal *n,* we now perform another replica trick inside the replica trick =)

$$\mathbb{E}\left[X^n\right] = \mathbb{E}\left[e^{n\log X}\right] \approx \mathbb{E}[1 + n\log X] \approx 1 + n\mathbb{E}[\log X] \approx e^{n\mathbb{E}[\log X]}$$

To obtain

$$\mathbb{E}\left[Z_N^n\right] \propto \int dm d\hat{m} e^{inN\left(\hat{m}m + \beta\frac{1}{2}m^2\right)} e^{nN\mathbb{E}[\log(2\cosh(\beta h - i\hat{m}))]}$$

From

$$\mathbb{E}\left[Z_N^n\right] \propto \int dm d\hat{m} e^{inN\left(\hat{m}m + \beta\frac{1}{2}m^2\right)} e^{nN\mathbb{E}[\log(2\cosh(\beta h - i\hat{m}))]}$$

take the limit $N \to \infty$ and the solution is dominated by the saddle point integration

$$\mathbb{E}\left[Z_N^n\right] \propto e^{nN\,\mathrm{Extr}_{m,\hat{m}}\left\{i\left(\hat{m}m + \beta\frac{1}{2}m^2\right) + \mathbb{E}[\log(2\cosh(\beta h - i\hat{m}))]\right\}}$$

Extremizing the arguments, one finds the necessary condition $i\hat{m} = \beta m$

Therefore,

$$\boxed{\mathbb{E}\left[Z_N^n\right] \propto e^{nN\,\mathrm{Extr}_{m}\left\{-\beta\frac{1}{2}m^2 + \mathbb{E}[\log(2\cosh(\beta(h+m)))]\right\}}}$$

From

$$\mathbb{E}\left[Z_N^n\right] \propto \int dm d\hat{m} e^{inN\left(\hat{m}m + \beta\frac{1}{2}m^2\right)} e^{nN\mathbb{E}[\log(2\cosh(\beta h - i\hat{m}))]}$$

take the limit $N \to \infty$ and the solution is dominated by the saddle point integration

$$\mathbb{E}\left[Z_N^n\right] \propto e^{nN\,\mathrm{Extr}_{m,\hat{m}}\left\{i\left(\hat{m}m + \beta\frac{1}{2}m^2\right) + \mathbb{E}[\log(2\cosh(\beta h - i\hat{m}))]\right\}}$$

Extremizing the arguments, one finds the necessary condition $i\hat{m} = \beta m$

Therefore,

$$\boxed{\mathbb{E}\left[Z_N^n\right] \propto e^{nN\,\mathrm{Extr}_{m}\left\{-\beta\frac{1}{2}m^2 + \mathbb{E}[\log(2\cosh(\beta(h+m)))]\right\}}}$$

Recall the purpose of this replica trick: we're after the average free energy, which is a proxy of typical value of minimal cost.

$$\begin{aligned} e_{\min} &= -\mathbb{E}\left[\lim_{\beta\to\infty}\frac{1}{N}\partial_\beta \log Z_N(\beta)\right] \\ &= -\lim_{\beta\to\infty}\partial_\beta \underbrace{\mathbb{E}\left[\frac{1}{N}\log Z_N(\beta)\right]}_{\text{average free energy } \Phi_N(\beta)} \end{aligned}$$

From the identity $$\mathbb{E}[\log Z] = \lim_{n\to 0} \frac{\mathbb{E}\left[Z^n\right] - 1}{n}$$ we obtain

$$\boxed{\mathbb{E}\left[\frac{1}{N}\log Z_N\right] \to \mathrm{Extr}_m\left\{\Phi_{\mathrm{RS}}(m)\right\} \quad , \Phi_{\mathrm{RS}}(m) \equiv -\beta\frac{1}{2}m^2 + \mathbb{E}\left[\log\left(2\cosh(\beta(h+m))\right)\right]}$$

"Replica Symmetric Potential"

The extremum point satisfies

$$m^* = \mathbb{E}\left[\tanh\left(\beta\left(h + m^*\right)\right)\right]$$

From the identity $\mathbb{E}[\log Z] = \lim_{n \to 0} \frac{\mathbb{E}\left[Z^n\right] - 1}{n}$ we obtain

$$\mathbb{E}\left[\frac{1}{N}\log Z_N\right] \to \mathrm{Extr}_m\left\{\Phi_{\mathrm{RS}}(m)\right\} \qquad , \Phi_{\mathrm{RS}}(m) \equiv -\beta\frac{1}{2}m^2 + \mathbb{E}\left[\log\left(2\cosh(\beta(h+m))\right)\right]$$

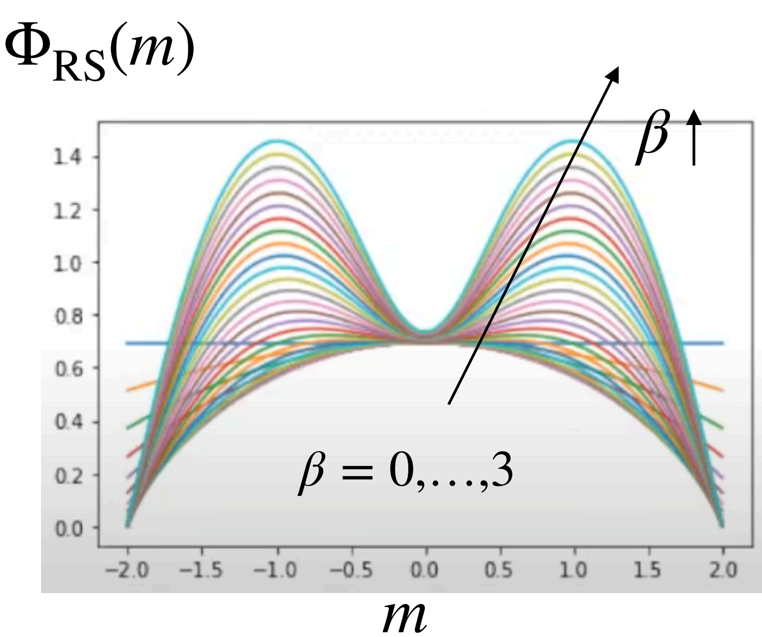
"Replica Symmetric Potential"

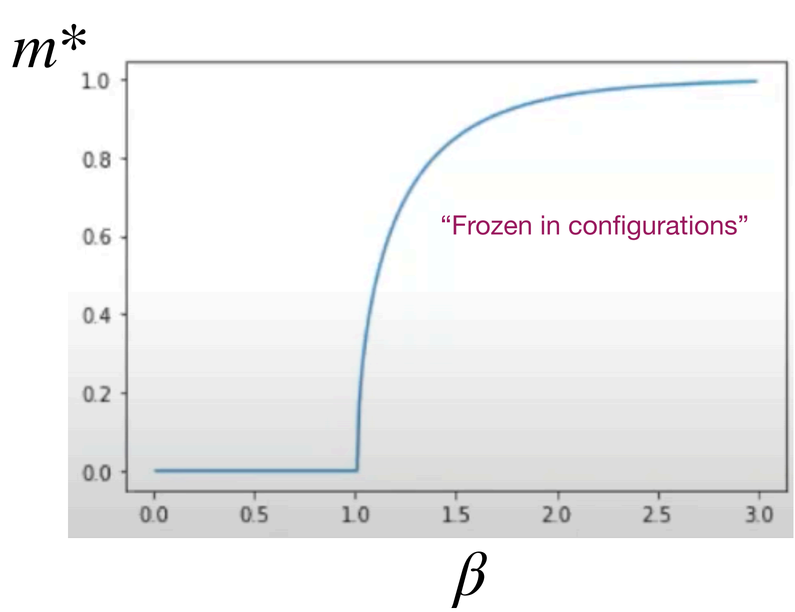
The extremum point satisfies

$$m^* = \mathbb{E}\left[\tanh\left(\beta\left(h + m^*\right)\right)\right]$$

Phase Transitions in the Random Field Ising Model

$$h \sim \mathcal{N}(0, \Delta = 0.1)$$





From the identity $\mathbb{E}[\log Z] = \lim_{n \to 0} \dfrac{\mathbb{E}\left[Z^n\right] - 1}{n}$ we obtain

$$\mathbb{E}\left[\frac{1}{N}\log Z_N\right] \to \mathrm{Extr}_m \left\{\Phi_{\mathrm{RS}}(m)\right\} \qquad , \Phi_{\mathrm{RS}}(m) \equiv -\beta\frac{1}{2}m^2 + \mathbb{E}\left[\log\left(2\cosh(\beta(h+m))\right)\right]$$

"Replica Symmetric Potential"

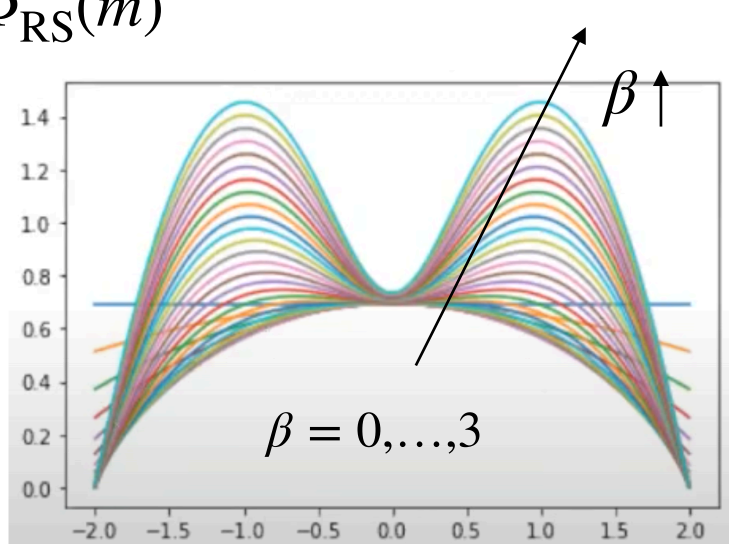
The extremum point satisfies

$$m^* = \mathbb{E}\left[\tanh\left(\beta\left(h + m^*\right)\right)\right]$$

Phase Transitions in the Random Field Ising Model

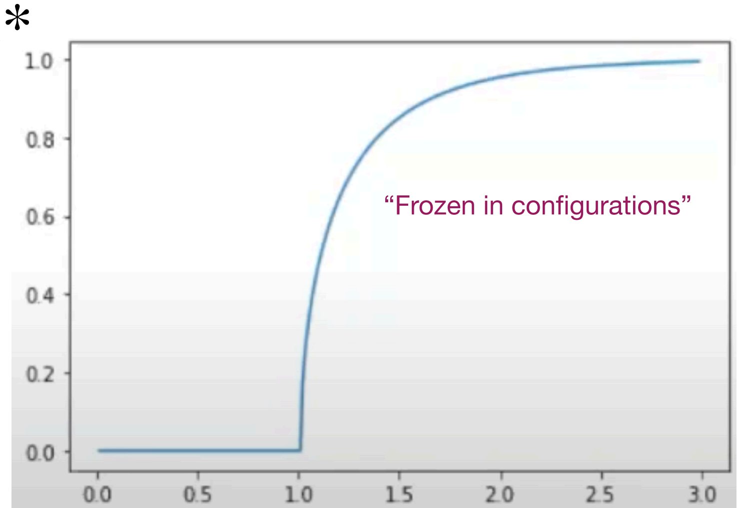
$$h \sim \mathcal{N}(0, \Delta = 0.1)$$

$\Phi_{\mathrm{RS}}(m)$



m

m^*

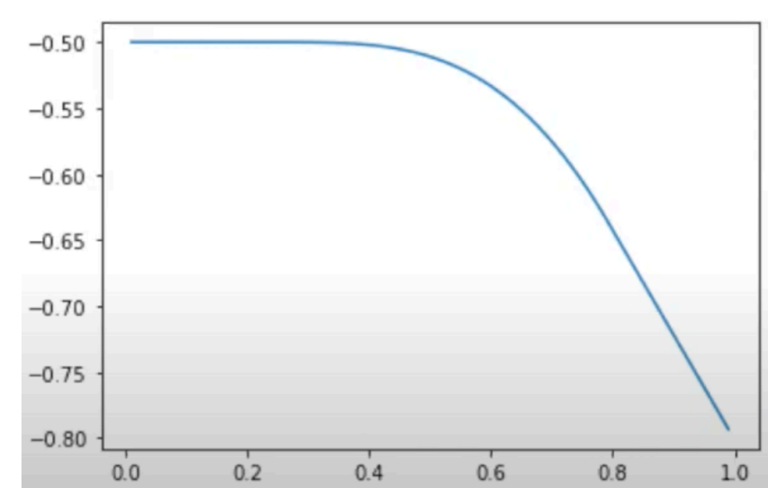


β

e^{min} vs variance Δ of *the random field*

Disorder-induced Freezing

$$\beta \to \infty$$



Δ