From Spin Glasses to Deep Networks

Quantum Connections Summer School
11.-12. June 2024

John Hertz
Nordita and Niels Bohr Institute

Outline

I. The prototypical spin glass: the Sherrington-Kirkpatrick model

II. Another kind of model: the p-spin glass

III. Neural networks and how to train them

IV. The learning transition in deep recurrent networks (and a quantum connection)

The Sherrington-Kirkpatrick Model

$$E = -\frac{1}{2}\sum_{ij} J_{ij}S_iS_j - b\sum_i S_i \qquad S_i = \pm 1$$

J_{ij} independent and Gaussian with mean 0 and variance $J^2/(N-1)$

"infinite-range" Ising model: Mean field theory exact (but nontrivial!)

Heuristic theory $(b = 0)$: assume frozen magnetisations $m_i = \langle S_i \rangle$

net field on spin i: $h_i = \sum_j J_{ij}m_j, \quad \Longrightarrow m_i = \tanh \beta h_i \quad (\beta = 1/T)$

order parameter (Edwards & Anderson) $q = \mathrm{var}(m_i)$

$$\mathrm{var}(h_i) = \sum_{j \neq i} \frac{J^2}{N-1} q = J^2 q \quad \Longrightarrow q = \int \frac{dh}{\sqrt{(2\pi J^2 q)}} \mathrm{e}^{-\frac{h^2}{2J^2q}} \tanh^2 \beta h$$

Single spin in a field with self-consistently determined variance

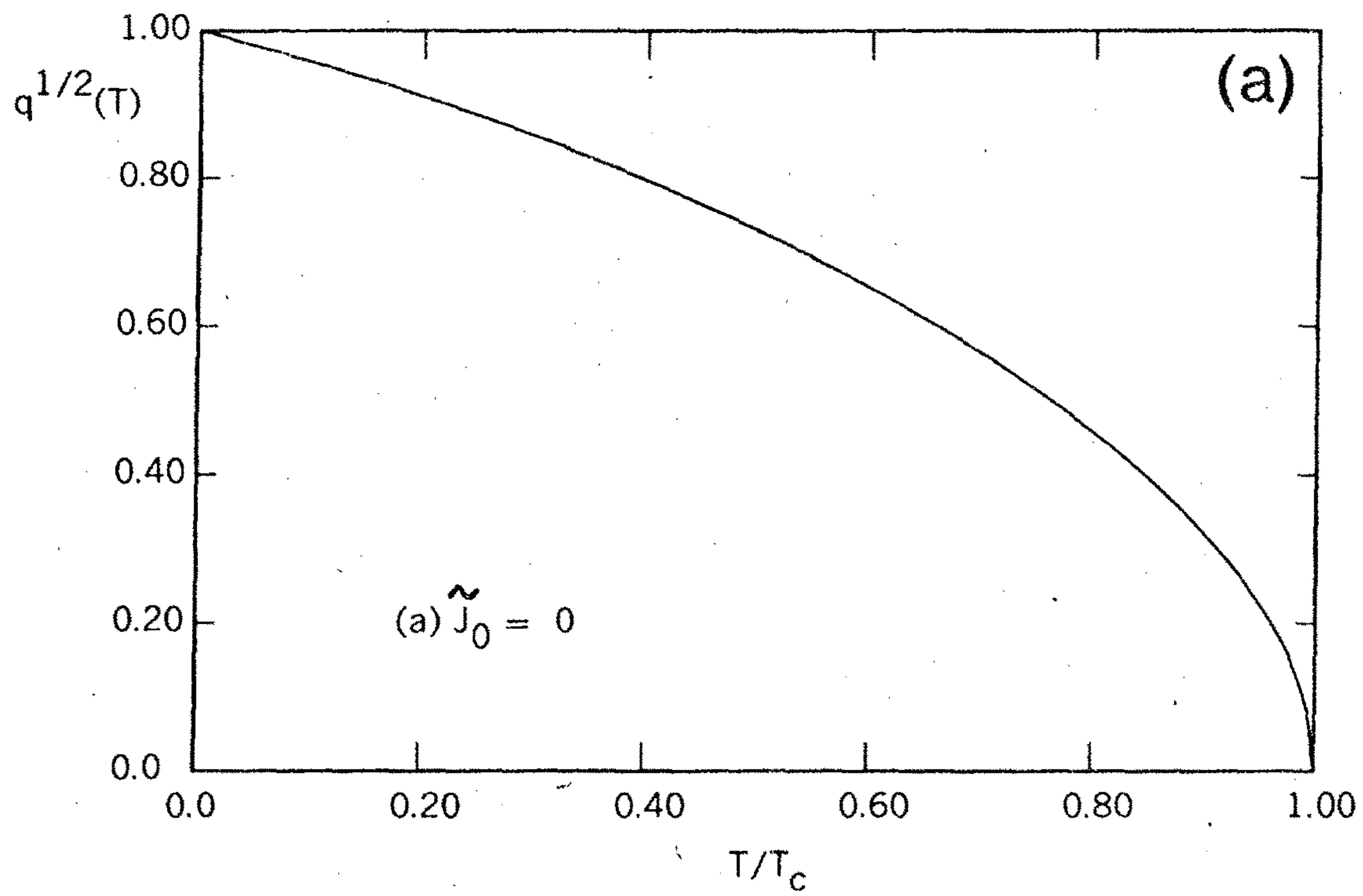
The phase transition and the glass phase:

Expand the tanh:

$$q = (\beta J)^2 q - 2(\beta J)^4 q^2 + \cdots$$

nontrivial solutions for $T < T_g = J$

$$q = \frac{T_g - T}{T_g} + \mathrm{O}((T_g - T)^2)$$



(from Sherrington and Kirkpatrick, Phys Rev B17, 4384-4403 (1978))

But this solution is wrong!

first problem: negative entropy at low T (S&K)

even more serious problem (de Almeida & Thouless): $\sum_{\langle ij \rangle} \langle S_i S_j \rangle^2 < 0$ everywhere below T_g !

Replicas!

For systems with quenched randomness, doing the stat mech requires *replicas.*

The partition function Z for a particular sample depends on the set of quenched variables $\mathbb{J} = \{J_{ij}\}$ for that sample: $Z = Z[\mathbb{J}]$.

But the relevant quantity to average over realisations of $\mathbb{J}$ is not $Z[\mathbb{J}]$, but $\log Z[\mathbb{J}]$

Use $\log Z[\mathbb{J}] = \lim_{n \to 0} \frac{Z[\mathbb{J}]^n - 1}{n}$, because $Z[\mathbb{J}]^n$ is easier to average than $\log Z[\mathbb{J}]$

Here, $$Z^n[\mathbb{J}] = \prod_a \sum_{\{S^a\}} \exp\left[\sum_{\langle ij \rangle} \beta J_{ij} S_i^a S_j^a\right]$$ (a: "replica index")

The averaged replica partition function

$$Z^n[\mathbb{J}] = \sum_{\{S^a\}} \exp\left[\sum_{\langle ij\rangle}\sum_a J_{ij} S_i^a S_j^a\right] = \prod_{\langle ij\rangle}\sum_{\{S_a\}} \exp\left[J_{ij}\sum_a S_i^a S_j^a\right]$$

From now on, set $\beta = 1$, i.e., measure J relative to T

Averaging: for each pair $\langle ij\rangle$, a 1-d integral: $\int dJ_{ij}\sqrt{\frac{N}{2\pi J^2}}\exp\left[-\frac{NJ_{ij}^2}{2J^2} + J_{ij}\sum_a S_i^a S_j^a\right]$

Complete the square in the exponent $\Longrightarrow$

$$[Z^n[\mathbb{J}]]_{\mathbb{J}} = \sum_{\{S^a\}} \exp\left[\frac{J^2}{2N}\sum_{\langle ij\rangle}\left(\sum_a S_i^a S_j^a\right)^2\right] = \sum_{\{S^a\}} \exp\left[\frac{J^2}{2N}\sum_{ab}\sum_{\langle ij\rangle} S_i^a S_i^b S_j^a S_j^b\right]$$

Replica energy: $E = -\frac{J^2}{2N}\sum_{ab}\sum_{\langle ij\rangle} S_i^a S_i^b S_j^a S_j^b$

a model with no disorder!

The order parameter

In a ferromagnet, the order parameter is $m = \langle S_i \rangle$

Here, it is $q_{ab} = \langle S_i^a S_i^b \rangle$ A matrix! (and how to take the limit $n \to 0$?)

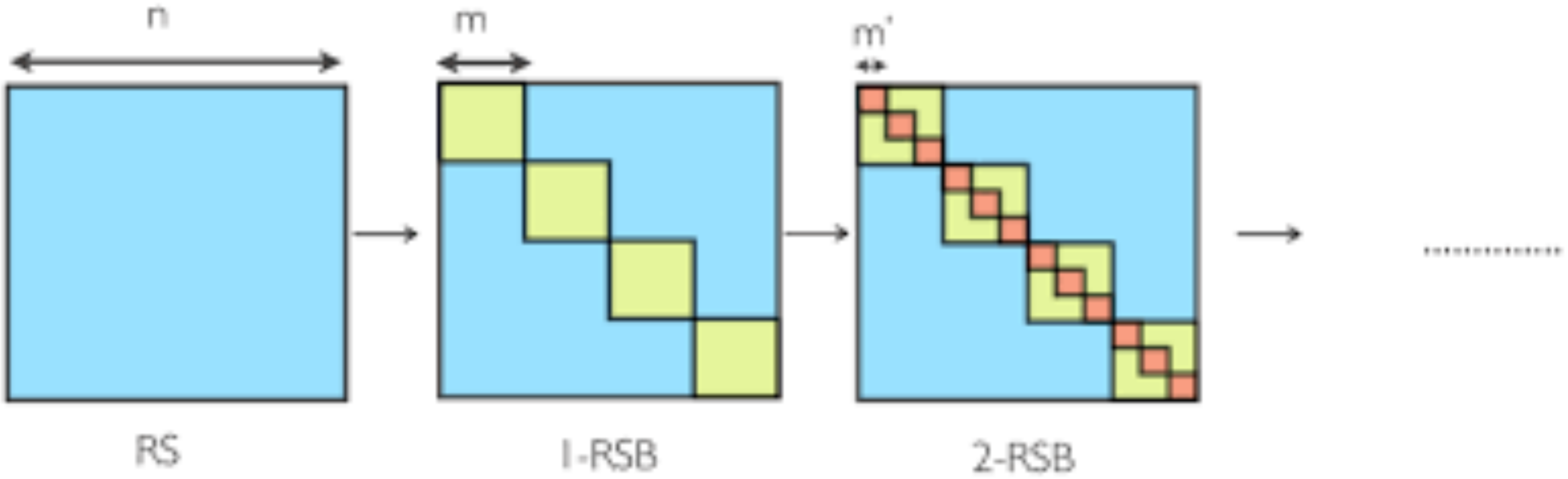
S&K assumed "replica symmetry": $q_{ab} = q$ $(a \neq b)$
(sounds innocent)

leads back to $\Longrightarrow q = \int \frac{dh}{\sqrt{(2\pi J^2 q)}} \mathrm{e}^{-\frac{h^2}{2J^2 q}} \tanh^2 h$

and negative entropy and other unacceptable features

Replica symmetry breaking is necessary below T_g (Parisi and others)

RSB is hierarchical



Langevin dynamics

The simplest model: harmonically bound fully overdamped particle driven by Gaussian white noise:

$$\dot{x} = -rx + h(t) + \xi(t) \qquad \langle \xi(t)\xi(t') \rangle = 2T\delta(t-t')$$

$h(t)$: external driving field, used to define the response function (susceptibility)

$$G(t-t') = \frac{\delta \langle x(t) \rangle}{\delta h(t')}$$

Here, $G^{-1} = \frac{d}{dt} + r$, i.e., $G_0(t-t') = \Theta(t-t')\mathrm{e}^{-r(t-t')}$ or $G(\omega) = (-\mathrm{i}\omega + r)^{-1}$

Correlation function

To calculate $C(t-t') = \langle x(t)x(t')\rangle$,   Fourier transform the eqn of motion $(h=0)$:

$$[-\mathrm{i}\omega + r]x(\omega) = \xi(\omega)$$

$$C(\omega) = \langle |x(\omega)|^2\rangle = \frac{\langle |\xi(\omega)|^2\rangle}{\omega^2+r^2} = \frac{2T}{\omega^2+r^2} \quad (=2T|G(\omega)|^2)$$

$$\Longrightarrow C(t-t') = \frac{T}{r}\mathrm{e}^{-r|t-t'|}$$

Note: $C(t=t') = \dfrac{T}{r} = TG(\omega=0)$

(fluctuation-response relation in equilibrium stat mech)

Fluctuation-dissipation theorem (FDT)

General relation between G and C, valid for *any* model that satisfies detailed balance

$$TG(t) = -\dot{C}(t) \quad (t > 0) \qquad TG(-t) = \dot{C}(t) \quad (t < 0)$$

Correlations in spontaneous correlations decay in the same way as relaxation after external perturbation.

Fourier transform:

$$T\frac{G(\omega) - G(-\omega)}{\mathrm{i}\,\omega} = \frac{2T}{\omega}\mathrm{Im}\,G(\omega) = C(\omega)$$

(Here, can check this using $\mathrm{Im}\,G(\omega) = \omega/(\omega^2 + r^2)$.)

A linear model

Consider the model (Gaussian spins)

$$\dot{S}_i = -r_0 S_i + \sum_j J_{ij} S_j(t) + h_i(t) + \xi_i(t) \qquad \langle \xi_i(t)\xi_j(t')\rangle = 2T\delta_{ij}\delta(t-t')$$

with SK couplings $\mathrm{var}(J_{ij}) = \dfrac{J^2}{N-1} \quad \forall i, j$

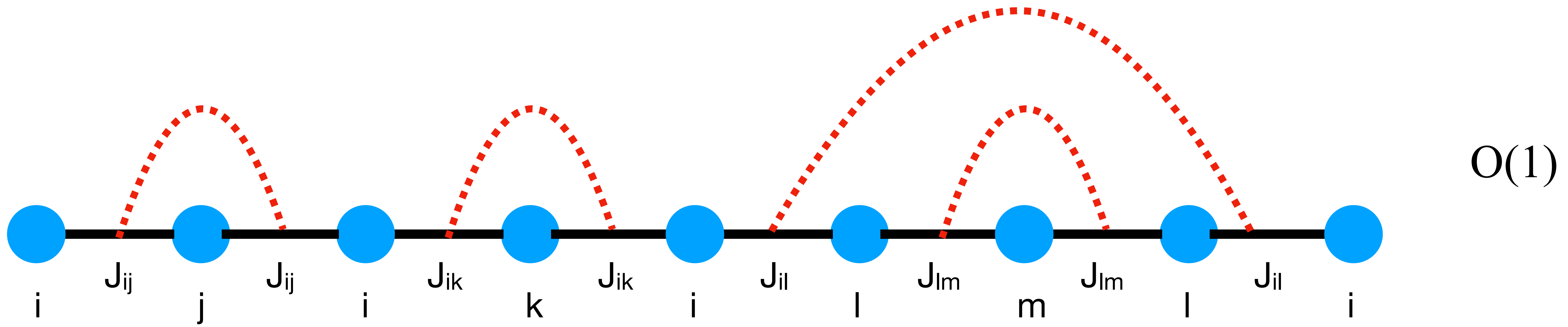
response function (unaveraged): $g_{ij}^{-1}(\omega) = (-\mathrm{i}\omega + r_0)\delta_{ij} - \sum_j J_{ij}$

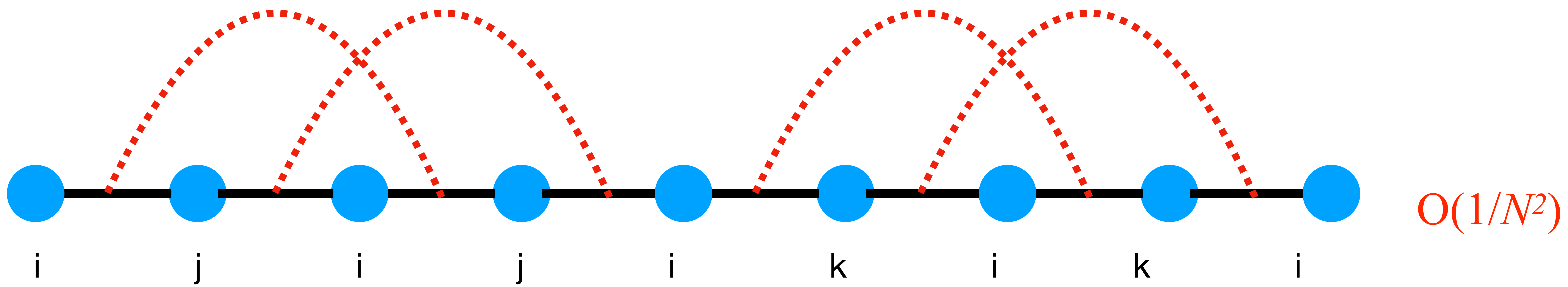
Want averaged single-site response and correlation functions G and C

Expand $g_{ii}(\omega)$: $g_{ii}(\omega) = g_0 + g_0 \sum_j J_{ij} g_0 J_{ji} g_o - g_0 \sum_{jk} J_{ij} g_0 J_{jk} g_0 J_{ki} g_0 + \cdots$

$$g_0 = g_o(\omega) = \frac{1}{-\mathrm{i}\omega + r_0}$$

In pictures:





From Dyson equation

$G(\omega) = G_0(\omega) + G_0(\omega)\Sigma_0(\omega)G(\omega)$, we have
(Σ_0 is all diagrams that can't be disconnected by cutting out a G_0 circle)

$$\Sigma_0(\omega) = J^2 G(\omega),$$

i.e.,
$$G(\omega) = \frac{1}{-i\omega + r_0 - J^2 G(\omega)}$$
(quadratic equation for $G(\omega)$)

In time domain, $G(t - t') = G_0(t - t') + \int_{t_1}^{t} dt_2 \int_{t_2}^{t_1} dt_2 G_0(t - t_1)\Sigma_0(t_1 - t_2)G(t_2 - t')$

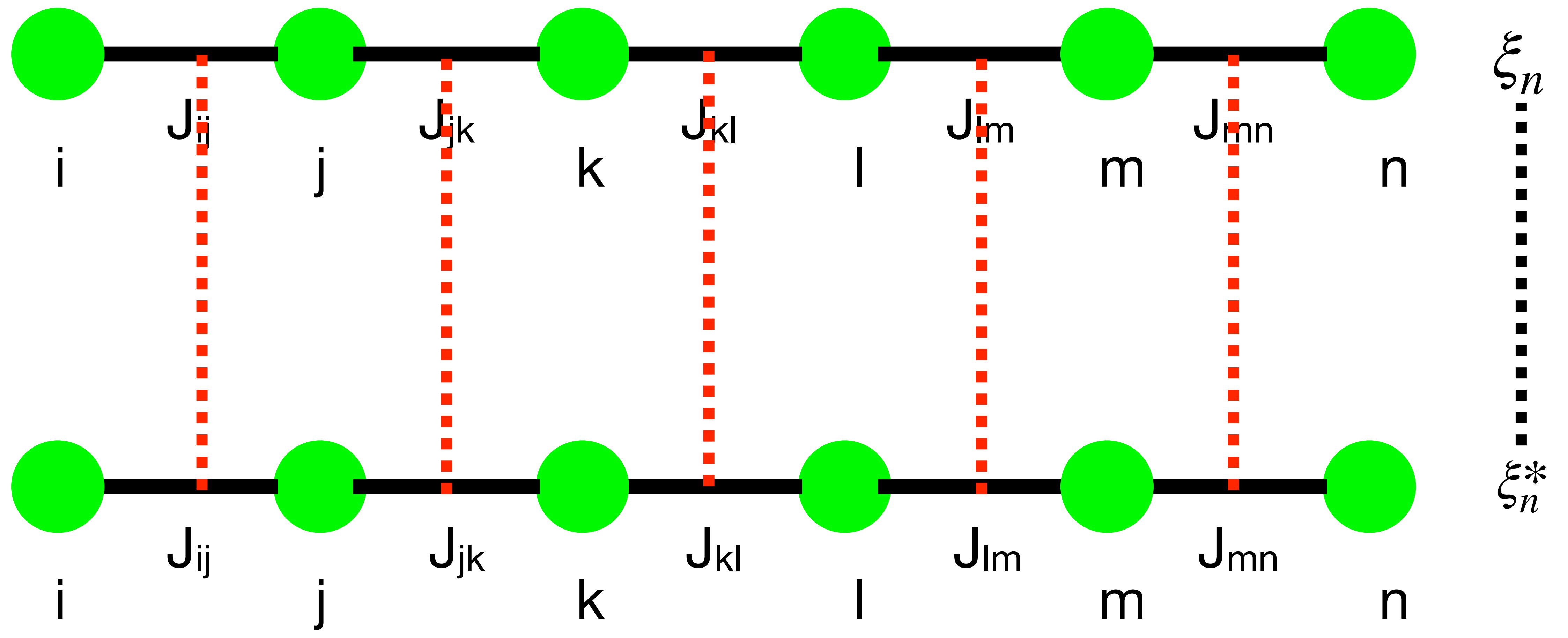
$\Sigma_0(t_1 - t_2)$ is effectively a delayed self-interaction

Correlation function:

(unaveraged, $h = 0$) $S_i(t) = \sum_j \int dt' g_{ij}(t-t')\xi_j(t')$

square this equation, take Fourier transform and average:

$$\langle C_{ii}(\omega)\rangle_{\mathbb{J},\xi} = \langle\, | g_0\xi_i - g_0 \sum_j J_{ij} g_0 \xi_j + g_0 \sum_{jk} J_{ij} g_0 J_{jk} g_0 \xi_k + \cdots |^2 \rangle_{\mathbb{J},\xi}$$



$$\Longrightarrow C(\omega) = \frac{2T|G(\omega)|^2}{1 - J^2|G(\omega)|^2}$$

Could have got this result more easily:

We could have got this result from the FDT: $\dfrac{2T}{\omega}\mathrm{Im}\, G(\omega) = C(\omega)$:

$$\mathrm{Im}\, G(\omega) = \mathrm{Im}\left(\frac{1}{-\mathrm{i}\omega + r_0 - J^2 G(\omega)}\right) = |\,G(\omega)\,|^2\,[\omega + J^2 \mathrm{Im}\, G(\omega)]$$

so
$$\mathrm{Im}\, G(\omega) = \frac{\omega\,|\,G(\omega)\,|^2}{1 - J^2\,|\,G(\omega)\,|^2}$$

and, using the FDT, $2T\dfrac{\mathrm{Im}\, G(\omega)}{\omega} = \dfrac{2T\,|\,G(\omega)\,|^2}{1 - J^2\,|\,G(\omega)\,|^2}$, in agreement with previous slide

Summary:

After averaging, the problem is reduced exactly to a single-site self-consistent problem with a a retarded self-interaction

$$\Sigma_0(t - t') = J^2 G(t - t')$$

and an effective (non-white) noise with variance

$$\langle \xi(t)\xi(t') \rangle = 2T\delta(t - t') + J^2 C(t - t')$$

Soft-spin model (Sompolinsky & Zippelius)

$$\dot{S}_i = -V'(S_i) + \sum_j J_{ij} S_j(t) + h_i(t) + \xi_i(t) \qquad \langle \xi_i(t) \xi_j(t') \rangle = 2T\delta_{ij}\delta(t - t')$$

with $V(S) = -\frac{1}{2} r S^2 + \frac{1}{4} u S^4$ (double-well potential)

(deep well limit: $r = u \to \infty$: Ising variables)

perturbation theory in u : $\Longrightarrow$ lots more Feynman diagrams

But our previous arguments leading to an effective single-site problem with retarded self-interaction $J^2 G(t - t')$ and effective noise $2T\delta(t - t') + J^2 C(t - t')$ still hold. The only difference is that G and C are now different because of the nonlinearity.

New G, relaxation rate

get a new self-energy contribution $\Sigma_1(\omega)$ from perturbation theory in $u \Longrightarrow$

New self-consistent equation for $G(\omega)$: $$G(\omega) = \frac{1}{-\mathrm{i}\omega - r - J^2 G(\omega) - \Sigma_1(\omega)}$$

Relaxation time:

$$\tau(\omega) \equiv \mathrm{i}\frac{\partial G^{-1}(\omega)}{\partial \omega} = 1 - \mathrm{i}J^2 \frac{\partial G(\omega)}{\partial \omega} - \mathrm{i}\frac{\partial \Sigma_1(\omega)}{\partial \omega}$$

so, $$\tau(\omega) = 1 + J^2 G^2(\omega)\tau(\omega) - \mathrm{i}\frac{\partial \Sigma_1(\omega)}{\partial \omega}$$

and solving for γ, $$\tau(\omega) = \frac{1 - \mathrm{i}\partial \Sigma_1 / \partial \omega}{1 - J^2 G^2(\omega)}$$

Critical slowing down approaching T_g from above:

Low-frequency/long-time limit: $\omega \to 0$: $-\mathrm{i}\partial\Sigma_1(\omega)/\partial\omega$ is finite

so $$\tau(0) \propto \frac{1}{1 - J^2 G^2(0)} \propto \frac{1}{T^2 - T_g^2}$$

with $T_g = JC(t=0),$ using FDT $TG(\omega = 0) = C(t=0)$.

(Ising spin limit: $C(t=0) = 1$)

This is all for $T > T_g$.

Critical slowing down right at T_g

Using $\tau(\omega) = \dfrac{1 - \mathrm{i}\partial\Sigma_1/\partial\omega}{1 - J^2G^2(\omega)}$ with $G^{-1}(\omega) = G^{-1}(0) - \mathrm{i}\omega\tau(\omega)$

leads (with $JG(0) = 1$ and $b \equiv 1 - \mathrm{i}d\Sigma_1(\omega)/d\omega\,|_{\omega=0}$) to

$$\tau = \frac{b}{1 - \dfrac{1}{(1 - \mathrm{i}\omega G(0)\tau)^2}}$$

Keeping terms of lowest order in $\omega\tau \Longrightarrow$ $\tau = \left(\dfrac{b}{-2i\omega G(0)}\right)^{1/2}$

Then the FDT implies $C(\omega) \propto |\omega|^{-1/2}$ and $C(t) \propto |t|^{-1/2}$

A little above T_g, C has $t^{-1/2}$ power law decay followed by crossover to $\exp(-t/\tau)$ for $t \gg \tau$.

What about $T < T_g$?

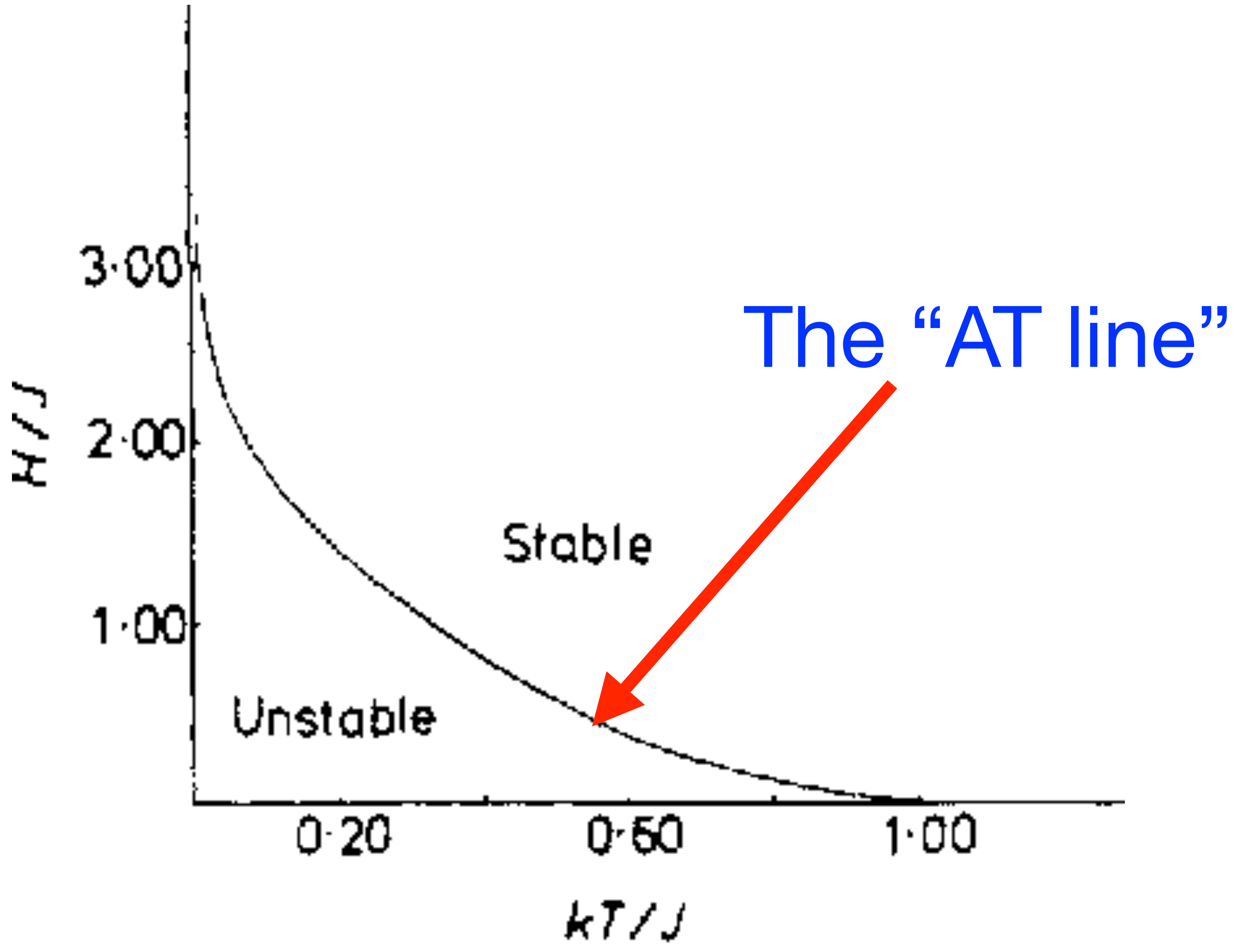
It depends on the history of the system (how it was cooled from disordered state).

For sudden quenches, the above theory can be extended to $T < T_g$ for times less than the age of the system (done by Sompolinsky and Zippelius).

For times long compared to the age of the system, complicated dynamics with many relaxation rates entering into the dynamics in succession (Sompolinsky). Theory is related to Parisi replica-symmetry-breaking solution for equilibrium theory.

SK glass in an external field

Everything in this lecture can be extended to systems in an external field.
T_g is now field-dependent.



(from de Almeida and Thouless, 1978)

$C(\omega) \propto \omega^{-\nu(H)}$ on the AT line

$$\nu(H) < 1/2, \;\; H > 0$$