

TOPICAL REVIEW

# Path integral methods for the dynamics of stochastic and disordered systems

To cite this article: John A Hertz *et al* 2017 *J. Phys. A: Math. Theor.* **50** 033001

View the [article online](#) for updates and enhancements.

## Related content

- [Symmetry and Collective Fluctuations in Evolutionary Games: Symmetry and collective fluctuations: large deviations and scaling in population processes](#)  
E Smith and S Krishnamurthy
- [Magnetic Excitations and Geometric Confinement: Magnetic vortex core motion and internal dynamics](#)  
G M Wysin
- [Extended Plefka expansion for stochastic dynamics](#)  
B Bravi, P Sollich and M Opper

## Recent citations

- [The prevalence of chaotic dynamics in games with many players](#)  
James B. T. Sanders *et al*
- [Statistical physics approaches to subnetwork dynamics in biochemical systems](#)  
B Bravi and P Sollich

## Topical Review

# Path integral methods for the dynamics of stochastic and disordered systems

John A Hertz<sup>1,2</sup>, Yasser Roudi<sup>1,3,4</sup> and Peter Sollich<sup>5</sup><sup>1</sup> Nordita, KTH Royal Institute of Technology and Stockholm University, Stockholm, Sweden<sup>2</sup> Department of Neuroscience and Pharmacology and Niels Bohr Institute, University of Copenhagen, Denmark<sup>3</sup> Kavli Institute for Systems Neuroscience and Centre for Neural Computation, NTNU, Trondheim, Norway<sup>4</sup> Simons Center for Systems Biology, Institute for Advanced Study, Princeton, NJ, USA<sup>5</sup> King's College London, Department of Mathematics, Strand, London WC2R 2LS, UKE-mail: [Yasser.roudi@ntnu.no](mailto:Yasser.roudi@ntnu.no)

Received 25 April 2016, revised 16 September 2016

Accepted for publication 27 September 2016

Published 12 December 2016



CrossMark

**Abstract**

We review some of the techniques used to study the dynamics of disordered systems subject to both quenched and fast (thermal) noise. Starting from the Martin–Siggia–Rose/Janssen–De Dominicis–Peliti path integral formalism for a single variable stochastic dynamics, we provide a pedagogical survey of the perturbative, i.e. diagrammatic, approach to dynamics and how this formalism can be used for studying soft spin models. We review the supersymmetric formulation of the Langevin dynamics of these models and discuss the physical implications of the supersymmetry. We also describe the key steps involved in studying the disorder-averaged dynamics. Finally, we discuss the path integral approach for the case of hard Ising spins and review some recent developments in the dynamics of such kinetic Ising models.

Keywords: path integral methods, disordered systems, spin glasses, dynamics

**1. Introduction**

Studying the statistical properties of variables or classical fields subject to stochastic forces has a long history in physics. A key tool in this effort is the Langevin equation [1]. Originally introduced for describing the Brownian motion of a particle in a fluid, over the years, this has also been used for studying the dynamics of a variety of other systems including the standard

models of critical phenomena, e.g. those described by the Ginzburg–Landau Hamiltonian [2] and, later on, the relaxation of spin glass models with soft spins [3, 4].

In order to study the dynamical version of the Ginzburg–Landau model in the presence of thermal noise one has to appeal to a perturbative analysis. A useful pedagogical description of the early methods used in such a perturbative treatment can be found in the seminal book by Ma [2] (chapter 5) among other places. A major advance was made by Martin, Siggia and Rose (MSR) in 1973 [5] who realized that one can study classical variables subject to stochastic noise by representing them as Heisenberg operators with appropriately defined commutation relations with *conjugate fields*. These observations allowed MSR to write down a generating functional that naturally lends itself to a perturbative treatment and the use of other field theoretic tools. The most important early development that followed was the work by De Dominicis, Peliti and Janssen [6–8]. They realized that the conjugate field, which was a fundamental insight and step in the MSR theory, and was introduced in some sense *by hand* there, arises naturally if one takes a different point of view. Here one starts the analysis by writing the probability of a path taken by the stochastic variables or fields, which can be used to construct a generating functional expressed as a functional path integral. We refer the reader specifically to [9], which in addition to describing various extensions of the earlier work provides an instructive comparison between the MSR operator based formalism and the functional integral method of De Dominicis, Peliti and Janssen. This latter path integral formalism, which we describe in more detail in this review, has been used to study systems subject to both fast and quenched noise [3]. It paved the way for a more complete understanding of the ensemble averaged dynamics of these systems [4, 10] and, later on, the single sample [11, 12] dynamics of spin glasses and related models.

Despite the remarkable power of the path integral approach, a pedagogical introduction to the method and its key results is lacking. The aim of this paper is to provide a review to fill this gap. The material presented here has previously been used by the authors in a number of lectures and courses (see e.g. the online lectures in [13]); parts of it can also be found in [14].

Some of the material has also been reviewed and discussed elsewhere. The diagrammatic approach to the study of Langevin equations has been described in several classical books [2, 15, 16]. The supersymmetric method applied to dynamics is covered in [15], for example. In this review we aim to combine these classical results with more novel material, including results on the dynamics of hard spin models, in a coherent and consistent notation. By going into more depth in the calculations, we also hope to provide a useful resource, both for newcomers and more experienced researchers in the field. We also recommend that readers consult a recent review by Chow and Buice [17], which covers much of the same material that we do, studying some different model systems from a somewhat different perspective, and in particular demonstrating applications of these methods to the FitzHugh–Nagumo model of a neuron.

In what follows, we start by introducing the path integral formalism for the dynamics of a random variable that evolves according to a stochastic differential equation. We then illustrate how to perform a perturbation analysis and construct Feynman diagrams within this approach. We also show how this treatment can be cast into a supersymmetric form and discuss the physical interpretation of the supersymmetries that are revealed by this approach. Subsequently, we extend the treatment to systems with quenched interactions, especially spin glasses. Finally, we discuss how the path integral formulation can be used to study the dynamics of hard spin models with quenched interactions through a systematic expansion in the strength of the couplings. Before diving into the path integral formulation, however, we review in the next section the two main approaches (Ito and Stratonovich) for interpreting a stochastic differential equation.

## 2. Ito versus Stratonovich

Consider the linear Langevin equation

$$\dot{\phi}(t) = -\mu\phi(t) + h(t) + \zeta(t), \quad (1)$$

where  $h(t)$  is a small field used to probe linear response and  $\zeta$  is zero mean Gaussian noise with correlation function  $\langle \zeta(t)\zeta(t') \rangle = 2T\delta(t-t')$ , where  $\delta(\cdot)$  is the Dirac delta-function. Equation (1) can of course be integrated directly from an initial condition at time  $t = 0$

$$\phi(t) = \phi(0)e^{-\mu t} + \int_0^t dt' e^{-\mu(t-t')} [\zeta(t') + h(t')] \quad (2)$$

and by averaging over the noise and, uncorrelated with it,  $\phi(0)$ , we get the correlation function (for  $h = 0$ )<sup>6</sup>

$$C(t, t') = \langle \phi(t)\phi(t') \rangle = \frac{T}{\mu} e^{-\mu|t-t'|} + \left( \langle \phi^2(0) \rangle - \frac{T}{\mu} \right) e^{-\mu(t+t')} \quad (3)$$

and the response function

$$R(t, t') = \frac{\delta \langle \phi(t) \rangle}{\delta h(t')} = \Theta(t-t') e^{-\mu(t-t')}, \quad (4)$$

where  $\Theta(x)$  is the Heaviside step function, with  $\Theta(x) = 1$  for  $x > 0$  and  $\Theta(x) = 0$  for  $x < 0$ .

We will see below that in the diagrammatic perturbation theory for treating *nonlinear* Langevin equations, we will need the equal time response  $R(t, t)$ . This is *a priori* undefined, however, since the response function has a step discontinuity at  $t' = t$ .

An alternative perspective on this issue is provided by writing the response as a correlation function with the noise. Using the fact that the average over the distribution of  $\zeta$  is with probability weight  $P[\zeta] \sim \exp[-(4T)^{-1} \int dt \zeta^2(t)]$ , and integrating by parts, we can write

$$R(t, t') = \left\langle \frac{\delta \phi(t)}{\delta \zeta(t')} \right\rangle = \frac{1}{2T} \langle \phi(t)\zeta(t') \rangle. \quad (5)$$

For the equal-time response  $R(t, t)$  we therefore require  $\langle \phi(t)\zeta(t) \rangle$ , an equal-time product of fluctuating quantities.

There are two conventions for assigning values to such quantities, due to Stratonovich and to Ito. We will briefly review these conventions and refer the reader to [18] (chapter 4) for a more comprehensive treatment.

### 2.1. Stratonovich convention

The idea of the Stratonovich convention is that, physically,  $\zeta$  is a noise process with non-zero correlation time, so we should really write  $\langle \zeta(t)\zeta(t') \rangle = C_\zeta(t-t')$  with  $C_\zeta$  an even function, decaying quickly to zero for  $|t-t'|$  greater than some small correlation time  $\tau_0$ , and whose integral is  $\int dt C_\zeta(t) = 2T$ . Then, setting the field  $h$  to zero as it is no longer needed at this point:

<sup>6</sup> We use brackets,  $\langle \cdot \rangle$ , to denote averages over all noise histories and, where relevant (as in equation (3)), initial conditions. Whether the initial condition average is included or not will be clear from the context. Other kinds of brackets, with subscripts, that appear later in the paper will denote averages over specific distributions that appear in the theoretical development.

$$\phi(t) = \phi(0)e^{-\mu t} + \int_0^t dt' e^{-\mu(t-t')} \zeta(t') \quad (6)$$

so

$$\langle \phi(t)\zeta(t) \rangle = \int_0^t dt' e^{-\mu(t-t')} C_\zeta(t-t') \simeq \int_0^t dt' C_\zeta(t-t') \simeq T, \quad (7)$$

where the second equality follows from the fast decay of  $C_\zeta$  compared to the macroscopic timescales of  $\mathcal{O}(1/\mu)$  so that we can approximate  $e^{-\mu(t-t')} = 1$ . When considering  $\langle \phi(t)\zeta(t') \rangle$  with  $t > t'$  (more precisely,  $t - t' \gg \tau_0$ ) on the other hand, all of the ‘mass’ of the correlation function is captured in the integration range; it therefore acts just like a  $\delta$ -function and we get  $\langle \phi(t)\zeta(t') \rangle = 2T \exp[-\mu(t-t')]$  as expected from (4) and (5).

To summarize, in the Stratonovich convention, the equal-time value  $R(t,t) = T/(2T) = 1/2$  of the response function is *half* of that obtained in the limit  $t' \rightarrow t - 0$ . It is therefore also called the *midpoint rule*; see below.

## 2.2. Ito convention

The Ito convention effectively assumes that the noise  $\zeta(t)$  acts ‘after  $\phi(t)$  has been updated’, so it sets  $\langle \phi(t)\zeta(t) \rangle = \lim_{t' \rightarrow t+0} \langle \phi(t)\zeta(t') \rangle = 0$ , and hence also  $R(t,t) = 0$ .

## 2.3. Discretization

We will later look at path integral representations of the dynamics and so need a discretization of the stochastic process  $\phi(t)$ . We will now see that Ito and Stratonovich can be viewed as corresponding to different discretization methods.

Let us discretize time  $t = n\Delta$ , with  $\Delta$  a small time step eventually to be taken to zero, and write  $\phi_n = \phi(t = n\Delta)$  and  $h_n = h(t = n\Delta)$ . The noise variables over the interval  $\Delta$  are  $\zeta_n = \int_{n\Delta}^{(n+1)\Delta} dt \zeta(t)$ , with  $\langle \zeta_m \zeta_n \rangle = 2T\Delta\delta_{mn}$ , where  $\delta_{mn}$  is the Kronecker symbol. Then a suitable discrete version of (1) is

$$\phi_{n+1} - \phi_n = \Delta[(1-\lambda)(-\mu\phi_n + h_n) + \lambda(-\mu\phi_{n+1} + h_{n+1})] + \zeta_n \quad (8)$$

for any  $\lambda \in [0, 1]$ ; here we are evaluating (the non-noise part of) the right-hand side of (1) as a weighted combination of the values at the two ends of the interval  $\Delta$ . It is easy to solve this linear recursion exactly: from

$$\phi_{n+1}[1 + \Delta\lambda\mu] = \phi_n[1 + \Delta(\lambda-1)\mu] + \Delta[(1-\lambda)h_n + \lambda h_{n+1}] + \zeta_n, \quad (9)$$

setting

$$c = \frac{1 + \Delta(\lambda-1)\mu}{1 + \Delta\lambda\mu} \quad (10)$$

and assuming the initial condition  $\phi_0 = 0$ , we have

$$\phi_n = \sum_{m=0}^{n-1} c^{n-m-1} \frac{\Delta[(1-\lambda)h_m + \lambda h_{m+1}] + \zeta_m}{1 + \Delta\lambda\mu}, \quad (11)$$

$$\langle \phi_n \rangle = \frac{\Delta}{1 + \Delta\lambda\mu} \left\{ \lambda h_n + \sum_{m=1}^{n-1} \frac{c^{n-m-1}}{1 + \Delta\lambda\mu} h_m + c^{n-1}(1-\lambda)h_0 \right\}. \quad (12)$$

From this we can read off the response function  $R_{nm} = \partial\langle \phi_n \rangle / \partial(\Delta h_m)$ ; setting  $n = t/\Delta$ ,  $m = t'/\Delta$  and taking  $\Delta \rightarrow 0$  we then get for the continuous time response

$$R(t, t') = \begin{cases} 0 & \text{for } t < t', \\ \lambda & \text{for } t = t', \\ \exp[-\mu(t - t')] & \text{for } t > t'. \end{cases} \quad (13)$$

The value of  $\lambda$  only affects the equal-time response; we see that  $\lambda = 1/2$  gives the Stratonovich convention, while  $\lambda = 0$  gives Ito. Note that in the discretization (8),  $\lambda = 1/2$  corresponds to evaluating the change in  $\phi$  at the midpoint of the interval between  $n$  and  $n + 1$  (hence ‘midpoint rule’). Ito ( $\lambda = 0$ ) on the other hand just evaluates at the earlier time. Note also that for different times,  $t \neq t'$ , the two discretization schemes are equivalent as expected.

The two conventions for multiplying equal-time fluctuations are, for the systems we will look at, simply different ways of describing the *same* time evolution  $\phi(t)$ . Cases where the noise strength is coupled to  $\phi$ , such as in  $\dot{\phi} = f(\phi) + g(\phi)\zeta$ , are more serious: a convention for the equal-time product  $g(\phi)\zeta$  has to be adopted, and the two conventions here actually give *different* stochastic processes  $\phi(t)$ ; the corresponding Fokker–Planck equations differ by a non-trivial drift term [18].

For the simple cases we consider, where the noise power  $T$  does not depend on  $t$  or  $\phi$ , Ito versus Stratonovich is basically a matter of taste—calculated expectation values of physical observables are the same in the two approaches. Stratonovich is the more ‘physical’ because it corresponds to a noise process  $\zeta$  with small but non-zero correlation time; it also obeys all the usual rules for transformation of variables etc. Ito is more obvious from the discretized point of view—it is very much what one might naively program in a simulation. We will also see below that it can lead to technical simplifications in calculations.

### 3. The Martin–Siggia–Rose/ Janssen–De Dominicis–Peliti path integral formulation

Now consider the nonlinear Langevin equation

$$\dot{\phi} = f(\phi) + h + \zeta \quad (14)$$

and assume for simplicity that  $\phi(0) = 0$ . It is straightforward to extend the formalism to systems with several components  $\phi_i$ ; the inclusion of distributions of initial values is discussed briefly in section 6. We discretize as in (8), abbreviating  $f_n = f(\phi_n)$ :

$$\phi_{n+1} - \phi_n = \Delta[(1 - \lambda)(f_n + h_n) + \lambda(f_{n+1} + h_{n+1})] + \zeta_n. \quad (15)$$

The plan from here is to write down a path integral for this process and evaluate the effects of nonlinearities in  $f(\phi)$  perturbatively. Let  $\Delta$  be fixed for now and let  $M$  be the largest value of the index  $n$  that we are interested in. Abbreviate  $\phi = (\phi_1 \dots \phi_M)$  and let  $\psi = (\psi_1 \dots \psi_M)$  be a vector of conjugate variables; then the relevant generating function is

$$Z[\psi] = \int d\phi P[\phi] \exp\left(i \sum_{n=1}^M \psi_n \phi_n\right), \quad (16)$$

where  $P[\phi]$  is the probability of the entire history  $\{\phi_n\}$ . Averages can be computed by differentiation with respect to the  $\psi_n$  at  $\psi = 0$ .

Quenched disorder can also be treated: one averages the generating function over the quenched disorder before performing the perturbative expansion. In contrast to the case of equilibrium statistical mechanics, performing the average is unproblematic since, because  $P[\phi]$  is a normalized distribution,  $Z[0] = 1$ , independent of the parameters of the model [3]. We return to this approach in section 9.

The  $M$  variables  $\phi_n$  ( $n = 1 \dots M$ ) depend on the  $M$  noise variables  $\zeta_n$  ( $n = 0 \dots M - 1$ ). We know the distribution of the latter,

$$P[\zeta] = (4\pi T\Delta)^{-M/2} \exp\left[-\frac{1}{4T\Delta} \sum_{n=0}^{M-1} \zeta_n^2\right] \quad (17)$$

so we can find the distribution of the former:

$$P[\phi] = P(\zeta)J[\phi], \quad (18)$$

where  $J[\phi] = |\partial\zeta/\partial\phi|$  is the Jacobian. Using (15) to express the  $\zeta$  in terms of the  $\phi$ ,

$$\zeta_n = \phi_{n+1} - \phi_n - \Delta[(1 - \lambda)(f_n + h_n) + \lambda(f_{n+1} + h_{n+1})] \quad (19)$$

we thus have

$$Z[\psi] = (4\pi T\Delta)^{-M/2} \int d\phi J[\phi] \exp\left[i \sum_{n=1}^M \psi_n \phi_n + \right. \\ \left. - \frac{1}{4T\Delta} \sum_{n=0}^{M-1} (\phi_{n+1} - \phi_n - \Delta[(1 - \lambda)(f_n + h_n) + \lambda(f_{n+1} + h_{n+1})])^2\right], \quad (20)$$

where  $d\phi$  means  $\prod_n d\phi_n$ . The square can be decoupled using conjugate integration variables  $\hat{\phi} = \hat{\phi}_0 \dots \hat{\phi}_{M-1}$ :

$$Z[\psi] = \int \frac{d\phi d\hat{\phi}}{(2\pi)^M} J[\phi] \exp\left[i \sum_{n=1}^M \psi_n \phi_n + \sum_{n=0}^{M-1} (-T\Delta\hat{\phi}_n^2 \right. \\ \left. + i\hat{\phi}_n \{-\phi_{n+1} + \phi_n + \Delta[(1 - \lambda)(f_n + h_n) + \lambda(f_{n+1} + h_{n+1})]\})\right]. \quad (21)$$

Now it is time to evaluate  $J[\phi]$ . Consider the matrix  $\partial\zeta/\partial\phi$ , remembering that the index for  $\zeta$  runs from 0 to  $M - 1$ , and the one for  $\phi$  from 1 to  $M$ . The diagonal elements  $\partial\zeta_n/\partial\phi_{n+1}$  are then, from (19),  $1 - \Delta\lambda f'_{n+1}$  (with  $f'_n = f'(\phi_n)$ ); the elements just below the diagonal are  $\partial\zeta_n/\partial\phi_n = -1 - \Delta(1 - \lambda)f'_n$ . All other elements are zero; in particular the matrix has all zeros above the diagonal and so its determinant is the product of the diagonal elements, giving

$$J[\phi] = \prod_{n=1}^M (1 - \Delta\lambda f'_n) = \exp\left[\sum_{n=1}^M \ln(1 - \Delta\lambda f'_n)\right] = \exp\left[-\Delta\lambda \sum_{n=1}^M f'_n\right], \quad (22)$$

where the last equality anticipates that  $\Delta$  will be made small so that we can discard  $\mathcal{O}(\Delta^2)$  terms in the exponent. Note that for the Ito convention ( $\lambda = 0$ ),  $J \equiv 1$  identically, which is one of the reasons for preferring Ito. The alternative to the above direct evaluation of  $J[\phi]$  is to represent the determinant  $|\partial\zeta/\partial\phi|$  as an integral over Grassmann variables; this will be discussed in section 7.

With  $J[\phi]$  evaluated, we have

$$Z[\psi] = \int \frac{d\phi d\hat{\phi}}{(2\pi)^M} \exp\left[\sum_{n=1}^M (i\psi_n \phi_n - \Delta\lambda f'_n) + \sum_{n=0}^{M-1} (-T\Delta\hat{\phi}_n^2 \right. \\ \left. + i\hat{\phi}_n \{-\phi_{n+1} + \phi_n + \Delta[(1 - \lambda)(f_n + h_n) + \lambda(f_{n+1} + h_{n+1})]\})\right]. \quad (23)$$

Defining the average over a (normalized, complex valued) measure

$$\langle \dots \rangle_S = \int \frac{d\phi d\hat{\phi}}{(2\pi)^M} \dots \exp(-S) \quad (24)$$

with the ‘action’

$$S = \sum_{n=0}^{M-1} (T\Delta\hat{\phi}_n^2 - i\hat{\phi}_n \{-\hat{\phi}_{n+1} + \phi_n + \Delta[(1-\lambda)f_n + \lambda f_{n+1}]\}) + \Delta\lambda \sum_{n=1}^M f'_n \quad (25)$$

one can also write

$$Z[\psi] = \left\langle \exp \left( i \sum_{n=1}^M \psi_n \phi_n + i \sum_{n=0}^{M-1} \hat{\phi}_n \Delta[(1-\lambda)h_n + \lambda h_{n+1}] \right) \right\rangle_S. \quad (26)$$

From this representation one has, in particular (taking all derivatives at  $\psi = h = 0$ , and adopting the convention  $\hat{\phi}_{-1} = 0$ ) the expressions for correlation and response functions

$$C_{nm} = \langle \phi_n \phi_m \rangle = \frac{\partial}{i\partial\psi_n} \frac{\partial}{i\partial\psi_m} Z = \langle \phi_n \phi_m \rangle_S, \quad (27)$$

$$R_{nm} = \frac{\partial \langle \phi_n \rangle}{\partial(\Delta h_m)} = \frac{\partial}{\partial(\Delta h_m)} \frac{\partial}{i\partial\psi_n} Z = \langle \phi_n i \{(1-\lambda)\hat{\phi}_m + \lambda\hat{\phi}_{m-1}\} \rangle_S. \quad (28)$$

It also follows that averages of any product of  $\hat{\phi}$ 's vanish. This is because (as remarked above)  $Z = 1$  for  $\psi = 0$ , whatever value the  $h_n$  take. As a result, derivatives of any order of  $Z$  with respect to  $h$  vanish when taken at  $\psi = 0$ . We can combine the expressions for correlation and response if we define new variables as

$$\eta_{1n} = \phi_n \quad (n = 1 \dots M), \quad (29)$$

$$\eta_{2n} = i \{(1-\lambda)\hat{\phi}_n + \lambda\hat{\phi}_{n-1}\} \quad (n = 0 \dots M-1). \quad (30)$$

Arranging these into a big vector  $\eta = (\eta_{11}, \dots, \eta_{1M}, \eta_{20}, \dots, \eta_{2M-1})$  gives

$$\langle \eta \eta^T \rangle_S = \begin{pmatrix} C & R \\ R^T & 0 \end{pmatrix} = G \quad (31)$$

and the calculation of  $G$  (the ‘propagator’ in quantum field theory) is often the main goal of the analysis. In the case of non-zero initial conditions one also wants to be able to calculate the means  $\langle \eta \rangle$ , as discussed in section 5.5 below. The fields  $\psi$  and  $h$  have served their purpose and will be set to zero from now on.

#### 4. Perturbation theory

To illustrate the perturbative expansion, consider a concrete example:

$$f(\phi) = -\mu\phi - \frac{g}{3!}\phi^3 \quad (32)$$

(the  $3!$  factor is for later convenience). For  $g = 0$ , we recover a solvable linear Langevin equation; correspondingly, the action in (25) becomes quadratic. For  $g \neq 0$ , we therefore separate this quadratic part and write

$$S = S_0 + S_{\text{int}}, \quad (33)$$



$$S_0 = \sum_{n=0}^{M-1} (T\Delta\hat{\phi}_n^2 - i\hat{\phi}_n\{-\phi_{n+1} + \phi_n - \Delta\mu[(1-\lambda)\phi_n + \lambda\phi_{n+1}]\}) + \Delta\mu\lambda M, \quad (34)$$

$$S_{\text{int}} = i\frac{g}{6}\Delta\sum_{n=0}^{M-1}\hat{\phi}_n[(1-\lambda)\phi_n^3 + \lambda\phi_{n+1}^3] - \frac{g}{2}\Delta\lambda\sum_{n=1}^M\phi_n^2. \quad (35)$$

Rearranging the first sum, the non-quadratic (or ‘interacting’ in a field theory) part of the action can be written in the simpler form

$$S_{\text{int}} = \frac{g}{6}\Delta\sum_{n=0}^{M-1}\eta_{2n}\eta_{1n}^3 - \frac{g}{2}\Delta\lambda\sum_{n=1}^M\eta_{1n}^2. \quad (36)$$

We have dropped the last term,  $\lambda\hat{\phi}_{M-1}\phi_M^3$  from the first sum; since (by causality) whatever we do with this last term does not affect any of the results for earlier times, we just have to make  $M$  larger than any times of interest for this omission to be irrelevant. Similarly, whether we have the sums start at  $n=0$  or  $n=1$  has a vanishing effect for  $\Delta \rightarrow 0$ , so in the following we will leave off the summation ranges.

Now for  $g=0$  we have  $S=S_0$ , and a corresponding normalized *Gaussian* measure over the  $\eta$  for which we denote averages by  $\langle \dots \rangle_0$  and the corresponding propagator by  $G_0 = \langle \eta\eta^T \rangle_0$ . For  $g \neq 0$ , our desired averages are then written as

$$\langle \dots \rangle_S = \langle \dots \exp(-S_{\text{int}}) \rangle_0. \quad (37)$$

Assuming  $g$  and hence  $S_{\text{int}}$  to be small, we thus arrive at the *perturbative expansion*

$$\langle \dots \rangle_S = \sum_{k=0}^{\infty} \frac{1}{k!} \langle \dots (-S_{\text{int}})^k \rangle_0 \quad (38)$$

which is a series expansion (in general asymptotic, non-convergent) in the nonlinearity parameter  $g$ . We can now evaluate each term in this series using an elementary fact about multivariate Gaussian statistics known to physicists as *Wick’s theorem*: the average of any product of zero mean Gaussian random variables is found by summing over all possible pairings; symbolically (leaving out all indices etc)

$$\langle \eta\eta \dots \eta \rangle_0 = \sum_{\text{all pairings}} \langle \eta\eta \rangle_0 \dots \langle \eta\eta \rangle_0. \quad (39)$$

Let us apply this to the simplest example. We know that  $\langle 1 \rangle_S = 1$ , so this should be true to all orders in the expansion (38). Up to  $\mathcal{O}(g)$  one has

$$\langle 1 \rangle_S = \langle 1 \rangle_0 - \langle S_{\text{int}} \rangle_0 + \dots \quad (40)$$

$$= 1 - g\Delta \sum_n \left\langle \frac{1}{6}\eta_{2n}\eta_{1n}^3 - \frac{1}{2}\lambda\eta_{1n}^2 \right\rangle_0. \quad (41)$$

Using Wick’s theorem, the fourth-order average is

$$\langle \eta_{2n}\eta_{1n}\eta_{1n}\eta_{1n} \rangle_0 = 3\langle \eta_{2n}\eta_{1n} \rangle_0 \langle \eta_{1n}\eta_{1n} \rangle_0 = 3R_{nn}^0 C_{nn}^0. \quad (42)$$

(Why the 3? There are 3 different pairings between the 4 variables—(1, 2) and (3, 4), (1, 3) and (2, 4), and (1, 4) and (2, 3)—but all give the same product of averages.) Here one sees how the equal time response function appears in the formalism. Inserting the value  $R_{nn}^0 = \lambda$  that we found earlier, (41) thus becomes

$$\langle 1 \rangle = \langle 1 \rangle_0 - g\Delta \sum_n \left( \frac{1}{2} \lambda C_{nn}^0 - \frac{1}{2} \lambda C_{nn}^0 \right) = 1 + \mathcal{O}(g^2) \tag{43}$$

as it should be. This illustrates how the non-trivial determinant  $J[\phi]$  that arises in the Stratonovich formalism ( $\lambda = 1/2$ ) and appears as the term proportional to  $\lambda$  in the interaction part of the action (36) is essential for maintaining the correct normalization. Similar cancellations occur at all orders in  $g$ . Ito ( $\lambda = 0$ ) is simpler here: the terms in the perturbation expansion of  $\langle 1 \rangle_S = 1 + \dots$  all vanish *individually*, rather than just cancelling each other out.

## 5. Diagrams

### 5.1. Basics

Diagrams are just a pictorial way of keeping track of the various terms in the perturbation expansion (38), as evaluated using Wick’s theorem. Having illustrated the equivalence between Ito and Stratonovich above, we stick to Ito ( $\lambda = 0$ ) for now, where the non-trivial part of the action is simply

$$S_{\text{int}} = \frac{g}{6} \Delta \sum_n \eta_{2n} \eta_{1n}^3. \tag{44}$$

To illustrate the diagrammatic notation, consider again the expansion for the normalization factor

$$\langle 1 \rangle = \langle 1 \rangle_0 - \langle S_{\text{int}} \rangle_0 + \frac{1}{2} \langle S_{\text{int}}^2 \rangle_0 + \dots. \tag{45}$$

We have dealt with the zeroth and first order terms above. The second order term is

$$\frac{1}{2} \left( -\frac{g}{6} \right)^2 \Delta^2 \sum_{m,n} \langle \eta_{2m} \eta_{1m}^3 \eta_{2n} \eta_{1n}^3 \rangle_0. \tag{46}$$

Represent each of the summed over time indices  $m$  and  $n$  by a vertex with four ‘legs’ that symbolize the four  $\eta$  factors with the corresponding time index. Each vertex comes with a factor  $-g/6$  from  $-S_{\text{int}}$ . Both of the time indices are summed over and the result multiplied by  $\Delta$ ; in the limit  $\Delta \rightarrow 0$  these scaled sums of course become time integrals. The time indices are not fixed by our choice of observable to average, and are therefore called ‘internal’—we will see external vertices in a moment. Now, having drawn the vertices, we can connect the legs in a number of ways, one for each pairing that Wick’s theorem gives for the average in (46): a line between a vertex at  $m$  and one at  $n$  stands for a pair average  $\langle \eta_{\alpha m} \eta_{\beta n} \rangle$  ( $\alpha, \beta = 1$  or  $2$ ). E.g. the diagram where the legs from both vertices are connected to legs from the same vertex only



represents all the pairings where  $\eta_n$ ’s are connected to  $\eta_n$ ’s only, and  $\eta_m$ ’s to  $\eta_m$ ’s only. At each vertex there are three choices for pairings of this form, so this diagram has the value

$$\begin{aligned} & \frac{1}{2} \left( -\frac{g}{6} \right)^2 \Delta^2 \sum_{m,n} 3 \langle \eta_{2m} \eta_{1m} \rangle_0 \langle \eta_{1m} \eta_{1m} \rangle_0 \times 3 \langle \eta_{2n} \eta_{1n} \rangle_0 \langle \eta_{1n} \eta_{1n} \rangle_0 \\ & = \frac{1}{2} \left\{ \left( -\frac{g}{6} \right) \Delta \sum_n 3 \langle \eta_{2n} \eta_{1n} \rangle_0 \langle \eta_{1n} \eta_{1n} \rangle_0 \right\}^2. \end{aligned} \tag{48}$$

This illustrates an important fact: if a diagram separates into subparts which are not connected, its value is just the product of the two diagrams separately (apart from the overall factor of 1/2 here, which comes from the expansion of  $\exp(-S_{\text{int}})$ ).

The diagram above certainly does not exhaust all the Wick pairings of (46). So what are the other diagrams corresponding to (46)? We can have either two  $mn$ -pairings, one  $mn$  and one  $mm$ ; or four  $mn$  pairings. Altogether, one would therefore write (46) in diagrams as:

$$\text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} + \text{Diagram 4} \tag{49}$$

We write down the value of the last of these: start with  $\eta_{2m}$ . From the diagram, this must be connected to a  $\eta$  on the other vertex, i.e. either  $\eta_{2n}$  or one of the three  $\eta_{1n}$ . In the first case, each of the remaining  $\eta_{1m}$  legs is connected to a  $\eta_{1n}$  leg; there are  $3 \times 2 \times 1 = 6$  ways of making those connections (pairings). In the second case, the  $\eta_2$  factor from the second vertex,  $\eta_{2n}$ , must be connected to one of the three  $\eta_{1m}$  vertices, and the remaining two  $\eta_{1m}$  and  $\eta_{1n}$  legs can be connected to each other in two ways. Thus the value of the diagram is:

$$\frac{1}{2} (-g/6)^2 \Delta^2 \sum_{m,n} \{ 6 \langle \eta_{2m} \eta_{2n} \rangle_0 (\langle \eta_{1m} \eta_{1n} \rangle_0)^3 + 18 \langle \eta_{2m} \eta_{1n} \rangle_0 \langle \eta_{1m} \eta_{2n} \rangle_0 (\langle \eta_{1m} \eta_{1n} \rangle_0)^2 \}. \tag{50}$$

(Exercise: evaluate the remaining second order diagram and, as a check, count all pairings. The diagram just above dealt with  $6 + 18 = 24$  pairings, while the first one corresponded to  $3 \times 3 = 9$  pairings. Since we have 8  $\eta$ 's in total, there are  $7 \times 5 \times 3 \times 1 = 8!/(4!2!4!) = 105$  pairings overall, hence the remaining diagram must correspond to  $105 - 24 - 9 = 72$  pairings.)

In terms of diagrams, it is now quite easy to understand that  $\langle 1 \rangle_S = 1$  as it should to all orders in  $g$ . Consider an arbitrary diagram in the expansion; let us assume it is connected, otherwise consider the subdiagrams separately. Now within the Ito convention the response function  $\langle \eta_{1n} \eta_{2m} \rangle_0$  is non-zero only for  $n > m$  (for Stratonovich, the value for  $n = m$  is also non-zero; for  $n < m$  it is zero either way from causality). So to make the diagram non-zero, we have to connect the  $\eta_{2n_1}$  from a given vertex, with time index  $n_1$ , say, to the  $\eta_{1n_2}$  leg at another vertex,  $n_2$ . The  $\eta_{2n_2}$  from that vertex must in turn be connected to  $\eta_{1n_3}$  on another vertex and so on. Eventually, because we have a finite number of vertices, we must come back to our original vertex  $n_1$ . In the 'ring' sequence  $n_1, n_2, n_3, \dots, n_1$  there are at least two time indices that are in the 'wrong' order for the response function to be non-zero, so that the diagram contains at least one vanishing factor and thus vanishes itself.

The moral of the story so far is: diagrams factorize over disconnected sub-diagrams, and the sub-diagrams with only internal (summed-over) vertices vanish. The latter are also called 'vacuum diagrams' in field theory.

### 5.2. Diagrams for correlator/response; Dyson equation

Next we look at the diagrams for the propagator, which encapsulates correlation and response functions,  $\langle \eta_{\alpha i} \eta_{\beta j} \rangle_S$  where  $\alpha, \beta \in \{1, 2\}$ . We now have two ‘external’ vertices with fixed time indices  $i$  and  $j$ ; the propagator is represented as a double line between these two vertices. It is also called the ‘full’ or ‘dressed’ propagator, in contrast to the ‘bare’ propagator that is represented by the single lines in the diagrams and results from the pairings in Wick’s theorem. To zeroth order in  $g$ , full and bare propagator are obviously equal. To first order, we have, from (38),

$$G_{\alpha i, \beta j} = \langle \eta_{\alpha i} \eta_{\beta j} \rangle_S = \langle \eta_{\alpha i} \eta_{\beta j} \rangle_0 - \langle \eta_{\alpha i} \eta_{\beta j} S_{\text{int}} \rangle_0 + \dots \quad (51)$$

$$= \langle \eta_{\alpha i} \eta_{\beta j} \rangle_0 - (g/6)\Delta \sum_n \langle \eta_{\alpha i} \eta_{\beta j} \eta_{2n} \eta_{1n}^3 \rangle_0 + \dots \quad (52)$$

In diagrams, we have two new contributions from the first order term, depending on whether we pair up  $\eta_i$  with  $\eta_j$  or with one of the  $\eta_n$ . We can therefore write the full propagator as

$$\begin{array}{c} \text{====} \\ i \quad j \end{array} = \begin{array}{c} \text{---} \\ i \quad j \end{array} + \begin{array}{c} \text{---} \\ i \quad j \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \eta \\ \text{---} \\ \text{---} \end{array} + \begin{array}{c} \text{---} \\ i \quad n \quad j \end{array} + \dots \quad (53)$$

The second of these has a disconnected part with only internal vertices, so vanishes. The same argument applies to higher order diagrams as well: we only need to consider *connected* diagrams<sup>7</sup>. Thus, evaluating the surviving diagram,

$$G_{\alpha i, \beta j} = G_{\alpha i, \beta j}^0 - (g/6)\Delta \sum_n \{3G_{\alpha i, 1n}^0 G_{1n, 1n}^0 G_{2n, \beta j}^0 + 3G_{\alpha i, 2n}^0 G_{1n, 1n}^0 G_{1n, \beta j}^0\}. \quad (54)$$

If we define a matrix  $\Sigma^1$  by

$$\Sigma_{\gamma m, \delta n}^1 = -(g/2)\Delta^{-1} \delta_{mn} G_{1n, 1n}^0 (\delta_{\gamma, 1} \delta_{\delta, 2} + \delta_{\gamma, 2} \delta_{\delta, 1}) \quad (55)$$

(where  $\delta_{ij}$  is the Kronecker symbol as before) and agree to absorb a factor of  $\Delta$  into matrix products, so that

$$(AB)_{\alpha i, \gamma k} = \Delta \sum_{\beta j} A_{\alpha i, \beta j} B_{\beta j, \gamma k} \quad (56)$$

then we can write (54) in matrix form simply as

$$G = G_0 + G_0 \Sigma^1 G_0. \quad (57)$$

The way factors of  $\Delta$  are absorbed here ensures that matrix multiplications become time integrals in the natural way for  $\Delta \rightarrow 0$ , while the factor  $\Delta^{-1} \delta_{mn}$  in  $\Sigma^1$  becomes  $\delta(t - t')$ .

To first order in  $g$ , the above is the whole story for the propagator. Now look at the second order. Among the diagrams we have ones such as

<sup>7</sup> Digression on equilibrium statistical mechanics: equilibrium averages can be evaluated by a diagrammatic expansion very similar to the one here. The main difference is that the partition function (whose perturbative expansion is the same as the one for  $\langle 1 \rangle_S$  above) is not automatically normalized. Averages are thus written as  $\langle \dots \rangle = \langle \dots e^{-S_{\text{int}}} \rangle_0 / \langle e^{-S_{\text{int}}} \rangle_0$  and the denominator, when expanded, ensures again that disconnected diagrams vanish. Equivalently, one can think of averages as derivatives of the log partition function; the perturbative expansion for  $\ln Z$  consists of just the *connected* diagrams from the expansion of  $Z$ .

$$\text{Diagram 1} \quad \text{Diagram 2} \quad (58)$$

These two only differ in how the internal vertices are labeled; since the latter are summed over, we can lump the two diagrams together into one *unlabeled* diagram. This just gives a factor of 2, which cancels exactly the prefactor  $1/2!$  from the second order expansion of the exponential in (38). Again, the same happens at higher orders: at  $\mathcal{O}(g^k)$  we have a prefactor of  $1/k!$  but also  $k$  internal vertices which can be labeled in  $k!$  different ways, so the unlabeled diagram has a prefactor of one. Thus, the unlabeled diagram

$$(59)$$

has the value (bearing in mind that the  $\eta_2$  components at each of the internal vertices must be connected either to a  $\eta_1$  leg at the other internal vertex, or to an external vertex, and that there are three choices at each vertex for which pair of  $\eta_1$ 's to connect to each other)

$$\begin{aligned} & 3 \times 3 \times (-g/6)^2 \Delta^2 \sum_{mn} (G_{\alpha i, 1m}^0 G_{1m, 1m}^0 G_{2m, 1n}^0 G_{1n, 1n}^0 G_{2n, \beta j}^0 \\ & + G_{\alpha i, 1m}^0 G_{1m, 1m}^0 G_{2m, 2n}^0 G_{1n, 1n}^0 G_{1n, \beta j}^0 \\ & + G_{\alpha i, 2m}^0 G_{1m, 1m}^0 G_{1m, 1n}^0 G_{1n, 1n}^0 G_{2n, \beta j}^0 \\ & + G_{\alpha i, 2m}^0 G_{1m, 1m}^0 G_{1m, 2n}^0 G_{1n, 1n}^0 G_{1n, \beta j}^0). \end{aligned} \quad (60)$$

Using the definition (55), one sees that in matrix form this is simply

$$G_0 \Sigma^1 G_0 \Sigma^1 G_0. \quad (61)$$

To make this simple form more obvious we have included in (60) the term in the second line, which vanishes because it contains a zero  $G_{22}^0$  factor. We have an example here of a ‘one-particle reducible’ (1PR) diagram, which can be cut in two by cutting just one bare propagator line, namely, the one in the middle. The result illustrates that the value of such diagrams factorizes into the pieces they can be cut into; e.g. the diagram

$$(62)$$

has the value  $G_0 \Sigma^1 G_0 \Sigma^1 G_0 \Sigma^1 G_0$ . If we sum up all the diagrams of this form, we get

$$G = G_0 + G_0 \Sigma^1 G_0 + G_0 \Sigma^1 G_0 \Sigma^1 G_0 + G_0 \Sigma^1 G_0 \Sigma^1 G_0 \Sigma^1 G_0 + \dots = [G_0^{-1} - \Sigma^1]^{-1} \quad (63)$$

or

$$G^{-1} = G_0^{-1} - \Sigma^1. \quad (64)$$

The inverses are relative to the appropriate unit element  $\mathbf{I}$  for our redefined matrix multiplication, which has elements  $I_{\alpha i, \beta j} = \Delta^{-1} \delta_{\alpha \beta} \delta_{ij}$ . (So  $A^{-1}A = \mathbf{I}$  means, because of the extra factor  $\Delta$  in the matrix multiplication, that  $A^{-1}$  is  $\Delta^{-2}$  times the conventional matrix inverse of  $A$ .)

The expression (64) is an example of a *resummation*: we have managed to sum up an infinite subseries from among all the diagrams for the propagator. What if we want to sum up *all* the diagrams?

Again we can classify them into one-particle irreducible (1PI) diagrams, which cannot be cut in two by cutting a single line, and 1PR diagrams that factorize into their 1PI components.

For example, to second order

$$\text{====} = \text{---} + \text{---} \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \end{array} + \text{---} \begin{array}{c} \circ \quad \circ \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \end{array} + \text{---} \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \end{array} + \text{---} \begin{array}{c} \circ \quad \circ \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \end{array} + \mathcal{O}(g^3), \quad (65)$$

$$= \left( \text{---}^{-1} - \left[ \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \end{array} + \begin{array}{c} \circ \quad \circ \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \circ \quad \circ \end{array} + \begin{array}{c} \circ \quad \circ \\ \diagup \quad \diagdown \\ \circ \end{array} \right] \right)^{-1} + \mathcal{O}(g^3). \quad (66)$$

More generally, if we denote the values of the different 1PI diagrams (defined analogously to  $\Sigma^1$ , i.e. without the external  $G_0$  legs) by  $\Sigma^1, \Sigma^2, \dots$  then we can get all possible (1PI and 1PR) diagrams by ‘stringing’ together all possible combinations of 1PI diagrams:

$$G = G_0 \sum_{k=0}^{\infty} \sum_{i_1 \dots i_k} \Sigma^{i_1} G_0 \dots \Sigma^{i_k} G_0 = G_0 \sum_{k=0}^{\infty} (\Sigma G_0)^k = (G_0^{-1} - \Sigma)^{-1}, \quad (67)$$

where  $\Sigma = \sum_{i=1}^{\infty} \Sigma^i$ . Hence we see that the full propagator, with all diagrams summed up, can be written in the general form

$$G^{-1} = G_0^{-1} - \Sigma, \quad (68)$$

where  $\Sigma$ , the so-called ‘self-energy’, is the sum of all 1PI diagrams. Equation (68) is called the *Dyson equation*. An alternative form that is often useful is

$$G_0^{-1} G = \mathbf{I} + \Sigma G. \quad (69)$$

Let us write this out in terms of the separate blocks corresponding to correlation and response functions: we have

$$G_0 = \begin{pmatrix} C_0 & R_0 \\ R_0^T & 0 \end{pmatrix} \Rightarrow G_0^{-1} = \begin{pmatrix} 0 & (R_0^{-1})^T \\ R_0^{-1} & -R_0^{-1} C_0 (R_0^{-1})^T \end{pmatrix} \quad (70)$$

and  $G^{-1}$  has the same structure, so that also the 11 block of  $\Sigma$  must vanish,

$$\Sigma = \begin{pmatrix} 0 & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}. \quad (71)$$

This can also be shown diagrammatically: a non-zero contribution to  $\Sigma_{11}$  would correspond to diagrams where the internal vertices that make connections to the two external vertices do so via  $\eta_1$  legs. This leaves all the  $\eta_2$  legs to be connected amongst the internal vertices, and then the same argument as for vacuum diagrams can be applied. Writing out (69) we have thus

$$(R_0^{-1})^T R^T = \mathbf{I} + \Sigma_{12} R^T, \quad (72)$$

$$0 = 0, \quad (73)$$

$$R_0^{-1} C - R_0^{-1} C_0 (R_0^{-1})^T R^T = \Sigma_{12}^T C + \Sigma_{22} R^T, \quad (74)$$

$$R_0^{-1} R = \mathbf{I} + \Sigma_{12}^T R, \quad (75)$$

where we have abused the notation by writing  $\mathbf{I}$  also for the non-zero  $M \times M$  sub-blocks of the original  $2M \times 2M$  matrix  $\mathbf{I}$ . The first and last of these equations are both equivalent to

$$R^{-1} = R_0^{-1} - \Sigma_{12}^T \tag{76}$$

implying that  $\Sigma_{12}^T$  acts like a self-energy for the reponse function. Rearranging, the components of the Dyson equation reduce to

$$R_0^{-1}R = \Sigma_{12}^T R + I, \tag{77}$$

$$R_0^{-1}C = \Sigma_{12}^T C + [R_0^{-1}C_0(R_0^{-1})^T + \Sigma_{22}]R^T. \tag{78}$$

Using (76), the last equation can also be solved explicitly for  $C$  as

$$C = R[R_0^{-1}C_0(R_0^{-1})^T + \Sigma_{22}]R^T. \tag{79}$$

In the above we have defined  $\Sigma$  to be *plus* the sum of all 1PI diagrams; the opposite sign convention also appears in the literature.

### 5.3. Self-consistency

Of course in general one cannot sum up all the diagrams for the self-energy. One must make some approximation. In (64) we used just the lowest order diagram to approximate  $\Sigma$ . Can we easily improve the approximation? Yes, if we replace  $G_0$  in the expression for  $\Sigma$  by the full propagator  $G$ ; diagrammatically,

$$\Sigma = \text{tadpole diagram} \tag{80}$$

Which diagrams does this correspond to? This is easiest to find out order by order in  $g$ . To first order,  $\Sigma$  and  $G$  are

$$\Sigma = \text{tadpole} = \text{double line} = \text{single line} + \text{tadpole on line} \tag{81}$$

Re-inserting  $G$  into  $\Sigma$  we get its form to second order and therefore also  $G$

$$\Sigma = \text{tadpole} + \text{two tadpoles} \tag{82}$$

$$\text{double line} = \text{single line} + \text{tadpole on line} + \text{two tadpoles on line} + \text{three tadpoles on line} \tag{83}$$

and then we can iterate to get

$$\Sigma = \text{tadpole} + \text{two tadpoles} + \text{three tadpoles} + \text{four tadpoles} + \dots \tag{84}$$

So the simple operation of replacing  $G_0$  by  $G$  effectively sums up an infinite series of ‘tadpole’ diagrams in the expansion for the self-energy. This is called ‘mean field theory’. It gives exact results for many models with weak long-ranged interactions; for such models the diagrams that have not been included are negligible in the thermodynamic limit. The approximation is also called ‘self-consistent one-loop’ or ‘Hartree–Fock’ approximation.

### 5.4. Diagrammatic conventions

We have used a particular way of drawing diagrams above that does not distinguish between the physical field  $\phi = \eta_1$  and the conjugate or ‘response’ field  $\hat{\phi} = \eta_2$ . This approach has the

virtue of making all diagrams look essentially like in a static (equilibrium)  $\phi^4$  field theory. It does however mean that each diagram can group a number of terms, involving response or correlation functions depending where on each vertex the  $\hat{\phi}$  variables are located in any given Wick pairing. If one wants to avoid this one can e.g. mark on each vertex the  $\hat{\phi}$ -leg by an outgoing arrow. Similarly in the response part of the propagator—the  $\eta_1\eta_2$  sector in our previous convention—one would mark the  $\hat{\phi}$  end by an arrow. The arrows then represent the flow of time because in all contractions of the type  $\langle \hat{\phi}\phi \rangle$  the  $\hat{\phi}$  must be at a time before the  $\phi$ .

The rules for constructing diagrams are modified only slightly in this more detailed diagrammatic representation: an arrow from a vertex cannot connect back to the same vertex as this would give an equal-time response function, which vanishes in the Ito convention. At each vertex one has to consider the possible choices for how the leg with an arrow can be connected to other vertices or external nodes. Diagrams are non-vanishing only if the arrows give a consistent flow of time, i.e. arrows cannot form closed loops nor can two legs with outgoing arrows be connected.

Within the above convention, a propagator *without* an arrow on it is always a correlation function. In our example, the diagrammatic expansion up to second order of the correlation function then reads

$$\begin{aligned}
 \text{---} &= \text{---} + \text{---} \circlearrowleft + \text{---} \circlearrowright + \text{---} \circlearrowleft \circlearrowleft + \text{---} \circlearrowright \circlearrowright + \text{---} \circlearrowleft \circlearrowright \\
 &+ \text{---} \circlearrowleft \circlearrowright + \text{---} \circlearrowright \circlearrowleft + \text{---} \circlearrowleft \circlearrowright \circlearrowleft \\
 &+ \text{---} \circlearrowleft \circlearrowright \circlearrowleft + \text{---} \circlearrowright \circlearrowleft \circlearrowright + \mathcal{O}(g^3)
 \end{aligned} \tag{85}$$

while for the response function there are fewer diagrams that contribute:

$$\text{---} \Rightarrow = \text{---} \Rightarrow + \text{---} \circlearrowleft \Rightarrow + \text{---} \circlearrowright \circlearrowright \Rightarrow + \text{---} \circlearrowleft \circlearrowright \Rightarrow + \mathcal{O}(g^3) . \tag{86}$$

Note that the expansions (85) and (86) are written as the direct analogs of (65) and differ from the latter only through the addition of the appropriate arrows. More compact expressions can be obtained in terms of the self-energy. For the correlation function specifically, the identity (79) may be more efficient for use in practical calculations. This is illustrated in the example in section 6 below, see (106) there.

### 5.5. Non-zero fields and initial values

So far we have discussed dynamics without applied fields  $h_n$  and with zero initial value  $\phi_0$ . These restrictions can be lifted, as we now explain. We continue to use the Ito convention ( $\lambda = 0$ ). It suffices to discuss non-zero fields, because a non-zero initial value  $\phi_0 = c$  can be produced by setting  $h_0 = c/\Delta$ . This gives  $\phi_1 = c + \mathcal{O}(\Delta^{1/2})$ , which in the continuous time limit  $\Delta \rightarrow 0$  approaches  $c$ . In the same limit the difference between  $\phi_0$  and  $\phi_1$  is immaterial, so the above choice of  $h_0$  effectively fixes a non-zero initial condition.



For the perturbative approach to work, we need a quadratic action  $S_0$  as our baseline while all non-quadratic terms are contained in the interacting part  $S_{\text{int}}$ . If the field  $h_n$  is non-zero, this generates linear terms in the action. We then have two choices: either we include these linear terms in  $S_{\text{int}}$  and treat them perturbatively, or we transform variables by expanding the full action around a stationary or saddle point.

The first approach is relatively straightforward: one now has an extra vertex, with only one  $\hat{\phi}$ -leg, and the perturbative expansion is jointly in  $g$  and the field amplitude. For the propagator all contributing diagrams must have a total number of legs on the internal vertices that is even so the number of field vertices must also be even, hence the expansion is effectively in  $g$  and  $h^2$ . At  $\mathcal{O}(h^2)$  one has the extra diagram

$$\text{---}\bullet\text{---}\bullet\text{---}, \tag{87}$$

where each dot with a *single* connection is an  $h$ -vertex<sup>8</sup>, while at  $\mathcal{O}(h^2g)$  one gets

$$\text{---}\bullet\text{---}\bullet\text{---} + \text{---}\bullet\text{---}\bullet\text{---} + \text{---}\bullet\text{---}\bullet\text{---}. \tag{88}$$

We could represent the fact that all field vertices have a single  $\hat{\phi}$ -leg by an arrow pointing away from the vertex (see section 5.4). For the response function part of the propagator there would be an arrow pointing away from one of the external vertices as well so the  $\mathcal{O}(h^2)$  diagram in (87) vanishes as it contains an edge with opposing arrows. To a correlator  $C(t, t')$  the diagram contributes in the continuous time limit

$$\int dt_1 dt_2 R_0(t, t_1)h(t_1)R_0(t', t_2)h(t_2). \tag{89}$$

The higher order diagrams can be evaluated similarly.

In the presence of a field even the mean  $\langle\phi_n\rangle$  is in general non-zero. The perturbative expansion for this consists of diagrams with one external vertex and therefore requires an odd number of internal field vertices. The contributions to  $\mathcal{O}(h)$  and  $\mathcal{O}(hg)$  are

$$\text{=}\bullet\text{---} = \text{---}\bullet\text{---} + \text{---}\bullet\text{---}\bullet\text{---}. \tag{90}$$

One sees that most of the new diagrams (87) and (88) in the propagator are products of diagrams like these. In fact one can convince oneself that if one considers the *connected* propagator, defined as  $\langle\eta_{\alpha i}\eta_{\beta j}\rangle - \langle\eta_{\alpha i}\rangle\langle\eta_{\beta j}\rangle$ , then as in the case without a field only the connected diagrams remain [19]; the lowest order one of these is the third diagram in (88).

Next we look at the saddle point approach, where one defines new variables relative to a stationary point of the action, thus eliminating any linear terms in the transformed action. Looking at (23) with  $\psi_n = 0$ , the saddle point conditions with respect to  $\phi_n$  and  $\hat{\phi}_n$  give

$$\begin{aligned} 0 &= \hat{\phi}_n - \hat{\phi}_{n-1} + \hat{\phi}_n \Delta f'_n, \\ 0 &= -2T\hat{\phi}_n + i[-\phi_{n+1} + \phi_n + \Delta(f_n + h_n)]. \end{aligned}$$

The ‘initial’ (at  $n = M - 1$ ) condition for the first of these equations, which results from stationarity with respect to  $\phi_M$ , is  $\hat{\phi}_{M-1} = 0$ , and solving backwards in time then shows that  $\hat{\phi}_n = 0$  for all  $n$  at the saddle point. With this the second equation is just

<sup>8</sup> It is common in the literature to use a different symbol, such as a cross, to represent the external field. Here, in the interest of a more uniform notation, we stick with dots, and the number of lines connected at a dot indicates whether it means  $h$ ,  $g$ , or a coupling of some still-different order, such as the cubic vertex in (93).

$$\phi_{n+1} - \phi_n = \Delta(f_n + h_n) \tag{91}$$

which is the discrete time version of the deterministic (noise-free) time evolution  $\dot{\phi} = f(\phi) + h$ . If we call the solution of this  $\phi^*$ , then we need to use as new variables  $\delta\phi = \phi - \phi^*$ . In terms of these, by making use of the saddle point conditions, the (Ito) action reads

$$S = \sum_{n=0}^{M-1} \{T\Delta\hat{\phi}_n^2 - i\hat{\phi}_n[-\delta\phi_{n+1} + \delta\phi_n + \Delta\delta f_n]\}. \tag{92}$$

The difference  $\delta f_n = f_n - f_n^* = f(\phi_n^* + \delta\phi_n) - f(\phi_n^*)$  needs to be expanded in  $\delta\phi_n$  to read off the resulting vertices in the interacting part of the action. For our example (32),

$$\delta f = -\mu\delta\phi - \frac{g}{3!}(3\phi^{*2}\delta\phi + 3\phi^*\delta\phi^2 + \delta\phi^3). \tag{93}$$

If we include the  $-\mu\delta\phi$  term in the unperturbed action  $S_0$  then the latter has the same form as originally except for the change of variable from  $\phi$  to  $\delta\phi$ , and the same diagrammatic expansion technique can be applied. The difference is that the interaction part  $S_{\text{int}}$  now contains three different kinds of vertices resulting from the terms proportional to  $g$  in (93), with two, three and four legs respectively and different time-dependent prefactors from the time dependence of  $\phi^*$ .

One point to bear in mind is that even though we have defined  $\delta\phi$  relative to the saddle point solution, its average is not in general zero. To  $\mathcal{O}(g)$ , for example, one has for  $\langle\delta\phi(t)\rangle$  the diagram

$$\text{---}\bullet\text{---} = \text{---}\bullet\text{---}\bigcirc \tag{94}$$

coming from the  $-(g/2)\phi^*\delta\phi^2$  term in (93), which evaluates to

$$\langle\delta\phi(t)\rangle = -\frac{g}{2} \int dt' R_0(t, t')\phi^*(t')C_0(t', t'). \tag{95}$$

### 6. A one-particle example

Let us apply the diagrammatic formalism developed above to our one-particle model (32). We want to take a distribution over the initial values  $\phi(0)$  into account here. But this only takes a small extension of the formalism: if the initial distribution is a zero mean Gaussian, we can simply include the average over  $\phi(0)$  in the unperturbed average  $\langle\dots\rangle_0$ ; the measure is still Gaussian, so we can apply all of the above formalism except that the values of  $C_0$  and  $R_0$ , i.e. the components of the bare propagator are affected by the presence of uncertainty in the initial condition. If the distribution has non-Gaussian parts, we include the Gaussian part as above and the remainder is put into the non-trivial part of the action  $S_{\text{int}}$ , giving a new kind of vertex in the diagrams. A non-zero mean in a Gaussian initial distribution would also be put into  $S_{\text{int}}$ .

Now let us write down the Dyson equation for our model. Having derived all relations previously so that they have the obvious limits for  $\Delta \rightarrow 0$ , we work directly with continuous times. The bare response function is, from (4),

$$R_0(t, t') = \Theta(t - t')e^{-\mu(t-t')}. \tag{96}$$

The inverse of  $R_0$  is the operator  $\partial_t + \mu$ , since when applied to  $R_0$  it gives  $\delta(t - t')$ . The other quantity we need to write down the Dyson equation (78) is  $R_0^{-1}C_0(R_0^{-1})^T$ . To find this, it is easiest to start from the fact that

$$\phi(t) = \int_0^\infty dt_1 R_0(t, t_1) [\zeta(t_1) + \phi(0)\delta(t_1)]. \quad (97)$$

The lower boundary of the integral here is meant as  $0 - \epsilon$ ; the same applies to all integrals that follow. Averaging the product  $\phi(t)\phi(t')$  gives

$$C_0(t, t') = \int_0^\infty dt_1 dt_2 R_0(t, t_1) [2T\delta(t_1 - t_2) + \langle \phi^2(0) \rangle \delta(t_1)\delta(t_2)] R_0(t', t_2) \quad (98)$$

so  $C_0 = R_0 M R_0^T$  with  $M(t_1, t_2)$  given by the square brackets; thus

$$(R_0^{-1} C_0 (R_0^{-1})^T)(t, t') = 2T\delta(t - t') + \langle \phi^2(0) \rangle \delta(t)\delta(t'). \quad (99)$$

Now we can write down the Dyson equation (78):

$$\left( \frac{\partial}{\partial t} + \mu \right) R(t, t') = \int_0^\infty dt'' \Sigma_{12}(t'', t) R(t'', t') + \delta(t - t'), \quad (100)$$

$$\begin{aligned} \left( \frac{\partial}{\partial t} + \mu \right) C(t, t') &= \int_0^\infty dt'' \Sigma_{12}(t'', t) C(t'', t') \\ &+ \int_0^\infty dt'' [2T\delta(t - t'') + \langle \phi^2(0) \rangle \delta(t)\delta(t'')] \\ &+ \Sigma_{22}(t, t'')] R(t', t''). \end{aligned} \quad (101)$$

To first order in  $g$  the self-energy is, from (55),

$$\Sigma_{12}(t, t') = -(g/2)\delta(t - t')C_0(t, t) \quad \Sigma_{22}(t, t') = 0. \quad (102)$$

Within this approximation, the Dyson equation becomes

$$\left( \frac{\partial}{\partial t} + \mu \right) R(t, t') = -(g/2)C_0(t, t)R(t, t') + \delta(t - t'), \quad (103)$$

$$\left( \frac{\partial}{\partial t} + \mu \right) C(t, t') = -(g/2)C_0(t, t)C(t, t') + 2TR(t', t) + \langle \phi^2(0) \rangle \delta(t)R(t', 0). \quad (104)$$

From the first equation,

$$R(t, t') = \Theta(t - t') \exp \left[ -\mu(t - t') - (g/2) \int_{t'}^t dt'' C_0(t'', t'') \right] \quad (105)$$

and with this the second equation for  $C$  can also be solved (compare (79)) to give

$$C(t, t') = \langle \phi^2(0) \rangle R(t, 0)R(t', 0) + 2T \int_{t'}^t dt'' R(t, t'')R(t', t''). \quad (106)$$

For the simplest case where  $C_0$  is time-translation invariant, corresponding to  $\langle \phi^2(0) \rangle = C_0(t, t) = T/\mu$ , we see that the effect of  $g$  in the response function  $R$  is just to replace  $\mu \rightarrow \mu + gT/(2\mu)$ . At long times the effect on  $C$  is similar though at short times  $C$  will not be time-translation invariant.

If we make our first order (one-loop) approximation self-consistent, the only change is to replace  $C_0$  by  $C$  in (102) and correspondingly in the Dyson equation, so that

$$R(t, t') = \Theta(t - t') \exp \left[ -\mu(t - t') - (g/2) \int_{t'}^t dt'' C(t'', t'') \right]. \quad (107)$$

This has a simple interpretation: it corresponds to replacing our original nonlinear force term (32) by

$$f(\phi) \approx -\mu\phi - \frac{g}{3!} \langle \phi^2 \rangle \phi \tag{108}$$

with  $\langle \phi^2(t) \rangle = C(t, t)$  to be determined self-consistently. (The non-self-consistent version instead sets  $\langle \phi^2(t) \rangle = C_0(t, t)$ .) Assuming that we can find a solution with  $C(t, t) = c = \text{constant}$ , we get for  $C$  by inserting (107) into (106) and setting  $\tilde{\mu} = \mu + cg/2$

$$C(t, t') = \langle \phi^2(0) \rangle e^{-\tilde{\mu}(t+t')} + \frac{T}{\tilde{\mu}} [e^{-\tilde{\mu}|t-t'|} - e^{-\tilde{\mu}(t+t')}] \tag{109}$$

For  $\langle \phi^2(0) \rangle = T/\tilde{\mu}$  we then indeed get a time-translationally invariant  $C(t, t') = (T/\tilde{\mu}) \exp(-\tilde{\mu}|t-t'|)$ . This has  $C(t, t) = T/\tilde{\mu}$  and so the self-consistent equation determining  $c$  and  $\tilde{\mu}$  is

$$\frac{T}{\tilde{\mu}} = c \quad \Rightarrow \quad \tilde{\mu} = \mu + \frac{gT}{2\tilde{\mu}} \tag{110}$$

This is the self-consistent analog of our previous result  $\tilde{\mu} = \mu + (gT)/(2\mu)$ , to which it reduces when expanded to first order in  $g$ .

We see that for our particular model the self-consistent approximation gives a more sensible result than the ‘vanilla’ first order approximation: it allows a time-translation invariant solution for both  $C$  and  $R$


$$R(t, t') = \Theta(t-t') e^{-\tilde{\mu}(t-t')} \tag{111}$$

$$C(t, t') = \frac{T}{\tilde{\mu}} e^{-\tilde{\mu}|t-t'|} \tag{112}$$

which also obeys the fluctuation–dissipation theorem (FDT),  $R(t, t') = (1/T)(\partial/\partial t')C(t, t')$  for  $t > t'$ .

### 6.1. Mode-coupling theory

What if we want to improve the approximation to the self-energy further? The systematic approach is to include the lowest-order diagram not so far taken into account. We have the only first-order diagram already; the second-order ‘tadpole’ diagrams are also taken into account through self-consistency. The only missing second order diagram is therefore the ‘watermelon’ diagram



$$\tag{113}$$

To work out what contribution to  $\Sigma$  this gives, let us revert temporarily to discrete time notation and label the left and right vertex  $m$  and  $n$ , respectively. The elements  $\Sigma_{1m,2n}$  of  $\Sigma_{12}$  correspond to those pairings where a  $\eta_{1m}$  leg from vertex  $m$  is attached to an external vertex that would be on the left, and the  $\eta_{2n}$  leg from vertex  $n$  attached to an external vertex on the right. Internally (among the remaining legs) we thus have two  $\eta_{1m}\eta_{1n}$  pairings and one  $\eta_{2m}\eta_{1n}$  pairing. To work out the prefactor of the diagram, note that there are three choices for the externally attached  $\eta_{1m}$ ; there are three choices for which of the  $\eta_{1n}$  to pair up with  $\eta_{2m}$ ; and two more choices for how to make the two remaining  $\eta_{1m}\eta_{1n}$  pairings. Thus, the diagram gives for  $\Sigma_{12}$

$$\Sigma_{1m,2n} = 3 \times 3 \times 2(-g/6)^2 (C_{mn}^0)^2 R_{nm}^0 = (g^2/2)(C_{mn}^0)^2 R_{nm}^0 \tag{114}$$

For  $\Sigma_{22}$ , we have both the  $\eta_{2m}$  and  $\eta_{2n}$  legs attached externally, and 6 choices for how to connect the three  $\eta_{1m}$  and  $\eta_{1n}$  legs internally, giving

$$\Sigma_{2m,2n} = 6(-g/6)^2(C_{mn}^0)^3 = (g^2/6)(C_{mn}^0)^3. \quad (115)$$

We can again sum an infinite series of additional diagrams by replacing bare  $(C_0, R_0)$  by full quantities  $(C, R)$  here. Reverting to continuous time notation and including the first-order contribution in  $\Sigma_{12}$ , we thus get for the self-energy in the self-consistent two-loop approximation

$$\Sigma_{12}(t, t') = -(g/2)\delta(t - t')C(t, t) + (g^2/2)C^2(t, t')R(t', t), \quad (116)$$

$$\Sigma_{22}(t, t') = (g^2/6)C^3(t, t'). \quad (117)$$

A final comment: up to the order which we have considered, the self energy components  $\Sigma_{12}(t, t')$  and  $\Sigma_{22}(t, t')$  are simple *functions* of the correlation and response functions. This is not normally true once higher order diagrams are taken into account. For example, if we went to third order in  $g$  we would have to include the diagram


(118)

in the self energy; all other third order diagrams are automatically included by self-consistency. This diagram now has one internal vertex whose time index is not fixed by the two time indices that the self-energy carries. It therefore gives a contribution to  $\Sigma_{..}(t, t')$  which has an integral over this ‘internal’ time; e.g. for  $\Sigma_{22}$  we get a contribution

$$\Sigma_{22}(t, t') \propto C(t, t') \int dt'' C(t, t'')C(t', t'')[R(t, t'')C(t', t'') + R(t', t'')C(t, t'')]. \quad (119)$$

Diagrams of even higher order contain additional time integrals; in general, then, the self-energy is a *functional* of the response and correlation functions.

An approximation, like that in (116) and (117), in which we keep only diagrams for  $\Sigma(t, t')$  that are functions of  $G(t, t')$  and  $C(t, t')$  has come to be called a *mode coupling approximation*. These have been studied extensively, often because they are exact for some mean-field (infinite-range) models [20].

## 7. Ghosts and supersymmetry

### 7.1. Using Grassmann variables

In section 4, we discussed how the results of the perturbation theory are independent of the value of  $\lambda$  chosen in the discretization of the Langevin equation: the choice of  $\lambda$  changes the Jacobian but does not affect correlation functions and other observables. As alluded to already in section 3, there is another way of including the Jacobian in the path integral using Grassmann variables, which is both conceptually interesting and can often simplify notation. This observation was first made, and its consequences on dynamics explored, by Feigel’man and Tsvelik [21, 22], who followed earlier work by Parisi and Sourlas [23] on supersymmetric properties of an equilibrium system in a random external field. The current section is devoted to describing this approach.

Before getting into the details of the supersymmetric formalism, we need to familiarize ourselves with Grassmann variables. Grassmann variables are characterized by the fact that they *anticommute* with each other. Multiplication of them is also associative; i.e. for Grassmann variables  $\xi_i$ , we have

$$\xi_i(\xi_j\xi_k) = (\xi_i\xi_j)\xi_k \quad (\text{associative}), \quad (120)$$

$$\xi_i \xi_j = -\xi_j \xi_i \text{ (anti-commutative).} \quad (121)$$

A consequence of anti-commutation is that

$$\xi_i^n = 0 \quad \forall n > 1. \quad (122)$$

Grassmann variables can be added to each other and also multiplied by complex numbers: one says formally that they form an algebra, i.e. a vector field over the complex numbers endowed with a multiplication. This means that, given (122), the most general functions of one and two Grassmann variables can be written respectively as

$$f(\xi_1) = c_0 + c_1 \xi_1, \quad (123)$$

$$f(\xi_1, \xi_2) = c_0 + c_1 \xi_1 + c_2 \xi_2 + c_{12} \xi_1 \xi_2, \quad (124)$$

where  $c_0, c_1, c_2$  and  $c_{12}$  are arbitrary complex numbers.

Integration and differentiation for Grassmann variables are defined by

$$\frac{d}{d\xi_i} \xi_j = \delta_{ij}, \quad \int d\xi = 0, \quad \int d\xi \xi = 1, \quad (125)$$

and these lead to the following formulae that we will use later:

$$\int d\xi (a + b\xi) = b, \quad \int d\xi_1 d\xi_2 \xi_2 \xi_1 = 1, \quad (126)$$

$$\frac{d}{d\xi} f(\xi) = \frac{d}{d\xi} (a + b\xi) = b, \quad \frac{d}{d\xi_1} \xi_2 \xi_1 = -\xi_2. \quad (127)$$

As a consequence of the above, and using two independent sets of Grassmann variables  $\xi_n$  and  $\bar{\xi}_n$ , one has the following important representation of the determinant of a matrix  $A$ :

$$|A| = \int D[\xi \bar{\xi}] \exp \left\{ \sum_{mn} \bar{\xi}_m A_{mn} \xi_n \right\}, \quad (128)$$

where  $D[\xi \bar{\xi}] = \prod_n d\xi_n d\bar{\xi}_n$ . The integrand on the right-hand side of (128) defines a Gaussian measure for Grassmann variables under which we have  $\langle \xi_m \bar{\xi}_n \rangle = -(A^{-1})_{mn}$ .

Employing the representation (128) for the determinant of a matrix, we can write the Jacobian that appears in the transformation (18) in the following way. First, we recall that the non-zero elements of the Jacobian are

$$\partial \eta_n / \partial \phi_{n+1} = 1 - \Delta \mathcal{L}'_{n+1}, \quad (129)$$

$$\partial \eta_n / \partial \phi_n = -1 - \Delta(1 - \lambda) f'_n \quad (130)$$

and therefore

$$J[\phi] = \int D[\xi \bar{\xi}] \exp \left\{ \sum_n \bar{\xi}_n (1 - \Delta \mathcal{L}'_{n+1}) \xi_{n+1} + \sum_n \bar{\xi}_n [-1 - \Delta(1 - \lambda) f'_n] \xi_n \right\}, \quad (131)$$

where  $\xi_n$  (with  $n = 1, \dots, M$ ) and  $\bar{\xi}_n$  (with  $n = 0, \dots, M - 1$ ) are Grassmann numbers.

The dynamical partition function is defined as before in (26) but with the average now also involving integration over the Grassmann variables,

$$\langle \dots \rangle_S = \int \frac{d\phi d\hat{\phi}}{(2\pi)^M} \prod_n d\xi_n d\bar{\xi}_n \dots \exp(-S) \equiv \int D[\phi\hat{\phi}] D[\xi\bar{\xi}] \dots \exp(-S). \quad (132)$$

The action reads

$$S = \sum_n (T\Delta\hat{\phi}_n^2 - i\hat{\phi}_n \{-\phi_{n+1} + \phi_n + \Delta[(1-\lambda)f_n + \lambda f_{n+1}]\}) - \sum_n \bar{\xi}_n (1 - \Delta\lambda f'_{n+1}) \xi_{n+1} - \sum_n \bar{\xi}_n (-1 - \Delta(1-\lambda)f'_n) \xi_n, \quad (133)$$

where the second line replaces the last term in (25). In the continuous time limit  $\Delta \rightarrow 0$  this can be written as

$$S = \int dt \{T\hat{\phi}^2 + i\hat{\phi}[\partial_t \phi - f(\phi)]\} - \int dt \bar{\xi}[\partial_t - f'(\phi)]\xi. \quad (134)$$

Let us see how the inclusion of the Grassmann ‘ghosts’ works out for our example case of  $f(\phi) = -\mu\phi - (g/3!)\phi^3$ . Going back to discretized time temporarily will help us understand how to treat equal-time correlations. With  $f(\phi)$  as given, the action can be written as

$$S = S_0 + S_{\text{int}}, \quad (135)$$

$$S_0 = \sum_n (T\Delta\hat{\phi}_n^2 - i\hat{\phi}_n \{-\phi_{n+1} + \phi_n - \mu\Delta[(1-\lambda)\phi_n + \lambda\phi_{n+1}]\}) - \sum_n \{\bar{\xi}_n (1 + \Delta\lambda\mu)\xi_{n+1} + \bar{\xi}_n [-1 + \Delta(1-\lambda)\mu]\xi_n\}, \quad (136)$$

$$S_{\text{int}} = i\frac{g}{3!}\Delta \sum_n \hat{\phi}_n [(1-\lambda)\phi_n^3 + \lambda\phi_{n+1}^3] - \Delta\frac{g}{2} \sum_n [\bar{\xi}_n \lambda\phi_{n+1}^2 \xi_n + \bar{\xi}_n (1-\lambda)\phi_n^2 \xi_n]. \quad (137)$$

The coefficient matrix  $A$  appearing in the ghost term of the bare action  $S_0$  has entries  $1 + \Delta\lambda\mu$  on the main diagonal and  $-1 + \Delta(1-\lambda)\mu$  on the diagonal below. This matrix is easily inverted to show that the ghost covariance is

$$\langle \xi_m \bar{\xi}_n \rangle_0 = -\frac{1}{1 + \Delta\lambda\mu} \left( \frac{1 - \Delta(1-\lambda)\mu}{1 + \Delta\lambda\mu} \right)^{m-n-1} = -\exp[-\mu(m-n-1)\Delta] \quad (138)$$

for  $m > n$  and 0 otherwise; the last expression applies for  $\Delta \rightarrow 0$ . The ghost correlator is therefore causal and reads in the continuum limit:

$$\langle \xi(t) \bar{\xi}(t') \rangle_0 = -\Theta(t-t') \exp[-\mu(t-t')]. \quad (139)$$

While this is  $\lambda$ -independent, the dependence on  $\lambda$  reappears in how equal-time Wick contractions are treated in the perturbative expansion in powers of  $-S_{\text{int}}$ . To see this, note that up to vanishing corrections the interacting action is

$$S_{\text{int}} = i\frac{g}{3!}\Delta \sum_n [\lambda\hat{\phi}_{n-1} + (1-\lambda)\hat{\phi}_n] \phi_n^3 - \Delta\frac{g}{2} \sum_n [\lambda\bar{\xi}_{n-1} + (1-\lambda)\bar{\xi}_n] \phi_n^2 \xi_n. \quad (140)$$

The square brackets in the first line, including the factor  $i$ , define what we previously called the response field  $\eta_{2n}$ , which has equal-time correlator with  $\phi_n$  of  $R_m^0 = \lambda$ . Similarly we could now define the combination in square brackets in the second line as a new Grassmann

response field  $\tilde{\xi}_n$ , which has equal-time correlator with  $\xi$  of  $\langle \xi_n \tilde{\xi}_n \rangle_0 = \lambda \langle \xi_n \tilde{\xi}_{n-1} \rangle_0 = -\lambda$  using (138). If for simplicity of notation we do not distinguish between  $\tilde{\xi}$  and  $\bar{\xi}$  (and similarly  $\eta_2$  and  $\hat{\phi}$ ) then the upshot of this discussion is that for  $\Delta \rightarrow 0$  one can work directly with the continuous-time version

$$S_{\text{int}} = \frac{g}{3!} \int dt (i\hat{\phi}\phi^3 - 3\bar{\xi}\phi^2\xi) \quad (141)$$

of the interacting action, provided that one remembers that the equal-time ghost correlator has to be set to  $\langle \xi(t)\bar{\xi}(t) \rangle_0 = -\lambda$  in any Wick contractions, and similarly  $\langle \phi(t)i\hat{\phi}(t) \rangle_0 = \lambda$ . Note that there are never any contractions between ordinary and Grassmann variables because the average of any single Grassmann variable over a (Grassmann) Gaussian vanishes.

To see how the perturbation theory with ghosts works in practice, consider the response function to first order in the perturbation:

$$\langle \phi(t)\hat{\phi}(t') \rangle_S = \langle \phi(t)\phi(t') \exp[-S_{\text{int}}] \rangle_0 = \langle \phi(t)\hat{\phi}(t') [1 - S_{\text{int}}] \rangle_0 + \mathcal{O}(g^2) \quad (142)$$

$$\begin{aligned} &= \langle \phi(t)\hat{\phi}(t') \rangle_0 - \frac{ig}{3!} \int dt'' \langle \phi(t)\hat{\phi}(t')\hat{\phi}(t'')\phi^3(t'') \rangle_0 \\ &\quad + \frac{g}{2} \int dt'' \langle \phi(t)\hat{\phi}(t')\phi^2(t'')\bar{\xi}(t'')\xi(t'') \rangle_0 + \mathcal{O}(g^2). \end{aligned} \quad (143)$$

The physical piece of this is the first term and the contraction without equal-time response factors in the first integral:

$$\langle \phi(t)\hat{\phi}(t') \rangle = \langle \phi(t)\hat{\phi}(t') \rangle_0 - \frac{ig}{2} \int dt'' \langle \phi(t)\hat{\phi}(t'') \rangle_0 \langle \phi(t'')\hat{\phi}(t') \rangle_0 \langle \phi^2(t'') \rangle_0. \quad (144)$$

When  $\lambda \neq 0$ , i.e. for any convention other than Ito, there are two other non-zero contractions of the first integral:

$$\begin{aligned} &-\frac{g}{2} \langle \phi(t)\hat{\phi}(t') \rangle_0 \int dt'' \langle \phi^2(t'') \rangle_0 \langle \phi(t'')i\hat{\phi}(t'') \rangle_0 \\ &\quad - g \int dt'' \langle \phi(t)\phi(t'') \rangle_0 \langle \phi(t'')\hat{\phi}(t') \rangle_0 \langle \phi(t'')i\hat{\phi}(t'') \rangle_0. \end{aligned} \quad (145)$$

In addition the two possible Wick contractions of the ghost term in the second line of (143) give, bearing in mind that  $\langle \bar{\xi}\xi \rangle_0 = -\langle \xi\bar{\xi} \rangle_0$ ,

$$\begin{aligned} &-\frac{g}{2} \langle \phi(t)\hat{\phi}(t') \rangle_0 \int dt'' \langle \phi^2(t'') \rangle_0 \langle \xi(t'')\bar{\xi}(t'') \rangle_0 \\ &\quad - g \int dt'' \langle \phi(t)\phi(t'') \rangle_0 \langle \phi(t'')\hat{\phi}(t') \rangle_0 \langle \xi(t'')\bar{\xi}(t'') \rangle_0. \end{aligned} \quad (146)$$

Because  $\langle \phi i\hat{\phi} \rangle_0 = \lambda$  at equal-time while the ghost correlator has the *opposite* sign  $\langle \xi\bar{\xi} \rangle_0 = -\lambda$ , the terms in (145) and (146) exactly cancel each other as they should.

In other words, whether or not we include the Jacobian, the extra terms (145) do not appear in the perturbation theory, either because they are simply equal to zero, or because they cancel with the additional terms (146) that arise from the ghost correlation functions from the Grassmann representation of the Jacobian. As will be elaborated in the next section, this is formally a consequence of a symmetry of the theory that represents the fact that the path probabilities are normalized to one.



## 7.2. Manifestly supersymmetric description

Let us now define the *superfield*  $\Phi$  as

$$\Phi = \phi + \bar{\theta}\xi + \bar{\xi}\theta + i\bar{\theta}\theta\hat{\phi}, \quad (147)$$

where  $\theta$  and  $\bar{\theta}$  are themselves Grassmann variables and are sometimes referred to as Grassmann time. The action (134) can then be written in terms of  $\Phi$  in a compact way that reveals unexpected symmetries of the problem.

To see this, it is convenient to define

$$V(\phi) = -\int^{\phi} d\phi' f(\phi'), \quad (148)$$

so  $f(\phi) = -dV(\phi)/d\phi$ . The existence of the potential  $V$  is crucial for the supersymmetric description; it cannot be used for systems of more than one variable when these are non-equilibrium in the sense that the drift  $f$  cannot be written as a gradient.

In our current one-dimensional example where a potential can always be defined we have, by Taylor expansion around  $V(\phi)$  and throwing away terms that vanish because of (122),

$$V(\Phi) = V(\phi) - (\bar{\theta}\xi + \bar{\xi}\theta + i\bar{\theta}\theta\hat{\phi})f(\phi) - \frac{1}{2}(\bar{\theta}\xi + \bar{\xi}\theta + i\bar{\theta}\theta\hat{\phi})^2 f'(\phi) \quad (149)$$

$$= V(\phi) - (\bar{\theta}\xi + \bar{\xi}\theta + i\bar{\theta}\theta\hat{\phi})f(\phi) + \bar{\theta}\bar{\xi}\xi f'(\phi), \quad (150)$$

which leads to

$$\int d\theta d\bar{\theta} V(\Phi) = -i\hat{\phi}f(\phi) + \bar{\xi}f'(\phi)\xi, \quad (151)$$

giving us two of the terms in (134).

The remaining terms can be written in terms of derivatives of  $\Phi$ . We have

$$\partial_{\theta}\Phi = -\bar{\xi} - i\bar{\theta}\hat{\phi}, \quad (152)$$

$$\partial_{\bar{\theta}}\Phi = \xi + i\theta\hat{\phi}, \quad (153)$$

$$\theta\partial_t\Phi = \theta\partial_t\phi + \theta\bar{\theta}\partial_t\xi. \quad (154)$$

If we then evaluate the quantity

$$T \int d\theta d\bar{\theta} (\partial_{\theta}\Phi)(\partial_{\bar{\theta}}\Phi), \quad (155)$$

we find that we get the  $T\hat{\phi}^2$  term in the action (134). Similarly, we find

$$\int d\theta d\bar{\theta} (\partial_{\theta}\Phi)(\theta\partial_t\Phi) = -i\hat{\phi}\partial_t\phi + \bar{\xi}\partial_t\xi, \quad (156)$$

which are the negatives of the terms in (134) involving time derivatives. Putting all these results together and defining  $d\tau \equiv dt d\theta d\bar{\theta}$ , we can write the action in the form

$$S = \int d\tau \{ \partial_{\theta}\Phi [T\partial_{\bar{\theta}}\Phi - \theta\partial_t\Phi] + V(\Phi) \}. \quad (157)$$

It will be handy to introduce the notation

$$D = T\partial_{\bar{\theta}} - \theta\partial_t, \quad (158)$$

$$\bar{D} = \partial_\theta \tag{159}$$

so that

$$S = \int d\tau [\bar{D}\Phi D\Phi + V(\Phi)]. \tag{160}$$

Up to here the formalism is general enough—subject to the existence of the potential  $V$ —that it can describe also non-stationary dynamics, e.g. relaxation to equilibrium. From now on we restrict ourselves further by assuming we have a stationary state. The supersymmetric action then has several symmetries; the obvious one is time translation invariance. But it is also invariant under other ‘translations’ that involve shifts in the Grassmann times  $\theta$  and  $\bar{\theta}$ . In what follows, we identify these invariances and investigate their physical meanings [15, 24].

Consider a ‘translation’ generated by the operator

$$D' \equiv \partial_{\bar{\theta}}. \tag{161}$$

This produces a shift  $\bar{\theta} \rightarrow \bar{\theta} + \epsilon$  and a change in  $\Phi$ :

$$\Phi \rightarrow \Phi + \epsilon \partial_{\bar{\theta}}\Phi, \tag{162}$$

where  $\epsilon$  is a Grassmann ‘infinitesimal’ that acts like a separate Grassmann variable. Now, by using (153) (or simply by substituting  $\bar{\theta} \rightarrow \bar{\theta} + \epsilon$  in the definition (147) of the superfield), we find

$$\Phi \rightarrow \Phi + \epsilon\xi + i\epsilon\theta\hat{\phi} = (\phi + \epsilon\xi) + \bar{\theta}\xi + (\bar{\xi} + i\epsilon\hat{\phi})\theta + i\bar{\theta}\theta\hat{\phi}. \tag{163}$$

But according to the definition of the superfield, whatever appears in the expression (163) not multiplied by  $\theta$ ,  $\bar{\theta}$  or  $\bar{\theta}\theta$  should be identified as the new  $\phi$ , and whatever appears multiplied (from the right) by  $\theta$  should be identified as the new  $\bar{\xi}$ . The component fields therefore transform as

$$\phi \rightarrow \phi + \epsilon\xi, \tag{164}$$

$$\xi \rightarrow \xi, \tag{165}$$

$$\bar{\xi} \rightarrow \bar{\xi} + i\epsilon\hat{\phi}, \tag{166}$$

$$\hat{\phi} \rightarrow \hat{\phi}. \tag{167}$$

This transformation is called supersymmetric because it mixes ‘bosonic’ degrees of freedom (i.e.  $\phi$  and  $\hat{\phi}$ ) with ‘fermionic’ (i.e. Grassmann) ones.

To verify that such a transformation is indeed a symmetry of the model, we proceed analogously to what we would do to test, say, rotational invariance in an ordinary field theory with vector fields. We start by performing the ‘rotation’ (164)–(167) on the superfield components in each term of the integrand in (160), i.e. we substitute the transformed  $\Phi$  and carry out the required derivatives with respect to time and Grassmann time. In general, this leads to a change in the integrand, which we then integrate over  $\tau$  (i.e. over  $\theta$ ,  $\bar{\theta}$  and  $t$ ). If the result is zero we have a symmetry. As an example, let us see how the shift generated by  $D'$  affects the ‘kinetic’ term  $\bar{D}\Phi D\Phi$  in the action. From (152) we see, using (166) and (167), that

$$\bar{D}\Phi \rightarrow \bar{D}\Phi - i\epsilon\hat{\phi}. \tag{168}$$

For the change in  $D\Phi$ , we see from (153) that the first term  $\partial_{\bar{\theta}}\Phi$  does not change as it involves only  $\xi$  and  $\hat{\phi}$ , and these do not change under (165) and (167). From (154), (164) and (165), the change in  $\theta\partial_t\Phi$  is

$$\theta\partial_t\Phi \rightarrow \theta\partial_t\Phi - \epsilon\theta\dot{\xi}. \quad (169)$$

Putting these contributions together, the change in  $\bar{D}\Phi D\Phi$  is

$$(-\bar{\xi} - i\bar{\theta}\hat{\phi})\epsilon\theta\dot{\xi} - i\epsilon\hat{\phi}[T(\xi + i\theta\hat{\phi}) - \theta\dot{\phi} - \theta\bar{\theta}\dot{\xi}]. \quad (170)$$

Only terms proportional to  $\theta\bar{\theta}$  will survive the integration over  $\theta$  and  $\bar{\theta}$ , but these cancel:

$$-i\epsilon\theta\bar{\theta}\hat{\phi}\dot{\xi} + i\epsilon\theta\bar{\theta}\hat{\phi}\dot{\xi} = 0. \quad (171)$$

A similar calculation shows that the ‘potential’ term  $\int d\tau V(\Phi)$  is also unchanged. Thus, the action  $S$  is invariant under  $D' = \partial_{\bar{\theta}}$ .

An analogous calculation shows that the kinetic term is *not* invariant under shifts generated by  $\bar{D} = \partial_{\theta}$ . However, we can try combining a shift in  $\theta$  with one in time, using the generator

$$\bar{D}' = \partial_{\theta} + \alpha\bar{\theta}\partial_t \quad (172)$$

and see whether there is a value of  $\alpha$  for which  $\bar{D}'\Phi D\Phi$  is invariant. Now the transformation of  $\xi$  acquires a new term proportional to  $\partial_t\phi$ ,

$$\xi \rightarrow \xi + \epsilon(i\hat{\phi} - \alpha\partial_t\phi), \quad (173)$$

and  $\hat{\phi}$  is also changed, proportional to  $\partial_t\bar{\xi}$ :

$$\hat{\phi} \rightarrow \hat{\phi} - i(\alpha\partial_t\bar{\xi})\epsilon. \quad (174)$$

The remaining variables  $\phi$  and  $\bar{\xi}$  transform according to

$$\phi \rightarrow \phi + \bar{\xi}\epsilon, \quad (175)$$

$$\bar{\xi} \rightarrow \bar{\xi}, \quad (176)$$

Then, after some algebra, we find that, under our trial  $\bar{D}'$ ,  $\bar{D}'\Phi D\Phi$  is changed by

$$\begin{aligned} & -\bar{\xi}[(1 - \beta\theta\bar{\theta}\partial_t)(i\hat{\phi} - \alpha\dot{\phi}) + (\alpha - \beta)\theta\partial_t\bar{\xi}]\epsilon - i\bar{\theta}\hat{\phi}(i\hat{\phi} - \alpha\dot{\phi})\epsilon \\ & - i(\alpha - \beta)\hat{\phi}\bar{\theta}\partial_t\bar{\xi}\epsilon - \alpha\bar{\theta}(\partial_t\bar{\xi})\epsilon[\xi + \theta(i\hat{\phi} - \beta\dot{\phi})] \end{aligned} \quad (177)$$

(here  $\beta = 1/T$ ). Integrating over  $\theta$  and  $\bar{\theta}$  leaves

$$-\beta\bar{\xi}\partial_t(i\hat{\phi} - \alpha\dot{\phi})\epsilon - \alpha(\partial_t\bar{\xi})\epsilon(i\hat{\phi} - \beta\dot{\phi}) - i(\alpha - \beta)\hat{\phi}\partial_t\bar{\xi}\epsilon, \quad (178)$$

so we see that with the choice  $\alpha = \beta$  this just reduces to

$$-\beta\partial_t[\bar{\xi}(i\hat{\phi} - \dot{\phi})]\epsilon. \quad (179)$$

This vanishes on integration over  $t$  (for stationary initial and final states), proving the invariance of this part of the action. Again the proof of invariance for the  $V(\Phi)$  term is similar to that for  $D'$ , so the total action is invariant under  $\bar{D}'$ .

The reader who is uneasy with the formal manipulations using superfields here can check these results by applying the transformations (164)–(167) for  $D'$  and (173)–(176) for  $\bar{D}'$  directly to  $\phi$ ,  $\hat{\phi}$ ,  $\xi$  and  $\bar{\xi}$ , in the form (134) of the action not using superfields.

To see the meaning of these supersymmetries, we consider the supercorrelation function

$$Q(1, 2) = \langle \Phi(1)\Phi(2) \rangle, \quad (180)$$

where 1 stands for  $(t_1, \theta_1, \bar{\theta}_1)$  and analogously for 2. When we use (147) and expand the product, many terms vanish, either because the averages are of products of Grassmann variables with ordinary ones or of a pair of  $\hat{\phi}$ 's. The remaining terms are

$$Q(1, 2) = \langle \phi_1 \phi_2 \rangle + \bar{\theta}_1 \theta_2 \langle \xi_1 \bar{\xi}_2 \rangle + \theta_1 \bar{\theta}_2 \langle \bar{\xi}_1 \xi_2 \rangle + i \bar{\theta}_1 \theta_1 \langle \hat{\phi}_1 \phi_2 \rangle + i \bar{\theta}_2 \theta_2 \langle \phi_1 \hat{\phi}_2 \rangle, \quad (181)$$

where  $\phi_1$  means  $\phi(t_1)$ , etc.

From the invariance of the action under  $D'$  (translations in  $\bar{\theta}$ ), we have

$$\begin{aligned} D'Q(1, 2) &= (D'_1 + D'_2)Q(1, 2) = (\partial_{\bar{\theta}_1} + \partial_{\bar{\theta}_2})Q(1, 2) \\ &= \theta_2 (\langle \xi_1 \bar{\xi}_2 \rangle + i \langle \phi_1 \hat{\phi}_2 \rangle) + \theta_1 (-\langle \bar{\xi}_1 \xi_2 \rangle + i \langle \hat{\phi}_1 \phi_2 \rangle) = 0. \end{aligned} \quad (182)$$

The vanishing of the term proportional to  $\theta_2$  says that the ghost correlation function  $\langle \xi_1 \bar{\xi}_2 \rangle$  has to be the negative of the response function  $i \langle \phi_1 \hat{\phi}_2 \rangle$  (as in (139)). The vanishing of the term proportional to  $\theta_1$  says the same thing if we notice that the ghost correlation function here is  $\langle \bar{\xi}_1 \xi_2 \rangle$ , not  $\langle \xi_1 \bar{\xi}_2 \rangle$ . Thus, invariance under  $D'$ , through its enforcement of the cancellation of disconnected diagrams, is just conservation of probability.

Analogously, for the  $\bar{D}'$  symmetry (172), we have

$$\bar{D}'Q(1, 2) = (\bar{D}'_1 + \bar{D}'_2)Q(1, 2) = (\partial_{\theta_1} + \partial_{\theta_2} + \beta \bar{\theta}_1 \partial_{t_1} + \beta \bar{\theta}_2 \partial_{t_2})Q(1, 2) = 0. \quad (183)$$

This gives

$$(\bar{\theta}_1 - \bar{\theta}_2) (\langle \phi_1 \hat{\phi}_2 \rangle - \langle \hat{\phi}_1 \phi_2 \rangle) + \beta \bar{\theta}_1 \partial_{t_1} \langle \phi_1 \phi_2 \rangle + \beta \bar{\theta}_2 \partial_{t_2} \langle \phi_1 \phi_2 \rangle = 0. \quad (184)$$

For  $t_1 > t_2$ ,  $\langle \hat{\phi}_1 \phi_2 \rangle$  vanishes, so we have, also using  $\partial_{t_2} \langle \phi_1 \phi_2 \rangle = -\partial_{t_1} \langle \phi_1 \phi_2 \rangle$  from time translation invariance,

$$(\bar{\theta}_1 - \bar{\theta}_2) (\langle \phi_1 \hat{\phi}_2 \rangle + \beta \partial_{t_1} \langle \phi_1 \phi_2 \rangle) = 0. \quad (185)$$

Thus,

$$\langle \phi_1 \hat{\phi}_2 \rangle = -\beta \partial_{t_1} \langle \phi_1 \phi_2 \rangle, \quad (186)$$

which is the FDT.

To summarize, the theory has three invariances: time translation,  $D'$  and  $\bar{D}'$ .  $D'$  expresses conservation of probability, and  $\bar{D}'$  expresses the FDT, i.e. the fact that the system is in equilibrium.

So far, we have treated a single-site problem. It is straightforward to extend the formalism to multiple degrees of freedom  $\phi_i$ . However, as emphasized before, the supersymmetric construction is permitted only when the drift is the negative gradient of a potential,  $f_i = -\partial V / \partial \phi_i$ . Otherwise the  $\bar{D}'$  supersymmetry fails. This means that the FDT is not obeyed; the system, even though it may possess a steady state, is not in equilibrium. (Of course, it is well-known from simple arguments making no reference to supersymmetry that models with non-gradient drifts, as arising e.g. from asymmetric coupling matrices, do not satisfy the FDT.)

### 7.3. Superdiagrams

In addition to the insight it provides into the symmetries of the problem, the supersymmetric formulation can also be of practical advantage in calculations. For example, in diagrammatic perturbation theory, one needs only draw the diagrams for the static problem. Another example can be found in Biroli's analysis [11] of dynamical TAP equations for the  $p$ -spin glass, where the entire structure of the argument and the equations could be carried over from the static treatment of Plefka [25]. Here we sketch how to do diagrammatic perturbation theory in the superfield language and show, in a simple example, how it reduces to the diagrams we had in the MSR formalism with the  $\langle \phi(t) \phi(t') \rangle$  and  $\langle \phi(t) \hat{\phi}(t') \rangle$  correlation functions [20].

We write our standard model (14) in the form

$$S = S_0 + S_{\text{int}}, \quad (187)$$

with

$$S_0 = \int d\tau \left( \bar{D}\Phi D\Phi + \frac{1}{2}\mu\Phi^2 \right), \quad (188)$$

and

$$S_{\text{int}} = \frac{g}{4!} \int d\tau \Phi^4. \quad (189)$$

To do perturbation theory, we can expand  $\exp(-S_{\text{int}})$  in  $g$  and apply Wick's theorem to evaluate the resulting averages, just as we did in section 5. Wick's theorem holds also for Grassmann variables although in principle one has to be careful with sign changes that arise from changing the order of the variables when performing contractions. For example,  $\langle \xi_1 \bar{\xi}_1 \xi_2 \bar{\xi}_3 \rangle = \langle \xi_1 \bar{\xi}_1 \rangle \langle \xi_2 \bar{\xi}_3 \rangle - \langle \xi_1 \bar{\xi}_3 \rangle \langle \xi_2 \bar{\xi}_1 \rangle$ . Fortunately this is not an issue in our context as we only need averages of powers of the superfield  $\Phi$ , and from (147)  $\Phi$  only contains products of *pairs* of Grassmann variables: commuting such pairs through each other never gives any minus signs.

The first thing we need for the perturbation theory is the correlation function or propagator of the non-interacting system. Integrating (188) by parts, we get

$$S_0 = \int d\tau \left( -\Phi \bar{D}D\Phi + \frac{1}{2}\mu\Phi^2 \right). \quad (190)$$

It is convenient to write this as

$$S_0 = \frac{1}{2} \int d\tau [(-\Phi([\bar{D}, D]_- + [\bar{D}, D]_+)\Phi + \mu\Phi^2), \quad (191)$$

where  $[\dots]_-$  and  $[\dots]_+$  denote the commutator and anti-commutator respectively. It is then simple to show that

$$[\bar{D}, D]_+ = -\partial_t \quad (192)$$

and

$$[\bar{D}, D]_- = 2T\partial_\theta\partial_{\bar{\theta}} + 2\theta\partial_\theta\partial_t - \partial_t. \quad (193)$$

The term in (191) involving the anticommutator can be neglected because it vanishes on time integration, so we can write  $S_0$  in the appealing form

$$S_0 = \frac{1}{2} \int d\tau \Phi(-D^{(2)} + \mu)\Phi, \quad (194)$$

with<sup>9</sup>

$$D^{(2)} \equiv [\bar{D}, D]_-. \quad (195)$$

From this, we identify

$$Q_0^{-1} = -D^{(2)} + \mu \quad (196)$$

<sup>9</sup> We could of course add any term proportional to  $\partial_t$  to  $D^{(2)}$  without changing the action, i.e. we could put *any* coefficient  $a$  in front of the  $\partial_t$  term in (193) and the action would remain unchanged. Choosing  $a = 2$  would correspond to including the anticommutator term from (191). However, the choice  $a = 1$ , i.e.  $D^{(2)} = [\bar{D}, D]_-$ , will prove convenient when we want to invert this operator to find the unperturbed superfield correlation function.

as the inverse of the free propagator:

$$(-D^{(2)} + \mu)Q_0(1, 2) = \delta(1, 2) \equiv \delta(t_1 - t_2)(\bar{\theta}_1 - \bar{\theta}_2)(\theta_1 - \theta_2), \quad (197)$$

where we have used the fact that  $(\bar{\theta}_1 - \bar{\theta}_2)(\theta_1 - \theta_2)$  acts as a delta-function in the Grassmann times. Now, multiplying by  $(D^{(2)} + \mu)$  from the left, using the fact that  $(D^{(2)})^2 = \partial_t^2$ , and Fourier transforming in time, we arrive at

$$Q_0(\theta_1, \bar{\theta}_1, \theta_2, \bar{\theta}_2; \omega) = \frac{1}{\omega^2 + \mu^2} [2T + (\bar{\theta}_1 - \bar{\theta}_2)(\theta_1 - \theta_2)\mu + (\bar{\theta}_1 + \bar{\theta}_2)(\theta_1 + \theta_2)i\omega]. \quad (198)$$

Back in the time domain, this is

$$Q_0(1, 2) = \frac{T}{\mu} \exp(-\mu|t|) + (\bar{\theta}_2 - \bar{\theta}_1)\theta_2 \exp(-\mu t)\Theta(t) + (\bar{\theta}_1 - \bar{\theta}_2)\theta_1 \exp(\mu t)\Theta(-t), \quad (199)$$

where  $t = t_1 - t_2$ . It is comforting to confirm that we can get this result in another, simpler way, using the representation (181) with the equalities implied by  $D'$  invariance (182):

$$Q(1, 2) = \langle \phi_1 \phi_2 \rangle + (\bar{\theta}_2 - \bar{\theta}_1)(\theta_2 \langle \phi_1 i\hat{\phi}_2 \rangle - \theta_1 \langle i\hat{\phi}_1 \phi_2 \rangle). \quad (200)$$

Using the free correlation and response functions (3) and (4) here then leads to (199).

To see how the diagrammatics work in this formalism, consider the second-order watermelon graph (113) for the self-energy,

$$\Sigma(1, 2) = \left(-\frac{g}{4!}\right)^2 \times 4 \times 4 \times 3! \times Q(1, 2)^3 = \frac{g^2}{3!} Q(1, 2)^3. \quad (201)$$

(Here we are doing the resummed expansion in which  $\Sigma$  is a functional of the full correlation function  $Q$ , not just  $Q_0$ , as in section 6.)

From the representation (200), the part of  $Q(1, 2)$  with no Grassmann times multiplying it is the correlation function, and the parts multiplied by  $\pm(\bar{\theta}_2 - \bar{\theta}_1)\theta_{1,2}$  are the retarded and advanced response functions, respectively. Expanding  $Q(1, 2)^3$ , we get

$$\Sigma(1, 2) = \frac{g^2}{3!} [\langle \phi_1 \phi_2 \rangle^3 + 3 \langle \phi_1 \phi_2 \rangle^2 (\bar{\theta}_2 - \bar{\theta}_1)(\theta_2 \langle \phi_1 i\hat{\phi}_2 \rangle - \theta_1 \langle i\hat{\phi}_1 \phi_2 \rangle)]. \quad (202)$$

This has the same form as (200). We note that the components of  $\Sigma(1, 2)$  here are just the  $\mathcal{O}(g^2)$  contributions to the self-energies  $\Sigma_{12}$ ,  $\Sigma_{21}$  and  $\Sigma_{22}$  that we found in the conventional MSR theory (116) and (117), including the factor of 3 in  $\Sigma_{12}$  and  $\Sigma_{21}$ .

It is useful to discuss in more detail how perturbation-theoretic corrections to  $Q(1, 2)$  in the supersymmetric formulation correspond to results obtained in the conventional MSR theory. The two theories differ superficially in two ways: (1) In the conventional theory we have to keep track of two kinds of correlations functions ( $\langle \phi \phi \rangle$  and  $i \langle \phi \hat{\phi} \rangle$ ) while in the supersymmetric formulation we have only one (super)field and need only draw the diagrams we would have in statics. (2) In the supersymmetric theory both real and Grassmann times of intermediate vertices in the graphs are integrated over, while in the conventional theory only ordinary times are integrated over. To compare the two ways of doing the calculation, we consider the first term obtained in expanding the Dyson equation in  $\Sigma(1, 2)$ :

$$\Delta Q(1, 4) = \int d\tau_2 d\tau_3 Q(1, 2)\Sigma(2, 3)Q(3, 4). \quad (203)$$

To resolve  $\Delta Q(1, 4)$  into components, we use first the fact that Grassmann factors of the form  $(\bar{\theta}_2 - \bar{\theta}_1)\theta_2$  are idempotent under convolution, e.g.

$$\int d\theta_2 d\bar{\theta}_2 (\bar{\theta}_2 - \bar{\theta}_1)\theta_2(\bar{\theta}_3 - \bar{\theta}_2)\theta_3 = (\bar{\theta}_3 - \bar{\theta}_1)\theta_3. \quad (204)$$

That means that the retarded part of  $\Delta Q(1, 4)$  (the part proportional to  $(\bar{\theta}_4 - \bar{\theta}_1)\theta_4$ ) is, in the notation we have used earlier ( $\langle\langle\phi(t_1)\phi(t_2)\rangle\rangle = C(t_1, t_2)$ ,  $i\langle\phi(t_1)\hat{\phi}(t_2)\rangle = R(t_1, t_2)$ )

$$\int dt_2 dt_3 R(t_1, t_2) \left[ \frac{g^2}{3!} 3(C(t_2, t_3))^2 R(t_2, t_3) \right] R(t_3, t_4), \quad (205)$$

and analogously for the advanced part.

We also get a contribution to  $\Delta Q(1, 4)$  from the first term in  $\Sigma(2, 3)$ . Since that term contains no Grassmann times, this contribution will contain a factor

$$\int d\theta_2 d\bar{\theta}_2 Q(1, 2) \cdot \int d\theta_3 d\bar{\theta}_3 Q(3, 4). \quad (206)$$

These integrations pick out factors of the retarded function  $R(t_1, t_2)$  and the advanced function  $R(t_4, t_3)$ , respectively, so we find a contribution, involving no Grassmann times, of

$$\int dt_2 dt_3 R(t_1, t_2) \left[ \frac{g^2}{3!} (C(t_2, t_3))^3 \right] R(t_4, t_3). \quad (207)$$

These are exactly the second-order contributions to  $R(t_1, t_4)$  and  $C(t_1, t_4)$  that we would find in the conventional formulation from expanding the Dyson equation to first order in the self-energies  $\Sigma_{12}$  and  $\Sigma_{22}$  in (116) and (117), again up to  $\mathcal{O}(g)$  terms not written explicitly here. Thus, because of the algebra of the Grassmann times in the supersymmetric formulation, the results of the multiplication and convolution of the supercorrelation functions, when reduced to components, reproduce the terms found in the conventional MSR theory. This result extends to all graphs in perturbation theory, because it depends only on (1) the idempotency of factors like  $(\bar{\theta}_2 - \bar{\theta}_1)\theta_2$  and (2) the fact that multiple correlator lines between a pair of interaction vertices, like those in (201), combine as in (202).

## 8. An interacting example

To generalize the discussion so far to more interesting interacting models one can for example add a linear interaction between different degrees of freedom or ‘soft spins’  $\phi_i$ . This gives the Langevin equation of motion

$$\partial_t \phi_i = -\mu \phi_i - \frac{g}{3!} \phi_i^3 + \sum_j J_{ij} \phi_j + h_i + \zeta_i(t), \quad (208)$$

where we assume that there are no self-interactions, hence  $J_{ii} = 0$ . If the couplings are otherwise symmetric,  $J_{ij} = J_{ji}$ , then this dynamics obeys detailed balance because it represents noisy gradient descent  $\partial_t \phi_i = -\partial_{\phi_i} H + \zeta_i(t)$  on the energy function

$$H = \sum_i \left( \frac{\mu}{2} \phi_i^2 + \frac{g}{4!} \phi_i^4 \right) - \frac{1}{2} \sum_{ij} J_{ij} \phi_i \phi_j - \sum_i h_i \phi_i \quad (209)$$

which can be thought of as a soft spin version of the Sherrington–Kirkpatrick model [26]. The diagrammatic technique in its MSR incarnation (including fermionic ghost variables, if one chooses to follow the Stratonovich convention) can be applied irrespective of any such

restriction, i.e. whether or not the system obeys detailed balance. For a supersymmetric treatment interaction symmetry is necessary, on the other hand.

The generating functional has the form (26) with action, written directly in the continuous time limit for  $\hbar = 0$  and using the Ito convention,

$$S = \int dt \sum_i \left[ T \hat{\phi}_i^2 + i \hat{\phi}_i \left( \partial_t \phi_i + \mu \phi_i + \frac{g}{3!} \phi_i^3 - \sum_j J_{ij} \phi_j \right) \right]. \quad (210)$$

In this action the interaction gives an additional vertex with two legs,  $\int dt \sum_{ij} J_{ij} \hat{\phi}_i(t) \phi_j(t)$ . (As before we do not introduce a new symbol for this vertex as the meaning is clear from the number of legs attached.) The diagrammatic expansion now becomes a joint expansion in both  $g$  and the interaction amplitude; formally one could set  $J_{ij} = J \hat{J}_{ij}$ , consider the  $\hat{J}_{ij}$  fixed and expand in  $J$ . To illustrate the new diagrams that appear we restrict ourselves to  $g = 0$ . The expansion of the propagator to second order in  $J$  is then simply

$$\text{====} = \text{——} + \text{—} \bullet \text{—} + \text{—} \bullet \bullet \text{—} + \mathcal{O}(J^3).$$

All propagators and vertices now carry site indices in addition to the time index and the ‘sector’ label 1 or 2 for physical ( $\phi$ ) and conjugate ( $\hat{\phi}$ ) fields. If we focus on the response function part of the propagator, the expansion becomes

$$\text{====} \blacktriangleright = \text{—} \blacktriangleright + \text{—} \bullet \blacktriangleright + \text{—} \bullet \bullet \blacktriangleright + \mathcal{O}(J^3). \quad (211)$$

Writing out the diagrams this translates to

$$R_{ij}(t, t') = R_{ij0}(t, t') + \int dt_1 R_{i l 0}(t, t_1) J_{lk} R_{k j 0}(t_1, t') \\ + \int dt_1 dt_2 R_{i n 0}(t, t_2) J_{nm} R_{m l 0}(t_2, t_1) J_{lk} R_{k j 0}(t_1, t') + \mathcal{O}(J^3), \quad (212)$$

where all internal site indices ( $k, l, m, n$ ) are to be summed over. Here  $R_{ij0}(t, t')$  is the response function of the unperturbed dynamics, which because of the absence of interactions is diagonal in the site indices,  $R_{ij0}(t, t') = \delta_{ij} R_0(t, t')$ . The time integrals become simple products in frequency space, giving for the Fourier transform  $R_{ij}(\omega) = \int dt R_{ij}(t, t') e^{i\omega(t-t')}$

$$R_{ij}(\omega) = R_{ij0}(\omega) + R_{i l 0}(\omega) J_{lk} R_{k j 0}(\omega) \\ + R_{i n 0}(\omega) J_{nm} R_{m l 0}(\omega) J_{lk} R_{k j 0}(\omega) + \mathcal{O}(J^3) \quad (213)$$

or in matrix form

$$\mathbf{R}(\omega) = \mathbf{R}_0(\omega) + \mathbf{R}_0(\omega) \mathbf{J} \mathbf{R}_0(\omega) + \mathbf{R}_0(\omega) \mathbf{J} \mathbf{R}_0(\omega) \mathbf{J} \mathbf{R}_0(\omega) + \mathcal{O}(J^3) \quad (214)$$

$$= [\mathbf{R}_0^{-1}(\omega) - \mathbf{J}]^{-1}. \quad (215)$$

Inserting  $\mathbf{R}_0^{-1}(\omega) = (\mu - i\omega) \mathbf{I}$  where  $\mathbf{I}$  is the identity matrix gives  $\mathbf{R}(\omega) = [(\mu - i\omega) \mathbf{I} - \mathbf{J}]^{-1}$ . Given that we are considering  $g = 0$  where the dynamics is purely linear, this is easily seen to be the exact result. This was possible to obtain here because we were able to sum up all diagrams, or equivalently because only a single diagram (the bare quadratic vertex) contributes to the self-energy.

Up to here the discussion applies for general  $J_{ij}$ . To simplify further one needs to make assumptions on the statistics of these interactions. One interesting case is that of a soft-spin Sherrington–Kirkpatrick model for which the  $J_{ij}$  are zero-mean Gaussian random variables, uncorrelated for different index pairs  $ij$ , except that the symmetry in the interaction matrix is imperfect:



$$\overline{J_{ij}^2} = \frac{J^2}{N}, \quad \overline{J_{ij}J_{ji}} = \frac{\kappa J^2}{N}. \quad (216)$$

The symmetry parameter has the value  $\kappa = 1$  for a fully symmetric matrix, while  $\kappa = 0$  gives a fully asymmetric matrix.

For the local response function  $R_{ii}$  one can in general simplify (213) to

$$R_{ii}(\omega) = R_0(\omega) + R_0(\omega)J_{im}R_0(\omega)J_{mi}R_0(\omega) + \mathcal{O}(J^3), \quad (217)$$

where the term of first order in  $J_{ij}$  vanishes due to the lack of self-interactions. For the soft-spin SK model, the sum  $\sum_m J_{im}J_{mi}$  in the second order term has average  $(N-1)\kappa J^2/N = \kappa J^2 + \mathcal{O}(1/N)$  while its variance is  $\mathcal{O}(1/N)$ . For large  $N$  it is therefore self-averaging, i.e. equal to  $\kappa J^2$  with probability one. Hence

$$R_{ii}(\omega) = R_0(\omega) + \kappa J^2 R_0^3(\omega) + \mathcal{O}(J^3). \quad (218)$$

As one might have expected, because this model has each node interacting with all others the nodes become equivalent, making the local response functions  $R_{ii}$  independent of  $i$ . It remains true at higher orders that the local response only depends on the overall coupling amplitude  $J$  and the symmetry parameter  $\kappa$ . This can be shown by a separate diagrammatic argument [27–29] or by explicit averaging of the path generating function  $Z$  over the disorder, i.e. the statistics of the  $J_{ij}$ . This latter approach is the subject of the following section.

## 9. Quenched averaged dynamics of soft spin models

In the previous sections, we described how the dynamics of a set of interacting scalar variables  $\phi_i$  that evolve according to a stochastic differential equation can be studied perturbatively using the path integral formalism. When the evolution equations depend on random quantities such as the  $J_{ij}$  above, then instead of studying a single system with fixed  $J_{ij}$  one can consider an ensemble of systems, each with a different set of interactions  $J_{ij}$  drawn from some distribution. The resulting ensemble averages are expected to reflect the behavior of a typical sample as far as macroscopic quantities such as the average local response  $(1/N)\sum_i R_{ii}$  are concerned: these are *self-averaging*, i.e. to leading order dependent only on the overall statistics of the interactions in the system. Where the interactions are weak and long-ranged as in the soft-spin SK model, all spins also become equivalent so even the local responses  $R_{ii}$  are self-averaging as we saw above. This would not be true, for example, for systems with interactions on a network with finite connectivity.

We illustrate the approach for the dynamics given by (208), with dynamic action in (210) and assuming Gaussian statistics for the  $J_{ij}$  as specified in (216). The derivation that we outline here was first described in [4, 30] for symmetric interactions; the role of the degree of asymmetry in the interactions was analysed later in [31].

Using the fact that the average of  $\exp(i\mathbf{x} \cdot \mathbf{z})$  over a vector of zero mean Gaussian random variables  $\mathbf{z}$  with covariance  $\mathbf{A}^{-1}$  is  $\exp(-\mathbf{x} \cdot \mathbf{A}^{-1}\mathbf{x}/2)$ , the  $\mathbf{J}$ -average (denoted by an overline) of the part of the generating functional  $Z$  that depends on the  $J_{ij}$  reads

$$I \equiv \prod_{i < j} \overline{\exp\left[i \int dt (\hat{\phi}_i \phi_j J_{ij} + \hat{\phi}_j \phi_i J_{ji})\right]} \quad (219)$$

$$= \prod_{i < j} \exp \left\{ -\frac{J^2}{2N} \int dt dt' [\hat{\phi}_i(t) \hat{\phi}_i(t') \phi_j(t) \phi_j(t') + \hat{\phi}_j(t) \hat{\phi}_j(t') \phi_i(t) \phi_i(t') + 2\kappa \hat{\phi}_i(t) \phi_j(t) \hat{\phi}_j(t') \phi_i(t')] \right\} \quad (220)$$

$$= \prod_{i \neq j} \exp \left\{ -\frac{J^2}{2N} \int dt dt' [\hat{\phi}_i(t) \hat{\phi}_i(t') \phi_j(t) \phi_j(t') + \kappa \hat{\phi}_i(t) \phi_i(t') \phi_j(t) \hat{\phi}_j(t')] \right\}. \quad (221)$$

One can drop the restriction  $i \neq j$  for large  $N$  as including the  $i = j$  terms only gives subleading corrections. A mean-field decoupling of the quartic terms gives

$$I = \prod_i \exp(-L_i), \quad (222)$$

$$L_i = \frac{J^2}{2} \int dt dt' [C(t, t') \hat{\phi}_i(t) \hat{\phi}_i(t') - i\kappa R(t, t') \hat{\phi}_i(t) \phi_i(t')]. \quad (223)$$

Here

$$C(t, t') \equiv \frac{1}{N} \sum_j \phi_j(t) \phi_j(t'), \quad R(t, t') \equiv \frac{i}{N} \sum_j \phi_j(t) \hat{\phi}_j(t') \quad (224)$$

are the average local correlation and response functions of the system. They are in principle dependent on the specific trajectory of the system, but from the law of large numbers will assume deterministic values for  $N \rightarrow \infty$ . This heuristic argument can be justified by a formal calculation introducing conjugate order parameters to  $C$  and  $R$  and making a saddle point argument [4, 30, 31]. Because  $I$  now factorizes over sites  $i$ , so does the entire ensemble-averaged partition function  $\bar{Z}$ . The contributions from  $L_i$  give an effective noise with correlation function  $J^2 C(t, t')$ , and a delay term with memory kernel  $\kappa R(t, t')$ .

All spins have the same action, so we can drop the index  $i$  and write down the effective dynamics as

$$\partial_t \phi = -\mu \phi - \frac{g}{3!} \phi^3 + \kappa J^2 \int_0^t dt' R(t, t') \phi(t') + h(t) + \zeta_{\text{eff}}(t), \quad (225)$$

$$\langle \zeta_{\text{eff}}(t) \zeta_{\text{eff}}(t') \rangle = 2T \delta(t - t') + J^2 C(t, t'). \quad (226)$$

The correlation and response need to be found self-consistently from this, and to facilitate calculation of the response we have added back in the field term  $h(t)$ .

Let us now take a moment to study the dynamics of the bare model ( $g = 0$ ) that arises from the above equation. At  $g = 0$ , and assuming that for long times a stationary regime is reached where  $R(t, t') = R(t - t')$ , the equation of motion in frequency space is

$$-i\omega \phi(\omega) = -\mu \phi(\omega) + \kappa J^2 R(\omega) \phi(\omega) + h(\omega) + \zeta_{\text{eff}}(\omega). \quad (227)$$

Averaging both sides over the noise term  $\zeta_{\text{eff}}$  gives for the mean  $m(\omega) = \langle \phi(\omega) \rangle$

$$(-i\omega + \mu - \kappa J^2 R(\omega)) m(\omega) = h(\omega). \quad (228)$$

From this, we find the response function

$$R^{-1}(\omega) = \left[ \frac{\partial m(\omega)}{\partial h(\omega)} \right]^{-1} = \frac{\partial h(\omega)}{\partial m(\omega)} = -i\omega + \mu - \kappa J^2 R(\omega) = R_0^{-1}(\omega) - \kappa J^2 R(\omega). \quad (229)$$

To leading order in  $J^2$  this gives  $R^{-1}(\omega) = R_0^{-1}(\omega) - \kappa J^2 R_0(\omega)$ , which after inverting and re-expanding to  $\mathcal{O}(J^2)$  agrees with the perturbative result (218) as it should.

The critical dynamics of system can be understood via the low frequency behavior of the response function. In fact, one can define a characteristic response time scale as

$$\tau = \frac{\int dt (t - t') R(t - t')}{\int dt R(t - t')} = \left. \frac{1}{iR} \frac{\partial R}{\partial \omega} \right|_{\omega=0}. \quad (230)$$

Taking the  $\omega$ -derivative of both sides of (229) and multiplying by  $R$  gives

$$-\left( \frac{1}{R} \frac{\partial R}{\partial \omega} \right) = -iR - J^2 R^2 \left( \frac{1}{R} \frac{\partial R}{\partial \omega} \right) \quad (231)$$

and evaluating at  $\omega = 0$  one finds

$$-i\tau = -iR(0) - iJ^2 R^2(0)\tau. \quad (232)$$

The response timescale can therefore be expressed as

$$\tau = \frac{R(0)}{1 - J^2 R^2(0)}. \quad (233)$$

In other words, when  $R(0) = J^{-1}$ , the relaxation time scale of the system diverges and the system exhibits critical slowing down. From (229) one sees that the critical value of  $J$  obeys  $J = \mu - J^2/J$ , hence  $J = \mu/2$ . Using the explicit solution of (229), which reads

$$R(\omega) = (2J^2)^{-1} [\mu - i\omega - \sqrt{(\mu - i\omega)^2 - 4J^2}] \quad (234)$$

shows further that at criticality the response function has a singularity  $\propto \omega^{1/2}$  for  $\omega \rightarrow 0$ . In the time domain this corresponds to a power law tail  $R(t - t') \propto (t - t')^{-3/2}$ , which is responsible for the diverging mean timescale.

## 10. Path integrals for hard spin models

So far we have considered models in which the dynamical variables took continuous values. Here we show that, with some modifications, the same approach can be used to study the dynamics of models involving Ising spins.

Consider a system of  $N$  binary spin variables,  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$  with  $\sigma_i = \pm 1$ . We consider synchronous dynamics, where time  $t$  advances in discrete steps and at each time step all spins are updated. Specifically, we assume that the spins decide their states according to the following probability distribution

$$P[\boldsymbol{\sigma}(t + 1) | \boldsymbol{\sigma}(t)] = \prod_i \frac{\exp[h_i(t)\sigma_i(t + 1)]}{2 \cosh(h_i(t))}, \quad (235)$$

$$h_i(t) = h_i^{\text{ext}} + \sum_j J_{ij} \sigma_j(t). \quad (236)$$

If the external fields  $h_i^{\text{ext}}$  are constant in time as written above and if the couplings are symmetric, i.e.  $J_{ij} = J_{ji}$ , this dynamics reaches an equilibrium state where detailed balance is satisfied and configurations are visited according to the Boltzmann distribution  $p(\boldsymbol{\sigma}) \propto \exp(-H)$  with the Hamiltonian [32]

$$H = -\sum_i h_i^{\text{ext}} \sigma_i - \sum_i \text{lc}(h_i(\boldsymbol{\sigma})), \quad (237)$$

where we have introduced the abbreviation  $\text{lc}(x) \equiv \ln(2 \cosh(x))$  and emphasize that it is the total fields  $h_i(\boldsymbol{\sigma})$ , which depend on the configuration, that appear inside the log cosh.

An alternative to synchronous dynamics that is commonly studied in the literature is continuous time Glauber dynamics [33], in which spins are updated individually and at random (exponentially distributed) time intervals. This is achieved by prescribing that in any infinitesimal time interval  $\delta t \rightarrow 0$ , every spin is flipped with probability

$$\delta t \left[ 1 - \frac{1}{2} \sigma_i(t) \tanh(h_i(t)) \right]. \quad (238)$$

Again, for constant external fields and symmetric couplings an equilibrium Boltzmann distribution is reached, but this time with the familiar Ising Hamiltonian

$$H = -\sum_i h_i^{\text{ext}} \sigma_i - \sum_{i < j} J_{ij} \sigma_i \sigma_j. \quad (239)$$

For derivations of these equilibrium distributions see [34, 35]. Although the Glauber dynamics is the one that leads to the usual Ising Hamiltonian, in what follows we focus on the synchronous case for two reasons: (i) the path integral formulation is slightly simpler in terms of notation and (ii) the synchronous case has been the focus of recent work on the dynamics of hard spin models, including our own on the inverse problem of inferring the couplings in the model from the statistics of the spin history [36–38].

### 10.1. Path integral formulation

The generating functional for the dynamics is defined as

$$Z[\psi] = \left\langle \exp \left( \sum_i \psi_i(t) \sigma_i(t) \right) \right\rangle, \quad (240)$$

where the average denoted by  $\langle \dots \rangle$  is over the distribution of trajectories generated according to the probability distribution (235) and the  $\psi_i$  are fields that allow one to obtain the statistics of the  $\sigma_i$  by taking derivatives. The key to writing a path integral representation of the generating functional in this case is to work with the local fields  $h_i(t)$ , which are continuous for  $N \rightarrow \infty$ , and not the original spin variables:

$$\begin{aligned} Z[\psi] &= \text{Tr} \int D[h] \prod_t \delta(\mathbf{h}(t) - \mathbf{h}^{\text{ext}} - \mathbf{J}\boldsymbol{\sigma}(t)) \\ &\times \exp \left( \sum_t [\psi(t+1) + \mathbf{h}(t)] \cdot \boldsymbol{\sigma}(t+1) - \sum_{i,t} \text{lc}(h_i(t)) \right), \end{aligned} \quad (241)$$

where  $\mathbf{J}$  is the interaction matrix,  $\text{Tr}$  indicates a sum over all spin trajectories  $\{\boldsymbol{\sigma}(t)\}$  and similarly  $D[h]$  an integral over all field trajectories  $\{\mathbf{h}(t)\}$ . The discrete time variable range is  $t = 0, 1, \dots, T-1$  where  $T$  is the final time.

The delta function in (241) is introduced to enforce the definition (236) and can be written as a Fourier transform, leading to

$$Z[\psi] = \text{Tr} \int D[h\hat{h}] \exp(-S_\sigma), \quad (242)$$

$$S_\sigma = -\sum_t \mathbf{i}\hat{\mathbf{h}}(t) \cdot [\mathbf{h}(t) - \mathbf{h}^{\text{ext}} - \mathbf{J}\boldsymbol{\sigma}(t)] - \sum_t [\psi(t+1) + \mathbf{h}(t)] \cdot \boldsymbol{\sigma}(t+1) + \sum_{i,t} \text{lc}(h_i(t)). \quad (243)$$

This expression can be used in several ways. One is to average over the distribution of  $\mathbf{J}$ ; the second is to derive mean-field equations for the system. We do not go into the first route here, i.e. the quenched averaged dynamics of the system, as this is both very involved and also discussed in detail elsewhere [39]. However, as an example of how one can use the path integral formulation (242) of the generating functional, we consider in the following subsections the simple saddle point approximation to the path integral as well as possible improvements to this.

### 10.2. Saddle point

To derive the saddle point equations, we first perform the trace in (242), which for a single timestep involves

$$\begin{aligned} & \sum_{\boldsymbol{\sigma}(t)} \exp[-\mathbf{i}\hat{\mathbf{h}}(t) \cdot \mathbf{J}\boldsymbol{\sigma}(t) + (\psi(t) + \mathbf{h}(t-1)) \cdot \boldsymbol{\sigma}(t)] \\ &= \prod_i \sum_{\sigma_i = \pm 1} \exp\left(\left[\psi_i(t) + h_i(t-1) - \mathbf{i} \sum_j \hat{h}_j(t) J_{ji}\right] \sigma_i(t)\right) \end{aligned} \quad (244)$$

$$= \prod_i 2 \cosh(\psi_i(t) + h_i(t-1) - \mathbf{i}(\mathbf{J}^T \hat{\mathbf{h}}(t))_i). \quad (245)$$

Using this we can write

$$Z[\psi] = \int D[h\hat{\mathbf{h}}] \exp(-S), \quad (246)$$

$$\begin{aligned} S &= -\sum_t \mathbf{i}\hat{\mathbf{h}}(t) \cdot [\mathbf{h}(t) - \mathbf{h}^{\text{ext}}] - \sum_{i,t} [\text{lc}(\psi_i(t+1) \\ &+ h_i(t) - \mathbf{i}(\mathbf{J}^T \hat{\mathbf{h}}(t+1))_i) - \text{lc}(h_i(t))]. \end{aligned} \quad (247)$$

The saddle point equations for stationarity of  $S$  with respect to the  $h_i(t)$  and  $\hat{h}_i(t)$  are then

$$\frac{\partial S}{\partial h_i(t)} = -\mathbf{i}\hat{h}_i(t) - \tanh(\psi_i(t+1) + h_i(t) - \mathbf{i}(\mathbf{J}^T \hat{\mathbf{h}}(t+1))_i) + \tanh(h_i(t)) = 0, \quad (248)$$

$$\frac{\partial S}{\partial \hat{h}_i(t)} = -\mathbf{i}[h_i(t) - h_i^{\text{ext}}] + \mathbf{i} \sum_j J_{ij} \tanh(\psi_j(t) + h_j(t-1) - \mathbf{i}(\mathbf{J}^T \hat{\mathbf{h}}(t))_j) = 0. \quad (249)$$

These equations can be written in a simpler form in terms of the magnetizations in the system biased by  $\psi$ , which are generally given by  $m_i(t) = \partial_{\psi_i(t)} \ln Z$ . In the saddle point approximation  $Z \approx Z_s \equiv \exp(-S)$ , and because  $S$  is stationary with respect to  $h$  and  $\hat{h}$ ,

$$m_i^s(t) = -\partial_{\psi_i(t)} S = \tanh(\psi_i(t) + h_i^s(t-1) - \mathbf{i}(\mathbf{J}^T \cdot \hat{\mathbf{h}}^s(t))_i) \equiv \mu_i(t). \quad (250)$$

Here we have used 's' superscripts to denote saddle point values and introduced  $\mu_i(t)$  as a convenient abbreviation for later use. The saddle point equations then simplify to

$$\mathbf{i}\hat{h}_i^s(t) + m_i^s(t+1) - \tanh(h_i^s(t)) = 0, \quad (251)$$

$$h_i^s(t) = h_i^{\text{ext}} + \sum_j J_{ij} m_j^s(t). \quad (252)$$

For the physical dynamics we want the solution at  $\psi = 0$ , which we denote with a ‘0’ superscript. One can show that this has  $\hat{h}_i^0(t) = 0$ . Indeed, if one considers the dynamics over a finite number of timesteps  $t = 1, 2, \dots, T$ , then both  $\mathbf{h}(t)$  and  $\hat{\mathbf{h}}(t)$  are defined over the range  $t = 0, 1, \dots, T - 1$  and accordingly one finds that in the first saddle point equation (248) taken at  $t = T - 1$ , the  $\hat{\mathbf{h}}(t + 1)$ -term inside the tanh is absent. For  $\psi = 0$  this equation then dictates  $\hat{h}_i^0(T - 1) = 0$ , and working backwards in time from there one finds recursively  $\hat{h}_i^0(t) = 0$  for all  $t$ . Accordingly (251) and (252) simplify to the standard mean field equations for the magnetizations,

$$m_i^0(t + 1) = \tanh(h_i^0(t)), \quad (253)$$

$$h_i^0(t) = h_i^{\text{ext}} + \sum_j J_{ij} m_j^0(t). \quad (254)$$

Note that the saddle point value of the action (247) is then zero for  $\psi = 0$ , as expected from the normalization condition  $Z[0] = 1$ .

### 10.3. Beyond the saddle point: naive approach

One can go beyond the saddle point approximation and take into account the Gaussian fluctuations around the saddle point to obtain a better estimate of the generating functional and hence of the equations of motion for the  $m_i(t)$ . The Gaussian corrections to the log generating functional are

$$A[\psi] = \ln \int D[h\hat{h}] \exp\left(-\frac{1}{2}[\delta h \ \hat{h}]^T \cdot \partial^2 S[\delta h \ \hat{h}]\right) = -\frac{1}{2} \ln |\partial^2 S|, \quad (255)$$

where  $\delta h$  indicates the deviation of  $h$  from its saddle point value. We denote by  $\partial^2 S$  the matrix of the second derivatives of  $S$  with respect to  $h$  and  $\hat{h}$  calculated at the saddle point; this has entries

$$\frac{\partial^2 S}{\partial h_i(t) \partial h_j(t')} = \delta_{ij} \delta_{t,t'} [\mu_i^2(t + 1) - \tanh^2(h_i^s(t))], \quad (256)$$

$$\frac{\partial^2 S}{\partial h_i(t) \partial \hat{h}_j(t')} = -i \delta_{ij} \delta_{t,t'} + i \delta_{t+1,t'} J_{ji} [1 - \mu_i^2(t + 1)], \quad (257)$$

$$\frac{\partial^2 S}{\partial \hat{h}_i(t) \partial \hat{h}_j(t')} = \delta_{t,t'} \sum_k J_{ik} J_{jk} [1 - \mu_k^2(t)]. \quad (258)$$

From the corrected log generating functional  $\ln Z[\psi] = \ln Z_s[\psi] + A[\psi]$  we obtain a corrected expression for the magnetization

$$m_i(t) = \left. \frac{\partial \ln Z[\psi]}{\partial \psi_i(t)} \right|_{\psi=0} = m_i^0(t) + \left. \frac{\partial A[\psi]}{\partial \psi_i(t)} \right|_{\psi=0}. \quad (259)$$

To calculate the  $\psi$ -derivative of  $A$  at  $\psi = 0$  we can either appeal to numerical methods or evaluate  $A$  approximately. One relatively simple approximation, which will give us an intuitive feeling for the corrections and turns out to capture the lowest order corrections in  $J$ , is to only keep the equal-time ( $t = t'$ ) elements of the matrix  $\partial^2 S$ , i.e. to discard the second

term in (257). The matrix then separates into blocks corresponding to the different timesteps  $t$ . Each of those blocks is of size  $2N \times 2N$ . Let us order the elements so that the top left  $N \times N$  sub-block contains the entries from (256). This block, which we will denote by  $\alpha_t$ , is diagonal and vanishes at the physical saddle point, where  $\psi_i(t) = 0$  and  $\hat{h}_i^s(t) = \hat{h}_i^0(t) = 0$ . The two off-diagonal blocks (257) are  $-\mathbf{I}$  within our approximation. If we call the bottom right  $N \times N$  block  $\beta_t$ , then we have by a standard determinant block identity

$$A[\psi] = -\frac{1}{2} \sum_t \ln |\mathbf{I} + \alpha_t \beta_t| = -\frac{1}{2} \sum_t \text{Tr} \ln (\mathbf{I} + \alpha_t \beta_t). \quad (260)$$

Near the physical saddle point  $\alpha_t$  is small so we can linearize the logarithm to get

$$A[\psi] \approx -\frac{1}{2} \sum_t \text{Tr}(\alpha_t \beta_t) \quad (261)$$

$$= -\frac{1}{2} \sum_{it} [\mu_i^2(t+1) - \tanh^2(h_i^s(t))] \sum_k J_{ik}^2 [1 - \mu_k^2(t)]. \quad (262)$$

Linearizing then also the diagonal entries of  $\alpha_t$  that appear in the first square bracket, and accordingly setting the last factor to its value at the physical saddle point gives

$$\begin{aligned} A[\psi] &\approx -\sum_{it} (\psi_i(t+1) - \mathbf{i}(\mathbf{J}^T \cdot \hat{\mathbf{h}}^s(t+1))_i) [1 - m_i^0(t+1)^2] \\ &\quad \times m_i^0(t+1) \sum_k J_{ik}^2 [1 - m_k^0(t)^2]. \end{aligned} \quad (263)$$

When we take the derivative of this with respect to  $\psi_i(t)$ , we do in principle get a term from the dependence of  $\hat{\mathbf{h}}^s$  on  $\psi$ . But as  $\hat{\mathbf{h}}^s$  is multiplied by a factor of  $\mathbf{J}$ , this will give us a contribution to the derivative that is of higher order in  $J$  than the main term from the explicit dependence on  $\psi_i(t)$ . Discarding this higher order contribution—as our second approximation in addition to neglecting correlations between different time steps in  $\partial^2 S$ —yields for the derivative

$$A_i(t) \equiv \frac{\partial A[\psi]}{\partial \psi_i(t)} \approx -[1 - m_i^0(t)^2] m_i^0(t) \sum_k J_{ik}^2 [1 - m_k^0(t-1)^2]. \quad (264)$$

Here we have implicitly already set  $\psi = 0$  because we linearized around the physical saddle point throughout. Using this result in (259), we obtain the corrected magnetization as

$$m_i(t) = m_i^0(t) - [1 - m_i^0(t)^2] \left\{ m_i^0(t) \sum_k J_{ik}^2 [1 - m_k^0(t-1)^2] \right\}. \quad (265)$$

This is the result of our naive approach to including Gaussian corrections in the saddle point integral. Note that because of the normalization  $Z[0] = 1$ , we do not expect Gaussian corrections to  $\ln Z[0]$  at the physical saddle point. The approximation we made in evaluating  $A[\psi]$  as a sum of local-in-time terms preserves this requirement:  $A[0] = 0$  as is clear from (260) given that  $\alpha_t = 0$  when  $\psi = 0$ .

#### 10.4. Beyond the saddle point: Plefka and Legendre transform approaches

It is instructive to compare (265) with the so-called dynamical TAP equations for our system [12]

$$m_i^{\text{TAP}}(t) = \tanh \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j^{\text{TAP}}(t) - m_i^{\text{TAP}}(t) \sum_k J_{ik}^2 [1 - m_k^{\text{TAP}}(t-1)^2] \right), \quad (266)$$

which for stationary magnetizations reduce to the better-known equilibrium TAP equations [40, 41]. The negative term inside the hyperbolic tangent is known as the Onsager correction, which improves on the naive estimate of the mean effective field acting on a spin.

One now observes that the term inside the curly brackets on the right-hand side of our simpler corrected equation (265) is exactly of the form of this Onsager correction, so there is a close relation between the two approaches. However, there are two significant differences. First, while the Onsager term in (266) appears inside the  $\tanh$ , in (265) it is outside, as if we had linearized in the correction. This is important: it means there is nothing on the right-hand side of (265) to stop the magnetizations from going outside the physical range  $[-1, 1]$ . The origin of this difference is the fact that while the dynamical TAP method corrects the naive estimate of the *field* acting on a spin, the Gaussian fluctuation approach described in the previous subsection directly estimates corrections for the magnetizations themselves.

The second difference between (265) and (266) is that in the latter it is the corrected magnetizations  $m_i^{\text{TAP}}$  itself that appear in the Onsager term, making the approximation self-consistent. In our naive method, on the other hand, the Onsager term is evaluated at the uncorrected saddle-point magnetizations.

The dynamical TAP equations (266) can be derived by the Plefka approach [25]. This is an elegant method that captures corrections to the effective fields as explained above, ensuring in contrast to (265) that the magnetizations remain in the physical range even when corrections to the naive mean field estimate are large. This is achieved by working not with the generating functional directly (or the Helmholtz free energy in the equilibrium case), but with its Legendre transform with respect to the magnetizations (the analogue of the equilibrium Gibbs free energy). In this way, one essentially approximates the system with an independent spin model in which each spin feels an effective field  $h_i^{\text{eff}}$ . This leads to expressions of the form  $m_i(t) = \tanh(h_i^{\text{eff}}(t - 1))$  for the magnetizations, which therefore always lie between  $-1$  and  $1$ . The effective fields are determined so as to best match the Gibbs free energy of the original model, by a perturbation expansion to second order in  $J$ .

Interestingly, the philosophy of working with the Legendre transform can also be applied to the approach we have described above, which initially has a rather different starting point, namely a saddle point approximation with Gaussian corrections. We show this in the Appendix, where we demonstrate that by switching to the Gibbs free energy and keeping only first order terms in the corrections  $A_i$ , one can retrieve exactly the dynamical TAP equations. We refer the reader to [42, 43] for applications of the same idea in the context of the equilibrium Ising and Potts models.

## 11. Summary and discussion

Path integral methods and the associated perturbative and diagrammatic tools are among the most powerful methods used in almost every part of theoretical physics. Disordered systems, such as spin glasses, neural networks, models of the immune system etc are no exception and, over the years, path integrals have played a significant role in the study of equilibrium and out-of-equilibrium properties of these systems. In disordered systems, by definition, interactions in a single system do not have any trivial structure; there is symmetry only in their statistics, as described by the distribution from which they are drawn. These systems are typically subject to external stochastic forces with both rapidly changing and quenched random components. Furthermore, these sources of randomness are expected to play crucial roles in the physics and should therefore be included explicitly in studying such systems. The path integral and diagrammatic methods treated in textbooks on field theoretic tools for other physical systems usually do not deal with disordered systems. Therefore, we tried in this



paper to describe some of the key field theoretic and path integral techniques in a manner that can be applied directly to disordered systems.

We started by studying the dynamics of a scalar degree of freedom evolving according to a Langevin equation and showed how one can do perturbative calculations for this model and represent them in diagrammatic form. We then discussed the supersymmetries of the action that features in the path integral description of this system, the physical intuition and meaning behind them, and how they can help in doing diagrammatic perturbation theory by reducing the graphs to be considered to those in simpler, well-understood equilibrium calculations. Our next step was to study the dynamics of systems of interacting ‘soft-spin’ variables subject to Langevin equations with random interactions, first using a perturbative treatment for a single sample of the interactions, and then via a conceptually different approach based on averaging the generating function in its path integral formulation over the disorder.

Finally we switched to hard-spin models, focusing on the Ising model with synchronous update dynamics, for a single sample with arbitrary couplings. Here, as opposed to the way we treated soft-spin models, the path integral is not written in terms of the spins directly but in terms of the fields acting on them. We discussed how, from such a path integral formulation, one can derive approximate equations for the mean magnetizations at the saddle point level and compute the Gaussian fluctuation corrections around the saddle point. We showed that a naive calculation of these corrections yields equations of motion for the magnetizations that can lead to unphysical predictions. These issues can be avoided by going to improved approximation schemes like dynamical TAP equations, as described recently in other papers using a path integral formulation [12, 44], as well as several alternative approaches [38, 40, 45]. We closed the paper with the intriguing observation, however, that issues with naively corrected mean-field equations can also be cured within the general approach presented in this paper, using a Legendre transform to switch to the Gibbs free energy. It is tempting to infer that the Legendre transform implicitly achieves a resummation of the most important diagrams, but we have not been able to show this explicitly and leave this point as an open question for future work.

Although we have tried to cover what we view as the key concepts in this review, we have had to leave out several important issues to maintain a coherent focus. We list two of these in the following.

*The use of dynamical models in inference.* The path integral approach described here is designed initially for a *forward* problem: given the interactions in a system (as well as external fields etc where relevant), predict the dynamics of the order parameters of interest. In recent years, there has been strong interest in the *inverse* problem: given observations of the order parameter dynamics, find the interactions; for a review see e.g. [46]. This interest has been generated by recent advances in recording technology in various fields of science, allowing massive data sets to be gathered that invite researchers to attempt to reverse-engineer the underlying systems from the observed data.

The path integral formulation for forward problems described here is a natural first step in finding inverse equations, which in the simplest case consist of an inversion of the forward equations. Furthermore, in some cases, the path integral method can be immediately applied to the inverse problem itself. For instance, in the presence of hidden degrees of freedom, when trying to reconstruct the interactions in a spin model by observing only some of the spins and not the others, then calculating the likelihood of the data involves tracing over the trajectory of the hidden spins. This can then be done using the path integral methods discussed in this review [47].

*The extended Plefka expansion.* In a spin system at equilibrium, the magnetization of each spin is really the only parameter of interest, with spin–spin correlations then determined indirectly from the linear response of the system. This is not the case for out-of-equilibrium

dynamics. Here out-of-equilibrium correlation and response functions need to be calculated in addition to the magnetizations in order to achieve a full statistical description of the system. So while at equilibrium and also in the simplest form of the dynamical TAP approach for Ising models [12] it makes sense to perform a Plefka expansion by fixing only the magnetization (and in the dynamics, the conjugate fields), in the full dynamical treatment response and correlation functions should also be fixed as order parameters. This is even more important for soft spin models, where even at equilibrium one would want to use at least the means and variances of the local degrees of freedom as order parameters. Such ‘extended’ Plefka expansions have been used for the so-called  $p$ -spin spherical spin glass model in [11] and, more recently, for general systems of coupled Langevin equations [48] and Ising spin glasses [44]. They should be a productive route for further progress in the field, e.g. by extending them further to incorporate inference from observed data [49].

Other interesting current work in this field includes that of Cooper and Dawson [50], Poletini [51], and Nemoto *et al* [52]. These lie beyond the scope of our elementary treatment of the subject, but we hope that our introduction will equip readers with the tools to explore these and other more advanced problems.

## Acknowledgments

This work has been partially supported by the Marie Curie Initial Training Network NETADIS (FP7, grant 290038). YR was also financially supported by the Kavli Foundation, Norwegian Research Council, Center of Excellence Scheme (grant number 223262) and a Starr Foundation fellowship at the Simons Center for Systems Biology, Institute for Advanced Study. The authors are also grateful to Luca Peliti for fruitful discussions. PS acknowledges the stimulating research environment provided by the EPSRC Centre for Doctoral Training in Cross-Disciplinary Approaches to Non-Equilibrium Systems (CANES, EP/L015854/1).

## Appendix

The Legendre transform approach starts from the saddle point value of the log generating functional

$$\begin{aligned} \ln Z_s[\psi] = & \sum_t i\hat{h}^s(t) \cdot [\mathbf{h}^s(t) - \mathbf{h}^{\text{ext}}] \\ & + \sum_{it} \{\text{lc}(\psi_i^s(t+1) + h_i^s(t) - i(\mathbf{J}^T \hat{\mathbf{h}}^s(t+1))_i) - \text{lc}(h_i^s(t))\}. \end{aligned} \quad (\text{A.1})$$

We first rewrite this expression in a form that will simplify some of the algebra below, by decomposing the function  $\text{lc}(\cdot)$  in (A.1) as

$$\text{lc}(x) = H_2[\tanh(x)] + x \tanh(x), \quad (\text{A.2})$$

$$H_2[x] \equiv -\frac{1+x}{2} \ln \frac{1+x}{2} - \frac{1-x}{2} \ln \frac{1-x}{2}. \quad (\text{A.3})$$

Here  $H_2[x]$  is simply the entropy of a single spin with magnetization  $x$ . This decomposition gives, bearing in mind the definition of  $\mu_i(t)$  in (250),

$$\begin{aligned} \ln Z_s[\psi] &= \sum_{it} H_2[\mu_i(t)] + \sum_t i \hat{\mathbf{h}}^s(t) \cdot [\mathbf{h}^s(t) - \mathbf{h}^{\text{ext}}] \\ &+ \sum_t \boldsymbol{\mu}(t+1) \cdot [\boldsymbol{\psi}(t+1) + \mathbf{h}^s(t)] - i \sum_k \hat{h}_k^s(t) \sum_i \mu_i(t) J_{ki} - \sum_{it} \text{lc}(h_i^s(t)). \end{aligned} \quad (\text{A.4})$$

The second sum in the second line now cancels the second sum in the first line because of the saddle point equation (252), and we end up with the expression

$$\ln Z_s[\psi] = \sum_{i,t} H_2[\mu_i(t)] + \sum_t \boldsymbol{\mu}(t+1) \cdot [\boldsymbol{\psi}(t+1) + \mathbf{h}^s(t)] - \sum_{it} \text{lc}(h_i^s(t)). \quad (\text{A.5})$$

The Legendre transform of this expression with respect to the magnetizations  $m_i(t)$  is

$$\Gamma_s[m] = \ln Z_s[\psi] - \sum_t \boldsymbol{\psi}(t) \cdot \mathbf{m}(t), \quad (\text{A.6})$$

where the  $\psi$  are to be expressed in terms of the  $m$  by solving

$$m_i(t) = \frac{\partial \ln Z_s[\psi]}{\partial \psi_i(t)}. \quad (\text{A.7})$$

We already know that  $\partial \ln Z_s[\psi] / \partial \psi_i(t) = m_i^s(t) = \mu_i(t)$ , so at this saddle point level one has  $m_i(t) = \mu_i(t)$ . The second saddle point equation (252) then becomes  $\mathbf{h}^s(t) = \mathbf{h}^{\text{ext}} + \mathbf{J}\mathbf{m}(t)$  and inserting this gives

$$\Gamma_s[m] = \sum_{it} H_2[m_i(t)] + \sum_{it} m_i(t+1) \left[ h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right] - \sum_{it} \text{lc} \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right). \quad (\text{A.8})$$

The equation of state is, by standard Legendre transform properties,

$$-\psi_i(t) = \frac{\partial \Gamma_s[m]}{\partial m_i(t)}, \quad (\text{A.9})$$

which yields

$$\begin{aligned} -\psi_i(t) &= -\tanh^{-1}(m_i(t)) + h_i^{\text{ext}} + \sum_j J_{ij} m_j(t-1) \\ &+ \sum_j J_{ji} m_j(t+1) - \sum_j J_{ji} \tanh(h_j^s(t)). \end{aligned} \quad (\text{A.10})$$

At  $\psi_i(t) = 0$  this is naturally solved by  $m_i(t+1) = \tanh(h_i^{\text{ext}} + \sum_j J_{ij} m_j(t))$ . So far we have not achieved anything new: we have merely provided an alternative way of obtaining the naive mean-field equations (253) and (254) we already had.

Now let us consider the Legendre transform of the generating functional including the Gaussian correction,  $\ln Z[\psi] = \ln Z_s[\psi] + A[\psi]$

$$\Gamma[m] = \ln Z_s[\psi] + A[\psi] - \sum_t \boldsymbol{\psi}(t) \cdot \mathbf{m}(t), \quad (\text{A.11})$$

where as before  $\psi$  is to be treated as a function of the magnetizations, as determined by the condition

$$m_i(t) = \frac{\partial \ln Z[\psi]}{\partial \psi_i(t)} = \mu_i(t) + A_i(t). \quad (\text{A.12})$$

Here we have used the earlier definition  $A_i(t) = \partial A / \partial \psi_i(t)$ .

Expressing the right-hand side of (A.11) in terms of  $m$ , and keeping only terms linear  $A_i$ , we obtain, as shown below,

$$\begin{aligned} \Gamma[m] = & \sum_{it} H_2[m_i(t)] + \sum_{it} m_i(t+1) \left[ h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right] \\ & - \sum_{it} \text{lc} \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right) + A[m]. \end{aligned} \quad (\text{A.13})$$

Remarkably, all the  $A_i$  terms have cancelled here and the only difference to the saddle point result  $\Gamma_s[m]$  in (A.8) is the naive addition of the correction term  $A$ , expressed in terms of  $m$ .

From  $\Gamma[m]$  we can now derive the equation of state for the magnetizations from

$$-\psi_i(t) = \frac{\partial \Gamma[m]}{\partial m_i(t)} = -\tanh^{-1}(m_i(t)) + h_i^{\text{ext}} + \sum_j J_{ij} m_j(t-1) + \frac{\partial A}{\partial m_i(t)} \quad (\text{A.14})$$

$$+ \sum_k J_{ki} \left[ m_k(t+1) - \tanh \left( h_k^{\text{ext}} + \sum_j J_{kj} m_j(t) \right) \right]. \quad (\text{A.15})$$

If we are only interested in corrections to *quadratic* order in  $J$  for the field acting on each spin, we can drop the terms in the second line of the equation above as the term in square brackets is already  $\mathcal{O}(J^2)$ . In the physical limit  $\psi \rightarrow 0$  one then obtains

$$m_i(t) = \tanh \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j(t-1) + \frac{\partial A}{\partial m_i(t)} \right). \quad (\text{A.16})$$

We can now appreciate the role played by the Legendre transform: the corrections we obtain act on the effective fields and so are inside the tanh. As explained in the main text this makes more physical sense.

To evaluate  $\partial A / \partial m_i(t)$ , one can start from the expression (262) for  $A$ . Replacing  $h_i^s(t)$  in this using the saddle point equation (252) gives

$$A = -\frac{1}{2} \sum_{it} \left[ \mu_i^2(t+1) - \tanh^2 \left( h_i^{\text{ext}} + \sum_j J_{ij} \mu_j(t) \right) \right] \sum_k J_{ik}^2 [1 - \mu_k^2(t)]. \quad (\text{A.17})$$

As  $A$  is already  $\mathcal{O}(J^2)$ , replacing all  $\mu_i(t)$  by  $m_i(t)$  in this expression only gives a negligible correction of  $\mathcal{O}(J^4)$ . Also the first square bracket is small at the physical saddle point, of  $\mathcal{O}(J^2)$ , so the derivative of the final factor with respect to  $m_i(t)$  can be dropped. Finally the derivative of the tanh can also be neglected as it is of order  $J$ . One thus finds to leading order the simple result

$$\frac{\partial A}{\partial m_i} = -m_i(t) \sum_k J_{ik}^2 (1 - m_k^2(t-1)) + \mathcal{O}(J^3), \quad (\text{A.18})$$

Combined with (A.16) this yields the dynamical TAP equation (266).

It remains to show (A.13). Using the expression (A.5) for  $\ln Z_s$  and  $m_i(t) = \mu_i(t) + A_i(t)$  in (A.11), we can write  $\Gamma$  as

$$\Gamma[\mathbf{m}] = \sum_{i,t} H_2[\mu_i(t)] + \sum_t \boldsymbol{\mu}(t) \cdot \mathbf{h}^s(t-1) - \sum_{it} \text{lc}(h_i^s(t)) - \sum_i \psi_i(t) A_i(t) + A. \quad (\text{A.19})$$

We have shifted the time index  $t$  by one in the second and fourth sum for later convenience. Now we express the right-hand side of (A.19) in terms of  $m$  and keep only terms linear in  $A_i$ .

For the last sum we need the following identity, which can be obtained from the definition of  $\mu_i(t)$  in (250) together with  $m_i(t) = \mu_i(t) + A_i(t)$ :

$$\psi_i(t) = \tanh^{-1}(m_i(t) - A_i(t)) - h_i^{\text{ext}} - \sum_j J_{ij} m_j(t-1) + \sum_j J_{ij} A_j(t-1) + i(\mathbf{J}^T \hat{\mathbf{h}}^s(t))_i. \quad (\text{A.20})$$

To linear order in  $A_i$ , the various terms on the right-hand side of (A.19) are then

$$H_2[\mu_i(t)] = H_2[m_i(t)] + A_i(t) \tanh^{-1}(m_i(t)), \quad (\text{A.21})$$

$$\begin{aligned} \mu_i(t) h_i^s(t-1) &= m_i(t) \left[ h_i^{\text{ext}} + \sum_j J_{ij} m_j(t-1) \right] - A_i(t) \sum_j J_{ij} m_j(t-1) \\ &\quad - m_i(t) \sum_j J_{ij} A_j(t-1) - h_i^{\text{ext}} A_i(t), \end{aligned} \quad (\text{A.22})$$

$$-\text{lc}(h_i^s(t)) = -\text{lc} \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right) + \tanh \left( h_i^{\text{ext}} + \sum_k J_{ik} m_k(t) \right) \left[ \sum_j J_{ij} A_j(t) \right], \quad (\text{A.23})$$

$$\begin{aligned} -\psi_i(t) A_i(t) &= -A_i(t) \tanh^{-1}(m_i(t)) + A_i(t) \sum_j J_{ij} m_j(t-1) - i A_i(t) \sum_j J_{ji} \hat{h}_j^s(t) \\ &\quad + h_i^{\text{ext}} A_i(t). \end{aligned} \quad (\text{A.24})$$

Putting all these together we notice that the second term on the right-hand side of the first equation above and the first term of the last equation, the second terms in the second and last equations, as well as the last terms in the second and last equations cancel each other, yielding

$$\begin{aligned} \Gamma[m] &= \sum_{it} H_2[m_i(t)] + \sum_{it} m_i(t) \left[ h_i^{\text{ext}} + \sum_j J_{ij} m_j(t-1) \right] - \sum_{it} \text{lc} \left( h_i^{\text{ext}} + \sum_j J_{ij} m_j(t) \right) \\ &\quad - \sum_{ijt} J_{ij} A_j(t-1) m_i(t) + \sum_{it} \tanh \left( h_i^{\text{ext}} + \sum_k J_{ik} m_k(t) \right) \left[ \sum_j J_{ij} A_j(t) \right] \\ &\quad - i \sum_{ijt} A_j(t) J_{ij} \hat{h}_i^s(t) + A + \mathcal{O}(A_i^2). \end{aligned} \quad (\text{A.25})$$

Using the fact that from the first saddle point equation (251)

$$-i \hat{h}_i^s(t) = \mu_i(t+1) - \tanh(h_i^s(t)) = m_i(t+1) - \tanh(h_i^s(t)) - A_i(t+1), \quad (\text{A.26})$$

we can write the terms in the second line of (A.25), together with the first term in the third line, as

$$\sum_{ijt} \left\{ -\tanh(h_i^s(t)) + \tanh \left( h_i^{\text{ext}} + \sum_k J_{ik} m_k(t) \right) \right\} J_{ij} A_j(t) + \mathcal{O}(A_i^2). \quad (\text{A.27})$$

As the term in curly braces is itself a correction term that is non-zero only because of the difference between  $\mu_i(t)$  and  $m_i(t)$ , this overall expression can be neglected as subleading. The remaining terms of (A.25) then give exactly (A.13) as claimed. Note that keeping only terms

linear in  $A_i$  in this calculation—which then eventually cancel—in some sense plays the same role as expanding to second order in  $J^2$  in the Plefka method, as  $A_i$  is of  $\mathcal{O}(J^2)$ .

## References

- [1] Langevin P 1908 *CR Acad. Sci. Paris* **146** 530
- [2] Ma S K 1976 *Modern Theory Of Critical Phenomena* (Boulder, CO: Westview Press)
- [3] De Dominicis C 1978 *Phys. Rev. B* **18** 4913
- [4] Sompolinsky H and Zippelius A 1981 *Phys. Rev. Lett.* **47** 359
- [5] Martin P C, Siggia E and Rose H 1973 *Phys. Rev. A* **8** 423
- [6] De Dominicis C 1976 *J. Phys. Colloq.* **37** C1-247–C1-253
- [7] Janssen H K 1976 *Z. Phys. B* **23** 377–80
- [8] De Dominicis C and Peliti L 1978 *Phys. Rev. B* **18** 353
- [9] Jensen R V 1981 *J. Stat. Phys.* **25** 183–210
- [10] Coolen A and Sherrington D 1993 *Phys. Rev. Lett.* **71** 3886
- [11] Biroli G 1999 *J. Phys. A: Math. Gen.* **32** 8365
- [12] Roudi Y and Hertz J 2011 *J. Stat. Mech.* P03031
- [13] Roudi Y 2013 Lectures on disordered systems dynamics, netadis I summer school on statistical physics approaches to networks across disciplines [http://videlectures.net/netadis2013\\_roudi\\_disordered\\_dynamics](http://videlectures.net/netadis2013_roudi_disordered_dynamics)
- [14] Fischer K H and Hertz J A 1993 *Spin Glasses* (Cambridge: Cambridge University Press)
- [15] Zinn-Justin J 2002 *Quantum Field Theory and Critical Phenomena* (Oxford: Oxford University Press)
- [16] De Dominicis C and Giardinà I 2006 *Random Fields and Spin Glasses: A Field Theory Approach* (Cambridge: Cambridge University Press)
- [17] Chow C C and Buice M A 2015 *J. Math. Neurosci.* **5** 8
- [18] Gardiner C W *et al* 1985 *Handbook of Stochastic Methods* vol 3 (Berlin: Springer)
- [19] Lancaster T and Blundell S J 2014 *Quantum Field Theory For The Gifted Amateur* (Oxford: Oxford University Press)
- [20] Bouchaud J P, Cugliandolo L, Kurchan J and Mézard M 1996 *Physica A* **226** 243–73
- [21] Feigel'man M and Tselik A 1982 *J. Exp. Theor. Phys.* **56** 823
- [22] Feigel'man M and Tselik A 1983 *Phys. Lett. A* **95** 469–73
- [23] Parisi G and Sourlas N 1979 *Phys. Rev. Lett.* **43** 744
- [24] Kurchan J 1992 *J. Phys. I* **2** 1333–52
- [25] Plefka T 1982 *J. Phys. A: Math. Gen.* **15** 1971
- [26] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
- [27] Feigel'man M and Tselik A 1979 *Zh. Eksp. Teor. Fiz* **77** 2524–38
- [28] Hertz J and Klemm R 1979 *Phys. Rev. B* **20** 316
- [29] Verbaarschot J, Weidenmüller H and Zirnbauer M 1984 *Ann. Phys., NY* **153** 367–88
- [30] Sompolinsky H and Zippelius A 1982 *Phys. Rev. B* **25** 6860
- [31] Crisanti A and Sompolinsky H 1987 *Phys. Rev. A* **36** 4922
- [32] Peretto P 1984 *Biol. Cybern.* **50** 51–62
- [33] Glauber R J 1963 *J. Math. Phys.* **4** 294–307
- [34] Coolen A 2001 *Handbook of Biological Physics* vol 4 (Amsterdam: Elsevier) pp 531–96
- [35] Coolen A C C, Kühn R and Sollich P 2005 *Theory of Neural Information Processing Systems* (Oxford: Oxford University Press)
- [36] Roudi Y and Hertz J 2011 *Phys. Rev. Lett.* **106** 048702
- [37] Mézard M and Sakellariou J 2011 *J. Stat. Mech.* L07001
- [38] Mahmoudi H and Saad D 2014 *J. Stat. Mech.* P07001
- [39] Coolen A 2000 arXiv:cond-mat/0006011
- [40] Kappen H and Spanjers J 2000 *Phys. Rev. E* **61** 5658
- [41] Thouless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593–601
- [42] Negele J W and Orland H 1988 *Quantum Many-Particle Systems* (Boulder, CO: Westview)
- [43] Kholodenko A 1990 *J. Stat. Phys.* **58** 355–70
- [44] Bachschmid-Romano L, Battistin C, Oppen M and Roudi Y 2016 *J. Phys. A: Math. Theor.* **49** 434003
- [45] Aurell E and Mahmoudi H 2012 *Phys. Rev. E* **85** 031119
- [46] Roudi Y, Dunn B and Hertz J 2015 *Curr. Opin. Neurobiol.* **32** 38–44

- [47] Dunn B and Roudi Y 2013 *Phys. Rev. E* **87** 022127
- [48] Bravi B, Sollich P and Opper M 2016 *J. Phys. A: Math. Theor.* **49** 194003
- [49] Bravi B and Sollich P 2016 arXiv:[1603.05538](https://arxiv.org/abs/1603.05538)
- [50] Cooper F and Dawson J F 2016 *Ann. Phys., NY* **365** 118–54
- [51] Poletini M 2015 *J. Phys. A: Math. Theor.* **48** 365005
- [52] Nemoto T, Hidalgo E G and Lecomte V 2016 arXiv:[1607.04752](https://arxiv.org/abs/1607.04752)