Dynamics of Chaotic Random Neural Networks: Theoretical Methods

Kuljit S. Virk

March 10, 2024

Abstract

These notes describe my own version of the various theoretical methods I learned to study random neural networks as an instance of spin glasses. My interest is not in spin glass itself, but the use of this model in understanding various random networks in neuroscience. Therefore, the motivation of spin glass model does not arise from the solid state systems, but from the cable equation describing the equivalent circuit model of a neuron and its interconnection to other neurons. I derive the mapping to the basic spin glass model starting from an equivalent circuit model of neurons. I then digress into the fundamentals of path integral method to analyze dynamical equations and how to derive effective dynamical equations of the mean field. After the derivation of the mean field equation, I derive the equations of motion for the second-order correlation function and the special case when that equation can be described in terms of a Newtonian system with a potential function. Following this, I very briefly review the theory of Lyapunov exponents that is essential to analyzing the stability of the mean field equations and to understand its various chaotic regimes. The theoretical discussion is accompained by numerical simulations and calculations that illustrate the use of these methods and how the asymptotics that they describe emerges in the large N limit.

1 Introduction

The mathematical model of a neural network starts by abstracting the neuron to a state function that evolves in time and takes continuous values in a finite interval. The state of the i^{th} neuron is denoted by $S_i(t)$, where $-1 \leq S_i(t) \leq 1$, and $i = 1 \ldots N$. Each neuron receives inputs from one or more neurons on the network. The stimulation from all these neurons defines a *field* $h_i(t) \in (-\infty, +\infty)$ at the location of the neuron i. This field represents the cumulative stimulation that forms the input to the neuron. The stimulation h is a sum of the outputs from many neurons, and is thus unbounded, while the state function S is bounded.

The first modeling assumption is to relate the state function $S_i(t)$ to the stimulation $h_i(t)$ with a *neuron response function* or also called a *switching function* $\phi(x)$ so that

$$S_i(t) \quad = \quad \phi(h_i(t)), \tag{1}$$

where $\phi(x) \in [-1, +1]$, and typically $\phi(x) = \tanh(gx)$ with g being a *parameter* that controls the degree of non-linearity of the response of a neuron. At a deeper phenomenological modeling with membrane potentials (inputs) and firing rates (outputs), the dynamics of the fields h_i is mapped to an equivalent circuit model in which the total input from other neurons is linearly combined with couplings J_{ij}. Using the mapping (1), the second major modeling assumption yields a system of first order differential equations for the dynamics of h_i,

$$\frac{dh_i}{dt} \quad = \quad -h_i + \sum_{j=1}^{N} J_{ij}\phi(h_i). \tag{2}$$

Here J_{ij} represent the *synaptic efficacy* or the coupling of the output of the *presynaptic* neuron j to the input of the *postsynaptic* neuron i. Also, there is no self-coupling, *i.e.* $J_{ii} = 0$, and in addition, J_{ij} are random. In general, J_{ij} do not depend on the distance between neurons in the physical space, which makes the model infinite-ranged in general. However, for a particular case where that is important, the matrix elements J_{ij} would be modeled to have a decay with increasing distance between neurons i and j.

It is crucial to note that the random couplings J_{ij} are *quenched*. They are entirely random but remain fixed in time. A given system will correspond to a sample of $\{J_{ij}\}$ drawn from the underlying distribution, but the system itself will follow *deterministic evolution* in time with the given couplings. These systems are considered closed and their models do not include couplings with a bath. A statistical study of this topic therefore concerns averages over infinite set of systems, each realizing a sample of the couplings. This so-called *quenched averaging* is in contrast to *annealing* in which the couplings, and other parameters of the system evolve dynamically due to their interaction with the environment. This interaction is modeled using, for example, the Langevin equations or Lindblad formalsim. The interaction leads to *stochastic evolution* in time, and the statistics of the time-series corresponding to various physical observables is studied. Averaging over different systems may be performed, but will add no new information when the ergodic hypothesis holds.

There are several additional implicit assumptions made already in arriving at (2). Each neuron is assumed to have the *same* degree of non-linearity g. This simplifies the analysis since g is a parameter that plays a crucial role in determining the nature of the dynamics of $h_i(t)$ and therefore reduces the dimensionality of the phase space to be investigated. Similarly, the dynamical equation originates from a physical model in which time carries dimensions. Here time is an evolution parameter that is rendered dimensionless by scaling it to units of RC constant of neurons. Thus the capacitive and resistive coupling across the membrane of each neuron is also assumed to be nearly identical in order to arrive at such a simplified model. Finally, the couplings J_{ij} are drawn from the same underlying probability as well.

In a subsequent set of notes, I will relax these assumptions and show the impact of deviating away from them. Here these assumptions mean that the model applies to a to a single cluster of closely related neurons, which itself becomes a node in a larger network of clusters of neurons. From an opposite approach, one can decompose a network into coupled clusters of neurons, with high intra-cluster homogeneity and heterogeneity encoded in cluster-cluster couplings. This will be addressed in a later set of notes. Here, my focus is to establish as rigorously and in as much detail as I can the path to fully understanding the dynamics governed by (2).

Having recognized the assumption of homogeneity of the neuron properties within the network, a further step is taken in which the couplings J_{ij} are assigned random values from a Gaussian distribution,

$$P(J_{ij}) \quad = \quad \frac{1}{\sqrt{2\pi}J} e^{-J_{ij}^2/2J^2}. \tag{3}$$

The modeling of this probability distribution as a Gaussian has a deeper justification for the $N \to \infty$ limit. The justification arises from the analysis by Kirkpatrick and Sherrington [1] who showed that under this limit, physical and non-trivial dynamics requires $J \sim N^{-1/2}$, which survives under scaling of the Lagrangian in the derivation of effective dynamics, while higher order correlations' contributions vanish.

Thus the dynamics of the system is governed by two real-valued parameters (g, J) as well as the initial conditions, $h_i(0)$ or equivalently the neuron state variables $S_i(0)$. When the dynamical analysis is performed in the limit of large N and time t, the important parameter space is one-dimensional according to the value of the product gJ while the initial conditions are encoded into another one-dimensional parameter, the Edwards Anderson order parameter that I will denote by, Δ_0. Thus the mathematical analysis of this system's dynamics for $N \to \infty$ and $t \to \infty$ will be performed in a two-dimensional space spanned by the tuple: (gJ, Δ_0).

2 Dynamical Mean Field Theory

The long time properties of the system have been studied using the dynamical mean field theory (DMFT). The average in the mean field theory is over all possible couplings J_{ij}, which is equivalent to all possible networks described by the Gaussian distribution in (3). The "mean field" does not correspond to the spatial average of the state variables.

In fact, since the system governed by (1)-(3) leads to random orientations of $S_i(t)$, it follows that their mean behavior is not interesting,

$$\bar{S}(t) \quad \equiv \quad \lim_{N\to\infty} \frac{1}{N} \sum_i S_i(t) = 0. \tag{4}$$

The randomness also implies that,

$$\frac{1}{2N^2} \sum_{i \neq j} S_i(t) S_j(t) \quad = \quad 0. \tag{5}$$

However the correlation in time for a single neuron does not vanish,

$$C(\tau) \equiv \lim_{t\to\infty} \frac{1}{N}\sum_{i=1}^{N} \langle S_i(t)S_i(t+\tau)\rangle \neq 0, \tag{6}$$

where $\langle\cdot\rangle$ denotes average over all possible realizations of the network couplings J_{ij}. Equations (4)-(6) can be interpreted as a *freezing* of the network into a state of completely random variables $S_i(t)$ with zero mean. If a neuron at time t is found to be in a particular state, then (6) implies that with high probability, it will be found in the *same state* at a later time $t+\tau$. Thus the network consists of neurons whose states are completely random but fixed in time, or vary over some time scale. The randomness and slow variation is where the glassy nature of the network arises, and hence its analysis as a spin glass. The generality of this statement over all possible network dynamics is ensured by the average over the distributions of couplings, and therefore over all possible networks that can be formed with zero mean Gaussian random couplings.

The correlation function $C(\tau)$ defined in (6) is the ***Edwards-Anderson order parameter*** introduced in [2] that often forms the starting point of various analyses of spin glasses and related networks [1, 3, 4]. The time dependence of Δ describes how the state potentially loses memory over time since $\Delta(\tau)\leq\Delta_0$. A constant $\Delta(\tau)$ describes a truly frozen state while a decreasing $\Delta(\tau)$ describes a neuron that loses correlation to its own state at a long time and thus a vanishing memory of its state over time. Many other types of solutions are possible depending on the paramters of the dynamical equation, and will be explored in the sections below.

After averaging over all J_{ij} in (6), the time correlation is expected to be independent of the neuron and thus, using (1) it is possible to write the order parmeter as,

$$C(\tau) \equiv \langle \phi(h_i(t))\phi(h_i(t+\tau))\rangle\,, \tag{7}$$

where stationarity in the long time limit is implied due to the independence of C on t on the right hand side. Finally since the dynamical model is stated directly in terms of the fields and not the state variables, it is more convenient to study the order parameter

$$\Delta(\tau) = \langle h_i(t)h_i(t+\tau)\rangle\,. \tag{8}$$

With these quantities, the mean field limit of (2) replaces the second term in (2) with a randomly fluctuation field. The effective equation for the mean field is derived in section §5, and it gives the equation for the dynamics of the field in the form,

$$\frac{d}{dt}h_i = -h_i(t)+\eta_i(t), \tag{9}$$

where $\eta_i(t)$ is a Gaussian random field given by,

$$\eta_i(t) = \sum_j J_{ij}S_j(t),$$

Its time correlation is also derived explicitly from the functional formalism in section §5, but it can also be determined considering the average with the distribution $P(\{J_{ij}\})$ as,

$$\langle\eta_i(t)\eta_i(t+\tau)\rangle = \frac{1}{N}\sum_{j,k=1}^{N}\int\prod_{nm}\{dJ_{nm}P(J_{nm})\}\,J_{ij}J_{ik}\left[S_j(t)S_k(t+\tau)\right]_{\mathbf{J}}\,,$$

where P is the probability distribution (3), and the subscript "$\mathbf{J}$" indicates solutions computed with a fixed matrix $\mathbf{J}$. The integration over J_{nm} averages the product $S_j(t)S_k(t+\tau)$ over distributions of all J_{nm} except for J_{ij} and J_{jk} for which the integral is more complex. In the limit $N\to\infty$, it is clear that leaving out two of the couplings at a time should give essentially the same result as if all couplings were averaged over. Thus we can replace the integrals over all $N-2$ couplings with the correlation between $S_j(t)$ and $S_j(t+\tau)$, that is,

$$\langle\eta_i(t)\eta_i(t+\tau)\rangle = \frac{1}{N}\sum_{j=1}^{N}\int dJ_{ij}dJ_{jk}P(J_{ij})P(J_{ik})J_{ij}^2\,\langle S_j(t)S_j(t+\tau)\rangle\,.$$

As discussed above, the correlation is indepndent of i on the left hand side and j on the right hand side and thus can be written simply using (6) as

$$\begin{aligned}\langle \eta_i(t)\eta_i(t+\tau)\rangle &= C(\tau)\sum_{j=1}^{N}\int dJ_{ij}dJ_{jk}P(J_{ij})P(J_{ik})J_{ij}^2,\\ &= C(\tau)NJ^2,\end{aligned}$$

where we peform the integral over the two couplings with the probability distribution (3). Referring to the analysis in [1], define $\tilde{J} = \sqrt{N}J$, to obtain,

$$\langle \eta_i(t)\eta_i(t+\tau)\rangle = \tilde{J}^2 C(\tau). \tag{10}$$

The mean field equation (9) also yields the equation for the dynamics of $\Delta(\tau)$. Differentiating (8) with respect to τ and substituting (9) for $dh/d\tau$, it follows that

$$\frac{d\Delta}{d\tau} = -\langle h_i(t)h_i(t+\tau)\rangle + \langle h_i(t)\eta_i(t+\tau)\rangle .$$

Differentiating again, and using the relation $h_i = \eta_i - \dot{h}_i$,

$$\begin{aligned}\frac{d^2\Delta}{d\tau^2} &= \langle h_i(t)h_i(t+\tau)\rangle - \langle h_i(t)\eta_i(t+\tau)\rangle + \left\langle h_i(t)\frac{d}{d\tau}\eta_i(t+\tau)\right\rangle\\ &= \langle h_i(t)h_i(t+\tau)\rangle - \langle \eta_i(t)\eta_i(t+\tau)\rangle + \left\langle \dot{h}_i(t)\eta_i(t+\tau)\right\rangle + \left\langle h_i(t)\frac{d}{d\tau}\eta_i(t+\tau)\right\rangle .\end{aligned}$$

Noting that $d\eta(t+\tau)/d\tau = d\eta(t+\tau)/dt$ in the last term,

$$\frac{d^2\Delta}{d\tau^2} = \langle h_i(t)h_i(t+\tau)\rangle - \langle \eta_i(t)\eta_i(t+\tau)\rangle + \frac{d}{dt}\langle h_i(t)\eta_i(t+\tau)\rangle .$$

The last term is zero by assumption of stationarity. Therefore, substituting (8) and (10), it follows that

$$\frac{d^2\Delta}{dt^2} = \Delta(t) - \tilde{J}^2 C(t). \tag{11}$$

This is the main equation for analysis of mean field theory and can be found in the work by Sompolinski [4]. The right hand side of this equation $C(t)$ depends on Δ itself. Therefore the equation must be solved self-consistently. In fact, it follows from the analysis in the next section that C is a function only of $\Delta(t)$ and its equal-time value $\Delta_0 = \Delta(0)$, which implies that its time dependence is also arising purely from the time-dependnce of $\Delta(t)$. In the next section, the path integral formalism is used to evaluate C as the function

$$C = C(\Delta, \Delta_0).$$

By evaluating the anti-derivative of the right hand side with respect to Δ in (11), it follows that

$$\frac{d^2\Delta}{dt^2} = -\frac{\partial V}{\partial \Delta}. \tag{12}$$

The equation shows that the dynamics of Δ follows a force law in some abstract space in which Δ resides, and is thus conservative with a potential function defined by V,

$$V(\Delta) = -\frac{1}{2}\Delta^2 + \tilde{J}^2\int d\Delta\, C(\Delta, \Delta_0). \tag{13}$$

This form lends itself to the analysis of the dynamics of Δ in the parameter space $(g\tilde{J}, \Delta_0)$, as was first presented by Sompolinski [4]. Note that this result would not be possible if the switching non-linearity g was dependent on neurons, or had a distribution rather than being the same constant for all neurons. In fact, this form of the dynamics is very particular to otherwise identical neurons interacting via random couplings. A chemical environment

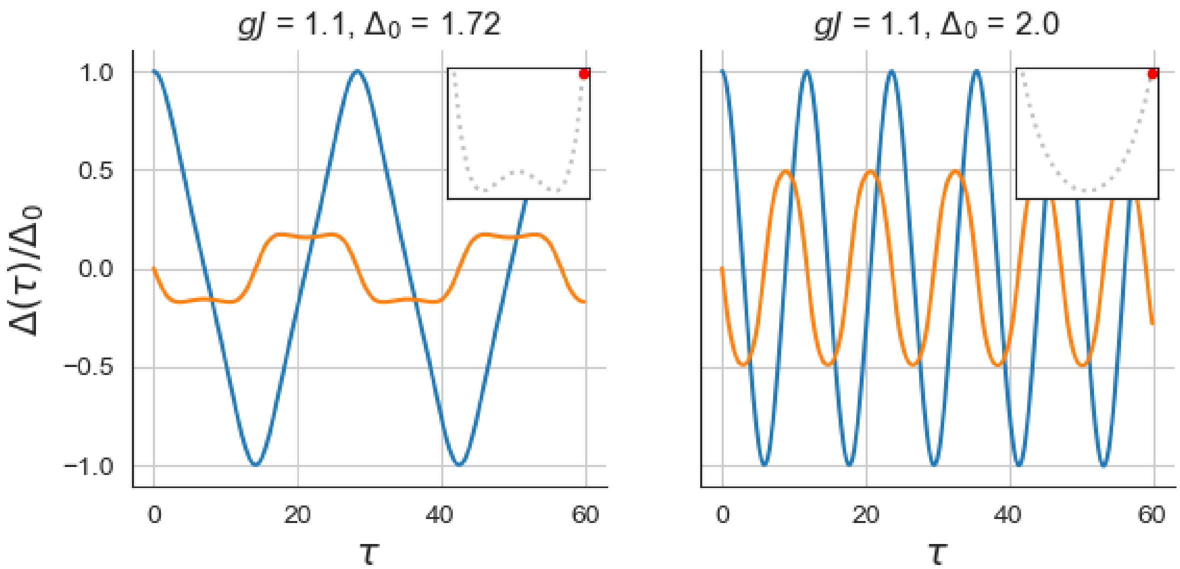

Figure 1: Dynamics of $\Delta(\tau)$ with the stable potentials shown in blue line, while the derivative $d\Delta/d\tau$ is shown in orange. The panel title shows the values of gJ and Δ_0, and the inset shows the corresponding $V(\Delta, \Delta_0)$. The red dot in the inset shows the initial value of the potential. The Jupyter Notebook for this calculation is ~/Research/RandomNeuralNets.ipynb.

surrounding real neurons will define a length scale over which such homegeneity can be assumed. Finally, since an equation for Δ exists in closed form, it is solved first and then $C(t)$ is determined by $C(t) = \Delta - \ddot{\Delta}$.

Figure 1 shows two examples of the dynamics of $\Delta(\tau)$ with $gJ > 1$ and two values Δ_0. The parameter $gJ > 1$ is the only one where the initial value $\Delta_0 \neq 0$ yields a bounded solution [4]. It is clear that the shape of the potential has a significant qualitative influeence on the EA parameter. In particular the case $\Delta_0 = 1.72$ shows an approximately square shaped $\dot{\Delta}(\tau)$, while a slightly larger value results in a more ordinary oscillatory function. With the value Δ_0, the field $h(t)$ switches linearly between full correlation to fully anti-correlated to its initial state, except near the extrema points where it changes the direction smoothly but rapidly.

3 Calculation of the potential

Introduce the Fourier transform,

$$\phi(x) = \int \frac{dk}{2\pi} e^{ikx} \hat{\phi}(k),$$

and the anti-derivative via the capitalized letters,

$$\Phi(x) = \int_0^x \phi(x')dx'.$$

Replace x by $h(t)$ and write the η-averaged time-correlation between arbitrary switching functions ϕ and ψ as,

$$\mathcal{C}(\Delta, \Delta_0, \phi, \psi) \equiv \langle \phi(h(t))\psi(h(t')) \rangle = \int \frac{dk}{2\pi} \int \frac{dk'}{2\pi} \hat{\phi}(k)\hat{\psi}(k') \left\langle e^{ikh(t)+ik'h(t')} \right\rangle. \quad (14)$$

The average is computed for a given t, t' and thus for a fixed $\tau = t - t'$. The values of the field at two different times then define two different variables $h_i(t)$ and $h_i(t+\tau)$ that depend on the particular sample of the stochastic field $\eta_i(t)$. We will consider fields with zero mean, and thus set

$$\begin{aligned} h_i(t) &= \alpha x + \kappa z, \\ h_i(t+\tau) &= \alpha y + \kappa z, \end{aligned}$$

where x, y, z are each normal random variables. It is clear that,

$$\alpha^2+\kappa^2 \quad = \quad \Delta_0 = \langle h_i(t)h_i(t)\rangle \,, \tag{15}$$

$$\kappa^2 \quad = \quad \Delta(\tau) = \langle h_i(t)h_i(t+\tau)\rangle \,. \tag{16}$$

The average over all J_{ij} is equivalent to performing an average over all possible h_i and thus all three variables x, y and z. Thus

$$\left\langle e^{ikh(t)+ik'h(t')}\right\rangle = \int dz\,\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}e^{i(k+k')\kappa z}\int dx\frac{e^{-\frac{1}{2}x^2+ik\alpha x}}{\sqrt{2\pi}}\int dy\frac{e^{-\frac{1}{2}y^2+ik'\alpha y}}{\sqrt{2\pi}} \tag{17}$$

$$= e^{-\frac{1}{2}(\alpha^2k^2+\alpha^2k'^2+\kappa^2(k+k')^2)} \tag{18}$$

$$= e^{-\frac{1}{2}\Delta_0(k^2+k'^2)-\Delta kk'}. \tag{19}$$

The above derivation defines two important relationships used below,

$$\left\langle e^{ikh(t)+ik'h(t')}\right\rangle = e^{-\frac{1}{2}\Delta_0(k^2+k'^2)-\Delta kk'} = \int dz\,\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}e^{i(k+k')\kappa z}\int dx\frac{e^{-\frac{1}{2}x^2+ik\alpha x}}{\sqrt{2\pi}}\int dy\frac{e^{-\frac{1}{2}y^2+ik'\alpha y}}{\sqrt{2\pi}}. \tag{20}$$

Note that the integration with respect to Δ results in replacing the fields by their anti-derivatives,

$$\begin{aligned}\int_0^{\Delta(\tau)} d\Delta'\,\langle\phi(h(t))\psi(h(t'))\rangle &= \int\frac{dk}{2\pi}\int\frac{dk'}{2\pi}\,\frac{\hat\phi(k)}{ik}\,\frac{\hat\psi(k')}{ik'}e^{-\frac{1}{2}\Delta_0(k^2+k'^2)-\Delta(\tau)kk'},\\ &= \int\frac{dk}{2\pi}\int\frac{dk'}{2\pi}\,\hat\Phi(k)\,\hat\Psi(k')e^{-\frac{1}{2}\Delta_0(k^2+k'^2)-\Delta(\tau)kk'},\\ &= \langle\Phi(h(t))\Psi(h(t'))\rangle\,.\end{aligned}$$

The derivative with respect to Δ follows a similar pattern,

$$\frac{d^n}{d\Delta^n}\langle\Phi(h(t))\Psi(h(t'))\rangle = \left\langle\Phi^{(n)}(h(t))\Psi^{(n)}(h(t'))\right\rangle.$$

The right hand side of (14) may now be written more explicitly and thus the function $\mathcal{C}$ can be written in terms of the switching functions and the correlations directly,

$$\mathcal{C}(\Delta,\Delta_0,\phi,\psi) = \int du\int dv\,\phi(u)\psi(\mathrm{sgn}(\Delta)v)\int\frac{dk}{2\pi}\int\frac{dk'}{2\pi}\,e^{-\frac{1}{2}\Delta_0(k^2+k'^2)-|\Delta(\tau)|kk'+iku+ik'v}, \tag{21}$$

Note that $\mathrm{sgn}(\Delta)$ can be moved to u instead of v. From the above analysis of the integrals and derviatives with respect to Δ, it follows also that

$$\int d\Delta\mathcal{C}(\Delta,\Delta_0,\phi,\psi) = \mathcal{C}(\Delta,\Delta_0,\Phi,\Psi).$$

We now recall the definitions (15) and (21), and the relations in (20) to write an alternative expression,

$$\begin{aligned}&\mathcal{C}(\Delta,\Delta_0,\Phi,\Psi)\\ &= \int dz\,\frac{e^{-z^2/2}}{\sqrt{2\pi}}\int dx\,\frac{e^{-x^2/2}}{\sqrt{2\pi}}\Phi\left(\sqrt{\Delta_0^2-\Delta^2}x+\sqrt{\Delta}z\right)\int dy\,\frac{e^{-y^2/2}}{\sqrt{2\pi}}\Psi\left[\mathrm{sgn}(\Delta)\left(\sqrt{\Delta_0-\Delta}y+\sqrt{|\Delta|}z\right)\right].\end{aligned}$$

The integrals are over $-\infty$ to $+\infty$.

An alternative expression follows by solving (14) integrals over k' and k. The integral over k' gives,

$$\int\frac{dk'}{2\pi}\,e^{-\frac{1}{2}\Delta_0\left(k'^2+\frac{2}{\Delta_0}(\Delta k-iy)k'\right)} = \sqrt{\frac{1}{2\pi\Delta_0}}\exp\left[\frac{\Delta^2k^2-v^2-2i|\Delta|kv}{2\Delta_0}\right]$$

Collect terms to prepare integration over k,

$$\begin{aligned}
I &= \sqrt{\frac{1}{2\pi\Delta_0}} \int \frac{dk}{2\pi} e^{-\frac{1}{2}\Delta_0 k^2 + ikx + \frac{\Delta^2 k^2 - v^2 - 2i|\Delta| kv}{2\Delta_0}} \\
&= \sqrt{\frac{1}{2\pi\Delta_0^2}} \int \frac{dk}{2\pi} \exp\left[\frac{1}{2\Delta_0}\left\{-\Delta_0^2 k^2 + \Delta^2 k^2 + 2i\left(\Delta_0 u - |\Delta| v\right) k - v^2\right\}\right] \\
&= \sqrt{\frac{1}{2\pi\Delta_0^2}} \int \frac{dk}{2\pi} \exp\left[\frac{-1}{2\Delta_0}\left\{(\Delta_0^2 - \Delta^2) k^2 - 2i\left(\Delta_0 u - |\Delta| v\right) k\right\} - \frac{v^2}{2\Delta_0}\right].
\end{aligned}$$

The integral over k is well defined because $\Delta^2 \leq \Delta_0^2$. Perform the integral by completing the squares and obtain,

$$\begin{aligned}
I &= \frac{1}{2\pi} \frac{1}{\sqrt{\Delta_0}} \frac{\sqrt{\Delta_0}}{\sqrt{\Delta_0^2 - \Delta^2}} \exp\left[-\frac{1}{2\Delta_0} \frac{\left(\Delta_0 u - |\Delta| v\right)^2}{(\Delta_0^2 - \Delta^2)} - \frac{v^2}{2\Delta_0}\right] \\
&= \frac{1}{2\pi\sqrt{\Delta_0^2 - \Delta^2}} \exp\left[-\frac{1}{2\Delta_0}\left\{\frac{\Delta_0^2 u^2 + \Delta^2 v^2 - 2\Delta_0 |\Delta| uv + \left(\Delta_0^2 - \Delta^2\right) v^2}{\Delta_0^2 - \Delta^2}\right\}\right] \\
&= \frac{1}{2\pi\sqrt{\Delta_0^2 - \Delta^2}} \exp\left[-\frac{1}{2\Delta_0}\left\{\frac{\Delta_0^2\left(u^2 + v^2\right) - 2|\Delta| uv}{\Delta_0^2 - \Delta^2}\right\}\right].
\end{aligned}$$

The correlation is now given by,

$$\mathcal{C}(\Delta, \Delta_0, \phi, \psi) = \int\int \frac{dxdy}{2\pi\sqrt{\Delta_0^2 - \Delta^2}} \exp\left[\frac{-\Delta_0\left(u^2 + v^2\right) + 2\,|\Delta|\,uv}{2\left(\Delta_0^2 - \Delta^2\right)}\right] \phi(u)\psi(\mathrm{sgn}(\Delta) v).$$

The integrals are again over $-\infty$ to $+\infty$.

4 Stability Analysis

For stability analysis, it is convenient to scale the fields to simplify the equations into more canonical form, by dividing the fields by $\tilde{J}$,

$$\begin{aligned}
h &\mapsto \frac{h}{\tilde{J}}, \\
\Delta(t) &\mapsto \frac{1}{\tilde{J}^2}\Delta(t).
\end{aligned}$$

Therefore, under this transformation, the dynamical equation becomes,

$$\begin{aligned}
\dot{h}(t) &= h(t) + \sum_{ij} \mathcal{J}_{ij} S_j(t), \\
\mathcal{J}_{ij} &= \frac{J_{ij}}{\tilde{J}}.
\end{aligned}$$

The equation for $\Delta(t)$ takes the form

$$\ddot{\Delta}(t) = -\frac{\partial V}{\partial \Delta},$$

where the potential is now written as

$$V(\Delta) = -\frac{1}{2}\Delta^2 + \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left[\int dx \frac{e^{-x^2/2}}{\sqrt{2\pi}} \frac{1}{\tilde{J}} \Phi(\tilde{J}\sqrt{\Delta_0 - \Delta}\, x + \tilde{J}\sqrt{\Delta} z)\right]^2.$$

Thus with the choice $\phi(x) = \psi(gx)$, and $\psi(x) = d\Psi/dx$,

$$V(\Delta) = -\frac{1}{2}\Delta^2 + \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left[\int dx \frac{e^{-x^2/2}}{\sqrt{2\pi}} \frac{1}{g\tilde{J}} \Psi(g\tilde{J}\sqrt{\Delta_0 - \Delta}\, x + g\tilde{J}\sqrt{\Delta} z) \right]^2.$$

This identifies $g\tilde{J}$ as one of the parameters. Next, define

$$r(t) = \frac{\Delta(t)}{\Delta_0}.$$

The potential is defined as $V(\Delta) = \Delta_0^2 W(r)$, where

$$W(r) = -\frac{1}{2}r^2 + U(r), \tag{22}$$

$$U(r) = \int dz \frac{e^{-z^2/2}}{\sqrt{2\pi}} W^2(r, z), \tag{23}$$

$$W(r, z) = \frac{1}{g\tilde{J}} \int dx \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, \Psi \left[g\tilde{J}\Delta_0^{1/2} \left(\sqrt{1-r}x + \sqrt{r}z \right) \right]. \tag{24}$$

The ODE for Δ can be written as

$$\Delta_0 \frac{d^2 r}{dt^2} = -\frac{\partial V(\Delta)}{\partial \Delta} = \Delta_0^2 \frac{\partial W(r)}{\partial r} \frac{\partial r}{\partial \Delta} = \Delta_0 \frac{\partial W(r)}{\partial r},$$

and rearranging the terms, this yields,

$$\frac{d^2 r}{dt^2} = -\frac{\partial W(r)}{\partial r} = r + \frac{\partial U}{\partial r} \tag{25}$$

$$r(0) = 1, \tag{26}$$

$$\dot{r}(0) = 0. \tag{27}$$

How does this ensure that $0 \leq r(t) \leq \Delta_0$? Since the equation corresponds to conservative motion, the value of the potential at initial time, where $\Delta/\Delta_0 = 1$ can be taken as the total energy. The total energy is conserved, and therefore Δ moves in a way that the total energy never exceeds $V(\Delta_0)$. The form of the potential above always yields some 4th or similar even order polynomial. This can also be seen by Taylor expansion. In that case the class of potentials is very clear and can probably be classified according to the number of minima etc and how Δ will move on them. From the potentials up to 4th order, it is clear that Δ decreases towards zero and if it crosses zero it is simply retracing itself back to Δ_0 on the symmetric portion, and will go no higher due to conservation law. Therefore conservation law and the topology of the potential is paramount to restricting all the solutions that can occur.

5 Functional formalism to derive mean field theory

Generating Functional

In this section, we will see how to evaluate $C(\tau)$ by a functional integral formalism, which also yields the mean field equation directly. The basic approach is to start with the generating functional for dynamical correlations. The functional is defined over all possible trajectories $h_i(t)$ as

$$Z[\mathbf{u}; \{J_{ij}\}] = \int \prod_i \left\{ Dh_i(t) \delta \left[\dot{h}_i(t) + h_i(t) - \sum_j J_{ij} S_j(t) \right] \right\} e^{\int dt \sum_i u_i(t) h_i(t)}. \tag{28}$$

The delta function ensures that the paths that contribute obey exactly the dynamical equation (2). The fields $u_i(t)$ are infinitesimal external couplings that are set to zero at the end of any derivation. First note that,

$$Z[0; \{J_{ij}\}] = 1,$$

$$\Rightarrow \left[\frac{1}{Z} \frac{\delta Z}{\delta u_i} \right]_{u_i=0} = \left[\frac{\delta \log Z}{\delta u_i} \right]_{u_i=0} = \left[\frac{\delta Z}{\delta u_i} \right]_{u_i=0}.$$

Thus due to the normalization, it is sufficient to consider derivatives of Z directly. Now, consider the derivatives,

$$\begin{aligned}\left[\frac{\delta Z}{\delta u_i}\right]_{u_i=0} &= \langle h_i(t)\rangle \\ \left[\frac{\delta^2 Z}{\delta u_i \delta u_j}\right]_{u_i=0} &= \langle h_i(t)h_j(t)\rangle\end{aligned}$$

Transform to Action

Next, substitute into (28), the representation of the delta functional in the form

$$\delta(x) = \int D\widehat{h}e^{i\widehat{h}x},$$

from which the following result follows fairly directly,

$$\begin{aligned} Z\left[\mathbf{u},\widehat{\mathbf{u}};\{J_{ij}\}\right] &= \int D\mathbf{h}D\widehat{\mathbf{h}}\; e^{\Omega_0[\mathbf{h},\widehat{\mathbf{h}}]+\int dt\,\mathbf{u}(t)\cdot\mathbf{h}(t)+i\widehat{\mathbf{u}}(t)\cdot\widehat{\mathbf{h}}(t)} \exp\left[i\int dt\,\widehat{h}_i(t)J_{ij}\phi(h_j(t))\right], \qquad (29)\\ \Omega_0[\mathbf{h},\widehat{\mathbf{h}}] &= -i\int dt\,\widehat{\mathbf{h}}(t)\cdot\left[\dot{\mathbf{h}}(t)+\mathbf{h}(t)\right],\end{aligned}$$

where where $\mathbf{J}$ is the matrix with elements J_{ij}. Since the path integrals are unrestricted, the variables $\mathbf{h}$ are independent of $\mathbf{J}$, while the restriction to the dynamics in (2) for a specific $\mathbf{J}$ is explicitly contained in Ω_0 and the conjugate fields $\widehat{\mathbf{h}}$. To interpret the physics of the conjugate field, consider driving the field with an external forces $f_i(t)$ in (2), which *do not* depend on $h_i(t)$. Then it is easy to check that

$$\Omega_0[\mathbf{h},\widehat{\mathbf{h}}] \mapsto -i\int dt\,\widehat{\mathbf{h}}(t)\cdot\left[\dot{\mathbf{h}}(t)+\mathbf{h}(t)+\mathbf{f}(t)\right],$$

and therefore, the conjugate fields enter into the definition of the *response function to external stimuli,*

$$\frac{\delta\langle h_i(t)\rangle}{\delta f_j(t')} = \left[\frac{\delta Z}{\delta f_j(t)\delta u_i(t')}\right]_{\mathbf{u},\hat{\mathbf{u}}=\mathbf{0}} = \left\langle h_i(t)\hat{h}_j(t')\right\rangle.$$

Thus the cross-average of $\mathbf{h}$ and $\hat{\mathbf{h}}$ defines the response of the field average to any external force driving the dynamics in (2).

Average over couplings

The generating functional is averaged over the distribution of J_{ij}, and the effective mean field theory is the one that follows from this averaged functional, defined as

$$\bar{Z}[\mathbf{u},\hat{\mathbf{u}}] = \prod_{i\neq j}\int\frac{dJ_{ij}}{(2\pi J^2)^{N/2}}e^{-\frac{J_{ij}^2}{2J^2}}Z\left[\mathbf{u},\widehat{\mathbf{u}};\{J_{ij}\}\right].$$

The average of Z over J_{ij} follows easily since the dependence of Z on couplings is linear and exists only in the last term of (29). Integrating that term under the distribution shown above,

$$\prod_{i\neq j}\int\frac{dJ_{ij}}{(2\pi J^2)^{N/2}}e^{-\frac{1}{2J^2}\left[J_{ij}^2+2iJ^2\int dt\,\widehat{h}_i(t)J_{ij}S_j(t)\right]} = \exp\left[-\frac{J^2}{2}\sum_{i\neq j}\int dt\int dt'\,\widehat{h}_i(t)\widehat{h}_i(t')\phi(h_j(t))\phi(h_j(t'))\right].$$

This term yields a four-field interaction between pairs that are *non-local* in time, with fields in each pair at equal time, $\hat{h}_i(t)\phi(h_j(t))$. This term can be further simplified to bilinear terms by

$$J^2\sum_{i\neq j}\widehat{h}_i(t)\,\widehat{h}_i(t')\phi(h_j(t))\phi(h_j(t')) = J^2\sum_{i=1}^{N}\widehat{h}_i(t)\,\widehat{h}_i(t')\left[\sum_{j=1}^{N}\phi(h_j(t))\phi(h_j(t'))\right] - J^2\sum_{i=1}^{N}\widehat{h}_i(t)\,\widehat{h}_i(t')\phi(h_i(t))\phi(h_i(t')).$$

The last term accounts for the $i = j$ terms that are contributed by the first two but are absent in the left hand side. As the system size increases J must scale as $\tilde{J}/\sqrt{N}$ in order to have non-trivial yet physical results [1]. Substitution of this yields

$$-\frac{\tilde{J}^2 N}{2}\left[\frac{1}{N}\sum_i \int dt \int dt'\, \widehat{h}_i(t)\widehat{h}_i(t')\right]\left[\frac{1}{N}\sum_{j=1}^{N}\phi(h_j(t))\phi(h_j(t'))\right] + \frac{\tilde{J}^2}{2}\left[\frac{1}{N}\sum_i \int dt \int dt'\, \widehat{h}_i(t)\widehat{h}_i(t')S_i(t)S_i(t')\right].$$

All the sums in the square brackets in the above expression are $O(1)$, but the first term scales overall as N multiple higher than the second. Thus under the large N limit the onsite contributions (or spatially local contributions) can be neglected in the unbiased sum over all pairs of contributions.

Making use of stationarity,

$$C_N(t-t') \quad = \quad \frac{1}{N}\sum_{j=1}^{N}\phi(h_j(t))\phi(h_j(t')).$$

So that the generating functional is,

$$Z[C] \quad = \quad \int D\mathbf{h}\int D\hat{\mathbf{h}}\ \exp\left\{\sum_{j=1}^{N} -i\int dt\, \widehat{h}_j(t)\left[\dot{h}_j(t) + h_j(t)\right] - \frac{\tilde{J}^2}{2}\int dt \int dt'\, \widehat{h}_j(t)\widehat{h}_j(t')C_N(t-t')\right\}. \quad (30)$$

Auxiliary field for self-consistent corrrelation

The last term in the exponential above is a four-field interaction. The interaction is non-local in time, and it is also a spatially non-local interaction between two pairs that are spatially local and temporally non-local. We introduce auxiliary fields to decouple the expression into products of two fields that depend on two time coordinates. The decoupling into two-field product allows us to use saddle point approximations, and Gaussian integral formalism. In the present case, we begin by introducing the following identity operation through the delta function,

$$\begin{aligned} Z[C] \quad = \quad & \int D\mathbf{h}\int D\hat{\mathbf{h}}\int DQ(t,t')\,\delta\left(Q(t,t') - C_N(t-t')\right) \times \\ & \exp\left\{\sum_{j=1}^{N} -i\int dt\, \widehat{h}_j(t)\left[\dot{h}_j(t) + h_j(t)\right] - \frac{\tilde{J}^2}{2}\int dt \int dt'\, \widehat{h}_j(t)\widehat{h}_j(t')Q(t,t')\right\}. \end{aligned}$$

Again using the Fourier representation of the delta functional, an effective action is derived. We will skip the explicit step and write the fairly straightforward result,

$$\begin{aligned} Z[C] \quad = \quad & \int D\mathbf{h}\int D\hat{\mathbf{h}}\int DQ(t,t')\int D\tilde{Q}(t,t') \\ & \prod_{j=1}^{N}\exp\left[-i\int dt\, \widehat{h}_j(t)\left\{\dot{h}_j(t) + h_j(t)\right\} - \frac{\tilde{J}^2}{2}\int dt \int dt'\, \widehat{h}_j(t)\widehat{h}_j(t')Q(t,t')\right] \\ & \times \exp\left[-i\int dt \int dt'\ \left\{\tilde{Q}(t,t')Q(t,t') - \tilde{Q}(t,t')C_N(t-t')\right\}\right]. \end{aligned}$$

The product of exponentials in the second line contains identical exponents. Thus it is useful to define

$$\begin{aligned} & \mathcal{A}_{eff}[Q,\tilde{Q},C] \\ = \quad & \log\int Dh\int D\hat{h}\ e^{-i\int dt\,\widehat{h}(t)\left\{\dot{h}(t)+h(t)\right\}} \times \\ & \exp\left[-\frac{1}{N}\int dt \int dt'\, \frac{\tilde{J}^2}{2}\widehat{h}(t)\widehat{h}(t')Q(t,t') + i\left\{\tilde{Q}(t,t')Q(t,t') - \tilde{Q}(t,t')C(t-t')\right\}\right]. \end{aligned}$$

The generator functional now takes the form,

$$Z[C] \quad = \quad \lim_{N\to\infty}\int\int DQD\tilde{Q}\ \exp\left[N\mathcal{A}_{eff}[Q,\tilde{Q},C]\right]. \quad (31)$$

Solution at the saddle point

In the limit of large N in the exponent, it makes sense to do saddle point approximation. The conditions are

$$\left[\frac{\delta \mathcal{A}_{eff}}{\delta \tilde{Q}}\right]_{Q^*,\tilde{Q}^*} = \left[\frac{\delta \mathcal{A}_{eff}}{\delta Q}\right]_{Q^*,\tilde{Q}^*} = 0.$$

We obtain the equations,

$$\begin{aligned} Q^*(t,t') &= C(t-t') = -i\,\langle \phi(h(t))\phi(h(t'))\rangle_{Q^*,\tilde{Q}^*}\,, \\ \tilde{Q}^*(t,t') &= \frac{-i\tilde{J}^2}{2}\left\langle \widehat{h}(t)\widehat{h}(t')\right\rangle_{Q^*,\tilde{Q}^*}. \end{aligned}$$

The subscripts indicate the the quantities on the right hand side, which are averages over J_{ij}, must be evaluated self-consistently. They will generate themselves at the saddle point. Evaluating $\bar{Z}_{eff} = e^{\mathcal{A}_{eff}}$ at the above saddle point, one finds,

$$\bar{Z}_{eff}[C] = \int\int Dh D\widehat{h} \exp\left[\int dt\, i\widehat{h}(t)\left(\dot{h}(t)+h(t)\right) - \frac{1}{2}\int dt \int dt'\, \widehat{h}(t)\left[\tilde{J}^2 C(t-t')\right]\widehat{h}(t')\right]. \tag{32}$$

This describes an effective dynamics of a *single* field.

Interpretation

To interpret the single field equation above, it helps to retrace the steps above for the stochastic differential equation

$$\dot{h}(t) = -h(t) + \eta(t), \tag{33}$$

by introducing a generating functional that is then averaged over η instead of J_{ij}. Thus

$$\bar{Z} = \int\int Dh D\hat{h}\, e^{\int dt\, i\hat{h}(t)(\dot{h}(t)+h(t))} \int D\eta e^{-\frac{1}{2}\int dt \int dt'\, \eta(t)\Xi(t,t')\eta(t') - \int dt \hat{h}(t)\eta(t))}.$$

The Gaussian integral over η has the standard result,

$$\int D\eta e^{-\frac{1}{2}\int dt\int dt'\, \eta(t)\tilde{\Xi}(t,t')\eta(t') - \int dt\hat{h}(t)\eta(t))} = e^{-\frac{1}{2}\int\int dt dt'\, \hat{h}(t)\hat{h}(t')\Xi(t,t')},$$

where the two-time operator and its inverse are defined as follows,

$$\int dt''\, \Xi(t,t'')\tilde{\Xi}(t'',t') = \delta(t-t').$$

Thus the operator appearing on the right hand side is non-local time correlation matrix of the random field $\eta(t)$. Comparison with the expression (32) shows that (2) may equivalently be modeled as (33) in which $\eta(t)$ is a Gaussian random field with correlation,

$$\langle \eta(t+\tau)\eta(t)\rangle = \tilde{J}^2 C(\tau).$$

This completes the derivation of the dynamical mean field in the form of a Langevin equation in which the statistics of the noise term is as described in the previous sections.

References

[1] Scott Kirkpatrick and David Sherrington. Infinite-ranged models of spin-glasses. *Physical Review B*, 17(11):4384–4403, 1978.

[2] SF Edwards and Philip W Anderson. Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5:965, 1975.

[3] H. Sompolinsky and Annette Zippelius. Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses. *Physical Review B*, 25(11):6860–6875, 1982.

[4] H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61(3):259–262, 1988.