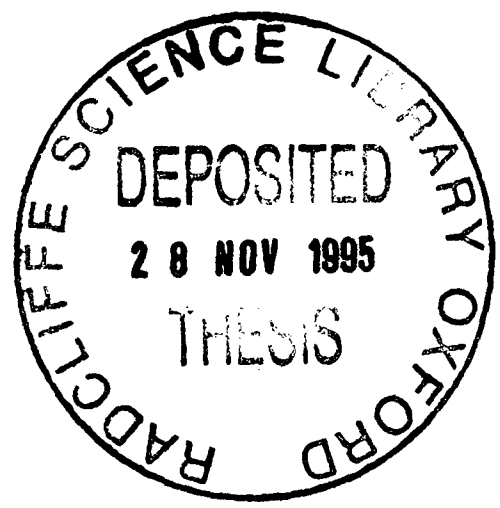
Dynamics of Neural Networks and Disordered Spin Systems

Stephen Nicholas Laughton
Brasenose College

Department of Physics - Theoretical Physics
University of Oxford



Thesis submitted for the degree of Doctor of Philosophy
in the University of Oxford

Trinity term 1995

Dynamics of Neural Networks and Disordered Spin Systems

Stephen Nicholas Laughton
Brasenose College

Department of Physics
University of Oxford

Abstract

I obtain a number of results for the dynamics of several disordered spin systems, of successively greater complexity. I commence with the generalised Hopfield model trained with an intensive number of patterns, where in the thermodynamic limit macroscopic, deterministic equations of motion can be derived exactly for both the synchronous discrete time and asynchronous continuous time dynamics. I show that for symmetric embedding matrices Lyapunov functions exist at the macroscopic level of description in terms of pattern overlaps. I then show that for asymmetric embedding matrices several types of bifurcation phenomena to complex non-transient dynamics occur, even in this simplest model. Extending a recent result of Coolen and Sherrington, I show how the dynamics of the generalised Hopfield model trained with extensively many patterns and non-trivial embedding matrix can be described by the evolution of a small number of overlaps and the disordered contribution to the 'energy', upon calculation of a noise distribution by the replica method. The evaluation of the noise distribution requires two key assumptions: that the flow equations are self averaging, and that equipartitioning of probability occurs within the macroscopic sub-shells of the ensemble. This method is inexact on intermediate time scales, due to the microscopic information integrated out in order to derive a closed set of equations. I then show how this theory can be improved in a systematic manner by introducing an order parameter function - the joint distribution of spins and local alignment fields, which evolves in time deterministically, according to a driven diffusion type equation. I show how the coefficients in this equation can be evaluated for the generalised Sherrington-Kirkpatrick model, both within the replica symmetric ansatz, and using Parisi's ultrametric ansatz for the replica matrices, upon making once again the two key assumptions (self averaging and equipartitioning). Since the order parameter is now a continuous function, however, the assumption of equipartitioning within the macroscopic sub-shells is much less restricting.

Thesis submitted for the degree of Doctor of Philosophy
in the University of Oxford

Trinity term 1995

Acknowledgment

I would like to thank both my supervisors for their help and encouragement: Ton Coolen - without whom none of this work would have been carried out, and David Sherrington. Thanks also to all the occupants of the various offices I have inhabited during my time here - Magnus, Andreas, Franck, Nigel, Richard, William, Rashmi and Harvey, for putting up with me; and of course to everyone in Theoretical Physics. Special thanks to all the computers for bearing the strain of my loose code, and to everyone who ran their jobs at 'niceness' 0.

Thanks to the EPSRC for financial support under grant number 92310910, to Brasenose college for spending it, and of course the humble tax payer for funding my overlong education - you will be repayed many times over.

Thanks to the Howard street running team for keeping me insane, and for endless hours of fun with Luke and Brian.

Finally thanks to my parents for starting it all.

Gestern liebt' ich,
Heute leid' ich,
Morgen sterb' ich;
Dennoch denk' ich
Heut und morgen
Gern an gestern.

-G. E. Lessing

Preface

This Thesis is split into three parts. Each deals with a specific way of choosing the bonds in a disordered spin system. Before proceeding with the original work which I present in application for my D. Phil., I present an overview of the methods used in all three parts. The brief introduction is included in order that physicists familiar only with the equilibrium statistical mechanics of ordered systems, and those not acquainted with neural network theory can better understand the following work without reference to a large amount of background material. At the end of the introduction the maps and flow equations for the evolution of a general set of macroscopic order parameters $\boldsymbol{\Omega}$ are derived. These will form the starting point for the majority of the following work.

The work contained in this thesis was carried out between Michelmas term 1992 and Trinity term 1995, at the sub-department of Theoretical Physics, Oxford University, and was first published in the following papers, co-authored by Ton Coolen and David Sherrington:

1. Lyapunov Functions for Separable Stochastic Neural Networks with Detailed Balance [1]
2. Quasi-Periodicity and Bifurcation Phenomena in Ising Spin Neural Networks with Asymmetric Interaction [2]
3. Order-Parameter Flow in Symmetric and Non-Symmetric Fully Connected Attractor Neural Networks near Saturation [3]
4. Order-Parameter Flow in the SK Spin Glass II: Inclusion of Microscopic Memory Effects [4]
5. Local Field Distribution Evolution in Spin-Glasses, *in preparation*, [5]

The derivation of the equilibrium properties of the generalised Hopfield model near saturation presented in Chapter 4 was obtained in Hilary term 1994 independently of, and prior to the publication of the same results by Cugliandolo and Tsodyks in [6].

Health warning

Many of the topics addressed in this thesis are of a highly mathematical nature. Many of the equations are highly complicated, for which I make no apology. In this field often only the starting point, and interpretation of results require physical intuition, the bulk of the calculation is a complicated mathematical exercise. The result is a large number of equations, which at first sight cannot easily be understood. Many mathematical steps are included in order that the calculations can be reconstructed by others. I hope however that the complicated nature of the results does not distract the reader from their importance.

Even Jehovah, after Moses had the
commandments committed to stone;
probably thought: "I always forget
the things I really wanted to say."

-Christopher Morley

Contents

List of Figures

Chapter 1

Introduction

Statistical mechanics is concerned with drawing together our understanding of how the microscopic properties of atoms and molecules combine to give the macroscopic bulk properties we are familiar with. Whilst in the past most systems studied obeyed detailed balance (in equilibrium the forward process balancing the reverse for all pairs of states), as most physical systems do, recently interesting cases of systems which do not obeying detailed balance have come to light (for a number of general examples see [7, 8]). The lack of detailed balance means that the conventional methods of equilibrium statistical mechanics, where all thermodynamic quantities can be derived from a generating function - the free energy, cannot be used. We are forced to study the dynamics directly.

Much of early statistical mechanics also assumed that interactions between microscopic elements were uniformly ordered. Although this makes the task of deriving relations technically easier (because of the high degree of symmetry), it is in contrast to our knowledge of real materials - impurities and structural defects cause disorder. The last twenty years has seen an explosion in the work done in trying to deal with this disorder.

Recently methods of statistical mechanics have begun to be applied to non-physical stochastic systems. Greatly simplified models of biological, ecological, and even socio-economical systems have been analysed in an attempt to derive some understanding of the generic behaviour of a class of problems. One system that has been studied in this way is the animal nervous system - neural networks. Here a large number of simple neurons are interconnected by many synapses. The way in which these synapses transmit the electrical signals received from the neurons to which they are connected determines how the network behaves.

These three topics will be briefly explained in the introduction, before a derivation of the evolution equation for a general set of macroscopic order parameters. This will form the basis of most of the work to follow. The introduction concludes with a brief discussion of each of the models studied in this thesis.

1.1 Microscopic Stochastic Processes

In conventional equilibrium statistical mechanics a rigorous definition of the microscopic dynamics is not essential. It suffices to assume that the system is in thermal equilibrium with a heat bath, and that thermal agitations cause changes in the microscopic state, so that the probability of

finding a specific state is proportional to $e^{-\beta\mathcal{H}(\boldsymbol{\sigma})}$, where β is the inverse temperature $1/T$, and $\mathcal{H}$ - the Hamiltonian - is a function of the microscopic state $\boldsymbol{\sigma}$. For a general introduction see text books such as [9]. For systems without detailed balance, a more thorough description of the microscopic dynamics is necessary if meaningful macroscopic equations are to be derived. This section is intended as a general introduction to stochastic processes for those familiar only with the more frequently encountered equilibrium results. For more rigorous treatments, and further background see text books such as [10] or [11] and references therein.

We consider a system of N Ising spins [12], $\sigma_i \in \{-1, 1\}$, which are coupled by (quenched) spin-spin (exchange) interactions J_{ij} (about which we currently make no assumptions of symmetry). Each spin experiences a local magnetic field $h_i = \sum_{j\neq i} J_{ij}\sigma_j + \theta_i$, due to the cumulative effect of all the other spins to which it is connected (i.e. those for which $J_{ij} \neq 0$). θ_i is an applied, external field which for simplicity we usually take equal to zero. In general the local field will be different for each spin, unless $\{\sigma_i\}$ and $\{J_{ij}\}$ are highly correlated.

We assume that each of the spins evolves in time according to a stochastic alignment to the local field [13]. Suppose that we know the state of the i^{th} spin at time t; in order that the process is Markovian, i.e. dependent only on the instantaneous state of the system, we define our elementary unit of time such that the following general equation describing the the state of the spin at a later is

$$\sigma_i(t+\delta) = \operatorname{sgn}\left[f(\beta, h_i(\boldsymbol{\sigma}(t))) + \eta_i(t)\right]. \tag{1.1}$$

The dynamics of the system is now determined by the relationship between the function f, and the distribution of the uncorrelated random numbers $\eta_i(t)$. β measures the relative strength of the deterministic and stochastic terms.

Clearly the stochasticity introduced through the η's means that we cannot know the exact state of the system at an arbitrary time. Instead we define a microscopic probability function $p_t(\boldsymbol{\sigma})$ - the probability of finding the microstate $\boldsymbol{\sigma}$ at time t, where $\boldsymbol{\sigma}$ is the vector $(\sigma_1, \ldots, \sigma_N)$. We now ask: what is the probability that $\sigma_i(t+\delta) = \sigma$ given the state $\boldsymbol{\sigma}'(t)$ at time t, written $p_{t+\delta}\left(\sigma_i = \sigma \,|\boldsymbol{\sigma}'(t)\right)$? Clearly this is given by

$$p_{t+\delta}\left(\sigma_i = \sigma \,|\boldsymbol{\sigma}'(t)\right) = \sigma \int_{-f(\beta,h_i)}^{\sigma\infty} d\eta_i\, \rho(\eta_i) \operatorname{sgn}\left[f(\beta, h_i(\boldsymbol{\sigma}'(t))) + \eta_i(t)\right],$$

where $\rho(\eta)$ is the measure of the distribution of η.

So far we have only considered the behaviour of a single spin. Suppose we now want to know the probability of finding a state $\boldsymbol{\sigma}$ of the whole sytem at time $t+\delta$ given state $\boldsymbol{\sigma}'$ at time t. This will obviously depend on the updating mechanism, i.e. how many spins have the opportunity to flip in time δ. If we divide the systems into clusters s_α spins which are updated simultaneously, then since the noise is uncorrelated amongst spins (as well as in time), after updating one cluster we have

$$p_{t+\delta}\left(\boldsymbol{\sigma} \,|\boldsymbol{\sigma}'(t)\right) = \prod_{i\in s_\alpha} \sigma_i \int_{-f(\beta,h_i)}^{\sigma_i\infty} d\eta_i\, \rho(\eta_i) \operatorname{sgn}\left[f(\beta, h_i(\boldsymbol{\sigma}'(t))) + \eta_i(t)\right].$$

We assume that the cluster to be updated is chosen at random, and therefore average over the clusters, and include a factor $\prod_{j\notin s_\alpha} \delta_{\sigma_j,\sigma_j'}$ to impose the fact that all spins in the other clusters remain unchanged, hence

$$p_{t+\delta}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma}'} p_t(\boldsymbol{\sigma}') \sum_\alpha \frac{|s_\alpha|}{N} \prod_{j\notin s_\alpha} \delta_{\sigma_j,\sigma_j'} \prod_{i\in s_\alpha} \sigma_i \int_{-f(\beta,h_i)}^{\sigma_i\infty} d\eta_i\, \rho(\eta_i) \operatorname{sgn}\left[f(\beta, h_i(\boldsymbol{\sigma}')) + \eta_i(t)\right].$$

The microscopic probability distribution therefore evolves according to a Markov chain

$$p_{t+\delta}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma}'} W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma}) p_t(\boldsymbol{\sigma}'), \tag{1.2}$$

where

$$W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma}) = \sum_{\alpha} \frac{s_\alpha}{N} \prod_{j \notin s_\alpha} \delta_{\sigma_j, \sigma_j'} \prod_{i \in s_\alpha} \sigma_i \int_{-f(\beta, h_i(\boldsymbol{\sigma}'))}^{\sigma_i \infty} d\eta_i \, \rho(\eta_i) \operatorname{sgn}\left[f(\beta, h_i(\boldsymbol{\sigma}')) + \eta_i\right]$$

is the transition probability for the process $\boldsymbol{\sigma}' \to \boldsymbol{\sigma}$ in time δ.

A convenient choice which is often made (for reasons that will become apparent later) is

$$W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma}) = \sum_{\alpha} \frac{s_\alpha}{N} \prod_{j \notin s_\alpha} \delta_{\sigma_j, \sigma_j'} \prod_{i \in s_\alpha} \frac{1}{2} \left[1 + \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}'))\right]. \tag{1.3}$$

So with η_i evenly distributed on the interval $[-1, +1]$, $f(\beta, h_i(\boldsymbol{\sigma})) = \tanh(\beta h_i(\boldsymbol{\sigma}))$, often used to mimic the non-linear transfer function in neural networks; or equivalently with $f(\beta, h_i(\boldsymbol{\sigma})) = \beta h_i(\sigma)$, $\rho(\eta_i) = \frac{\beta}{2}\left(1 - \tanh^2(\beta \eta_i)\right)$, which is qualitatively similar to a Gaussian distribution.

So far we have made no assumptions about the time step δ. If we define 1 time unit as the mean time required to update each spin, then we can imagine two extremes of updating mechanisms:

1. All spins are updated simultaneously. This will be known as parallel updating or synchronous dynamics. In this case all spins are in the same cluster and $\delta = 1$. The corresponding Markov equation is

$$p_{t+1}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma}'} \prod_{i=1}^{N} \frac{1}{2} \left[1 + \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}'))\right] p_t(\boldsymbol{\sigma}'). \tag{1.4}$$

2. Each spin is updated in a preordained, or random sequence. This will be known as sequential updating or asynchronous dynamics. In this case each spin is in a different cluster, so the corresponding Markov equation is

$$p_{t+\delta}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma}'} \frac{1}{N} \sum_{i=1}^{N} \prod_{j \neq i} \delta_{\sigma_j, \sigma_j'} \frac{1}{2} \left[1 + \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}'))\right] p_t(\boldsymbol{\sigma}').$$

Since in this case $\delta = \frac{1}{N}$ and we will usually be concerned with the thermodynamic limit ($N \to \infty$) it is more convenient to define the dynamics via a differential equation (the master equation):

$$\frac{d}{dt} p_t(\boldsymbol{\sigma}) = \sum_{i} \left\{w_i(F_i \boldsymbol{\sigma}) p_t(F_i \boldsymbol{\sigma}) - w_i(\boldsymbol{\sigma}) p_t(\boldsymbol{\sigma})\right\}. \tag{1.5}$$

Here F_i is the spin flip operator: $F_i \Phi(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N) = \Phi(\sigma_1, \ldots, -\sigma_i, \ldots, \sigma_N)$ (we are considering only single spin flips); and $w_i(F_i \boldsymbol{\sigma})$ (which follows directly from the appropriate version of (1.3)) is the transition rate for the process $\boldsymbol{\sigma} \to F_i \boldsymbol{\sigma}$:

$$w_i(\boldsymbol{\sigma}) = \frac{1}{2} \left[1 - \sigma_i \tanh \beta h_i(\boldsymbol{\sigma})\right]. \tag{1.6}$$

The reason for the specific choice for (1.3) now becomes clear, since for symmetric interactions we recover the Gibbs distribution in detailed balance equilibrium:

$$\frac{d}{dt} p_t = 0 \Rightarrow \frac{w_i(\boldsymbol{\sigma})}{w_i(F_i \boldsymbol{\sigma})} = \frac{p_t(F_i \boldsymbol{\sigma})}{p_t(\boldsymbol{\sigma})} = e^{-\beta(\mathcal{H}(F_i \boldsymbol{\sigma}) - \mathcal{H}(\boldsymbol{\sigma}))} \ \forall \ i.$$

with the standard Ising Hamiltonian $\mathcal{H}(\boldsymbol{\sigma}) = -\sum_{i<j} \sigma_i J_{ij} \sigma_j = -\frac{1}{2} \sum_i \sigma_i h_i(\boldsymbol{\sigma})$.

Case 2 is closer to physical reality where changes in spin state may be caused by some random collision within the material (though it takes no account of double spin flips); whereas case 1 might correspond to a neural network where the states of all the neurons are updated simultaneously, on application of a timing signal.

1.2 Disordered Systems and the Replica Method

The main problem encountered when studying disordered systems (i.e. when the interactions are not uniform but drawn randomly from some distribution), as opposed to simple ordered systems, is how to compare samples with different realisations of the disorder. Results which depend on the exact realisation of the microscopic disorder are useless. Any theory must describe any (sufficiently large) sample, independent of the exact realisation of the disorder. A concept central to disordered systems theory is that of *self averaging*. A self averaging quantity is one which, in the thermodynamic limit, is independent of the exact realisation of the disorder. Physically this means that the probability of finding a particular value of the desired quantity, as the realisation of the disorder is varied, is strongly peaked around a typical value. This allows us to average over the disorder at an early stage - which is technically important in deriving useful equations. The question remains of which quantities are self averaging.

In most physical models of magnetism the disorder is *site disorder*, i.e. it is caused by the magnetic atoms or impurities being placed at random in a substrate. The site disorder causes *bond disorder*, i.e. because the impurities are placed at random, the distance from their neighbours, and hence the strength of their interaction with them is random. The task of constructing a model is now one of choosing a distribution for the interactions J_{ij}.

Equilibrium statistical mechanics tells us that in detailed balance equilibrium many properties of the system can be calculated from the partition function $Z = \sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})}$, or equivalently the free energy per spin $f = -\frac{1}{\beta N} \ln[Z]$. It is not however immediately obvious how to apply the disorder average, what effect the average will have, or how to interpret the averaged quantities. There appear to be two candidates for averaging:

1. Averaging the partition function Z is equivalent to the annealed approximation where the interactions are treated as fast stochastic variables on the same footing as the spins

$$\langle Z \rangle_{J_{ij}} = \int dJ_{ij}\, P(J_{ij}) \sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma}, J_{ij})}.$$

2. Averaging the free energy, on the other hand is equivalent to the quenched approximation, where the disorder is static within the system

$$\langle f \rangle_{J_{ij}} = -\frac{1}{\beta N} \int dJ_{ij}\, P(J_{ij}) \ln \left[\sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma}, J_{ij})} \right].$$

Clearly the quenched approximation is most relevant when considering a solid with frozen, or slowly varying random magnetic disorder. We also require any quantities of interest to be

independent of the precise realization of the disorder, i.e. to be self averaging. Consider the standard Ising Hamiltonian $\mathcal{H}(\boldsymbol{\sigma}, J_{ij}) = -\frac{1}{2}\sum_{i,j}\sigma_i J_{ij}\sigma_j$. This is an extensive quantity if we require that the free energy per spin is intensive (this fixes the scaling with N of the interactions J_{ij}). Averaging the partition function Z over the disorder involves averaging the exponential of an extensive quantity. In this case the integral may become dominated by regions of $\{J_{ij}\}$ space which are rare, but have a large value of $\mathcal{H}(\boldsymbol{\sigma}, J_{ij})$. Hence Z is not self averaging for infinite range models, where there is an extensive number of disordered variables per spin. In averaging the free energy f we will not encounter the same problem. (**NB** it has never been strictly proved that the free energy is indeed self averaging for infinite range models, however this property is observed in numerical simulations.) Clearly then the free energy per spin f is the relevant quantity; calculating this however introduces the problem of averaging the logarithm of the partition function.

In order to average the logarithm of the partition function Edwards and Anderson proposed the *Replica Trick* [14] which was later applied by Sherrington and Kirkpatrick to an infinite range model (where mean field theory becomes exact) [15], to obtain an expression for f in terms of a saddle point integration. Consider the expansion of Z^n:

$$Z^n = e^{\ln[Z^n]} = e^{n\ln[Z]} = 1 + n\ln[Z] + \mathcal{O}(n^2\ln^2[Z]), \tag{1.7}$$

therefore $\ln[Z] = \lim_{n\to 0}\frac{Z^n-1}{n}$ making it possible to replace the average of a logarithm of Z with the average of a product of Z's. Note that as mentioned earlier, for an intensive free energy per spin $\ln[Z] \sim N$, hence the first discarded term in the expansion is $\mathcal{O}(n^2N^2)$. Therefore to be rigorous we should take $n \to 0$ *before* $N \to 0$, in order to get a well defined expansion. In practice the order of limits is reversed since we need to assume the mean-field approximation before taking $n \to 0$. This causes some very subtle complications which will be discussed later.

The different terms, known as *replicas*, appearing in the product $Z^a, Z^b \ldots Z^n$ are interpreted as systems in different thermal baths, with the same realisations of the disorder. We therefore have to calculate an object like

$$\lim_{n\to 0}\left\langle \sum_{\boldsymbol{\sigma}^1} e^{-\beta\mathcal{H}(\boldsymbol{\sigma}^1)} \sum_{\boldsymbol{\sigma}^2} e^{-\beta\mathcal{H}(\boldsymbol{\sigma}^2)} \ldots \sum_{\boldsymbol{\sigma}^n} e^{-\beta\mathcal{H}(\boldsymbol{\sigma}^n)} \right\rangle_{J_{ij}}.$$

In order to evaluate this we reverse the order of averaging, so that the disorder average is carried out before the spin average. The disorder average has the effect of coupling spins from different replicas, so that as well as the magnetisation $m^a = \frac{1}{N}\sum_i \sigma_i^a$, a relevant order parameter is found to be the spin-glass order parameter $q_{ab} = \frac{1}{N}\sum_i \sigma_i^a\sigma_i^b$. After carrying out the spin average, Z^n takes the form of a saddle point integration:

$$\int dm^a\, d\hat{m}^a\, dq_{ab}\, d\hat{q}_{ab}\; e^{-\beta N\Psi(\{m,\hat{m},q,\hat{q}\})},$$

where the integrals arise from the delta function representations of unity used to introduce the order parameters, and $\hat{m}^a$ and $\hat{q}_{ab}$, the conjugate parameters to m^a and q_{ab}, are used to impose those delta functions. In the limit $N \to \infty$ (with $n > 1$) this integral will be dominated by the values of $\{m, \hat{m}, q, \hat{q}\}$ that minimise Ψ. This is known as the saddle point method. Expanding about this minimum ($\tilde{\Psi}$), the integral is

$$\int d\mathcal{Q}\; e^{-\beta N\tilde{\Psi}(\{\mathcal{Q}\}) - \frac{1}{2}\beta N\mathcal{Q}\cdot\frac{\partial^2\Psi}{\partial\mathcal{Q}\partial\mathcal{Q}}\mathcal{Q}} \propto \frac{e^{-\beta N\tilde{\Psi}(\{\mathcal{Q}\})}}{\sqrt{\det\left(\frac{\partial^2\Psi}{\partial\mathcal{Q}\partial\mathcal{Q}}\right)}},$$

where $\{\mathcal{Q}\}$ represents a general set of order parameters and their conjugates, and $\frac{\partial^2\Psi}{\partial\mathcal{Q}\partial\mathcal{Q}}$ means the matrix $\frac{\partial^2\Psi}{\partial\mathcal{Q}^\mu\partial\mathcal{Q}^\nu}$. The first derivative is missing due to the extremalisation condition. From this we see that $\frac{\partial^2\Psi}{\partial\mathcal{Q}\partial\mathcal{Q}}$ must have a non-negative determinant, in order for the integral to be well defined. This is equivalent to requiring that the extremum is indeed a minimum, and not a maximum. Here we have implicitly assumed that (possibly imaginary valued) conjugate parameters have been removed in favour of true (real valued) order parameters. Extending the saddle point method into the complex plane is not trivial, since then the meaning of maximal and minimal in extrema is ill defined.

So far we have implicitly assumed that n is a positive integer. This is required in order to perform the spin average. The identity for $\ln[Z]$ however requires us to take the limit $n \to 0$ however. Clearly we cannot do this explicitly without further assumptions. There are also conceptual problems, i.e. how do we interpret an $n \times n$ matrix in the limit $n \to 0$? Upon making some ansatz for the form of the replica matrices, i.e. Sherrington and Kirkpatrick's *Replica Symmetry* [15] ($m^a = m \ \forall\ a$, $\hat{m}^a = \hat{m} \ \forall\ a$, $q_{ab} = q \ \forall\ a \neq b$ and $\hat{q}_{ab} = \hat{q} \ \forall\ a \neq b$ where the diagonal terms are excluded since $q_{aa} = 1$) or Parisi's *Ultrametric* structure [16, 17, 18, 19], $\tilde{\Psi}$ can be evaluated for a finite value of n, so that Z^n has the form $e^{-n\beta N\tilde{\Psi}+\mathcal{O}(n^2)}$, before taking the limit $n \to 0$, so that

$$f = -\frac{1}{\beta N}\ln[Z] = \frac{1}{\beta N}\lim_{n\to 0}\frac{1}{n}\left(1 - e^{-n\beta N\tilde{\Psi}}\right) = \tilde{\Psi}.$$

The order parameters are found to have physical interpretations $m = \langle\langle\sigma\rangle\rangle_J$ and $q = \langle\langle\sigma\rangle\langle\sigma\rangle\rangle_J$. Where the inner brackets indicate a site (or equivalently thermal) average, and the outer brackets an average over the disorder. A phase diagram of a spin-glass type model shows a paramagnetic region ($m = 0, q = 0$), a ferromagnetic region ($m \neq 0, q \neq 0$) and a spin glass region ($m = 0, q \neq 0$).

The replica method has another use, often encountered when attempting to derive dynamic laws, when we may wish to calculate the average of some quantity over a constrained region of phase space. Although the starting point is very different, the methodology is very similar. Consider the average of a function $\Phi(\boldsymbol{\sigma})$ over a constrained region of phase-space $W(\boldsymbol{\sigma})$

$$\langle\Phi(\boldsymbol{\sigma})\rangle_W \equiv \frac{\sum_{\boldsymbol{\sigma}}\Phi(\boldsymbol{\sigma})W(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}}W(\boldsymbol{\sigma})} = \frac{\sum_{\{\boldsymbol{\sigma}\}}\Phi(\boldsymbol{\sigma}^1)W(\boldsymbol{\sigma}^1)\prod_{a=2}^n W(\boldsymbol{\sigma}^a)}{\sum_{\{\boldsymbol{\sigma}\}}W(\boldsymbol{\sigma}^1)\prod_{a=2}^n W(\boldsymbol{\sigma}^a)} = \lim_{n\to 0}\sum_{\{\boldsymbol{\sigma}\}}\Phi(\boldsymbol{\sigma}^1)\prod_{a=1}^n W(\boldsymbol{\sigma}^a). \tag{1.8}$$

Here replica 1 is singled out as being 'special'. As in the equilibrium case we usually assume self-averaging, and reverse the order of averaging. This method then typically gives something of the form

$$\int d\hat{Q}\ \mathcal{I}(\{Q,\hat{Q}\})\ e^{-N\Psi(\{Q,\hat{Q}\})},$$

where $\mathcal{I}$ is an intensive term, and $\{Q\}$ are a general set of order parameters. As before the integral will be extremally dominated, leaving the intensive term to be evaluated in the saddle point equations defined by Ψ.

1.3 Neural Networks

The subject of neural networks is concerned with the study and simulation of the properties of densely interconnected information processing elements - both as a means of understanding better the properties and behaviour of the nervous system of animals, and as a possible means of developing new artificial computing machines. The operation of the brain is hugely complex, with, in the human brain, of the order of 10^9 neurons (the 'computing' elements) all operating in parallel, interconnected by around 10^4 synapses (the 'wiring') each. In common with the working of most complex systems, the brain has several levels of complexity. At the most basic level, the operation of individual neurons can be studied - their operation is found to be based on the transmission of electro-chemical signals. At a higher level, the interaction of several neurons can be studied - it is found that even a very small number of interconnected neurons can have very complex behaviour. At the highest level we can study the global properties of the brain reacting to inputs from the nervous system. Typical behaviour falls into two categories: associative recall (i.e. retrieval of patterns or sequences), and cognition (i.e. reaction to stimuli); behaviour of a real brain is obviously a mixture of both of these occurring simultaneously. Research into all these levels of activity is being carried out by workers in many different fields - biology, psychology, physics, computer science ..., in an attempt to develop a unified approach involving self-consistent and realistic models of the different levels of operation of the brain.

The study of neural networks in the physics community has largely been centred around the behaviour of the brain at its highest level, using the techniques of statistical mechanics. In this formulation the operation of individual neurons is given a simple form and the behaviour of the network studied for various choices of interactions between neurons. Because of the large number of simple neurons ($N \sim 10^9$), statistical mechanical techniques can be applied effectively to neural networks, and as with many physical models of this sort the exact functional form of the interactions is unimportant to a qualitative understanding of the bulk properties. Although the global behaviour of networks of this kind is robust to alterations in the exact form of microscopic activity, this formulation takes no account of the importance of individual elements in neural computation. We are solely interested in the behaviour that can occur or be induced, and the limits of performance of a typical realisation, rather than the operation of a specific example. This may give the impression of a lack of rigour, however such a complicated problem is intractable without a large number of simplifying assumptions. If insufficient approximations are made, the model will be just as complicated as the original system and no new insight can be gained by studying it.

Much of the study of information processing in neural networks has focussed on long-range stochastic Ising spin models, where there are a number of parallels to be drawn with Ising spin-glass models such as the S-K model. In such models spins represent neurons, interaction strengths between spins represent synaptic efficacies, and local (magnetic) alignment fields play the role of post-synaptic potentials. For a more general introduction into this field refer to recent textbooks like [20, 21, 22] or review papers like [23, 24, 25, 26]. Biological realism forces us to try to abandon symmetry of the neural interactions, and since in the usual type of model interaction symmetry is equivalent to having detailed balance (apart from some trivial exceptions), studying the dynamics directly is often the only route towards analytical results. The dynamics of such systems at a

macroscopic level (with a suitably chosen set of macroscopic order parameters) can be described in terms of a set of coupled non-linear mappings (for parallel microscopic dynamics) or coupled non-linear differential equations (for sequential microscopic dynamics).

1.3.1 Some Neuro-Biology

The central nervous system consists of a huge number of interconnected neurons each acting according to the electrical stimuli of the other neurons. The neuron basically consists of three important parts: the cell body, the axon connecting to other neurons, and the synapses connecting the axons from other neurons to the cell body. The basic operation of the neuron is, if active ('firing'), to emit spikes of electricity ($\sim$ few mV) via its axons to other neurons every few 10's of ms. These spikes of activity travel along the axons to other neurons causing ions to be released or inhibited at the synapses (excitatory or inhibitory respectively). These ions build up at the cell membrane between the synapse and the cell body causing an electrical potential. If the potential reaches a critical value the synapse emits a spike of electricity and the concentration of neuro transmitter, and hence post synaptic potential (PSP) is dissipated. The neuro transmitter will dissipate naturally, independent of the release of a spike, so that the impulses from other neurons must build up a sufficient concentration within a finite amount of time in order to trigger a spike. It is also observed that occasionally a neuron will fire when its PSP is below threshold, therefore we must allow for stochastic effects in our model. The relationship between any neuron's input and output (its transfer function) has a complicated non-linear form.

Hodgkin and Huxley [27] studied the operation of neurons at an immensely detailed level, and were able to derive complicated differential equations governing the currents and potentials present in a working neuron. It is however possible to derive simpler equations based on a few important variables, from which the macroscopic analysis can proceed:

- u_i : post synaptic potential built up at the cell membrane of the i^{th} neuron due to the concentration of the neuro-transmitter.
- v_i : time averaged potential of the spikes of electricity emitted by the i^{th} neuron.
- C_i : the concentration of the neuro-transmitter at the cell membrane of the i^{th} neuron.
- J_{ij} : the synaptic strength governing how much neuro-transmitter is releases at the cell membrane of the i^{th} neuron due to a potential at the synapse of the j^{th} neuron.
- $g(\gamma_i u_i)$: the transfer function governing how the output (i.e. rate of emission of electrical spikes) of the i^{th} neuron is related to its PSP.
- γ_i : the gain of the transfer function of the i^{th} neuron.
- T_i : the threshold that the PSP must exceed before a spike is emitted.
- η_i : a stochastic element introduced to take account of random fluctuations in the neuro transmitter concentration.
- τ_i : decay time of the PSP build up.
- $\tilde{\tau}_i$: decay time of the concentration of the neuro transmitter.
- I_i : currents applied externally to the network, as a means of input for example.

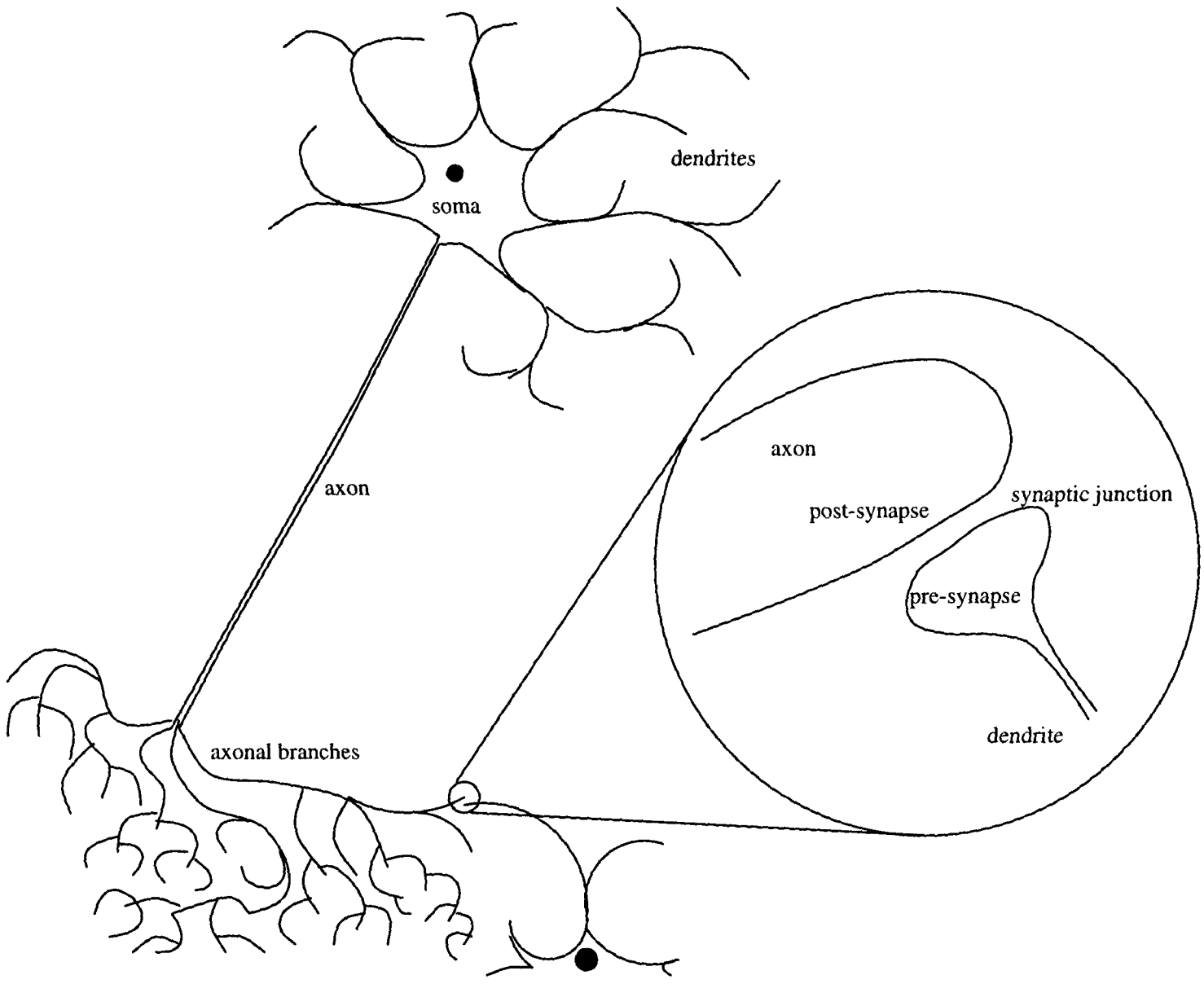


Figure 1.1: Sketch of a neuron

The operation of real (biological) neurons is obviously much too complicated as a starting point to model the large scale properties of a neural network. Whilst there has been much interest in deriving simplified equations directly from the Hodgkin-Huxley equations [28, 29], via 'integrate and fire' networks [30], the most useful set of equations has been derived by considering each neuron as a simple R-C circuit. In this formulation, we are not concerned with individual spikes of activity, but with short time averages, leading to continuous valued (soft spin) models [31]. Hopfield [32] noted that networks of continuous valued neurons have properties very similar to those of discrete valued neurons. These discrete valued neurons correspond to the rapid saturation of the transfer function, so that each neuron is either firing at its maximum rate, or resting. Hence neural networks can be modeled as a set of stochastic Ising spins, similar to spin-glasses, since the interactions (synaptic efficacies) have a quasi-random disordered form.

1.3.2 From circuit equation to Ising spins

The simplest microscopic equations we can envisage take the form of a set of first order differential equations - the *circuit equations.* These are derived by considering the neuron as a strictly electrical resistor - capacitor type circuit. The electrical spikes v_i arriving at the synapses induce a neuro transmitter concentration $C_i \propto v_i$, which causes a potential build up at the cell membrane ($\propto \sum_j J_{ij}C_j$) which acts as a capacitor. There is also a leakage of ions causing a discharging current plus the effect of any external currents. In the presence of noise the equations read

$$\frac{du_i}{dt} = \sum_j J_{ij}v_j - \frac{u_i}{\tau_i} + I_i \qquad\qquad v_i = g_i(\gamma_i(u_i - T_i + \eta_i)) \tag{1.9}$$

with solutions

$$u_i(t) = e^{-\frac{t}{\tau_i}} \int_0^t dt' e^{\frac{t'}{\tau_i}} \sum_j J_{ij} g_j(\gamma_j(u_j(t') - T_j + \eta_j(t'))) + I_i(t').$$

Consider $u_i(t+\Delta)$ with $\Delta \gg \tau$, omitting the external currents I_i

$$\begin{aligned} u_i(t+\Delta) &= e^{-\frac{t+\Delta}{\tau}} \int_0^{t+\Delta} dt'\, e^{\frac{t'}{\tau}} \sum_j J_{ij} g_j(\gamma_j(u_j(t') - T_j + \eta_j(t'))) \\ &\simeq \Delta \sum_j J_{ij} g_j(\gamma_j(u_j(t) - T_j + \eta_j(t))) \end{aligned} \tag{1.10}$$

$g(\gamma u)$ has the sigmoid shape therefore $v_i = \lim_{\gamma\to\infty} g(\gamma u_i) = \mathrm{sgn}(u_i) \to \sigma_i \in \{-1, 1\}$. Hence

$$\sigma_i(t+\Delta) = \mathrm{sgn}\Delta \left(\sum_j J_{ij}\sigma_j(t) + \tilde{\eta}_j(t) \right)$$

Writing $h_i = \sum_j J_{ij}\sigma_j$, this has the same form as (1.1) and therefore obeys the Markov equation (1.4) with transition probabilities (1.3).

Assuming the simplest model: that the build up of concentration of neuro transmitter (C_i) was proportional to the rate of arrival of spikes (v_i) at the synapse i.e. $C_i \propto v_i$, this leads to a familiar spin model which can be studied by statistical mechanical techniques. It seems however more plausible that the rate of build up of concentration is proportional to the spike rate i.e. $\frac{d}{dt}C_i \propto v_i$. We also include a decay term to represent dissipation of the concentration, as well as the decay of the PSP due to leakage of current across the cell membrane, and assume that the neuro transmitter is discharged whenever the neuron fires. With the other relationships as before, the equations governing the PSP's and spike rates of the neurons become

$$\begin{aligned} \frac{dC_i}{dt} &= \sum_{j=1}^{N} J_{ij}v_j - \frac{C_i}{\tilde{\tau}_i} \quad \Rightarrow \quad C_i(t) = e^{-\frac{t}{\tilde{\tau}_i}} \int^t dt'\, e^{\frac{t'}{\tilde{\tau}_i}} \sum_{j=1}^{N} J_{ij}v_j(t') \\ \frac{du_i}{dt} &= C_i - \frac{u_i}{\tau_i} + I_i \qquad\qquad v_i = g(\gamma_i(u_i - T_i + \eta_i)). \end{aligned} \tag{1.11}$$

Taking the $\gamma \to \infty$ limit as before, when $g(\gamma u)$ becomes a step function, we see that for $\Delta \gg \tau$

$$u_i(t+\Delta) \simeq \Delta C_i(t);$$

taking $\Delta = 1$, $\lim_{\gamma\to\infty} g(\gamma u) = \mathrm{sgn}(u)$, and $v_i \to \sigma_i$ we get

$$\sigma_i(t+1) = \mathrm{sgn}\left[\int^t dt'\, e^{\frac{t'-t}{\tilde{\tau}_i}} \sum_j J_{ij}\sigma_j(t') + \tilde{\eta}_i(t)\right].$$

i.e. The effect of the finite rate of build up of the neuro transmitter concentration is to introduce a short time average over the recent states of the neurons, with an exponentially decaying weight. Integrating over $\tilde{\eta}$ with the same measure as before gives

$$p(\sigma_i(t+1) = \tilde{\sigma}|\boldsymbol{\sigma}(t)) = \frac{1}{2}\left[1 + \tilde{\sigma}\tanh\beta\int^t dt'\, e^{\frac{t'-t}{\tilde{\tau}_i}} \sum_j J_{ij}\sigma_j(t')\right]. \tag{1.12}$$

This can be used as a starting point for further investigations, in a similar vein as the first order models.

1.3.3 The Hopfield model

The Hopfield model [33] incorporates Hebb's rule [34] (that the interaction between two neurons is strengthened whenever they fire simultaneously) and has (infinite range) interactions J_{ij} of the form $J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$, where $\xi_i^\mu \in \{-1, 1\}$ represents a single bit of information of the μ^{th} of p patterns stored by the network. In analysing the Hopfield model using statistical mechanical techniques, we are not interested in any particular realisation of the ξ's, rather in a 'typical' choice. Therefore we assume the ξ's are drawn at random from an uncorrelated, unbiased distribution, this allows us to average over the pattern variables ξ in an analogous way to averaging over the the disorder in the S-K model.

The equilibrium properties of the Hopfield model were calculated by Amit, Gutfreund and Sompolinsky, both for the case of $\alpha = \frac{p}{N} \to 0$ [35] and for $\alpha = \frac{p}{N} \sim \mathcal{O}(1)$ [36, 37]. Their results show that the Hopfield model appears to reproduce the general properties displayed by neural systems (a large number of attractors and a phase transition between retrieval and non-retrieval states as α and the noise β are varied). The biological reality of the asymmetry of synaptic connections, however is not embodied in the model. Because of this a generalisation of the Hopfield model was proposed [38], allowing for asymmetry of interactions, causing cyclical as well as static non-transient behaviour.

1.4 Deterministic Macroscopic Order Parameter Flow

From the stochastic microscopic description, our aim is derive deterministic laws for macroscopic observable quantities. In order to do this we have to derive a method of systematically removing microscopic effects whilst retaining the essential characteristics of the dynamics. There are essentially two approaches to this:

1. **Path Integral Method**

   This method involves writing down formal solutions to the equations (1.4) and (1.5) [39, 40] as a path integral representation of a generating functional.

2. **Macroscopic Probability Distribution Method**

   This method involves introducing a set of macroscopic order parameters $\boldsymbol{\Omega} = (\Omega^1, \ldots, \Omega^c)$ with their own macroscopic probability distribution $\mathcal{P}_t(\boldsymbol{\Omega}) = \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))$. Our aim is to choose the macroscopic order parameters such that the macroscopic probability distribution evolves deterministically, allowing us to derive deterministic laws for the macroscopic order parameters.

This thesis will be concerned solely with the second method.

1.4.1 Continuous Time

Consider first the master equation (1.5) with transition rates (1.6). Multiplying each side by $\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))$, and summing over all micro-states $\boldsymbol{\sigma}$ gives

$$\frac{d}{dt}\mathcal{P}_t(\boldsymbol{\Omega}(\boldsymbol{\sigma})) = \sum_{\boldsymbol{\sigma}} \sum_{k=1}^{N} p_t(\boldsymbol{\sigma}) w_k(\boldsymbol{\sigma}) \left\{\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(F_k\boldsymbol{\sigma})) - \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))\right\}.$$

Now carrying out a Taylor expansion in $\boldsymbol{\Delta}_i = \boldsymbol{\Omega}(F_i\boldsymbol{\sigma}) - \boldsymbol{\Omega}(\boldsymbol{\sigma})$ (the Kramers-Moyal expansion [10]) we see that

$$\frac{d\mathcal{P}_t(\boldsymbol{\Omega})}{dt} = \sum_{l\geq 1} \frac{(-1)^l}{l!} \sum_{\mu_1=1}^{c} \cdots \sum_{\mu_l=1}^{c} \frac{\partial^l}{\partial\Omega^{\mu_1}\ldots\partial\Omega^{\mu_l}} \left\{\mathcal{P}_t(\boldsymbol{\Omega}(\mathbf{s})) \left\langle \sum_{j=1}^{N} w_j(\boldsymbol{\sigma})\Delta_j^{\mu_1}(\boldsymbol{\sigma})\ldots\Delta_j^{\mu_l}(\boldsymbol{\sigma}) \right\rangle_{\boldsymbol{\Omega},t}\right\} \tag{1.13}$$

where

$$\langle f(\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},t} = \frac{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\,\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))\,f(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\,\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))} \qquad\qquad \Delta_j^{\mu}(\sigma) = \Omega^{\mu}(F_j\boldsymbol{\sigma}) - \Omega^{\mu}(\boldsymbol{\sigma}).$$

If $\lim_{N\to\infty} c\Delta\sqrt{N} = 0$ then it can be shown that only the first term $l = 1$ in the expansion is relevant, giving a Liouville form

$$\frac{d}{dt}\mathcal{P}_t(\boldsymbol{\Omega}(\boldsymbol{\sigma})) = -\sum_{\mu=1}^{c} \frac{\partial}{\partial\Omega^{\mu}} \left\{\mathcal{P}_t(\boldsymbol{\Omega}(\boldsymbol{\sigma})) \left\langle \sum_{j=1}^{N} w_j(\boldsymbol{\sigma})\Delta_j^{\mu}(\boldsymbol{\sigma}) \right\rangle_{\boldsymbol{\Omega};t}\right\}$$

with solutions $\mathcal{P}_t(\boldsymbol{\Omega}(\boldsymbol{\sigma})) = \int d\boldsymbol{\Omega}_0\, \mathcal{P}_0(\boldsymbol{\Omega}_0(\boldsymbol{\sigma}))\ \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}^*(t))$ where

$$\frac{d\boldsymbol{\Omega}^*(t)}{dt} = \left\langle \sum_{j=1}^{N} w_j(\boldsymbol{\sigma})\boldsymbol{\Delta}_j(\boldsymbol{\sigma}) \right\rangle_{\boldsymbol{\Omega}^*;t} \tag{1.14}$$

1.4.2 Discrete Time

Similarly consider the Markov equation (1.4) with transition probabilities (1.3). Multiplying both sides by the density of states factor $\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))$ and summming over all micro-states $\boldsymbol{\sigma}$ we get

$$\begin{aligned}
\mathcal{P}_{t+1}(\boldsymbol{\Omega}) &= \sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))\, W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma})\, p_t(\boldsymbol{\sigma}') \\
&= \int d\boldsymbol{\Omega}' \sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}(\boldsymbol{\sigma}))\, \delta(\boldsymbol{\Omega}' - \boldsymbol{\Omega}'(\boldsymbol{\sigma}'))\, W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma})\, p_t(\boldsymbol{\sigma}') \\
&= \int d\boldsymbol{\Omega}' \left\langle W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma})\right\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t} \mathcal{P}_t(\boldsymbol{\Omega}')
\end{aligned} \tag{1.15}$$

where

$$\langle f(\boldsymbol{\sigma}',\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t} = \frac{\sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'}\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))\,\delta(\boldsymbol{\Omega}'-\boldsymbol{\Omega}'(\boldsymbol{\sigma}'))\,f(\boldsymbol{\sigma}',\boldsymbol{\sigma})\,p_t(\boldsymbol{\sigma}')}{\sum_{\boldsymbol{\sigma}'}\delta(\boldsymbol{\Omega}'-\boldsymbol{\Omega}'(\boldsymbol{\sigma}'))p_t(\boldsymbol{\sigma}')}.$$

Using the transition probabilities (1.3) we see that

$$\begin{aligned}\langle W(\boldsymbol{\sigma}'\to\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t} &= \sum_{\boldsymbol{\sigma}}\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))\left\langle\prod_i\frac{e^{\beta\sigma_i h_i(\boldsymbol{\sigma}')}}{2\cosh\beta\sigma_i h_i(\boldsymbol{\sigma}')}\right\rangle_{\boldsymbol{\Omega}';t}\\ &= \left(\frac{N}{2\pi}\right)^c\int d\mathbf{K}e^{iN\mathbf{K}\cdot\boldsymbol{\Omega}}\sum_{\boldsymbol{\sigma}}e^{-iN\mathbf{K}\cdot\boldsymbol{\Omega}(\boldsymbol{\sigma})}\left\langle e^{\beta\sum_i\sigma_i h_i(\boldsymbol{\sigma}')-\sum_i\ln[2\cosh\beta h_i(\boldsymbol{\sigma}')]}\right\rangle_{\boldsymbol{\Omega}';t}\end{aligned}$$

In order to derive deterministic laws we require that $\langle W(\boldsymbol{\sigma}'\to\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t} = \delta\left[\boldsymbol{\Omega}-\mathbf{F}(\boldsymbol{\Omega}')\right]$. If in the limit $N\to\infty$ the expression for $\langle W(\boldsymbol{\sigma}'\to\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t}$ has a saddle point form so that

$$\frac{\int d\boldsymbol{\Omega}\,f(\boldsymbol{\Omega})\,\langle W(\boldsymbol{\sigma}'\to\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t}}{\int d\boldsymbol{\Omega}\,\langle W(\boldsymbol{\Omega}'\to\boldsymbol{\Omega})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t}} = \frac{\int d\boldsymbol{\Omega}\,d\mathbf{K}\,f(\boldsymbol{\Omega})\,e^{N\Psi(\boldsymbol{\Omega},\mathbf{K},\boldsymbol{\Omega}';t)}}{\int d\boldsymbol{\Omega}\,d\mathbf{K}\,e^{N\Psi(\boldsymbol{\Omega},\mathbf{K},\boldsymbol{\Omega}';t)}} = f(\boldsymbol{\Omega}^*(\boldsymbol{\Omega}';t))$$

where $\boldsymbol{\Omega}^*$ solves the saddle point problem $\frac{\partial\Psi}{\partial\boldsymbol{\Omega}} = \frac{\partial\Psi}{\partial\mathbf{K}} = 0$, then $\langle W(\boldsymbol{\sigma}'\to\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},\boldsymbol{\Omega}';t} = \delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}^*(\boldsymbol{\Omega}';t))$ as required, $\mathcal{P}_t(\boldsymbol{\Omega}) = \int d\boldsymbol{\Omega}'\,\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}^*(\boldsymbol{\Omega}';t))\,\mathcal{P}_t(\boldsymbol{\Omega}')$ and

$$\boldsymbol{\Omega}_{t+1} = \boldsymbol{\Omega}^*(\boldsymbol{\Omega}_t;t) \tag{1.16}$$

1.4.3 Removing Explicit Time Dependence

Equations (1.14) and (1.16) still contain explicit time dependence, stemming from the presence of $p_t(\boldsymbol{\sigma})$ in the average $\langle f(\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega};t}$. In order to calculate the average therefore, we are forced to solve the microscopic equations (1.4,1.5) which is exactly what we are trying to avoid. There are two ways of avoiding this difficulty:

1. Choosing the order parameters $\boldsymbol{\Omega}$ such that the local fields, and hence $\sum_{j=1}^N w_j(\boldsymbol{\sigma})\Delta_j^\mu(\boldsymbol{\sigma})$ and $\boldsymbol{\Omega}^*(\boldsymbol{\Omega}_t;t)$, depend on the microscopic variables $\boldsymbol{\sigma}$ only through the macroscopic variables $\boldsymbol{\Omega}$. In this case the microscopic probability distribution, and hence the explicit time dependence automatically drops out. This is the underlying approach employed in [38, 41, 42].

2. Choosing the order parameters $\boldsymbol{\Omega}$ such that equipartitioning of probability occurs within the subshells corresponding to a particular value of $\boldsymbol{\Omega}$, i.e. $p_t(\boldsymbol{\sigma}) = \frac{\mathcal{P}_t(\boldsymbol{\Omega}(\boldsymbol{\sigma}))}{\sum_{\boldsymbol{\sigma}'}\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}'))}$, in which case the average becomes $\frac{\sum_{\boldsymbol{\sigma}}\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))f(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}}\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))}$. This is the underlying approach employed in [1, 43, 44, 45, 46]. **NB** although we may be able to choose a state displaying equipartitioning as initial configurations, additional restrictions such as extreme dilution [43] may be required in order for this to be true for arbitrary time.

1.5 The Models

So far the development of the ideas to be embraced by this thesis have been essentially model independent, i.e. the formalism has been the same for spin-glass models, neural network models,

or indeed any model that can be mapped onto an Ising spin model with infinite range, disordered interactions. Essentially the only distinguishing feature of each model is the choice of interactions J_{ij}. The choice of the J_{ij}'s and the interpretation of the final results are the only times when physical insight distinguishes between the models. In this section we discuss the models - i.e. the reasoning behind the choice of interactions, how this effects the choice of order parameters, and the meaning of the results obtained. For a more thorough introduction to spin-glass and neural network models refer to [47] or [48].

In spin models of Neural Networks, the disorder is truly bond disorder. Here the synaptic efficacies i.e. interactions between the neurons are (quasi-) random. In spin-glasses, the disorder is site disorder. i.e. the magnetic impurities are placed randomly within a substrate. The random distances between atoms however, causes the interaction to be disordered, hence the problem is mapped onto a bond-disorder model.

1.5.1 S-K Model

The S-K model [15, 49] is the seminal, and perhaps most important infinite range spin-glass model. It is essentially the infinite range, Ising spin version of the Edwards-Anderson model [14], where magnetic impurities placed randomly on a lattice cause disordered interactions. The S-K model uses symmetric, infinite range interactions of the form

$$J_{ij} = \frac{\tilde{J}_0}{N} + \frac{\tilde{J} z_{ij}^s}{\sqrt{N}} \tag{1.17}$$

where the $z_{ij}^s = z_{ji}^s$ are drawn from a Gaussian distribution with zero mean and unit width; and $J_{ii} = 0$.

Using the replica method, outlined in Section 1.2, Sherrington and Kirkpatrick were able to derive an analytic expression for the free energy, via a saddle point integration, in terms of the magnetisation $m = \langle\langle\sigma\rangle\rangle$ and the spin-glass order parameter $q = \langle\langle\sigma\rangle\langle\sigma\rangle\rangle$. The inner brackets represent a site, or equivalently thermal average, whilst the outer brackets represent the disorder average. Their original formulation was within the so called replica symmetric (RS) ansatz, which was however later shown to be unstable in the low temperature region [50]. The RS ansatz turns out to be equivalent to assuming ergodicity [51], which can be broken in the low T phase, due to the extensively high 'hills' between 'valleys' in the energy landscape. Later Parisi proposed that q should in fact be an order parameter function $q(x)$ [19], measuring the overlap between two pure states.

A method for studying the dynamics in terms of flow equations for the order-parameters m and r (the disorder contribution to the energy) has recently been proposed [45].

1.5.2 The Generalised S-K Model

The generalised S-K model [52, 53, 54] is an attempt to retain the simplicity of interactions, and hence ease of averaging, of the fully symmetric S-K model, while introducing non-symmetric terms of variable strength. Biological realism in neural networks forces us to abandon the symmetry of interactions which is equivalent to lack of detailed balance. Hence standard equilibrium techniques cannot be employed, and we are forced to study the dynamics directly.

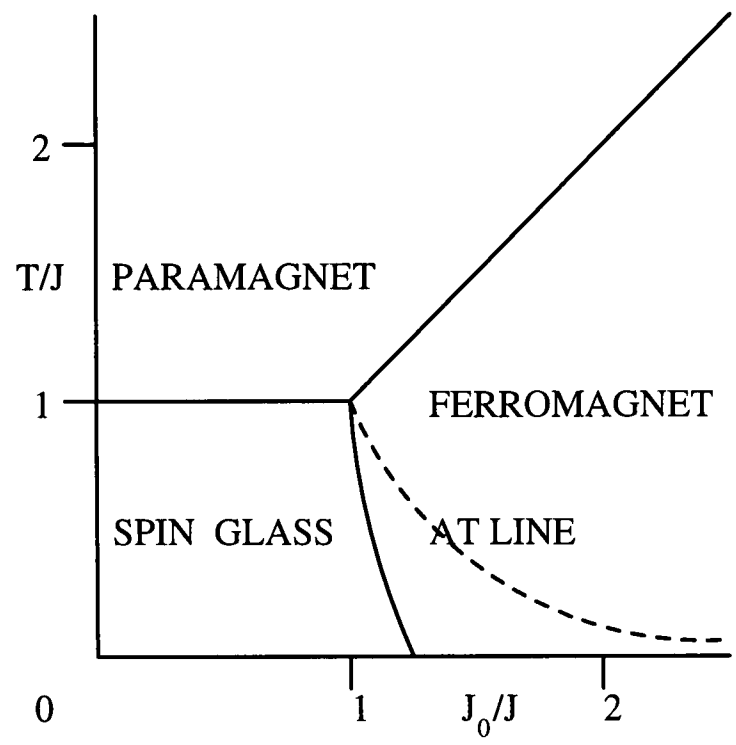


Figure 1.2: Phase diagram for the Sherrington-Kirkpatrick spin-glass model

The generalised S-K model has (asymmetric) interactions given by

$$J_{ij} = \frac{\tilde{J}_0}{N} + \frac{\tilde{J}\left(z_{ij}^s + k z_{ij}^a\right)}{\sqrt{N(1+k^2)}} \tag{1.18}$$

where the $z_{ij}^s = z_{ji}^s$ and $z_{ij}^a = -z_{ji}^a$ are drawn from a Gaussian distribution with zero mean and unit width; and $J_{ii} = 0$. $k = 0$ corresponds to the original S-K model; $k = 1$ corresponds to the fully asymmetric S-K model [53] where J_{ij} and J_{ji} are fully independent; and $k \to \infty$ corresponds to the fully antisymmetric S-K model.

Due to the lack of detailed balance, and the lack of an accepted formalism for studying the dynamics, there are few results for this model, except for $k = 0$. Some results may however be found in [53, 54, 55, 56, 57].

1.5.3 The Generalised Hopfield Model far from saturation

The Generalised Hopfield model [33] embedded with a finite number of patterns is in many ways an intermediate step between the uniform interactions of the infinite range ferro-magnet, and disordered systems such as the S-K model and the fully connected Hopfield model (i.e. with an extensive number of patterns). The interactions are given by

$$J_{ij} = \frac{1}{N} \sum_{\mu,\nu=1}^{p} \xi_i^\mu A_{\mu\nu} \xi_j^\nu, \tag{1.19}$$

where the $\xi_i^\mu \in \{-1, +1\}$ are drawn at random from an unbiased distribution and represent items of information stored by the network. $J_{ii} = 0$, and $\alpha = \frac{p}{N} \to 0$ in the limit $N \to \infty$.

Although the interactions are infinite range, due to the form of J_{ij} (1.19) there are only a finite number of disordered variables per spin. This makes disorder averaging much simpler than

for the case $\alpha = \frac{p}{N} \sim \mathcal{O}(1)$. Consider a site average of the form $\frac{1}{N}\sum_{i=1}^{N} F(\boldsymbol{\xi}_i)$, $\boldsymbol{\xi}_i \equiv (\xi_i^1, \ldots, \xi_i^p)$. Now, since there are a finite number, p, of patterns, there is a finite number of combinations of ξ_i^μ (2^p). Therefore in the limit $N \to \infty$ each of the combination will appear an infinite number of times and a site average can be replaced by an average over the 2^p possible values of ξ_i^μ. i.e.

$$\lim_{N\to\infty,\alpha=\frac{p}{N}\to 0} \frac{1}{N}\sum_{i=1}^{N} F(\boldsymbol{\xi}_i) = \langle F(\boldsymbol{\xi})\rangle_{\boldsymbol{\xi}} \equiv 2^{-p} \sum_{\boldsymbol{\xi}\in\{-1,+1\}^p} F(\boldsymbol{\xi}).$$

The equilibrium properties of this model with $A_{\mu\nu} = \delta_{\mu\nu}$ were studied in [35]. Relevant order parameters are found to be the overlaps $m^\mu = \langle\langle \xi^\mu \sigma\rangle\rangle$. We therefore choose our dynamic order parameters also to be the overlaps

$$\mathbf{m} = (m^1, \ldots, m^p) \qquad\qquad m^\mu = \frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu \sigma_i \tag{1.20}$$

If the number of patterns, p, is not too large ($p \ll \sqrt{N}$) [26], then in the thermodynamic limit the local field h_i can be written in terms of $\mathbf{m}$ only, and deterministic equations for the evolution in time of a set of macroscopic order parameters can be derived for both parallel [41, 42] and sequential [38, 42] dynamics. With this choice of order parameters $\Delta_i^\mu(\boldsymbol{\sigma}) = m^\mu(F_i\boldsymbol{\sigma}) - m^\mu(\boldsymbol{\sigma}) = -\frac{2}{N}\xi_i^\mu\sigma_i$. Equations (1.14) and (1.16) are then

$$\frac{d}{dt}\mathbf{m} = \langle \boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\rangle - \mathbf{m} \tag{1.21}$$

and

$$\mathbf{m}_{t+1} = \langle \boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_t)\rangle \tag{1.22}$$

respectively.

1.5.4 The Generalised Hopfield Model near saturation

The generalised Hopfield model [33] near saturation has interactions given by

$$J_{ij} = \sum_{\mu,\nu=1}^{p} \xi_i^\mu A_{\mu\nu}\xi_j^\nu \tag{1.23}$$

where the $\xi_i^\mu \in \{-1, +1\}$ are drawn at random from an unbiased distribution and represent items of information stored by the network. $J_{ii} = 0$, and $\alpha = \frac{p}{N} \sim \mathcal{O}(1)$ in the limit $N \to \infty$.

Since there are now an extensive number of random variables per spin, the disorder average cannot trivially be carried out as for the finite p case. Also, since there are an extensive number of patterns, there are an extensive number of overlaps. Hence we cannot simply use the overlaps as our order parameters, since we require the number of order parameters $c \ll \sqrt{N}$ for deterministic flow equations. This is as we might expect, since with an extensive number of order parameters, we have not integrated out the microscopic stochasticity, hence we can only expect stochastic equations. However, closer inspection of the definition of the order parameters reveals that only a finite number of them can be simultaneously condensed (i.e. $\mathcal{O}(1)$). After a moments thought, this becomes obvious - the spins cannot be simultaneously aligned to an extensive number of patterns. The effect of the uncondensed patterns (which are $\mathcal{O}(\frac{1}{\sqrt{N}})$) must still be taken into account however, since cumulatively they will have a sizable effect. This is the so-called condensed ansatz.

For $A_{\mu\nu} = \delta_{\mu\nu}$ we recover the Hopfield model, the equilibrium properties of which were studied in the seminal papers [36, 37]. A method for studying the dynamics in terms of flow equations for the order-parameters $\mathbf{m}$ (the condensed overlaps) and r the uncondensed contribution to the energy has recently been proposed [44, 58]. In order to generalise this to arbitrary $\mathbf{A}$ we choose as our dynamic order parameters the condensed overlaps $\mathbf{m}(\boldsymbol{\sigma}) = (m^1(\boldsymbol{\sigma}), \ldots, m^c(\boldsymbol{\sigma}))$, and a state variable $r(\boldsymbol{\sigma})$ to represent the uncondensed overlaps:

$$\begin{aligned} m^\mu(\boldsymbol{\sigma}) &= \frac{1}{N}\sum_i \xi_i^\mu \sigma_i \quad \mu = 1, \ldots, c \\ \alpha r(\boldsymbol{\sigma}) &= \sum_{[\mu,\nu>c \mid \mu>c,\ \nu\leq c \mid \mu\leq c,\ \nu>c]} \left[\frac{1}{N}\sum_j \xi_j^\mu \sigma_j\right] A_{\mu\nu} \left[\frac{1}{N}\sum_j \xi_j^\nu \sigma_j\right]. \end{aligned} \tag{1.24}$$

For symmetric systems $r(\boldsymbol{\sigma})$ is proportional to the disorder-dependent contribution to the Hamiltonian (the disorder-independent contribution being a function of $\mathbf{m}$). Since $r(\boldsymbol{\sigma})$ contains a symmetric sum, it only depends on the symmetric part $\mathbf{A}^s$ of the matrix $\mathbf{A}$. As in [44, 58, 59] we hereby build in the correct equilibrium behaviour. For non-symmetric systems, however, no such guide for choosing r is available and the only motivations are analogy with the symmetric case and (*a posteriori*) success of the resulting theory.

Part I

The Generalised Hopfield Model Far From Saturation

Chapter 2

Macroscopic Lyapunov functions

2.1 Introduction

In this chapter, we show how Lyapunov functions $\mathcal{L}(\mathbf{m})$ of the macroscopic dynamic order parameters that evolve according to (1.21, 1.22) (describing the evolution of the pattern overlaps m^μ, $\mu = 1 \ldots p$ in the generalised Hopfield model trained with a finite number of patterns) can be constructed, for arbitrary symmetric matrices $\mathbf{A}$.

From the equations of motion (1.21, 1.22), we can see that $\mathbf{m}_t \in [-1, +1]^p$ for all $t \geq 0$. This is immediately obvious for the discrete time case since $-1 \leq \langle \xi^\mu \tanh(\beta \boldsymbol{\xi} \cdot \mathbf{A}\mathbf{m}_t) \rangle \leq +1$. For the continuous time case, we can see for the same reason that each component of $\mathbf{m}$ will remain within $[-1, +1]$. We mention this despite the fact that due to its physical interpretation (1.20) $\mathbf{m} \in [-1, +1]^p$, since it is used in subsequent analysis which only builds on the equations of motion (1.21, 1.22).

At present, apart from general (microscopic) results like the H-theorem [10], the only Lyapunov functions that have been constructed for Ising spin neural networks apply to the zero-temperature case [33, 60], to strictly positive-definite or negative-definite interaction matrices [25, 61], or to specific zero-temperature models with delays [62]. In this chapter we address the problem of how to construct macroscopic Lyapunov functions corresponding to the dynamical laws which (in the thermodynamic limit) govern the evolution of order parameters, for arbitrary finite stochastic noise levels, and for both sequential and parallel microscopic dynamics. We consider only symmetric networks (i.e. with detailed balance). For both types of dynamics (sequential and parallel) we find a direct relation between the form of the macroscopic Lyapunov functions (which describe dynamical processes) and the saddle point integration that results from performing equilibrium statistical mechanical studies. This emphasises the equivalence of thermodynamic stability and dynamic stability previously demonstrated for the special case $\mathbf{A} = \mathbb{1}$ and sequential dynamics [25]. This study clearly confirms the intuitive picture of visualising the dynamics of symmetric neural networks as the minimisation of some state variable, to be interpreted as a dynamic 'free energy'.

2.2 The H - Theorem

The H-theorem guarantees the approach of the microscopic probability distribution $p_t(\boldsymbol{\sigma})$ to an equilibrium probability distribution $p_e(\boldsymbol{\sigma})$, for any stochastic process defined by a continuous time master equation which obeys detailed balance; provided that $p_e(\boldsymbol{\sigma})$ exists and is positive for each microscopic state $\boldsymbol{\sigma}$. It can quite easily be proved [10] that $\mathcal{L}_t \equiv \sum_{\boldsymbol{\sigma}} p_e(\boldsymbol{\sigma}) f\left(\frac{p_t(\boldsymbol{\sigma})}{p_e(\boldsymbol{\sigma})}\right)$ is a monotonically decreasing function for such processes, and is bounded from below, for any arbitrary convex function $f(x)$, i.e. $\forall x \geq 0: \ f(x) \geq 0, \ f''(x) > 0$. If the equilibrium probability distribution has the Boltzmann form, $p_e(\boldsymbol{\sigma}) \sim e^{-\beta \mathcal{H}(\boldsymbol{\sigma})}$, it is usual to make the choice $f(x) = x \ln[x]$. $\mathcal{L}_t$ now becomes

$$\mathcal{L}_t = \beta \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma}) \left[\mathcal{H}(\boldsymbol{\sigma}) + \frac{1}{\beta} \ln[p_t(\boldsymbol{\sigma})]\right]. \tag{2.1}$$

This is not a true Lyapunov function for the evolving macroscopic state vector $\mathbf{m}$, since it is a function of the microscopic probability distribution. However, intuitively one may hope to use (2.1) as a starting point for constructing a Lyapunov function for the macroscopic laws (1.21, 1.22).

If the Hamiltonian $\mathcal{H}(\boldsymbol{\sigma})$ depends on the microscopic state $\boldsymbol{\sigma}$ only through the values of the order parameters $\mathbf{m}$ (as is the case for the class of models considered here), by putting $\mathcal{H}(\boldsymbol{\sigma}) = N\epsilon(\mathbf{m}(\boldsymbol{\sigma}))$ we obtain

$$\frac{\mathcal{L}_t}{\beta N} = \int d\mathbf{m}\, \mathcal{P}_t(\mathbf{m}) \left[\epsilon(\mathbf{m}) + \frac{\sum_{\boldsymbol{\sigma}} \delta\left[\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})\right] p_t(\boldsymbol{\sigma}) \frac{1}{\beta N} \ln[p_t(\boldsymbol{\sigma})]}{\sum_{\boldsymbol{\sigma}} \delta\left[\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})\right] p_t(\boldsymbol{\sigma})}\right], \tag{2.2}$$

where $\mathcal{P}_t(\mathbf{m}) \equiv \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\, \delta\left[\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})\right]$ is the macroscopic probability distribution. For the case where the order parameters $\mathbf{m}$ evolve in time deterministically in the thermodynamic limit, $\mathcal{P}_t(\mathbf{m}) \to \delta\left[\mathbf{m} - \mathbf{m}_t\right]$ where $\mathbf{m}_t$ evolves according to (1.21, 1.22). The quantity (2.2), however, cannot be reduced to a function of $\mathbf{m}_t$ only without using additional information about the underlying microscopic states, or by making additional assumptions or approximations, because of the appearance of the entropic term $\ln[p_t(\boldsymbol{\sigma})]$. Such microscopic information would require knowledge of the solution of the equations (1.4, 1.5) which is exactly what one tries to avoid in deriving the macroscopic equations (1.21, 1.22).

An approximation which could be made would be to (incorrectly) assume equipartitioning of probability in the $\mathbf{m}$-subshells of the ensemble, which would imply making the replacement $p_t^{-1}(\boldsymbol{\sigma}) \sim \sum_{\boldsymbol{\sigma}} \delta\left[\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})\right] = e^{-Nc^*(\boldsymbol{\sigma})}$ $(N \to \infty)$ in (2.2), resulting in the appealing expression

$$\frac{\mathcal{L}_t}{\beta N} \ \to \ \frac{\mathcal{L}_t^{\text{equip}}}{\beta N} = \epsilon(\mathbf{m}) + \frac{c^*(\mathbf{m})}{\beta} \qquad \text{for } N \to \infty. \tag{2.3}$$

$c^*(\mathbf{m})$ is the Legendre transform of the cumulant generating function, derived from large deviations theory [25]:

$$c^*(\mathbf{m}) = \sup_{\mathbf{x}} \left(\mathbf{m} \cdot \mathbf{x} - \langle \ln\left[\cosh(\mathbf{x} \cdot \boldsymbol{\xi})\right] \rangle\right).$$

Unfortunately, out of equilibrium the assumption of subshell-equipartitioning of probability is unsustainable. $p_t(\boldsymbol{\sigma})$ is the full complicated solution of the microscopic laws (1.4 ,1.5) and will exhibit equipartitioning in the macroscopic subshells *only in equilibrium*. To see this, assume

that at $t = 0$ we prepare an ensemble distribution $p_0(\boldsymbol{\sigma})$ obeying equipartitioning within the $\mathbf{m}$ subshells, i.e. $p_0(\boldsymbol{\sigma}) = f[\mathbf{m}(\boldsymbol{\sigma})]$. According to the master equation (1.5) we now find

$$\left.\frac{d}{dt}p_t(\boldsymbol{\sigma})\right|_{t=0} = \frac{1}{2}\sum_i \left\{ f\left[\boldsymbol{m}(\boldsymbol{\sigma}) - \frac{2}{N}\boldsymbol{\xi}_i\sigma_i\right] - f\left[\boldsymbol{m}(\boldsymbol{\sigma})\right]\right\} + \frac{1}{2}\sum_i \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}))\left\{ f\left[\boldsymbol{m}(\boldsymbol{\sigma}) - \frac{2}{N}\boldsymbol{\xi}_i\sigma_i\right] + f\left[\boldsymbol{m}(\boldsymbol{\sigma})\right]\right\}.$$

Expanding in powers of N, using $h_i(\boldsymbol{\sigma}) = \boldsymbol{\xi}_i \cdot \mathbf{A}\boldsymbol{m}(\boldsymbol{\sigma}) - \frac{1}{N}\boldsymbol{\xi}_i \cdot \mathbf{A}\boldsymbol{\xi}_i\sigma_i$ we get

$$\begin{aligned}\left.\frac{d}{dt}p_t(\boldsymbol{\sigma})\right|_{t=0} = & -\nabla_{\boldsymbol{m}} f\left[\boldsymbol{m}(\boldsymbol{\sigma})\right] \cdot \left[\boldsymbol{m}(\boldsymbol{\sigma}) + \langle \boldsymbol{\xi} \tanh(\beta\boldsymbol{\xi} \cdot \mathbf{A}\boldsymbol{m}(\boldsymbol{\sigma}))\rangle_{\boldsymbol{\xi}}\right] \\ & - f\left[\boldsymbol{m}(\boldsymbol{\sigma})\right]\frac{\beta}{N}\sum_i \left[\boldsymbol{\xi}_i \cdot \mathbf{A}\boldsymbol{\xi}_i\right]\left[1 - \tanh^2(\beta\boldsymbol{\xi}_i \cdot \mathbf{A}\boldsymbol{m}(\boldsymbol{\sigma}))\right] \\ & + f\left[\boldsymbol{m}(\boldsymbol{\sigma})\right]\sum_i \sigma_i \tanh(\beta\boldsymbol{\xi}_i \cdot \mathbf{A}\boldsymbol{m}(\boldsymbol{\sigma})) + \mathcal{O}(N^{-1}). \end{aligned} \tag{2.4}$$

Equipartitioning at $t = 0$ is sustained for $t > 0$ only if the above expression depends on the micro-variables $\boldsymbol{\sigma}$ only through the macro-variables $\boldsymbol{m}(\boldsymbol{\sigma})$. The last of the above terms violates this requirement (except for the trivial case $p = 1$ where we recover the ∞-range ferromagnet). Therefore equipartitioning at $t = 0$ does not imply equipartitioning at $t > 0$. This is in contrast to a formulation in terms of sublattice magnetisations $m_{\boldsymbol{\xi}} = \frac{1}{I_{\boldsymbol{\xi}}}\sum_{i\in I_{\boldsymbol{\xi}}}\sigma_i$, where the sublattices $I_{\boldsymbol{\xi}}$ consist of all those sites for which $\boldsymbol{\xi}_i = \boldsymbol{\xi}$ [25]. Deterministic evolution of these order parameters (which are akin to the magnetisation in the ∞-range ferromagnet), however, requires $p \ll \log N$, rather than the much greater number afforded by our description at the level of overlaps, requiring only $p^2 \ll N$.

Therefore the H-theorem cannot be used to prove *a priori* that (2.3) is a true macroscopic Lyapunov function in the thermodynamic limit for the macroscopic laws (1.21, 1.22). Nevertheless we will see in Section 2.4 that for separable systems this is indeed the case, and can even be generalised to discrete-time parallel dynamics.

2.3 Strictly Positive or Strictly Negative Matrices A

It is reasonably easy to find a macroscopic Lyapunov function, $\mathcal{L}(\mathbf{m})$, for the differential equation (1.21), when the symmetric matrix $\mathbf{A}$ is either positive definite ($\mathcal{L}^+(\mathbf{m})$) or negative definite ($\mathcal{L}^-(\mathbf{m})$) [25, 35, 61]:

$$\mathcal{L}^{\pm}(\mathbf{m}) \equiv \pm\frac{1}{2}\mathbf{m} \cdot \mathbf{A}\mathbf{m} \mp \frac{1}{\beta}\left\langle \ln\left[\cosh(\beta\boldsymbol{\xi} \cdot \mathbf{A}\mathbf{m})\right]\right\rangle, \tag{2.5}$$

since

$$\frac{d}{dt}\mathcal{L}^{\pm} = \pm\left(\mathbf{m} - \langle \boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi} \cdot \mathbf{A}\mathbf{m})\rangle\right) \cdot \mathbf{A}\frac{d\mathbf{m}}{dt} = \mp\frac{d\mathbf{m}}{dt} \cdot \mathbf{A}\frac{d\mathbf{m}}{dt} \leq 0.$$

$\frac{d}{dt}\mathcal{L}^{\pm}$ can only be zero when $\frac{d}{dt}\mathbf{m} = 0$, since $\mathbf{A}$ is strictly positive or strictly negative definite. It is also easy to convince oneself that $\mathcal{L}^{\pm}$ is bounded from below, since $\mathbf{m}$ is only defined on the interval $[-1, +1]^p$, and the expression (2.5) contains no singularities. Hence $\mathcal{L}^{\pm}$ is a Lyapunov function for the macroscopic equations (1.21) derived for asynchronous updating, and strictly positive or strictly negative matrices $\mathbf{A}$ respectively.

For synchronous updating we know from a microscopic analysis that at $T = 0$ the network will again settle into an equilibrium configuration for $\mathbf{A}$ positive definite, and a period two cycle for $\mathbf{A}$ negative definite [20]. Here we show that the Lyapunov function $\mathcal{L}^+$ (2.5) of the sequential case is also a Lyapunov function of the macroscopic equations (1.22) derived for the parallel case for $\mathbf{A}$ positive definite.

First note that the macroscopic parallel dynamics (1.22) can be written

$$\mathbf{m}_{t+1} = \mathbf{m}_t - \mathbf{A}^{-1} \cdot \boldsymbol{\nabla}_{\mathbf{m}} \mathcal{L}^+(\mathbf{m}_t).$$

Writing $\frac{\partial}{\partial x^\mu}\mathcal{L}^+(\mathbf{x}) = \partial_\mu \mathcal{L}^+(\mathbf{x})$, and using the following identities

$$\mathcal{L}^+(\mathbf{x}+\mathbf{y}) = \mathcal{L}^+(\mathbf{x}) + \sum_\mu y^\mu \partial_\mu \mathcal{L}^+(\mathbf{x}) + \int_0^1 \lambda\, d\lambda \int_0^1 d\rho \sum_{\mu,\nu} y^\mu y^\nu \partial^2_{\mu\nu} \mathcal{L}^+(\mathbf{x}+\lambda\rho\mathbf{y})$$

$$\partial^2_{\rho\alpha}\mathcal{L}^+(\mathbf{x}) = A_{\rho\alpha} - \beta \sum_{\mu,\eta} \Gamma_{\mu\eta}(\mathbf{x}) A_{\eta\rho} A_{\mu\alpha} \qquad \Gamma_{\mu\eta}(\mathbf{x}) = \left\langle \xi^\mu \xi^\eta \left(1 - \tanh^2(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{x})\right)\right\rangle$$

(where the symmetric matrix $\boldsymbol{\Gamma}(\mathbf{x})$ has only non-negative eigenvalues), we obtain for symmetric and invertible $\mathbf{A}$

$$\begin{aligned}
\Delta\mathcal{L}^+ &= \mathcal{L}^+(\mathbf{m}_{t+1}) - \mathcal{L}^+(\mathbf{m}_t) = \mathcal{L}^+\left(\mathbf{m}_t - \nabla \mathbf{A}^{-1}\mathcal{L}^+_{\mathbf{m}_t}\right) - \mathcal{L}^+(\mathbf{m}_t) \\
&= -\boldsymbol{\nabla}_{\mathbf{m}}\mathcal{L}^+ \cdot \left\{ \int_0^1 \lambda\, d\lambda \int_0^1 d\rho\, \mathbf{A}^{-1} + \beta\boldsymbol{\Gamma}(\mathbf{m}_t - \lambda\rho\mathbf{A}^{-1}\boldsymbol{\nabla}_{\mathbf{m}}\mathcal{L}^+) \right\} \boldsymbol{\nabla}_{\mathbf{m}}\mathcal{L}^+ \\
&\leq 0. \qquad (2.6)
\end{aligned}$$

Equality only holds when $\partial_\mu \mathcal{L}^+ = 0\ \forall\ \mu$, which reduces to exactly the condition for a fixed point $\mathbf{m} = \langle \boldsymbol{\xi} \tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\rangle$. Hence $\mathcal{L}^+$ decreases monotonically until a fixed point is reached.

2.4 Construction for Arbitrary Separable Models

In equilibrium we know that the properties of the network can be determined by using a generating function - the free energy for the sequential case, and Peretto's pseudo free energy [60] for the parallel case. Consider the free energy per neuron, f, with $\mathcal{H}$ as either the Ising Hamiltonian or Peretto's pseudo Hamiltonian (as appropriate)

$$f = -\frac{1}{\beta N} \ln Z = -\frac{1}{\beta N} \ln \left[\sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})} \right].$$

If the (extensive) Hamiltonian depends on the system state $\boldsymbol{\sigma}$ only through the macroscopic variables $\{m^\mu\}$ then we can write $\mathcal{H}(\boldsymbol{\sigma}) = N\epsilon(\mathbf{m}(\boldsymbol{\sigma}))$. Along with the density of microscopic states $\mathcal{D}(\mathbf{m}) = \sum_{\boldsymbol{\sigma}} \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})) = e^{-Nc^*(\mathbf{m})}$ $(N \to \infty)$ [25] we can now write the free energy for large N as

$$f = -\frac{1}{\beta N} \ln \left[\int d\mathbf{m}\, e^{-N(c^*(\mathbf{m}) + \beta\epsilon(\mathbf{m}))} \right] = \frac{1}{\beta} \min_{\mathbf{m}\in\mathcal{R}^p} \mathcal{L}(\mathbf{m}) = \frac{1}{\beta} c^*(\tilde{\mathbf{m}}) + \epsilon(\tilde{\mathbf{m}}) \qquad (2.7)$$

where $\tilde{\mathbf{m}}$ is the value of $\mathbf{m}$ which minimises $c^*(\mathbf{m}) + \beta\epsilon(\mathbf{m})$ (since the integral will be extremally dominated). We will prove that $\frac{1}{\beta}c^*(\mathbf{m}) + \epsilon(\mathbf{m})$, with $\epsilon(\mathbf{m}(\boldsymbol{\sigma}))$ as appropriate to the dynamics, is a Lyapunov function for both types of macroscopic laws (1.21,1.22) with symmetric $\mathbf{A}$.

We will first prove a property of $c^*(\mathbf{m})$, namely that for physically realisable $\mathbf{m}$ (i.e. $m^\mu = \frac{1}{N}\sum_i \xi_i^\mu \sigma_i$, $|m^\mu| \leq 1$) the supremum is satisfied by a finite critical point rather than at $\mathbf{x} = \infty$. This result is required in the proofs to follow. If we write $\mathbf{x} = \hat{\mathbf{x}}x$, where $\hat{\mathbf{x}}$ is a unit vector in the direction of $\mathbf{x}$ and x is its modulus, and use the identity $\ln[\cosh(z)] = -\ln 2 + |z| + \ln\left[1 + e^{-2|z|}\right]$, then the quantity to be maximised becomes

$$x\left(\mathbf{m}\cdot\hat{\mathbf{x}} - \langle|\boldsymbol{\xi}\cdot\hat{\mathbf{x}}|\rangle\right) + \ln 2 - \left\langle \ln\left[1 + e^{-2x|\boldsymbol{\xi}\cdot\hat{\mathbf{x}}|}\right]\right\rangle . \tag{2.8}$$

We can write $\mathbf{m}$ in terms of the sublattice magnetisations $m_{\boldsymbol{\xi}} \in [-1,+1]$, $\mathbf{m} = \left\langle \boldsymbol{\xi} m_{\boldsymbol{\xi}}\right\rangle$, where $m_{\boldsymbol{\xi}}(\boldsymbol{\sigma}) = \frac{1}{I_{\boldsymbol{\xi}}}\sum_{i\in I_{\boldsymbol{\xi}}}\sigma_i$ and the sublattices $I_{\boldsymbol{\xi}}$ consist of those sites i for which $\boldsymbol{\xi}_i = \boldsymbol{\xi}$. As a result for large x the expression in (2.8) behaves as

$$x\left\langle \boldsymbol{\xi}\cdot\hat{\mathbf{x}}\; m_{\boldsymbol{\xi}} - |\boldsymbol{\xi}\cdot\hat{\mathbf{x}}|\right\rangle + \ln 2 = x\left\langle |\boldsymbol{\xi}\cdot\hat{\mathbf{x}}|\left(m_{\boldsymbol{\xi}}\,\mathrm{sgn}(\boldsymbol{\xi}\cdot\hat{\mathbf{x}}) - 1\right)\right\rangle + \ln 2.$$

Clearly this is not maximised in the limit $x \to \infty$ for physical $m_{\boldsymbol{\xi}}$ ($\in [-1,+1]$). Hence the supremum is bounded so must be realised by a finite critical point.

2.4.1 Parallel Dynamics

The microscopic properties of the models with synchronous updating (1.22) are obtained by using the Peretto pseudo-Hamiltonian [60]

$$\mathcal{H}(\boldsymbol{\sigma}) = -\frac{1}{\beta}\sum_i \ln\left[2\cosh\beta\sum_{j\neq i} J_{ij}\sigma_j\right] = -\frac{N}{\beta}\left\langle \ln\left[2\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{Am}(\boldsymbol{\sigma}))\right]\right\rangle + \mathcal{O}(1)$$

so that

$$\epsilon_{par}(\mathbf{m}) = -\frac{1}{\beta}\left\langle \ln\left[2\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{Am})\right]\right\rangle ,$$

and

$$\mathcal{L}_{par}(\mathbf{m}) = c^*(\mathbf{m}) - \left\langle \ln\left[2\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{Am})\right]\right\rangle . \tag{2.9}$$

THEOREM 1 *$\mathcal{L}_{par}(\mathbf{m})$ is a monotonically decreasing function of the process (1.22) if $\mathbf{A}^\dagger = \mathbf{A}$, only stationary when in a period-two cycle.*

PROOF

Consider $\mathcal{L}_{par}(\mathbf{m}_{t+1})$. Taking the gradient of the argument of the supremum, and substituting the dynamics (1.22), we find that it is realised amongst those $\mathbf{x}$ satisfying $\langle\boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{Am}_t)\rangle = \langle\boldsymbol{\xi}\tanh\boldsymbol{\xi}\cdot\mathbf{x}\rangle$. Since the supremum is realised at a finite critical point we can write for those $\mathbf{x}$ satisfying the extremisation criterion

$$0 = \langle\boldsymbol{\xi}\left[\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{Am}) - \tanh(\boldsymbol{\xi}\cdot\mathbf{x})\right]\rangle = \left\langle \boldsymbol{\xi}\int_{\boldsymbol{\xi}\cdot\mathbf{x}}^{\beta\boldsymbol{\xi}\cdot\mathbf{Am}} dz\,\left(1 - \tanh^2 z\right)\right\rangle .$$

Then using the substitution $z = \boldsymbol{\xi}\cdot(\mathbf{x} + \lambda(\beta\mathbf{Am} - \mathbf{x}))$, and taking the inner product with $\beta\mathbf{Am} - \mathbf{x}$ gives

$$0 = \left\langle [\boldsymbol{\xi}\cdot(\beta\mathbf{Am} - \mathbf{x})]^2\int_0^1 d\lambda\,\left(1 - \tanh^2\left[\boldsymbol{\xi}\cdot(\mathbf{x} - \lambda(\beta\mathbf{Am} - \mathbf{x}))\right]\right)\right\rangle ,$$

which implies that $\mathbf{x} = \beta \mathbf{A}\mathbf{m}_t$, since the integral can never be zero. If we use this to calculate $\Delta\mathcal{L}_{par} = \mathcal{L}_{par}(\mathbf{m}_{t+1}) - \mathcal{L}_{par}(\mathbf{m}_t)$ then

$$\begin{aligned}
\Delta\mathcal{L}_{par} &= \beta\mathbf{m}_{t+1}\cdot\mathbf{A}\mathbf{m}_t - \langle \ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_t)\right]\rangle - \langle \ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_{t+1})\right]\rangle \\
&\quad - \sup_{\mathbf{y}}\left(\mathbf{m}_t\cdot\mathbf{y} - \langle \ln\left[\cosh(\boldsymbol{\xi}\cdot\mathbf{y})\right]\rangle\right) + \langle \ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_t)\right]\rangle \\
&= \mathbf{m}_t\cdot(\beta\mathbf{A}\mathbf{m}_{t+1}) - \langle \ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_{t+1})\right]\rangle - \sup_{\mathbf{y}}\left(\mathbf{m}_t\cdot\mathbf{y} - \langle \ln\left[\cosh(\boldsymbol{\xi}\cdot\mathbf{y})\right]\rangle\right) \\
&\leq 0. \qquad (2.10)
\end{aligned}$$

$\Delta\mathcal{L}_{par} = 0$ requires $\mathbf{y} = \beta\mathbf{A}\mathbf{m}_{t+1}$, the supremum condition for $\mathbf{y}$ is $\mathbf{m}_t = \langle \boldsymbol{\xi}\tanh(\boldsymbol{\xi}\cdot\mathbf{y})\rangle$, hence a stationary $\mathcal{L}_{par}$ implies

$$\mathbf{m}_t = \langle \boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m}_{t+1})\rangle = \mathbf{m}_{t+2}.$$

Therefore $\mathcal{L}_{par}$ is a monotonically decreasing function of the macroscopic dynamics (1.22) for $\mathbf{A} = \mathbf{A}^\dagger$, and is only stationary when the system is in a period-two cycle. Since $\mathbf{m}$ is only defined on $[-1,+1]^p$, and $\mathcal{L}_{par}$ has no poles, it must be bounded from below, and hence satisfies the necessary conditions to be a Lyapunov function of the macroscopic synchronous dynamics (1.22).

2.4.2 Sequential Dynamics

For the asynchronous Glauber dynamics (1.21) we must use the usual Ising Hamiltonian

$$\mathcal{H}(\boldsymbol{\sigma}) = -\frac{1}{2N}\sum_{i,j\neq i} J_{ij}\sigma_i\sigma_j = -\frac{1}{2N}\sum_{i,j}\sum_{\mu,\nu}\sigma_i\xi_i^\mu A_{\mu\nu}\xi_j^\nu\sigma_j + \text{const.} = -\frac{N}{2}\mathbf{m}\cdot\mathbf{A}\mathbf{m} + \text{const.}$$

so that

$$\epsilon_{seq}(\mathbf{m}) = -\frac{1}{2}\mathbf{m}\cdot\mathbf{A}\mathbf{m},$$

and

$$\mathcal{L}_{seq}(\mathbf{m}) = c^*(\mathbf{m}) - \frac{\beta}{2}\mathbf{m}\cdot\mathbf{A}\mathbf{m}. \qquad (2.11)$$

THEOREM 2 *$\mathcal{L}_{seq}(\mathbf{m})$ is monotonically decreasing function of the process (1.21) if $\mathbf{A}^\dagger = \mathbf{A}$, only stationary when in a fixed point.*

PROOF

The supremum is realised amongst those $\mathbf{x}$ satisfying $\mathbf{m} = \langle \boldsymbol{\xi}\tanh\boldsymbol{\xi}\cdot\mathbf{x}\rangle$, since the argument of $\mathcal{L}_{seq}$ remains bounded. If we substitute this relation into $\mathcal{L}_{seq}(\mathbf{m})$ we get

$$\begin{aligned}
\mathcal{L}_{seq}(\mathbf{m}) &= \langle \boldsymbol{\xi}\cdot\mathbf{x}\tanh(\boldsymbol{\xi}\cdot\mathbf{x}) - \ln\left[\cosh(\boldsymbol{\xi}\cdot\mathbf{x})\right]\rangle - \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} \\
&= \left\langle \int_0^{\boldsymbol{\xi}\cdot\mathbf{x}} dy\, y\,(1-\tanh^2 y)\right\rangle - \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m}.
\end{aligned}$$

Differentiating with respect to one component of $\mathbf{m}$ we find using $\mathbf{A}^\dagger = \mathbf{A}$

$$\frac{\partial\mathcal{L}_{seq}(\mathbf{m})}{\partial m^\alpha} = \sum_\gamma \left\langle \boldsymbol{\xi}\cdot\mathbf{x}\,\xi^\gamma(1-\tanh^2(\boldsymbol{\xi}\cdot\mathbf{x}))\right\rangle \frac{\partial x^\gamma}{\partial m^\alpha} - \beta(\mathbf{A}\mathbf{m})^\alpha.$$

The supremum, however is realised amongst those $\mathbf{x}$ satisfying $\mathbf{m} = \langle \boldsymbol{\xi}\tanh\boldsymbol{\xi}\cdot\mathbf{x}\rangle$. Differentiating this equation gives

$$\delta_{\mu\alpha} = \frac{\partial m^\mu}{\partial m^\alpha} = \sum_\nu \left\langle \xi^\mu\xi^\nu(1-\tanh^2(\boldsymbol{\xi}\cdot\mathbf{x}))\right\rangle \frac{\partial x^\nu}{\partial m^\alpha}.$$

Combining these previous two equations we see that the $\mathbf{x}$ realising the supremum satisfies $\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m}) = \mathbf{x} - \beta\mathbf{A}\mathbf{m}$. Substituting this into our expression for $\mathcal{L}_{seq}(\mathbf{m})$ gives

$$\mathcal{L}_{seq}(\mathbf{m}) = \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} + \mathbf{m}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m}) - \langle \ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m} + \boldsymbol{\xi}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m}))\right]\rangle \,.$$

Now $\frac{d\mathcal{L}_{seq}(\mathbf{m})}{dt} = \nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m})\cdot\frac{d\mathbf{m}}{dt}$, using the equation of motion for $\mathbf{m}$ (1.21) this gives

$$\begin{aligned}\frac{d\mathcal{L}_{seq}(\mathbf{m})}{dt} &= -\mathbf{m}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m}) + \langle\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m})\cdot\boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\rangle \\ &= -\mathcal{L}_{seq}(\mathbf{m}) - \langle\ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m} + \boldsymbol{\xi}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m}))\right]\rangle \\ &\quad + \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} + \langle\nabla_{\mathbf{m}}\mathcal{L}_{seq}(\mathbf{m})\cdot\boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\rangle \,.\end{aligned}$$

If we let $\boldsymbol{\xi}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq} = \lambda_{\boldsymbol{\xi}}$, then the maximum of $\frac{d\mathcal{L}_{seq}}{dt} + \mathcal{L}_{seq}$ with respect to variation of $\lambda_{\boldsymbol{\xi}}$, is satisfied for $\lambda_{\boldsymbol{\xi}} = 0\ \forall\boldsymbol{\xi}$. Hence we can write

$$\begin{aligned}\frac{d\mathcal{L}_{seq}(\mathbf{m})}{dt} &\leq -\mathcal{L}_{seq}(\mathbf{m}) + \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} - \langle\ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\right]\rangle \\ &\leq -\mathcal{L}_{seq}(\mathbf{m}) + \left((\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} - \langle\ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\right]\rangle) - \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m}\right) \\ &\leq \left((\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m} - \langle\ln\left[\cosh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\right]\rangle) - \frac{1}{2}\beta\mathbf{m}\cdot\mathbf{A}\mathbf{m}\right) \\ &\quad - \left(\sup_{\mathbf{x}}\left(\mathbf{m}\cdot\mathbf{x} - \langle\ln\left[\cosh(\mathbf{x}\cdot\boldsymbol{\xi})\right]\rangle\right) - \frac{\beta}{2}\mathbf{m}\cdot\mathbf{A}\mathbf{m}\right) \\ &\leq 0 \quad . \end{aligned} \tag{2.12}$$

$\mathcal{L}_{seq}(\mathbf{m})$ is only stationary when the above mentioned maximum of $\frac{d\mathcal{L}_{seq}}{dt} + \mathcal{L}_{seq}$ is realised, i.e. $\boldsymbol{\xi}\cdot\nabla_{\mathbf{m}}\mathcal{L}_{seq} = 0$ for all $\boldsymbol{\xi}$. From this we conclude that $\nabla_{\mathbf{m}}\mathcal{L}_{seq} = 0$ and hence $\mathbf{x} = \beta\mathbf{A}\mathbf{m}$. Combination with the supremum criterion $\mathbf{m} = \langle\boldsymbol{\xi}\tanh(\boldsymbol{\xi}\cdot\mathbf{x})\rangle$ subsequently gives

$$\mathbf{m} = \langle\boldsymbol{\xi}\tanh(\beta\boldsymbol{\xi}\cdot\mathbf{A}\mathbf{m})\rangle \tag{2.13}$$

which is a fixed point of the dynamics (1.21). The only remaining constraint on $\mathcal{L}_{seq}(\mathbf{m})$ for it to be a Lyapunov function is that it is bounded from below, which is obviously the case, since $\mathbf{m}$ only exists in the range $[-1,+1]^p$, and $\mathcal{L}_{seq}(\mathbf{m})$ has no poles. Hence $\mathcal{L}_{seq}(\mathbf{m})$ is a Lyapunov function of the macroscopic asynchronous dynamics (1.21).

2.5 Conclusion

We have shown that for Ising spin models of neural networks with long-range separable symmetric interactions of the form (1.19), Lyapunov functions of macroscopic variables (the pattern overlaps m^{μ}, $\mu = 1\ldots p$ evolving according to (1.21, 1.22) exist. We have generalised existing Lyapunov functions of the 'free energy' type to finite temperatures and arbitrary separable symmetric embedding matrices. We have shown that the proposed Lyapunov functions correspond exactly to the scalar surfaces that are encountered in the saddle-point integration resulting from equilibrium statistical mechanical studies, emphasising the equivalence of thermodynamic and dynamic stability. We can therefore interpret the dynamics of the present type of symmetric network as (not necessarily gradient) descent on a 'free energy' surface, for both parallel and sequential updating

of the individual spins. The macroscopic dynamical equations which form the basis of our calculations are only strictly valid for $p \ll \sqrt{N}, N \to \infty$, but we may suspect that similar results can be obtained for the equations governing the behaviour for $p = \alpha N$ as suggested by recent dynamical studies [58].

Chapter 3

Quasi-periodicity and bifurcation phenomena

3.1 Introduction

In Chapter 2 we saw how Lyapunov functions can be constructed for the equations (1.21, 1.22) (describing the evolution of the pattern overlaps in the generalised Hopfield model trained with a finite number of patterns) if the embedding matrix $\mathbf{A}$ is symmetric, guaranteeing approach to equilibrium. For asymmetric $\mathbf{A}$ however, exact results are usually impossible to obtain and we must resort to the qualitative techniques of bifurcation theory, and numerical iteration of the dynamical equations.

This chapter is devoted to the study of these deterministic macroscopic equations. For asymmetric interactions (i.e. systems without detailed balance) we expect complex behaviour, with long period cycles or perhaps even chaotic trajectories. We illustrate typical behaviour with specific examples, and show that the complex behaviour that a network exhibits can be qualitatively understood using bifurcation theory, which can be used to reduce a complex high dimensional system to a low dimensional set of equations containing only essential terms from which generic behaviour can be deduced.

Much work has already been done on the complex behaviour of the dynamics of neural networks, and it has been shown that chaotic behaviour is possible in systems with continuous (Langevin type) neurons [63, 64, 65, 66, 67, 68, 69, 70]. Complex behaviour has also been found in systems with discrete-valued neurons [71, 72], though chaos is not possible due to the finite number of microscopic configurations of the system. In all these studies, however, description is at the microscopic level of individual neurons. In contrast, this chapter aims at identifying and studying complexity at the deterministic and low-dimensional level of order parameter evolution.

In the limit of high noise the only stable fixed point is the origin, representing a paramagnetic macroscopic state. As the noise is varied bifurcations to non-trivial behaviour occur. The form of bifurcation from the trivial fixed point depends solely on one relevant eigenvalue of the embedding matrix which determines the neuronal interactions.

3.2 Bifurcation Analysis

A brief introduction to the theory of bifurcation analysis for studying sets of non-linear differential equations, or maps, can be found in Appendix A. A more thorough discussion of methods can be found in [73, 74]. The basic tenet of bifurcation theory is that sufficiently close to a bifurcation (i.e. where there is a qualitative change in the form of solutions to a set of dynamic equations), the behaviour of the full complicated set of equations can be reduced to a few important degrees of freedom. First the relevant degrees of freedom must be identified by local linear analysis and the full set of equations reduced to the equations on the centre manifold (the manifold to which the solution is confined sufficiently close to the bifurcation point). Then unnecessary non-linear terms removed as far as possible, to leave the normal form. The normal form will fall into one of a small number of categories with well documented types of behaviour. Within these categories the precise values of numerical constants determined by the model do not affect the qualitative behaviour. Rather it is the form of the equations which is important.

In the following sections we carry out a bifurcation theory analysis for the equations (1.21, 1.22). Assuming the matrix $\mathbf{A}$ has only one real critical eigenvalue, or a complex conjugate pair, we show that to cubic order centre manifold reduction does not alter the expansion of the dynamic equations, hence we need only carry out the reduction to normal form for the cases $p = 1$ and $p = 2$, since these will lead to the interesting and well understood *steady state* and *Hopf* bifurcations respectively. All other cases without further degeneracy in the critical eigenvalues will therefore to low orders at least, behave in a similar manner.

3.2.1 Linear Analysis

The first step to analysing the dynamics is to linearise the equations of motion (1.21, 1.22) about the trivial fixed point $\mathbf{m} = 0$:

$$\frac{d\mathbf{m}}{dt} = \beta\mathbf{A}\mathbf{m} - \mathbf{m} + \mathcal{O}(\mathbf{m}^3) \qquad \mathbf{m}_{t+1} = \beta\mathbf{A}\mathbf{m}_t + \mathcal{O}(\mathbf{m}^3).$$

A bifurcation from the trivial fixed point occurs when there is a loss of linear stability, this occurs as $\beta \to \beta_c + \delta\beta$, where $\delta\beta$ is an infinitesimal increment, and β_c is given by

$$\beta_c = \frac{1}{\max \Re(\alpha(\mathbf{A}))} \qquad \beta_c = \frac{1}{\max |\alpha(\mathbf{A})|} \tag{3.1}$$

for the continuous time and discrete time cases respectively, where $\alpha(\mathbf{A})$ are the eigenvalues of the matrix $\mathbf{A}$. **NB** due to the Hartman - Grobman theorem [75], these results give strong bounds on the region where the only long time solution is the trivial fixed point, however these are only true locally, i.e. when the solution is already close to the origin. Bounds for the global stability of the trivial fixed point will be derived in Section 3.3.

We can now classify the bifurcations according to the forms of the eigenvalues α. Bifurcations where the relevant eigenvalue α is a real number are known as *steady state* bifurcations. Bifurcations with a complex conjugate pair of eigenvalues are known as *Hopf* bifurcations. There is a further special case relevant to the discrete time case only, when the relevant eigenvalue is a negative number, these are known as *pitchfork* or *period doubling* bifurcations.

3.2.2 Centre Manifold Reduction

Here we derive the first few terms in the expansion of $\mathbf{h}(\mathbf{y})$ the map defining the centre manifold close to the critical point. We consider first the expansion of the dynamic equations (1.21,1.22)

$$\left.\begin{array}{c}\frac{dm^\nu}{dt}+m^\nu \\ m^\nu_{t+1}\end{array}\right\} = \langle \xi^\nu \tanh \beta \boldsymbol{\xi}\mathbf{A}\mathbf{m}_t\rangle \simeq \beta\left(\mathbf{A}\mathbf{m}_t\right)^\nu\left\{1-\frac{\beta^2}{3}\left[3\left(\mathbf{m}_t\mathbf{A}^\dagger\mathbf{A}\mathbf{m}_t\right)-2\left(\left(\mathbf{A}\mathbf{m}_t\right)^\nu\right)^2\right]\right\} + \mathcal{O}(\mathbf{m}^5).$$

Consider now a matrix $\mathbf{T}$ such that $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ has the block diagonal form $\begin{pmatrix}\mathbf{A}_1 & 0\\ 0 & \mathbf{A}_2\end{pmatrix}$, where $\beta\mathbf{A}_1$ has only critical eigenvalues - i.e. with real parts equal to one in the continuous time case, and modulus equal to one in the discrete time case. We label the number of critical eigenvalues n_c. We also re-label the co-ordinate $(\mathbf{T}^{-1}\mathbf{m}_t)^\nu = y_t^\nu$ if $\nu \le n_c$ and $(\mathbf{T}^{-1}\mathbf{m}_t)^\nu = z_t^\nu$ if $\nu > n_c$. Our equations of motion then become

$$\left.\begin{pmatrix}\frac{dy^\nu}{dt}+y^\nu\\ \frac{dz^\nu}{dt}+z^\nu\end{pmatrix}\right\} = \beta\begin{pmatrix}(\mathbf{A}_1)_{\nu\mu} & 0\\ 0 & (\mathbf{A}_2)_{\nu\mu}\end{pmatrix}\begin{pmatrix}y_t^\mu\\ z_t^\mu\end{pmatrix}\left\{1-\beta^2\left(\mathbf{y}_t\mathbf{A}_1^\dagger\mathbf{A}_1\mathbf{y}_t+\mathbf{z}_t\mathbf{A}_2^\dagger\mathbf{A}_2\mathbf{z}_t\right)\right\}$$
$$+\frac{2\beta^3}{3}\begin{pmatrix}T^{-1}_{\nu\eta}\left(T_{\eta\lambda}\left(\sum_{\rho\le n_c}(A_1)_{\lambda\rho}y_t^\rho+\sum_{\rho>n_c}(A_2)_{\lambda\rho}z_t^\rho\right)\right)^3\\ T^{-1}_{\nu\eta}\left(T_{\eta\lambda}\left(\sum_{\rho\le n_c}(A_1)_{\lambda\rho}y_t^\rho+\sum_{\rho>n_c}(A_2)_{\lambda\rho}z_t^\rho\right)\right)^3\end{pmatrix}+\mathcal{O}(y^5,z^5).$$

If we substitute $\mathbf{z}=\mathbf{h}(\mathbf{y})$ then equations (A.4, A.5) can be used to define the centre manifold. We propose a power series expansion $\mathbf{h}(\mathbf{y}) = \sum_{\mu,\nu}\boldsymbol{\Phi}_{\mu\nu}y^\nu y^\mu + \sum_{\mu,\nu,\lambda}\boldsymbol{\Theta}_{\mu\nu\lambda}y^\nu y^\mu y^\lambda + \mathcal{O}(\mathbf{y}^4)$, where the constant and linear terms are missing due to the boundary conditions $\mathbf{h}(0)=0$ and $\partial_\nu \mathbf{h}(0)=0\,\forall\,\nu$. By comparing terms in y^n $(n=2,3,\ldots)$ we obtain a set of simultaneous equations determining the expansion coefficients $\boldsymbol{\Phi},\boldsymbol{\Theta}$ etc. The lowest orders of which are, in the continuous time case

$$\mathcal{O}(y^2): \quad \sum_{\mu,\nu}[\boldsymbol{\Phi}_{\mu\nu}y^\nu((\beta\mathbf{A}_1-\mathbb{1})\mathbf{y})^\mu+\boldsymbol{\Phi}_{\mu\nu}y^\mu((\beta\mathbf{A}_1-\mathbb{1})\mathbf{y})^\nu] = \beta(\mathbf{A}_2-\mathbb{1})\sum_{\mu,\nu}\boldsymbol{\Phi}_{\mu\nu}y_t^\mu y_t^\nu$$
$$\mathcal{O}(y^3): \quad \sum_{\mu,\nu,\lambda}\left[\Theta^\rho_{\mu\nu\lambda}y^\nu y^\lambda((\beta\mathbf{A}_1-\mathbb{1})\mathbf{y})^\mu+\Theta^\rho_{\mu\nu\lambda}y^\mu y^\lambda((\beta\mathbf{A}_1-\mathbb{1})\mathbf{y})^\nu+\Theta^\rho_{\mu\nu\lambda}y^\nu y^\mu((\beta\mathbf{A}_1-\mathbb{1})\mathbf{y})^\lambda\right]$$
$$= \frac{2\beta^3}{3}T^\dagger_{\rho\nu}\left(T_{\nu\eta}A^1_{\eta\lambda}y_t^\lambda\right)^3 \qquad \rho>n_c,$$

and in the discrete time case

$$\mathcal{O}(y^2): \qquad \beta^2\sum_{\mu,\nu}\boldsymbol{\Phi}_{\mu\nu}\left(\mathbf{A}_1\mathbf{y}_t\right)^\mu\left(\mathbf{A}_1\mathbf{y}_t\right)^\nu = \beta\mathbf{A}_2\sum_{\mu,\nu}\boldsymbol{\Phi}_{\mu\nu}y_t^\mu y_t^\nu$$
$$\mathcal{O}(y^3): \qquad \beta^3\sum_{\mu,\nu,\lambda}\Theta^\rho_{\mu\nu\lambda}\left(\mathbf{A}_1\mathbf{y}_t\right)^\mu\left(\mathbf{A}_1\mathbf{y}_t\right)^\nu\left(\mathbf{A}_1\mathbf{y}_t\right)^\lambda = \frac{2\beta^3}{3}T^\dagger_{\rho\nu}\left(T_{\nu\eta}A^1_{\eta\lambda}y_t^\lambda\right)^3 \qquad \rho>n_c.$$

Because the lowest order terms in $\mathbf{h}(\mathbf{y})$ are of order $\mathbf{y}^2$ we can see that to order $\mathbf{y}^3$ the equations of motion on the centre manifold are simply given by

$$\left.\begin{array}{c}\frac{dy^\nu}{dt}+y^\nu\\ y^\nu_{t+1}\end{array}\right\} = \beta(A_1)_{\nu\mu}y_t^\mu\left\{1-\beta^2\mathbf{y}_t\mathbf{A}_1^\dagger\mathbf{A}_1\mathbf{y}_t\right\}+\frac{2\beta^3}{3}T^{-1}_{\nu\eta}\left(T_{\eta\lambda}\left(\sum_{\rho\le n_c}(A_1)_{\lambda\rho}y_t^\rho\right)\right)^3+\mathcal{O}(\mathbf{y}^4); \tag{3.2}$$

i.e. to cubic order the forms of the equations of motion on the centre manifold are the same form as the full equations of motion. Therefore we carry out all the following stages of analysis for the $p = 1$ and $p = 2$ cases only, since the equations of motion for $p > 2$ displaying steady state and Hopf bifurcations respectively will be the same form as these up to cubic order, with slightly different numerical constants.

3.2.3 Normal Form for One Pattern

For a network trained with one pattern, the embedding matrix **A** is simply a number and hence symmetric, the Lyapunov functions of Chapter 2 therefore apply, and approach to equilibrium is guaranteed. We can investigate the behaviour, however, using bifurcation theory to discover how the fixed point of the dynamics deviates from the trivial fixed point. This will also serve as a useful introduction to the methods which will be applied to networks trained with more patterns. First the equations of motion are expanded in powers of m.

$$\frac{dm}{dt} \simeq (\beta A - 1)m - \frac{(\beta A m)^3}{3} + \ldots \qquad m_{t+1} \simeq \beta A m_t - \frac{(\beta A m_t)^3}{3} + \ldots$$

We can see that there will be a steady state bifurcation when $\beta = \beta_c = \frac{1}{A}$ for the continuous time case, and $\beta = \beta_c = \frac{1}{|A|}$ for the discrete time case.

Since the equations are one dimensional, we do not have to perform the centre manifold reduction, and can at once move on to the normal form reduction. We now apply a series of non-linear co-ordinate changes $m \to m + \phi^k(m)$ where $\phi^k(m) = a_k m^k$. The only terms which cannot be removed are those in the kernel (i.e. those with a zero eigenvalue) of the functions $L(\phi^k(m))$, $\tilde{L}(\phi^k(m))$ (see Appendix A). Applying (A.7) we see that

$$L(\phi^k(m)) = (\beta A - 1)a_k(1-k)m^k \qquad \tilde{L}(\phi^k(m)) = \beta A a_k m^k(1 - (\beta A)^{k-1}),$$

therefore $L(\phi^k(m)) = 0$ requires k=1 or $\beta A = 1$, and $\tilde{L}(\phi^k(m)) = 0$ requires $|\beta A|$=1, which are exactly the conditions satisfied at criticality. Hence no non-linear terms can be removed. To proceed further we expand the equations of motion ($\frac{dm}{dt}$,$m_{t+1} - m_t = V(\beta, m)$) about the critical point $m = 0$, $\beta_c = \frac{1}{A}$, or $\beta_c = \frac{1}{|A|}$ for the discrete time case, subject to the conditions at criticality $V(\beta_c, 0) = \frac{d}{dm}V(\beta_c, 0) = 0$.

$$\left.\begin{matrix}\frac{dm}{dt} \\ m_{t+1} - m_t\end{matrix}\right\} = V(\beta, m) \simeq (\beta - \beta_c)\frac{\partial V}{\partial \beta}(\beta_c, 0) + m\frac{\partial V}{\partial m}(\beta_c, 0) + \frac{m^2}{2}\frac{\partial^2 V}{\partial m^2}(\beta_c, 0)$$
$$+ m(\beta - \beta_c)\frac{\partial^2 V}{\partial m \partial \beta}(\beta_c, 0) + \frac{(\beta - \beta_c)^2}{2}\frac{\partial^2 V}{\partial \beta^2}(\beta_c, 0) + \ldots.$$

We note that the first two non-zero derivatives are $\frac{\partial^2 V(\beta_c,0)}{\partial m \partial \beta} = A$ and $\frac{\partial^3 V(\beta_c,0)}{\partial m^3} = -2$. Therefore to leading non vanishing orders, our equations of motion close to criticality (the normal forms) for $A > 0$ are

$$\left.\begin{matrix}\frac{dm}{dt} \\ m_{t+1} - m_t\end{matrix}\right\} \simeq m_t\left(\left(\frac{\beta}{\beta_c} - 1\right) - \frac{1}{3}m_t^2\right). \tag{3.3}$$

Normal forms such as these give rise to pitchfork bifurcations. We can see that fixed points of the motion are $m_t = 0$ and $m_t = \pm\sqrt{3(\frac{\beta}{\beta_c} - 1)}$, and that there stabilities are given by

$$\frac{\beta}{\beta_c} - 1 - m^2 = \begin{cases} \frac{\beta}{\beta_c} - 1 & \text{for } m_t = 0 \\ -2(\frac{\beta}{\beta_c} - 1) & \text{for } m_t = \pm\sqrt{3(\frac{\beta}{\beta_c} - 1)} \end{cases}.$$

Hence $m_t = 0$ is stable for $\frac{\beta}{\beta_c} < 1$, and $m_t = \pm\sqrt{3(\frac{\beta}{\beta_c} - 1)}$ is stable for $\frac{\beta}{\beta_c} > 1$.

Interesting things happen when $A < 0$ in the discrete time case, since then $\beta_c = \frac{\mathrm{sgn}[A]}{A}$ so the normal form is

$$m_{t+1} - m_t \simeq m_t \left(\mathrm{sgn}[A] \left(\frac{\beta}{\beta_c} - 1 \right) - \frac{1}{3} m_t^2 \right) \tag{3.4}$$

hence for $A < 0$ the system will settle into a two-cycle, oscillating between the two stable fixed points encountered for $A > 0$. For this reason a *Pitchfork* bifurcation is often known as a *period doubling* bifurcation. An infinite cascade of period doubling bifurcations as successive iterates of the quadratic map lose stability is responsible for the transition to chaos investigated in [76, 77]. This however will not occur in this system since the two cycle is always stable above the critical point.

The analysis for networks trained with more than one pattern displaying steady state bifurcations will obviously be more complicated since the process of centre manifold reduction must first be carried out. However similar behaviour will occur and the equations will be of the same general form with different numerical constants, since up to cubic order the centre manifold reduction adds no new terms. This is of course dependent on the relevant eigenvalue of the matrix $\mathbf{A}$ (the one with the largest modulus) being a real number. If the relevant eigenvalue is a complex conjugate pair, the normal form is two-dimensional, and the system will display a Hopf bifurcation.

Figure 3.1 shows steady state bifurcations described by equation (3.3) for two matrices:

$$\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \alpha = \begin{cases} 3 \\ 1+i \\ 1-i \end{cases} \qquad |\alpha| = \begin{cases} 3 \\ \sqrt{2} \\ \sqrt{2} \end{cases} .$$

for which we will see a steady state bifurcation to a fixed point at $\beta = \frac{1}{3}$, see Figure 3.1; and

$$\mathbf{A} = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \alpha = \begin{cases} -3 \\ 1+i \\ 1-i \end{cases} \qquad |\alpha| = \begin{cases} 3 \\ \sqrt{2} \\ \sqrt{2} \end{cases} ,$$

for which we will see a steady state bifurcation to a two-cycle at $\beta = \frac{1}{3}$, see Figure 3.1.

The results were obtained by iterating the dynamics for 1000 steps before recording subsequent points on the trajectory, until the system returned to the first recorded point.

3.2.4 Normal Form for Two Patterns

For networks trained with only two patterns there are four free parameters, in the embedding matrix. We can write the embedding matrix in terms of the Pauli matrices and the unit matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.5}$$

The free parameters are now α, γ, δ and β (which provides an overall scaling). The eigenvalues of these matrices are $\pm i$, 1, ± 1 and ± 1 respectively. The eigenvalues of $\mathbf{A}$ are given by

$$\omega = 1 \pm \sqrt{\gamma^2 + \delta^2 - \alpha^2}$$

We can see that the eigenvalues are:

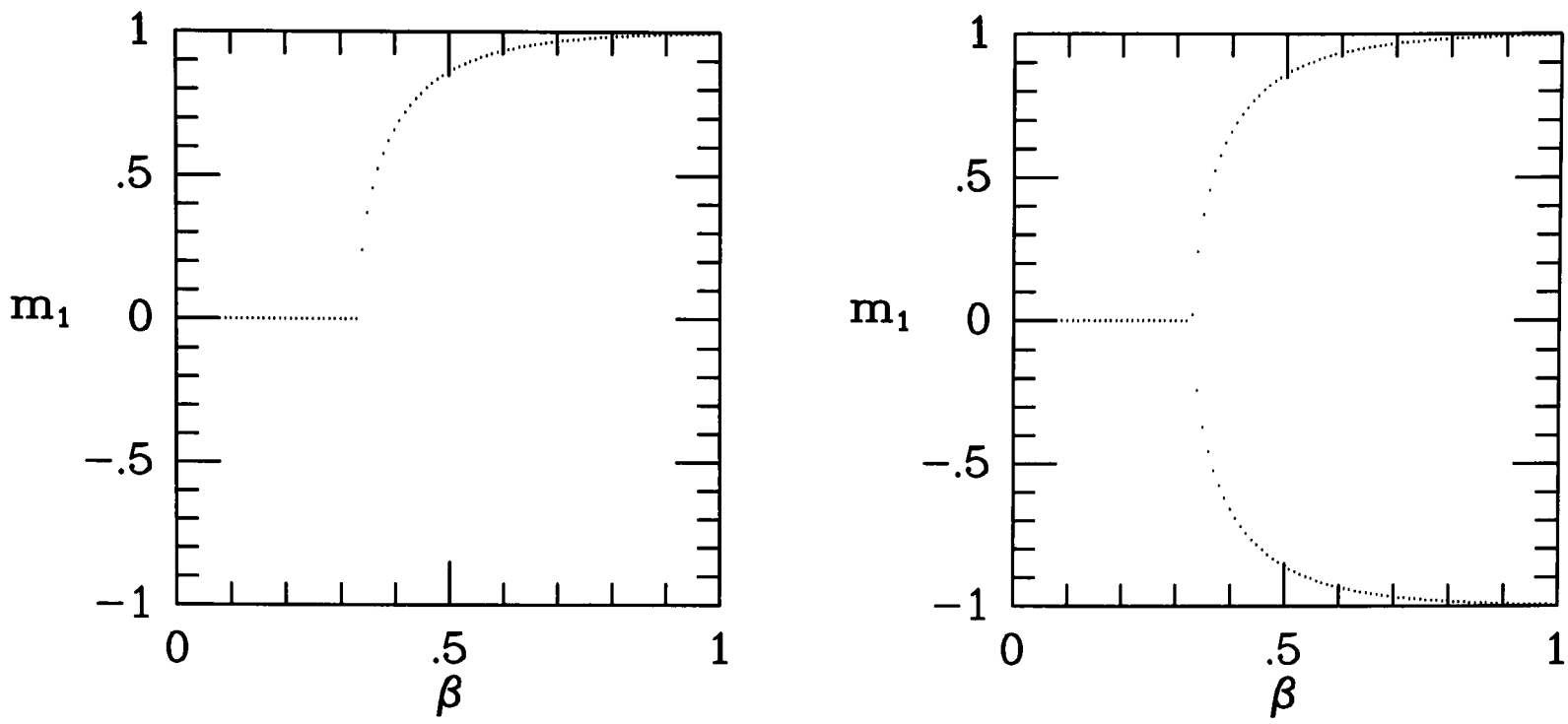


Figure 3.1: Steady state bifurcation diagrams as a function of β for p = 3. The relevant eigenvalue of the embedding matrix is real and positive for the left hand diagram whilst real and negative for the right hand diagram

1. Pure real if $\gamma^2 + \delta^2 \geq \alpha^2$.

2. Real and imaginary if $\gamma^2 + \delta^2 < \alpha^2$.

We can therefore classify the bifurcation behaviour accordingly.

1. Steady state bifurcations.

2. Hopf bifurcations; with resonances when $1 \pm \sqrt{\gamma^2 + \delta^2 - \alpha^2}$ is a root of unity.

The case of steady state bifurcations has already been discussed under one pattern.

For cases where the relevant eigenvalues at criticality form a complex conjugate pair the system undergoes a Hopf bifurcation. In this case the behaviour is slightly different for the continuous time case and the discrete time case. In the discrete time case if the eigenvalues exhibit resonance, i.e. if there exists an integer n such that $\lambda_c^n = 1$, then terms are introduced into the normal form which cause effects such as mode locking. Non-linear effects cause the system to lock into one frequency over a range of the external driving frequencies. This behaviour of resonant systems can be understood, since if n exists then the dynamics is effectively undergoing a steady state bifurcation in $m_{t+n} = F(m_t)$, hence the system will display a period n cycle. As parameters are varied the system can pass through several resonances, causing the devils staircase structure.

Consider a matrix $\mathbf{A}$ given by (3.5) with $\gamma = \delta = 0$. This has conjugate eigenvalues $1 \pm i\alpha$ and there will be a Hopf bifurcation at $\beta_c = 1$ in the continuous time case, and $\beta_c = \frac{1}{\sqrt{\alpha^2+1}}$ in the discrete time case. We proceed with the bifurcation analysis by expanding the dynamical equations (1.21, 1.22), keeping only the first non-linear term since higher order terms will be negligible at the critical point, and since we can always consider higher order terms later as perturbations. The eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix}$ are given by

$$\lambda_+ = 1 + i\alpha \quad z = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \qquad \text{and} \qquad \lambda_- = 1 - i\alpha \quad \bar{z} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

We now define a matrix $\mathbf{U}$ which diagonalises $\mathbf{A}$, and define new variables $\mathbf{x}$

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \mathbf{U}^\dagger \mathbf{m},$$

so that equation (3.2) for this matrix becomes

$$\begin{aligned} \left.\begin{matrix} \mathbf{x}_{t+1} \\ \frac{d\mathbf{x}}{dt} + \mathbf{x} \end{matrix}\right\} &= \beta \begin{pmatrix} 1+i\alpha & 0 \\ 0 & 1-i\alpha \end{pmatrix} \mathbf{x}_t \left[1 - \beta^2(1+\alpha^2)\mathbf{x}^2\right] + \frac{2\beta^3}{3}\mathbf{U} \begin{pmatrix} (\mathbf{AUx})_1^3 \\ (\mathbf{AUx})_2^3 \end{pmatrix} \\ &= \beta \begin{pmatrix} 1+i\alpha & 0 \\ 0 & 1-i\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &\quad + \frac{\beta^3}{6} x_1^3 \begin{pmatrix} (-1-7i)\,\alpha^3 + (-9+3i)\,\alpha^2 + (3-3i)\,\alpha + (-5-i) \\ (1+i)\,\alpha^3 + (3-3i)\,\alpha^2 + (-3-3i)\,\alpha + (-1+i) \end{pmatrix} \\ &\quad + \frac{\beta^3}{6} x_1^2 x_2 \begin{pmatrix} (-3+3i)\,\alpha^3 + (3+3i)\,\alpha^2 + (-3+3i)\,\alpha + (3+3i) \\ (-3+9i)\,\alpha^3 + (-3+3i)\,\alpha^2 + (-3+9i)\,\alpha + (-3+3i) \end{pmatrix} \\ &\quad + \frac{\beta^3}{6} x_1 x_2^2 \begin{pmatrix} (-3-9i)\,\alpha^3 + (-3-3i)\,\alpha^2 + (-3-9i)\,\alpha + (-3-3i) \\ (3+3i)\,\alpha^3 + (-3+3i)\,\alpha^2 + (3+3i)\,\alpha + (-3+3i) \end{pmatrix} \\ &\quad + \frac{\beta^3}{6} x_2^3 \begin{pmatrix} (-1+i)\,\alpha^3 + (-3-3i)\,\alpha^2 + (3-3i)\,\alpha + (1+i) \\ (-1+7i)\,\alpha^3 + (-9-3i)\,\alpha^2 + (3+3i)\,\alpha + (-5+i) \end{pmatrix} \\ &\quad + \mathcal{O}(\mathbf{x}^5). \end{aligned} \tag{3.6}$$

We now apply non-linear co-ordinate changes $x \to x + \phi(x)$ to reduce the equations to normal form. In order to find the non-removable terms we need to find the zero eigenvalues of $L(\phi^k(\mathbf{m}))$, $\tilde{L}(\phi^k(\mathbf{m}))$. First, transforming co-ordinates $\phi^k(\mathbf{x}) = \mathbf{U}^\dagger \phi^k(\mathbf{m})$, then

$$\begin{aligned} L(\phi^k(\mathbf{x})) &= \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \begin{pmatrix} \phi_1^k(\mathbf{x}) \\ \phi_2^k(\mathbf{x}) \end{pmatrix} - \begin{pmatrix} \frac{\partial \phi_1^k(\mathbf{x})}{\partial x_1} & \frac{\partial \phi_1^k(\mathbf{x})}{\partial x_2} \\ \frac{\partial \phi_2^k(\mathbf{x})}{\partial x_1} & \frac{\partial \phi_2^k(\mathbf{x})}{\partial x_2} \end{pmatrix} \begin{pmatrix} \lambda_+ x_1 \\ \lambda_- x_2 \end{pmatrix} \\ \tilde{L}(\phi^k(\mathbf{x})) &= \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \begin{pmatrix} \phi_1^k(x_1, x_2) \\ \phi_2^k(x_1, x_2) \end{pmatrix} - \begin{pmatrix} \phi_1^k(\lambda_+ x_1, \lambda_- x_2) \\ \phi_2^k(\lambda_+ x_1, \lambda_- x_2) \end{pmatrix}. \end{aligned}$$

We can see that suitable basis vectors are

$$\xi_+^{k,l}(x_1, x_2) = \begin{pmatrix} x_1^l x_2^{k-l} \\ 0 \end{pmatrix} \qquad \text{and} \qquad \xi_-^{k,l}(x_1, x_2) = \begin{pmatrix} 0 \\ x_1^l x_2^{k-l} \end{pmatrix} \qquad \text{where} \quad l = 0, 1, 2, \ldots, k,$$

and that the corresponding eigenvalues of $L(\phi(\mathbf{x}))$, $\tilde{L}(\phi(\mathbf{x}))$, with $\lambda_\pm = \gamma \pm i\omega$ for the continuous time case and $\lambda_\pm = ae^{\pm i2\pi\psi b}$ for the discrete time case are

$$\alpha_\pm^{k,l} = (1-k)\gamma + i\omega(k - 2l \pm 1) \qquad \text{and} \qquad \tilde{\alpha}_\pm^{k,l} = ae^{\pm i2\pi\psi b}(1 - a^{k-1}e^{-i2\pi\psi b(k-2l\pm1)}).$$

At the critical point $\gamma = 0$ and $a = b = 1$. Zero eigenvalues in the continuous time case therefore requires $k = 2l \pm 1$. i.e. k must be odd, and the eigenvectors for the irreducible part of the equations of motion are

$$\xi_+^k(\mathbf{x}) = \begin{pmatrix} x_1 |x_1\, x_2|^{\frac{k-1}{2}} \\ 0 \end{pmatrix} \qquad \text{and} \qquad \xi_-^k(\mathbf{x}) = \begin{pmatrix} 0 \\ x_2 |x_1\, x_2|^{\frac{k-1}{2}} \end{pmatrix} \qquad \text{where} \quad k = 3, 5, \ldots.$$

The normal form is therefore

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \gamma + i\omega & 0 \\ 0 & \gamma - i\omega \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \sum_{j=1}^{\infty} |x_1\, x_2|^j \begin{pmatrix} a_j x_1 \\ \bar{a}_j x_2 \end{pmatrix},$$

where a_j, $\bar{a}_j$ are the coefficients in the equations of motion. If we write $x_{1,2} = r\, e^{\mp i\theta}$ then $\frac{\dot{x}_{1,2}}{x_{1,2}} = \frac{\dot{r}}{r} \mp i\dot{\theta}$ so in polar co-ordinates the equations of motion become

$$\dot{r} = r\left(\gamma + \sum_{j=1}^{\infty} \Re(a_j) r^{2j}\right) \qquad \text{and} \qquad \dot{\theta} = -\omega - \sum_{j=1}^{\infty} \Im(a_j) r^{2j}.$$

In our case $\gamma = \beta - 1$, $\omega = \alpha\beta$ and $a_1 = \frac{\beta^3}{2}(\alpha^2+1)((1-\alpha) + i(1+\alpha))$, so our equations of motion close to the critical point $\beta_c = 1$ are

$$\begin{aligned} \dot{r} &= (\frac{\beta}{\beta_c} - 1)r + \frac{\beta^3}{2}(\alpha^2+1)(1-\alpha)r^3 + \mathcal{O}(r^5) \\ \dot{\theta} &= -\alpha\beta - \frac{\beta^3}{2}(\alpha^2+1)(1+\alpha)r^2 + \mathcal{O}(r^4). \end{aligned} \tag{3.7}$$

Notice that both equations are independent of θ. Close to the critical point where higher order terms can be omitted, the system will settle into a cycle of fixed radius (the Hopf radius). For $\alpha > 1$ this is given by $r_H = \sqrt{\frac{2(\frac{\beta}{\beta_c}-1)}{\beta^3(\alpha^2+1)(\alpha-1)}}$; with angular velocity approximately $\dot{\theta} = -\alpha\beta - (\frac{\beta}{\beta_c}-1)\frac{(1+\alpha)}{(1-\alpha)}$. For $\alpha < 1$ higher order terms are needed to determine the Hopf radius.

In the discrete time case things are a little more complicated, since the conditions for zero eigenvalues are that $\psi(k - 2l \pm 1) = n$ where n is an integer. If ψ is irrational this can only occur for $k - 2l \pm 1 = 0$, however if ψ is rational $= \frac{p}{q}$ then we also have solutions $k - 2l \pm 1 = nq$. If this is satisfied for some value s of n then the basis vectors for removing the s^{th} terms are

$$\xi^s_+ = \begin{pmatrix} x_2^{s-1} \\ 0 \end{pmatrix} \qquad \text{and} \qquad \xi^s_- = \begin{pmatrix} 0 \\ x_1^{s-1} \end{pmatrix}.$$

Therefore the normal form is

$$\begin{pmatrix} x_1(t+1) \\ x_2(t+1) \end{pmatrix} = \begin{pmatrix} ae^{i2\pi\psi b} & 0 \\ 0 & ae^{-i2\pi\psi b} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \sum_{j=1}^{\infty} |x_1\, x_2|^j \begin{pmatrix} a_j x_1 \\ \bar{a}_j x_2 \end{pmatrix} + \sum_{k=2}^{\infty} \begin{pmatrix} a^*_k x_2^k \\ \bar{a}^*_k x_1^k \end{pmatrix},$$

where a_j, $\bar{a}_j$, a^*_k and $\bar{a}^*_k$ are the coefficients for the non-resonant, and resonant terms respectively. In terms of polar co-ordinates the equations of motion are

$$\begin{aligned} r_{t+1} &= r_t \left| a + \sum_{j=1}^{\infty} a_j r_t^{2j} + \sum_{k=3}^{\infty} a^*_k r_t^{k-1} e^{ik\theta_t} \right| \\ \theta_{t+1} - \theta_t &= -2\pi\psi b - \arg\left(a + \sum_{j=1}^{\infty} a_j e^{-i2\pi\psi b} r_t^{2j} + \sum_{k=2}^{\infty} a^*_k r_t^k e^{-i(k+1)\theta_t} \right). \end{aligned}$$

For our case $a = \beta\sqrt{\alpha^2+1}$, $2\pi\psi b = \arctan(\alpha)$ and $a_1 = \frac{\beta^3}{2}(\alpha^2+1)((1-\alpha) + i(1+\alpha))$, giving equations of motion close to the critical point $\beta_c = \frac{1}{\sqrt{\alpha^2+1}}$ in the absence of strong resonances

$$\begin{aligned} r_{t+1} &= r_t \left| \frac{\beta}{\beta_c} - \frac{\beta^3}{2}(\alpha^2+1)((1-\alpha) + i(\alpha+1))r_t^2 + \mathcal{O}(r^4) \right| \\ \theta_{t+1} - \theta_t &= -\arctan(\alpha) - \arg\left(1 + \frac{\beta^2}{2}(\alpha^2+1)((1-\alpha) + i(\alpha+1))e^{-i\arctan(\alpha)} r_t^2 + \mathcal{O}(r^4) \right). \end{aligned} \tag{3.8}$$

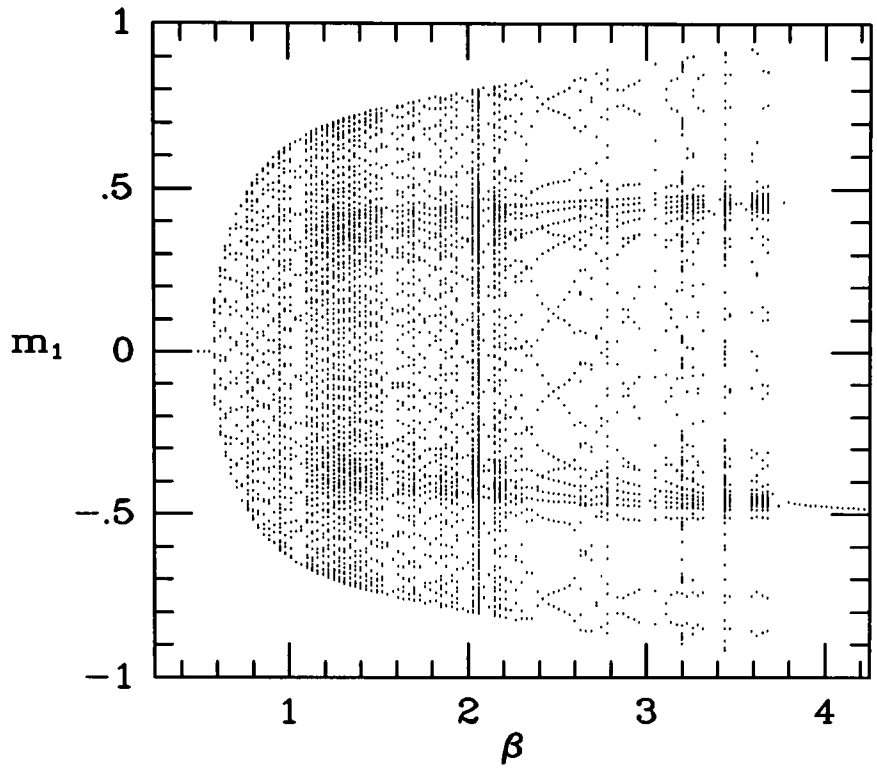


Figure 3.2: Hopf bifurcation diagram as a function of β for p = 3

This predicts a steady state Hopf radius, close to the critical point where higher order terms can be neglected given by $r_t^2 = \frac{-(\alpha-1)\pm\sqrt{\frac{2}{\beta^2}-(\alpha+1)^2}}{(\alpha^2+1)^{\frac{3}{2}}}$.

Figure 3.2 shows a bifurcation diagram for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad \alpha = \begin{Bmatrix} 0.324718 \\ 1.66236 + 0.56228i \\ 1.66236 - 0.56228i \end{Bmatrix} \qquad |\alpha| = \begin{Bmatrix} 0.324718 \\ 1.75488 \\ 1.75488 \end{Bmatrix} . \tag{3.9}$$

Here we expect to see a Hopf bifurcation to quasi-periodic behavior at $\beta = \frac{1}{1.75488} = 0.56984$.

In Figure 3.2 we clearly see the bifurcation from trivial fixed point to quasi-periodic behaviour, for a network with the embedding matrix (3.9). The period is obviously dependent on β, as illustrated by the variation in density of points, and the periodic windows are clearly visible. This behaviour is described by equations (3.8).

The existence of resonances will cause non-linear θ_t terms in the arg(...) term, causing phenomena such as mode locking, and could lead to a transition to chaos, if the θ_t map loses invertibility. These equations are strictly only valid in a small interval around the critical point. As we move away from the critical point, the resonances, initially points on the unit circle will grow giving Arnold tongues [78]. It is this effect that produces mode locking.

3.3 Exact Results

Analytical results for the equations (1.21, 1.22) are difficult, if not impossible, to obtain due to the non-linearity of the mapping and the lack of symmetries in the matrix **A**. Some results have been found for specific choices of the matrix **A** i.e. Chapter 2 and [79]. It is possible, however, to derive some bounds on β and the matrix **A**, for which we can predict the behaviour. For $p = 1$ or 2 some exact results, valid for all β and **A**, can be found.

3.3.1 Bounds for Low and High β

It is clear from equations (1.21, 1.22) that in the limit $\beta \to 0$, the network will always settle into the trivial fixed point $\mathbf{m} = 0$. We wish to derive some rigorous bound for β, dependent on the properties of the matrix $\mathbf{A}$, such that the network will always settle into the trivial fixed point.

Consider two vectors $\mathbf{m}$ and $\mathbf{k}$ related by the relation $\mathbf{m} = \langle \boldsymbol{\xi} \tanh(\beta \boldsymbol{\xi} \cdot \mathbf{k}) \rangle_{\boldsymbol{\xi}}$. Since tanh is an odd function and the ξ's are independent, unbiased random variables with values ± 1, we can write

$$m^\mu = \langle \tanh(\beta k^\mu + \beta \sum_{\rho \neq \mu} \xi^\rho k^\rho) \rangle.$$

Now $\frac{\partial m^\mu}{\partial k^\mu} \geq 0$, and if $k^\mu = 0$ then $m^\mu = 0$ (the $\boldsymbol{\xi}$'s are randomly distributed over $\{-1, +1\}$). From this we can see that

1. $\mathrm{sgn}[m^\mu] = \mathrm{sgn}[k^\mu]$,
2. $|m^\mu| = \langle \tanh(\beta |k^\mu| + \beta \sum_{\rho \neq \mu} \xi^\rho k^\rho) \rangle$.

Since the ξ's are randomly ± 1, we can replace ξ in any expression with $-\xi$. Therefore

$$|m^\mu| = \frac{1}{2} \langle \tanh \beta |k^\mu| + \beta \sum_{\rho \neq \mu} \xi^\rho k^\rho \rangle + \frac{1}{2} \langle \tanh \beta |k^\mu| - \beta \sum_{\rho \neq \mu} \xi^\rho k^\rho \rangle.$$

This is a function of the general form $f(|x| + y) + f(|x| - y)$, which equals $f(|x| + |y|) + f(|x| - |y|)$. Therefore

$$\left\langle \tanh \left(\beta |k^\mu| + \beta \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \right) \right\rangle$$
$$= \tanh \beta |k^\mu| + \left\langle \int_0^1 d\lambda \, \beta \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \left\{ 1 - \tanh^2 \left(\beta |k^\mu| + \beta \lambda \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \right) \right\} \right\rangle,$$

so that

$$|m^\mu| = \tanh \beta |k^\mu|$$
$$+ \left\langle \int_0^1 d\lambda \, \beta \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \left\{ \tanh^2 \left(\beta |k^\mu| - \beta \lambda \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \right) - \tanh^2 \left(\beta |k^\mu| + \beta \lambda \left| \sum_{\rho \neq \mu} \xi^\rho k^\rho \right| \right) \right\} \right\rangle.$$

The second term is always positive, hence $|m^\mu| \leq \tanh(\beta |k^\mu|) \leq \beta |k^\mu|$ and $\mathbf{m}^2 \leq \beta^2 \mathbf{k}^2$.

Now if we set $\mathbf{k} = \mathbf{A}\mathbf{m}$, and $\mathbf{m} \to \frac{d\mathbf{m}}{dt} + \mathbf{m}$, where $\mathbf{m}$ solves the continuous time dynamics (1.21) then we have $\left(\mathbf{m} + \frac{d\mathbf{m}}{dt}\right)^2 = \mathbf{m}^2 + \frac{d\mathbf{m}^2}{dt} + \left(\frac{d\mathbf{m}}{dt}\right)^2 \leq \beta^2 \mathbf{m} \mathbf{A}^\dagger \mathbf{A} \mathbf{m}$ so

$$\frac{d\mathbf{m}^2}{dt} \leq \beta^2 \mathbf{m} \mathbf{A}^\dagger \mathbf{A} \mathbf{m} - \mathbf{m}^2 \tag{3.10}$$

Hence for $\beta^2 < \frac{1}{\alpha_{max}(\mathbf{A}^\dagger \mathbf{A})}$ then $\frac{d\mathbf{m}^2}{dt} < 0$ and as a result $\lim_{t \to \infty} m = 0$ (where $\alpha_{max}(\mathbf{A}^\dagger \mathbf{A})$ is the largest eigenvalue of $\mathbf{A}^\dagger \mathbf{A}$).

Similarly in the discrete time case (1.22) we choose $\mathbf{m} \to \mathbf{m}_{t+1}$ and $\mathbf{k} = \mathbf{A}\mathbf{m}_t$, so

$$\frac{\mathbf{m}_{t+1}^2}{\mathbf{m}_t^2} \leq \beta^2 \frac{\mathbf{m}_t (\mathbf{A}^\dagger \mathbf{A}) \mathbf{m}_t}{\mathbf{m}_t^2} \quad \forall \mathbf{m}_t \tag{3.11}$$

Therefore for $\beta^2 < \frac{1}{\alpha_{max}(\mathbf{A}^\dagger\mathbf{A})}$ then $\frac{\mathbf{m}_{t+1}^2}{\mathbf{m}_t^2} < 1$ hence $\lim_{t\to\infty} \mathbf{m}_t = 0$.

Notice these are weaker bounds than the ones given by considering the linearisation of the equations of motion. They are however global bounds, rather than bounds only valid in the vicinity of the origin.

In the limit $\beta \to \infty$ we also expect to be able to tell something about the dynamics since

$$\lim_{\beta\to\infty} \langle \boldsymbol{\xi} \tanh(\beta\boldsymbol{\xi}\mathbf{A}\mathbf{m}) \rangle = \frac{1}{2^p} \sum_{\boldsymbol{\xi}=\{\pm1\}^p} \boldsymbol{\xi} \operatorname{sgn}(\boldsymbol{\xi}\mathbf{A}\mathbf{m}) \tag{3.12}$$

In the continuous time case this leads to sharp corners in the trajectories, when $\boldsymbol{\xi}\mathbf{A}\mathbf{m}$ changes sign. In the discrete time case, since $\operatorname{sgn}(\boldsymbol{\xi}\mathbf{A}\mathbf{m}_t)$ has only two possible values ±1, and $\boldsymbol{\xi}$ has 2^p possibilities, we expect cycles of maximum period 2^p. However, the sgn function is the limit of a tanh, so there it has a third possible value 0, when $\boldsymbol{\xi}\mathbf{A}\mathbf{m}_t = 0$. This admits a further 2^p possibilities, therefore in the limit $\beta \to \infty$ we expect a limit cycle with maximum period 2^{p+1}.

3.3.2 One Pattern

The dynamics of a network trained with only one pattern (i.e. p=1) is given by one autonomous non-linear differential, or difference equation

$$\frac{dm}{dt} = \tanh(\beta A m) - m \qquad \text{and} \qquad m_{t+1} = \tanh(\beta A m_t).$$

This is a special case of the general equations (1.21, 1.22), for which $\mathbf{A}$ is simply a number, therefore it is necessarily symmetric, and provides only an extra scaling for β. The Lyapunov functions in Chapter 2 apply, and the system is guaranteed to reach an equilibrium configuration. Steady state bifurcations will occur from the trivial fixed point to non-trivial fixed points in the continuous time case, and to fixed points or period two cycles (for $A < 0$) in the case of discrete time dynamics.

3.3.3 Two Patterns

A network trained with two patterns (i.e. p=2) is the simplest we expect to display interesting behaviour. It can be analysed directly, since there are a reasonably small number of free parameters, and from these results we hope to infer general statements about networks with larger numbers of patterns.

The network may display steady state bifurcations if the relevant eigenvalue of $\mathbf{A}$ is real or Hopf bifurcations to complex cyclical behaviour if the relevant eigenvalues are a complex conjugate pair, as discussed in the previous section. It is easier to consider the order parameters $\mathbf{m}$ in terms of polar co-ordinates r, θ since this is the natural co-ordinate frame for the Hopf bifurcation. The non-linear θ terms in the mapping are going to give the interesting behaviour such as mode locking.

We consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix}$ since this is guaranteed to have a complex conjugate pair of eigenvalues, though the extension of the following work to general $\mathbf{A}$ is not difficult. $\mathbf{A}$ has

conjugate eigenvalues $1 \pm i\alpha$ and there will be a Hopf bifurcation from the trivial fixed point at $\beta_c = 1$ in the continuous time case, and $\beta_c = \frac{1}{\sqrt{\alpha^2+1}}$ in the discrete time case. We will only consider the (discrete time) difference equation (1.22), since the differential equation cannot display the property of mode locking. The appropriate map is

$$\mathbf{m}_{t+1} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tanh\beta\left[(1-\alpha)m_t^1 + (\alpha+1)m_t^2\right] + \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tanh\beta\left[(\alpha+1)m_t^1 + (\alpha-1)m_t^2\right]. \tag{3.13}$$

We now change to polar co-ordinates, the natural co-ordinate frame in which to consider Hopf bifurcations. Putting $(m^1, m^2) = r(\cos\theta, \sin\theta)$ gives

$$\begin{aligned} r_{t+1}^2 &= \frac{1}{2}\tanh^2 a_t r_t + \frac{1}{2}\tanh^2 b_t r_t \\ \theta_{t+1} &= \arctan\left(\frac{\tanh a_t r_t - \tanh b_t r_t}{\tanh a_t r_t + \tanh b_t r_t}\right), \end{aligned} \tag{3.14}$$

with the abbreviations $a_t = \beta((1-\alpha)\cos\theta_t + (\alpha+1)\sin\theta_t)$ and $b_t = \beta((\alpha+1)\cos\theta_t + (\alpha-1)\sin\theta_t)$.

We are only interested in the non-transient behaviour, and the dependence of r_{t+1} and θ_{t+1} on r_t and θ_t in the long time limit. For non-chaotic behaviour we expect the system to settle into a trajectory given by an invariant curve, i.e. any point on the curve is mapped to another point on the curve. Since a point on a 2-d curve can be fully parametrised by one co-ordinate we need only study the angular dependence (θ_t). We choose θ_t as our co-ordinate since it uniquely describes a point on the invariant curve, whereas r_t may not if there are symmetries present. It is the non-linear θ_t dependence that causes mode locking.

Close to the bifurcation point we expect the radius of the curve to be small, if we expand r_{t+1}^2 in terms of r_t we find

$$r_{t+1}^2 \simeq (a^2+b^2)r_t^2 - \frac{4}{3}(a^4+b^4)r_t^4 + \mathcal{O}(r_t^6).$$

Now, $a^2 + b^2$ is independent of θ_t, so angular dependence only enters through the second, and higher terms. As we move away from the bifurcation point, and higher terms need to be taken into account, the angular dependence will become stronger, and higher order θ_t terms will appear. This will have the effect of making the trajectory more convoluted and less circular as we move away from the bifurcation point, and higher order angular harmonics become important (Figure 3.3).

Consider now the derivative

$$\frac{1}{r_t}\frac{\partial r_{t+1}^2}{\partial\theta_t} = b_t \tanh(a_t r_t)\operatorname{sech}^2(a_t r_t) - a_t \tanh(b_t r_t)\operatorname{sech}^2(b_t r_t),$$

where we have used the identities $\frac{da}{d\theta_t} = b, \frac{db}{d\theta_t} = -a$. The derivative can only be zero if $a_t = 0$ or $b_t = 0$, or $a_t = \pm\infty$ and $b_t = \pm\infty$. For this to occur for all values of θ_t requires $\beta = 0, r_t = 0$ or $\alpha = \pm 1$ for the first cases, or $\beta = \infty$ or $\alpha = \pm\infty$ for the second.

$\beta = 0$ or $r_t = 0$ are trivial cases. $\alpha = \pm 1$ represents the case of strong resonances, since at the bifurcation point the eigenvalues are $\tilde{\lambda} = \frac{\pm 1 \pm i}{\sqrt{2}}$, and $\tilde{\lambda}^8 = 1$. We can therefore consider this as steady state bifurcation in $r_{t+8} = f(r_t, \theta_t)$. For β or $\alpha \to \infty$ the tanh functions become sgn functions; all trajectories then involve hopping from corner to corner, all at $r_t = 1$. Hence, except very close to the bifurcation point, and a small number of exceptional cases, r_t will oscillate as a function of θ_t.

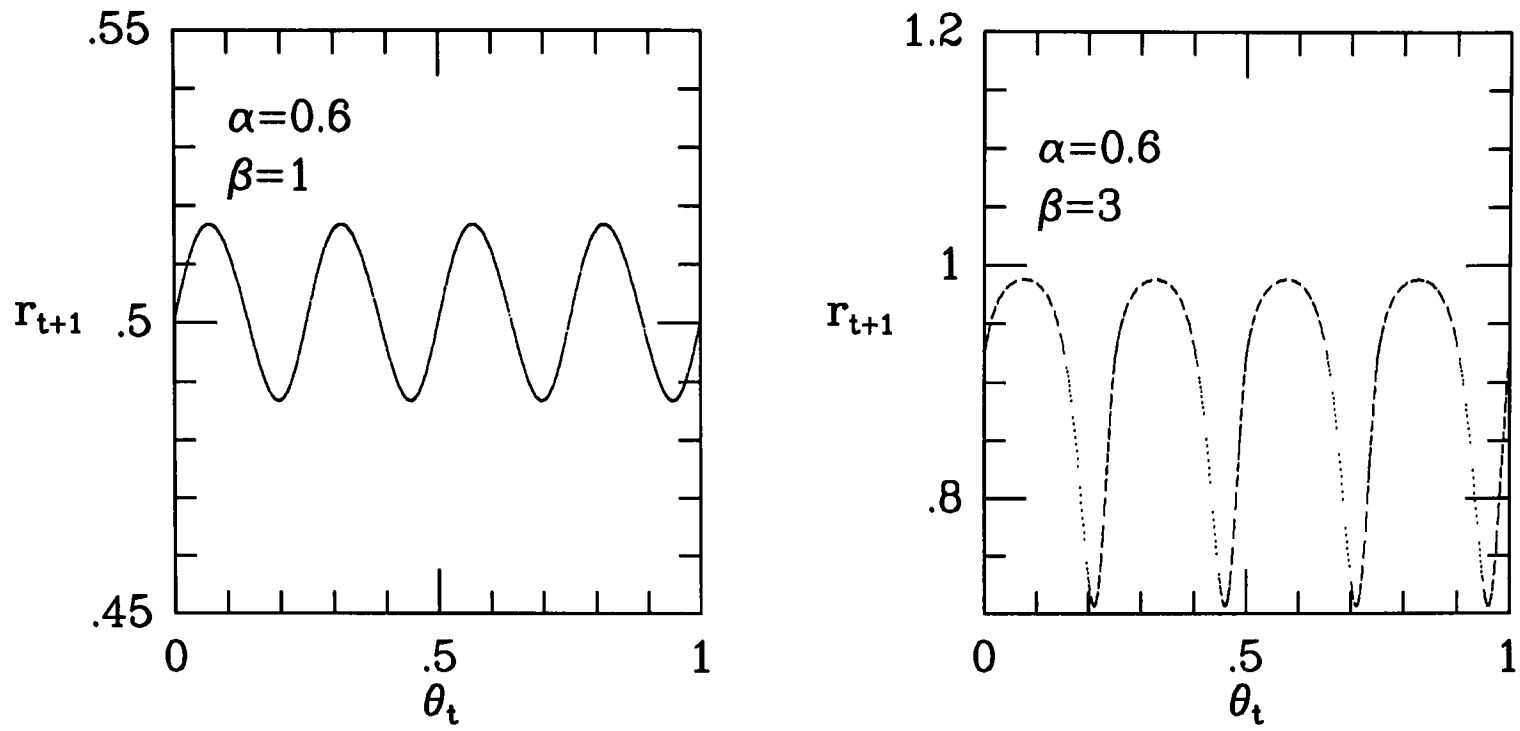


Figure 3.3: r_{t+1} against θ_t for $\beta = 1$ and $\beta = 3$ showing the increasing angular dependence of r_t as we move away from the bifurcation point

Next we wish to investigate the θ_t dependence of θ_{t+1}. Consider the derivative

$$\frac{\partial \tan\theta_{t+1}}{\partial\theta_t} = (1-\tan^2\theta_{t+1})\frac{\partial\theta_{t+1}}{\partial\theta_t} = \frac{2a_tr_t\tanh a_tr_t\,\mathrm{sech}^2 b_tr_t + 2b_tr_t\tanh b_tr_t\,\mathrm{sech}^2 a_tr_t}{(\tanh a_tr_t+\tanh b_tr_t)^2}.$$

We see that this is zero if $a_tr_t = b_tr_t = 0$, or $a_t = \pm\infty, b_t = \pm\infty$, which corresponds to $\beta = 0$ or $r_t = 0$ and $\beta = \infty$, or $\alpha = \pm\infty$ respectively; and infinite if $a_t = -b_t$, which occurs for $\tan\theta_t = -\frac{1}{\alpha}$ (the divergence is an artifact of the mathematics caused by the divergence in the tan function, and has no effect on the real behaviour).

If we examine the derivative closer we see that in the limit $\beta \to \infty$ there is also a divergence when a_t or b_t, but not both are zero, since $\mathrm{sech}^2(0) = 1$, and $\tanh(\infty) = 1$, so the linear a_t or b_t term will dominate. However the previous analysis suggests that the derivative is zero everywhere for $\beta \to \infty$, in this limit we therefore expect the function to be flat, except for intervals of vanishing measure where it is discontinuous. If we examine what happens to θ_{t+1} as θ_t is varied, in the limit $\beta \to \infty$, with $\alpha \neq 1$ we see that the tanh functions become sgn functions, so

$$\theta_{t+1} = \arctan\left(\frac{\mathrm{sgn}((1-\alpha)\cos\theta_t + (1+\alpha)\sin\theta_t) - \mathrm{sgn}((1+\alpha)\cos\theta_t - (-1+\alpha)\sin\theta_t)}{\mathrm{sgn}((1-\alpha)\cos\theta_t + (1+\alpha)\sin\theta_t) + \mathrm{sgn}((1+\alpha)\cos\theta_t - (-1+\alpha)\sin\theta_t)}\right).$$

Hence the function is step like with values $\frac{n\pi}{2}$, with the risers occurring when $\tan\theta = \frac{\alpha-1}{\alpha+1}$ or $\tan\theta = \frac{\alpha+1}{\alpha-1}$.

Therefore as we increase β from the bifurcation point, we expect the return map of θ_{t+1} against θ_t to go from a straight, diagonal line (since there is no θ_t dependence exactly at the bifurcation point), to a series of four rounded steps, to a series of four sharp steps (Figure 3.4). As β is increased the points visited by the dynamics tend to be only at the corners of the steps. Exactly at $\beta = \infty$, the function becomes non-invertible. This results in 4-cycles for $\alpha > 1$, fixed points for $\alpha < 1$, and eight-cycles for $\alpha = 1$. Exactly at $\beta = \infty$ and $\alpha = 1$, the return map only just touches the diagonal. The behaviour is therefore indeterminate, and depends on the order in which the limits $\beta \to \infty$ and $\alpha \to 0$ are taken.

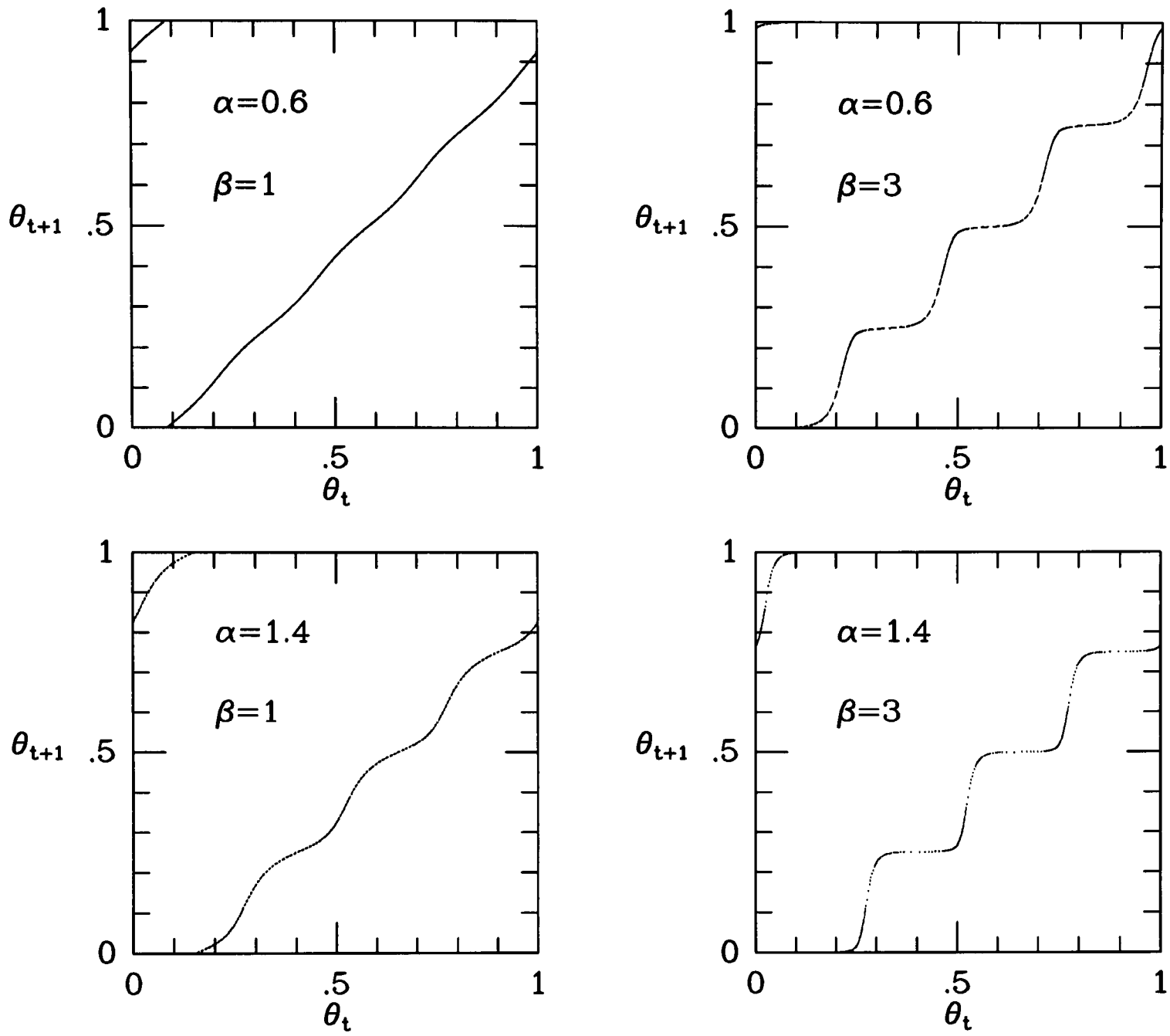


Figure 3.4: θ_{t+1} against θ_t for $\beta = 1$ and $\beta = 3$, $\alpha = 0.6$ and $\alpha = 1.4$ obtained from iterating the dynamics directly. Away from the bifurcation point the maps become stair-like

The behaviour can be determined graphically by considering the return maps (Figure 3.5). For $\alpha < 1$, the (near) vertical parts of the steps occur for $\theta < \frac{\pi}{2}$. As β is increased the steps sharpen, and hence cross the diagonal, creating four stable and four unstable fixed points. For $\alpha > 1$ the (near) vertical parts of the steps occur for $\theta > \frac{\pi}{2}$, therefore the steps will always lie below the diagonal. As β is increased and the steps sharpen, a four cycle will emerge, since each iteration will take you on to the next step. For $\alpha = 1$, the (near) vertical parts of the steps occur for $\theta = \frac{\pi}{2}$ therefore the corners of the steps will just touch the diagonal. For large β the constriction between step and diagonal is small, and therefore an iteration may not be enough to bring you to the next step. This method does not determine the behaviour well for $\beta \to \infty$ and $\alpha \to 1$ where the order of limits is important, however it is a good method to visualise the stabilisation and destabilisation, of the fixed points, four cycles, and eight cycles. Analytic estimates of their domains of stability will be derived in the next section.

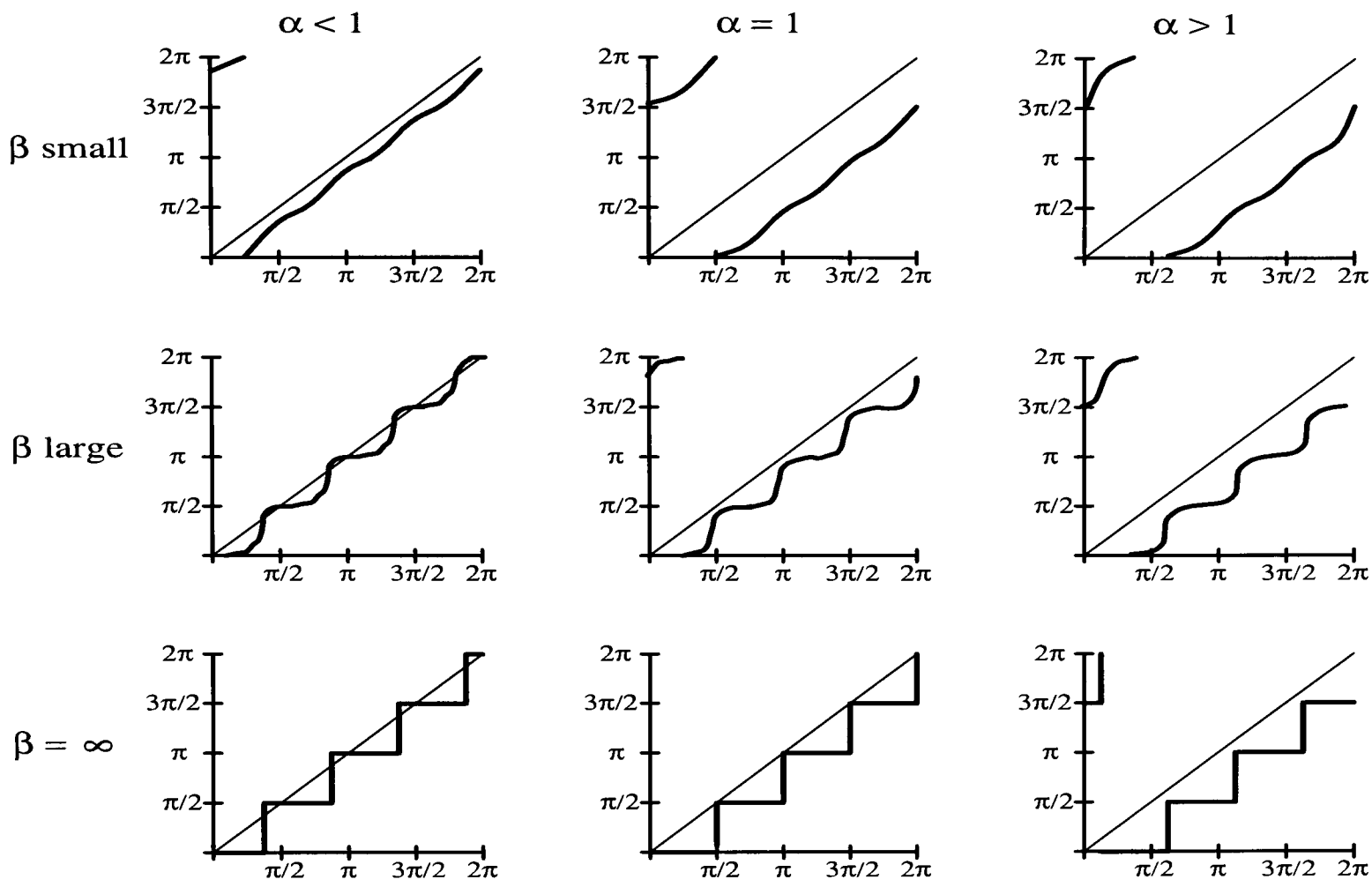


Figure 3.5: Figurative plots of return maps for $\alpha < 1$, $\alpha = 1$, $\alpha > 1$, and various values of β

3.4 Stability Analysis and Arnold Tongues

In this section we consider in detail the different periodic orbits generated by the non-linear mappings. If a complex pair of eigenvalues of the linearised equations of motion leading to a Hopf bifurcation are the n^{th} root of unity, then the behaviour close to the trivial fixed point can be considered as a steady state bifurcation in $\mathbf{m}_{t+n} = f(\mathbf{m}_t)$. These resonances cause mode locking. At the bifurcation point, non-resonant eigenvalues form points of zero measure on the unit circle forming the range of eigenvalues in the complex plane. As we move away from the bifurcation point these points grow, forming Arnold tongues - finite regions in parameter space where the system locks into a given period orbit. As these tongues grow (as some parameter is varied) they may fill all space, leave some space for quasi-periodic orbits or even overlap, causing chaos. These tongues are caused by non-linear angular terms which become important as we move away from the bifurcation point. This section explores how the tongues grow as parameters are varied in our system, and how the stability of various periodic orbits can be calculated. We use the $p = 2$ system with the matrix $\begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix}$ as our prototypical model, since it has only two parameters to adjust. We do however present numerical evidence that the $p = 3$ case also displays similar behaviour.

Figure 3.6 shows the winding number (i.e. average rotation per iteration) over a range of β for $0.1 \leq \alpha \leq 1.9$ in steps of 0.1.

Networks with $\alpha < 1$ tend towards $w = 0$ as β increases and $\alpha > 1$ tend towards $w = 0.25$ as β increases. Exactly at $\alpha = 1$, $w = 0.125$ is a stable orbit for all β. Notice also several regions of

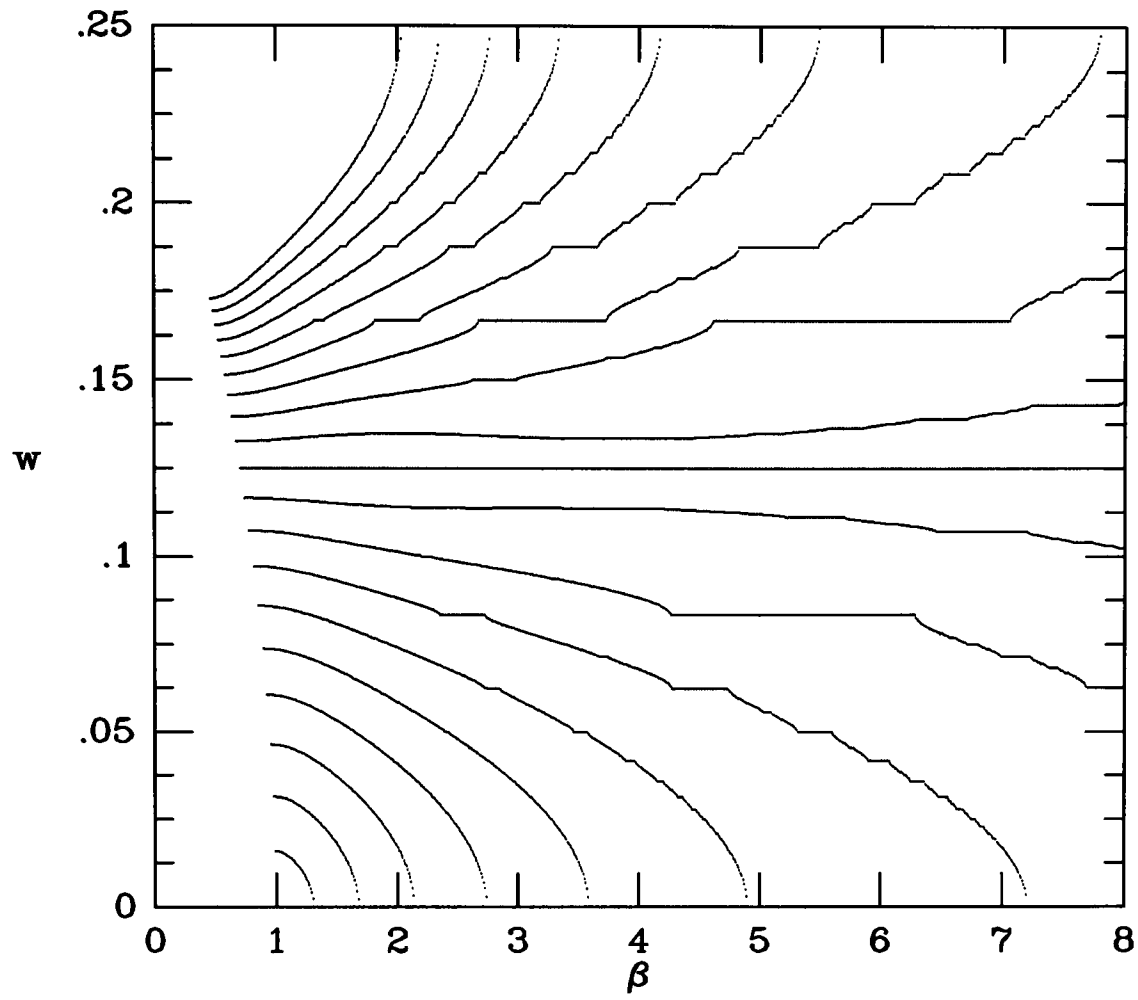


Figure 3.6: Winding number as a function of β for α from 0.1 at the bottom to 1.9 at the top, in steps of 0.1

mode locking where $1/w$ is either integer or a rational fraction, corresponding to periodic orbits. We can gain much insight into the problem by considering the equations of motion at $\beta = \infty$, then considering how the behaviour will change for finite β. For $\beta = \infty$ the tanh functions become sgn functions, and the equations of motion (3.13) become

$$\mathbf{m}_{t+1} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \mathrm{sgn}\left[(1-\alpha)m^1 + (1+\alpha)m^2\right] + \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \mathrm{sgn}\left[(1+\alpha)m^1 + (-1+\alpha)m^2\right].$$

Since the mapping now contains sgn functions, any point within a certain boundary will be mapped to a single point. Basin boundaries are given by $\frac{m^1}{m^2} = \frac{\alpha+1}{\alpha-1}$ and $\frac{m^1}{m^2} = \frac{1-\alpha}{1+\alpha}$. The behaviour of points exactly on a boundary is marginal, they will first be mapped into a basin, and then onto an appropriate point.

We can investigate the behaviour by considering what happens to the basin boundaries as α is varied, as depicted in Figure 3.7. Starting at $\alpha = 0$ the basin boundaries are the diagonals $m^1 = \pm m^2$. Consider the upper quadrant, all points here are mapped to $(0, 1)$, hence there are four equivalent fixed points, since all points within each quadrant are mapped to a point within that quadrant. As α is increased the boundary lines rotate anti-clockwise. At $\alpha = 1$ the basin boundaries intersect the axis, and hence also the fixed points. Here the behaviour changes, since the points to which the dynamics maps are on the boundaries and hence the behaviour marginal. The points on the axis will in fact me mapped clockwise to points $(\pm\frac{1}{2}, \pm\frac{1}{2})$. The trajectory is now in another basin, and will again be mapped clockwise to a point on the next axis. Hence an

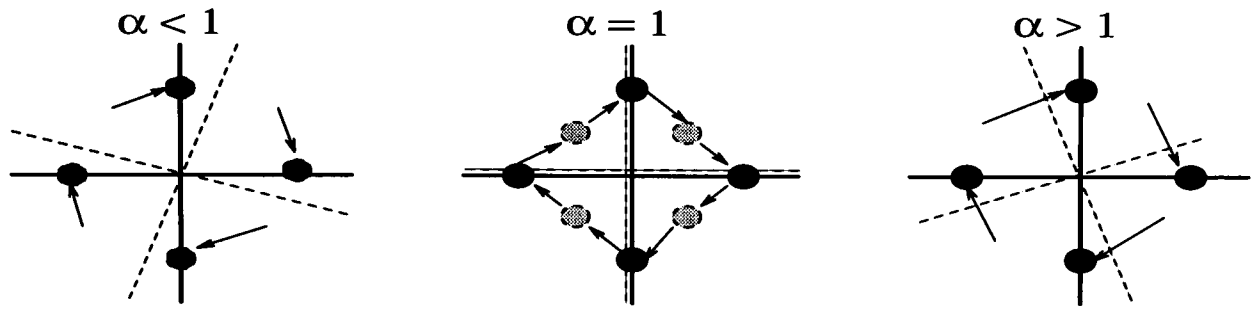


Figure 3.7: Basin boundaries and trajectories in $m^1 - m^2$ plane, for $\alpha < 1$ $\alpha = 1$ and $\alpha = 1$. The dotted lines represent the basin boundaries, and the dark dots represent points that are visited by the dynamics. The shaded dots represent points on the cycle that appear only when the basin boundaries intersect the axis'. The arrows show how points in each region are mapped to points on the cycle

8-cycle is established. For $\beta = \infty$ this only occurs for $\alpha = 1$. As α is increased further past 1, the boundaries continue to rotate anti-clockwise, once they have passed the axis' however, the points mapped to are no longer in the same quadrant, therefore a 4-cycle results.

As β is decreased the points visited by the dynamics deviate slightly from the axis'. Behaviour need not be restricted to fixed points, 4-cycles and 8-cycles, but also a wealth of integer period cycles and quasi-periodic cycles may appear. The basin boundaries now become meaningless, and the type of behaviour must be determined by a stability analysis, using the Jacobian of the mapping ($\frac{\partial m_{t+1}^{\alpha}}{\partial m_t^{\gamma}}$). If we consider some finite value of β and increase α from zero, we go from fixed points through quasi-periodic and other period cycles, to an 8-cycle, then through another region of quasi-periodic and periodic cycles, to a 4-cycle.

For $\alpha = 1 - \epsilon$, the fixed points are of the form $(1 - \delta, \mathrm{sgn}(\epsilon)\delta)$ and its analogs close to the other axis', for large β, where δ is a function of ϵ and β. As β increases δ decreases to zero. For ϵ close to zero, there are not only points close to the axis' visited by the dynamics, but also points in between. For moderate β these are close to $(\pm\frac{1}{2}, \pm\frac{1}{2})$, but as β is increased these become closer to the axis', so that points visited by the dynamics are like $(1 - \delta_1, \delta_1)$ and $(1 - \delta_2, -\delta_2)$. δ_1 and δ_2 decrease as β is increased, and at $\beta = \infty$ these two points merge on the axis, and another points is created at $(\pm\frac{1}{2}, \pm\frac{1}{2})$.

We commence the stability analysis by calculating the Jacobian of the mapping using (3.13)

$$\mathbf{J} = \begin{pmatrix} \frac{\beta(1-\alpha)}{2}\left(1 - \tanh^2(\beta A)\right) & \frac{\beta(1+\alpha)}{2}\left(1 - \tanh^2(\beta A)\right) \\ +\frac{\beta(1+\alpha)}{2}\left(1 - \tanh^2(\beta B)\right) & -\frac{\beta(1-\alpha)}{2}\left(1 - \tanh^2(\beta B)\right) \\ \\ \frac{\beta(1-\alpha)}{2}\left(1 - \tanh^2(\beta A)\right) & \frac{\beta(1+\alpha)}{2}\left(1 - \tanh^2(\beta A)\right) \\ -\frac{\beta(1+\alpha)}{2}\left(1 - \tanh^2(\beta B)\right) & +\frac{\beta(1-\alpha)}{2}\left(1 - \tanh^2(\beta B)\right) \end{pmatrix}, \tag{3.15}$$

with $A = (1 - \alpha)m^1 + (1 + \alpha)m^2)$ and $B = ((1 + \alpha)m^1 + (-1 + \alpha)m^2)$. In principal from this we can calculate the critical stability line of any cycle in β - ϵ space. In general though, this is difficult, apart from certain cycles for which symmetries simplify the problem.

3.4.1 Fixed Points, 4-Cycles and 8-Cycles

First we consider the case of $\alpha = 1 - \epsilon$ giving rise to fixed points. Since all fixed points are equivalent, we take the one near $(1,0)$, and use the postulated form $(1-\delta+\mathcal{O}(\delta^2), -\delta+\mathcal{O}(\delta^2))$, this implies that $\tanh\beta\left((2-\epsilon)m^1 - \epsilon m^2\right) \simeq 1+\mathcal{O}(\delta^2)$ and $\tanh\beta\left(\epsilon m^1 + (2-\epsilon)m^2\right) \simeq 1-2\delta+\mathcal{O}(\delta^2)$. Using these results and keeping only linear terms in δ the Jacobian becomes $2\beta\delta\begin{pmatrix} \epsilon & 2-\epsilon \\ \epsilon & 2-\epsilon \end{pmatrix}$. This assumes that β is large so that $\delta \ll 1$. Now critical stability occurs when the maximum absolute value of any eigenvalue of the Jacobian is one, which is given by $4\beta\delta = 1$. In order to obtain a value for δ in terms of β and ϵ only we now expand $\tanh\beta\left(\epsilon m^1 + (2-\epsilon)m^2\right)$ around the fixed point $(1-\delta, -\delta)$, assuming $\delta < \epsilon$, and that $\beta\epsilon$ is large giving

$$\delta \simeq \frac{1-\tanh\beta\epsilon}{2(1-\beta(1-\tanh^2\beta\epsilon))}.$$

Since the numerator is necessarily positive, and δ is positive, a restriction is placed on the range of β and ϵ for which this expression is valid: $\epsilon > \frac{1}{\beta}\tanh^{-1}(1-\frac{1}{\beta})$, as well as the condition for the expansion to be valid $\beta\delta \ll 1$, $\delta < \epsilon$, and $\beta\epsilon \gg 1$. Hence

$$\frac{2\beta(1-\tanh\beta\epsilon)}{(1-\beta(1-\tanh^2\beta\epsilon))} - 1 = 0 \tag{3.16}$$

defines a relationship between β and ϵ, valid in the high β limit giving the stability boundary of the fixed point. This equation must be solved numerically.

If we let β go to infinity along the line defined by (3.16) we can determine the asymptotic form of the line defining the stability in this limit. Writing $\tanh x \simeq 1 - 2e^{-2x}$ for $x \to \infty$ this gives an equation $8\beta e^{-2\beta\epsilon} - 4\beta e^{-4\beta\epsilon} = 1$. Assuming that $e^{-4\beta\epsilon}$ is negligible compared with $e^{-2\beta\epsilon}$ this gives a form like $\epsilon \sim \frac{\log 8\beta}{2\beta}$, so $\alpha \sim 1 - \frac{\log 8\beta}{2\beta}$.

A similar relationship can be derived for the 4-cycle, with $\alpha = 1+\epsilon$, using similar arguments and the relations between successive points on the cycle. Since all points on the 4-cycle are equivalent the Jacobian need only be calculated at one point. The Jacobian at $(-\delta, 1-\delta)$ is $2\beta\delta\begin{pmatrix} 2+\epsilon & \epsilon \\ -2-\epsilon & -\epsilon \end{pmatrix}$, with critical stability again given by $4\beta\delta = 1$. In order to calculate δ we expand the mapping at the transition $(-\delta, 1-\delta) \to (1-\delta, \delta)$ giving

$$\delta = \frac{1-\tanh\beta\epsilon}{2(1-\beta(\epsilon+1)(1-\tanh^2\beta\epsilon))},$$

where we have again assumed $\beta\delta \ll 1$, $\delta < \epsilon$ and $\beta\epsilon \gg 1$. The equation governing the stability for the 4-cycle, for large β is then

$$\frac{2\beta(1-\tanh\beta\epsilon)}{(1-\beta(\epsilon+1)(1-\tanh^2\beta\epsilon))} - 1 = 0 \tag{3.17}$$

We can derive the asymptotic functional form, in the same way as for the fixed points: $\epsilon \sim \frac{\log 4\beta(2+\epsilon)}{2\beta}$, so $\alpha \sim 1 + \frac{\log 8\beta}{2\beta}$.

We can calculate the stability line for the 8-cycle in a similar way, though in this case it is slightly more complex, since the 8-cycle contains two set of non-equivalent points, hence two Jacobians and two δ's must be calculated.

First we take the case $\alpha = 1+\epsilon$, we consider the $(1-\delta_1, -\delta) \to (\delta_2, -1+\delta_2) \to (-\delta_1, -1+\delta_1)$ part of the trajectory. The Jacobians at the first two points are

$$J^+(1-\delta_1,-\delta_1) = 2\beta\delta_2 \begin{pmatrix} -\epsilon & 2+\epsilon \\ -\epsilon & 2+\epsilon \end{pmatrix} + \mathcal{O}(\delta_2^2) \qquad J^+(\delta_2,-1+\delta_2) = 2\beta\delta_1 \begin{pmatrix} 2+\epsilon & \epsilon \\ -2-\epsilon & -\epsilon \end{pmatrix} + \mathcal{O}(\delta_1^2).$$

The critical stability line is therefore given by the eigenvalue of their product with the largest magnitude: $16\beta^2\delta_1\delta_2(\epsilon+1)^2 = 1$. We calculate δ_1 and δ_2 as before by expanding the dynamic equations, with $\delta_1\beta \ll 1$ though in contrast to before $\delta_2 > \epsilon$, and $\epsilon > \delta_1$. The equations for δ_1 and δ_2 are then

$$\delta_1 = e^{-2\beta(2\delta_2-\epsilon)} \qquad \delta_2 = \frac{1}{2}(1-\tanh\beta\epsilon) + \beta\delta_1(1-\tanh^2\beta\epsilon).$$

From the equations for δ_1 and δ_2 and the critical stability relation it is possible to derive an equation involving β and ϵ only, determining the critical stability of the 8-cycle for $\alpha > 1$ and large β.

If we use the critical stability equation to create a quadratic form for δ_2, we can again determine the asymptotic limit for the line, and find $\epsilon \sim \frac{\log 2\beta}{\beta} - \frac{\log\log 16\beta^2}{\beta}$ so $\alpha \sim 1 + \frac{\log 2\beta}{\beta} - \frac{\log\log 16\beta^2}{\beta}$.

Similarly for $\alpha = 1-\epsilon$ we study the $(1-\delta_1, \delta_1) \to (1-\delta_2, -\delta_2) \to (\delta_1, -1+\delta_1)$ part of the trajectory. The Jacobian at the first two points are

$$J^-(1-\delta_1,\delta_1) = 2\beta\delta_2 \begin{pmatrix} \epsilon & 2-\epsilon \\ \epsilon & 2-\epsilon \end{pmatrix} + \mathcal{O}(\delta_2^2) \qquad J^-(1+-\delta_2,-\delta_2) = 2\beta\delta_1 \begin{pmatrix} \epsilon & 2-\epsilon \\ \epsilon & 2-\epsilon \end{pmatrix} + \mathcal{O}(\delta_1^2)$$

The critical stability line is now given by $16\beta^2\delta_1\delta_2 = 1$. In order to calculate δ_1 and δ_2 we again assume that $\beta\delta_1 \ll 1$ and $\delta_2 > \epsilon$, giving the following expressions

$$\delta_1 = e^{-2\beta(2\delta_2-\epsilon)} \qquad \delta_2 = \frac{1}{2}(1-\tanh\beta\epsilon) + \beta\delta_1(\epsilon-1)(1-\tanh^2\beta\epsilon).$$

These again form a set of equations to be solved numerically giving the critical stability of the 8-cycle region for large β.

Again the asymptotic form of the solution can be found $\epsilon \sim \frac{\log 4\beta}{2\beta} - \frac{\log\log 16\beta^2}{2\beta}$ and $\alpha \sim 1 - \frac{\log 4\beta}{2\beta} + \frac{\log\log 16\beta^2}{2\beta}$.

3.4.2 Mode Locked Regions

Using this information we can construct a phase diagram for the network, showing the boundaries of stability of the different period cycles. The natural co-ordinates to use are T/T_c and $1/T_c$, where $T = 1/\beta$ and $T_c = \sqrt{1+\alpha^2}$ is the temperature at which the system bifurcates from the trivial fixed point. Figure 3.8 shows the phase diagram, with the dots calculated from the above stability analysis, and the lines taken between critical stability points from iterations of the mappings directly. The markers are from calculations of the asymptotic form.

The fit is only qualitative for moderate T, however for small T ($\beta \to \infty$) all three sets of data converge. In between the regions shown on the diagram are other regions corresponding to periodic orbits, which have not been shown in order that the diagram is comprehensible. In between these mode locked regions are regions of quasi-periodicity, which shrink as $T \to 0$.

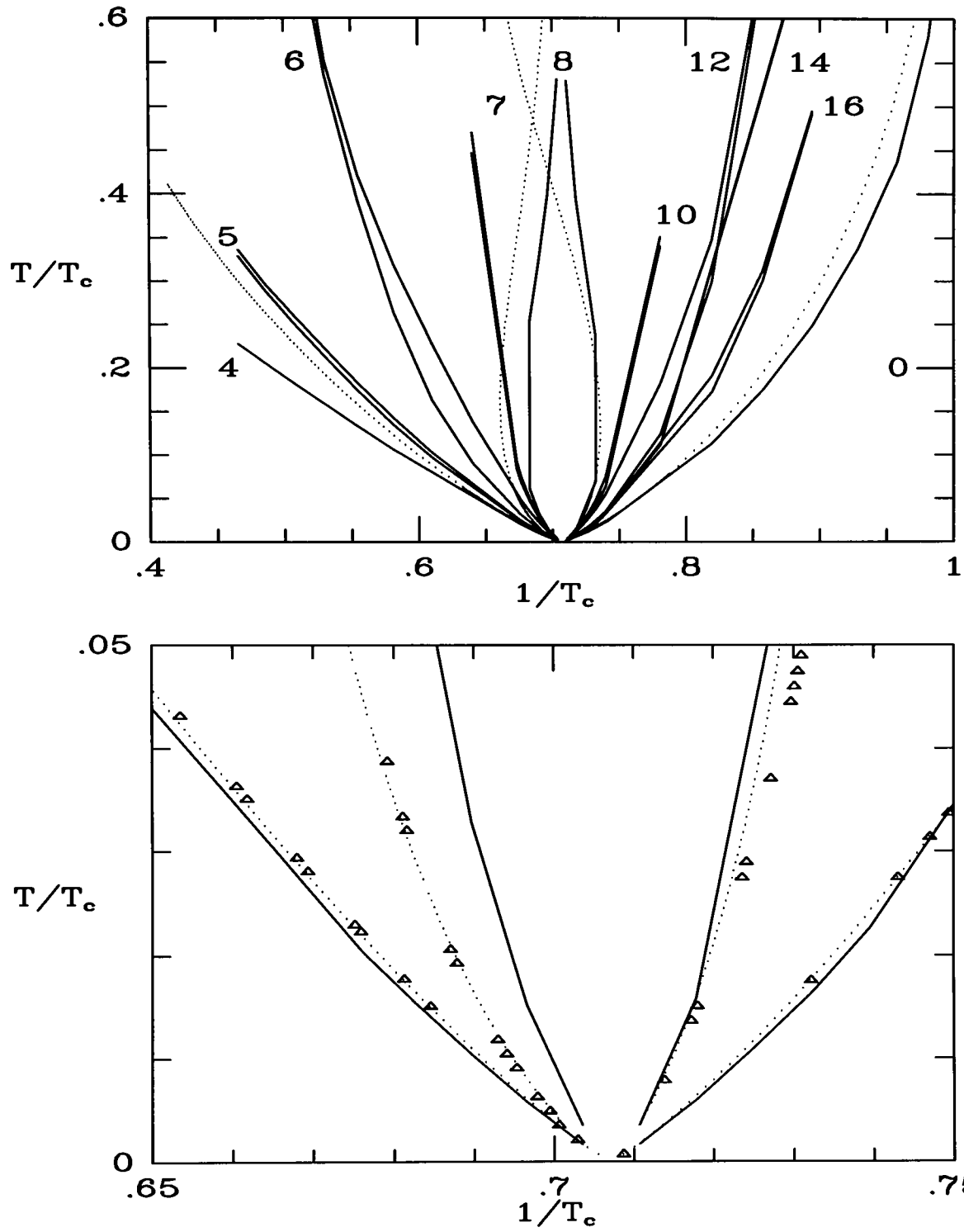


Figure 3.8: Phase diagram showing the different periodic trajectories. The number represents the period in the particular region. The dots mark points calculated from the preceding stability analysis; the triangle are points taken from the $\beta \to \infty$ asymptotic form of these equations; and the straight lines join points taken from direct iteration of the dynamic equations. The lines are meant as a guide between points only. A restricted number of periodic regions only is shown in order not to crowd the diagram.

We expect similar behaviour to occur in networks trained with any number of patterns, since the normal forms are the same up to cubic order. Figure 3.9 shows the winding number and 1/period for a network trained with three patterns, using the matrix (3.9), with an arbitrary linear combination ($m^1 + m^2$ and m^3) as the two co-ordinates. The winding number shows similar behaviour to the $p = 2$ case with many mode locked steps, showing that a Hopf bifurcation leads to motion with only two free variables. The period is measured to be the number of steps required to return arbitrarily close to the starting point. We see that in the mode locked steps the winding number and 1/period are the same. This is as we would expect. Out of the mode locked regions, i.e. in the quasi-periodic regime, the period is higher than the winding number would suggest, reflecting the fact that the period must be integer.

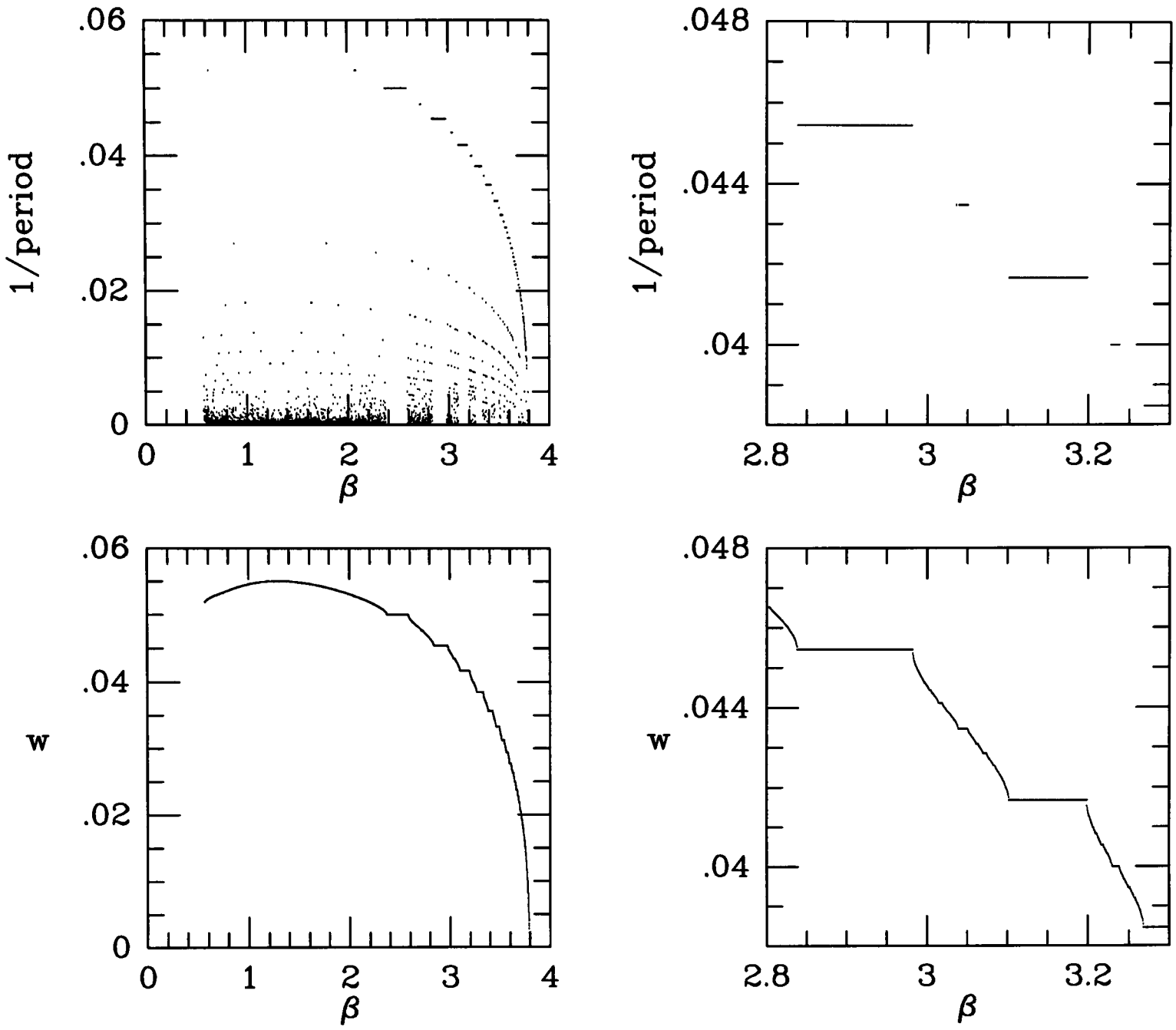


Figure 3.9: Winding number and 1/period for p=3

In principal we could construct a phase diagram for networks with arbitrary matrix **A**. Since there are p^2 free parameters the phase diagram would need to be in p^2 dimensional space. As we varied any one of the parameters, keeping β (the overall scaling factor) fixed, we would go through a range of periodic and quasi-periodic orbits. The precise shape of the various regions would depend on the numerical values in the embedding matrix **A**. There would however be certain regions which we can readily identify :

1. The fixed point region corresponding to symmetry of **A** when all eigenvalues are pure real.
2. The 4-cycle region corresponding to complete asymmetry of **A** when all eigenvalues are pure imaginary.
3. The 2^{p+1}-cycle region which for $\beta = \infty$ corresponds to the basin boundaries coinciding with the axis'

3.5 Conclusion

In this chapter we have studied the qualitative behaviour of stochastic Ising spin neural networks, with separable interactions given by (1.19) and general (possibly asymmetric) embedding matrix **A**. The macroscopic variables (pattern overlaps) which in the limit $p \ll \sqrt{N}\,, N \to \infty$ (where p is the number of patterns and N the number of neurons) evolve due to the deterministic laws (1.21, 1.22) for asynchronous and synchronous updating of the neurons respectively, show a variety of bifurcation phenomena if the noise (parametrised by an inverse temperature β) and the eigenvalues of the embedding matrix **A** are varied. We have shown that the qualitative behaviour of such types of neural networks can be analysed using the methods of bifurcation theory. These methods can be used to predict when and how the system will move away from the trivial fixed point and we have shown that the type of bifurcation is dependent only on the form of the eigenvalue of **A** with the largest real part (in the continuous time case) or the largest modulus (in the discrete time case). Away from the critical point non-trivial and quasi-periodic cycles appear, due to the non-linear angular terms becoming important. These are the *Arnold Tongues*, and the stability of these cycles can be calculated in the large β limit. Hence a phase diagram may be constructed in the p^2 dimensions of the free parameters of the embedding matrix **A**. This phase diagram shows how the regions where different periodic orbits are stable are affected by changes to the elements in **A**. In between the mode locked periodic phases are regions of quasi-periodic behaviour where the period of the orbit is irrational, hence the system returns close to its starting point, but never coincides with it.

We have supported our analysis with numerical simulations which illustrate the types of behaviour encountered. In this way we have shown the condition necessary for a recursive neural network of this separable type to show the properties of

1. Associative memory is $\lambda^* \in \Re$
2. Sequence retrieval is $\lambda^* \in \Re + \Im$

where λ^* is the eigenvalue of **A** with the largest real part in the continuous time case, or the largest modulus in the discrete time case. We have shown that in the case of sequence retrieval the period depends both on the eigenvalues of the embedding matrix **A** and the noise, parametrised by β. Hence for a network with a given amount of internal noise the eigenvalues of the embedding matrix **A** must be tuned in order to be able to retrieve a given sequence.

Part II

The Generalised Hopfield Model Near Saturation

Chapter 4

Detailed balance equilibrium and the free energy

4.1 Introduction

Here we derive the equilibrium properties of the generalised Hopfield model near saturation (i.e. trained with an extensive number p of patterns), and symmetric embedding matrix $\mathbf{A}$. Interactions J_{ij} are given by (1.23), with $A_{\mu\nu} = A_{\nu\mu}$. We use standard equilibrium statistical mechanical techniques to derive the phase diagram from the free energy, and use the replica method to allow us to average a logarithm. We obtain an integral of the form $\int dO\, d\hat{O}\, e^{N\Psi(\{O,\hat{O}\})}$ which in the thermodynamic limit ($N \to \infty$) will be dominated by the values of the order parameters $O, \hat{O}$ which maximise Ψ. The derivation closely follows that of [35], which in turn is similar to the corresponding calculation for the S-K spin-glass [15].

The calculation of the free energy for neural networks with an extensive number of patterns differs from that for the spin-glass since there are an extensive number of magnetisation like order parameters (the overlaps $m^\mu = \frac{1}{N}\sum_i \xi_i^\mu \sigma_i$). Hence the saddle point method cannot be used to solve the integral directly, with the (extensive number of) overlaps as integration variables, since fluctuations around the saddle point will be significant. However, it is easy to convince oneself that only a finite number overlaps can simultaneously be $\mathcal{O}(1)$, and that the rest must be $\mathcal{O}(\frac{1}{\sqrt{N}})$. Consider the energy of the Hopfield model (i.e. $\mathbf{A} = \mathbb{1}$) $E \sim \sum_{\mu=1}^{p} (m^\mu)^2$. For this to be physically meaningful (i.e. extensive) for $p = \alpha N$, requires only a finite number, c, of so-called *condensed* overlaps. The effect of the remaining $p - c$ *uncondensed* overlaps is parametrised by the order parameter $r = \frac{N}{p}\sum_{\mu>c}^{p} (m^\mu)^2$. We will see how this can be generalised to arbitrary symmetric $\mathbf{A}$.

4.2 Disorder Average

We assume the free energy to be self averaging with respect to the quenched disorder $\boldsymbol{\xi}$, so that for symmetric $\mathbf{A}$ we can attempt to calculate it using the replica method: $-\beta N f = \langle \ln[Z] \rangle_{\boldsymbol{\xi}} = \lim_{n\to 0} \langle \frac{Z^n - 1}{n} \rangle_{\boldsymbol{\xi}}$, where the angle brackets indicate an average over $\boldsymbol{\xi}$, and Z is the partition function for a particular realisation of the disorder $Z = \sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma}, J_{ij})}$. Using the standard Ising

Hamiltonian, and interactions given by (1.23) this gives

$$Z = \sum_{\boldsymbol{\sigma}} e^{\frac{\beta}{2}\sum_{i,j\neq i}\sigma_i J_{ij}\sigma_j} = \sum_{\boldsymbol{\sigma}} e^{\frac{\beta}{2N}\sum_{\mu,\nu}\sum_i \sigma_i \xi_i^\mu A_{\mu\nu}\sum_j \xi_j^\nu \sigma_j + \mathcal{O}(1)}.$$

There are now two strategies with which we can proceed.

4.2.1 Method I

The first is to use the Hubbard-Stratonivich transformation

$$e^{\frac{1}{2\beta}b^\mu A_{\mu\nu} b^\nu} = \int \frac{d\tilde{m}^1 \ldots d\tilde{m}^p}{\sqrt{\left(\frac{2\pi}{\beta}\right)^p \det \mathbf{A}}} e^{-\frac{\beta}{2}\tilde{m}^\mu A^{-1}_{\mu\nu}\tilde{m}^\nu + b^\mu \tilde{m}^\mu},$$

with $b^\mu = \frac{\beta}{\sqrt{N}}\sum_i \xi_i^\mu \sigma_i$, and $\mathbf{A}$ positive definite (any negative modes can be dealt with by making the appropriate b^μ imaginary) to give

$$Z = \left(\frac{\beta}{2\pi}\right)^{\frac{p}{2}} (\det \mathbf{A})^{-\frac{1}{2}} \sum_{\boldsymbol{\sigma}} \int \left(\prod_\mu d\tilde{m}^\mu\right) e^{-\frac{\beta}{2}\tilde{\mathbf{m}}\cdot\mathbf{A}^{-1}\tilde{\mathbf{m}} + \frac{\beta}{\sqrt{N}}\sum_i \sigma_i \boldsymbol{\xi}_i \tilde{\mathbf{m}}}.$$

At some stage we will have to make the condensed ansatz, i.e. assume that only a finite number, c, overlaps are $\mathcal{O}(1)$ and that the remaining $p-c$ are $\mathcal{O}(\frac{1}{\sqrt{N}})$. In this case the ξ^μ's which represent the quenched disorder take on a slightly different meaning, dependent on whether $\mu > c$ or $\mu \leq c$. Although we average over them both, this average will be carried out at different stages. To this end we will write $\xi_i^{\mu>c} = \eta_i^\mu$. Therefore

$$\langle Z^n \rangle_{\boldsymbol{\eta}} \propto \prod_a \left\langle \left\langle \int \left(\prod_\mu d\tilde{m}_a^\mu\right) e^{-\frac{\beta}{2}\sum_{\mu,\nu}\tilde{m}_a^\mu A^{-1}_{\mu\nu}\tilde{m}_a^\nu + \frac{\beta}{\sqrt{N}}\sum_{\mu\leq c}\sum_i \sigma_i^a \xi_i^\mu \tilde{m}_a^\mu + \frac{\beta}{\sqrt{N}}\sum_{\mu>c}\sum_i \sigma_i^a \eta_i^\mu \tilde{m}_a^\mu} \right\rangle_{\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\eta}}.$$

In order to proceed we follow the standard path of reversing the order of averaging, and carrying out the disorder average first. Averaging over η_i^μ with $\mu > c$ gives a term

$$\prod_{\mu>c} e^{\sum_i \ln\cosh \frac{\beta}{\sqrt{N}}\sum_a m_a^\mu \sigma_i^a}.$$

We can expand this using $\ln\cosh x \simeq \frac{x^2}{2}$ to give for the uncondensed patterns a term like

$$\int \left(\prod_a \prod_{\mu>c} d\tilde{m}_a^\mu\right) e^{-\frac{1}{2}\sum_{\mu,\nu>c}\sum_{a,b}\tilde{m}_a^\mu\left(\beta A^{-1}_{\mu\nu}\delta_{ab} - \beta^2\delta_{\mu\nu}\frac{1}{N}\sum_i \sigma_i^a\sigma_i^b\right)\tilde{m}_b^\nu - \beta\sum_a\sum_{\mu>c}\sum_{\nu\leq c}\tilde{m}_a^\mu A^{-1}_{\mu\nu}\tilde{m}_a^\nu},$$

where in the last term we have used the symmetry property of $\mathbf{A}$. In order to remove the terms containing σ's allowing us to integrate over $m^{\mu>c}$, we introduce a convenient integral representation of unity

$$1 = \int dq_{ab}\ \delta\left(q_{ab} - \frac{1}{N}\sum_i \sigma_i^a\sigma_i^b\right) = \int \frac{dq_{ab}\, d\hat{q}_{ab}}{2\pi} e^{-iN\hat{q}_{ab}\left(q_{ab} - \frac{1}{N}\sum_i \sigma_i^a\sigma_i^b\right)}.$$

In order to clarify notation, we split the matrices $\mathbf{A}$ and $\mathbf{A}^{-1}$ into four quadrants, dependent on whether each row (column) represents a condensed or an uncondensed pattern

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{cc} & \mathbf{A}_{cu} \\ \mathbf{A}_{uc} & \mathbf{A}_{uu} \end{pmatrix} \qquad \mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}^{-1}_{cc} & \mathbf{A}^{-1}_{cu} \\ \mathbf{A}^{-1}_{uc} & \mathbf{A}^{-1}_{uu} \end{pmatrix}.$$

NB $\mathbf{A}_{cc} \neq (\mathbf{A}_{cc}^{-1})^{-1}$.

Integrating over the uncondensed patterns then gives a term like

$$\frac{e^{\frac{\beta^2}{2}\sum_{a,b}\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cu}^{-1}\left(\beta\mathbf{A}_{uu}^{-1}\otimes\mathbb{I}-\beta^2\mathbb{I}\otimes\mathbf{q}\right)^{-1}_{ab}\mathbf{A}_{uc}^{-1}\tilde{\mathbf{m}}_b}}{\sqrt{\det\left(\beta\mathbf{A}_{uu}^{-1}\otimes\mathbb{I}-\beta^2\mathbb{I}_{uu}\otimes\mathbf{q}\right)}}\int\frac{dq_{ab}\,d\hat{q}_{ab}}{2\pi}e^{-iN\hat{q}_{ab}\left(q_{ab}-\frac{1}{N}\sum_i\sigma_i^a\sigma_i^b\right)},$$

so that

$$\langle Z^n\rangle=\left(\frac{\beta}{2\pi}\right)^{\frac{nc}{2}}\int\left(\prod_a\prod_{\mu\leq c}d\tilde{m}_a^\mu\prod_{a,b}\frac{dq_{ab}\,d\hat{q}_{ab}}{2\pi}\right)\frac{e^{-\frac{\beta}{2}\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cc}^{-1}\tilde{\mathbf{m}}_a+\frac{\beta}{2}\sum_{a,b}\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cu}^{-1}\left(\mathbf{A}_{uu}^{-1}\otimes\mathbb{I}-\beta\mathbb{I}\otimes\mathbf{q}\right)^{-1}_{ab}\mathbf{A}_{uc}^{-1}\tilde{\mathbf{m}}_b}}{\sqrt{(\det\mathbf{A})^n\det\left(\mathbf{A}_{uu}^{-1}\otimes\mathbb{I}-\beta\mathbb{I}\otimes\mathbf{q}\right)}}$$

$$\times e^{-iN\sum_{a,b}q_{ab}\hat{q}_{ab}}2^{nN}\left\langle\left\langle e^{\frac{\beta}{\sqrt{N}}\sum_{\mu\leq c}\sum_i\sigma_i^a\xi_i^\mu\tilde{m}_a^\mu+i\hat{q}_{ab}\sum_i\sigma_i^a\sigma_i^b}\right\rangle_{\boldsymbol{\sigma}}\right\rangle_{\boldsymbol{\xi}}. \tag{4.1}$$

If we rescale our integration variables so that $\tilde{m}_a^{\mu\leq c}\to\sqrt{N}\tilde{m}_a^\mu$, then we have an integral of the form

$$\langle Z^n\rangle=\int\left(\prod_a\prod_{\mu\leq c}d\tilde{m}_a^\mu\prod_{a,b}\frac{dq_{ab}\,d\hat{q}_{ab}}{2\pi}\right)\;e^{-\beta N\Psi_I(\tilde{\mathbf{m}},\mathbf{q},\hat{\mathbf{q}})}=e^{-\beta N\Psi_I(\tilde{\mathbf{m}}',\mathbf{q}',\hat{\mathbf{q}}')},$$

where

$$\begin{aligned}\Psi_I &= \frac{1}{2}\sum_a\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cc}^{-1}\tilde{\mathbf{m}}_a-\frac{1}{2}\sum_{a,b}\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cu}^{-1}\left(\boldsymbol{\Lambda}^{-1}\right)^{ab}\mathbf{A}_{uc}^{-1}\tilde{\mathbf{m}}_b+\frac{1}{2\beta N}\,\mathrm{Tr}\,\ln\left[(\mathbf{A}_{uu}\otimes\mathbb{I})\,\boldsymbol{\Lambda}\right]\\ &+\frac{i}{\beta}\sum_{a,b}\hat{q}_{ab}q_{ab}-\frac{1}{\beta N}\sum_i\ln\left[e^{\left\langle\left\langle\beta\sum_{\mu\leq c}\sum_a\sigma_i^a\xi_i^\mu\tilde{m}_a^\mu+i\sum_{a,b}\sigma_i^a\hat{q}_{ab}\sigma_i^b\right\rangle_{\boldsymbol{\sigma}}\right\rangle_{\boldsymbol{\xi}}}\right]-\frac{n}{\beta}\ln 2,\end{aligned} \tag{4.2}$$

and

$$\Lambda_{\mu\nu}^{ab}=\left(\mathbf{A}_{\mu\nu}^{-1}\delta_{ab}-\beta\delta_{\mu\nu}q_{ab}\right).$$

The primed order parameters are the values which minimise $\Psi_I(\tilde{\mathbf{m}},\mathbf{q},\hat{\mathbf{q}})$, hence

$$f=\lim_{n\to0}\frac{1}{n\beta N}\left(1-e^{-\beta N\Psi_I(\tilde{\mathbf{m}}',\mathbf{q}',\hat{\mathbf{q}}')}\right)\simeq\lim_{n\to0}\frac{1}{n}\Psi_I(\tilde{\mathbf{m}}',\mathbf{q}',\hat{\mathbf{q}}').$$

The extremal conditions for $\Psi_I(\tilde{\mathbf{m}},\mathbf{q},\hat{\mathbf{q}})$ $\frac{\partial\Psi_I}{\partial\tilde{m}_a^\mu}=\frac{\partial\Psi_I}{\partial q_{ab}}=\frac{\partial\Psi_I}{\partial\hat{q}_{ab}}=0$ give the saddle point equations

$$\begin{aligned}\left(A_{\nu\mu}^{-1}-A_{\nu\rho}^{-1}\left(\Lambda^{-1}\right)_{\rho\lambda}^{ba}A_{\lambda\mu}^{-1}\right)\tilde{m}_a^\mu &= \left\langle\left\langle\xi^\nu\sigma^b\right\rangle_{\boldsymbol{\sigma}}\right\rangle_{\boldsymbol{\xi}} \\ q_{ab} &= \left\langle\left\langle\sigma^a\sigma^b\right\rangle_{\boldsymbol{\sigma}}\right\rangle_{\boldsymbol{\xi}} \\ i\hat{q}_{ab} &= \frac{\beta}{2}\frac{\partial}{\partial q_{ab}}\sum_{a,b}\tilde{\mathbf{m}}_a\cdot\mathbf{A}_{cu}^{-1}\left(\boldsymbol{\Lambda}^{-1}\right)^{ab}\mathbf{A}_{uc}^{-1}\tilde{\mathbf{m}}_b \\ &\quad -\frac{1}{2N}\frac{\partial}{\partial q_{ab}}\,\mathrm{Tr}\,\ln\left[(\mathbf{A}_{uu}\otimes\mathbb{I})\,\boldsymbol{\Lambda}\right].\end{aligned} \tag{4.3}$$

Since the interactions are infinite range, we have made the mean field approximation (that all sites are equivalent), so that the brackets $\langle\langle\ldots\rangle_{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\xi}}$, indicate an average over the effective Hamiltonian defined by

$$\langle\langle f(\boldsymbol{\sigma},\boldsymbol{\xi})\rangle_{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\xi}}=\frac{\sum_{\boldsymbol{\sigma}}\sum_{\boldsymbol{\xi}}f(\boldsymbol{\sigma},\boldsymbol{\xi})\,e^{\beta\sum_{\mu\leq c}\sum_a\sigma^a\xi^\mu\tilde{m}_a^\mu+i\sum_{a,b}\sigma^a\hat{q}_{ab}\sigma^b}}{\sum_{\boldsymbol{\sigma}}\sum_{\boldsymbol{\xi}}e^{\beta\sum_{\mu\leq c}\sum_a\sigma^a\xi^\mu\tilde{m}_a^\mu+i\sum_{a,b}\sigma^a\hat{q}_{ab}\sigma^b}}.$$

4.2.2 Method II

In the first method however, the physical meaning of the m^μ's is lost. A more transparent way to proceed is to introduce the following integral representation of unity

$$1 = \int \left(\prod_\mu dm^\mu \right) \prod_\mu \delta(m^\mu - m^\mu(\boldsymbol{\sigma}, \boldsymbol{\xi}^\mu)) = \left(\frac{\sqrt{N}}{2\pi} \right)^p \int \left(\prod_\mu dm^\mu \, d\hat{m}^\mu \right) e^{i\sqrt{N} \sum_\mu \hat{m}^\mu \left(m^\mu - m^\mu(\boldsymbol{\sigma}, \boldsymbol{\xi}^\mu) \right)}$$

where $m^\mu(\boldsymbol{\sigma}, \boldsymbol{\xi}^\mu) = \frac{1}{N} \sum_i \xi_i^\mu \sigma_i$. Once again making the condensed ansatz, and writing $\xi_i^{\mu > c} = \eta_i^\mu$ so that

$$\langle Z^n \rangle_{\boldsymbol{\eta}} \propto \prod_a \left\langle \left\langle \int (d\mathbf{m}_a \, d\hat{\mathbf{m}}_a) \, e^{\frac{\beta}{2} N \mathbf{m}_a \cdot \mathbf{A}_{cc} \mathbf{m}_a + i\sqrt{N} \mathbf{m}_a \cdot \hat{\mathbf{m}}_a - \frac{i}{\sqrt{N}} \hat{\mathbf{m}}_a \cdot \sum_i \sigma_i^a \boldsymbol{\xi}_i - \frac{i}{\sqrt{N}} \sum_{\mu > c} \hat{m}_a^\mu \sum_i \sigma_i^a \eta_i^\mu} \right\rangle_{\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\eta}}.$$

Carrying out the average over $\boldsymbol{\eta}$ gives a term

$$\prod_{\mu > c} e^{\sum_i \ln \left[\cos \frac{1}{\sqrt{N}} \sum_a \hat{m}_a^\mu \sigma_i^a \right]} \simeq \prod_{\mu > c} e^{-\frac{1}{2} \sum_{a,b} \hat{m}_a^\mu \frac{1}{N} \sum_i \sigma_i^a \sigma_i^b \hat{m}_b^\mu}.$$

Again we introduce the integral representation of unity

$$1 = \int dq_{ab} \, \delta \left(q_{ab} - \frac{1}{N} \sum_i \sigma_i^a \sigma_i^b \right) = \int \frac{dq_{ab} \, d\hat{q}_{ab}}{2\pi} e^{-iN\hat{q}_{ab} \left(q_{ab} - \frac{1}{N} \sum_i \sigma_i^a \sigma_i^b \right)}.$$

Integrating over $\hat{m}^{\mu > c}$ then gives a term like

$$\frac{e^{-\frac{1}{2} N \mathbf{m}_a \cdot \left(\mathbb{1} \otimes \mathbf{q}^{-1} - \beta \mathbf{A}_{uu} \otimes \mathbb{1} \right)_{ab} \mathbf{m}_b + N\beta \sum_a \sum_{\mu > c} \sum_{\nu \leq c} m_a^\mu A_{\mu\nu} m_b^\nu}}{\sqrt{((2\pi)^n \det \mathbf{q})^{(p-c)}}},$$

carrying out the integral over $m^{\mu > c}$, and rescaling $\hat{m}_a^{\mu \leq c} \to \sqrt{N} \hat{m}_a^{\mu \leq c}$ then gives

$$\langle Z^n \rangle \propto \int \left(\prod_a d\mathbf{m}_a \, d\hat{\mathbf{m}}_a \prod_{a,b} \frac{dq_{ab} \, d\hat{q}_{ab}}{2\pi} \right) \frac{e^{N \left(\frac{\beta}{2} \mathbf{m}_a \mathbf{A}_{cc} \mathbf{m}_b + i \hat{\mathbf{m}}_a \mathbf{m}_a + \beta^2 \mathbf{m}_a \mathbf{A}_{cu} \left(\mathbb{1} \otimes \mathbf{q}^{-1} - \beta \mathbf{A}_{uu} \otimes \mathbb{1} \right)^{-1}_{ab} \mathbf{A}_{uc} \mathbf{m}_b \right)}}{\sqrt{(\det \mathbf{q})^{(p-c)} \det \left(\mathbb{1} \otimes \mathbf{q}^{-1} - \beta \mathbf{A}_{uu} \otimes \mathbb{1} \right)}}$$
$$\times e^{-iN \sum_{a,b} q_{ab} \hat{q}_{ab}} \, 2^{nN} \left\langle \left\langle e^{-i \sum_\mu \sum_a \hat{m}_a^\mu \sum_i \sigma_i^a \xi_i^\mu + i \sum_{a,b} \hat{q}_{ab} \sum_i \sigma_i^a \sigma_i^b} \right\rangle_{\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\xi}}. \tag{4.4}$$

This is again a saddle point form

$$\langle Z^n \rangle \propto \int \left(\prod_a \prod_{\mu \leq c} dm_a^\mu \, d\hat{m}_a^\mu \prod_{a,b} \frac{dq_{ab} \, d\hat{q}_{ab}}{2\pi} \right) e^{-\beta N \Psi_{II}(\mathbf{m}, \hat{\mathbf{m}}, \mathbf{q}, \hat{\mathbf{q}})} = e^{-\beta N \Psi_{II}(\mathbf{m}', \hat{\mathbf{m}}', \mathbf{q}', \hat{\mathbf{q}}')},$$

where

$$\begin{aligned} \Psi_{II} \;=\; & \frac{1}{2} \sum_a \mathbf{m}_a \cdot \mathbf{A}_{cc} \mathbf{m}_a + i \sum_a \hat{\mathbf{m}}_a \cdot \mathbf{m}_a + \frac{\beta}{2} \sum_{a,b} \mathbf{m}_a \cdot \mathbf{A}_{cu} \left(\tilde{\Lambda}^{-1} \right)_{uu}^{ab} \mathbf{A}_{uc} \mathbf{m}_b \\ & - \frac{1}{2\beta N} \operatorname{Tr} \ln \left[(\mathbb{1} \otimes \mathbf{q}) \, \tilde{\boldsymbol{\Lambda}} \right] + \frac{i}{\beta} \sum_{a,b} \hat{q}_{ab} q_{ab} \\ & - \frac{1}{\beta N} \sum_i \ln \left[e^{\left\langle \left\langle -i \sum_{\mu \leq c} \sum_a \sigma_i^a \xi_i^\mu \hat{m}_a^\mu + i \sum_{a,b} \sigma_i^a \hat{q}_{ab} \sigma_i^b \right\rangle_{\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\xi}}} \right] - \frac{n}{\beta} \ln 2, \end{aligned} \tag{4.5}$$

and

$$\tilde{\Lambda}^{ab}_{\mu\nu} = \left(\delta_{\mu\nu} q_{ab} - \beta \mathbf{A}^{-1}_{\mu\nu} \delta_{ab}\right).$$

The primed order parameters are the values which minimise $\Psi_{II}(\tilde{\mathbf{m}}, \mathbf{q}, \hat{\mathbf{q}})$, hence

$$f = \lim_{n\to 0} \frac{1}{n\beta N} \left(1 - e^{-\beta N \Psi_{II}(\mathbf{m}', \hat{m}', \mathbf{q}', \hat{\mathbf{q}}')}\right) \simeq \lim_{n\to 0} \frac{1}{n} \Psi_{II}(\mathbf{m}', \hat{\mathbf{m}}', \mathbf{q}', \hat{\mathbf{q}}').$$

The extremal conditions for $\Psi_{II}(\mathbf{m}, \hat{\mathbf{m}}, \mathbf{q}, \hat{\mathbf{q}})$ $\frac{\partial \Psi_{II}}{\partial m^{\mu}_{a}} = \frac{\partial \Psi_{II}}{\partial \hat{m}^{\mu}_{a}} = \frac{\partial \Psi_{I}}{\partial q_{ab}} = \frac{\partial \Psi_{II}}{\partial \hat{q}_{ab}} = 0$ give the saddle point equations

$$\begin{aligned}
m^{\mu}_{a} &= \langle\langle \xi^{\mu} \sigma^{a} \rangle_{\boldsymbol{\sigma}} \rangle_{\boldsymbol{\xi}} \\
-i\hat{m}^{\mu}_{a} &= A_{\mu\nu} m^{\nu}_{a} + \beta A_{\mu\lambda} \tilde{\Lambda}^{ab}_{\lambda\rho} A_{\rho\nu} m^{\nu}_{b} \\
q_{ab} &= \left\langle \left\langle \sigma^{a} \sigma^{b} \right\rangle_{\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\xi}} \\
-i\hat{q}_{ab} &= \frac{\beta^2}{2} \frac{\partial}{\partial q_{ab}} \sum_{a,b} \mathbf{m}_{a} \mathbf{A}_{cu} \left(\tilde{\Lambda}^{-1}\right)^{ab}_{uu} \mathbf{A}_{uc} \mathbf{m}_{b} - \frac{1}{2N} \frac{\partial}{\partial q_{ab}} \operatorname{Tr} \ln \left[(\mathbb{1} \otimes \mathbf{q})\, \tilde{\boldsymbol{\Lambda}}\right].
\end{aligned} \tag{4.6}$$

Again we have made the mean field approximation, and the average over the effective Hamiltonian $\langle\langle \ldots \rangle_{\boldsymbol{\sigma}} \rangle_{\boldsymbol{\xi}}$ is now defined by

$$\langle\langle f(\boldsymbol{\sigma}, \boldsymbol{\xi}) \rangle_{\boldsymbol{\sigma}} \rangle_{\boldsymbol{\xi}} = \frac{\sum_{\boldsymbol{\sigma}} \sum_{\boldsymbol{\xi}} f(\boldsymbol{\sigma}, \boldsymbol{\xi})\, e^{-i \sum_{\mu \le c} \sum_{a} \sigma^{a} \xi^{\mu} \hat{m}^{\mu}_{a} + i \sum_{a,b} \sigma^{a} \hat{q}_{ab} \sigma^{b}}}{\sum_{\boldsymbol{\sigma}} \sum_{\boldsymbol{\xi}} e^{-i \sum_{\mu \le c} \sum_{a} \sigma^{a} \xi^{\mu} \hat{m}^{\mu}_{a} + i \sum_{a,b} \sigma^{a} \hat{q}_{ab} \sigma^{b}}}.$$

From equations (4.4) and (4.6) we notice that the two methods are equivalent, once we realise that

$$\Lambda^{ab}_{\mu\nu} = -\tilde{\Lambda}^{ab}_{\mu\nu} \qquad \text{and} \qquad \left(A^{-1}_{\nu\mu} - A^{-1}_{\nu\rho} \left(\Lambda^{-1}\right)^{ba}_{\rho\lambda} A^{-1}_{\lambda\mu}\right) \tilde{m}^{\mu}_{a} = m^{\nu}_{b}.$$

Since the methods are equivalent, we proceed using method I.

4.3 Replica Symmetry

In order to calculate the free energy, we assume *Replica Symmetry* (RS) i.e. that $\tilde{m}^{\mu}_{a} = \tilde{m}^{\mu}$, $q_{ab} = \delta_{ab} + q(1 - \delta_{ab})$, $i\hat{q}_{ab} = \frac{p}{2N} \beta^2 r (1 - \delta_{ab}) = \frac{\alpha \beta^2}{2} r (1 - \delta_{ab})$ for all a, b. Although it is known that in the low temperature region the RS saddle points are unstable [50] signifying a breaking of ergodicity, and replica-symmetry breaking (RSB) techniques have been developed [16, 17] these will not be discussed here.

We now have to calculate the eigenvalues and inverse of $(\delta_{ab}(A_{\mu\eta} - \beta\delta_{\mu\eta}) - \beta q_{ab}\delta_{\mu\eta})$ in the limit $n \to 0$. Now $Tr \ln \boldsymbol{\Lambda} = \sum_{a} \ln \lambda_{a}$ where λ_{a} are the eigenvalues of Λ. Considering the trace over replica space, and using $q_{ab} = \delta_{ab} + q(1 - \delta_{ab})$, we see that the replica matrix has the form

$$\boldsymbol{\Lambda} = \begin{pmatrix}
\mathbf{A}^{-1}_{uu} - \beta\mathbb{1} & -\beta q \mathbb{1} & \ldots & -\beta q \mathbb{1} \\
-\beta q \mathbb{1} & \mathbf{A}^{-1}_{uu} - \beta\mathbb{1} & \ldots & -\beta q \mathbb{1} \\
\vdots & \vdots & \ddots & \vdots \\
-\beta q \mathbb{1} & -\beta q \mathbb{1} & \ldots & \mathbf{A}^{-1}_{uu} - \beta\mathbb{1}
\end{pmatrix}.$$

This has one eigenvalue $\lambda_1 = \mathbf{A}_{uu}^{-1} - \beta\mathbb{1} - (n-1)\beta q\mathbb{1}$ and $n-1$ eigenvalues $\lambda_{2\ldots n} = \mathbf{A}_{uu}^{-1} - \beta\mathbb{1} + \beta q\mathbb{1}$. Therefore the trace has the form

$$\begin{aligned}
\mathrm{Tr}\,\ln[\Lambda] &= \mathrm{Tr}\,\ln\left[\mathbf{A}_{uu}^{-1} - \beta(1-q+nq)\mathbb{1}\right] + (n-1)\,\mathrm{Tr}\,\ln\left[\mathbf{A}_{uu}^{-1} - \beta(1-q)\right] \\
&= \mathrm{Tr}\,\ln\left[\mathbb{1} - n\beta q\left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)^{-1}\right] + n\,\mathrm{Tr}\,\ln\left[\mathbf{A}^{-1} - \beta(1-q)\mathbb{1}\right] \\
\lim_{n\to 0}\frac{1}{n}\mathrm{Tr}\,\ln[\Lambda] &= -\beta q\,\mathrm{Tr}\left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)^{-1} + \mathrm{Tr}\,\ln\left[\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right].
\end{aligned}$$

Similarly we need to know the inverse of $\boldsymbol{\Lambda}$. Writing it in the form $\Lambda_{\mu\nu}^{ab} = \delta_{ab}(A_{\mu\nu}^{-1} - \beta\delta_{\mu\nu}(1-q)) - \beta q\delta_{\mu\nu}$ we then postulate an inverse of the form $\delta_{bc}B_{\nu\eta} + C_{\nu\eta}$. This then defines B and C, by

$$\delta_{ac}\left(A_{\mu\nu}^{-1}B_{\nu\eta} - \beta(1-q)B_{\nu\eta}\right) + (A_{\mu\nu}^{-1} - \beta\delta_{\mu\nu}(1-q))C_{\nu\eta} - \beta qB_{\mu\eta} - n\beta qC_{\mu\eta} = \delta_{ac}\delta_{\mu\eta},$$

therefore

$$\begin{aligned}
(A_{\mu\nu}^{-1} - \beta(1-q)\delta_{\mu\nu})B_{\nu\eta} = \delta_{\mu\eta} &\implies \mathbf{B} = \left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)^{-1} \\
(A_{\mu\nu}^{-1} - \beta\delta_{\mu\nu}(1-q+nq))C_{\nu\eta} - \beta qB_{\mu\eta} = 0 &\implies \mathbf{C} = \beta q(\mathbf{A}_{uu}^{-1} - \beta(1-q+nq)\mathbb{1})^{-1}\mathbf{B}
\end{aligned}$$

and

$$\begin{aligned}
\left(\Lambda^{-1}\right)_{\mu\nu}^{ab} &= \delta_{ab}\left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)_{\mu\nu}^{-1} \\
&\quad + \beta q\left((\mathbf{A}_{uu}^{-1} - \beta(1-q+nq)\mathbb{1})^{-1}\left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)^{-1}\right)_{\mu\nu}. \qquad (4.7)
\end{aligned}$$

Lastly we have to calculate the spin average within the RS ansatz which we achieve by introducing an auxiliary integration variable. Now $\sum_{a,b\neq a}\sigma^a\sigma^b = (\sum_a \sigma^a)^2 - n$. Therefore this term is of the form

$$\begin{aligned}
&-\ln\left[e^{-\frac{n}{2}\alpha r\beta^2}\prod_a\sum_{\sigma^a} e^{\left\langle\beta\sigma^a\sum_{\mu\le c}\xi^\mu\tilde{m}^\mu + \frac{\beta^2\alpha r}{2}\left(\sum_a\sigma^a\right)^2\right\rangle_{\boldsymbol{\xi}}}\right] \\
&= -\ln\left[e^{-\frac{n}{2}\alpha r\beta^2}\prod_a\sum_{\sigma^a}\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,e^{\left\langle\beta(z\sqrt{\alpha r}+\sum_{\mu\le c}\xi^\mu\tilde{m}^\mu)\sigma^a\right\rangle_{\boldsymbol{\xi}}}\right] \\
&= -\ln\left[e^{-\frac{n}{2}\alpha r\beta^2}\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\left\langle 2\cosh\beta(z\sqrt{\alpha r}+\sum_{\mu\le c}\xi^\mu\tilde{m}^\mu)\right\rangle_{\boldsymbol{\xi}}^n\right],
\end{aligned}$$

dividing by n and taking the limit $n\to 0$ this term becomes

$$\begin{aligned}
&-\lim_{n\to 0}\frac{1}{n\beta}\ln\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}e^{n\ln\left\langle 2\cosh\beta(z\sqrt{\alpha r}+\sum_{\mu\le c}\xi^\mu\tilde{m}^\mu)\right\rangle_{\boldsymbol{\xi}} - \frac{1}{2}n\alpha r\beta^2} \\
&\simeq \frac{1}{2}\alpha\beta r - \frac{1}{\beta}\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\left\langle\ln\left[2\cosh\beta(z\sqrt{\alpha r}+\sum_{\mu\le c}\tilde{m}^\mu\xi^\mu)\right]\right\rangle_{\boldsymbol{\xi}}
\end{aligned}$$

where in the last step we have used the fact that $e^x \simeq 1+x$ and $\ln(1+x)\simeq x$.

Therefore the full RS free energy becomes

$$\begin{aligned} f_{RS} &= \frac{1}{2}\tilde{\mathbf{m}}\cdot\left(\mathbf{A}_{cc}^{-1}-\mathbf{A}_{cu}^{-1}\left(\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right)^{-1}\mathbf{A}_{uc}\right)\tilde{\mathbf{m}}-\frac{q}{2N}\operatorname{Tr}\left(\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right)^{-1} \\ &\quad+\frac{1}{2\beta N}\operatorname{Tr}\ln\left[\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right]+\frac{1}{2}\alpha\beta r(1-q) \\ &\quad-\frac{1}{\beta}\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\left\langle\ln\left[2\cosh\beta(z\sqrt{\alpha r}+\sum_{\mu\leq c}\tilde{m}^{\mu}\xi^{\mu})\right]\right\rangle_{\boldsymbol{\xi}}. \end{aligned} \tag{4.8}$$

Using the symmetry property of $\mathbf{A}$ the saddle point equations are

$$\begin{aligned} &\left(\delta_{\kappa\mu}\delta_{\mu\eta}-A_{\kappa\mu}^{-1}\left(\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right)_{\mu\eta}^{-1}\right)A_{\eta\lambda}^{-1}\tilde{m}^{\lambda} \\ &\quad=\int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\left\langle\xi^{\kappa}\tanh\beta\left(z\sqrt{\alpha r}+\boldsymbol{\xi}\cdot\tilde{\mathbf{m}}\right)\right\rangle_{\boldsymbol{\xi}} \\ q &= \int\frac{dz}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\left\langle\tanh^2\beta\left(z\sqrt{\alpha r}+\boldsymbol{\xi}\cdot\tilde{\mathbf{m}}\right)\right\rangle_{\boldsymbol{\xi}} \\ \alpha r &= \frac{q}{N}\operatorname{Tr}\left(\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right)^{-1}\left(\mathbf{A}_{uu}^{-1}-\beta(1-q)\mathbb{I}\right)^{-1}-\tilde{\mathbf{m}}\cdot\mathbf{A}_{cu}^{-1}\frac{\partial\boldsymbol{\Lambda}^{-1}}{\partial q}\mathbf{A}_{uc}^{-1}\tilde{\mathbf{m}}. \end{aligned} \tag{4.9}$$

We can see that if $A_{\mu\nu}=\delta_{\mu\nu}$ i.e. the Hopfield model, this reduces exactly to the result derived in [36, 37]. The results are obviously dependent upon the (symmetric) choice of $\mathbf{A}$. In particular the presence of the $\mathbf{A}_{cu}$ and $\mathbf{A}_{uc}$ terms coupling the condensed and non-condensed patterns causes a great deal of complexity.

4.4 Conclusion

We have shown how to calculate the free energy of the generalised Hopfield model trained with an extensive number of patterns, with interactions given by (1.23), and arbitrary symmetric $\mathbf{A}$, using the replica method. This extension of the work of [36, 37] (who considered the case $\mathbf{A}=\mathbb{I}$), shows that the coupling of condensed and uncondensed patterns, by the off block diagonal elements $\mathbf{A}_{cu}$ and $\mathbf{A}_{uc}$ adds a great deal of complexity.

Upon making the condensed ansatz (that only a finite number of patterns are $\mathcal{O}(1)$), we have evaluated the free energy using the replica symmetric ansatz, and evaluated the saddle point equations.

Some results for explicit choices of $\mathbf{A}$ can be found in [6], who independently obtained the same results.

Chapter 5

Dynamics of a finite number of order parameters

5.1 Introduction

Recently a method has been proposed for deriving a closed set of equations governing the evolution of macroscopic order parameters in the Hopfield neural network model near saturation [44, 58]. This method, based on the systematic removal of microscopic memory effects, has subsequently been applied to other disordered spin systems [45, 46], and is understood to be exact at least (*i*) for short times (upon appropriate choice of initial conditions), (*ii*) in equilibrium, and (*iii*) in the limit where the disorder is removed (i.e. for attractor neural networks far from saturation). For an overview of the method and its present applications refer to [59]. Although for intermediate time-scales the procedure is not exact (manifested in an overall slowing down of the dynamics, which the theory does not account for) it does capture the essential characteristics of the flows and recovers the well known equilibrium properties of the archetypal disordered spin systems [15, 36] as stable fixed points of the dynamic equations, including the full replica formalism. Furthermore, in contrast to phenomenological strategies for deriving flow equations, based on, or inspired by, time-dependent Landau-Ginzburg type equations as in [25, 80], or based on making an (incorrect) Gaussian ansatz for the local field distribution as in [81] (with various degrees of success in explaining dynamical phenomena), the present theory is derived from microscopic principles and generates explicitly the non-Gaussian shape of the local field distribution which is ultimately responsible for the spin-glass type features of the dynamics.

In this chapter we apply the theory in [44, 58, 59] to the more general and technically more complicated case of fully connected neural networks near saturation with (*i*) arbitrary separable interactions, which (*ii*) need not be symmetric, and with (*iii*) more than one condensed pattern. Since the absence of symmetry of the neural interactions implies absence of detailed balance, our analysis includes systems for which equilibrium statistical mechanics does not apply, so that hitherto there has been no theory available with which they could be studied. This latter aspect was in fact one of the main motivations behind the development of the theory in [44, 58].

Following the key assumptions of the theory in [44, 58, 59]: (i) self-averaging of the macroscopic flow with respect to the disorder, and (ii) equipartitioning of probability within the macroscopic sub-shells of the ensemble, the distribution of intrinsic noise components of the alignment fields

is calculated with the replica method. In our analysis we make the replica-symmetric ansatz. In the region where replica symmetry is stable, numerical simulations on large systems (40,000 spins), using the method proposed by Kohring [82], show that for the models considered our equations capture the essential features of the flow, even for non-symmetric choices for the neural interactions (i.e. without detailed balance). One specific implication of this result is that we now have a general theory with which to calculate, for instance, the storage capacity of a class of attractor neural networks in which the stored attractors are limit-cycles as opposed to fixed-points. For symmetric systems, the fixed points of our flow equations are shown to reproduce the thermodynamic equilibrium equations derived for the present class of networks with separable interactions in Chapter 4.

We make the choice (1.24) for the macroscopic order parameters comprising the condensed overlaps m^μ $\mu = 1 \ldots c \ll \sqrt{N}$, and the disordered contribution to the 'energy' r, then the corresponding macroscopic probability distribution $\mathcal{P}_t(\mathbf{m}, r)$ is:

$$\mathcal{P}_t(\mathbf{m}, r) = \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\, \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\, \delta(r - r(\boldsymbol{\sigma})).$$

The essential point to notice is that the local fields $h_i(\boldsymbol{\sigma})$ can be written in terms of the condensed overlaps, plus a noise term

$$h_i(\boldsymbol{\sigma}) \;=\; \sum_{\mu,\nu \leq c} \xi_i^\mu A_{\mu\nu} m^\nu(\boldsymbol{\sigma}) + z_i^s(\boldsymbol{\sigma}) + z_i^a(\boldsymbol{\sigma})$$

$$z_i^{s,a} \;=\; \sum_{[\mu,\nu>c \;|\; \mu>c,\; \nu\leq c \;|\; \mu\leq c,\; \nu>c]} \xi_i^\mu A_{\mu\nu}^{s,a} \cdot \left[\frac{1}{N}\sum_{j\neq i} \xi_j^\nu \sigma_j\right], \qquad (5.1)$$

where $\frac{1}{p}\sum_i \sigma_i z_i^s = r(\boldsymbol{\sigma}) - \frac{1}{p}\,\mathrm{Tr}\,\mathbf{A} + \mathcal{O}(\frac{1}{\sqrt{N}})$. Therefore using (1.14) we have deterministic dynamic equations

$$\begin{aligned}
\frac{d}{dt}\mathbf{m} &= \int dz^s\, dz^a \left\langle \mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a]\boldsymbol{\zeta} \tanh\beta(\boldsymbol{\zeta}\cdot\mathbf{A}_{cc}\mathbf{m} + z^s + z^a)\right\rangle_{\boldsymbol{\zeta}} - \mathbf{m} \\
\frac{1}{2}\frac{d}{dt}r &= \frac{1}{\alpha}\int dz^s\, dz^a \left\langle \mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] z^s \tanh\beta(\boldsymbol{\zeta}\cdot\mathbf{A}_{cc}\mathbf{m} + z^s + z^a)\right\rangle_{\boldsymbol{\zeta}} - r + \frac{1}{p}\,\mathrm{Tr}\,\mathbf{A} \qquad (5.2)
\end{aligned}$$

given in terms of a distribution $\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a]$ of the intrinsic noise contributions to the alignment fields within the sublattices:

$$\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] = 2^c \left\langle \frac{1}{N}\sum_i \delta(z^s - z_i^\sigma(\boldsymbol{\sigma}))\, \delta(z^a - z_i^a(\boldsymbol{\sigma}))\, \delta_{\boldsymbol{\zeta}\boldsymbol{\xi}_i}\right\rangle_{\mathbf{m},r;t} \qquad \int dz^s\, dz^a \left\langle \mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a]\right\rangle_{\boldsymbol{\zeta}} = 1, \qquad (5.3)$$

where we have introduced the subshell-average

$$\langle \Phi \rangle_{\mathbf{m},r;t} = \frac{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\, \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\, \delta(r - r(\boldsymbol{\sigma}))\, \Phi(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\, \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\, \delta(r - r(\boldsymbol{\sigma}))}.$$

The task of solving the dynamics is now one of finding the distribution (5.3), which at this stage still carries an explicit time dependence, due to the appearance of the microscopic probability distribution $p_t(\boldsymbol{\sigma})$ in the sub-shell average.

So far the theory is exact for $N \to \infty$, within the condensed ansatz. However, the noise distributions depend explicitly on t through the microscopic probability distribution $p_t(\boldsymbol{\sigma})$, requiring us to solve the master equation (1.5), which is exactly what we want to avoid. We now close the macroscopic equations (5.2), following [44, 58, 59], by assuming:

1. The flow equations (5.2), and hence the noise distributions, are self averaging with respect to the disorder, i.e. the variables ξ_i^μ, allowing us to average over them.

2. We can assume equipartitioning of probability within the $(\mathbf{m}, r)$ sub-shells of the ensemble as far as the calculation of the $\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a]$ is concerned.

NB We do necessarily not assume that the microscopic probability distribution $p_t(\boldsymbol{\sigma})$ obeys equipartitioning, rather that probability fluctuations within the sub-shells are negligible once the average over patterns has been carried out. These two assumptions, the validity of which can be verified only by comparison of the predictions of the resulting theory with numerical simulations, close the flow equations (5.2) and reduce our problem to that of calculating

$$\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] = \left\langle \frac{\sum_{\boldsymbol{\sigma}} \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\, \delta(r - r(\boldsymbol{\sigma}))\, \delta(z^s - z_i^s(\boldsymbol{\sigma}))\, \delta(z^a - z_i^a(\boldsymbol{\sigma}))}{\sum_{\boldsymbol{\sigma}} \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\, \delta(r - r(\boldsymbol{\sigma}))} \right\rangle_{\boldsymbol{\eta}} \quad \text{for } \boldsymbol{\xi}_i = \boldsymbol{\zeta} \qquad (5.4)$$

Here we have redefined the pattern components to be averaged over to differentiate them from the condensed patterns: $\eta_j^\mu = \xi_j^\mu$ for $\mu > c$.

5.2 The Intrinsic Noise Distributions

5.2.1 Replica Approach

We calculate the noise distributions using the replica identity (1.8). Upon insertion of integral representations for the various delta functions we can write the intrinsic noise distributions as

$$\begin{aligned} \mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] \;=\; & \lim_{n\to 0} \left(\frac{N}{2\pi}\right)^{2n} \int dx\, dy\, e^{ixz^s + iyz^a} \int \left(\prod_a d\hat{\mathbf{m}}_a\, d\hat{r}^a\right) e^{iN\sum_a (\hat{\mathbf{m}}_a \cdot \mathbf{m}_a + \hat{r}^a r)} \\ & \times \left\langle \left\langle e^{-ixz^s(\boldsymbol{\sigma}^1) - iyz^a(\boldsymbol{\sigma}^1) - iN\sum_a (\hat{\mathbf{m}}_a \cdot \mathbf{m}(\boldsymbol{\sigma}^a) + \hat{r}^a r(\boldsymbol{\sigma}^a))} \right\rangle_{\boldsymbol{\eta}} \right\rangle_{\boldsymbol{\sigma}^a} \end{aligned} \qquad (5.5)$$

with the (Roman) replica indices $a = 1, \ldots, n$ labeling the replicas. We now define n $(p - c)$ quantities $z_a^\mu = \frac{1}{\sqrt{N}} \sum_{k\neq i} \eta_k^\mu \sigma_k^a \sim \mathcal{O}(1)$, $(\mu > c)$, so that the average in (5.5) becomes

$$\left\langle \left\langle e^{-i\left(\sum_{\mu>c} \eta_i^\mu \left(\sum_{\nu>c} (xA^s_{\mu\nu} + yA^a_{\mu\nu}) \frac{z_1^\nu}{\sqrt{N}} + \sum_{\nu\leq c} (xA^s_{\mu\nu} + yA^a_{\mu\nu}) m^\nu\right) + \sum_{\mu\leq c, \nu>c} \zeta^\mu (xA^s_{\mu\nu} + yA^a_{\mu\nu}) \frac{z_1^\nu}{\sqrt{N}}\right)} \right.\right.$$
$$\left.\left. e^{-iN\sum_a \left(\hat{\mathbf{m}}^a \cdot \mathbf{m}(\boldsymbol{\sigma}^a) + \frac{\hat{r}^a}{\alpha N} \left(\sum_{\mu,\nu>c} \left(z_a^\mu + \frac{\eta_i^\mu \sigma_i^a}{\sqrt{N}}\right) A^s_{\mu\nu} \left(z_a^\nu + \frac{\eta_i^\nu \sigma_i^a}{\sqrt{N}}\right) + 2\sqrt{N} \sum_{\mu>c, \nu\leq c} \left(z_a^\mu + \frac{\eta_i^\mu \sigma_i^a}{\sqrt{N}}\right) A^s_{\mu\nu} m^\nu\right)\right)} \right\rangle_{\boldsymbol{\eta}} \right\rangle_{\boldsymbol{\sigma}}$$

Assuming that $\sum_{\nu>c} \eta_i^\mu A^s_{\mu\nu} \eta_i^\nu \sim \mathcal{O}(1)$, so that this term can be neglected, we can average over $\{\eta_i^\mu\}$ and $\{z_i^\mu\}$ upon introduction of the density of states

$$\mathcal{D}(\mathbf{z}, \boldsymbol{\sigma}) = \left\langle \delta(\mathbf{z} - \frac{1}{\sqrt{N}} \sum_{k\neq i} \boldsymbol{\eta}_k \otimes \boldsymbol{\sigma}_k) \right\rangle_{\boldsymbol{\eta}} = \frac{e^{-\frac{1}{2}\mathbf{z}\cdot(\mathbf{q}^{-1}\otimes \mathbb{1})\mathbf{z}}}{\sqrt{(2\pi)^{n(p-c)} \det(\mathbf{q} \otimes \mathbb{1})}} \qquad (\text{for } N \to \infty), \qquad (5.6)$$

where $q_{ab}(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{k\neq i} \sigma_k^a \sigma_k^b$, and $\mathbb{1}_{\mu\nu} = \delta_{\mu\nu}$. The average in (5.5) therefore becomes

$$\int d\mathbf{z}\, \mathcal{D}(\mathbf{z}, \boldsymbol{\sigma}) e^{-\frac{i}{\alpha}\sum_a \hat{r}^a \mathbf{z}_a \cdot \mathbf{A}^s_{uu} \mathbf{z}_a} e^{-i\sum_a \sum_{\mu\leq c, \nu>c} \left(\zeta^\mu (xA^s_{\mu\nu} + yA^a_{\mu\nu}) \frac{\delta_{a1}}{\sqrt{N}} + \frac{2}{\alpha}\sqrt{N} \hat{r}^a m^\mu A^s_{\mu\nu}\right) z_a^\nu}$$

$$\times \prod_{\mu>c} \cos\left(\left(\sum_{\nu>c}\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)\frac{z^\nu_1}{\sqrt{N}}+\sum_{\nu\leq c}\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)m^\nu\right)\right.$$
$$\left.+\frac{2}{\alpha}\sum_a \hat{r}^a\left(\sum_{\nu>c}A^s_{\mu\nu}\frac{\sigma^a_i z^\nu_a}{\sqrt{N}}+\sum_{\nu\leq c}A^s_{\mu\nu}\sigma^a_i m^\nu\right)\right).$$

To evaluate the integral we expand the cosine: $\cos x = 1-\frac{x^2}{2}+\mathcal{O}(x^4)$. The requirement that the $\mathcal{O}(x^4)$ terms vanish places a restriction on the matrix $\mathbf{A}$. Splitting the symmetric and anti-symmetric parts of $\mathbf{A}$ into condensed-condensed, condensed-uncondensed, uncondensed-condensed and uncondensed-uncondensed sub-matrices

$$\mathbf{A}=\begin{pmatrix} A^s_{cc} & A^s_{cu}\\ A^s_{uc} & A^s_{uu}\end{pmatrix}+\begin{pmatrix} A^a_{cc} & A^a_{cu}\\ A^a_{uc} & A^a_{uu}\end{pmatrix}$$

shows the restrictions on $\mathbf{A}$ to be $\sum_{\nu>c}A^{s,a}_{\mu\nu}z^\nu\sim\mathcal{O}(1)$ and $\sum_{\nu\leq c}A^{s,a}_{\mu\nu}m^\nu\sim\mathcal{O}(\frac{1}{\sqrt{N}})$, with in both cases $\mu>c$. i.e. the condensed-uncondensed and uncondensed-condensed interactions must be $\mathcal{O}(\frac{1}{\sqrt{N}})$, and the uncondensed-uncondensed interactions either sparse, or $\mathcal{O}(\frac{1}{\sqrt{N}})$.

Assuming the matrix $\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}^s_{uu}$ to be positive definite, the result of the Gaussian integral has the form e^{Φ} where

$$\Phi = \ln\left[\frac{e^{-\frac{1}{2}\boldsymbol{\Upsilon}\cdot(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}^s_{uu})^{-1}\boldsymbol{\Upsilon}}}{\sqrt{\det(\mathbf{q}\otimes\mathbb{1})\det(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}^s_{uu})}}\right]$$
$$\times\sum_{\mu>s}\ln\left[1-\frac{1}{2}\frac{\int d\mathbf{z}\, e^{-\frac{1}{2}\mathbf{z}\cdot(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}^s_{uu})\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}\,\Gamma^2_\mu(x,y,\mathbf{m},\hat{\mathbf{r}},\mathbf{z})}{\int d\mathbf{z}\, e^{-\frac{1}{2}\mathbf{z}\cdot(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}^s_{uu})\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}}\right]$$
$$\Upsilon^\nu_a = \sum_{\mu\leq c}\zeta^\mu\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)\frac{\delta_{a1}}{\sqrt{N}}+\frac{2}{\alpha}\sqrt{N}\hat{r}^a m^\mu A^s_{\mu\nu}$$
$$\Gamma_\mu(x,y,\mathbf{m},R,\mathbf{z}) = \left(\sum_{\nu>c}\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)\frac{z^\nu_1}{\sqrt{N}}+\sum_{\nu\leq c}\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)m^\nu\right)$$
$$+\frac{2}{\alpha}\sum_a\hat{r}^a\left(\sum_{\nu>c}A^s_{\mu\nu}\frac{\sigma^a_i z^\nu_a}{\sqrt{N}}+\sum_{\nu\leq c}A^s_{\mu\nu}\sigma^a_i m^\nu\right).$$

We split Φ into extensive part ($N\Omega$) which will determine the saddle point equations, and the intensive part ($-\mathcal{R}$) which when we take $n\to 0$ will determine the noise distribution.

$$\Omega = -\frac{2}{\alpha^2}\sum_{a,b}\sum_{\mu,\rho\leq c}\sum_{\nu,\eta>c}\hat{r}^a m^\mu A^s_{\mu\nu}\left(\Lambda^{-1}\right)^{ab}_{\nu\eta}A^s_{\eta\rho}m^\rho\hat{r}^b-\frac{1}{2N}\det(\mathbf{q}\otimes\mathbb{1})\det(\boldsymbol{\Lambda})$$
$$\mathcal{R} = \frac{1}{\alpha}\sum_b\sum_{\mu,\rho\leq c}\sum_{\nu,\eta>c}\left(\zeta^\mu\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)\right)\left(\Lambda^{-1}\right)^{1b}_{\nu\eta}A^s_{\eta\rho}m^\rho\hat{r}^b$$
$$+\frac{1}{\alpha}\sum_a\sum_{\mu,\rho\leq c}\sum_{\nu,\eta>c}\hat{r}^a m^\mu A^s_{\mu\nu}\left(\Lambda^{-1}\right)^{a1}_{\nu\eta}\left(\zeta^\rho\left(xA^s_{\rho\eta}+yA^a_{\rho\eta}\right)\right)^\dagger$$
$$+\frac{1}{2N}\sum_{\mu,\rho\leq c}\sum_{\nu,\eta>c}\left(\zeta^\mu\left(xA^s_{\mu\nu}+yA^a_{\mu\nu}\right)\right)\left(\Lambda^{-1}\right)^{11}_{\nu\eta}\left(\zeta^\rho\left(xA^s_{\rho\eta}+yA^a_{\rho\eta}\right)\right)^\dagger$$
$$+\frac{1}{2}\sum_{\mu>c}\frac{\int d\mathbf{z}\, e^{-\frac{1}{2}\mathbf{z}\cdot(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}_{uu})\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}\,\Gamma^2_\mu(x,y,\mathbf{m},\hat{\mathbf{r}},\mathbf{z})}{\int d\mathbf{z}\, e^{-\frac{1}{2}\mathbf{z}\cdot(\mathbf{q}^{-1}\otimes\mathbb{1}+\frac{2i}{\alpha}\hat{\mathbf{r}}\otimes\mathbf{A}_{uu})\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}}$$
$$\Lambda^{ab}_{\mu\nu} = q^{-1}_{ab}\delta_{\mu\nu}+\frac{2i}{\alpha}\hat{r}^a\delta_{ab}A^s_{\mu\nu}. \qquad (5.7)$$

Notice that since $\boldsymbol{\Lambda}$ is symmetric in the indices μ, ν, the antisymmetric parts in the first two lines of the expression for $\mathcal{R}$ cancel, and symmetric and antisymmetric parts in the third line decouple.

In order to facilitate the spin average, we now introduce the following representation of unity

$$1 = \int \prod_{a,b} dq_{ab}\, \delta(q_{ab} - q_{ab}(\boldsymbol{\sigma})) = \left(\frac{N}{2\pi}\right)^{n^2} \int \prod_{a,b} dq_{ab}\, d\hat{q}_{ab}\; e^{iN \sum_{a,b} \hat{q}_{ab}(q_{ab} - q_{ab}(\boldsymbol{\sigma}))}$$

Hence the noise distributions (5.5) become

$$\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] \sim \lim_{n\to 0} \int d\hat{\mathbf{r}}\, d\hat{\mathbf{m}}\, d\mathbf{q}\, d\hat{\mathbf{q}}\, e^{N\Psi(\hat{\mathbf{r}},\hat{\mathbf{m}},\mathbf{q},\hat{\mathbf{q}})} \int dx\, dy\, e^{ixz^s + iyz^a} \left\langle e^{-\mathcal{R}(x,y) - i\sum_a \sum_{\mu\le c} \zeta^\mu \hat{m}_a^\mu \sigma^a} \right\rangle_{\boldsymbol{\sigma}}$$

where

$$\begin{aligned} \Psi &= i\sum_a (\mathbf{m}\hat{\mathbf{m}}_a + r\hat{r}^a) + i\sum_{a,b} q_{ab}\hat{q}_{ab} + \Omega(\hat{\mathbf{r}}, \mathbf{q}) \\ &\quad + \frac{1}{N}\sum_{k\neq i} \ln \left\langle \left\langle e^{-i\left(\sum_a \sum_{\mu\le c} \xi_k^\mu \hat{m}_a^\mu \sigma_k^a + \sum_{a,b} \sigma_k^a \hat{q}_{ab} \sigma_k^b\right)} \right\rangle_{\boldsymbol{\xi}} \right\rangle_{\boldsymbol{\sigma}} . \end{aligned} \tag{5.8}$$

The integrals over $\hat{\mathbf{r}}, \hat{\mathbf{m}}, \mathbf{q}, \hat{\mathbf{q}}$ are to be performed by the saddle point method, upon making the mean field approximation to remove site dependence. The saddle point relations in turn allow us to evaluate the (intensive) integrals over x and y determining the noise distribution.

5.2.2 Replica-Symmetric Saddle-Points

We assume that the relevant saddle points are replica symmetric, and take the opportunity to change variables to

$$\begin{aligned} q_{ab} &= \delta_{ab} + q(1-\delta_{ab})\ \forall a,b & \hat{q}_{ab} &= \frac{i}{2}\lambda^2(1-\delta_{ab})\ \forall a,b \\ \hat{r}^a &= \frac{i\alpha}{2}\rho\ \forall a & \hat{m}_a^\mu &= i\mu^\mu\ \forall a \end{aligned} \tag{5.9}$$

(**NB** the diagonal elements $\hat{q}_{aa}$ vanish automatically).

The matrix $\delta_{ab} + q(1-\delta_{ab})$ has one eigenvalue $1 + q(n-1)$ and $n-1$ eigenvalues $1-q$ hence we obtain

$$\ln[\det(\mathbf{q}\otimes \mathbb{1})] + \ln[\det(\boldsymbol{\Lambda})] = \mathrm{Tr}\ln\left[\delta_{\mu\nu} - \rho A^s_{\mu\nu}(1-q+nq)\right] + (n-1)\,\mathrm{Tr}\ln\left[\delta_{\mu\nu} - \rho A^s_{\mu\nu}(1-q)\right]$$

and

$$\begin{aligned} \left(\Lambda^{-1}\right)^{ab}_{\mu\nu} &= \delta_{ab}(1-q)\left(\delta_{\mu\nu} - \rho A^s_{\mu\nu}(1-q)\right)^{-1} \\ &\quad + q\left(\delta_{\mu\lambda} - \rho A^s_{\mu\lambda}(1-q)\right)^{-1}\left(\delta_{\lambda\nu} - \rho A^s_{\lambda\nu}(1-q+nq)\right)^{-1}, \end{aligned}$$

while the last term in Ψ (5.8) becomes

$$\ln\left\langle \int Dw\, e^{n\ln\left[\cosh\left(\sum_{\nu\le c}\xi^\nu\mu^\nu + \lambda w\right)\right]} \right\rangle_{\boldsymbol{\xi}} \qquad \left(Dw = e^{-\frac{1}{2}w^2}\frac{dw}{\sqrt{2\pi}}\right).$$

The saddle point equations, determined by $\frac{\partial\Psi}{\partial\hat{m}_a^\mu}=0$, $\frac{\partial\Psi}{\partial\hat{q}_{ab}}=0$, $\frac{\partial\Psi}{\partial\hat{r}^a}=0$ and $\frac{\partial\Psi}{\partial q_{ab}}=0$ then give

$$\begin{aligned}
m^\mu &= \frac{\left\langle \int Dw\, \xi^\mu \tanh\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\cosh^n\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}}{\left\langle \int Dw\, \cosh^n\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}}\\
q &= \frac{\left\langle \int Dw\, \tanh^2\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\cosh^n\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}}{\left\langle \int Dw\, \cosh^n\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}}\\
r &= \frac{1}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)(1-q+nq)\mathbf{A}^s_{uu}\right)\right.\\
&\qquad \left.\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
&\quad +\frac{1}{\alpha n}\frac{\partial}{\partial\rho}\rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\boldsymbol{\Lambda}^{-1}(q,\rho)\mathbf{A}^s_{uc}\mathbf{m}\\
\lambda^2 &= \frac{\alpha q\rho^2}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
&\quad +\frac{1}{n(n-1)}\rho^2\frac{\partial}{\partial q}\mathbf{m}\cdot\mathbf{A}^s_{cu}\boldsymbol{\Lambda}^{-1}(q,\rho)\mathbf{A}^s_{uc}\mathbf{m}
\end{aligned}\tag{5.10}$$

and the extensive exponent Ψ, evaluated at the RS saddle point, is given by

$$\begin{aligned}
\frac{1}{n}\Psi_{RS}(\boldsymbol{\mu},\rho,q,\lambda) &= -\mathbf{m}\cdot\boldsymbol{\mu}-\frac{1}{2}\alpha\rho-\frac{1}{2}(n-1)\lambda^2 q-\frac{1}{2nN}\mathrm{Tr}\left[\ln\left(\mathbb{1}-\rho\mathbf{A}^s_{uu}(1-q+nq)\right)\right]\\
&\quad -\frac{n-1}{2nN}\mathrm{Tr}\ln\left[\mathbb{1}-\rho\mathbf{A}^s_{uu}(1-q)\right]+\frac{\rho^2}{n}\mathbf{m}\cdot\mathbf{A}^s\sum_{a,b}\boldsymbol{\Lambda}^{-1\,ab}\mathbf{A}^s\mathbf{m}\\
&\quad +\frac{1}{n}\ln\left\langle\int Dw\; e^{n\ln\left[\cosh\left(\sum_{\mu\le c}\xi^\mu\mu^\mu+\lambda w\right)\right]}\right\rangle_{\boldsymbol{\xi}}.
\end{aligned}\tag{5.11}$$

$\boldsymbol{\Lambda}^{-1}$ has the general form $\mathbf{B}\delta_{ab}+\mathbf{C}$. Such matrices have the properties

$$\lim_{n\to0}\frac{1}{n}\sum_{a,b}\left[\mathbf{B}\delta_{ab}+\mathbf{C}\right]=\mathbf{B}\qquad\qquad \lim_{n\to0}\frac{1}{n(n-1)}\sum_{a,b}\left[\mathbf{B}\delta_{ab}+\mathbf{C}\right]=-\mathbf{B}.$$

Therefore in the limit $n\to0$ our saddle point equations become

$$\begin{aligned}
m^\mu &= \left\langle \xi^\mu\int Dw\,\tanh\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}\\
q &= \left\langle \int Dw\,\tanh^2\left(\sum_{\nu\le c}\xi^\nu\mu^\nu+\lambda w\right)\right\rangle_{\boldsymbol{\xi}}\\
r &= \frac{1}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)^2\mathbf{A}^s_{uu}\right)\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
&\quad +\frac{1-q}{\alpha}\frac{\partial}{\partial\rho}\rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\left(\mathbb{1}-\rho\mathbf{A}^s_{uu}(1-q)\right)^{-1}\mathbf{A}^s_{uc}\mathbf{m}\\
\lambda^2 &= \frac{\alpha q\rho^2}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
&\quad +\rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-2}\mathbf{A}^s_{uc}\mathbf{m}
\end{aligned}\tag{5.12, 5.13}$$

These equations are to be solved for $\boldsymbol{\Lambda}$ positive definite, which is necessary for the Gaussian integral to be well defined, possibly placing restrictions on the eigenvalues of $\mathbf{A}$.

5.2.3 Shape of the Intrinsic Noise Distribution

The shape of the noise distribution is now determined by the intensive terms in the integrand, with the values of $\boldsymbol{\mu}, q, \rho, \lambda$ evaluated in the saddle point. We leave the details of the calculation of $\mathcal{R}$ to Appendix B.1 the result is

$$\begin{aligned}
\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] &\sim \int dx\, dy\, e^{ixz^s+iyz^a} \left\langle e^{-\mathcal{R}(x,y)+\sum_{\nu\leq c}\zeta^\nu\mu^\nu\sum_a\sigma^a}\right\rangle_{\boldsymbol{\sigma}} \qquad (5.14)\\
&\sim \int \frac{dx}{2\pi}\frac{dy}{2\pi}\, e^{ixz^s+iyz^a}\, e^{-\frac{1}{2}y^2\mathcal{G}} \times \\
&\left\langle e^{\sum_{\nu\leq c}\zeta^\nu\mu^\nu\sum_a\sigma^a - ix\mathcal{B} - ix\rho\left(\alpha\tilde{r}\sigma^1+\frac{\tilde{\lambda}^2}{\rho^2}\sum_{a>1}\sigma^a\right)-x^2\mathcal{E}-i\rho x\sum_a\mathcal{F}\sigma^a}\int Dw\, e^{w\Gamma\sum_a\sigma^a}\right\rangle_{\boldsymbol{\sigma}},
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{G} &= -\alpha\tilde{p} - 2\rho\mathbf{m}\cdot\mathbf{A}^a_{cu}\mathbf{A}^a_{uu}\mathbf{A}^{s\,-1}_{uu}\left(\tilde{\mathcal{R}}+(n-1)\mathcal{A}\right)\mathbf{m} - \rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\tilde{\mathcal{P}}\mathbf{A}^s_{uc}\mathbf{m} - \mathbf{m}\cdot\mathbf{A}^a_{cu}\mathbf{A}^a_{uc}\mathbf{m}\\
\mathcal{B} &= \rho\boldsymbol{\zeta}\cdot\left(\tilde{\mathcal{R}}_{cc}+(n-1)\mathcal{A}_{cc}\right)\mathbf{m}\\
\mathcal{E} &= \frac{1}{2}\mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uv}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}+\frac{\alpha\tilde{r}}{2}\\
\mathcal{F} &= \mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uc}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}\\
\Gamma^2 &= \lambda^2 = \tilde{\lambda}^2+\rho^2\mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uc}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}\\
\frac{\tilde{\lambda}^2}{\rho^2} &= \frac{1}{N}\sum_{\mu,\nu,\lambda>c}A^s_{\lambda\mu}\left(\Lambda^{-1}\right)^{a\neq b}_{\mu\nu}A^s_{\nu\lambda} = \frac{1}{N}\operatorname{Tr}\mathcal{A}\\
&= \frac{\alpha q}{p}\operatorname{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
\alpha\tilde{r} &= \frac{1}{N}\sum_{\mu,\nu,\lambda>c}A^s_{\lambda\mu}\left(\Lambda^{-1}\right)^{a=b}_{\mu\nu}A^s_{\nu\lambda} = \frac{1}{N}\operatorname{Tr}\tilde{\mathcal{R}}\\
&= \frac{1}{p}\operatorname{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)(1-q+nq)\mathbf{A}^s_{uu}\right)\right.\\
&\qquad\left.\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
\alpha\tilde{p} &= \frac{1}{N}\sum_{\mu,\nu,\lambda>c}A^a_{\lambda\mu}\left(\Lambda^{-1}\right)^{11}_{\mu\nu}A^a_{\nu\lambda} = \frac{1}{N}\operatorname{Tr}\tilde{\mathcal{P}}. \qquad (5.15)
\end{aligned}$$

Carrying out the spin average in (5.14) gives

$$\begin{aligned}
&\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] = \qquad (5.16)\\
&\lim_{n\to 0}\int\frac{dx}{2\pi}\frac{dy}{2\pi}\, e^{ixz^s+iyz^a}\, e^{-\frac{1}{2}y^2\mathcal{G}}\, e^{-x^2\mathcal{E}-ix\mathcal{B}}\times\\
&\int Dw\,\cosh^{n-1}\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu - ix\frac{\tilde{\lambda}^2}{\rho} - i\rho x\mathcal{F}+w\Gamma\right)\cosh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu - ix\rho\alpha\tilde{r} - i\rho x\mathcal{F}+w\Gamma\right).
\end{aligned}$$

We can carry out the integral over y immediately (assuming $\mathcal{G}$ to be positive), however, the $\cosh^{n-1}(\ldots)$ term will cause a divergence for $n\to 0$ since it has an imaginary argument. However, if we define

$$\Delta = \alpha\rho\tilde{r} - \frac{\tilde{\lambda}^2}{\rho} = \frac{\alpha\rho(1-q)}{p}\operatorname{Tr}\left[\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\mathbf{A}^s_{uu}\right] \qquad (5.17)$$

the cosh terms in (5.16) become

$$\cos(\Delta x)\cosh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu - ix\frac{\tilde{\lambda}^2}{\rho} - i\rho x\mathcal{F} + w\Gamma\right) - i\sin(\Delta x)\sinh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu - ix\frac{\tilde{\lambda}^2}{\rho} - i\rho x\mathcal{F} + w\Gamma\right),$$

so that the possible divergence is eliminated by the contour shift $y' = w - \frac{ix}{\Gamma}\left(\frac{\tilde{\lambda}^2}{\rho} + \rho\mathcal{F}\right)$, giving

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] &= (2\pi\mathcal{G})^{-\frac{1}{2}} e^{-\frac{1}{2}z^a\mathcal{G}^{-1}z^a}\int\frac{dx}{2\pi}e^{ix(z^s-\mathcal{B})-\frac{\Delta}{2\rho}x^2} \\ &\times\frac{1}{2}\int Dy' e^{-ixy'\frac{\Gamma}{\rho}}\left[\left(e^{i\Delta x}+e^{-ix\Delta}\right)-\left(e^{i\Delta x}-e^{-ix\Delta}\right)\tanh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu + y'\Gamma\right)\right],\end{aligned} \tag{5.18}$$

where we have used the relations $\frac{\tilde{\lambda}^2}{\rho} + \rho\mathcal{F} = \frac{\Gamma^2}{\rho}$ and $\mathcal{E} - \frac{1}{2}\left(\frac{\tilde{\lambda}^2}{\rho} + \rho\mathcal{F}\right) = \frac{\Delta}{2\rho}$. The remaining integral over x is easily performed using $1 + \frac{\Gamma^2}{\rho\Delta} = \frac{2\mathcal{E}\rho}{\Delta}$ and $-\frac{1}{2\mathcal{E}} + \frac{\rho}{\Delta} = \frac{\Gamma^2}{2\rho\mathcal{E}\Delta}$ and (again) a rescaling $y = \sqrt{\frac{2\mathcal{E}\rho}{\Delta}}\left(y' - \frac{\Gamma}{2\mathcal{E}\rho}(z^s \pm \Delta - \mathcal{B})\right)$, giving the final result

$$\begin{aligned}\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] &= \\ &\frac{e^{-\frac{1}{2}z^a\mathcal{G}^{-1}z^a}}{4\pi\sqrt{2\mathcal{G}\mathcal{E}}}\left[e^{-\frac{(z^s-\mathcal{B}+\Delta)^2}{4\mathcal{E}}}\left(1-\int Dy\tanh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu + y\Gamma\sqrt{\frac{\Delta}{2\mathcal{E}\rho}} + (z^s+\Delta-\mathcal{B})\frac{\Gamma^2}{2\mathcal{E}\rho}\right)\right.\right. \\ &\left.\left.+e^{-\frac{(z^s-\mathcal{B}-\Delta)^2}{4\mathcal{E}}}\left(1+\int Dy\tanh\left(\sum_{\nu\leq c}\zeta^\nu\mu^\nu + y\Gamma\sqrt{\frac{\Delta}{2\mathcal{E}\rho}} + (z^s-\Delta-\mathcal{B})\frac{\Gamma^2}{2\mathcal{E}\rho}\right)\right)\right].\end{aligned} \tag{5.19}$$

Figures 5.1 and 5.2 show the noise distribution as a function of z^s for 4 values of z^a and $\boldsymbol{\zeta} = \begin{pmatrix}1\\1\end{pmatrix}$ as predicted by our theory (solid line) and as measured using spin simulations (histograms). The simulations were carried out for a system of 40,000 spins (neurons), using the block diagonal matrix (5.29) (see Section 5.3). Figure 5.1 corresponds to the values of parameters: $a = 1$ (giving rise to cyclical behaviour), $\alpha = 0.0256$, $m^1 = -0.2173$, $m^2 = -0.5511$ and $r = 3.700631$, whilst figure (5.2) corresponds the parameter values $a = 2$ (converging towards the fixed point $\mathbf{m} = 0, r > 1$), $\alpha = 0.0512$, $m^1 = -0.00735$, $m^2 = 0.00435$ and $r = 4.173877$.

From Figures 5.1 and 5.2 we can see that in the first case when cyclical behaviour is observed our model appears to underestimate the distribution at extreme values of z^a whilst overestimating for z^a around zero. This could be partially caused by the finite size of the bins of the histogram in the z^a direction, and by other finite size effects. (The number of spins needed to obtain good statistics given that one must specify (i) a value of $\boldsymbol{\zeta}$, (ii) an interval of z^s, (iii) an interval of z^a, is unfeasibly large.)

Both sets of distributions 'look' approximately Gaussian. In fact the results of numerous simulations have revealed that it extremely difficult to force a network with non-symmetric interactions to iterate towards a state where the noise distribution is strongly non-Gaussian. In contrast non-retrieval states ($\mathbf{m} = 0$) in the symmetric case [44] have double peaked noise distributions. Although the theory does not describe the shape of the distributions exactly in these cases [83], it shows a good qualitative fit to the clearly non-Gaussian shape, in contrast to the time dependent Ginzburg-Landau approaches [80, 81]. The lack of non-Gaussian shaped noise distributions in networks with non-symmetric interactions implies that symmetry plays an important role in building up the correlations between the system state and the uncondensed patterns

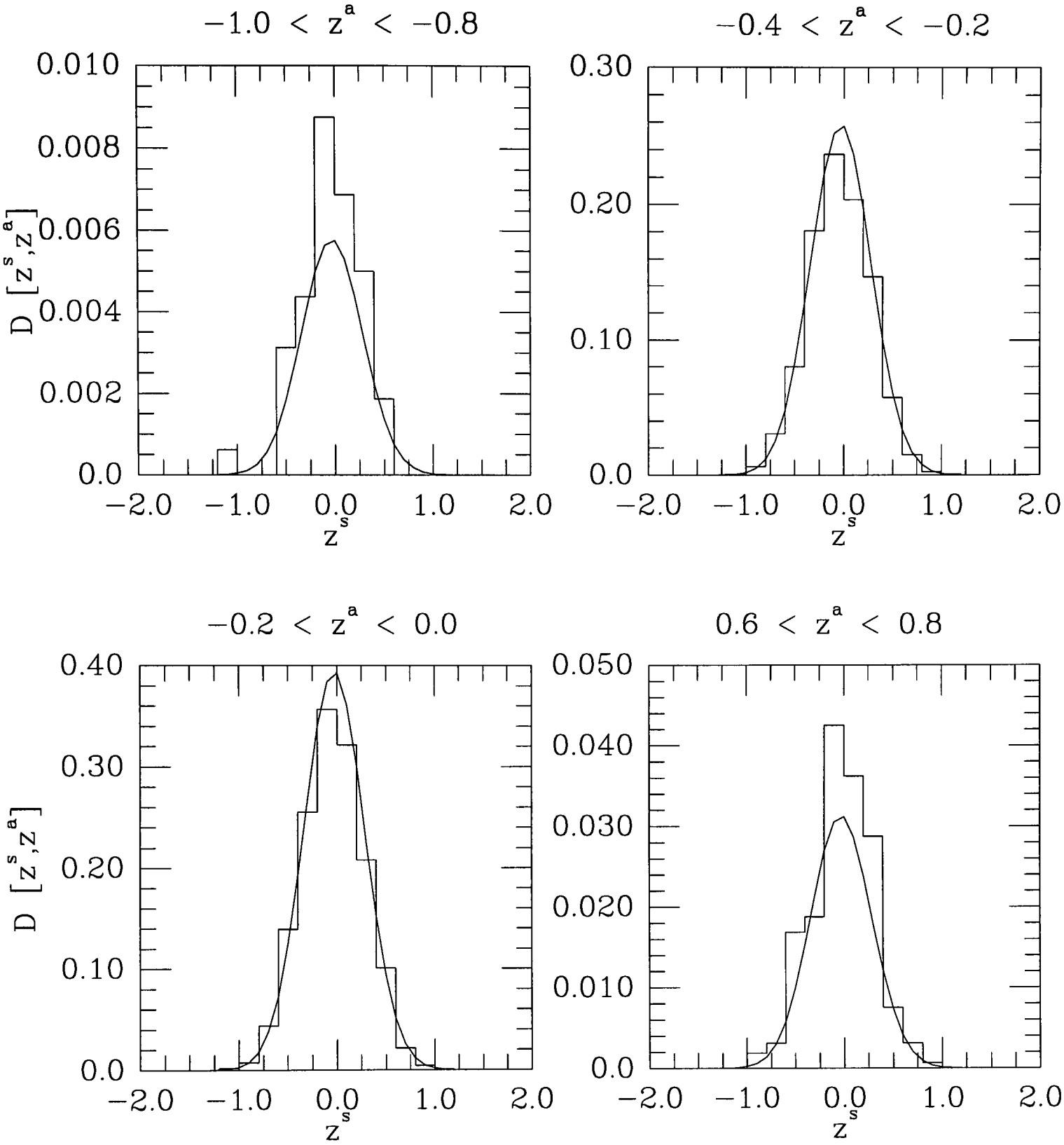


Figure 5.1: Plots showing sections through the noise distribution at values of z^a near $z^a = -0.9$, $z^a = -0.3$, $z^a = -0.1$ and $z^a = 0.7$, for a model with block diagonal $\mathbf{A}$ (5.29) (see section 5.3) with $a = 1$ (giving rise to cyclical behaviour), $\alpha = 0.0256$, $m^1 = -0.2173$, $m^2 = -0.5511$ and $r = 3.700631$. The solid lines are theoretical predictions, whilst the histograms are taken from simulations by counting the number of sites with z^s, z^a in a range ± 0.1 from the centre.

at the microscopic level, or equivalently that the macroscopic dynamics does not lead the system to a region of $(\mathbf{m}, r)$ space where the noise distribution is strongly non-Gaussian ($q \to 1$).

5.2.4 Special Cases and Critical Surfaces

The noise distribution (5.19) in general can only be calculated numerically. For specific choices of $\mathbf{A}$ however, it simplifies considerably. In the absence of condensed-uncondensed couplings for instance, $\mathcal{G} = -\alpha\tilde{p}$, $\mathcal{B} = 0$, $\mathcal{E} = \frac{\alpha\tilde{r}}{2}$, $\mathcal{F} = 0$ and $\Gamma^2 = \tilde{\lambda}^2$. In the absence of antisymmetric components to $\mathbf{A}$ we find that $\mathcal{G} = 0$, and the Gaussian function of z^a becomes $\delta(z^a)$. For $\mathbf{A} = \mathbb{1}$ (the Hopfield model) the noise distribution reduces to the result given in [45].

There are also several regions (in the space of $(\mathbf{m}, r)$) where the noise distribution takes on a

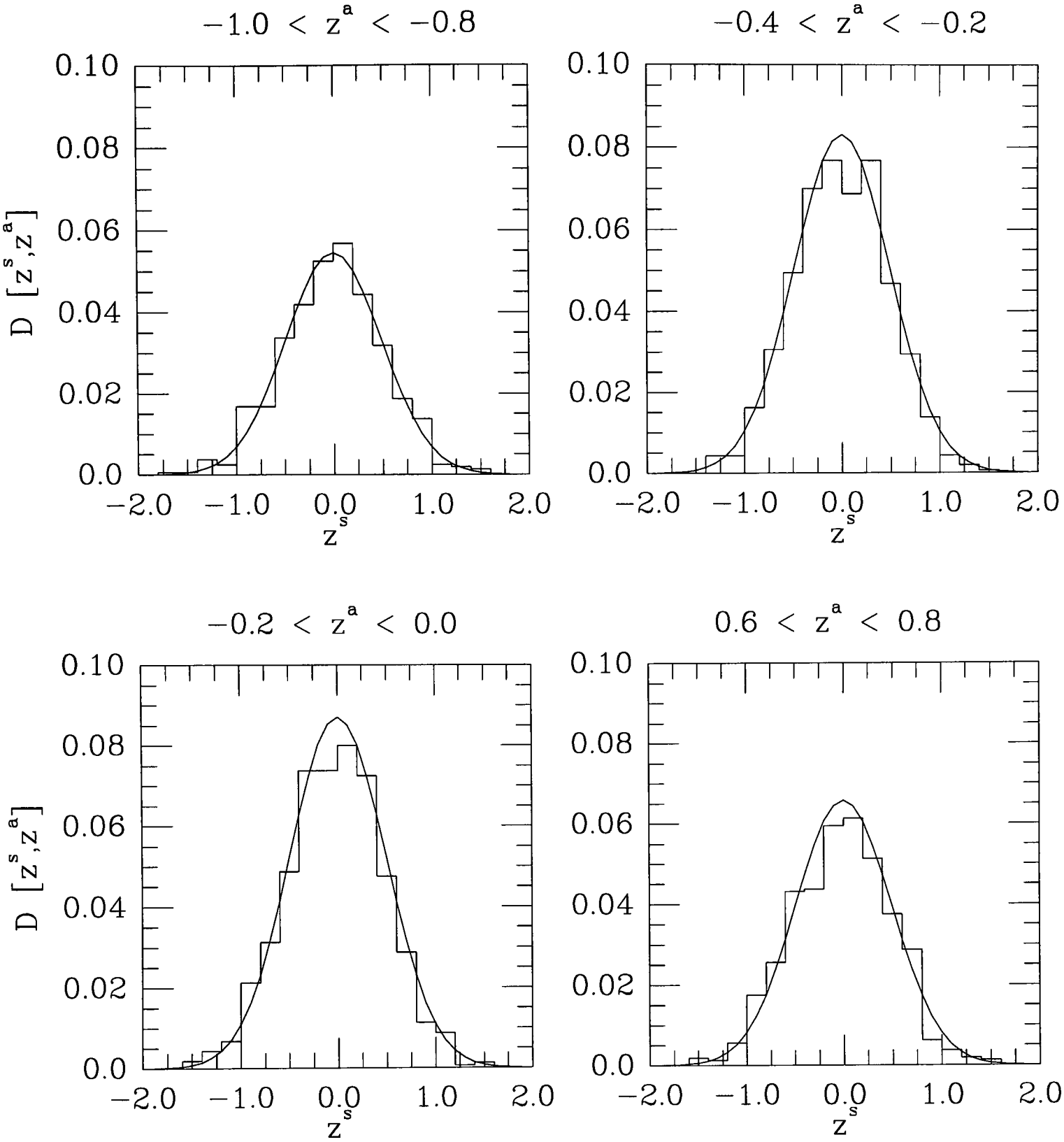


Figure 5.2: Plots showing sections through the noise distribution at values of z^a near $z^a = -0.9$, $z^a = -0.3$, $z^a = -0.1$ and $z^a = 0.7$, for a model with block diagonal $\mathbf{A}$ (5.29) with $a = 2$ (see section 5.3) (converging towards the fixed point $\mathbf{m} = 0, r > 1$), $\alpha = 0.0512$, $m^1 = -0.00735$, $m^2 = 0.00435$ and $r = 4.173877$. The solid lines are theoretical predictions, whilst the histograms are taken from simulations by counting the number of sites with z^s, z^a in a range ± 0.1 from the centre.

special form, or where significant transitions occur:

1. Gaussian Noise

   In previous papers the noise distribution has often been assumed to have a Gaussian shape [81]. Here we can see that this requires $\Delta = 0$, which implies that

$$q\delta_{\mu\nu} = \delta_{\mu\nu} - \rho(1-q)^2 A^s_{\mu\nu}$$

   requiring $A^s_{\mu\nu} = A\delta_{\mu\nu}$ and $\rho = \frac{1}{A(q-1)}$. Tracing this back through the saddle point equations leads to the requirement $r = 1$, which explains why the Gaussian noise approximation is only reasonable in the region of $(\mathbf{m}, r)$ space where retrieval occurs.

2. $q = 0$, $\mathbf{m} = 0$

From the saddle point equations we can see that $\mathbf{m} = 0$ implies $\boldsymbol{\mu} = 0$. Expanding the equation for q around $q = 0$ therefore gives

$$q \simeq \lambda^2 \int Dw\, w^2 \simeq \frac{\alpha q \rho^2}{p} \mathrm{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1} - \rho\mathbf{A}^s_{uu}\right)^{-2}\right].$$

The only solution is $q = 0$ if $1 > \frac{\alpha\rho^2}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1} - \rho\mathbf{A}^s_{uu}\right)^{-2}\right]$. The saddle point equation for r at $q = 0$: $r = r_c = \frac{1}{p}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\left(\mathbb{1} - \rho\mathbf{A}^s_{uu}\right)^{-1}\right]$ subsequently defines the value of r below which, $q = \mathbf{m} = 0$ is the only solution.

In this region of phase space ($\mathbf{m} = 0$, $r < r_c$) the noise distribution is a sum of two Gaussians:

$$\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] = \frac{e^{\frac{(z^a)^2}{2\alpha\tilde{p}}}}{4\pi\alpha\sqrt{-\tilde{r}\tilde{p}}}\left(e^{-\frac{(z^s+\alpha\rho\tilde{r})^2}{2\alpha\tilde{r}}} + e^{-\frac{(z^s-\alpha\rho\tilde{r})^2}{2\alpha\tilde{r}}}\right). \tag{5.20}$$

3. Freezing Line

The freezing line occurs when the number of microscopic states contributing to the macroscopic state goes from an extensive large number to an extensively small number (equivalent to a transition to a macroscopic state with negative entropy):

$$\lim_{N\to\infty}\frac{1}{N}\ln\left[\sum_{\boldsymbol{\sigma}}\delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))\,\delta(r - r(\boldsymbol{\sigma}))\right] = 0. \tag{5.21}$$

This is the first sign that something has gone wrong with the replica method. We can relate the freezing line to the saddle point problem encountered in the calculation of the intrinsic noise distribution. In order to average over the patterns $\boldsymbol{\xi}$ we use $\ln[Z] = \lim_{n\to 0}\frac{1}{n}(Z^n - 1)$ to cast (5.21) into the familiar form

$$\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{nN}\left(\left\langle\left\langle\prod_{a=1}^{n}\delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}^a))\,\delta(r - r(\boldsymbol{\sigma}^a))\right\rangle_{\boldsymbol{\xi}}\right\rangle_{\boldsymbol{\sigma}} - 1\right) = \lim_{n\to 0}\frac{1}{n}\Psi = -\ln[2].$$

Using (5.8) this leads in the RS approximation to

$$\begin{aligned} -\ln[2] \;=\; & -\mathbf{m}\cdot\boldsymbol{\mu} - \frac{1}{2}\alpha\rho + \frac{1}{2}\lambda^2 q + \frac{\rho q}{2N}\mathrm{Tr}\left[\mathbf{A}^s_{uu}\left(\mathbb{1} - \rho\mathbf{A}^s_{uu}(1-q)\right)^{-1}\right] \\ & - \frac{1}{2N}\mathrm{Tr}\left[\log\left(\mathbb{1} - \rho\mathbf{A}^s_{uu}(1-q)\right)\right] + \rho^2(1-q)\mathbf{m}\cdot\mathbf{A}^s\left(\mathbb{1} - \rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\mathbf{A}^s_{uc}\mathbf{m} \\ & + \left\langle\int Dy\,\ln\left[\cosh\left(\sum_{\mu\le c}\xi^\mu\mu^\mu + \lambda y\right)\right]\right\rangle_{\boldsymbol{\xi}}. \end{aligned} \tag{5.22}$$

This freezing line defined by (5.22) occurs close to, but not exactly on, the $q = 1$ line. The above equation is to be solved numerically along with the saddle point equations.

4. $q = 1$

In the $q \simeq 1$ region of $(\mathbf{m}, r)$ space we can use the asymptotic form of the saddle point equations to derive a simpler expression for the noise distribution. We expand the saddle

point equations in $\epsilon = 1 - q$, in order to determine where the $q = 1$ line occurs. We first note that $\lambda \propto \rho$. From the saddle point equation relating r and ρ we note that $\rho \propto \epsilon^{-1}$ and hence $\lambda \propto \epsilon^{-1}$. Therefore $\tanh(\lambda y + \boldsymbol{\xi}\cdot\boldsymbol{\mu}) = \mathrm{sgn}(\lambda y + \boldsymbol{\xi}\cdot\boldsymbol{\mu})$, which we use to simplify our saddle point equations for $\mathbf{m}$ and q:

$$\begin{aligned}
m^\mu &= \left\langle \xi^\mu \mathrm{erf}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\mu}}{\sqrt{2}\lambda}\right)\right\rangle_{\boldsymbol{\xi}} \\
-\epsilon &= q - 1 = \left\langle \int Dy\, \tanh^2(\lambda y + \boldsymbol{\xi}\cdot\boldsymbol{\mu}) - 1 \right\rangle_{\boldsymbol{\xi}} \\
&= \left\langle \left[\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}\frac{1}{\lambda}\tanh(\lambda y + \boldsymbol{\xi}\cdot\boldsymbol{\mu})\right]\right\rangle_{\boldsymbol{\xi}} - \frac{2}{\lambda}\left\langle \int \frac{dy}{\sqrt{2\pi}}\, y\, e^{-\frac{1}{2}y^2}\tanh(\lambda y + \boldsymbol{\xi}\cdot\boldsymbol{\mu})\right\rangle_{\boldsymbol{\xi}} \\
&= -\frac{2}{\lambda}\left\langle \int_{-\frac{\boldsymbol{\xi}\cdot\boldsymbol{\mu}}{\lambda}}^{\infty} y\, Dy \;-\; \int_{-\infty}^{-\frac{\boldsymbol{\xi}\cdot\boldsymbol{\mu}}{\lambda}} y\, Dy \right\rangle_{\boldsymbol{\xi}} + \mathcal{O}(\epsilon^2) \\
&= -\frac{4}{\lambda\sqrt{2\pi}}\left\langle e^{-\frac{1}{2}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\mu}}{\lambda}\right)^2}\right\rangle_{\boldsymbol{\xi}} + \mathcal{O}(\epsilon^2). \qquad (5.23)
\end{aligned}$$

These equations are to be used along with the relationship between λ and ρ, to define the $q = 1$ line in the $(\mathbf{m}, r)$ plane. However for more than one condensed pattern the average over $\boldsymbol{\xi}$ makes the expression for $\mathbf{m}$ non-invertible, and the relation between λ and ρ is complicated for general $\mathbf{A}$.

Since $\tilde{\lambda}$ and ρ diverge faster than the term containing the integration variable y, as $q \to 1$, the noise distribution becomes

$$\begin{aligned}
\mathcal{D}_{\boldsymbol{\zeta}}[z^s, z^a] = \frac{e^{-\frac{1}{2}z^a\mathcal{G}^{-1}z^a}}{\sqrt{8\pi^2\mathcal{G}\mathcal{E}}}\Bigg[& e^{-\frac{(z^s-\mathcal{B}+\Delta)^2}{4\mathcal{E}}}\Theta\left(-\frac{1}{\epsilon}\sum_{\nu\le c}\zeta^\nu\mu^\nu - (z^s+\Delta-\mathcal{B})\frac{\Gamma^2}{2\epsilon\mathcal{E}\rho}\right) \\
& + e^{-\frac{(z^s-\mathcal{B}-\Delta)^2}{4\mathcal{E}}}\Theta\left(\frac{1}{\epsilon}\sum_{\nu\le c}\zeta^\nu\mu^\nu + (z^s-\Delta-\mathcal{B})\frac{\Gamma^2}{2\epsilon\mathcal{E}\rho}\right)\Bigg], \qquad (5.24)
\end{aligned}$$

where $\epsilon = 1 - q$ and $\mathcal{G}, \mathcal{E}, \mathcal{B}, \rho, \Delta$ are evaluated at $q = 1$. Note, however, that the above analysis extends the noise distribution into the region where replica symmetry is unstable.

5. AT Surface

The A-T surface [50] signals the instability of the replica symmetric solution via bifurcations of the form

$$q_{ab} \to q + \delta q_{ab} \qquad \hat{q}_{ab} \to \frac{i}{2}\lambda^2 + i\delta\hat{q}_{ab} \qquad (5.25)$$

with $\sum_{\alpha\neq\beta}\delta q_{ab} = 0$, $\sum_{a\neq b}\delta\hat{q}_{ab} = 0$ (the so-called replicon mode). Using (5.8) we can expand about the replica symmetric solution in first non-vanishing orders giving

$$\Psi - \Psi_{RS} \simeq \frac{1}{2}\sum_{\alpha\neq\beta}\sum_{c\neq d}\delta q_{ab}\delta q_{cd}\frac{\partial^2\Omega}{\partial q_{ab}\partial q_{cd}} + \frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\hat{q}_{ab}\delta\hat{q}_{cd}\frac{\partial^2\Phi}{\partial\hat{q}_{ab}\partial\hat{q}_{cd}} - \sum_{a\neq b}\delta q_{ab}\delta\hat{q}_{ab},$$

where

$$\begin{aligned}
\Omega &= \frac{1}{N}\ln\int d\mathbf{z}\,\mathcal{D}(\mathbf{z},\mathbf{q})\, e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z} + \rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}} \\
\Phi &= \ln\left\langle\left\langle e^{\boldsymbol{\xi}\cdot\boldsymbol{\mu}\sum_a\sigma^a + \frac{1}{2}\sum_{a,b\neq a}(\lambda^2+2\delta\hat{q}_{ab})\sigma^a\sigma^b}\right\rangle_{\boldsymbol{\xi}}\right\rangle_{\boldsymbol{\sigma}}.
\end{aligned}$$

The stability of the RS solution can be written as an eigenvalues problem

$$\begin{pmatrix} \partial^2\Omega & -1 \\ -1 & \partial^2\Phi \end{pmatrix} \begin{pmatrix} \delta q_{ab} \\ \delta \hat{q}_{ab} \end{pmatrix} = \Lambda \begin{pmatrix} \delta q_{ab} \\ \delta \hat{q}_{ab} \end{pmatrix}$$

The critical stability will occur when the eigenvalue Λ is zero, so the A-T surface defines a solution of $\partial^2\Omega\ \partial^2\Phi = 1$. We leave the details of the calculation to Appendix B.2, where it is shown that the A-T surface is given by

$$1 = \frac{\rho^2}{N}\left(\frac{\rho^2}{p}\left(\mathrm{Tr}\,\tilde{\mathcal{R}}\right)^2 + 2\rho\,\mathrm{Tr}\left(\mathbf{A}^s\tilde{\mathcal{R}}\right) + \mathrm{Tr}\left(\mathbf{A}^s\mathbf{A}^s\right)\right)\left\langle \int Dy \cosh^{-4}\left(\boldsymbol{\xi}\cdot\boldsymbol{\mu} + \lambda y\right)\right\rangle_{\boldsymbol{\xi}} \tag{5.26}$$

($\mathrm{Tr}\,\tilde{\mathcal{R}} = N\Delta$.) Replica symmetry is stable as long as the right hand side is greater than one.

5.3 Replica-Symmetric Order Parameter Flow

Having derived an expression for the noise distribution we now obtain differential equations for the flow of the order parameters $(\mathbf{m}, r)$ by inserting (5.19) into (5.2). Upon making specific choices for the matrix $\mathbf{A}$ we can test the validity of our theory by comparing results from solving the macroscopic equations with results of performing microscopic spin simulations. As pointed out in [59, 83] the present theory appears not to be exact due to the assumption of equipartitioning within the energy subshells. However as with the previous study [44], we expect the equations to capture the essential features of the flow.

We can write the dynamic equations in a more compact form by changing variables to

$$\tilde{z}^a = \mathcal{G}^{-\frac{1}{2}} z^a \qquad\qquad x' = \frac{z^s - \mathcal{B} \pm \Delta}{(2\mathcal{E})^{\frac{1}{2}}}.$$

The flow equations then become

$$\begin{aligned} \frac{d\mathbf{m}}{dt} &= \left\langle \int D\tilde{z}^a \int Dx' \int Dy'\, \boldsymbol{\zeta}\, M_{\boldsymbol{\zeta}}(\tilde{z}^a, x', y') \right\rangle_{\boldsymbol{\zeta}} - \mathbf{m} \\ \frac{1}{2}\frac{dr}{dt} &= \left\langle \int D\tilde{z}^a \int Dx' \int Dy'\, R_{\boldsymbol{\zeta}}(\tilde{z}^a, x', y') \right\rangle_{\boldsymbol{\zeta}} - (r - \frac{1}{p}\,\mathrm{Tr}\,\mathbf{A}) \end{aligned} \tag{5.27}$$

where

$$\begin{aligned} M_{\boldsymbol{\zeta}}(\tilde{z}^a, x', y') = {} & \frac{1}{2}\left(1 - \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + y'\sqrt{\frac{\Gamma^2\Delta}{2\mathcal{E}\rho}} + x'\frac{\Gamma^2}{(2\mathcal{E})^{\frac{1}{2}}\rho}\right)\right)\tanh\beta\left(\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \mathcal{G}^{\frac{1}{2}}\tilde{z}^a + U^-\right) \\ & + \frac{1}{2}\left(1 + \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + y'\sqrt{\frac{\Gamma^2\Delta}{2\mathcal{E}\rho}} + x'\frac{\Gamma^2}{(2\mathcal{E})^{\frac{1}{2}}\rho}\right)\right)\tanh\beta\left(\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \mathcal{G}^{\frac{1}{2}}\tilde{z}^a + U^+\right) \\ R_{\boldsymbol{\zeta}}(\tilde{z}^a, x', y') = {} & \frac{1}{2}\left(1 - \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + y'\sqrt{\frac{\Gamma^2\Delta}{2\mathcal{E}\rho}} + x'\frac{\Gamma^2}{(2\mathcal{E})^{\frac{1}{2}}\rho}\right)\right)U^-\tanh\beta\left(\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \mathcal{G}^{\frac{1}{2}}\tilde{z}^a + U^-\right) \\ & + \frac{1}{2}\left(1 + \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + y'\sqrt{\frac{\Gamma^2\Delta}{2\mathcal{E}\rho}} + x'\frac{\Gamma^2}{(2\mathcal{E})^{\frac{1}{2}}\rho}\right)\right)U^+\tanh\beta\left(\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \mathcal{G}^{\frac{1}{2}}\tilde{z}^a + U^+\right) \\ U^\pm = {} & (2\mathcal{E})^{\frac{1}{2}}x' + \mathcal{B} \mp \Delta \end{aligned} \tag{5.28}$$

5.3.1 Comparison with Numerical Simulations

We make the following specific choice for $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} 1 & a & & & \emptyset & \\ -a & 1 & & & & \\ & & 1 & a & & \\ & & -a & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \quad a \\ & & \emptyset & & & -a \quad 1 \end{pmatrix} \tag{5.29}$$

and assume that the first two patterns are condensed. The choice (5.29), without condensed-uncondensed coupling terms is made purely for computational convenience. Without such terms the saddle points have unique solutions, if such terms were included they could become much more difficult to solve (increasing the already large amount of computer time necessary). By variation of a we are still in a position to investigate the agreement between theory and simulations for both symmetric and asymmetric $\mathbf{A}$.

We notice that for $a = 0$ we obtain the Hopfield model, and as $a \to \infty$ the interactions become predominantly antisymmetric. Because the saddle point equations contain only the symmetric part of the matrix, which is in this case the unit matrix, they simplify considerably

$$\begin{aligned} \begin{pmatrix} m^1 \\ m^2 \end{pmatrix} &= \frac{1}{2}\int D\hat{y}\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\tanh\left(\mu^1+\mu^2+\lambda\hat{y}\right)+\begin{pmatrix} 1 \\ -1 \end{pmatrix}\tanh\left(\mu^1-\mu^2+\lambda\hat{y}\right)\right] \\ q &= \frac{1}{2}\int D\hat{y}\left[\tanh^2\left(\mu^1+\mu^2+\lambda\hat{y}\right)+\tanh^2\left(\mu^1-\mu^2+\lambda\hat{y}\right)\right] \\ \frac{\lambda^2}{\alpha\rho^2} &= \frac{q}{(1-\rho(1-q))^2} \qquad \tilde{r} = r = \frac{1-\rho(1-q)^2}{(1-\rho(1-q))^2} \qquad \mathcal{G} = \alpha a^2 r. \end{aligned} \tag{5.30}$$

The absence of coupling between condensed and uncondensed patterns leads to $\mathcal{B} = 0$, $\Delta = \frac{\alpha\rho(1-q)}{1-\rho(1-q)}$ $2\mathcal{E} = \alpha r$ and $\mathbf{Am} = \begin{pmatrix} m^1 + am^2 \\ -am^1 + m^2 \end{pmatrix}$. From these expressions the dynamics of the order parameters $(\mathbf{m}, r)$ according to (5.27) can be calculated numerically.

In Figures 5.3, 5.4 and 5.5 we show the results of Monte-carlo spin simulations, along with the results of the theory in the $m^1 - m^2$ and $m^1 - r$ planes, comprising of flow lines of the simulations along with arrows indicating the theoretically predicted magnitude and direction of the derivatives $\frac{dm^1}{dm^2}$ and $\frac{dm^1}{dr}$. Each simulation is carried out for system of 40,000 spins with each of the following sets of parameters: a=0, 1, 2; T=0, 1; p=1024, 2048.

As we can see the arrows give a good qualitative description of the direction of the actual flow. Closer analysis of the time derivatives in Figures 5.6 and 5.7, however, shows deviations from the simulations noticeable for states giving rise to non-retrieval (i.e. flowing towards the fixed point $\mathbf{m} = 0, r > 1$), manifested in an overall slowing effect. The derivatives $\frac{dm^2}{dm^1}$ and $\frac{dr}{dm^1}$ however show a better fit. It is not clear whether this is due to the effects of replica symmetry breaking or a manifestation of a more fundamental error within the formulation of the model. Such measurements were made for the Hopfield model [83] where small but significant discrepancies were found between the shape of the noise distribution found from the model, and from simulations. Such effects were also observed in a toy model [46] where an exact solution was possible.

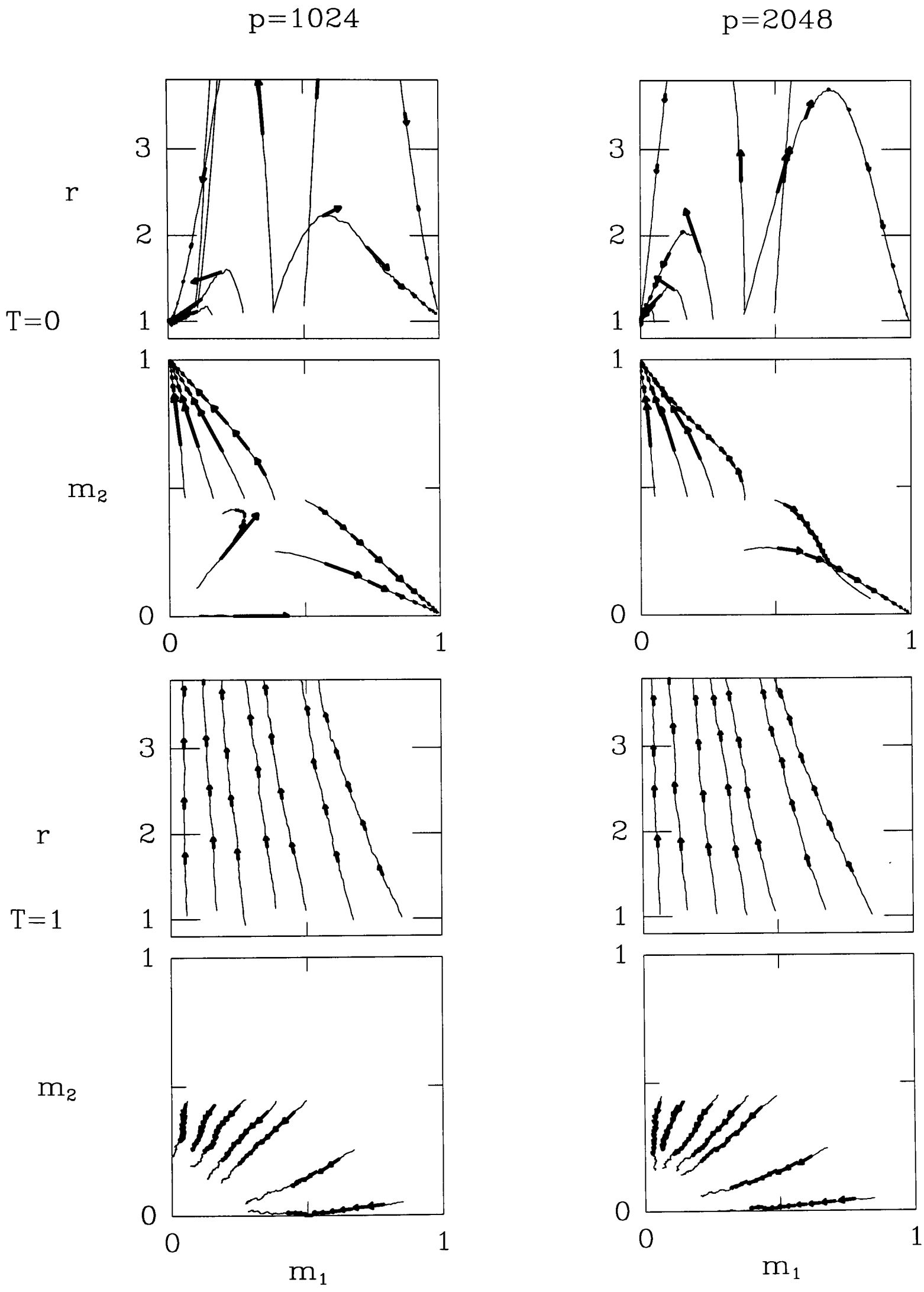


Figure 5.3: Plots showing the evolution of m^1, m^2 and r for a system of 40,000 spins and embedding matrix (5.29) with $a = 0$. The solid lines are the results of spin simulations, and the arrows show the instantaneous values of $\frac{dm^2}{dm^1}$ and $\frac{dr}{dm^1}$ calculated at intervals of 1/2 iteration per spin. The lengths of the arrows are proportional to the magnitude of the derivatives.

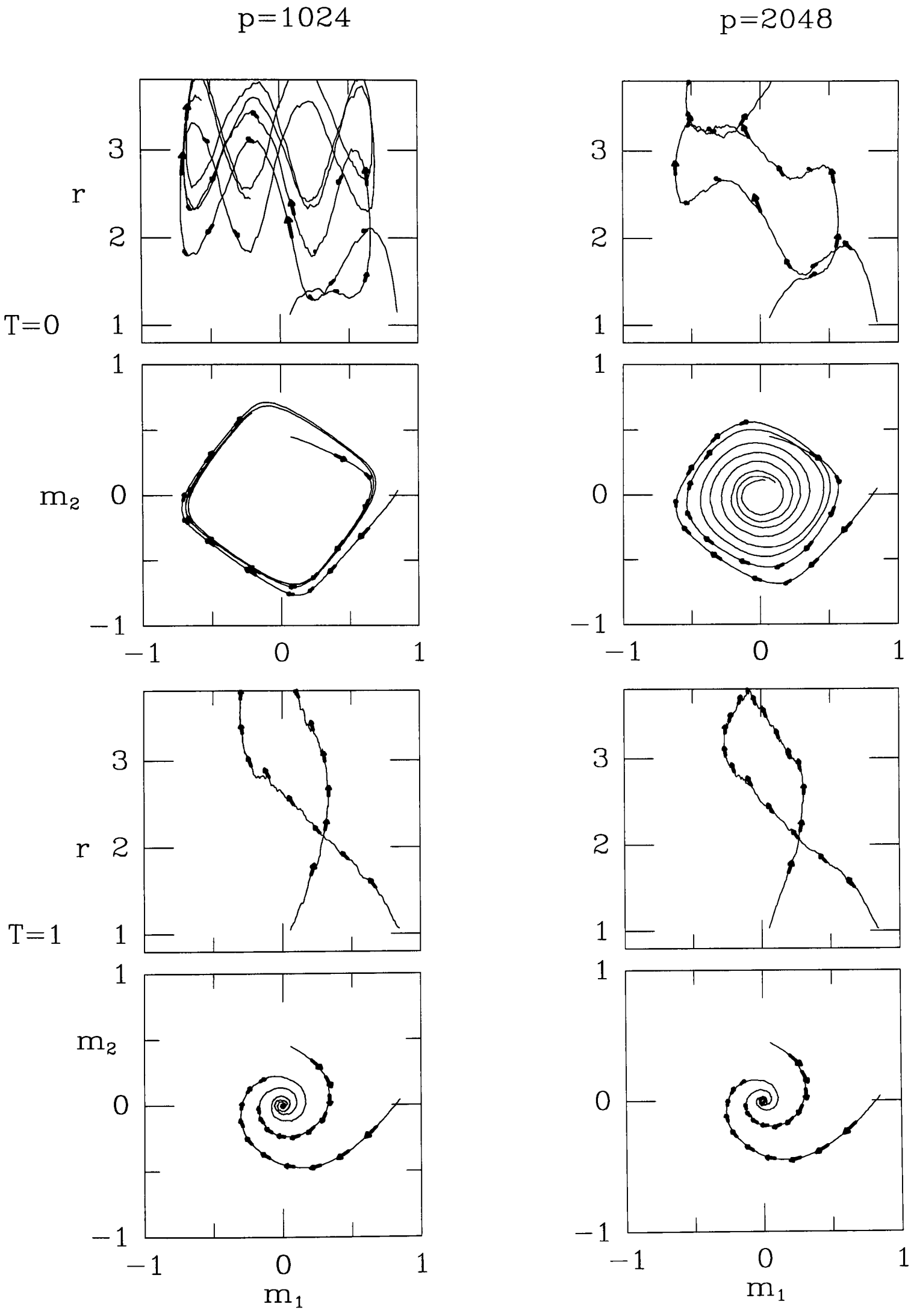


Figure 5.4: Plots showing the evolution of m^1, m^2 and r for a system of 40,000 spins and embedding matrix (5.29) with $\alpha = 1$. The solid lines are the results of spin simulations, and the arrows show the instantaneous values of $\frac{dm^2}{dm^1}$ and $\frac{dr}{dm^1}$ calculated at intervals of 1/2 iteration per spin. The lengths of the arrows are proportional to the magnitude of the derivatives.

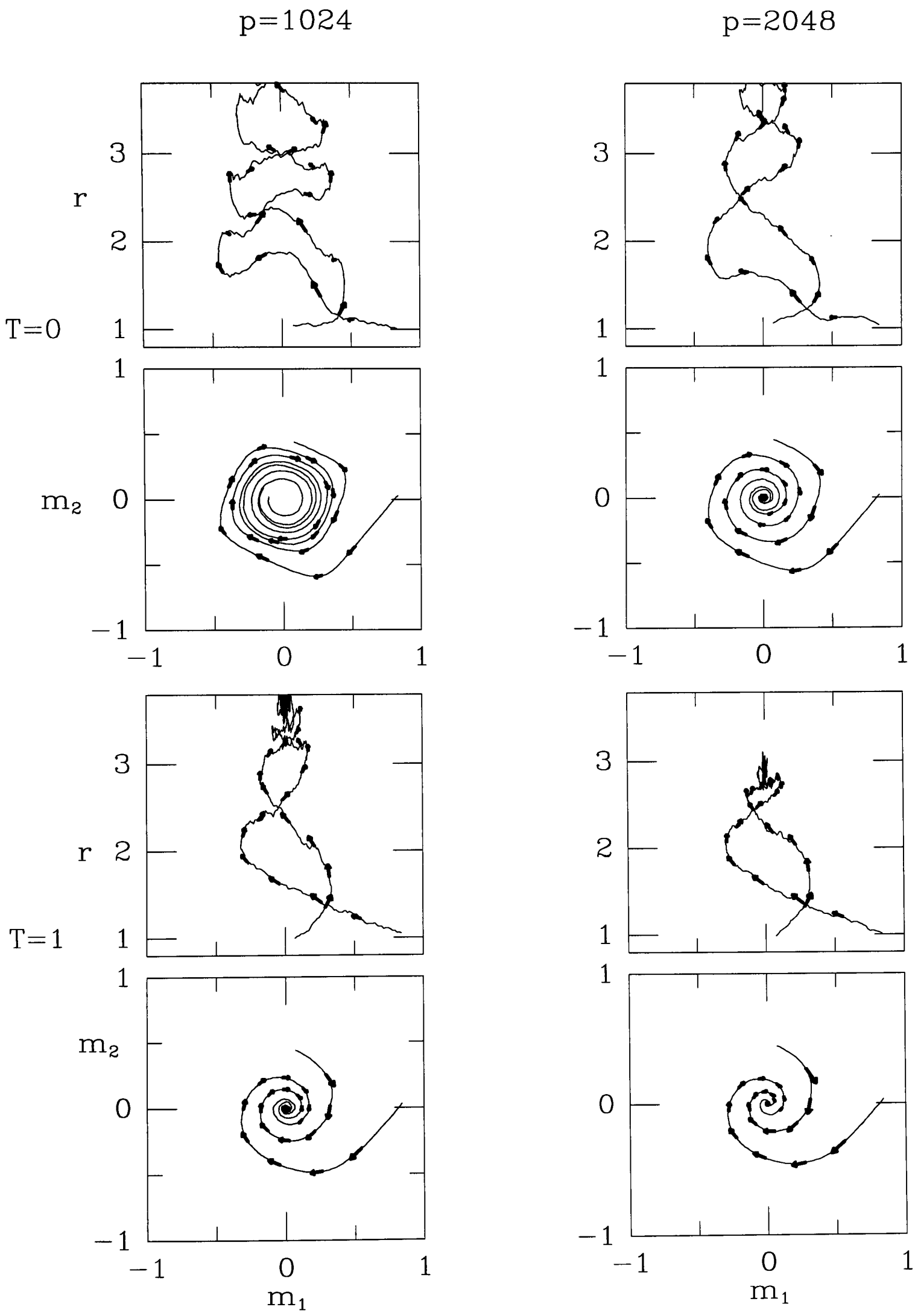


Figure 5.5: Plots showing the evolution of m^1, m^2 and r for a system of 40,000 spins and embedding matrix (5.29) with $a = 2$. The solid lines are the results of spin simulations, and the arrows show the instantaneous values of $\frac{dm^2}{dm^1}$ and $\frac{dr}{dm^1}$ calculated at intervals of 1/2 iteration per spin. The lengths of the arrows are proportional to the magnitude of the derivatives.

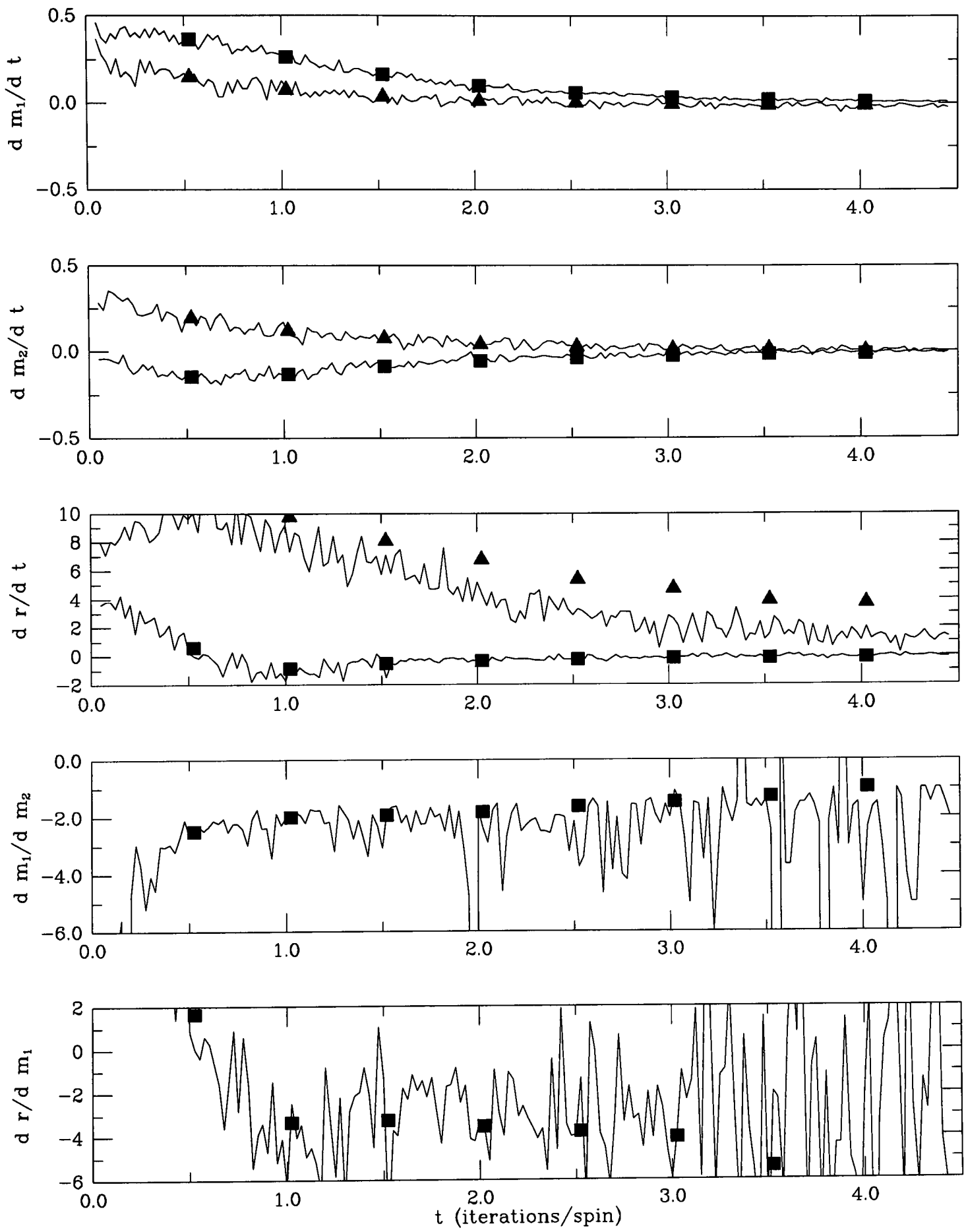


Figure 5.6: Plots showing the derivatives $\frac{dm^1}{dt}, \frac{dm^2}{dt}, \frac{dr}{dt}, \frac{dm^1}{dm^2}$ and $\frac{dr}{dm^1}$ as a function of time measured in iterations per spin, for initial conditions leading to retrieval, i.e. $\mathbf{m} \neq 0$ (squares), and non-retrieval, i.e. $\mathbf{m} = 0$ (triangles). The lines show data taken from numerical simulations whilst the points are theoretical data. The system has 40,000 spins and embedding matrix (5.29) with $a = 0$, $T = 0$, and $p = 1024$.

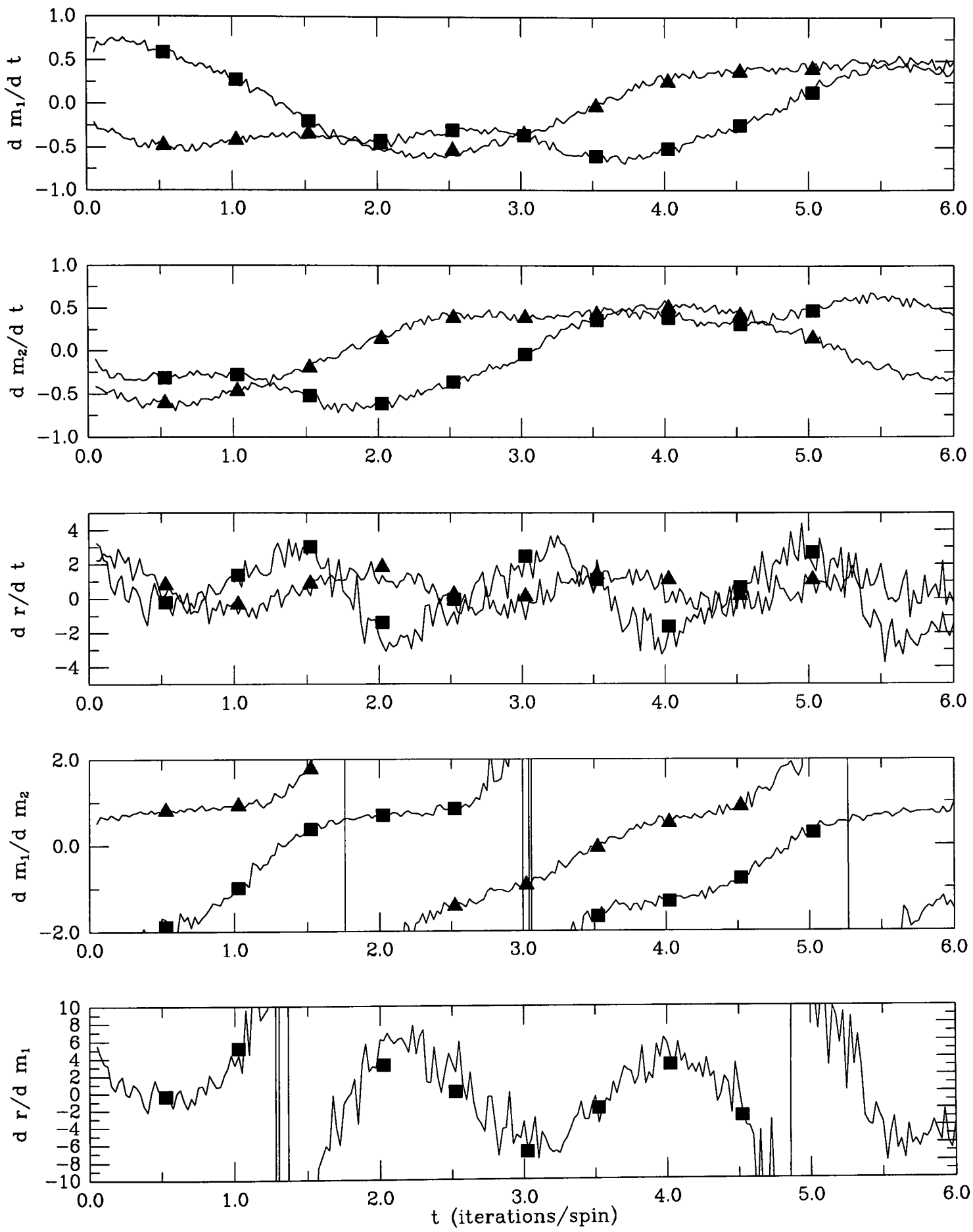


Figure 5.7: Plots showing the derivatives $\frac{dm^1}{dt}, \frac{dm^2}{dt}$, $\frac{dr}{dt}$, $\frac{dm^1}{dm^2}$ and $\frac{dr}{dm^1}$ as a function of time measured in iterations per spin. The lines show data taken from numerical simulations whilst the points are theoretical data. The system has 40,000 spins and embedding matrix (5.29) with $a = 1$, $T = 0$, and $p = 1024$ (squares) and $p = 2048$ (triangles).

5.3.2 Fixed-Points of the Flow

In the absence of asymmetric interactions we expect the fixed points of the flow (5.27) to correspond to the equilibrium saddle point equations derived in Chapter 4. In order to show that this is indeed the case we carry out the following co-ordinate changes on the flow equations

$$x = x'\sqrt{\frac{\Gamma^2}{2\mathcal{E}\rho^2}} + y'\sqrt{\frac{\Delta}{2\mathcal{E}\rho}} \qquad\qquad y = y'\sqrt{\frac{\Gamma^2}{2\mathcal{E}\rho^2}} - x'\sqrt{\frac{\Delta}{2\mathcal{E}\rho}},$$

the flow equations then take the form

$$\begin{aligned}
\frac{d\mathbf{m}}{dt} &= \left\langle \int Dx \int Dy\, \boldsymbol{\zeta}\, M_{\boldsymbol{\zeta}}(\mathbf{m}, r, x, y) \right\rangle_{\boldsymbol{\zeta}} - \mathbf{m} \\
\frac{1}{2}\frac{dr}{dt} &= \left\langle \int Dx \int Dy\, R_{\boldsymbol{\zeta}}(\mathbf{m}, r, x, y) \right\rangle_{\boldsymbol{\zeta}} - (r - \frac{1}{p}\operatorname{Tr}\mathbf{A})
\end{aligned} \tag{5.31}$$

where

$$\begin{aligned}
M_{\boldsymbol{\zeta}}(\mathbf{m}, r, x, y) &= \frac{1}{2}\left(1 - \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right)\tanh\left(\beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \beta U^-\right) \\
&\quad + \frac{1}{2}\left(1 + \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right)\tanh\left(\beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \beta U^+\right) \\
R_{\boldsymbol{\zeta}}(\mathbf{m}, r, x, y) &= \frac{1}{2}\left(1 - \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right)U^-\tanh\left(\beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \beta U^-\right) \\
&\quad + \frac{1}{2}\left(1 + \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right)U^+\tanh\left(\beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} + \beta U^+\right) \\
U^\pm &= \frac{\Gamma}{\rho}x - \sqrt{\frac{\Delta}{\rho}}y - \mathcal{B} \mp \Delta.
\end{aligned}$$

We now use the two identities:

$$\begin{aligned}
\tanh u &= \frac{1}{2}(1 - \tanh u)\int Dy\, \tanh\left(u + yz - z^2\right) \\
&\quad + \frac{1}{2}(1 + \tanh u)\int Dy\, \tanh\left(u + yz + z^2\right) \\
u\tanh u + z^2 &= \frac{1}{2}(1 - \tanh u)\int Dy\, \left(u + yz - z^2\right)\tanh\left(u + yz - z^2\right) \\
&\quad + \frac{1}{2}(1 + \tanh u)\int Dy\, \left(u + yz + z^2\right)\tanh\left(u + yz + z^2\right),
\end{aligned} \tag{5.32}$$

with $u = \boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x$ and $z^2 = \beta\Delta$, to show upon choosing $\rho = \beta$ and $\boldsymbol{\zeta}\cdot\boldsymbol{\mu} = \beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} - \beta\mathcal{B}$ the flow (5.31) is indeed at a fixed point

$$\begin{aligned}
\frac{d\mathbf{m}}{dt} &= \left\langle \int Dx\, \boldsymbol{\zeta}\, \tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right\rangle_{\boldsymbol{\zeta}} - \mathbf{m} = 0 \\
\frac{1}{2}\frac{dr}{dt} &= \left\langle \int Dx \left(\frac{\boldsymbol{\zeta}\cdot\boldsymbol{\mu}}{\beta} + \frac{\Gamma x}{\beta} - \boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m}\right)\tanh\left(\boldsymbol{\zeta}\cdot\boldsymbol{\mu} + \Gamma x\right)\right\rangle_{\boldsymbol{\zeta}} + \frac{\Delta}{\alpha} - (r - \frac{1}{p}\operatorname{Tr}\mathbf{A}) \\
&= \frac{1}{\alpha\beta}\mathbf{m}\cdot(\boldsymbol{\mu} - \beta\mathbf{A}\mathbf{m}) + \frac{\Gamma^2}{\alpha\beta}\left(1 - \left\langle \int Dx\, \tanh^2\left(\boldsymbol{\xi}\boldsymbol{\mu} + \Gamma x\right)\right\rangle_{\boldsymbol{\zeta}}\right) + \frac{\Delta}{\alpha} - (r - \frac{1}{p}\operatorname{Tr}\mathbf{A}) \\
&= \frac{1}{\alpha\beta}\mathbf{m}\cdot(\boldsymbol{\mu} - \beta\mathbf{A}\mathbf{m}) + \frac{\Gamma^2}{\alpha\beta}(1 - q) + \frac{\Delta}{\alpha} - (r - \frac{1}{p}\operatorname{Tr}\mathbf{A}) = 0,
\end{aligned} \tag{5.33}$$

where the equalities are consequences of the saddle point equations (5.13). By comparing the dynamic saddle point equations with the equilibrium saddle point equations we can see that the equilibrium equations derived in Chapter 4 correspond to:

$$\begin{aligned}
\boldsymbol{\mu} &= \beta\left(\mathbf{A}_{cc}^{-1} - \mathbf{A}_{cu}^{-1}\left(\mathbf{A}_{uu}^{-1} - \beta(1-q)\mathbb{1}\right)^{-1}\mathbf{A}_{uc}^{-1}\right)^{-1}\mathbf{m} \\
&= \beta\mathbf{A}_{uc}^{s}\mathbf{m} - \beta\rho(1-q)\mathbf{A}_{cu}^{s}\left(\mathbb{1} - \rho(1-q)\mathbf{A}_{uu}^{s}\right)^{-1}\mathbf{A}_{uc}^{s}\mathbf{m} \qquad (5.34) \\
\rho &= \beta \qquad (5.35)
\end{aligned}$$

Careful analysis of the equilibrium saddle point equations reveal that these indeed imply $\boldsymbol{\zeta}\cdot\boldsymbol{\mu} = \beta\boldsymbol{\zeta}\cdot\mathbf{A}\mathbf{m} - \beta\mathcal{B}$. Hence the present dynamic formalism correctly recovers the equilibrium phase diagram of Chapter 4 as fixed points of the flow.

5.4 Conclusions

In this chapter we have generalised a recent theory [44, 58] to describe the dynamics of the Hopfield model near saturation to systems with (i) arbitrary separable interactions, (ii) which need not be symmetric, and (iii) more than one condensed pattern. The theory, valid within the condensed ansatz, describes the evolution of macroscopic order parameters: the condensed overlaps m^{μ} ($\mu = 1,\ldots,c$, where c is the number of condensed overlaps), and the disordered contribution to the energy r. The theory is based on the systematic removal of microscopic memory effects, and requires the two assumptions of self averaging with respect to the microscopic realisations of the stored patterns, and equipartitioning of probability within the macroscopic subshells of the ensemble. Whilst these assumptions are correct in detailed balance equilibrium there is no reason *a priori* to believe they will carry over to non-equilibrium, non-symmetric cases. Indeed the results of previous studies [44, 59, 83] suggests that the assumption of equipartitioning within the $(\mathbf{m}, r)$ sub-shells is not always valid, leading to an overall slowing down effect in the flows that lead to non-retrieval states ($\mathbf{m} = 0$). The closure of the dynamic laws requires the calculation of an intrinsic noise distribution which can be calculated using the replica method. We have computed the region over which the replica symmetric solution is stable. We have shown that the theory captures the essential characteristics of the flow for both symmetric and asymmetric interactions, whilst close observation of the flows and the noise distribution itself [83] show that the agreement between experiment and theory is not perfect.

Furthermore, in those regions where detailed balance holds the correct equilibrium equations derived in Chapter 4 are recovered as stable fixed points of the dynamics. The theory has the additional advantage that the noise distribution can be calculated exactly within the limitations of the theory, without having to make ad-hoc assumptions such as the Gaussian approximation, which is clearly not true in the regions where $q \to 1$. Our results imply that the present theory provides the first systematic method for analysing non-symmetric attractor neural networks near saturation, which cannot be analysed within equilibrium statistical mechanics due to the absence of detailed balance.

Part III

The Generalised S-K Model

Chapter 6

Dynamic local field distribution: towards exact results

6.1 Introduction

In Chapter 5 we saw how to describe the Glauber dynamics of the generalised Hopfield [33] neural network model near saturation, in terms of the deterministic flow equations for a small number of macroscopic state variables: a finite number of condensed pattern overlaps m^μ ($\mu = 1, \ldots, c \ll \sqrt{N}$) and the disordered contribution to the 'energy', r, which accounts for the uncondensed patterns. The method for deriving the dynamic equations, upon calculating the 'noise' distribution to the local field using the replica identity (1.8), proposed by Coolen and Sherrington (CS) [44, 58], has subsequently been applied to a number of disordered spin models [46, 59] including the S-K model [45]. All these derivations were originally carried out within the replica symmetric ansatz, although generalisation to include replica symmetry breaking using Parisi's scheme [17, 19] is not difficult in principle [84].

The CS theory involves systematic removal of microscopic 'memories' in favour of a description in terms of the macroscopic quantities m and r only, and in order to derive a closed set of equations requires two key assumptions:

1. That the flow equations are self averaging.
2. That equipartitioning of the microscopic probability distribution $p_t(\boldsymbol{\sigma})$ within the macroscopic sub-shells occurs.

The two parameter CS theory is the simplest which is exact in the limits:

1. Near $t = 0$, for appropriate initial conditions, before microscopic correlations have had a chance to build up.
2. In detailed balance equilibrium, when the familiar replica results of [36] (Hopfield model) and [15] (S-K model) are recovered as fixed points of the flow equations.
3. In the limit of no disorder: $\alpha \to 0$ (Hopfield model) or $J \to 0$ (S-K model).

Although choosing m and r as our order parameters guarantees accuracy in each of these limits, this simple theory fails to account for a slowing down encountered on intermediate time scales.

The inaccuracy of the theory can be traced back to the assumption of equipartitioning within the macroscopic sub-shells. Equivalently, we are eliminating too much microscopic information. Clearly we need more parameters to define our macroscopic state. It is not immediately clear, however, which parameters should be included in order to achieve the required improvement in the theory. Both the magnetisation and the energy must be included, either implicitly or explicitly, in order to retain the accuracy of the theory in the limits mentioned above. Also, since we wish to retain the spirit of the original CS theory, where only (one time) state variables are used, we don't want to introduce objects such as correlation and response functions used in other formalisms [39, 40, 53, 55].

In this chapter we show how a continuous function, the joint spin-field distribution, can be used as a dynamic object. We show that the appropriate evolution equations are deterministic in the thermodynamic limit, and evaluate the coefficients for the generalised S-K model with interactions given by (1.18).

The choice of an order parameter function means that microscopic information discarded by the old theory is included in the new one. Although the same assumptions of (i) self averaging flow equations, and (ii) equipartitioning within the macroscopic sub-shells are again needed in order to close the equations derived within the replica formalism, the second assumption is now much less restricting, since much more information is specified by defining the macroscopic shell.

6.2 Definitions and macroscopic flow equations

In equilibrium only the energy is required to define the macroscopic sub-shell, hence a naive theory may choose only m and r (m is included since we need an observable) as our macro-variables as in Chapter 5. Closer analysis, however shows that too much microscopic information is being integrated out (i.e. equipartitioning does not hold). Since the master equation (1.5) is defined in terms of σ_i and $h_i(\boldsymbol{\sigma})$, a better choice might be the function

$$\rho^{\varsigma}(h;\boldsymbol{\sigma}) = \frac{1}{N}\sum_{j}\delta_{\varsigma,\sigma_j}\,\delta(h-h_j(\boldsymbol{\sigma})). \tag{6.1}$$

This choice is as least as good as the one in the old theory since the magnetisation and energy can be written

$$m(\boldsymbol{\sigma}) = \sum_{\varsigma}\varsigma\int dh\,\rho^{\varsigma}(h;\boldsymbol{\sigma}) \qquad\qquad E(\boldsymbol{\sigma}) = -\frac{1}{2}\sum_{\varsigma}\varsigma\int dh\,h\,\rho^{\varsigma}(h;\boldsymbol{\sigma}), \tag{6.2}$$

and since $\rho^{\varsigma}(h)$ is a continuous function it contains significantly more microscopic information.

Our macroscopic probability distribution is now given by

$$\mathcal{P}_t(\rho^{\varsigma}(h)) = \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\,\delta\left(\rho^{\varsigma}(h) - \frac{1}{N}\sum_{j}\delta_{\varsigma,\sigma_j}\,\delta(h-h_j(\boldsymbol{\sigma}))\right).$$

Substituting this into (1.13) we acquire the Liouville form, giving deterministic flow of the order parameter function, if

1. $c \ll \sqrt{N}$. Where c is the number of points at which we evaluate our function. Hence in order to derive equations for an order parameter function we need to take $c \to \infty$ *after* $N \to \infty$.

2. The distribution is sufficiently smooth so that there are no derivatives $\frac{\partial}{\partial h}\rho^{\varsigma}(h)$ that scale with N

In order to derive evolution equations for $\rho^{\varsigma}(h)$, from (1.14) we need first to calculate the discrete derivatives $\Delta_i^{\mu}(\boldsymbol{\sigma}) = F_i\rho^{\varsigma}(h^{\mu};\boldsymbol{\sigma}) - \rho^{\varsigma}(h^{\mu};\boldsymbol{\sigma})$, which we do using a Taylor expansion

$$\begin{aligned}\Delta_i^{\mu}(\boldsymbol{\sigma}) &= \frac{1}{N}\sum_j \left\{\delta_{\varsigma,F_i\sigma_j}\delta\left(h^{\mu}-h_j(\boldsymbol{\sigma})+2J_{ji}\sigma_i\right)-\delta_{\varsigma,\sigma_j}\delta\left(h^{\mu}-h_j(\boldsymbol{\sigma})\right)\right\}\\ &= \frac{1}{N}\sum_j \left(\delta_{\varsigma,F_i\sigma_j}-\delta_{\varsigma,\sigma_j}\right)\delta\left(h^{\mu}-h_j(\boldsymbol{\sigma})\right)\\ &\quad+\frac{2}{N}\sum_j \delta_{\varsigma,F_i\sigma_j}\left[J_{ji}\sigma_i\,\delta'\left(h^{\mu}-h_j(\boldsymbol{\sigma})\right)+J_{ji}^2\,\delta''\left(h^{\mu}-h_j(\boldsymbol{\sigma})\right)+\mathcal{O}(J_{ji}^3)\right] \qquad (6.3)\end{aligned}$$

where the primes indicate derivatives, and $\delta_{\varsigma,F_i\sigma_j} = \delta_{ij}\left(\delta_{\varsigma,-\sigma_j}-\delta_{\varsigma,\sigma_j}\right)+\delta_{\varsigma,\sigma_j}$. Here μ labels one of the c values of h at which we evaluate $\rho^{\varsigma}(h)$.

Using (6.3) in (1.14), and in the absence of self interactions, we have the following flow equation for $\rho_t^{\varsigma}(h)$:

$$\begin{aligned}\frac{d}{dt}\rho_t^{\varsigma}(h^{\mu}) &= \frac{1}{2}\left(1+\varsigma\tanh(\beta h^{\mu})\right)\rho_t^{-\varsigma}(h^{\mu})-\frac{1}{2}\left(1-\varsigma\tanh(\beta h^{\mu})\right)\rho^{\varsigma}(h^{\mu})\\ &\quad+\frac{\partial}{\partial h^{\mu}}\left(h^{\mu}\rho_t^{\varsigma}(h^{\mu})-\frac{1}{N}\sum_{i,j}J_{ji}\left\langle\tanh(\beta h_i(\boldsymbol{\sigma}))\delta_{\varsigma,\sigma_j}\delta(h^{\mu}-h_j(\boldsymbol{\sigma}))\right\rangle_{\rho^{\varsigma}(h);t}\right)\\ &\quad+\frac{\partial^2}{\partial h^{\mu 2}}\left(\frac{1}{N}\sum_{i,j}J_{ji}^2\left\langle\left[1-\sigma_i\tanh(\beta h_i(\boldsymbol{\sigma}))\right]\delta_{\varsigma,\sigma_j}\delta(h^{\mu}-h_j(\boldsymbol{\sigma}))\right\rangle_{\rho^{\varsigma}(h);t}\right)\\ &\quad+\mathcal{O}(N^{-1}\sum_{i,j}J_{ij}^3), \qquad (6.4)\end{aligned}$$

where the sub-shell average is defined as

$$\left\langle f(\boldsymbol{\sigma},\mathbf{h}(\boldsymbol{\sigma}))\right\rangle_{\rho^{\varsigma}(h);t} = \frac{\sum_{\boldsymbol{\sigma}}p_t(\boldsymbol{\sigma})\,f(\boldsymbol{\sigma},\mathbf{h}(\boldsymbol{\sigma}))\prod_{\nu}\prod_{\varsigma}\delta\left(\rho^{\varsigma}(h^{\nu})-\frac{1}{N}\sum_j\delta_{\varsigma,\sigma_j}\delta(h^{\nu}-h_j(\boldsymbol{\sigma})\right)}{\sum_{\boldsymbol{\sigma}}p_t(\boldsymbol{\sigma})\prod_{\nu}\prod_{\varsigma}\delta\left(\rho^{\varsigma}(h^{\nu})-\frac{1}{N}\sum_j\delta_{\varsigma,\sigma_j}\delta(h^{\nu}-h_j(\boldsymbol{\sigma})\right)}. \qquad (6.5)$$

So far the equation (6.4) is exact in the thermodynamic limit, however, the sub-shell average (6.5) needed to evaluate the coefficients still contains an explicit dependence on time, and on the disorder. As in Chapter 5, and following the prescription of the CS theory, we make the two vital assumptions:

1. Self-averaging of the flow equations.

2. Equipartitioning of probability within the macroscopic sub-shells.

Hence we need to evaluate terms like

$$\left\langle f(\boldsymbol{\sigma},\mathbf{h}(\boldsymbol{\sigma}))\right\rangle_{\rho^{\varsigma}(h)} = \left\langle\frac{\sum_{\boldsymbol{\sigma}}f(\boldsymbol{\sigma},\mathbf{h}(\boldsymbol{\sigma}))\prod_{\nu}\prod_{\varsigma}\delta\left(\rho^{\varsigma}(h^{\nu})-\frac{1}{N}\sum_j\delta_{\varsigma,\sigma_j}\delta(h^{\nu}-h_j(\boldsymbol{\sigma})\right)}{\sum_{\boldsymbol{\sigma}}\prod_{\nu}\prod_{\varsigma}\delta\left(\rho^{\varsigma}(h^{\nu})-\frac{1}{N}\sum_j\delta_{\varsigma,\sigma_j}\delta(h^{\nu}-h_j(\boldsymbol{\sigma})\right)}\right\rangle_{J_{ij}}$$

which we do using the familiar replica identity (1.8) to give

$$\lim_{n\to 0}\left\langle\sum_{\boldsymbol{\sigma}^1}\cdots\sum_{\boldsymbol{\sigma}^n}f(\boldsymbol{\sigma}^1,\mathbf{h}(\boldsymbol{\sigma}^1))\prod_a\prod_{\nu}\prod_{\varsigma}\delta\left(\rho^{\varsigma}(h^{\nu})-\frac{1}{N}\sum_j\delta_{\varsigma,\sigma_j^a}\delta(h^{\nu}-h_j(\boldsymbol{\sigma}^a)\right)\right\rangle_{J_{ij}}. \qquad (6.6)$$

NB This approach to closing the equations is quite different from that of Horner [85] who also considered the evolution of the joint spin-field distribution. Upon arriving at (6.4) Horner then set to derive equations for double distribution quantities such as

$$\sum_{i,j\neq i} J_{ij} \left\langle \delta_{\varsigma',\sigma_i}\, \delta(h^\nu - h_i(\boldsymbol{\sigma}))\; \delta_{\varsigma,\sigma_j} \delta(h^\mu - h_j(\boldsymbol{\sigma})) \right\rangle_{\rho^\varsigma(h);t}$$

which are needed to evaluate the coefficients for the single distribution evolution equation. The evolution equations for these objects, however, require triple distribution quantities, and so on, leading to an infinite hierarchy of equations. The hierarchy has to be closed at some stage. Upon introduction of the function $f(\varsigma,\varsigma',h^\mu,h^\nu)$ defined by

$$\sum_{i,j\neq i} J_{ij} \left\langle \delta_{\varsigma',\sigma_i}\, \delta(h^\nu - h_i(\boldsymbol{\sigma}))\; \delta_{\varsigma,\sigma_j} \delta(h^\mu - h_j(\boldsymbol{\sigma})) \right\rangle_{\rho^\varsigma(h);t} = f(\varsigma,\varsigma',h^\mu,h^\nu)\,\rho^\varsigma(h^\mu)\rho^{\varsigma'}(h^\nu),$$

Horner approximates the triple distributions by an appropriately symmetrised combinations of f's and ρ's. Whilst the advantage of this approach is its simplicitly - the evolution equations obtained can be iterated directly without solving a set of saddle point equations at each time step, the disadvantage is that it is not clear what approximations are being made by closing at a particular level. The choice is always somewhat arbitrary.

The major advantage of the approach presented in this chapter is the transparency of the assumptions above.

6.3 The replica calculation

In order to calculate sub-shell averages of the kind (6.6) we proceed in the now familiar fashion for replica calculations by reversing the order of averaging, carrying out the disorder average before the spin average. First, in order to remove the disorder dependence (through $h_i(\boldsymbol{\sigma})$) from within the constraining delta functions, we use the following representation of unity

$$1 = \prod_a \prod_k \int dH_k^a\; \delta(H_k^a - h_k(\boldsymbol{\sigma}^a)) = \prod_a \prod_k \int \frac{d\hat{h}_k^a\, dH_k^a}{2\pi}\; e^{i\hat{h}_k^a (H_k^a - h_k(\boldsymbol{\sigma}^a))}.$$

We are considering the generalised S-K model with interactions given by (1.18), hence we have both symmetric contributions to the disorder $z_{ij}^s = z_{ji}^s$ and antisymmetric contributions $z_{ij}^a = -z_{ji}^a$, where in each case the z's are Gaussianly distributed with zero mean and unit width.

After properly symmetrising the exponent, and using permutation symmetry of the spin labels, we see that the disorder average involves terms like

$$N \prod_{i,j<i} \int \left(\frac{dz_{ij}^s\, dz_{ij}^a}{2\pi}\, e^{-\frac{1}{2}\left(z_{ij}^{s\,2} + z_{ij}^{a\,2}\right)} \right) \left\{ \begin{array}{c} \left(\frac{\tilde{J}_0}{N} + \frac{\tilde{J}\left(z_{21}^s + k z_{21}^a\right)}{\sqrt{N(1+k^2)}} \right) \\ \\ \left(\frac{\tilde{J}_0}{N} + \frac{\tilde{J}\left(z_{21}^s + k z_{21}^a\right)}{\sqrt{N(1+k^2)}} \right)^2 \end{array} \right\}$$
$$\prod_a e^{-i\frac{\tilde{J}_0}{N}\hat{h}_i^a\sigma_j^a - i\frac{J}{\sqrt{N(1+k^2)}}\left(z_{ij}^s\left(\hat{h}_i^a\sigma_j^a + \hat{h}_j^a\sigma_i^a\right) + k z_{ij}^a\left(\hat{h}_i^a\sigma_j^a - \hat{h}_j^a\sigma_i^a\right)\right)},$$

where the upper term within the bracket corresponds to the coefficient of the first derivative in (6.4), whilst the lower term belongs to the diffusion (second derivative) term. The only $\mathcal{O}(1)$

terms in this average are

$$\left\{ \begin{array}{c} \tilde{J}_0 - i\frac{\tilde{J}^2}{(1+k^2)}\sum_a \left[(1+k^2)\hat{h}_2^a\sigma_1^a + (1-k^2)\hat{h}_1^a\sigma_2^a\right] \\ \\ \tilde{J}^2 \end{array} \right\}$$

$$\prod_{i,j<i} e^{-i\frac{\tilde{J}_0}{N}\sum_a \hat{h}_i^a\sigma_j^a - \frac{\tilde{J}^2}{2N(1+k^2)}\sum_{a,b}\left[(\hat{h}_i^a\sigma_j^a+\hat{h}_j^a\sigma_i^a)(\hat{h}_i^b\sigma_j^b+\hat{h}_j^b\sigma_i^b)+k^2(\hat{h}_i^a\sigma_j^a-\hat{h}_j^a\sigma_i^a)(\hat{h}_i^b\sigma_j^b-\hat{h}_j^b\sigma_i^b)\right]}.$$

Hence the diffusion term is trivially

$$\tilde{J}^2\sum_\sigma \langle 1-\sigma\tanh(\beta H)\rangle_{\rho_t^\sigma(H)}\,\rho_t^\varsigma(h),$$

where $\langle f(\sigma,H)\rangle_{\rho_t^\sigma(H)} = \int dH\,\rho_t^\sigma(H)\,f(\sigma,H)$. The term arising from the $\tilde{J}_0$ in the first derivative term is also readily evaluated:

$$\tilde{J}_0\sum_\sigma \langle\tanh(\beta H)\rangle_{\rho_t^\sigma(H)}\,\rho_t^\varsigma(h);$$

whilst the remainder of this coefficient is non-trivial.

Upon introduction of the following order parameters

$$m^a(\boldsymbol{\sigma}) = \frac{1}{N}\sum_i \sigma_i^a \qquad q_{ab}(\boldsymbol{\sigma}) = \frac{1}{N}\sum_i \sigma_i^a\sigma_i^b$$

$$Q_{ab}(\hat{\mathbf{h}}) = \frac{1}{N}\sum_i \hat{h}_i^a\hat{h}_i^b \qquad \mathcal{R}_{ab}(\hat{\mathbf{h}},\boldsymbol{\sigma}) = \frac{1}{N}\sum_i \hat{h}_i^a\sigma_i^b \qquad \mathcal{W}_{ab}(\hat{\mathbf{h}},\boldsymbol{\sigma}) = \frac{1}{N}\sum_i \hat{h}_i^a\hat{h}_i^b\sigma_i^a\sigma_i^b,$$

and their conjugates, via integral representations of unity, the full replica expression for (6.6) becomes

$$\lim_{n\to 0}\prod_{a=1}^n \int\left(\prod_{k=1}^N dH_k^a\right)\delta(h^\mu - H_2^1)\tanh(\beta H_1^1)\,e^{-N\frac{\tilde{J}^2}{2(1+k^2)}\sum_{a,b}\left((1+k^2)Q_{ab}q_{ab}+(1-k^2)\mathcal{R}_{aa}\mathcal{R}_{ab}\right)}$$

$$\int d\mathbf{m}\,d\hat{\mathbf{m}}\,d\mathbf{q}\,d\hat{\mathbf{q}}\,d\mathbf{Q}\,d\hat{\mathbf{Q}}\,d\mathcal{R}\,d\hat{\mathcal{R}}\,d\mathcal{W}\,d\hat{\mathcal{W}}\,e^{iN\left(\sum_a \hat{m}^a m^a + \sum_{a,b}\left(q_{ab}\hat{q}_{ab}+Q_{ab}\hat{Q}_{ab}+\mathcal{R}_{ab}\hat{\mathcal{R}}_{ab}+\mathcal{W}_{ab}\hat{\mathcal{W}}_{ab}\right)\right)}$$

$$\int\left(\prod_i\prod_a \frac{d\hat{h}_i^a}{2\pi}\right)e^{i\sum_a \hat{h}_i^a(H_i^a - \tilde{J}_0 m^a)}\left\langle \delta_{\varsigma,\sigma_2^1}\prod_\nu\prod_{\varsigma'}\delta\left(\rho^{\varsigma'}(h^\nu) - \frac{1}{N}\sum_j \delta_{\varsigma',\sigma_j^a}\,\delta(h^\nu - H_j^a)\right)\right.$$

$$e^{i\tilde{J}_0\sum_a \mathcal{R}_{aa}+\frac{\tilde{J}^2}{(1+k^2)}\sum_{a,b}\mathcal{W}_{ab}}\left(-i\frac{\tilde{J}^2}{(1+k^2)}\sum_a\left[(1+k^2)\hat{h}_2^a\sigma_1^a + (1-k^2)\hat{h}_1^a\sigma_2^a\right]\right)$$

$$\left.\prod_i e^{-i\left(\sum_a \hat{m}^a\sigma_i^a + \sum_{a,b}\left(\sigma_i^a\hat{q}_{ab}\sigma_i^b + \hat{h}_i^a\hat{Q}_{ab}\hat{h}_i^b + \hat{h}_i^a\hat{\mathcal{R}}_{ab}\sigma_i^b + \hat{h}_i^a\sigma_i^a\hat{\mathcal{W}}_{ab}\hat{h}_i^b\sigma_i^b\right)\right)}\right\rangle_{\boldsymbol{\sigma}}.$$

NB Here for ease of notation we have assumed that the replica matrices are symmetric under interchange of labels, i.e. $\mathcal{R}_{ab} = \mathcal{R}_{ba}$ etc.

In order to carry out the integral over H_k^a we need to apply the constraint $\delta(\rho^\varsigma(h^\nu) - \frac{1}{N}\sum_j \delta_{\varsigma,\sigma_j^a}\,\delta(h^\nu - H_j^a))$, which we do by introducing the integral representation of the delta function:

$$\prod_\nu\prod_\varsigma \delta(\rho^\varsigma(h^\nu) - \frac{1}{N}\sum_j \delta_{\varsigma,\sigma_j^a}\delta(h^\nu - H_j^a))$$

$$= \left(\frac{N}{2\pi}\right)^{2nc}\int\left(\prod_a\prod_\nu\prod_\varsigma dx_a^{\nu,\varsigma}\right)e^{iN\sum_a\sum_\nu\sum_\varsigma x_a^{\nu,\varsigma}\left(\rho^\varsigma(h^\nu) - \frac{1}{N}\sum_j \delta_{\varsigma,\sigma_j^a}\,\delta(h^\nu - H_j^a)\right)}.$$

This expression still contains a delta function in the exponent, which we want to avoid. In order to deal with this, we restrict the region over which $\rho^{\varsigma}(h)$ is to be known to $[-L, L]$. We will eventually let $L \to \infty$, such that $\frac{L}{c} \to 0$. Recalling that $\lim_{N\to\infty} \frac{c^2}{N} \to 0$, we see that this implies that $\frac{L}{\sqrt{N}} \to 0$, i.e. we are assuming that the width of the distribution scales as something less than $\sqrt{N}$. Dividing the region into c intervals of width Δ, so that $c\Delta = 2L$, changing variables so that $x_a^{\nu,\varsigma} = i\Delta\hat{\rho}_a^{\varsigma}(h^{\nu})$, and assuming $\rho^{\varsigma}(h)$ is smooth within an interval Δ so that we can convert a discrete sum to an integral, we have

$$\begin{aligned} & iN\sum_{\nu} x_a^{\nu,\varsigma}\rho^{\varsigma}(h^{\nu}) - i\sum_{\nu}\sum_{j} x_a^{\nu,\varsigma}\delta_{\varsigma,\sigma_j^a}\,\delta(h^{\nu} - H_j^a) \\ &= -\Delta N\sum_{\nu}\hat{\rho}_a^{\varsigma}(h^{\nu})\rho^{\varsigma}(h^{\nu}) + \Delta\sum_{\nu}\sum_{j}\hat{\rho}_a^{\varsigma}(h^{\nu})\delta_{\varsigma,\sigma_j^a}\,\delta(h^{\nu} - H_j^a) \\ &= -N\int dh\,\hat{\rho}_a^{\varsigma}(h)\,\rho^{\varsigma}(h) + \sum_{j}\delta_{\varsigma,\sigma_j^a}\hat{\rho}_a^{\varsigma}(H_j^a). \end{aligned}$$

Also $dx_a^{\nu,\varsigma} = i\delta\hat{\rho}_a^{\varsigma}(h)$.

We then have an integral with a saddle point form and an intensive remainder, which, evaluated in the saddle point determines the coefficient:

$$\begin{aligned} &\lim_{N\to\infty}\lim_{n\to 0}\int d\mathbf{m}\,d\hat{\mathbf{m}}\,d\mathbf{q}\,d\hat{\mathbf{q}}\,d\mathbf{Q}\,d\hat{\mathbf{Q}}\,d\mathcal{R}\,d\hat{\mathcal{R}}\,d\mathcal{W}\,d\hat{\mathcal{W}}\,d\mathcal{E}\,d\hat{\mathcal{E}}\,\delta\hat{\rho}^{\varsigma'}(h)\,e^{N\Omega+\sum_i \ln\left[\int d\hat{\mathbf{h}}_i\,d\mathbf{H}_i\,e^{\Phi(\hat{\mathbf{h}}_i,\mathbf{H}_i,\boldsymbol{\sigma}_i)}\right]+\mathcal{O}(1)} \\ &-i\tilde{J}^2\sum_a\left[\frac{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,\sigma^a\tanh(\beta H^1)\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}\,\frac{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,\hat{h}^a\,\delta_{\varsigma,\sigma^1}\,\delta(h-H^1)\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}\right. \\ &\left.+\frac{(1-k^2)}{(1+k^2)}\,\frac{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,\hat{h}^a\tanh(\beta H^1)\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}\,\frac{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,\sigma^a\,\delta_{\varsigma,\sigma^1}\,\delta(h-H^1)\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}}\int d\hat{\mathbf{h}}\,d\mathbf{H}\,e^{\Phi(\hat{\mathbf{h}},\mathbf{H},\boldsymbol{\sigma})}}\right] \end{aligned} \tag{6.7}$$

where

$$\begin{aligned} \Omega &= i\sum_a \hat{m}^a m^a + i\sum_{a,b}\left(q_{ab}\hat{q}_{ab} + Q_{ab}\hat{Q}_{ab} + \mathcal{R}_{ab}\hat{\mathcal{R}}_{ab} + \mathcal{W}_{ab}\hat{\mathcal{W}}_{ab}\right) \\ &\quad -\frac{\tilde{J}^2}{2}\sum_{a,b}\left(Q_{ab}q_{ab} + \frac{(1-k^2)}{(1+k^2)}\mathcal{R}_{ab}\mathcal{R}_{ab}\right) - \sum_a\sum_{\sigma}\int dh\,\hat{\rho}_a^{\sigma}(h)\rho^{\sigma}(h) \\ \Phi(\hat{\mathbf{h}}_i,\mathbf{H}_i,\boldsymbol{\sigma}_i) &= -\frac{1}{2}\sum_{a,b}\hat{h}_i^a\left(2i\hat{Q}_{ab} + 2i\sigma_i^a\hat{\mathcal{W}}_{ab}\sigma_i^b\right)\hat{h}_i^b + i\sum_a\hat{h}_i^a\left(H_i^a - \tilde{J}_0 m^a - \sum_{\beta}\hat{\mathcal{R}}_{ab}\sigma_i^b\right) \\ &\quad -i\sum_a\hat{m}^a\sigma_i^a - i\sum_{a,b}\sigma_i^a\hat{q}_{ab}\sigma_i^b + \sum_a\hat{\rho}^{\sigma_i^a}(H_i^a). \end{aligned} \tag{6.8}$$

The saddle point exponent is $N\Psi = N\Omega + \sum_i \ln\left[\int d\hat{\mathbf{h}}_i\,d\mathbf{H}_i\,e^{\Phi(\hat{\mathbf{h}}_i,\mathbf{H}_i,\boldsymbol{\sigma}_i}\right]$. By interchanging the order of limits, then, in the thermodynamic limit ($N \to \infty$) the integral will be dominated by the maximal (for $n > 1$) value of Ψ, determined by the extremum condition $\frac{\partial\Psi}{\partial m^a} = 0$ etc. The procedure is less clear if we adhere strictly to the order of limits [17]. Here problems occur for $n < 1$ when some maxima become minima (in the RS solution).

Four of the saddle points can be readily evaluated: since $\mathcal{W}_{ab}$ only appears in the combination $\mathcal{W}_{ab}\hat{\mathcal{W}}_{ab}$, $\frac{\partial\Psi}{\partial\mathcal{W}_{ab}} = 0$ gives $\hat{\mathcal{W}}_{ab} = 0$ which can be implemented immediately since $\mathcal{W}_{ab}$ does not

appear in the intensive terms, except in combinations that will vanish when we take $n \to 0$; secondly $\frac{\partial \Psi}{\partial Q_{ab}} = 0$ gives $i\hat{Q}_{ab} = \frac{\tilde{J}^2}{2} q_{ab}$; thirdly $\frac{\partial \Psi}{\partial q_{ab}} = 0$ gives $i\hat{q}_{ab} = \frac{\tilde{J}^2}{2} Q_{ab}$; and fourthly $\frac{\partial \Psi}{\partial \mathcal{R}_{ab}} = 0$ gives $i\hat{\mathcal{R}}_{ab} = \frac{\tilde{J}^2(1-k^2)}{(1+k^2)} \mathcal{R}_{ab}$. Applying these allows us to perform the integral over $\hat{h}^a_k$.

Since the S-K model has infinite range interactions, we can use mean-field theory (all spins are equivalent) allowing us to remove the site label i, therefore the saddle point exponent is

$$\begin{aligned} \Psi &= i \sum_a \hat{m}^a m^a + \frac{\tilde{J}^2}{2} \sum_{a,b} q_{ab} Q_{ab} - \frac{(1+k^2)}{2\tilde{J}^2(1-k^2)} \sum_{a,b} \hat{\mathcal{R}}_{ab} \hat{\mathcal{R}}_{ab} - \frac{1}{2} \ln \left[\det \left(\tilde{J}^2 \mathbf{q} \right) \right] \\ &\quad - \sum_a \sum_\sigma \int dh \, \hat{\rho}^\sigma_a(h) \rho^\sigma_t(h) + \ln \left\langle \int d\mathbf{H} \, \mathcal{M}(\mathbf{H}, \boldsymbol{\sigma}) \right\rangle_{\boldsymbol{\sigma}} \end{aligned} \tag{6.9}$$

where

$$\mathcal{M}(\mathbf{H}, \boldsymbol{\sigma}) = e^{-\frac{1}{2}(\mathbf{H} - \tilde{J}_0 \mathbf{m} - \hat{\mathcal{R}}\boldsymbol{\sigma})(\tilde{J}^2 \mathbf{q})^{-1}(\mathbf{H} - \tilde{J}_0 \mathbf{m} - \hat{\mathcal{R}}\boldsymbol{\sigma}) - i\hat{\mathbf{m}}\boldsymbol{\sigma} - \frac{\tilde{J}^2}{2}\boldsymbol{\sigma}\mathbf{Q}\boldsymbol{\sigma} + \sum_a \hat{\rho}^{\sigma^a}_a(H^a)}.$$

The intensive term is

$$\begin{aligned} &\tilde{J}^2 \sum_{a,b} \left[\left\langle \sigma^a \tanh(\beta H^1) \right\rangle_{\mathcal{M}} \left(\tilde{J}^2 \mathbf{q} \right)^{-1}_{ab} \left\langle \left(H^b - \tilde{J}_0 m^b - \sum_c \hat{\mathcal{R}}_{bc} \sigma^c \right) \delta_{\varsigma,\sigma^1} \, \delta(h - H^1) \right\rangle_{\mathcal{M}} \right. \\ &\left. + \frac{(1-k^2)}{(1+k^2)} \left\langle \left(H^a - \tilde{J}_0 m^a - \sum_c \hat{\mathcal{R}}_{ac} \sigma^c \right) \tanh(\beta H^1) \right\rangle_{\mathcal{M}} \left(\tilde{J}^2 \mathbf{q} \right)^{-1}_{ab} \left\langle \sigma^b \delta_{\varsigma,\sigma^1} \, \delta(h - H^1) \right\rangle_{\mathcal{M}} \right], \end{aligned} \tag{6.10}$$

where

$$\langle f(\mathbf{H}, \boldsymbol{\sigma}) \rangle_{\mathcal{M}} = \frac{\sum_{\boldsymbol{\sigma}} \int d\mathbf{H} \, \mathcal{M}(\mathbf{H}, \boldsymbol{\sigma}) \, f(\mathbf{H}, \boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}} \int d\mathbf{H} \, \mathcal{M}(\mathbf{H}, \boldsymbol{\sigma})}.$$

The saddle point equations, needed to evaluate the intensive terms, are in general complicated. The familiar magnetisation, and spin glass order parameter however take their familiar forms as averages over the measure $\mathcal{M}$

$$m^a = \langle \sigma^a \rangle_{\mathcal{M}} \qquad\qquad q_{ab} = \left\langle \sigma^a \sigma^b \right\rangle_{\mathcal{M}},$$

whilst

$$\rho^\varsigma_t(h) = \langle \delta_{\varsigma,\sigma^c} \, \delta(h - H^c) \rangle_{\mathcal{M}} \;\; \forall \; c. \tag{6.11}$$

6.3.1 Detailed balance equilibrium

In detailed balance (i.e. $k = 0$ corresponding to the original S-K model) equilibrium, we know that the probability distribution is given by the Gibbs form $p_e(\boldsymbol{\sigma}) \propto e^{-\beta \mathcal{H}(\boldsymbol{\sigma})}$. Therefore, in equilibrium, the constraint restricting states under consideration to those with the same joint distribution, must reduce to restricting consideration to those states with the same energy. For this reason we postulate that in equilibrium $\hat{\rho}^{\sigma^a}(H^a) = \frac{\beta}{2} \sigma^a H^a$, since in this case

$$-\sum_a \sum_\sigma \int dh \, \hat{\rho}^\sigma_a(h) \rho^\sigma_t(h) = -\frac{\beta}{2} \sum_a \sum_\sigma \sigma \int dh \, h \, \rho^\sigma_t(h) = \beta \sum_a E(\boldsymbol{\sigma}^a).$$

The conjugate term appearing in the measure $\mathcal{M}(\mathbf{H}, \boldsymbol{\sigma})$ gives $-\beta E(\boldsymbol{\sigma})$, which is exactly what we would expect from doing the equilibrium calculation, in which using $E = -\frac{1}{2} \sum_i \sigma_i h_i$ we calculate $\rho^\varsigma_{eq}(h)$ from

$$\rho^\varsigma_{eq}(h) = \frac{\frac{1}{N} \sum_k \sum_{\boldsymbol{\sigma}} e^{\frac{\beta}{2} \sum_i \sigma_i h_i(\boldsymbol{\sigma})}, \delta_{\varsigma,\sigma_k} \, \delta(h - h_k(\boldsymbol{\sigma}))}{\sum_{\boldsymbol{\sigma}} e^{\frac{\beta}{2} \sum_i \sigma_i h_i(\boldsymbol{\sigma})}}.$$

In order to verify this choice, we need to check that it corresponds to a fixed point of the flow equations (6.4). First, however, consider its effect on the saddle point exponent (6.9). Given this specific choice for $\hat{\rho}^{\sigma^a}(H^a)$, the integral over H^a in the logarithm term can be carried out, leaving

$$\ln \left\langle e^{\frac{\beta}{2}\boldsymbol{\sigma}(\tilde{J}_0\mathbf{m}+\hat{\mathcal{R}}\boldsymbol{\sigma})+\frac{\beta^2}{8}\tilde{J}^2\boldsymbol{\sigma}\mathbf{q}\boldsymbol{\sigma}-i\hat{\mathbf{m}}\boldsymbol{\sigma}-\frac{\tilde{J}^2}{2}\boldsymbol{\sigma}\mathbf{Q}\boldsymbol{\sigma}} \right\rangle_{\boldsymbol{\sigma}},$$

giving the saddle point equations

$$\begin{aligned}
m^a &= \frac{-2i}{\beta\tilde{J}_0}\hat{m}^a = \frac{\left\langle \sigma^a e^{\beta\tilde{J}_0\sum_a m^a\sigma^a+\frac{\beta^2\tilde{J}^2}{2}\sum_{a,b}\sigma^a q_{ab}\sigma^b}\right\rangle_{\boldsymbol{\sigma}}}{\left\langle e^{\beta\tilde{J}_0\sum_a m^a\sigma^a+\frac{\beta^2\tilde{J}^2}{2}\sum_{a,b}\sigma^a q_{ab}\sigma^b}\right\rangle_{\boldsymbol{\sigma}}} \\
q_{ab} &= \frac{-4}{\beta^2}Q_{ab} = \frac{2}{\tilde{J}^2\beta}\hat{\mathcal{R}}_{ab} = \frac{\left\langle \sigma^a\sigma^b e^{\beta\tilde{J}_0\sum_a m^a\sigma^a+\frac{\beta^2\tilde{J}^2}{2}\sum_{a,b}\sigma^a q_{ab}\sigma^b}\right\rangle_{\boldsymbol{\sigma}}}{\left\langle e^{\beta\tilde{J}_0\sum_a m^a\sigma^a+\frac{\beta^2\tilde{J}^2}{2}\sum_{a,b}\sigma^a q_{ab}\sigma^b}\right\rangle_{\boldsymbol{\sigma}}} \\
\rho_t^{\varsigma}(h) &= \frac{\left\langle \int d\mathbf{H}\,\delta_{\varsigma,\sigma^c}\,\delta(h-H^c)\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})+\beta\boldsymbol{\sigma}\mathbf{H}}\right\rangle_{\boldsymbol{\sigma}}}{\left\langle \int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})+\beta\boldsymbol{\sigma}\mathbf{H}}\right\rangle_{\boldsymbol{\sigma}}} \;\forall\; c.
\end{aligned} \tag{6.12}$$

i.e. we have the familiar equilibrium S-K saddle point equations of [15], with the distribution of [86].

We now have to show that the choice $\hat{\rho}^{\sigma^a}(H^a)$ corresponds to fixed points of the flow (6.4) $\frac{d}{dt}\rho_t^{\varsigma}(h) = 0\;\forall\; h,\; \varsigma$.

Let us first evaluate the first derivative with respect to h in the diffusive term. This gives

$$-\tilde{J}^2\sum_{\sigma}\langle 1-\sigma\tanh(\beta H)\rangle_{\rho_t^{\sigma}(H)}\times\left\langle \delta_{\varsigma,\sigma^c}\,\delta(h-H^c)\left[\left(\tilde{J}^2\mathbf{q}\right)^{-1}_{cb}\left(H^b-\tilde{J}_0m^b\right)-\beta\sigma^c\right]\right\rangle_{\mathcal{M}_{eq}} \tag{6.13}$$

where

$$\langle f(\mathbf{H},\boldsymbol{\sigma})\rangle_{\mathcal{M}_{eq}} = \frac{\left\langle \int d\mathbf{H} f(\mathbf{H},\boldsymbol{\sigma})\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})+\beta\boldsymbol{\sigma}\mathbf{H}}\right\rangle_{\boldsymbol{\sigma}}}{\left\langle \int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})+\beta\boldsymbol{\sigma}\mathbf{H}}\right\rangle_{\boldsymbol{\sigma}}}.$$

The terms which together form the first derivative are

$$\begin{aligned}
&\left[-\tilde{J}_0\left(\langle\tanh(\beta H)\rangle_{\mathcal{M}_{eq}}-m\right)+\left(h-\tilde{J}_0m\right)\right]\rho_t^{\varsigma}(h) \\
&\quad-\tilde{J}^2\left\langle\sigma^a\tanh(\beta H^1)\right\rangle_{\mathcal{M}_{eq}}\left(\tilde{J}^2\mathbf{q}\right)^{-1}_{ab}\left\langle\delta_{\varsigma,\sigma^1}\,\delta(h-H^1)\left(H^b-\tilde{J}_0m^b\right)\right\rangle_{\mathcal{M}_{eq}}
\end{aligned}$$

equations (6.13) and (6.14) sum to zero. The terms

$$\frac{1}{2}\left(1+\varsigma\tanh(\beta h^\mu)\right)\rho_t^{-\varsigma}(h^\mu)-\frac{1}{2}\left(1-\varsigma\tanh(\beta h^\mu)\right)\rho^{\varsigma}(h^\mu)=-\frac{1}{2}\sum_{\varsigma}\varsigma\frac{(1-\varsigma)e^{\beta h}+(1+\varsigma)e^{-\beta h}}{e^{\beta h}+e^{-\beta h}}\rho_t^{\varsigma}(h)$$

also sum to zero since $\rho_t^{\varsigma}(h)\sim e^{\varsigma\beta h}$.

Hence for the symmetric (S-K) model (i.e. with $k=0$), and for $\hat{\rho}^{\sigma^a}(H^a)=\frac{\beta}{2}\sigma^a H^a$, $\frac{d}{dt}\rho_t^{\varsigma}(h)=0\ \forall\ h,\ \varsigma$, with the familiar detailed balance equilibrium saddle point equations (6.12).

6.4 Replica symmetry

In order to proceed further in evaluating the saddle points we need to make some ansatz about the form of the replica matrices. As a first step we assume replica symmetry, and simultaneously make a change of variables so that

$$\begin{aligned} m^a &= m\ \forall\ a & q_{ab}&=\delta_{ab}+q(1-\delta_{ab})\ \forall\ a,b\\ \hat{m}^a &= i\mu\ \forall\ a & \left(\tilde{J}^2\mathbf{q}\right)^{-1}\hat{\mathcal{R}}&=(1-k^2)\left(R_0\delta_{ab}+R_1\right)\ \forall\ a,b\\ \hat{\rho}_a^{\sigma}(h) &= \ln\left[\chi^{\sigma}(h)\right]\ \forall\ a & -\frac{\tilde{J}^2}{2}\mathbf{Q}-\frac{1}{2}\hat{\mathcal{R}}\left(\tilde{J}^2\mathbf{q}\right)^{-1}\hat{\mathcal{R}}&=\frac{Q_0}{2}(1-\delta_{ab})\ \forall\ a,b. \end{aligned}$$

Parisi's ultrametric ansatz for the replica matrices, where the saddle point equations and coefficients for the evolution equation (6.4) are obtained as solutions of non-linear partial differential equations will be discussed in Chapter 7.

6.4.1 Saddle point equations

Using two Gaussian linearisations, to separate the spin and field terms in the logarithm into products over replica space, the saddle point exponent becomes

$$\begin{aligned}\frac{\Psi_{RS}}{n} &= -m\mu+\frac{1}{2}Q_0(q-1)-\frac{1}{2}\ln(1-q)-\frac{q}{2(1-q)}\\ &\quad-\tilde{J}^2(1-k^2)\left[(1-q)^2R_0^2+2q(1-q)R_0^2+2(1-q)^2R_0R_1\right]\\ &\quad-\sum_{\sigma}\int dH\,\rho_t^{\sigma}(H)\ln\chi^{\sigma}(H)+\left\langle\!\!\left\langle\ln\left\langle\int dH\,\mathcal{M}_{RS}(H,\sigma)\right\rangle_{\sigma}\right\rangle\!\!\right\rangle_{wy},\end{aligned}$$

where $\left\langle\!\!\left\langle f(w,y)\right\rangle\!\!\right\rangle_{wy}\equiv\int\!\!\int Dw\,Dy\,f(w,y)\quad\left(Dw\equiv e^{-\frac{1}{2}w^2}\frac{dw}{\sqrt{2\pi}}\right)$, and the measure $\mathcal{M}_{RS}(H,\sigma)$ is defined as

$$\mathcal{M}_{RS}(H,\sigma)=\chi^{\sigma}(H)\,e^{-\frac{\lambda_0^2+\lambda_1^2}{2}\left(H-\tilde{J}_0m\right)^2+\left(\sigma(1-k^2)R_0+w\frac{(1-k^2)R_1}{\sqrt{Q_0}}+y\sqrt{\lambda_1^2-\frac{(1-k^2)^2R_1^2}{Q_0}}\right)\left(H-\tilde{J}_0m\right)+\sigma\left(\mu+\sqrt{Q_0}w\right)}.$$

We now write the average over the measure $\mathcal{M}_{RS}(H,\sigma)$ as

$$\langle f^{\sigma}(H,w,y)\rangle_{\mathcal{M}}=\frac{\sum_{\sigma}\int dH\,f^{\sigma}(H,w,y)\mathcal{M}_{RS}(H,\sigma)}{\sum_{\sigma}\int dH\,\mathcal{M}_{RS}(H,\sigma)},$$

so that (obviously) $\left\langle\!\left\langle \langle f^\sigma(H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} = \sum_\sigma \int dH\, f^\sigma(H)\,\rho^\sigma(H)$. After having removed explicit w, or y dependence by partial integration, the replica symmetric saddle point equations are

$$
\begin{aligned}
m &= \left\langle\!\left\langle \langle\sigma\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} \qquad (6.15)\\
\mu &= -\tilde{J}_0(1-k^2)mR_0 + \tilde{J}_0(\lambda_0^2+\lambda_1^2)\left\langle\!\left\langle \left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\\
q &= \left\langle\!\left\langle \langle\sigma\rangle_{\mathcal{M}}^2 \right\rangle\!\right\rangle_{wy}\\
Q_0 &= \frac{q}{(1-q)^2} - 4\tilde{J}^2(1-k^2)R_0\left[qR_0+2(1-q)R_1\right]\\
&\quad -\frac{2q}{\tilde{J}^2(1-q)^3}\left\langle\!\left\langle \left\langle\left(H-\tilde{J}_0 m\right)^2\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} + \frac{(1+q)}{\tilde{J}^2(1-q)^3}\left\langle\!\left\langle \left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}}^2 \right\rangle\!\right\rangle_{wy}\\
R_0 &= \frac{1}{2\tilde{J}^2(1-q)^2}\left[\left\langle\!\left\langle \left\langle\sigma\left(H-\tilde{J}_0 m\right)\right\rangle_{\mathcal{M}}^2 \right\rangle\!\right\rangle_{wy} - \left\langle\!\left\langle \langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\right]\\
R_1 &= \frac{1}{2\tilde{J}^2(1-q)^2}\left[\left\langle\!\left\langle \left\langle\sigma\left(H-\tilde{J}_0 m\right)\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} - 2\tilde{J}^2(1-q^2)R_0\right]\\
\rho_t^\varsigma(h) &= \left\langle\!\left\langle \langle\delta_{\varsigma,\sigma}\,\delta(h-H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} \equiv \left\langle\!\left\langle \rho_t^\varsigma(h,w,y) \right\rangle\!\right\rangle_{wy},
\end{aligned}
$$

with the short hand $\lambda_0^2 = \frac{1-2q}{\tilde{J}^2(1-q)^2}$ and $\lambda_1^2 = \frac{q}{\tilde{J}^2(1-q)^2}$.

6.4.2 Flow equations

Calculation of the coefficients (6.10) to the the flow equation to be evaluated using the replica symmetric saddle point equations (6.15) involves tortuous algebra, eventually yielding RS flow equations of the form

$$
\begin{aligned}
\frac{d}{dt}\rho_t^\varsigma(h) &= \frac{1}{2}\left(1+\varsigma\tanh(\beta h)\right)\rho_t^{-\varsigma}(h) - \frac{1}{2}\left(1-\varsigma\tanh(\beta h)\right)\rho_t^\varsigma(h) \qquad (6.16)\\
&\quad +\frac{\partial}{\partial h}\left[\rho_t^\varsigma(h)\left\{\left(h-\tilde{J}_0 m\right)\Phi_1(t)+\Phi_2^\varsigma(t)\right\}\right]\\
&\quad +\frac{\partial}{\partial h}\left[\Phi_3^\varsigma(h,t)\right] + \frac{\partial^2}{\partial h^2}\left[\rho_t^\varsigma(h)\Phi_4(t)\right]
\end{aligned}
$$

where

$$
\begin{aligned}
\Phi_1 &= 1-\tilde{J}^2\lambda_0^2\left\langle\!\left\langle \langle\sigma\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} - \tilde{J}^2\lambda_1^2\left\langle\!\left\langle \langle\sigma\rangle_{\mathcal{M}}\langle\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\\
\Phi_2^\varsigma &= \tilde{J}_0 m - \tilde{J}_0\left\langle\!\left\langle \langle\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} - \varsigma\frac{\tilde{J}^2(1-k^2)}{(1+k^2)}\times\\
&\quad \left[\lambda_0^2\left\langle\!\left\langle \left\langle\left(H-\tilde{J}_0 m\right)\tanh(\beta H)\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} + \lambda_1^2\left\langle\!\left\langle \left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}}\langle\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\right]\\
&\quad +\varsigma\frac{2\tilde{J}^2(1-k^2)}{(1+k^2)}\left[(R_0+R_1)\left\langle\!\left\langle \langle\sigma\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} - R_1\left\langle\!\left\langle \langle\sigma\rangle_{\mathcal{M}}\langle\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\right]\\
\Phi_3^\varsigma &= \frac{\tilde{J}^2(1-k^2)}{(1+k^2)}\left\langle\!\left\langle \rho_t^\varsigma(h,w,y)\langle\sigma\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\times\\
&\quad \left[\left(\lambda_0^2+2\lambda_1^2\right)\left\langle\!\left\langle \left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}}\langle\tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy} - \lambda_1^2\left\langle\!\left\langle \left\langle\left(H-\tilde{J}_0 m\right)\tanh(\beta H)\right\rangle_{\mathcal{M}} \right\rangle\!\right\rangle_{wy}\right]
\end{aligned}
$$

$$
\begin{aligned}
&+\tilde{J}^2 \left\langle\!\!\left\langle \rho_t^{\varsigma}(h,w,y) \left\langle H - \tilde{J}_0 m \right\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \times \\
&\left[\left(\lambda_0^2 + 2\lambda_1^2\right) \left\langle\!\!\left\langle \langle \tanh(\beta H)\rangle_{\mathcal{M}} \langle\sigma\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} - \lambda_1^2 \left\langle\!\!\left\langle \langle \sigma \tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \right] \\
&-\frac{2\tilde{J}^2(1-k^2)}{(1+k^2)} \left\langle\!\!\left\langle \rho_t^{\varsigma}(h,w,y) \langle\sigma\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \times \\
&\left[(R_0 - R_1) \left\langle\!\!\left\langle \langle\sigma\rangle_{\mathcal{M}} \langle \tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} + R_1 \left\langle\!\!\left\langle \langle \sigma \tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \right] \\
\Phi_4 \;=\; & \tilde{J}^2 \left[1 - \left\langle\!\!\left\langle \langle \sigma \tanh(\beta H)\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \right]. \qquad (6.17)
\end{aligned}
$$

6.5 Analysis of the RS flow

We here present comparisons between numerical solutions of the flow equations (6.16) using (6.15), and monte-carlo simulations using 8,000 spins for the fully symmetric model ($k = 0$) and 5,600 spins for the fully asymmetric model ($k = 1$); as well as analysing some special cases of the flow equations.

The iteration of the flow equations (6.16) requires a formidable computational effort, since the saddle point equations have to be solved at each iterations step. Solution of the saddle point equations is equivalent to free energy minimisation of a continuous function. Even for a reasonably modest partitioning of the local field axis ($\simeq 15$), this already requires simultaneous solution of more than 30 coupled non-linear equations, involving a three layered Gaussian integration. Because of the difficulty in solving the saddle point equations, the only results presented here involve evolution from an initial state with spins chosen at random (to give the required magnetisation), rather than the more interesting cooling in an external field, and other experiments associated with spin glasses.

6.5.1 Initial Conditions

The spins in the initial state are drawn at random from a biased distribution in order to achieve a given initial magnetisation m. The initial probability density is therefore

$$
p_0(\boldsymbol{\sigma}) = \prod_k \left\{ \frac{1}{2} \sum_{\varsigma'} (1 + \varsigma' m) \delta_{\sigma_k, \varsigma'} \right\}.
$$

We use this to calculate the initial value of the distribution $\rho_0^{\varsigma}(h)$, assuming it to be self-averaging with respect to the interactions J_{ij}:

$$
\begin{aligned}
\rho_0^{\varsigma}(h) \;&=\; \left\langle \frac{1}{N} \sum_i \sum_{\boldsymbol{\sigma}} \prod_k \left\{ \frac{1}{2} \sum_{\varsigma'} (1 + \varsigma' m) \delta_{\sigma_k, \varsigma'} \right\} \delta_{\sigma_i, \varsigma}\, \delta(h - h_i(\boldsymbol{\sigma})) \right\rangle_{J_{ij}} \\
&=\; \frac{1}{2}(1 + \varsigma m) \frac{1}{N} \sum_i \int \frac{d\hat{h}}{2\pi} e^{i\hat{h}h} \left\langle \sum_{\boldsymbol{\sigma}} \prod_{k \neq i} \left\{ \frac{1}{2} \sum_{\varsigma'} (1 + \varsigma' m) \delta_{\sigma_k, \varsigma'} \right\} e^{-i\hat{h}\sigma_k \left(\frac{\tilde{J}_0}{N} + \tilde{J} \frac{z_{ik}^s + k z_{ik}^a}{\sqrt{N(1+k^2)}} \right)} \right\rangle_{z_{ik}},
\end{aligned}
$$

where the $z_{ik}^{s,a}$'s are drawn from an unbiased Gaussian distribution with unit width. z_{ik}^s is symmetric, whilst z_{ik}^a is antisymmetric. Because of the structure of the exponent, we can treat $\frac{z_{ik}^s + k z_{ik}^a}{\sqrt{(1+k^2)}}$

as one Gaussian variable. Performing the spin average, then the disorder average we get

$$\begin{aligned}
\rho_0^\varsigma(h) &= \frac{1}{2}(1+\varsigma m)\frac{1}{N}\sum_i \int \frac{d\hat{h}}{2\pi} e^{i\hat{h}h} \int \frac{dz_{ik}}{\sqrt{2\pi}} e^{-\frac{1}{2}{z_{ik}}^2} \\
&\qquad \prod_k \frac{1}{2}\left[(1+m)e^{-i\hat{h}\left(\frac{\tilde{J}_0}{N}+\frac{\tilde{J}}{\sqrt{N}}z_{ik}\right)} + (1-m)e^{i\hat{h}\left(\frac{\tilde{J}_0}{N}+\frac{\tilde{J}}{\sqrt{N}}z_{ik}\right)}\right] \\
&= \frac{1}{2}(1+\varsigma m)\int \frac{d\hat{h}}{2\pi} e^{i\hat{h}h} \prod_k e^{-\frac{\hat{h}^2\tilde{J}^2}{2N}} \frac{1}{2}\left[(1+m)e^{-i\frac{\hat{h}\tilde{J}_0 m}{N}} + (1-m)e^{i\frac{\hat{h}\tilde{J}_0 m}{N}}\right] \\
&= \frac{1}{2}(1+\varsigma m)\int \frac{d\hat{h}}{2\pi} e^{i\hat{h}h - \frac{\tilde{J}^2}{2}\hat{h}^2 + N\log\left[\cos\left(\frac{\tilde{J}_0\hat{h}}{N}\right) - im\sin\left(\frac{\tilde{J}_0\hat{h}}{N}\right)\right]} \\
&= \frac{1}{2}(1+\varsigma m)\frac{e^{-\frac{(h-\tilde{J}_0 m)^2}{2\tilde{J}^2}}}{\sqrt{2\pi\tilde{J}^2}}.
\end{aligned} \tag{6.18}$$

Therefore evaluating the saddle point equations (6.15) using (6.18) we see that initially $\chi^\varsigma(h) = \frac{1}{2}(1+\varsigma m)\ \forall\ h$, $\mu = R_0 = R_1 = Q_0 = 0$ and $q = m^2$.

6.5.2 Detailed balance equilibrium

By analogy with the full replica result, we expect that for the symmetric model (i.e. with $k = 0$) in detailed balance equilibrium $\chi^\sigma(h) = e^{\frac{1}{2}\sigma\beta h}$, will give fixed points of the RS flow equations (6.16), as well as recovering the RS saddle point equations.

Using $\chi^\sigma(h) = e^{\frac{1}{2}\sigma\beta h}$ we find that after one of the Gaussian integrals (over w) has been performed the equilibrium saddle point equations (6.15) are

$$\begin{aligned}
m &= \int Dy\ \tanh\beta\left(\tilde{J}_0 m + \tilde{J}\sqrt{q}y\right) & \mu &= \frac{\beta}{2}\tilde{J}_0 m \\
q &= \int Dy\ \tanh^2\beta\left(\tilde{J}_0 m + \tilde{J}\sqrt{q}y\right) & Q_0 &= 0 \\
R_0 &= \frac{\beta}{2} & R_1 &= 0,
\end{aligned} \tag{6.19}$$

so that the equilibrium field distribution is

$$\rho_{eq}^\varsigma(h) \equiv \left\langle \langle \delta_{\varsigma,\sigma}\,\delta(h-H)\rangle_{\mathcal{M}_{eq}}\right\rangle_y = \int Dy\ \frac{e^{-\frac{1}{2\tilde{J}^2(1-q)}\left(h-\tilde{J}_0 m\right)^2 + y\frac{\sqrt{q}}{\tilde{J}(1-q)}\left(h-\tilde{J}_0 m\right)+\varsigma\beta h}}{\sum_\sigma \int dH\, e^{-\frac{1}{2\tilde{J}^2(1-q)}\left(H-\tilde{J}_0 m\right)^2 + y\frac{\sqrt{q}}{\tilde{J}(1-q)}\left(H-\tilde{J}_0 m\right)+\sigma\beta H}}. \tag{6.20}$$

Using this we can calculate the terms appearing in the evolution equation (6.16). Firstly, the

$$\frac{1}{2}\left(1+\varsigma\tanh(\beta h)\right)\rho_t^{-\varsigma}(h) - \frac{1}{2}\left(1-\varsigma\tanh(\beta h)\right)\rho_t^{\varsigma}(h)$$

term is zero, because of the $e^{\varsigma\beta h}$ appearing in the equation for $\rho^\varsigma(h)$.

We now have to calculate the coefficients (6.17) using the equilibrium form of the field distribution (6.20), to confirm that it is indeed stationary. Now

$$\begin{aligned}
q &= \left\langle \langle \tanh(\beta H)\rangle_{\mathcal{M}_{eq}} \langle\sigma\rangle_{\mathcal{M}_{eq}}\right\rangle_y \\
\beta &= \lambda_0^2\left\langle\left\langle\left(H-\tilde{J}_0 m\right)\tanh(\beta H)\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y + \lambda_1^2\left\langle\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\langle\tanh(\beta H)\rangle_{\mathcal{M}_{eq}}\right\rangle_y,
\end{aligned}$$

so that

$$\frac{\Phi_1}{\Phi_4} = \lambda_0^2 \qquad \text{and} \qquad \frac{\Phi_2^\varsigma}{\Phi_4} = -\varsigma\beta.$$

Consider now $\frac{\partial}{\partial h}\rho^\varsigma(h)$

$$\begin{aligned}\frac{\partial \rho^\varsigma(h)}{\partial h} &= \left\langle \left\langle \delta_{\varsigma,\sigma}\,\delta(h-H)\left[-(\lambda_0^2+\lambda_1^2)(H-\tilde{J}_0 m)+\lambda_1 y+\sigma\beta\right]\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y \\ &= \left(-\lambda_0^2(h-\tilde{J}_0 m)+\varsigma\beta\right)\rho^\varsigma(h) - \lambda_1^2\left\langle \rho_t^\varsigma(h,y)\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y,\end{aligned}$$

hence

$$\left((h-\tilde{J}_0 m)\Phi_1+\Phi_2^\varsigma\right)\rho^\varsigma(h)+\Phi_4\frac{\partial}{\partial h}\rho^\varsigma(h) = -\lambda_1^2\left\langle \rho_t^\varsigma(h,y)\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y.$$

Also

$$\begin{aligned}\Phi_3^\varsigma(h) &= \tilde{J}^2\left\langle \rho_t^\varsigma(h,y)\langle\sigma\rangle_{\mathcal{M}_{eq}}\right\rangle_y\left(\lambda_0^2+\lambda_1^2\right)\left[\left\langle\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\langle\tanh(\beta H)\rangle_{\mathcal{M}_{eq}}\right\rangle_y\right. \\ &\qquad \left. -\left\langle\left\langle\left(H-\tilde{J}_0 m\right)\tanh(\beta H)\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y+\frac{\beta(1-q)}{(\lambda_0^2+\lambda_1^2)}\right] \\ &\quad +\tilde{J}^2\lambda_1^2\left\langle \rho_t^\varsigma(h,y)\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y\left[1-\left\langle\langle\sigma\tanh(\beta H)\rangle_{\mathcal{M}_{eq}}\right\rangle_y\right].\end{aligned}$$

The last term cancels with a similar term from $\Phi_4\frac{\partial}{\partial h}\rho^\varsigma(h)$, and the first term is identically zero, since

$$\begin{aligned}&\left\langle\left\langle\left(H-\tilde{J}_0 m\right)\tanh(\beta H)\right\rangle_{\mathcal{M}_{eq}}\right\rangle_y - \left\langle\left\langle H-\tilde{J}_0 m\right\rangle_{\mathcal{M}_{eq}}\langle\tanh(\beta H)\rangle_{\mathcal{M}_{eq}}\right\rangle_y \\ &= \frac{1}{\lambda_1}\int Dy\,\frac{\partial}{\partial y}\frac{\sum_\varsigma \varsigma\int dh\, e^{-\frac{\lambda_0^2+\lambda_1^2}{2}\left(h-\tilde{J}_0 m\right)^2+y\lambda_1\left(h-\tilde{J}_0 m\right)+\varsigma\beta h}}{\sum_\sigma\int dH\, e^{-\frac{\lambda_0^2+\lambda_1^2}{2}\left(H-\tilde{J}_0 m\right)^2+y\lambda_1\left(H-\tilde{J}_0 m\right)+\sigma\beta H}} \\ &= \frac{\beta}{\lambda_0^2+\lambda_1^2}(1-q).\end{aligned} \tag{6.21}$$

Hence in detailed balance equilibrium, (6.20) is indeed a fixed point of the flow equation (6.16) and $\chi^\sigma(h) = e^{\frac{1}{2}\sigma\beta h}$.

6.5.3 Flows for symmetric interactions ($k = 0$)

In Figure 6.1 we can see comparisons of flows from randomly chosen initial states, between monte-carlo simulations with 8,000 spins (solid lines) and solution of the RS dynamic equations (6.16) for the parameters $\tilde{J}_0 = 0, 1$ and $T = 0, 1$, with initial magnetisations $m_0 = 0, 0.3$. As we can see, in most cases the fit between theory and experiment for both the magnetisation m and energy E, both calculated from the distribution using (6.2), is almost perfect (within finite size effects in the simulations), on the intermediate time scales studied ($t \sim 0-6$ iterations/spin). There is no evidence of the inaccuracies encountered with the previous simple two-parameter (CS) theory, where the slowing down was not fully accounted for.

The one slight discrepancy is the case $T = 0$, $\tilde{J}_0 = 1$, $m_0 = 0.3$. Here the theoretical prediction strays slightly from the simulations. This can be traced back to two sources. Firstly we can see that close to the beginning of the flow, there is a significant fluctuation, caused by finite size

effects. This sets the flow off on a slightly higher path than the rest of the simulation would predict. Secondly, the solution of the saddle point equations (6.15) is difficult in the spin-glass region. The finite accuracy of the algorithm can build up over time, as the solution progresses, producing significant discrepancies. Indeed the algorithm fails entirely around $t = 4$. For these parameter values we have been unable to obtain results for even moderate times. It must be stressed however, that these discrepancies are manifestations of inaccuracies in the computation, rather than the theory. Within the limits of finite size effects, and computational accuracy, the theory appears to be if not exact, then a very good approximation. Applying this theory to the Hopfield neural network, where because of the simplicity of interactions greater system sizes can be used, or to a toy model [46] where an exact solution is possible, could help to further determine whether the present theory is exact or not.

Figure 6.2 shows comparisons with simulations of the distributions themselves, at intervals of 2 iterations/spin during the dynamics. Obviously there are finite size effects, since we need to count the number of sites with local field in some range in order to calculate the distribution. However, again fit here is very good, with no evidence of the accelerated broadening associated with the old theory.

From both Figures 6.1, 6.2 we notice that the majority of relaxation occurs within the first couple of iterations per spin. Past about $t = 2$ iterations/spin, the relaxation towards the equilibrium results is very slow. Inspecting 6.2 there is very little difference between the distributions themselves for $t = 2, 4, 6$, however crucial changes are occuring. These are particularly significant for the zero temperature state, where as mentioned in [85, 86] zero field distribution (i.e. $\rho^+(0) + \rho^-(0)$) is zero in equilibrium. In fact in equilibrium at $T = 0$, all spins must be aligned to their fields, hence the joint distribution is zero for $\mathrm{sgn}[h] \neq \varsigma$. Relaxation towards this state is however very slow. Due to the disorder it may take a very long time for the last few spins to flip [87].

6.5.4 Flows for fully asymmetric interactions ($k = 1$)

Complete asymmetry ($k = 1$) represents a special case. Many of the terms in the saddle point and flow equations drop out. We note that solutions exist of the form

$$\rho_t^+(h) + \rho_t^-(h) \propto e^{-\frac{(h-\tilde{J}_0 m_t)^2}{2\phi_t}} \tag{6.22}$$

with $\chi^+(h) + \chi^-(h) = \chi \ \forall \ h$. The evolution equations for m_t and ϕ_t are obtained by plugging (6.22) into (6.16) with $k = 1$, upon realising that

$$\left\langle\!\!\left\langle \left(\rho_t^+(h,w,y) + \rho_t^-(h,w,y)\right) \left\langle \left(H - \tilde{J}_0 m\right) \right\rangle_{\mathcal{M}} \right\rangle\!\!\right\rangle_{wy} \propto \left(h - \tilde{J}_0 m\right) \left(\rho_t^+(h) + \rho_t^-(h)\right),$$

with the current specific choice for $\chi^+(h) + \chi^-(h)$. These solutions are of the same form as obtained by Crisanti and Sompolinsky [53] using the path integral formalism. The local field distribution remains Gaussian. This is because with fully asymmetric bonds correlations cannot build up.

As with the fully symmetric case, comparison with monte-carlo simulations shows very good agreement. Figure 6.3 shows flow of the order parameters m and E as a function of time. Although

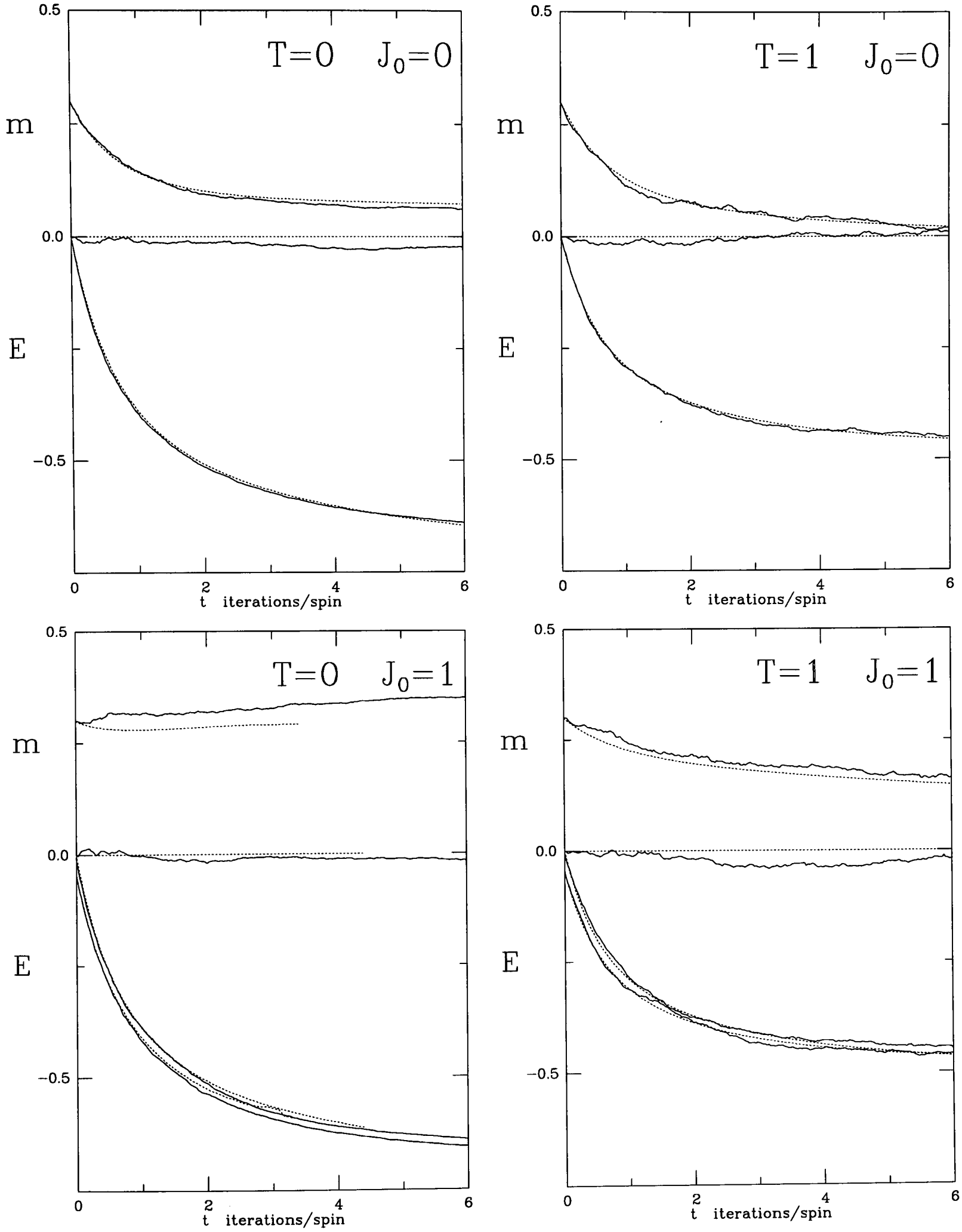
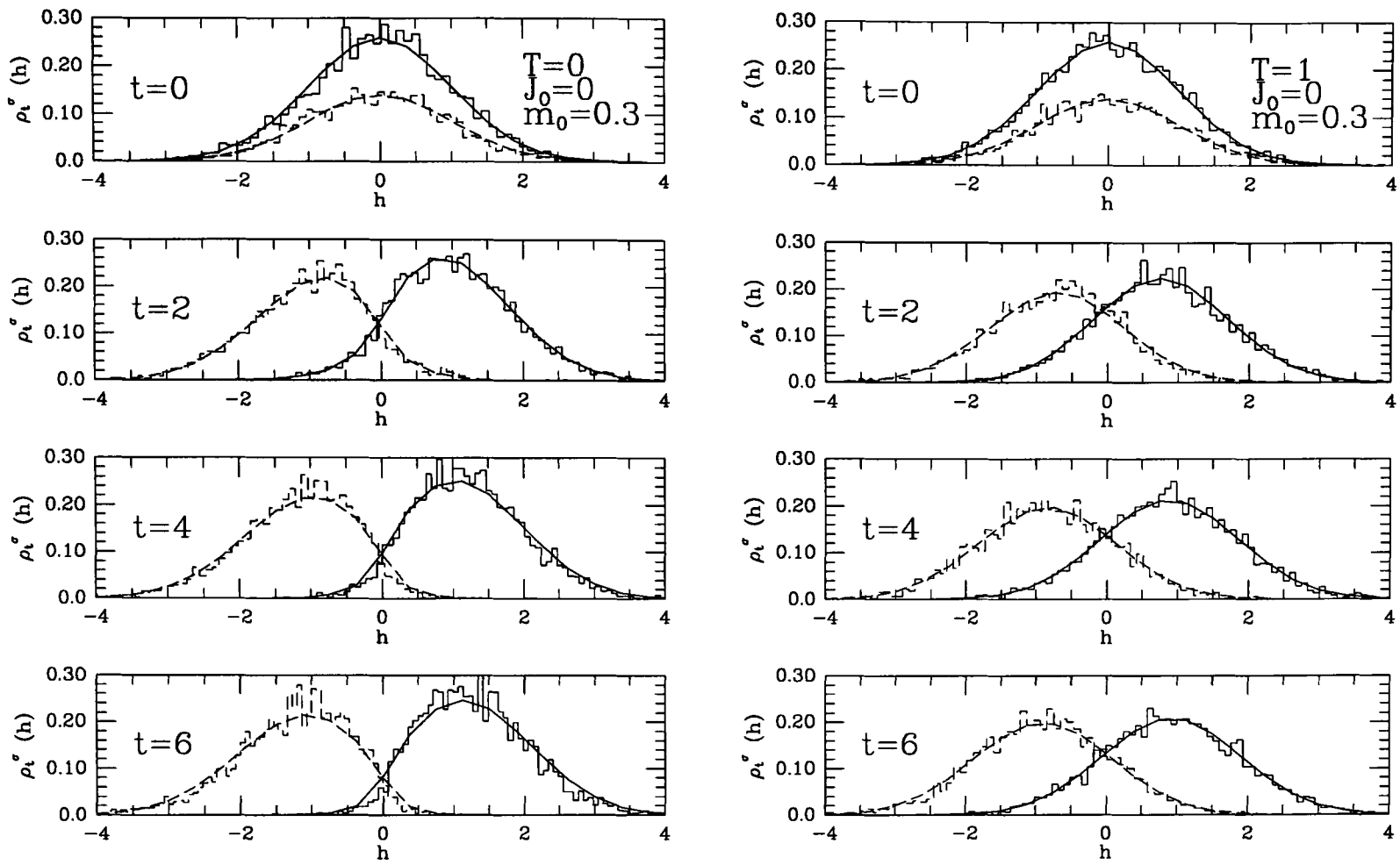


Figure 6.1: Comparison between monte-carlo simulation with 8,000 spins (solid line), and theory (dashed line), for the fully symmetric S-K model ($k = 0$) with $\tilde{J}_0 = 0, 1$ and $T = 0, 1$.



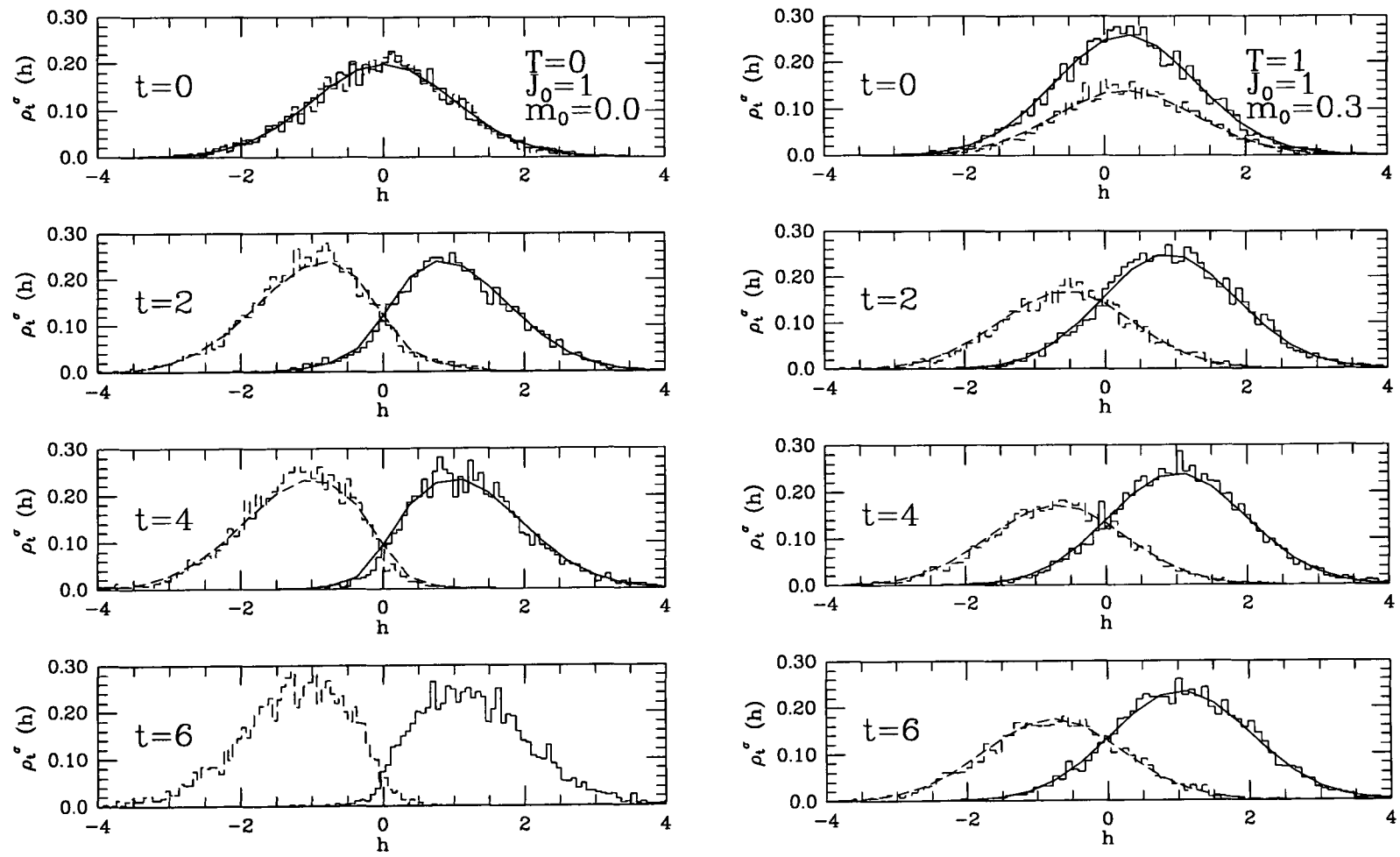


Figure 6.2: Comparison between monte-carlo simulation with 8,000 spins (histograms), and theory (lines), for the fully symmetric S-K model ($k = 0$) with $\tilde{J}_0 = 0, 1$, $T = 0, 1$ and $m_0 = 0$ or 0.3. Solid lines represent $\rho^+(h)$ whilst dashed lines represent $\rho^-(h)$

finite size effect here are larger, due to the smaller system size (since we are storing twice as many interactions), the fit between theory (dashed lines) and experiment (solid lines) is very good. Figure 6.4 shows the evolution of the field distributions as a function of time, with again an excellent fit between theory (lines) and experiment (histograms).

6.6 The A-T instability

The de Almeida-Thouless (AT) line [50] marks the instability of the replica symmetric saddle points, to the so-called replicon mode bifurcations. Unlike the original calculation, here we have to worry about fluctuations with respect to three replica matrices. Furthermore the A-T instability is not simply a function of a small number of parameters, but of a whole distribution. This makes calculation of the A-T line difficult.

First we make a convenient change of variables:

$$\mathcal{G} = -\frac{i}{2}\tilde{J}^2\mathbf{Q} - \frac{i}{2}\hat{\mathcal{R}}\left(\tilde{J}^2\mathbf{q}\right)^{-1}\hat{\mathcal{R}} \qquad \hat{\mathcal{E}} = \frac{i}{(1-k^2)}\left(\tilde{J}^2\mathbf{q}\right)^{-1}\hat{\mathcal{R}}$$

The replicon mode bifurcations are then of the form

$$q_{ab} \to q + \delta q_{ab} \qquad \mathcal{G}_{ab} = \frac{i}{2}Q_0 + i\delta\mathcal{G}_{ab} \qquad \hat{\mathcal{E}}_{ab} = iR + i\delta\hat{\mathcal{E}}_{ab}$$

with $\sum_{a\neq b}\delta q_{ab} = \sum_{a\neq b}\delta\hat{q}_{ab} = \sum_{a\neq b}\delta\hat{\mathcal{E}}_{ab} = 0$. Expanding around the RS saddle point we have

$$\begin{aligned}
\Psi - \Psi_{RS} \;=\; & \frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta q_{ab}\delta q_{cd}\frac{\partial^2\Psi}{\partial q_{ab}\partial q_{cd}} - \frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\mathcal{G}_{ab}\delta\mathcal{G}_{cd}\frac{\partial^2\Psi}{\partial\mathcal{G}_{ab}\partial\mathcal{G}_{cd}} \\
& -\frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\hat{\mathcal{E}}_{ab}\delta\hat{\mathcal{E}}_{cd}\frac{\partial^2\Psi}{\partial\hat{\mathcal{E}}_{ab}\partial\hat{\mathcal{E}}_{cd}} \\
& +\frac{i}{2}\sum_{a\neq b}\sum_{c\neq d}\delta q_{ab}\delta\mathcal{G}_{cd}\frac{\partial^2\Psi}{\partial q_{ab}\partial\mathcal{G}_{cd}} + \frac{i}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\mathcal{G}_{ab}\delta q_{cd}\frac{\partial^2\Psi}{\partial\mathcal{G}_{ab}\partial q_{cd}} \\
& +\frac{i}{2}\sum_{a\neq b}\sum_{c\neq d}\delta q_{ab}\delta\hat{\mathcal{E}}_{cd}\frac{\partial^2\Psi}{\partial q_{ab}\partial\hat{\mathcal{E}}_{cd}} + \frac{i}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\hat{\mathcal{E}}_{ab}\delta q_{cd}\frac{\partial^2\Psi}{\partial\hat{\mathcal{E}}_{ab}\partial q_{cd}} \\
& -\frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\mathcal{G}_{ab}\delta\hat{\mathcal{E}}_{cd}\frac{\partial^2\Psi}{\partial\mathcal{G}_{ab}\partial\hat{\mathcal{E}}_{cd}} - \frac{1}{2}\sum_{a\neq b}\sum_{c\neq d}\delta\hat{\mathcal{E}}_{ab}\delta\mathcal{G}_{cd}\frac{\partial^2\Psi}{\partial\hat{\mathcal{E}}_{ab}\partial\mathcal{G}_{cd}}.
\end{aligned}$$

In contrast to other analysis' of spin glasses, the parts Ψ dependent on q_{ab}, $\mathcal{G}_{ab}$ and $\hat{\mathcal{E}}_{ab}$ do not separate. We have to calculate the nine second derivatives of

$$\ln\left[\int\left(\prod_a \frac{dH^a}{\sqrt{2\pi\det(\tilde{J}^2\mathbf{q})}}\right)e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}\right].$$

Due the symmetry under interchange of the order of differentiation, we need only calculate six quantities.

Since the replicon mode has $\sum_{a\neq b}\delta q_{ab} = \sum_{a\neq b}\delta\mathcal{G}_{ab} = \sum_{a\neq b}\delta\hat{\mathcal{E}}_{ab} = 0$ the only important terms are those containing two delta functions $\delta_{ac}\delta_{bd}$ or $\delta_{ad}\delta_{bc}$.

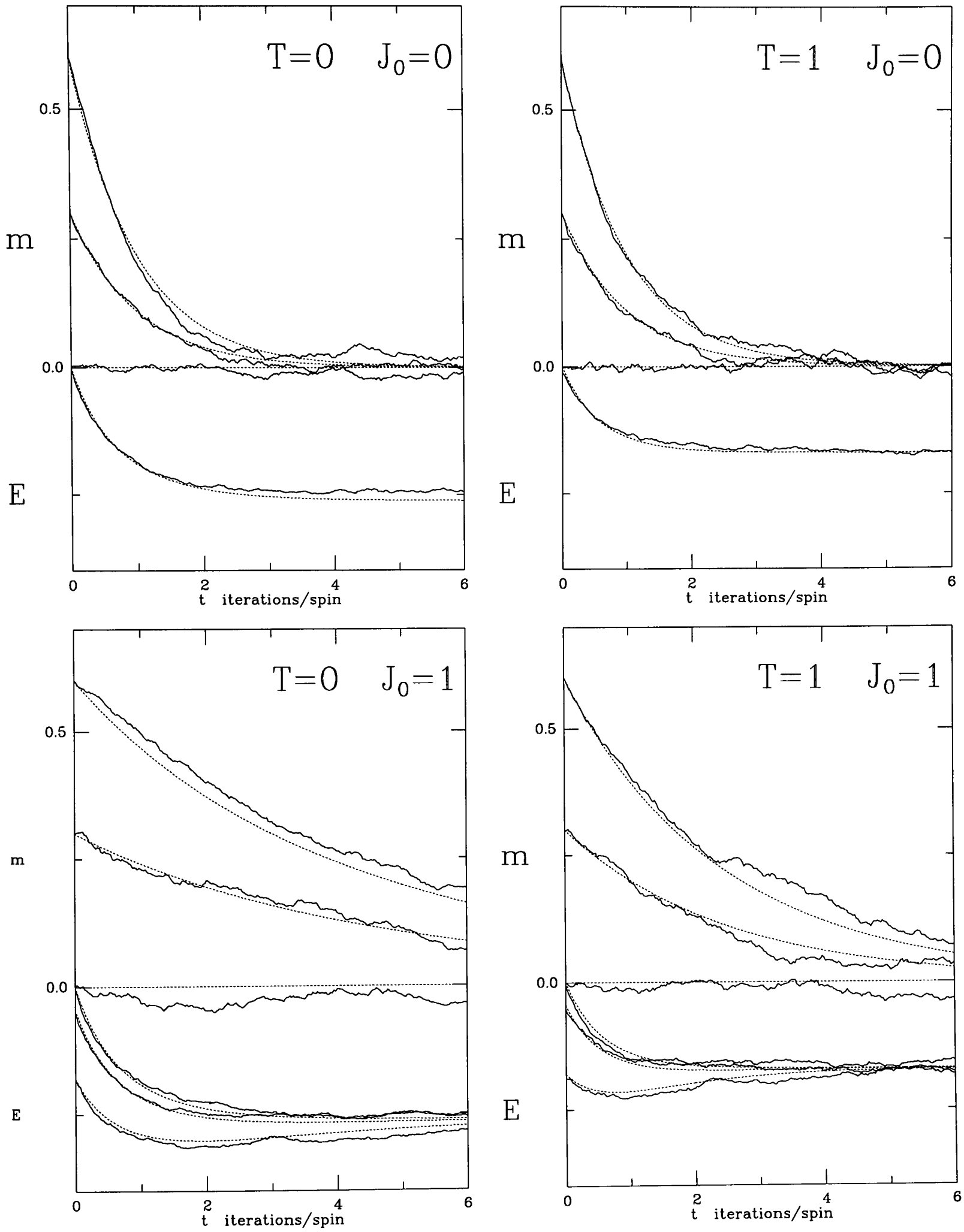


Figure 6.3: Comparison between monte-carlo simulation with 5,600 spins (solid line), and theory (dashed line) for the fully asymmetric S-K model ($k = 1$), with $\tilde{J}_0 = 0, 1$ and $T = 0, 1$.

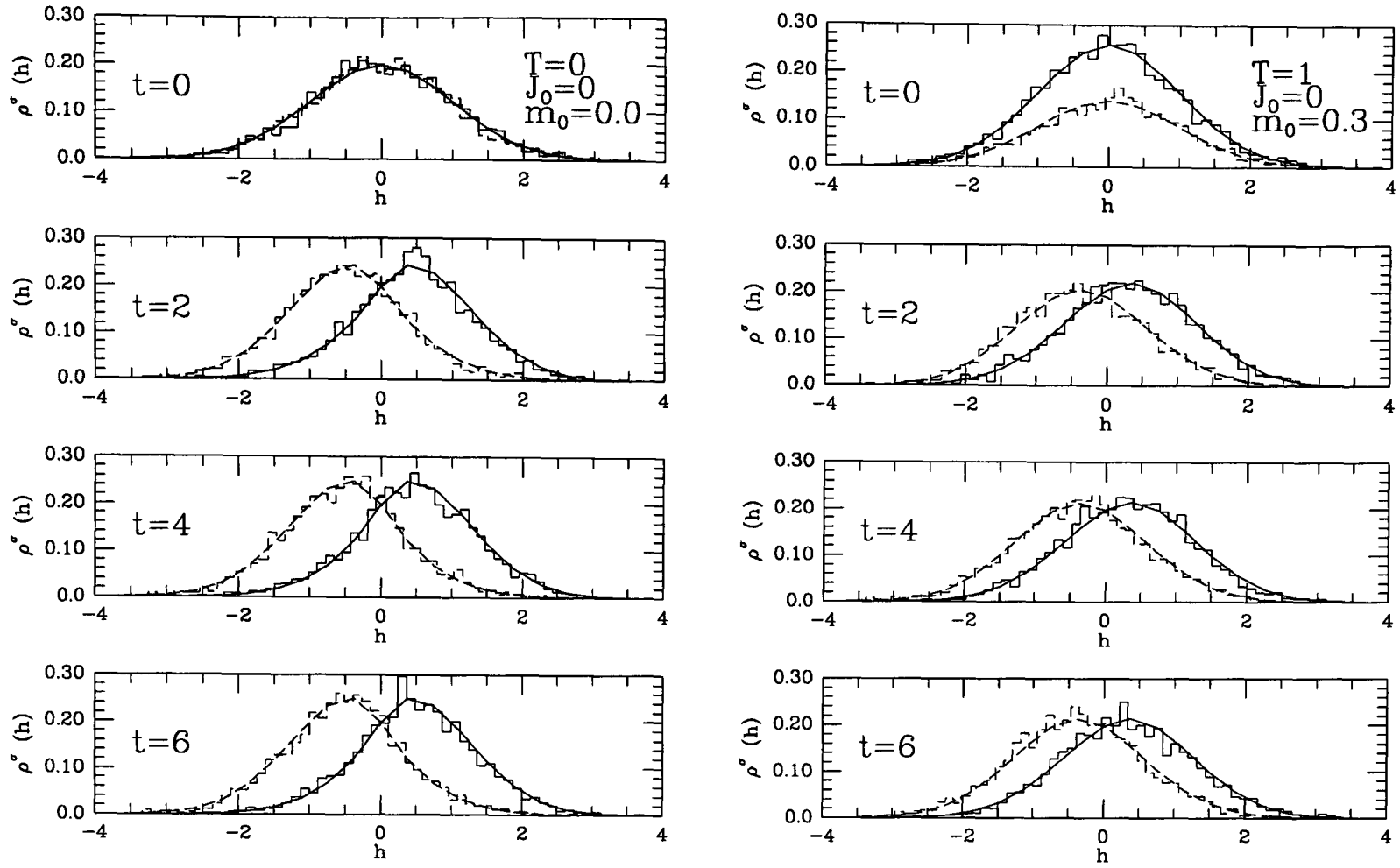


Figure 6.4: Comparison between monte-carlo simulation with 5,600 spins (histograms), and theory (lines) for the fully asymmetric S-K model ($k = 1$), with $\tilde{J}_0 = 0$, $T = 0$ and $m_0 = 0$, 0.3 or 0.6. Solid lines represent $\rho^+(h)$ whilst dashed lines represent $\rho^-(h)$

The A-T line, signifying instability of the RS solution to the saddle point equation, occurs when the the matrix of derivatives has determinant equal to zero:

$$\begin{vmatrix} \frac{\partial^2\Psi}{\partial q_{ab}\partial q_{cd}} & i\frac{\partial^2\Psi}{\partial q_{ab}\partial \mathcal{G}_{cd}} & i\frac{\partial^2\Psi}{\partial q_{ab}\partial \hat{\mathcal{E}}_{cd}} \\ i\frac{\partial^2\Psi}{\partial \mathcal{G}_{ab}\partial q_{cd}} & -\frac{\partial^2\Psi}{\partial \mathcal{G}_{ab}\partial \mathcal{G}_{cd}} & -\frac{\partial^2\Psi}{\partial \mathcal{G}_{ab}\partial \hat{\mathcal{E}}_{cd}} \\ i\frac{\partial^2\Psi}{\partial \hat{\mathcal{E}}_{ab}\partial q_{cd}} & -\frac{\partial^2\Psi}{\partial \hat{\mathcal{E}}_{ab}\partial \mathcal{G}_{cd}} & -\frac{\partial^2\Psi}{\partial \hat{\mathcal{E}}_{ab}\partial \hat{\mathcal{E}}_{cd}} \end{vmatrix} = 0 \qquad a \neq b \ , \ c \neq d \ . \tag{6.23}$$

We leave the details of the calculation of the coefficients to the Appendix C. Replica symmetry is stable as long as the matrix (6.23) has a negative determinant (since we are investigating a maximum of Ψ). Since there AT matrix is 3×3, there are three eigenvalues, and therefore possibly three zero's of the determinant. Starting from a region where replica symmetry is stable and allowing the dynamics to evolve, the AT 'line', should be interpreted as the first zero determinant of (6.23).

If we evaluate the matrix for the initial state $t = 0$, we find

$$\begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & (1-q)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Hence replica symmetry is only marginally stable already at $t = 0$, higher order derivatives are required to determine replica stability with respect to $\hat{\mathcal{E}}_{ab}$, since it it identically zero for zero time when the distributions are Gaussian. The upper left corner has negative determinant - indicating stability, since we are investigating the stability of a maximal extremum.

For all subsequent times the matrix $\hat{\mathcal{E}}_{ab}$ becomes non-trivial, and we have to evaluate the full determinant with elements given in Appendix C to determine whether the replica symmetric solutions are stable.

NB The A-T instability is no longer given by a line, but by an infinte dimensional surface defined by the distribution $\rho_t^{\varsigma}(h)$. There are still an infinite number of distributions giving rise to the same values of (say) the order parameters m and E, some of which may be RS stable, and some not.

6.7 Conclusion

We have presented a theory to describe the dynamics of the generalised S-K model, on finite time scales, in terms of the deterministic evolution of an order parameter function - the joint spin-field distribution. We have shown that upon making the two vital assumptions of self-averaging of the flow equations, and equipartitioning of probability within the macroscopic sub-shells, the coefficients can be calculated using a replica identity. The evaluation of the non-trivial coefficients involves solution of a dynamic replica saddle point problem. In detailed balance equilibrium the fixed points of the flow equations reproduce exactly the full replica saddle point equations of the equilibrium statistical mechanical calculation.

We have evaluated the saddle point equations, and the evolution equation coefficients within the replica symmetric ansatz. Comparison with simulations shows that the RS theory provides an excellent fit, on moderate time scales ($t \sim 0 - 6$ iterations/spin) for both the fully symmetric

($k = 0$), and fully asymmetric ($k = 1$) models. The critical slowing missed by the previous two parameter (CS) theory, is fully accounting for by the new theory.

The theory contains three replica matrices, hence the AT line, marking the instability of the RS solutions to replicon mode fluctuations, occurs when the determinant of a 3×3 matrix is zero. The coefficients in this matrix are evaluated in Appendix C.

The agreement between RS theory, and monte-carlo simulations is within the boundaries of finite size effects. The theory is either a very good approximation, or even exact. In order to determine which, the method could be applied to an exactly soluble toy model [46], or the Hopfield model near saturation, where the simplicity of interactions allows much greater system sizes.

It would be interesting to see if any insight could be gained into what approximations are being made in this theory, by comparing solutions with the exact (though formal) dynamic equations in terms of correlation and response functions of the path integral approach [39, 40, 53, 55].

This method could be used to investigate the complex temporal behaviour such as aging encountered in spin glasses, to determine whether these effects are genuinely manifestations of the breaking up of phase space (manifested in the mathematics by the breaking of replica symmetry), or whether a sufficiently complex RS theory could capture the essential features of the flow. A sketch of the full replica symmetry broken solution (*à la* Parisi) will be presented in Chapter 7. Considering the difficulty encountered in solving even the RS theory computationally however, solving the full RSB theory may be unrealistic.

Chapter 7

Aspects of replica symmetry breaking

7.1 Introduction

In Chapters 5 and 6 we saw how dynamic equations for the flow of order parameters in the generalised Hopfield and generalised S-K models could be derived, using the replica identity (1.8) to enable us to carry out an average over a constrained region of phase space. Because in both models the interactions are infinite range, we believe the mean field theory to be exact, and in both theories we assumed replica symmetry in order to evaluate the saddle point equations. It must be stressed however that this was only an ansatz - a trial solution which we apply in the absence of other information, the validity of which we must check *a posteriori*. It is well known that the RS phase is unstable at low temperatures [50] - replica symmetry is broken. Indeed even in the simplest imaginable model of this type - the infinite-range ferromagnet there is a breaking of symmetry at low temperatures between the positive and negative magnetisation phases; because a transition between these two phases requires flipping every spin, there is an extensive energy barrier between them, and on finite time scales the system is stuck in one or the other - ergodicity is broken. Things are complicated further in models with an extensive number of disordered interactions per site. Because the interactions are disordered, the energy landscape consists of a large number of 'valleys' and 'hills' (which in the thermodynamic limit become extensively large); the energy landscape is however locally smooth, since a single spin flip will have only a small effect. This leads to a rich variety of behaviour, with a large number ($\sim e^{N^{\frac{1}{4}}}$ [88]) of metastable states.

The spin-glass order parameter, which measures the degree of overlap between two systems with the same realisations of disorder but different thermal baths, becomes highly non-trivial at low temperatures when the system can settle into anyone of the non-equivalent valleys. Since when $N \to \infty$ the energy barriers between states can become extensive, for sufficiently low temperatures the system can get trapped in a finite region of phase space - i.e. ergodicity is broken. The value of the spin-glass order parameter obviously depends on exactly which valleys the two systems become stuck in.

These problems are manifested in the mathematics when the limit $n \to 0$ is taken - the spin average can be performed for n positive integer, the case $n \to 0$ must be treated as an analytic

continuation. In the regions of phase space where the replica symmetric saddle points are found to be unstable [50], replica symmetry breaking effects [16, 19] need to be taken into account. These are manifestations of the breaking of ergodicity [51].

7.2 The Ultrametric Structure

In order to derive what appears to be an exact solution [89] to the S-K model, Parisi introduced an (infinite) sequence of approximations to the matrix q_{ab} [16, 17, 18, 19], in which the $n \times n$ matrix is successively divided into $n_i \times n_i$ submatrices ($i = 1 \ldots K$):

$$\mathbf{q} = \begin{pmatrix} \mathbf{q}_1 & q_0\mathbb{1} & \ldots & q_0\mathbb{1} \\ q_0\mathbb{1} & \mathbf{q}_1 & \ldots & q_0\mathbb{1} \\ \ldots & \ldots & \ddots & \ldots \\ q_0\mathbb{1} & q_0\mathbb{1} & \ldots & \mathbf{q}_1 \end{pmatrix} \qquad (n \times n)$$

$$\mathbf{q}_1 = \begin{pmatrix} \mathbf{q}_2 & q_1\mathbb{1} & \ldots & q_1\mathbb{1} \\ q_1\mathbb{1} & \mathbf{q}_2 & \ldots & q_1\mathbb{1} \\ \ldots & \ldots & \ddots & \ldots \\ q_1\mathbb{1} & q_1\mathbb{1} & \ldots & \mathbf{q}_2 \end{pmatrix} \qquad (n_1 \times n_1)$$

$$\vdots$$

$$\mathbf{q}_K = \begin{pmatrix} 1 & q_K & \ldots & q_K \\ q_K & 1 & \ldots & q_K \\ \ldots & \ldots & \ddots & \ldots \\ q_K & q_K & \ldots & 1 \end{pmatrix} \qquad (n_K \times n_K). \tag{7.1}$$

Clearly $n > n_1 > n_2 \ldots > n_K > 1$; when we take the limit $n \to 0$ however, the inequalities are reversed so that $0 < n_1 < \ldots < n_{K-1} < n_K < 1$.

If we take the limit $K \to \infty$, and write $q(x) = q_i$ where $n_{i+1} \leq x \leq n_i$, the overlap order parameter q, then becomes an order parameter function $q(x)$ [19], and the free energy becomes the solution of a non-linear partial differential equation [17, 90, 91]. We are however left with the problem of how to interpret the function $q(x)$.

We can define the probability distribution of q_{ab}, $P(q)$ by

$$P(q) = \sum_{a,b} P_a P_b \, \delta(q - q_{ab}),$$

where P_a is the probability of finding a pure state a. A pure state is defined as one in which the values of intensive quantities of the form $\frac{1}{N}\sum_i A_i$ do not fluctuate in the thermodynamic limit. The probability of finding a value of the spin-glass order parameter, less than some q is then

$$x(q) = \int_{-\infty}^{q} dq' P(q').$$

It can be shown [19] that the inverse of this function, $q(x)$, is exactly the same as the function $q(x)$ introduced in the infinite sequence of approximations in the free energy calculation. Hence $q(x)$ has a physical interpretation in terms of some kind of probability law. Unfortunately, the weights

P_a are dependent on the exact realisation of the disorder, and hence $P(q)$ is not self averaging but itself follows some distribution law.

The breaking of replica symmetry is believed to be a mathematical manifestation of the infinite number of pure thermodynamic equilibrium states in the spin-glass phase. This is caused by the large number of free energy valleys [49], separated by barriers which diverge in the thermodynamic limit, but at varying rates. This can lead to an interpretation of x in terms of time scales over which states with mutual overlap $q(x)$ are accessible to each other [92]. All these time scales diverge in the thermodynamic limit, but at differing rates.

The nature of the spin glass phase is further elucidated [51, 93] by considering the mutual overlaps between any three states: q_{12}, q_{23} and q_{13}. It is found that at least two of these will be equal, say $q_{12} = q_{23} = q$, and the third $q_{13} = q'$ has greater or equal value: $q' \geq q$. i.e. The space is *ultrametric* - by grouping together all pure states with overlap greater than some value q we have divided space into disjoint clusters. Each cluster can itself be subdivided by grouping all states with overlap greater than q', where $q' > q$. This structure can be represented diagrammatically as a genealogical tree, with end points representing states, and branches clusters. The distance (i.e. overlap) between states is determined by the number of generations one has to go back to find a common ancestor.

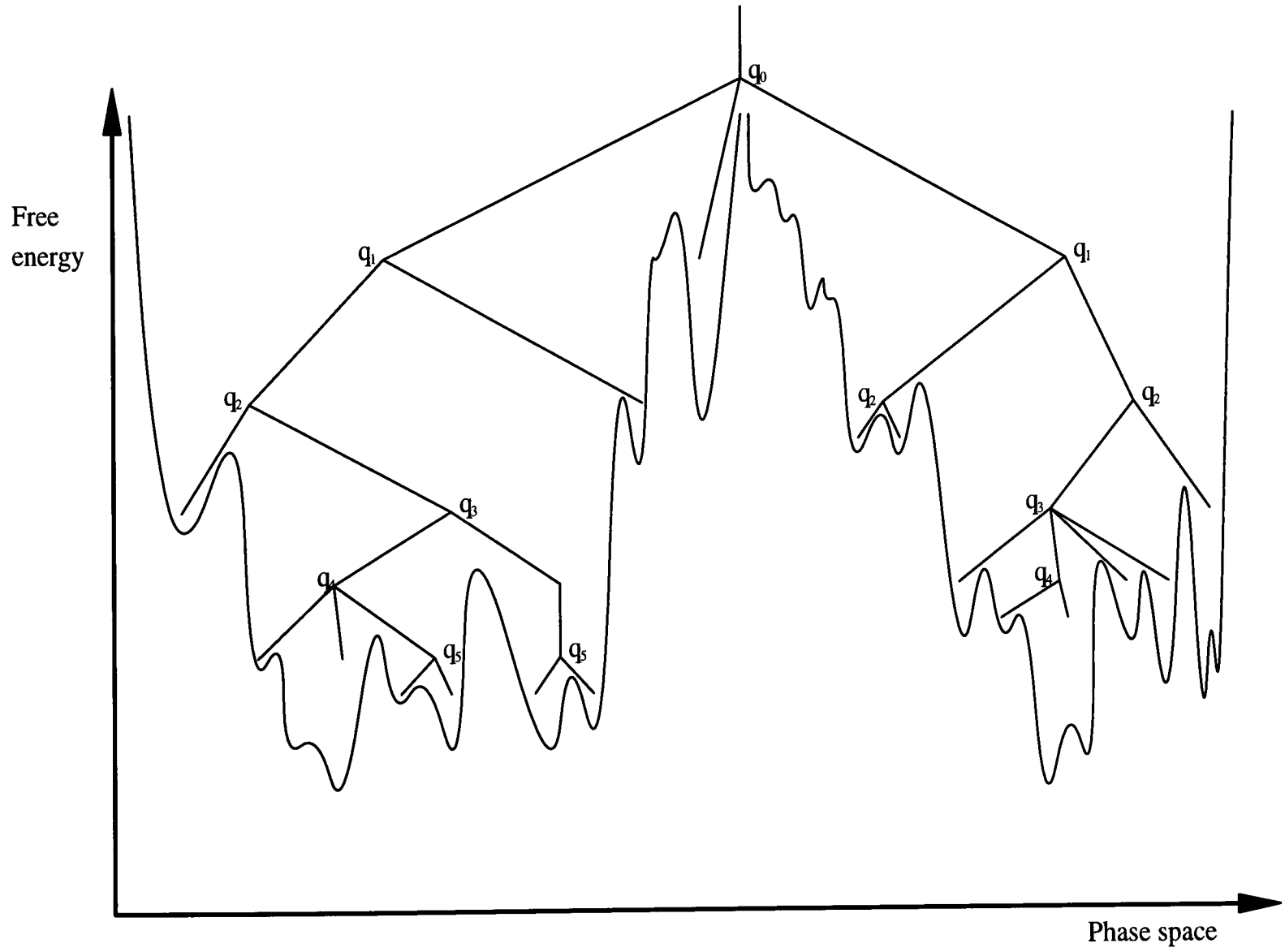


Figure 7.1: The ultrametric structure of phase space in a spin glass

The branches of the tree can be interpreted as a temperature scale. The clusters of low temperature states become states of the system at higher temperature. In the spin glass phase there are an infinite number of states (i.e. end points) for all temperatures; i.e. the tree has an

infinite number of branches.

7.3 Full RSB calculation for the field distribution evolution

In this section we sketch how to derive full RSB equations *à la* Parisi, for the evolution of the local field distribution in the S-K model, described in its replica symmetric form in Chapter 6. Generalisation to the asymmetric model (i.e. with $k \neq 0$) is trivial.

The instability of the RS solution, signified by the determinant of the A-T matrix, who's coefficients are calculated in Appendix C, becoming positive, means that we have to try a more complicated ansatz for the replica matrices. The failure of the RS scheme is due to the breaking up of phase space into many non-equivalent components. Because of the extensive energy barriers, these states are inaccessible to each other on finite time scales. This leads to the so-called ultrametric structure [51]. Parisi's ansatz [16, 17, 18, 19] involves a sequence of approximations for the replica matrices. In the limit of an infinite number of approximations in this sequence, the matrices become functions of the interval $0-1$. Rather than acquiring a definite value, the relevant order parameters follow a probability law. Parisi's solution, involving successive block diagonalisation can be graphically represented as

$$\mathbf{q}-\mathbb{1} = q_0 \Bigg(\overset{n}{\square} \Bigg) + (q_1-q_0) \Bigg(\overset{n_1}{\square\!\square\!\square} \Bigg) + (q_2-q_1) \Bigg(\overset{n_2}{\square\!\square\!\square\!\square\!\square} \Bigg) + \ldots + (q_K-q_{K-1}) \Bigg(\overset{n_K}{\square\!\square\!\square\!\square\!\square\!\square\!\square\!\square\!\square} \Bigg) - q_K \Bigg(\overset{1}{\square} \Bigg) \tag{7.2}$$

where the matrices are only non-zero for those elements within the boxes, when they are unity. For positive integral n, $n > n_1 > \ldots n_K > n_{K+1} = 1$, when we take the limit $n \to 0$, these inequalities will be reversed so that $0 < n_1 < \ldots n_K < 1$.

In our current dynamic formulation, we have three replica matrices q_{ab}, Q_{ab} and $\hat{\mathcal{R}}_{ab}$, which are not trivially related, rather than the one encountered in the original equilibrium calculation. This complicates matters on a technical level, but the underlying principles are the same.

In order to incorporate Parisi's ansatz we return to (6.7), after having made the relevant simplification to get rid of $\hat{q}_{ab}$ and $\mathcal{W}_{ab}$ etc. The logarithm term in the exponent (6.9) is

$$\ln \left\langle \int \left(\prod_a \frac{d\hat{h}^a\, dH^a}{2\pi} \right) e^{-\frac{\tilde{J}^2}{2}\hat{h}^a q_{ab} \hat{h}^b + i\hat{h}^a \left(H^a - \tilde{J}_0 m^a - \hat{\mathcal{R}}_{ab}\sigma^b \right) - i\hat{m}^a\sigma^a - \frac{\tilde{J}^2}{2}\sigma^a Q_{ab}\sigma^b + \hat{\rho}^{\sigma^a}(H^a)} \right\rangle_{\boldsymbol{\sigma}},$$

with an implied sum over consecutive indices. We first remove the diagonal terms of q_{ab}, $\hat{\mathcal{R}}_{ab}$ and Q_{ab}, since clearly $q_{aa} = 1$, and hence $Q_{aa} = 0$. At this stage it is convenient to explicitly remove the diagonal terms of the last block in the ultrametric structure of the matrices. This allows us to carry out a full sum at a later stage, rather than restricting to off-diagonal elements. Therefore $q_{aa} = 1 - q_K$, $\hat{\mathcal{R}}_{aa} = \hat{\mathcal{R}} - \hat{\mathcal{R}}_K$ and $Q_{aa} = -Q_K$. These terms give $e^{-\frac{\tilde{J}^2}{2}(1-q_K)(\hat{h}^a)^2 - i\hat{h}^a(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K)\sigma^a + \frac{\tilde{J}^2}{2}Q_K}$. We next split off the constant parts of the matrices q_n, $\hat{\mathcal{R}}_n$ and Q_n for which we can carry out a Gaussian linearisation giving

$$\log \left\langle \int \left(\prod_a \frac{d\hat{h}^a \, dH^a}{2\pi} \right) e^{-\frac{\tilde{J}^2}{2}\sum \hat{h}^a (q_{ab}-q_n)\hat{h}^b - i\sum \hat{h}^a (\hat{\mathcal{R}}_{ab}-\hat{\mathcal{R}}_n)\sigma^b - \frac{\tilde{J}^2}{2}\sum \sigma^a (Q_{ab}-Q_n)\sigma^b}$$
$$\prod_a e^{-\frac{\tilde{J}^2}{2}(1-q_K-q_n)(\hat{h}^a)^2 - i\hat{h}^a(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma^a + \frac{\tilde{J}^2}{2}(Q_K+Q_n) + \hat{\rho}^{\sigma^a}(H^a)}$$
$$\int D\tilde{w} \int D\tilde{y} \prod_a e^{i\hat{h}^a(H^a - \tilde{J}_0 m^a - ib\tilde{w} + cd\tilde{y}) + (b\tilde{w} - ic\tilde{y} - i\hat{m}^a)\sigma^a} \Bigg\rangle_{\boldsymbol{\sigma}},$$

where b, c, d solve

$$b^2 - c^2 d^2 = -\tilde{J}^2 q_n \qquad b^2 - c^2 d = -i\hat{\mathcal{R}}_n \qquad b^2 - c^2 = -\frac{\tilde{J}^2}{2} Q_n.$$

We take this opportunity to change variables $\mu^a = -i\hat{m}^a$, then following Duplantier [90] (see also [91, 94, 95]) we note that the m^a's and μ^a's can be used as generating fields for the $\hat{h}^a$'s and σ^a's:

$$\begin{aligned} \hat{h}^a &= \frac{i}{\tilde{J}_0}\frac{\partial}{\partial m^a} e^{i\sum_a \hat{h}^a(H^a - \tilde{J}_0 m^a - ib\tilde{w} + cd\tilde{y}) + \sum_a (b\tilde{w} - ic\tilde{y} + \mu^a)\sigma^a} \\ \sigma^a &= \frac{\partial}{\partial \mu^a} e^{i\sum_a \hat{h}^a(H^a - \tilde{J}_0 m^a - ib\tilde{w} + cd\tilde{y}) + \sum_a (b\tilde{w} - ic\tilde{y} + \mu^a)\sigma^a}. \end{aligned}$$

The non-diagonal quadratic terms in the exponent can then be written

$$e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}\sum (q_{ab}-q_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial m^b} + \frac{1}{\tilde{J}_0}\sum (\hat{\mathcal{R}}_{ab}-\hat{\mathcal{R}}_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial \mu^b} - \frac{\tilde{J}^2}{2}\sum (Q_{ab}-Q_n)\frac{\partial}{\partial \mu^a}\frac{\partial}{\partial \mu^b}}.$$

7.3.1 Saddle point term

In order to calculate the logarithm term that appears in the saddle point exponent (6.9) we carry out the integral over $\hat{h}^a$ to obtain

$$\ln \Bigg[e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}\sum (q_{ab}-q_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial m^b} + \frac{1}{\tilde{J}_0}\sum (\hat{\mathcal{R}}_{ab}-\hat{\mathcal{R}}_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial \mu^b} + \frac{\tilde{J}^2}{2}\sum (Q_{ab}-Q_n)\frac{\partial}{\partial \mu^a}\frac{\partial}{\partial \mu^b}}$$
$$\int D\tilde{w} \int D\tilde{y} \prod_{a=1}^{n} \left\langle \int dH^a \, e^{-\frac{\left(H^a - \tilde{J}_0 m^a - ib\tilde{w} + cd\tilde{y} - (\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma^a\right)^2}{2\tilde{J}^2(1-q_K-q_n)} + (b\tilde{w} - ic\tilde{y} + \mu^a)\sigma^a + \hat{\rho}^{\sigma^a}(H^a)} \right\rangle_{\sigma^a} \Bigg]$$
$$+ n\frac{\tilde{J}^2}{2}(Q_K + Q_n) - \frac{n}{2}\ln\left[2\pi\tilde{J}^2(1 - q_K - q_n)\right]$$
$$= \log \int D\tilde{w} \int D\tilde{y} \; e^{\log[G(n,\mathbf{m},\hat{\mathbf{m}})]} + n\frac{\tilde{J}^2}{2}(Q_K + Q_n) - \frac{n}{2}\ln\left[2\pi\tilde{J}^2(1 - q_K - q_n)\right].$$

Writing $\sum_a \frac{\partial g(\{m^a\})}{\partial m^a} = \frac{dg(\{m\})}{dm}$, we have a series of equations:

$$g(n_{K+1}, m, \mu) = \left\langle \int dH \, e^{-\frac{\left(H - \tilde{J}_0 m - ib\tilde{w} + cd\tilde{y} - (\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)} + (b\tilde{w} - ic\tilde{y} + \mu)\sigma + \hat{\rho}^{\sigma}(H)} \right\rangle_{\sigma}$$

$$
\begin{aligned}
g(n_K,m,\mu) &= e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1})\frac{\partial^2}{\partial m\partial\mu}+\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}} \\
&\qquad [g(n_{K+1},m,\mu)]^{\frac{n_K}{n_{K+1}}} \\
g(n_{K-1},m,\mu) &= e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_{K-1}-q_{K-2})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}(\hat{\mathcal{R}}_{K-1}-\hat{\mathcal{R}}_{K-2})\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_{K-1}-Q_{K-2})\frac{\partial^2}{\partial\mu^2}} \\
&\qquad [g(n_K,m,\mu)]^{\frac{n_{K-1}}{n_K}} \\
&\vdots \\
G &= g(n,m,mu) = [g(n_1,m,\mu)]^{\frac{n}{n_1}} .
\end{aligned}
$$

If we write $n_i \leq x \leq n_{i-1}$ then $\frac{n_{i-1}}{n_i} = \frac{x+dx}{x}$, and we have the recurrence relation

$$
g(x+dx,m,\mu) = e^{-\frac{\tilde{J}^2}{2\tilde{J}_0^2}dq(x)\frac{\partial^2}{\partial m^2}-\frac{1}{\tilde{J}_0}d\hat{\mathcal{R}}(x)\frac{\partial^2}{\partial m\partial\mu}+\frac{\tilde{J}^2}{2}dQ(x)\frac{\partial^2}{\partial\mu^2}}\; g^{1+\frac{dx}{x}}(x,m,\mu),
$$

with

$$
g(1,m,\mu) = \left\langle \int dH\, e^{-\frac{(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma .
$$

In the limit $K\to\infty$, $dx\to 0$, so we can use $\lim_{\frac{dx}{x}\to 0}\frac{Z^{\frac{dx}{x}}}{\frac{dx}{x}} = \log Z + \frac{1}{\frac{dx}{x}}$; therefore

$$
\frac{\partial g}{\partial x} = -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g}{\partial m^2} - \frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g}{\partial m\partial\mu} + \frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g}{\partial\mu^2} + \frac{g}{x}\log g, \tag{7.3}
$$

which upon writing $f(x,m,\mu) = \frac{\log g(x,m,\mu)}{x}$ gives the non-linear partial differential equation

$$
\begin{aligned}
\frac{\partial}{\partial x}f(x,m,\mu) &= -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\left[\frac{\partial^2 f}{\partial m^2}+x\left(\frac{\partial f}{\partial m}\right)^2\right] - \frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\left[\frac{\partial^2 f}{\partial m\partial\mu}+x\frac{\partial f}{\partial m}\frac{\partial f}{\partial\mu}\right] \\
&\quad +\frac{\tilde{J}^2}{2}\frac{dQ}{dx}\left[\frac{\partial^2 f}{\partial\mu^2}+x\left(\frac{\partial f}{\partial\mu}\right)^2\right].
\end{aligned} \tag{7.4}
$$

Hence, in the limit $n\to 0$ the logarithm term in the saddle point exponent (6.9) is

$$
\begin{aligned}
\lim_{n\to 0}\frac{1}{n}\log\int D\tilde{w}\int D\tilde{y}\; e^{\log[G]} &= \lim_{n\to 0}\frac{1}{n}\log\int D\tilde{w}\int D\tilde{y}\; e^{\log[g(n,m,\mu)]} \\
&= \lim_{n\to 0}\frac{1}{n}\log\int D\tilde{w}\int D\tilde{y}\; e^{nf(n,m,\mu)} \\
&= \int D\tilde{w}\int D\tilde{y}\; f(0,m,\mu),
\end{aligned}
$$

where $f(x,m,\mu)$ solves (7.4) with boundary conditions

$$
f(1,m,\mu) = \log\left\langle \int dH\, e^{-\frac{(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}(1)-\hat{\mathcal{R}}(0))\sigma)^2}{2\tilde{J}^2(1-q(1)-q(0))}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma .
$$

Following Parisi's scheme object like $\frac{1}{n}\sum_{a,b\neq a} q_{ab}Q_{ab} \to -\int dx\, q(x)Q(x)$, whilst objects with only one replica label remain replica symmetric. Hence our saddle point exponent (6.9) evaluated within Parisi's ansatz is

$$
\begin{aligned}
\frac{\Psi}{n} &= -\mu m - \frac{\tilde{J}^2}{2}\int dx\, q(x)Q(x) + \frac{1}{2\tilde{J}^2}\int dx\,\hat{\mathcal{R}}^2(x) - \frac{1}{2\tilde{J}^2}\hat{\mathcal{R}}^2 - \frac{1}{2}\log[1-q(1)-q(0)] \\
&\quad +\frac{\tilde{J}^2}{2}(Q(1)+Q(0)) - \sum_\sigma\int dh\,\hat{\rho}^\sigma(h)\rho_t^\sigma(h) + \int D\tilde{w}\int D\tilde{y}\, f(0,m,\mu).
\end{aligned} \tag{7.5}
$$

The saddle point equations are now given by the functional differentials $\frac{\delta\Psi}{\delta q(x)}$ etc, see [91, 94, 96], e.g. the equation for $\rho_t^{\varsigma}(h)$ is given by

$$\rho_t^{\varsigma}(h) = \int D\tilde{w} \int D\tilde{y}\, \frac{\delta f(0,m,\mu)}{\delta\hat{\rho}^{\varsigma}(h)},$$

where $\frac{\delta f(0,m,\mu)}{\delta\hat{\rho}^{\varsigma}(h)}$ solves (7.4) with boundary conditions

$$\frac{\delta f(1,m,\mu)}{\delta\hat{\rho}^{\varsigma}(h)} = \frac{e^{-\frac{\left(h-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}(1)-\hat{\mathcal{R}}(0))\varsigma\right)^2}{2\tilde{J}^2(1-q(1)-q(0))}+(b\tilde{w}-ic\tilde{y}+\mu)\varsigma+\hat{\rho}^{\varsigma}(h)}}{\sum_{\sigma}\int dH\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}(1)-\hat{\mathcal{R}}(0))\sigma\right)^2}{2\tilde{J}^2(1-q(1)-q(0))}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^{\sigma}(H)}}.$$

7.3.2 Intensive terms

For the non-trivial intensive terms in (6.10), giving the coefficients in the evolution equation (6.4), we have to evaluate objects like

$$-i\tilde{J}^2 \int \left(\prod_a d\hat{h}^a_{1,2}\, dH^a_{1,2}\right) \tanh(\beta H^1_1)\, \delta(h^{\mu} - H^1_2) \left\langle \delta_{\varsigma,\sigma^1_2} \sum_{a'} \left[\hat{h}^{a'}_2 \sigma^{a'}_1 + \hat{h}^{a'}_1 \sigma^{a'}_2\right] e^{\Phi_1+\Phi_2} \right\rangle_{\boldsymbol{\sigma}}$$

which conveniently factorises into parts dependent only on site 1, and parts dependent only on site 2.

The procedure is essentially the same as for the saddle point equations, only now we have to differentiate between the cases $a' = 1$ or $a' \neq 1$; $k = 1$ or $k = 2$; and, $\hat{h}^{a'}$ appearing in the integral or $\sigma^{a'}$ in the integral. As before we can use $\frac{\partial}{\partial m^{a'}}$ and $\frac{\partial}{\partial \mu^{a'}}$ to generate the $\hat{h}^{a'}$'s and $\sigma^{a'}$'s in the integral. Carrying out the integral over $\hat{h}^a_k$ therefore gives

$$e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}\sum_{a,b}(q_{ab}-q_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial m^b}+\frac{1}{\tilde{J}_0}\sum_{a,b}(\hat{\mathcal{R}}_{ab}-\hat{\mathcal{R}}_n)\frac{\partial}{\partial m^a}\frac{\partial}{\partial \mu^b}-\frac{\tilde{J}^2}{2}\sum_{a,b}(Q_{ab}-Q_n)\frac{\partial}{\partial \mu^a}\frac{\partial}{\partial \mu^b}}$$

$$\sum_{a'}\int D\tilde{w}\int D\tilde{y}\left\langle \int\left(\prod_{a=1}^{n} dH^a\right) f_{1,2}(\sigma^1,H^1)\left(-i\tilde{J}^2\right)\left\{\begin{array}{c} \frac{i}{\tilde{J}_0}\frac{\partial}{\partial m^{a'}} \\ \\ \frac{\partial}{\partial m u^{a'}} \end{array}\right\}\right.$$

$$\left.\prod_{a=1}^{n} e^{-\frac{\left(H^a-\tilde{J}_0 m^a - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}_{aa}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma^a\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu^a)\sigma^a+\hat{\rho}^{\sigma^a}(H^a)}\right\rangle, \qquad (7.6)$$

where $f_k(\sigma^1,H^1) = \tanh(\beta H^1)$ for $k=1$, and $\delta_{\varsigma,\sigma^1}\,\delta(h-H^1)$ for $k=2$.

Proceeding to apply Parisi's ansatz as before, we derive very similar partial differential equation as for the saddle point term (7.3). Differences are caused by the 'singled out' replicas 1 and a'. When $a'=1$ there is only one singled out replica label, and when $a'\neq 1$ there are two.

For $a'=1$ we have

$$g^0(n_{K+1},m,\mu) = \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^{\sigma}(H)}\right\rangle_{\sigma}$$

$$g^1_k(n_{K+1},m,\mu) = \left\{\begin{array}{c} \frac{\partial}{\partial m} \\ \frac{\partial}{\partial \mu} \end{array}\right\}\left\langle \int dH\, f_k(\sigma,H)\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^{\sigma}(H)}\right\rangle_{\sigma}$$

$$g^0(n_K,m,\mu) = e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1})\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}}$$

$$
\begin{aligned}
& & & \Big[g^0(n_{K+1},m,\mu)\Big]^{\frac{n_K}{n_{K+1}}} \\
g^1(n_K,m,\mu) &=& e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}\left(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1}\right)\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}} \\
& & & \Big[g^0(n_{K+1},m,\mu)\Big]^{\frac{n_K}{n_{K+1}}-1} g^1(n_{K+1},m,\mu) \\
& & \vdots & \\
g^0(n,m,\mu) &=& \Big[g^0(n_1,m,\mu)\Big]^{\frac{n}{n_1}} \\
g^1(n,m,\mu) &=& \Big[g^0(n_1,m,\mu)\Big]^{\frac{n}{n_1}-1} g^1(n_1,m,\mu).
\end{aligned}
$$

Hence we have two differential equations

$$
\begin{aligned}
\frac{\partial g^0}{\partial x} &=& -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g^0}{\partial m^2}-\frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g^0}{\partial m\partial\mu}+\frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g^0}{\partial\mu^2}+\frac{g^0}{x}\log g^0 \\
\frac{\partial g^1}{\partial x} &=& -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g^1}{\partial m^2}-\frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g^1}{\partial m\partial\mu}+\frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g^1}{\partial\mu^2}+\frac{g^1}{x}\log g^0,
\end{aligned}
\tag{7.7}
$$

with boundary conditions

$$
\begin{aligned}
g^0(1,m,\mu) &=& \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g_k^1(1,m,\mu) &=& \\
& & \left\{ \begin{array}{c} \frac{\partial}{\partial m} \\ \frac{\partial}{\partial\mu} \end{array} \right\} \left\langle \int dH\, f_k(\sigma,H)\, e^{-\frac{\left(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma .
\end{aligned}
$$

For the case $a' \neq 1$ we have

$$
\begin{aligned}
g^0(n_{K+1},m,\mu) &=& \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g_k^1(n_{K+1},m,\mu) &=& \\
& & \left\langle \int dH\, f_k(\sigma,H)\, e^{-\frac{\left(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g_k^2(n_{K+1},m,\mu) &=& \\
& & \left\{ \begin{array}{c} \frac{\partial}{\partial m} \\ \frac{\partial}{\partial\mu} \end{array} \right\} \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m-ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma^a\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g^0(n_K,m,\mu) &=& e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}\left(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1}\right)\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}} \\
& & \Big[g^0(n_{K+1},m,\mu)\Big]^{\frac{n_K}{n_{K+1}}} \\
g^1(n_K,m,\mu) &=& e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}\left(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1}\right)\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}} \\
& & \Big[g^0(n_{K+1},m,\mu)\Big]^{\frac{n_K}{n_{K+1}}-1} g^1(n_{K+1},m,\mu) \\
g^2(n_K,m,\mu) &=& e^{\frac{\tilde{J}^2}{2\tilde{J}_0^2}(q_K-q_{K-1})\frac{\partial^2}{\partial m^2}+\frac{1}{\tilde{J}_0}\left(\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_{K-1}\right)\frac{\partial^2}{\partial m\partial\mu}-\frac{\tilde{J}^2}{2}(Q_K-Q_{K-1})\frac{\partial^2}{\partial\mu^2}} \\
& & \Big[g^0(n_{K+1},m,\mu)\Big]^{\frac{n_K}{n_{K+1}}-2} g^1(n_{K+1},m,\mu) g^2(n_{K+1},m,\mu)
\end{aligned}
$$

$$
\begin{aligned}
&\vdots \\
g^0(n,m,\mu) &= \left[g^0(n_1,m,\mu)\right]^{\frac{n}{n_1}} \\
g^1(n,m,\mu) &= \left[g^0(n_1,m,\mu)\right]^{\frac{n}{n_1}-1} g^1(n_1,m,\mu) \\
g^2(n,m,\mu) &= \left[g^0(n_1,m,\mu)\right]^{\frac{n}{n_1}-2} g^1(n_1,m,\mu) g^2(n_1,m,\mu).
\end{aligned}
$$

Hence we have three differential equations

$$
\begin{aligned}
\frac{\partial g^0}{\partial x} &= -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g^0}{\partial m^2} - \frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g^0}{\partial m \partial \mu} + \frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g^0}{\partial \mu^2} + \frac{g^0}{x}\log g^0 \\
\frac{\partial g^1}{\partial x} &= -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g^1}{\partial m^2} - \frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g^1}{\partial m \partial \mu} + \frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g^1}{\partial \mu^2} + \frac{g^1}{x}\log g^0 \\
\frac{\partial g^2}{\partial x} &= -\frac{\tilde{J}^2}{2\tilde{J}_0^2}\frac{dq}{dx}\frac{\partial^2 g^2}{\partial m^2} - \frac{1}{\tilde{J}_0}\frac{d\hat{\mathcal{R}}}{dx}\frac{\partial^2 g^2}{\partial m \partial \mu} + \frac{\tilde{J}^2}{2}\frac{dQ}{dx}\frac{\partial^2 g^2}{\partial \mu^2} + \frac{g^1 g^2}{x g^0}\log g^0,
\end{aligned}
\tag{7.8}
$$

with boundary conditions

$$
\begin{aligned}
g^0(1,m,\mu) &= \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g_k^1(1,m,\mu) &= \\
&\left\langle \int dH\, f_k(\sigma,H)\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma \\
g_k^2(1,m,\mu) &= \\
&\left\{ \begin{array}{c} \frac{\partial}{\partial m} \\ \frac{\partial}{\partial \mu} \end{array} \right\} \left\langle \int dH\, e^{-\frac{\left(H-\tilde{J}_0 m - ib\tilde{w}+cd\tilde{y}-(\hat{\mathcal{R}}-\hat{\mathcal{R}}_K-\hat{\mathcal{R}}_n)\sigma\right)^2}{2\tilde{J}^2(1-q_K-q_n)}+(b\tilde{w}-ic\tilde{y}+\mu)\sigma+\hat{\rho}^\sigma(H)} \right\rangle_\sigma .
\end{aligned}
$$

The coefficient are then given by

$$
\begin{aligned}
&\frac{\tilde{J}^2}{\tilde{J}_0}\left[\int D\tilde{w}\int D\tilde{y}\; g^1_{m_1}(0,m,\mu)\int D\tilde{w}\int D\tilde{y}\; g^1_{\mu_2}(0,m,\mu)\right. \\
&\qquad \left. - \int D\tilde{w}\int D\tilde{y}\; g^2_{m_1}(0,m,\mu)\int D\tilde{w}\int D\tilde{y}\; g^2_{\mu_2}(0,m,\mu)\right] \\
&+\frac{\tilde{J}^2}{\tilde{J}_0}\left[\int D\tilde{w}\int D\tilde{y}\; g^1_{\mu_1}(0,m,\mu)\int D\tilde{w}\int D\tilde{y}\; g^1_{m_2}(0,m,\mu)\right. \\
&\qquad \left. - \int D\tilde{w}\int D\tilde{y}\; g^2_{\mu_1}(0,m,\mu)\int D\tilde{w}\int D\tilde{y}\; g^2_{m_2}(0,m,\mu)\right],
\end{aligned}
\tag{7.9}
$$

where g_{m_1} indicates the term containing the $\frac{\partial}{\partial m}$ and $f(\sigma_1^1, H_1^1)$, and g_{μ_2} indicates the term containing $\frac{\partial}{\partial \mu}$ and $f(\sigma_2^1, H_2^1)$.

At the present time, we regard the full replica symmetry broken calculation as far too difficult to solve computationally, requiring three extra order parameter functions.

7.3.3 Detailed balance equilibrium

Upon applying our ansatz for the equilibrium form of $\hat{\rho}^{\sigma^a}(H^a) = \frac{\beta}{2}\sigma^a H^a$, and using the relations (6.12), the differential equation (7.4) reduces exactly to that of Parisi's original equilibrium calculation [17].

The coefficients (7.9) with this specific choice of $\hat{\rho}^{\sigma^a}(H^a)$ should also give fixed points of the flow, when used in (6.4). Since they are derived from (6.10,) which we proved give fixed points of the flow upon applying the equilibrium form of $\hat{\rho}^{\sigma^a}(H^a)$ in Subsection 6.3.1, the coefficients (7.9) must indeed give fixed points of the flow when evaluated using Parisi's ansatz.

7.4 Conclusion

Following Parisi's scheme for the ultrametric structure of the replica matrices, the saddle point equations and coefficients in the equations of motion for the joint spin-field distribution of the generalised S-K model can be evaluated. The full RSB hierarchy leads to a set of non-linear partial differential equations which in general cannot be solved analytically.

Because there are three replica matrices rather than the normal one or two, the position of the A-T line, marking the instability of the RS solution, is difficult to evaluate. For the same reason solving the full RSB equations computationally would require formidable effort. It would however be interesting to compare the solution of the full RSB equations with Monte-Carlo simulations in the spin-glass region. This could provide a method for understanding whether the complex temporal behaviours of spin-glasses such, as aging, are purely due to RSB effects or whether they can be captured in a replica symmetric theory.

Chapter 8

Conclusion

In this thesis I have investigated the dynamics of a number of disordered spin systems (i.e. where the interactions are not uniform but drawn at random from a predetermined distribution), of successively greater complexity. Since we require results of the study of disordered systems to be independent of the exact realisation of the disorder, we need to identify quantities which are self averaging (i.e. peaked around a typical value for almost all realisations) allowing us to average over the disorder. Carrying out this average often involves the replica trick, which, though solving one problem, opens up a whole series of questions of interpretation.

Disordered spin systems are often used as models of neural networks, and of materials with magnetic impurities randomly distributed in a substrate. The motivation for a study of the dynamics in such systems is twofold. Firstly, many interesting dynamic phenomena such as aging, and long relaxation times are observed in such systems which cannot be captured by an equilibrium study, where we have already assumed the long time limit. Secondly, biological reality forces us to try and abandon the symmetry of interactions in neural networks. This is equivalent to lack of detailed balance, and hence a dynamical study is the only way to obtain results, since conventional equilibrium techniques require detailed balance.

I have attempted to present a consistent approach to the different models studied. Although the motivation and interpretation of the models may be very different, the approaches used are often similar. Often the only difference between the models is the choice of interactions, and hence the form of the disorder average.

The dynamics of the simplest model - the generalised Hopfield model trained with an intensive number of patterns, can be described by the patterns overlaps which, if the interactions are symmetric, evolve to an equilibrium state - signified by the existence of Lyapunov functions. For asymmetric interactions, complex trajectories exist even in this simplest model.

For more complex models, with an extensive number of disordered variables per site (i.e. the generalised Hopfield model trained with an extensive number of patterns or the generalised SK model), the overlaps (or magnetisation) are insufficient to describe the system state - a disordered contribution to the energy must be included at the very least. The question of how to derive closed sets of equations in these systems now arises, since they do not automatically close as they do in the simple models with intensive amounts of disorder per site.

In this thesis I have further developed a method for closing the equations based on the removal of microscopic effects by assuming equipartitioning of probability within the macroscopic sub-

shells. The dynamic equations will close if we choose the order parameters such equipartitioning holds. I have presented evidence that a simple two parameter model is not sufficient to describe the dynamics accurately, despite a good qualitative fit to simulations. Introducing an order parameter function - the joint distribution of spins and fields, provides much better results on intermediate time scales, as well as reproducing the well known equilibrium results in detailed balance equilibrium (as indeed the two parameter theory does too). The coefficients in the diffusion equation for the joint distribution, can be evaluated within the RS ansatz, or by using Parisi's ultrametric structure for the replica matrices.

Despite the quality of the fit on intermediate time scales, it is far from obvious whether this theory can predict, for instance, the dynamic exponents in the decay towards equilibrium, or the aging phenomena displayed by real glasses. This method does however provide a framework for studying such phenomena, and related interesting questions such as what role replica symmetry breaking plays in the long time dynamics.

In order to validate the theory involving an order parameter function, it would be interesting to apply the methods presented here to an exactly solvable model. This could enlighten us to the inaccuracies and approximations (if any) encountered in this approach.

The main result of this thesis (I hope) has been to show that the dynamics of complex systems can be approached using many methods originally employed to study their equilibrium states. Dynamic replica theories exist, with the saddle points defining the coefficients in flow equations for macroscopic order parameters. The choice of order parameters is of paramount importance if an accurate description is to be obtained. I have shown, that guided by physical insight, more complex choices lead to a more accurate description, culminating in the infinite number afforded by an order parameter function, which I hope leads towards exact results. Only then can I justify the horrendous waste of entropy.

> "Beware of the man who works hard to learn something, learns it, and finds himself no wiser than before," Bokonon tells us. "He is full of murderous resentment of people who are ignorant without having come by their ignorance the hard way."
>
> -*Kurt Vonnegut*, Cat's Cradle

Appendices

Appendix A

Bifurcation Theory

A.1 Introduction

Bifurcation theory is concerned with the qualitative description of dynamic equations that cannot be solved analytically, typically because they are non-linear, and/or because the system under consideration has a large number of degrees of freedom. Bifurcation theory is typically employed when the complexity of equations makes it impossible to derive exact solutions. The aim of bifurcation theory is not to derive exact, quantitative results, but to identify qualitative changes in the behaviour of the system, i.e. when and how new classes of solution emerge. The non-transient behaviour of a non-linear system may change from a stable fixed point to a complicated dynamics, as external control parameters are varied - we wish to be able to predict at which values of the control parameters this occurs. We consider here bifurcation from a stable fixed point, since this is the simplest case. Generalisation to bifurcations from more complex states is technically but not conceptually more difficult. In this appendix we outline the methods used in deriving the results in Chapter 3. For a more thorough review of methods see [73] or [74].

The basic tenet of bifurcation theory is that if a fixed point of a system loses its stability as a result of variation of a control parameter then the change in the behaviour of the full non-linear system is reflected in a change in behaviour of the system linearised about the fixed point. Close to the critical point all bifurcations can be classified into one of a small number of sets with well documented properties. The behaviour of all systems within each set is qualitatively the same, only the precise values of numerical constants are dependent on the actual system.

A local bifurcation theory analysis consists of three stages of calculation:

1. **Linearisation** about the fixed point determines which modes will lose stability creating new solutions as parameters are varied.

2. **Centre Manifold Reduction** reduces the full set of equations to those modes confined to the centre manifold - i.e. those with marginal stability. It is when these modes lose stability that new branches of solutions are created.

3. **Normal Form Reduction** systematically removes unnecessary non-linear terms in the equations, until only the essential terms describing the dynamics remain, leaving the *normal form*, with which we can classify the bifurcation.

We consider first-order differential equations, and first-order difference equations, without loss of generality, noting that we can create first order equations from equations of higher order by increasing the number of degrees of freedom.

$$\frac{d\mathbf{x}}{dt} = \mathbf{V}(\mu, \mathbf{x}) \qquad\qquad \mathbf{x}_{t+1} = \mathbf{F}(\mu, \mathbf{x}_t). \tag{A.1}$$

Here μ is a control parameter, and as it is varied the qualitative behaviour of the system changes. We consider systems whose non-transient behaviour is a fixed point (i.e. equilibrium) for some range of μ. We shift the co-ordinates of μ and $\mathbf{x}$ so that $\mu = 0$ is the value of the control parameter at which $\mathbf{x}$ deviates from its equilibrium value $\mathbf{x} = 0$. This is called the critical point. Local Bifurcation theory is concerned with the question of what happens near $\mathbf{x} = 0$ when μ is varied about $\mu = 0$. This will be addressed by considering the stability of solutions close to the critical point.

A.2 Linearisation

Close to $\mathbf{x} = 0$ we can create a Taylor series expansion for the dynamical equations (A.1)

$$\dot{\mathbf{x}} = \mathbf{V}(\mu, 0) + \partial_{\mathbf{x}}\mathbf{V}(\mu, 0) \cdot \mathbf{x} + \mathcal{O}(x^2) \qquad\qquad \mathbf{x}_{t+1} = \mathbf{F}(\mu, 0) + \partial_{\mathbf{x}}\mathbf{F}(\mu, 0) \cdot \mathbf{x}_t + \mathcal{O}(x_t^2), \tag{A.2}$$

where $\partial_{\mathbf{x}}\mathbf{V}$ indicates the matrix $\frac{\partial V^\mu}{\partial x^\nu}$. At the critical point, $\mu = 0$ the first, constant, term vanishes, and close to the equilibrium configuration $\mathbf{x} = 0$ we have a linear equation of motion. We now carry out a linear change in co-ordinates $\tilde{\mathbf{x}} = \mathbf{U}^\dagger\mathbf{x}$, where $\mathbf{U}^\dagger\partial_{\mathbf{x}}\mathbf{V}(0,0)\mathbf{U}$ ($\mathbf{U}^\dagger\partial_{\mathbf{x}}\mathbf{F}(0,0)\mathbf{U}$) is a diagonal matrix $\mathbf{\Lambda}$ ($\tilde{\mathbf{\Lambda}}$) with eigenvalues λ_i ($\tilde{\lambda}_i$). So that

$$\tilde{\mathbf{x}}(t) = e^{\mathbf{\Lambda}t} \cdot \tilde{\mathbf{x}}(0) \qquad\qquad \tilde{\mathbf{x}}_t = \tilde{\mathbf{\Lambda}}^t \cdot \tilde{\mathbf{x}}_0$$

We divide the eigenspace of $\mathbf{\Lambda}$ ($\tilde{\mathbf{\Lambda}}$), E_Λ, into three invariant subspaces according to the stability of each mode, characterised by the eigenvalues λ ($\tilde{\lambda}$)

$$\begin{aligned} E^s: &\quad \Re(\lambda) < 0 & \qquad \tilde{E}^s: &\quad |\tilde{\lambda}| < 1 \\ E^c: &\quad \Re(\lambda) = 0 & \qquad \tilde{E}^c: &\quad |\tilde{\lambda}| = 1 \\ E^u: &\quad \Re(\lambda) > 0 & \qquad \tilde{E}^u: &\quad |\tilde{\lambda}| > 1. \end{aligned}$$

E^s ($\tilde{E}^s$) is the subspace of stable trajectories, E^c ($\tilde{E}^c$) is the subspace of neutral trajectories, and E^u ($\tilde{E}^u$) is the subspace of unstable trajectories. Information about the linearised equations of motion, around equilibrium at the critical point is only relevant to the full non-linear problem when the motion is asymptotically stable. If this is not the case the full non-linear solutions will be driven away from the range of validity of the linearised equations. However, general theorems exist relating the qualitative features of the full non-linear solution to the behaviour of the linearised equations [75].

THEOREM 3 Hartman - Grobman *If $x = 0$ is a hyperbolic fixed point of (A.1), for some value of μ, then there exists a homeomorphism, valid in some neighbourhood of $x = 0$, mapping the solution of the full non-linear equation (A.1) to the linearised solution (A.2).*

i.e. Close to a hyperbolic fixed point (one with an empty centre sub-space) behaviour of the full non-linear equations and their linearised truncations are qualitatively the same.

A.3 Centre Manifold Reduction

Local bifurcation theory is clearly concerned with the centre sub-spaces E^c ($\tilde{E}^c$). Since the modes in the centre sub-space have neutral stability, we can define a centre manifold to which these neutral modes are confined. This reduces the dimensionality of the equations to that of the centre sub-space, i.e. to the number of critical modes. In general the centre manifold cannot be uniquely defined, even if it could, the non-linearity of the equations would make it impossible to find an analytic expression. Close to the bifurcation point however, it can be determined as a map, which in general must be evaluated as a series solution.

Consider the full non-linear equations at criticality. Without loss of generality we split the co-ordinates $\mathbf{x}$ into $\mathbf{y} \in E^c$ and $\mathbf{z} \in E^s \otimes E^u$, and separate the linear from the non-linear terms (to be denoted by $N_{1,2}$)

$$\begin{aligned} \frac{d\mathbf{y}}{dt} &= \mathbf{A}\mathbf{y} + \mathbf{N}_1(\mathbf{y},\mathbf{z}) & \mathbf{y}_{t+1} &= \mathbf{A}\mathbf{y}_t + \mathbf{N}_1(\mathbf{y}_t,\mathbf{z}_t) \\ \frac{d\mathbf{z}}{dt} &= \mathbf{B}\mathbf{z} + \mathbf{N}_2(\mathbf{y},\mathbf{z}) & \mathbf{z}_{t+1} &= \mathbf{B}\mathbf{z}_t + \mathbf{N}_2(\mathbf{y}_t,\mathbf{z}_t). \end{aligned} \tag{A.3}$$

Sufficiently close to equilibrium $\mathbf{x} = 0$, the centre manifold can be described by a function $\mathbf{h}[\mathbf{y}] = \mathbf{z}$. Since $\mathbf{x} = 0$ is on the centre manifold and tangent to E^c, $\mathbf{h}(0) = 0$, and $\partial_{y^\nu}\mathbf{h}(0) = 0\ \forall\ \nu$. We can derive an equation of motion for $\mathbf{x}^c(t) = (\mathbf{y}^c(t), \mathbf{z}^c(t))$, a point on the centre manifold, for small $\mathbf{y}^c(t)$

$$\begin{aligned} \frac{d\mathbf{y}^c}{dt} &= \mathbf{A}\mathbf{y}^c + \mathbf{N}_1(\mathbf{y}^c, \mathbf{h}[\mathbf{y}^c]) & \mathbf{y}^c_{t+1} &= \mathbf{A}\mathbf{y}^c_t + \mathbf{N}_1(\mathbf{y}^c_t, \mathbf{h}[\mathbf{y}^c_t]) \\ \frac{d\mathbf{z}^c}{dt} &= \partial_{\mathbf{y}}\mathbf{h}[\mathbf{y}^c]\frac{d\mathbf{y}^c}{dt} & \mathbf{z}^c_{t+1} &= \mathbf{h}[\mathbf{y}^c_{t+1}] \\ &= \partial_{\mathbf{y}}\mathbf{h}[\mathbf{y}^c]\left(\mathbf{A}\mathbf{y}^c + \mathbf{N}_1(\mathbf{y}^c, \mathbf{h}[\mathbf{y}^c])\right) & &= \mathbf{h}\left[\mathbf{A}\mathbf{y}^c_t + \mathbf{N}_1(\mathbf{y}^c_t, \mathbf{h}(\mathbf{y}^c_t))\right]. \end{aligned} \tag{A.4}$$

Combining with (A.3) we see that this equals

$$\mathbf{B}\mathbf{h}[\mathbf{y}^c] + \mathbf{N}_2(\mathbf{y}^c, \mathbf{h}[\mathbf{y}^c]) \qquad \mathbf{B}\mathbf{h}[\mathbf{y}^c_t] + \mathbf{N}_2(\mathbf{y}^c_t, \mathbf{h}[\mathbf{y}^c_t]) \tag{A.5}$$

giving an equation determining the center manifold near $\mathbf{x} = 0$. Having determined $\mathbf{h}(\mathbf{y}^c)$ we have decoupled the full dynamics to the dynamics on the centre manifold only, from which the normal forms can be derived, allowing us to determine the qualitative change in behaviour. In general $\mathbf{h}(\mathbf{y}^c)$ is determined as a power series to an arbitrary order.

Having achieved the centre manifold reduction we can extend the Hartman-Grobman theory to include non-hyperbolic equilibrium [97, 98]:

THEOREM 4 Shoshitaishvili *There exists a homeomorphism, valid in some neighbourhood of* $x = 0$, *mapping the solution of the full non-linear equation (A.1) to the decoupled, partially linearised solution (A.3), with* $\mathbf{N}_2 = 0$.

A.4 Poincare-Birkhoff Normal Forms

Our next task is to remove any unnecessary non-linear terms to leave the Normal form, which contains only the essential terms necessary to determine the dynamics close to the critical point.

Suppose that centre manifold reduction yields equations of the form

$$\frac{dx'}{dt} = V(\mu, x') = V^{(1)}(\mu, x') + V^{(2)}(\mu, x') + \ldots \qquad x'_{t+1} = F(\mu, x') = F^{(1)}(\mu, x'_t) + F^{(2)}(\mu, x'_t) + \ldots$$

where $V^{(k)}$, $F^{(k)}$ contain terms of order x'^k. Our goal is to apply non-linear co-ordinate changes $x' \to x$ to remove as many non-linear terms as possible. This is achieved by first attempting to remove $V^{(2)}$ then $V^{(3)}$ and so on. Consider the co-ordinate change $x = \Phi(x') = x' + \phi^{(k)}(x')$ with inverse $x' = \Phi^{-1}(x) = x - \phi^{(k)}(x) + \mathcal{O}(x^{2k-1})$, where $\phi^{(k)}$ satisfies $\phi^{(k)}(ax) = a^k \phi^{(k)}(x)$. Our aim is to choose $\phi^{(k)}$ such that $V^{(k)}$ is removed, if possible. Note that x has dimension n^c, i.e. the number of critical modes, which may be greater than one. In the following analysis therefore, we will treat x as a vector $\mathbf{x}$. Now

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \partial_{\mathbf{x}'} \boldsymbol{\Phi}(\mathbf{x}') \cdot \frac{d\mathbf{x}'}{dt} = \partial_{\mathbf{x}'} \boldsymbol{\Phi}(\boldsymbol{\Phi}^{-1}(\mathbf{x})) \cdot \mathbf{V}(\boldsymbol{\Phi}^{-1}(\mathbf{x})) \\ \mathbf{x}_{t+1} &= \boldsymbol{\Phi}(\mathbf{x}'_{t+1}) = \mathbf{F}(\boldsymbol{\Phi}^{-1}(\mathbf{x}_t)) + \boldsymbol{\phi}^{(k)}(\mathbf{F}(\boldsymbol{\Phi}^{-1}(\mathbf{x}_t))). \end{aligned}$$

Expanding the right hand sides gives

$$\begin{aligned} \mathbf{V}(\boldsymbol{\Phi}^{-1}(\mathbf{x})) &= \mathbf{V}(\mathbf{x} - \boldsymbol{\phi}^{(k)}(\mathbf{x}) + \mathcal{O}(x^{2k-1})) \\ &= \mathbf{V}(\mathbf{x}) - \partial_{\mathbf{x}} \mathbf{V}^{(1)}(\mathbf{x}) \boldsymbol{\phi}^{(k)}(\mathbf{x}) + \mathcal{O}(x^{k+1}) \\ \mathbf{F}(\boldsymbol{\Phi}^{-1}(\mathbf{x}_t)) &= \mathbf{F}((\mathbf{x} - \boldsymbol{\phi}^{(k)}(\mathbf{x}) + \mathcal{O}(x^{2k-1})) \\ &= \mathbf{F}(\mathbf{x}) - \partial_{\mathbf{x}} \mathbf{F}^{(1)}(\mathbf{x}) \boldsymbol{\phi}^{(k)}(\mathbf{x}) + \mathcal{O}(x^{k+1}), \end{aligned}$$

and

$$\partial_{\mathbf{x}} \boldsymbol{\Phi}(\boldsymbol{\Phi}^{-1}(\mathbf{x})) = \mathbb{1} + \partial_{\mathbf{x}} \boldsymbol{\phi}^{(k)}(\boldsymbol{\Phi}^{-1}(\mathbf{x})) = \mathbb{1} + \partial_{\mathbf{x}} \boldsymbol{\phi}^{(k)}(\mathbf{x}) + \mathcal{O}(x^{2k-2}).$$

Therefore the equations of motion are

$$\frac{d\mathbf{x}}{dt} = \mathbf{V}(\mathbf{x}) - \mathbf{L}(\boldsymbol{\phi}^{(k)}(\mathbf{x})) + \mathcal{O}(x^{k+1}) \qquad \mathbf{x}_{t+1} = \mathbf{F}(\mathbf{x}_t) - \tilde{\mathbf{L}}(\boldsymbol{\phi}^{(k)}(\mathbf{x}_t)) + \mathcal{O}(x_t^{k+1}) \tag{A.6}$$

where

$$\begin{aligned} \mathbf{L}(\boldsymbol{\phi}^{(k)}(\mathbf{x})) &= \partial_{\mathbf{x}} \mathbf{V}^{(1)}(\mathbf{x}) \boldsymbol{\phi}^{(k)}(\mathbf{x}) - \partial_{\mathbf{x}} \boldsymbol{\phi}^{(k)}(\mathbf{x}) \mathbf{V}^{(1)}(\mathbf{x}) \\ \tilde{\mathbf{L}}(\boldsymbol{\phi}^{(k)}(\mathbf{x})) &= \partial_{\mathbf{x}} \mathbf{F}^{(1)}(\mathbf{x}) \boldsymbol{\phi}^{(k)}(\mathbf{x}) - \boldsymbol{\phi}^{(k)}(\mathbf{F}^{(1)}(\mathbf{x})). \end{aligned} \tag{A.7}$$

Terms of order x^k are removed if $\boldsymbol{\phi}^{(k)}(\mathbf{x}) = \mathbf{L}^{-1}(\mathbf{V}^{(k)}(\mathbf{x}))$ which requires $\det \mathbf{L} \neq 0$. The essential terms to be retained are those $\mathbf{V}_c^{(k)}(\mathbf{x})$ in the kernel of $\mathbf{L}(\mathbf{V}, \boldsymbol{\phi}, \mathbf{x})$ (i.e. those for which $\mathbf{L}(\mathbf{V}, \boldsymbol{\phi}, \mathbf{x})$ has zero eigenvalues). Although this can be achieved in principle, it is more usual to remove unnecessary terms to finite order, and consider the remaining terms as perturbations.

We now have equation of motion determining the change in behaviour of the system at the critical point.

$$\frac{d\mathbf{x}}{dt} = \mathbf{V}_c(\mu, \mathbf{x}) = \mathbf{V}_c^{(1)}(\mu, \mathbf{x}) + \mathbf{V}_c^{(2)}(\mu, \mathbf{x}) + \ldots \qquad \mathbf{x}_{t+1} = \mathbf{F}_c(\mu, \mathbf{x}) = \mathbf{F}_c^{(1)}(\mu, \mathbf{x}_t) + \mathbf{F}_c^{(2)}(\mu, \mathbf{x}_t) + \ldots. \tag{A.8}$$

The change in behaviour is now determined by two criteria

1. The form of the relevant eigenvalue of the linearised equation at the critical point, eg. real, imaginary etc.

2. The non-vanishing orders in the expansion of $V_c(\mu, x)$ at the critical point.

Systems with a complex conjugate pair of relevant eigenvalues display Hopf bifurcations to cyclical behaviour, with a two dimensional normal form, whereas systems with real relevant eigenvalues display steady state bifurcations to fixed points (for sequential updating) or period two cycles (for synchronous updating), with a one-dimensional normal form. Steady state bifurcations can be further classified by the lowest non-vanishing orders in the expansion of the equations of motion at the critical point. Hopf bifurcations of systems with discretised time can be classified by their resonances, i.e. the value of m for which $\tilde{\lambda}^m = 1$, if $\tilde{\lambda}$ is the relevant eigenvalue of the linearised equations. These resonances introduce extra non-linear terms into the normal form causing mode-locking in quasi-periodic systems. This type of behaviour can also be considered as a steady state bifurcation of the map $\mathbf{x}_{t+m} = f(\mathbf{x}_t)$.

A.5 Classifying Normal Forms

We have shown that close to a hyperbolic equilibrium point, the dynamics of the linearised system is qualitatively the same as the full non-linear system. At a critical value of $\mu = \mu_c$ however hyperbolicity is lost, and we need to study what happens to the system as μ is varied about μ_c. We do this using the techniques of centre manifold reduction and normal form theory. We first consider only the neutral modes at μ_c and reduce the system to the appropriate centre manifold. We reduce the system to a simpler form, by applying appropriate co-ordinate changes yielding the normal form. We then include terms relevant as μ is varied away from μ_c, and study the bifurcations that occur. The advantage of using the normal form dynamics is that the bifurcation phenomena fall into distinct sets which are independent of the precise parameters of the system. The qualitative properties of such systems can therefore be worked out *a priori* without reference to any specific example. We classify the bifurcations according to the normal mode eigenvalues at criticality. In this section we describe the dynamics of the most common normal forms, close to the critical point.

Systems with only one critical eigenvalue have steady state bifurcations, to fixed point or low period cycles. Systems with a complex conjugate pair of critical eigenvalues however give rise to Hopf bifurcations, leading to more complex periodic or quasi-periodic behaviour.

A.5.1 Steady State Bifurcations

A steady state bifurcation has non-degenerate eigenvalues $\lambda = 0$ for flows, and $\tilde{\lambda} = 1$ for maps. This therefore yields a one dimensional normal form, with conditions at criticality

$$V(0,0) = 0\ ,\ \frac{\partial V}{\partial x}(0,0) = 0 \qquad F(0,0) = 0\ ,\ \frac{\partial F}{\partial x}(0,0) = 1.$$

These conditions become equivalent if we make the substitution $\hat{V}(\mu, x) = F(\mu, x) - x$. We can now consider both flows and maps as the same generic case. The Taylor expansion of the

centre-manifold dynamics then becomes

$$\left.\begin{array}{c}\dot{x}\\ x_{t+1}-x_t\end{array}\right\} = \mu\,\frac{\partial V}{\partial \mu}(0,0)+\frac{x_t^2}{2}\,\frac{\partial^2 V}{\partial x^2}(0,0)+\mu\,x_t\,\frac{\partial^2 V}{\partial \mu\,\partial x}(0,0)+\frac{\mu^2}{2}\,\frac{\partial^2 V}{\partial \mu^2}(0,0)+\ldots$$

with the left hand side $\dot{x}$ or $x_{t+1}-x_t$ as appropriate to the dynamics (continuous or discrete time).

Saddle-Node Bifurcation

A Saddle Node Bifurcation is the generic case of the steady state bifurcations, with $\frac{\partial V}{\partial \mu}(0,0)\neq 0$, $\frac{\partial^2 V}{\partial x^2}(0,0)\neq 0$. We truncate the expansion to first non-vanishing orders in μ and x giving

$$\mu\,\frac{\partial V}{\partial \mu}(0,0)\,(1+\mathcal{O}(\mu,x))+\frac{x^2}{2}\,\frac{\partial^2 V}{\partial x^2}(0,0)\,(1+\mathcal{O}(\mu,x)).$$

We now re-scale the co-ordinate to obtain the normal form

$$\dot{\tilde{x}} = \epsilon_1\tilde{\mu}+\epsilon_2\tilde{x}^2 = \tilde{V}(\tilde{\mu},\tilde{x}) \qquad \tilde{x}_{t+1} = \tilde{x}_t+\epsilon_1\tilde{\mu}+\epsilon_2\tilde{x}_t^2 = \tilde{F}(\tilde{\mu},\tilde{x})$$

where

$$\mu = \frac{2\tilde{\mu}}{\left|\frac{\partial V}{\partial \mu}(0,0)\,\frac{\partial^2 V}{\partial x^2}(0,0)\right|} \qquad x = \frac{2\tilde{x}}{\left|\frac{\partial^2 V}{\partial x^2}(0,0)\right|}$$

$$\epsilon_1 = \mathrm{sgn}\left(\frac{\partial V}{\partial \mu}(0,0)\right) \qquad \epsilon_2 = \mathrm{sgn}\left(\frac{\partial^2 V}{\partial x^2}(0,0)\right).$$

We can see from this that there are two fixed points which coalesce at criticality, given by $\tilde{x} = \pm\sqrt{-\frac{\epsilon_1}{\epsilon_2}\tilde{\mu}}$. These are only physical for positive arguments of the square root, and only one of the solutions is stable, since for stability $2\epsilon_2 x < 0$.

Transcritical Bifurcation

A transcritical bifurcation is one for which $\frac{\partial V}{\partial \mu}(0,0)=0$, but $\frac{\partial^2 V}{\partial \mu \partial x}(0,0)\neq 0$ and $\frac{\partial^2 V}{\partial x^2}(0,0)\neq 0$. We proceed exactly as before to obtain the normal form in rescaled variables

$$\dot{\tilde{x}} = \tilde{x}(\epsilon_1\tilde{\mu}+\epsilon_2\tilde{x}) = \tilde{V}(\tilde{\mu},\tilde{x}) \qquad \tilde{x}_{t+1} = \tilde{x}_j\,(1+\epsilon_1\tilde{\mu}+\epsilon_2\tilde{x}_t) = \tilde{F}(\tilde{\mu},\tilde{x})$$

where

$$\epsilon_1 = \mathrm{sgn}\left(\frac{\partial^2 V}{\partial \mu \partial x}(0,0)\right) \qquad \epsilon_2 = \mathrm{sgn}\left(\frac{\partial^2 V}{\partial x^2}(0,0)\right).$$

We can see that fixed point are $\tilde{x}=0$ and $\tilde{x}=-\frac{\epsilon_1}{\epsilon_2}\tilde{\mu}$, which are stable when $\epsilon_1\tilde{\mu}+2\epsilon_2\tilde{x}<0$.

A.5.2 Pitchfork or Period Doubling Bifurcation

Pitchfork bifurcations are specific to discrete time dynamics. They are similar to steady state bifurcations, with only one critical eigenvalue, except that $\frac{\partial F}{\partial x}(0,0)=-1$ rather than $+1$. We notice however that this is exactly the same as a steady state bifurcation for $F(0,F(0,x))$, i.e. the map acting twice on the original co-ordinate. We therefore make the substitution $\hat{V}(\mu,x) = f^2(\mu,x)-x$, then proceed as before.

Pitchfork bifurcations occur for systems in which $\frac{\partial^2 V}{\partial x^2}(0,0) = 0$ but $\frac{\partial^3 V}{\partial x^3}(0,0) \neq 0$, typically due to the inversion symmetry $-V(\mu, x) = V(\mu, -x)$, and $\frac{\partial^2 V}{\partial \mu \partial x}(0,0) \neq 0$. In this case the normal form is

$$\dot{\tilde{x}} = \tilde{x}(\epsilon_1 \tilde{\mu} + \epsilon_2 \tilde{x}^2) = \tilde{V}(\tilde{\mu}, \tilde{x}) \qquad \tilde{x}_{t+1} = \tilde{x}_t \left(1 + \epsilon_1 \tilde{\mu} + \epsilon_2 \tilde{x}_t^2\right) = \tilde{F}(\tilde{\mu}, \tilde{x})$$

with

$$\epsilon_1 = \text{sgn}\left(\frac{\partial^2 V}{\partial \mu \partial x}(0,0)\right) \qquad \epsilon_2 = \text{sgn}\left(\frac{\partial^3 V}{\partial x^3}(0,0)\right).$$

We can see that fixed point are $\tilde{x} = 0$ and $\tilde{x} = \pm\sqrt{-\frac{\epsilon_1}{\epsilon_2}\tilde{\mu}}$, which are stable when $\epsilon_1 \tilde{\mu} + 3\epsilon_2 \tilde{x}^2 < 0$. These are known as Pitchfork bifurcations due to the shape of the bifurcation diagrams - above the critical point, the solution oscillates between the two branches.

A.5.3 Hopf Bifuractions

Hopf Bifurcations are caused by a complex conjugate pair of eigenvalues reaching the criticality condition. They are slightly different from the cases studied so far, in that the normal form is a two-dimensional set of equations, usually expressed in polar co-ordinates. If for flows and maps respectively we have the complex conjugate pair of eigenvalues

$$\lambda(\mu) = \gamma(\mu) \pm i\omega(\mu) \qquad \tilde{\lambda}(\mu) = (1 + a(\mu))e^{i2\pi\psi(1+b(\mu))}$$

$$\gamma(0) = 0 \quad \omega(0) \neq 0 \quad \frac{d\gamma}{d\mu}(0) > 0 \qquad a(0) = b(0) = 0 \quad \frac{da}{d\mu}(0) > 0 \qquad 0 < \psi < \frac{1}{2},$$

then $V^{(1)}$ and $F^{(1)}$) are

$$\mathbf{V}^{(1)} = \begin{pmatrix} \gamma(\mu) & \omega(\mu) \\ -\omega(\mu) & \gamma(\mu) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\mathbf{F}^{(1)} = \begin{pmatrix} (1 + a(\mu)) \cos 2\pi\psi(1 + b(\mu)) & -(1 + a(\mu)) \sin 2\pi\psi(1 + b(\mu)) \\ (1 + a(\mu)) \sin 2\pi\psi(1 + b(\mu)) & (1 + a(\mu)) \cos 2\pi\psi(1 + b(\mu)) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It is convenient to introduce complex co-ordinates $\begin{pmatrix} z \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \mathbf{S} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ so that

$$\mathbf{S}\,\partial_{\mathbf{x}}\mathbf{V}^{(1)}\,\mathbf{S}^{-1} = \begin{pmatrix} \gamma - i\omega & 0 \\ 0 & \gamma + i\omega \end{pmatrix}$$

$$\mathbf{S}\,\partial_{\mathbf{x}}\mathbf{F}^{(1)}\,\mathbf{S}^{-1} = \begin{pmatrix} (1 + a)e^{i2\pi\psi(1+b)} & 0 \\ 0 & (1 + a)e^{-i2\pi\psi(1+b)} \end{pmatrix}. \tag{A.9}$$

We denote these eigenvalues by λ and $\bar{\lambda}$, and write $\boldsymbol{\phi}^{(k)}(z, \bar{z}) = \mathbf{S}\boldsymbol{\phi}^{(k)}(\mathbf{x})$. We can now write our equation for $\mathbf{L}(\boldsymbol{\phi}^{(k)}$ ($\tilde{\mathbf{L}}(\boldsymbol{\phi}^{(k)})$ explicitly

$$\mathbf{L}(\boldsymbol{\phi}^{(k)}) = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \begin{pmatrix} \phi_z^{(k)} \\ \phi_{\bar{z}}^{(k)} \end{pmatrix} - \begin{pmatrix} \frac{\partial \phi_z^{(k)}}{\partial z} & \frac{\partial \phi_z^{(k)}}{\partial \bar{z}} \\ \frac{\partial \phi_{\bar{z}}^{(k)}}{\partial z} & \frac{\partial \phi_{\bar{z}}^{(k)}}{\partial \bar{z}} \end{pmatrix} \begin{pmatrix} \lambda z \\ \bar{\lambda}\bar{z} \end{pmatrix}$$

$$\tilde{\mathbf{L}}(\boldsymbol{\phi}^{(k)}) = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \begin{pmatrix} \phi_z^{(k)}(z, \bar{z}) \\ \phi_{\bar{z}}^{(k)}(z, \bar{z}) \end{pmatrix} - \begin{pmatrix} \phi_z^{(k)}(\lambda z, \bar{\lambda}\bar{z}) \\ \phi_{\bar{z}}^{(k)}(\lambda z, \bar{\lambda}\bar{z}) \end{pmatrix}.$$

We can see that eigenvectors are given by $\xi_+^{(k,l)}(z,\bar{z}) = \begin{pmatrix} z^l\bar{z}^{k-l} \\ 0 \end{pmatrix}$, and $\xi_-^{(k,l)}(z,\bar{z}) = \begin{pmatrix} 0 \\ z^l\bar{z}^{k-l} \end{pmatrix}$ with $l = 0, 1, \ldots, k$, and the eigenvalues $\alpha_\pm^{(k,l)}(\mu)$ ($\tilde{\alpha}_\pm^{(k,l)}(\mu)$) are given by

$$\begin{aligned} \alpha_\pm^{(k,l)}(\mu) &= (1-k)\gamma(\mu) - i\omega(\mu)(k-2l\pm 1) \\ \tilde{\alpha}_\pm^{(k,l)}(\mu) &= (1+a(\mu))e^{\pm i2\pi\psi(1+b(\mu))}(1-(1+a)^{k-1}e^{-i2\pi\psi(1+b)(k-2l\pm 1)}). \end{aligned}$$

Now to reduce to the minimal representation we require $\det \mathbf{L} = 0$, i.e. we require at least one zero eigenvalue.

Continuous time

In the continuous time case $\det \mathbf{L} = 0$ requires $(1-k)\gamma(\mu) = 0$ and $(k-2l\pm 1)\omega(\mu) = 0$. We see that this can only be satisfied at the critical point ($\omega(0) = 0$), and for odd k. Natural basis vectors for V_c are therefore

$$\xi_+^{(k,\frac{k+1}{2})}(z,\bar{z}) = \begin{pmatrix} z|z|^{k-1} \\ 0 \end{pmatrix} \qquad \xi_-^{(k,\frac{k-1}{2})}(z,\bar{z}) = \begin{pmatrix} 0 \\ \bar{z}|z|^{k-1} \end{pmatrix} \qquad k = 3,5,7,\ldots$$

and the equation of motion, with all even non-linear terms removed is

$$\begin{pmatrix} \dot{z} \\ \dot{\bar{z}} \end{pmatrix} = \begin{pmatrix} \gamma - i\omega & 0 \\ 0 & \gamma + i\omega \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \end{pmatrix} + \sum_{j=1}^{\infty} \begin{pmatrix} a_j z|z|^{2j} \\ \bar{a}_j \bar{z}|z|^{2j} \end{pmatrix}.$$

Discrete time

The discrete time case is slightly more complicated. $\det \tilde{L} = 0$ requires $\tilde{\alpha} = 0$, which can only be achieved at criticality (a=b=0), or if $e^{-i2\pi\psi(k-2l\pm 1)} = 1$. Solutions vary according to whether ψ is rational or irrational, since we require $\psi(k-2l\pm 1) = m$, where m is an integer. If ψ is irrational this can only occur for $k-2l\pm 1 = 0$, giving similar solutions to the continuous time case

$$z_{t+1} = z_t \left(\lambda + \sum_{j=1}^{\infty} a_j |z_t|^{2j}\right).$$

If however ψ is rational $= \frac{p}{q}$ then we have solutions $k-2l\pm 1 = nq$, with n a positive or negative integer. If these resonant conditions are satisfied for $q = m$ then basis vectors for removing the m^{th} order terms are $\xi_+(z,\bar{z}) = \begin{pmatrix} \bar{z}^{m-1} \\ 0 \end{pmatrix}$ and $\xi_-(z,\bar{z}) = \begin{pmatrix} 0 \\ z^{m-1} \end{pmatrix}$. This introduces θ dependent terms of order θ_t^{m-1} on the right hand side of the equations of motion. Irrational ψ causes quasi-periodic solutions, since the system can never return exactly to its original state, and the resonances are responsible for phenomena such as mode locking in circle maps.

In polar co-ordinates the normal forms are

$$\begin{aligned} \dot{r} &= r\left(\gamma(\mu) + \sum_{i=1}^{\infty} a_i(\mu) r^{2i}\right) & r_{t+1} &= (1+a(\mu))r_t\left(1 + \sum_{i=1}^{\infty} a_i(\mu) r_t^{2i}\right) \\ \dot{\theta} &= \omega(\mu) + \sum_{i=1}^{\infty} b_i(\mu) r^{2i} & \theta_{t+1} &= \theta_t + 2\pi\psi(1+b(\mu)) + \sum_{i=1}^{\infty} b_i(\mu) r^{2i}. \end{aligned} \quad \text{(A.10)}$$

Notice that the right hand sides of the equations are independent of θ. For maps this is only strictly true if $\tilde{\lambda}$ satisfies the non-resonance conditions $\tilde{\lambda}^n(\mu) \neq 1$, for any integer n. If a resonance occurs at some integer m, then there will be a θ dependence of order θ_t^{m-1}.

The simplest non-trivial case is $a_1(\mu) \neq 0$, giving cycles of approximate radii $r = 0$ or $r = \sqrt{-\frac{\gamma(\mu)}{a_1}}$ for flows and $r = \sqrt{-\frac{a}{a_1}}$ for maps. Close to criticality frequencies are given by $\omega = \omega(\mu) + \sum_{i=1}^{\infty} b_i(\mu)(-\frac{\gamma}{a_1})^i$ for flows, and $\theta_{t+1} - \theta_t = 2\pi\psi(1 + b(\mu)) + \sum_{i=1}^{\infty} b_i(\mu)(-\frac{a}{a_1})^i$. Although the equations for $\dot{\theta}$ and $\theta_{t+1} - \theta_t$ have been assumed to be independent of θ, small perturbations introducing a non-linear dependence on θ cause a variety of phenomena to appear in the dynamics of the circle, such as model locking.

Quasi-Periodicity and Mode Locking

Quasi-Periodicity and Mode Locking are phenomena that have been observed in many non-linear systems, and have become well understood through the work of Jensen, Bak, Bohr [99, 100, 101]; Cvitanovic et al [102, 103]; Feigenbaum, Kadanoff, Shenker [104, 105]; and Rand, Ostlund, Sethna, Siggia [106]. Quasi-periodicity typically occurs when a non-linear system is driven externally at a frequency incommensurate with its natural frequency. If the non-linearity and the driving frequency are sufficiently strong, the system will tend to lock into rational multiples of the driving frequency. This gives behaviour such as the devil's staircase, with universal properties. Work has mainly centered on studies of the one dimensional circle map

$$\theta_{t+1} = f(\theta_t) \equiv \theta_t + \Omega - \frac{K}{2\pi}\sin(2\pi\theta_t) \qquad \text{mod } 1. \tag{A.11}$$

It has been shown numerically, that maps with the same general properties as (A.11) have universal properties, hence the exact form of the equations seems unimportant. Many real systems can be reduced, or at least approximated, to the one dimensional circle map, who's properties have been well documented both analytically and numerically. In contrast to other common one-dimensional maps used as models of complex non-linear systems (eg. the quadratic map [76, 77]) the circle map has two control parameters - Ω and K. Ω is the the ratio of the external driving frequency to the natural frequency of the system, and K is the strength of the non-linearity of the mapping. Two-frequency motion is topologicaly equivalent to motion on a two-torus, and it can be shown that the appearance of a third frequency in the system is unstable to chaos [107], signalled by the appearance of broad-band noise. If the frequencies are incommensurate, the trajectory will eventually cover the entire manifold, whereas trajectories with a rational ratio of frequencies will eventually close. The circle map (A.11) displays a transition to chaos at K=1, where it loses its invertibility, and gains a cubic inflection point.

The usual choice for the observable in the study of the circle map is the winding number w

$$w = \lim_{n\to\infty} \frac{f^n(\theta_0) - \theta_0}{n},$$

which gives the average rotation per iteration. If the cycle closes after a certain number, q, of iterations then w is rational ($= \frac{p}{q}$). If however the cycle never closes then w is irrational. Much of the work on this subject has centred on the transition to chaos of systems with irrational winding numbers. The behaviour of w, which characterises $f(\theta)$, as the driving frequency Ω is varied shows the *devil's staircase* structure. As Ω is increased there exist plateaux of finite width for which

w locks into a rational value. As Ω is increased, there is a discontinuous jump to the next step. There are three regions of different qualitative behaviour, for different values of the non-linearity parameter K.

1. For $K < 1$ the staircase structure is incomplete. i.e. there exist plateaux of finite width in Ω space locked into rational values of w. In between these plateaux, however, are regions in which the plateaux have zero measure.

2. At $K = 1$, the staircase becomes complete. i.e. for every rational value of the winding number, w, there exists a range of Ω of non-zero measure. In between any two plateaux a smaller plateau can be found with smaller range in Ω space. $K = 1$ is the transition to chaos, at which the map loses invertibility. Because the staircase structure is complete and self-similar at $K = 1$, the theory is renormalisable. Much work has been done on detecting the transition to chaos by the fixed points of the renormalisation scheme [99, 100, 101, 102, 103, 104, 105, 106].

3. For $K > 1$ the map is non-invertible, and the plateaux in the devils staircase begin to overlap. In this region the map exhibits hysteresis, where $w(\Omega)$ is dependent on how it is approached.

The plateaux are known as Arnold tongues [78, 108] and appear in many non-linear systems, where the system locks into a certain cycle over a range of parameter values. They are related to resonances of the eigenvalue $\tilde{\lambda}$.

Appendix B

Derivations for the Dynamics of the extensive p Hopfield model

This appendix contains the details of two calculations encountered in Chapter 5 where the dynamics of the generalised Hopfield model trained with an extensive number of patterns is discussed.

B.1 Calculation of the Intensive Contribution, $\mathcal{R}$, to the Intrinsic Noise distribution

In this appendix we carry out some of the manipulations necessary to calculate the intensive contribution $\mathcal{R}$ (5.7) to the intrinsic noise distribution. We first define

$$\begin{aligned}
g_a^\mu &= \frac{\int d\mathbf{z}\, e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}} z_a^\mu}{\int d\mathbf{z} e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}} = i\frac{\partial}{\partial\Upsilon_\alpha^\mu}\ln\left[\int d\mathbf{z} e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}\right] = -i\sum_b\sum_\nu\left(\Lambda^{-1}\right)_{\mu\nu}^{ab}\Upsilon_b^\nu \\
&= \rho\sqrt{N}\sum_b\sum_\nu\left(\Lambda^{-1}\right)_{\mu\nu}^{ab}A_{\nu\eta}^s m^\eta + \mathcal{O}(\frac{1}{N}) \\
g_{ab}^{\mu\nu} &= \frac{\int d\mathbf{z} e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}} z_a^\mu z_b^\nu}{\int d\mathbf{z} e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}} = -\frac{\partial^2}{\partial\Upsilon_a^\mu\partial\Upsilon_b^\nu}\ln\left[\int d\mathbf{z} e^{-\frac{1}{2}\mathbf{z}\cdot\boldsymbol{\Lambda}\mathbf{z}-i\boldsymbol{\Upsilon}\cdot\mathbf{z}}\right] + g_a^\mu g_b^\nu \\
&= \left(\Lambda^{-1}\right)_{\mu\nu}^{ab} - \left(\sum_c\sum_\eta\left(\Lambda^{-1}\right)_{\mu\eta}^{ac}\Upsilon_c^\eta\right)\left(\sum_d\sum_\rho\left(\Lambda^{-1}\right)_{\nu\rho}^{bd}\Upsilon_d^\rho\right),
\end{aligned}$$

and $\Gamma_\mu = \sum_a\left(\sum_{\nu>c}\Xi_{\mu\nu}^a z_a^\nu + \sum_{\nu\le c}\Pi_{\mu\nu}^a m^\nu(\boldsymbol{\sigma}^a)\right)$ where

$$\begin{aligned}
\Xi_{\mu\nu}^a &= \frac{1}{\sqrt{N}}\left(\left(xA_{\mu\nu}^s + yA_{\mu\nu}^a\right)\delta_{\alpha 1} + i\rho A_{\mu\nu}^s\sigma_i^a\right) \qquad \mu,\nu > c \\
\mathrm{H}_{\mu\nu}^a &= \left(\left(xA_{\mu\nu}^s + yA_{\mu\nu}^a\right)\delta_{\alpha 1} + i\rho A_{\mu\nu}^s\sigma_i^a\right) \qquad \mu > c, \nu \le c.
\end{aligned}$$

Then $\mathcal{R}$ (5.7) becomes

$$\begin{aligned}
\mathcal{R} &= ix\rho\sum_b\sum_{\mu,\rho\le c}\sum_{\nu,\eta>c}\zeta^\mu A_{\mu\nu}^s\left(\Lambda^{-1}\right)_{\nu\eta}^{1b}A_{\eta\rho}^s m^\rho + \frac{1}{2}x^2\sum_{\mu,\rho\le x}\sum_{\nu,\eta>c}\zeta^\mu A_{\mu\nu}^s\left(\Lambda^{-1}\right)_{\nu\eta}^{11}A_{\eta\rho}^s\zeta^\eta \qquad \text{(B.1)} \\
&\quad -\frac{1}{2}y^2\sum_{\mu,\rho\le x}\sum_{\nu,\eta>c}\zeta^\mu A_{\mu\nu}^a\left(\Lambda^{-1}\right)_{\nu\eta}^{11}A_{\eta\rho}^a\zeta^\eta + \frac{1}{2}\sum_{\mu>c}\Big[\Xi_{\mu\nu}^a\left(\Lambda^{-1}\right)_{\nu\eta}^{ab}\Xi_{\mu\eta}^b \\
&\quad + \left(-i\Xi_{\mu\nu}^a\left(\Lambda^{-1}\right)_{\nu\lambda}^{ac}\Upsilon_c^\lambda + \Pi_{\mu\lambda}^c m^\lambda\right)\left(-i\Xi_{\mu\eta}^b\left(\Lambda^{-1}\right)_{\eta\rho}^{bd}\Upsilon_d^\rho + \Pi_{\mu\rho}^d m^\rho\right)\Big].
\end{aligned}$$

The symmetric and antisymmetric terms decouple. We neglect terms of order $\frac{1}{N}$ and define auxiliary variables $\tilde{\lambda}$, $\mathcal{A}_{\mu\nu}$, $\tilde{r}$, $\tilde{\mathcal{R}}_{\mu\nu}$, $\tilde{p}$ and $\tilde{\mathcal{P}}_{\mu\nu}$ such that

$$\frac{1}{N}\sum_{\mu,\nu,\lambda>c} A^s_{\lambda\mu}\left(\Lambda^{-1}\right)^{a\neq b}_{\mu\nu} A^s_{\nu\lambda} = \frac{\tilde{\lambda}^2}{\rho^2} = \frac{1}{N}\operatorname{Tr}\mathcal{A} \qquad \frac{1}{N}\sum_{\mu,\nu,\lambda>c} A^s_{\lambda\mu}\left(\Lambda^{-1}\right)^{a=b}_{\mu\nu} A^s_{\nu\lambda} = \alpha\tilde{r} = \frac{1}{N}\operatorname{Tr}\tilde{\mathcal{R}}$$

$$\frac{1}{N}\sum_{\mu,\nu,\lambda>c} A^a_{\lambda\mu}\left(\Lambda^{-1}\right)^{11}_{\mu\nu} A^a_{\nu\lambda} = \alpha\tilde{p} = \frac{1}{N}\operatorname{Tr}\tilde{\mathcal{P}}$$

where

$$\begin{aligned}
\tilde{\lambda}^2 &= \frac{\alpha q\rho^2}{p}\operatorname{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right]\\
\tilde{r}_2 &= \frac{1}{p}\operatorname{Tr}\left[\mathbf{A}^s_{uu}\mathbf{A}^s_{uu}\left(\mathbb{1}-\rho(1-q)(1-q+nq)\mathbf{A}^s_{uu}\right)\right.\\
&\qquad \left.\left(\mathbb{1}-\rho(1-q+nq)\mathbf{A}^s_{uu}\right)^{-1}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\right].
\end{aligned}$$

We can then simplify the expression for $\mathcal{R}$ to

$$\begin{aligned}
\mathcal{R} &= ix\rho\boldsymbol{\zeta}\cdot\left(\tilde{\mathcal{R}}+(n-1)\mathcal{A}\right)\mathbf{m}+\frac{x^2}{2N}\boldsymbol{\zeta}\cdot\tilde{\mathcal{R}}\boldsymbol{\zeta}-\frac{y^2}{2N}\boldsymbol{\zeta}\cdot\tilde{\mathcal{P}}\boldsymbol{\zeta}+\frac{\alpha}{2}x^2\tilde{r}+ix\rho\left(\alpha\tilde{r}\sigma^1+\frac{\tilde{\lambda}^2}{\rho^2}\sum_{a>1}\sigma^a\right)\\
&\quad -\frac{\rho^2}{2}\left(\alpha\tilde{r}\sum_a\sigma^a\sigma^a+\frac{\tilde{\lambda}^2}{\rho^2}\sum_{a,b\neq a}\sigma^a\sigma^b\right)-\frac{\alpha}{2}y^2\tilde{p}\\
&\quad +\frac{1}{2}\rho^2x^2\mathbf{m}\cdot\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\mathbf{m}\\
&\quad +i\rho^3x\mathbf{m}\cdot\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\mathbf{m}\sum_a\sigma^a_i\\
&\quad -\frac{1}{2}\rho^4\mathbf{m}\cdot\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\left((n-1)\mathcal{A}+\tilde{\mathcal{R}}\right)\mathbf{m}\left(\sum_a\sigma^a\right)^2\\
&\quad -\frac{1}{2}y^2\rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\tilde{\mathcal{P}}\mathbf{A}^s_{uc}\mathbf{m}-y^2\rho\mathbf{m}\cdot\sum_b\mathbf{A}^a_{cu}\mathbf{A}^a_{uu}\left(\Lambda^{-1}\right)^{1b}\mathbf{A}^s_{uc}\mathbf{m}\\
&\quad +x^2\rho\mathbf{m}\cdot\mathbf{A}^s_{cu}\left(\tilde{\mathcal{R}}+(n-1)\mathcal{A}\right)\mathbf{m}\\
&\quad +i\rho^2x\mathbf{m}\cdot\mathbf{A}^s_{cu}\mathbf{A}^s_{uu}\left(\sum_{a,b}\sigma^a\left(\Lambda^{-1}\right)^{ab}+\sum_{c,b}\sigma^c\left(\Lambda^{-1}\right)^{1b}\right)\mathbf{A}^s_{uc}\mathbf{m}\\
&\quad -\rho^3\mathbf{m}\cdot\mathbf{A}^s_{cu}\mathbf{A}^s_{uu}\sum_c\sigma^c\sum_{a,b}\sigma^a\left(\Lambda^{-1}\right)^{ab}\mathbf{A}^s_{uc}\mathbf{m}+\frac{1}{2}x^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\mathbf{A}^s_{uc}\mathbf{m}\\
&\quad +i\rho x\mathbf{m}\cdot\mathbf{A}^s_{cu}\mathbf{A}^s_{uc}\mathbf{m}\sum_a\sigma^a-\frac{\rho^2}{2}\mathbf{m}\cdot\mathbf{A}^s_{cu}\mathbf{A}^s_{uc}\mathbf{m}\sum_{a,b}\sigma^a\sigma^b-\frac{1}{2}y^2\mathbf{m}\cdot\mathbf{A}^a_{cu}\mathbf{A}^a_{uc}\mathbf{m}. \qquad \text{(B.2)}
\end{aligned}$$

We drop terms of $\mathcal{O}(\frac{1}{N})$ and hence the remaining part of the integral is of the form

$$\begin{aligned}
&\mathcal{D}_{\boldsymbol{\zeta}}[z^s,z^a]\sim\int dx\,dy\,e^{ixz^s+iyz^a}e^{-\frac{1}{2}y^2\mathcal{G}}\times \qquad \text{(B.3)}\\
&\left\langle e^{\sum_{\nu\leq c}\xi^\nu\mu^\nu\sum_a\sigma^a-ix\mathcal{B}-ix\rho\left(\alpha\tilde{r}\sigma^1+\frac{\tilde{\lambda}^2}{\rho^2}\sum_{a>1}\sigma^a\right)-x^2\mathcal{E}-i\rho x\sum_a\mathcal{F}\sigma^a+\frac{1}{2}\Gamma^2\sum_{a,b}\sigma^a\sigma^b+\frac{n\rho^2\alpha\tilde{r}}{2}-\frac{n\tilde{\lambda}^2}{2}}\right\rangle_{\boldsymbol{\sigma}}
\end{aligned}$$

where

$$\mathcal{G}=-\alpha\tilde{p}-2\rho\mathbf{m}\cdot\mathbf{A}^a_{cu}\mathbf{A}^a_{uu}\mathbf{A}^s_{uu}{}^{-1}\left(\tilde{\mathcal{R}}+(n-1)\mathcal{A}\right)\mathbf{m}-\rho^2\mathbf{m}\cdot\mathbf{A}^s_{cu}\tilde{\mathcal{P}}\mathbf{A}^s_{uc}\mathbf{m}-\mathbf{m}\cdot\mathbf{A}^a_{cu}\mathbf{A}^a_{uc}\mathbf{m}$$

$$
\begin{aligned}
\mathcal{B} &= \rho\boldsymbol{\zeta}\cdot\left(\tilde{\mathcal{R}}_{cc}+(n-1)\mathcal{A}_{cc}\right)\mathbf{m}\\
\mathcal{E} &= \frac{1}{2}\mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uv}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}+\frac{\alpha\tilde{r}}{2}\\
\mathcal{F} &= \mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uc}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}\\
\Gamma^2 &= \tilde{\lambda}^2+\rho^2\mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uc}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}. \qquad \text{(B.4)}
\end{aligned}
$$

Using our definition of Λ^{-1} we see that

$$\tilde{\mathcal{R}}+(n-1)\mathcal{A}=(1-q+nq)\mathbf{A}^s\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-1}\mathbf{A}^s,$$

so that

$$
\begin{aligned}
&\mathbf{m}\cdot\left(\mathbf{A}^s_{cu}+\rho\left(\tilde{\mathcal{R}}_{cu}+(n-1)\mathcal{A}_{cu}\right)\right)\left(\mathbf{A}^s_{uc}+\rho\left(\tilde{\mathcal{R}}_{uc}+(n-1)\mathcal{A}_{uc}\right)\right)\mathbf{m}\\
&\qquad=\mathbf{m}\cdot\mathbf{A}^s_{cu}\left(\mathbb{1}-\rho(1-q)\mathbf{A}^s_{uu}\right)^{-2}\mathbf{A}^s_{uc}\mathbf{m},
\end{aligned}
$$

and therefore $\Gamma^2=\lambda^2$.

B.2 AT Surface

In this appendix we calculate the derivatives needed to determine the AT surface signalling the instability of the replica symmetric solutions encountered in the extremisation of the saddle point exponent (5.8). To calculate the derivatives we note that

$$\int dz\,\frac{\partial\mathcal{D}(\mathbf{z},\mathbf{q})}{\partial q_{ab}}F(\mathbf{z})=\frac{1}{2}\sum_{\mu}\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})\frac{\partial^2F(\mathbf{z})}{\partial z^\mu_a\partial z^\mu_b}.$$

Therefore

$$
\begin{aligned}
&\frac{N}{\rho^2}\frac{\partial^2\Omega}{\partial q_{ab}\partial q_{cd}}=\frac{1}{2}\sum_{\mu} \qquad \text{(B.5)}\\
&\frac{\partial}{\partial q_{ab}}\frac{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}\left((\mathbf{A}^s\mathbf{z}_c)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)\left((\mathbf{A}^s\mathbf{z}_d)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)}{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}}\\
&=\frac{1}{4}\sum_{\mu,\nu}\\
&\left\{\frac{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})\frac{\partial^2}{\partial z^\nu_a z^\nu_b}\left[e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}\left((\mathbf{A}^s\mathbf{z}_c)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)\left((\mathbf{A}^s\mathbf{z}_d)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)\right]}{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}}\right.\\
&-\rho^2\left(\frac{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}\left((\mathbf{A}^s\mathbf{z}_a)^\nu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right)\left((\mathbf{A}^s\mathbf{z}_b)^\nu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right)}{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}}\right)\\
&\left.\times\left(\frac{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}\left((\mathbf{A}^s\mathbf{z}_c)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)\left((\mathbf{A}^s\mathbf{z}_d)^\mu+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)}{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}}\right)\right\}.
\end{aligned}
$$

Using previous notation we write

$$g^\mu_a=\frac{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}z^\mu_a}{\int dz\,\mathcal{D}(\mathbf{z},\mathbf{q})e^{\frac{1}{2}\rho\mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}} \qquad \text{(B.6)}$$

$$
\begin{aligned}
g_{ab}^{\mu\nu} &= \frac{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}} z_a^\mu z_b^\nu}{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}} \\
g_{abc}^{\mu\nu\eta} &= \frac{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}} z_a^\mu z_b^\nu z_c^\eta}{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}} \\
g_{abcd}^{\mu\nu\eta\rho} &= \frac{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}} z_a^\mu z_b^\nu z_c^\eta z_d^\rho}{\int d\mathbf{z}\, \mathcal{D}(\mathbf{z},\mathbf{q}) e^{\frac{1}{2}\rho \mathbf{z}\cdot\mathbf{A}^s\mathbf{z}+\rho\sqrt{N}\mathbf{m}\cdot\mathbf{A}^s\mathbf{z}}}.
\end{aligned}
$$

(B.5) then becomes

$$
\begin{aligned}
&\frac{\rho^2}{4N}\Bigg\{\rho^2\sum_{\mu,\nu}\Bigg[\sum_{\alpha,\beta,\gamma,\delta} A^s_{\mu\alpha}A^s_{\mu\beta}A^s_{\nu\gamma}A^s_{\nu\delta}g^{\alpha\beta\gamma\delta}_{abcd} \\
&+\sqrt{N}\sum_{\alpha,\beta,\gamma}\left((\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\mu\alpha}A^s_{\nu\beta}A^s_{\nu\gamma}\left(g^{\alpha\beta\gamma}_{abc}+g^{\alpha\beta\gamma}_{abd}\right)+(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\nu\alpha}A^s_{\mu\beta}A^s_{\mu\gamma}\left(g^{\alpha\beta\gamma}_{acd}+g^{\alpha\beta\gamma}_{bcd}\right)\right) \\
&+N\sum_{\alpha,\beta}(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\nu\alpha}A^s_{\nu\beta}g^{\alpha\beta}_{ab} \\
&+\sum_{\alpha,\beta}(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\left(A^s_{\nu\alpha}A^s_{\mu\beta}g^{\alpha\beta}_{ac}+A^s_{\nu\alpha}A^s_{\mu\beta}g^{\alpha\beta}_{\alpha\delta}+A^s_{\nu\alpha}A^s_{\mu\beta}g^{\alpha\beta}_{bc}+A^s_{\nu\alpha}A^s_{\mu\beta}g^{\alpha\beta}_{bd}\right) \\
&+(\mathbf{m}\cdot\mathbf{A}^s)^\nu(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\mu\alpha}A^s_{\mu\beta}g^{\alpha\beta}_{cd} \\
&+N^{\frac{3}{2}}\sum_{\alpha}(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\left((\mathbf{m}\cdot\mathbf{A}^s)^\mu\left(A^s_{\nu\alpha}g^\alpha_a+A^s_{\nu\beta}g^\beta_b\right)+(\mathbf{m}\cdot\mathbf{A}^s)^\nu\left(A^s_{\mu\gamma}g^\gamma_c+A^s_{\mu\delta}g^\delta_d\right)\right) \\
&+N^2(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\Bigg] \\
&+\rho A^s_{\mu\nu}\delta_{bc}\left(A^s_{\nu\alpha}A^s_{\mu\delta}g^{\alpha\delta}_{ad}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\mu\delta}g^\delta_d+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\nu\alpha}g^\alpha_a+N(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right) \\
&+\rho A^s_{\mu\nu}\delta_{bd}\left(A^s_{\nu\alpha}A^s_{\mu\gamma}g^{\alpha\gamma}_{ac}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\mu\gamma}g^\gamma_c+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\nu\alpha}g^\alpha_a+N(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right) \\
&+\rho A^s_{\mu\nu}\delta_{ac}\left(A^s_{\nu\beta}A^s_{\mu\delta}g^{\beta\delta}_{bd}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\mu\delta}g^\delta_d+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\nu\beta}g^\beta_b+N(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right) \\
&+\rho A^s_{\mu\nu}\delta_{ad}\left(A^s_{\nu\beta}A^s_{\mu\gamma}g^{\beta\gamma}_{bc}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\mu\gamma}g^\gamma_c+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\nu\beta}g^\beta_b+N(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right) \\
&+A^s_{\mu\nu}A^s_{\mu\nu}\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right) \\
&-\rho^2\left(A^s_{\nu\alpha}A^s_{\nu\beta}g^{\alpha\beta}_{ab}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\nu\alpha}g^\alpha_a+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\nu A^s_{\nu\alpha}g^\alpha_b+N(\mathbf{m}\cdot\mathbf{A}^s)^\nu(\mathbf{m}\cdot\mathbf{A}^s)^\nu\right) \\
&\times\left(A^s_{\mu\gamma}A^s_{\mu\delta}g^{\gamma\delta}_{cd}+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\mu\gamma}g^\gamma_c+\sqrt{N}(\mathbf{m}\cdot\mathbf{A}^s)^\mu A^s_{\mu\gamma}g^\gamma_d+N(\mathbf{m}\cdot\mathbf{A}^s)^\mu(\mathbf{m}\cdot\mathbf{A}^s)^\mu\right)\Bigg\}.
\end{aligned}
\tag{B.7}
$$

We can evaluate (B.7) using

$$
\begin{aligned}
-i\frac{\partial\Omega}{\partial\Upsilon^\mu_a} &= g^\mu_a=\rho\sqrt{N}\sum_b\sum_\nu\left(\Lambda^{-1}\right)^{ab}_{\mu\nu}A^s_{\nu\eta}m^\eta \\
-\frac{\partial^2\Omega}{\partial\Upsilon^\mu_a\partial\Upsilon^\nu_b} &= g^{\mu\nu}_{ab}-g^\mu_a g^\nu_b=\left(\Lambda^{-1}\right)^{ab}_{\mu\nu} \\
i\frac{\partial^3\Omega}{\partial\Upsilon^\mu_a\partial\Upsilon^\nu_b\partial\Upsilon^\eta_c} &= g^{\mu\nu\eta}_{abc}-g^{\mu\nu}_{ab}g^\eta_c-g^{\mu\eta}_{ac}g^\nu_b-g^{\nu\eta}_{\beta\gamma}g^\mu_a+2g^\mu_a g^\nu_b g^\eta_c=0 \\
\frac{\partial^4\Omega}{\partial\Upsilon^\mu_a\partial\Upsilon^\nu_b\partial\Upsilon^\eta_c\partial\Upsilon^\delta_d} &= g^{\mu\nu\eta\delta}_{abcd}-g^{\mu\nu\eta}_{abc}g^\delta_d-g^{\mu\nu\delta}_{abd}g^\eta_c-g^{\mu\nu}_{ab}g^{\eta\delta}_{cd}+2g^{\mu\nu}_{ab}g^\eta_c g^\delta_d-g^{\mu\eta\delta}_{acd}g^\nu_b-g^{\mu\eta}_{ac}g^{\nu\delta}_{bd} \\
&\quad +2g^{\mu\eta}_{ac}g^\nu_b g^\delta_d-g^{\nu\eta\delta}_{bcd}g^\mu_a-g^{\mu\delta}_{ad}g^{\nu\eta}_{bc}+2g^{\nu\eta}_{bd}g^\mu_a g^\delta_d+2g^{\mu\delta}_{ad}g^\nu_b g^\eta_c+2g^\mu_a g^{\nu\delta}_{bd}g^\eta_c \\
&\quad +2g^\mu_a g^\nu_b g^{\eta\delta}_{cd}-6g^\mu_a g^\nu_b g^\eta_c g^\delta_d \\
&= 0 \\
\Rightarrow g^{\mu\nu\eta\delta}_{abcd} &= g^{\mu\nu}_{ab}g^{\eta\delta}_{cd}+g^{\mu\eta}_{ac}g^{\nu\delta}_{bd}+g^{\mu\delta}_{ad}g^{\nu\eta}_{bc}.
\end{aligned}
\tag{B.8}
$$

Since the replicon mode has the property $\sum_{a\neq b}\delta q_{ab} = 0$ [50] we need only consider terms containing exactly two delta function (in replica space). Therefore the only important terms in the above expression are

$$\frac{\rho^2}{4N}(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}) \times\sum_{\mu,\nu}\left\{\frac{\rho^2}{p}\left(\sum_{\alpha,\beta,\gamma,\delta}A^s_{\mu\alpha}A^s_{\mu\beta}A^s_{\nu\gamma}A^s_{\nu\delta}\hat{g}^{\alpha\beta}\hat{g}^{\gamma\delta}\right)+2\rho\sum_{\alpha,\beta}A^s_{\mu\nu}A^s_{\nu\alpha}A^s_{\mu\beta}\hat{g}^{\alpha\beta}+A^s_{\mu\nu}A^s_{\mu\nu}\right\}$$

where $\hat{g}^{\alpha\beta} = g^{\alpha\beta}_{aa}$.

Similarly we can show that the important terms in

$$\frac{\partial^2\Phi}{\partial\hat{q}_{ab}\partial\hat{q}_{cd}} = \frac{\partial}{\partial\hat{q}_{ab}}\frac{\left\langle\left\langle\sigma^c\sigma^d e^{\boldsymbol{\xi}\cdot\boldsymbol{\mu}\sum_a\sigma^a+\frac{1}{2}\sum_{a\neq b}(\lambda^2+2\delta\hat{q}_{ab})\sigma^a\sigma^b}\right\rangle_{\boldsymbol{\xi}}\right\rangle_{\boldsymbol{\sigma}}}{\left\langle\left\langle e^{\boldsymbol{\xi}\cdot\boldsymbol{\mu}\sum_a\sigma^a+\frac{1}{2}\sum_{a\neq b}(\lambda^2+2\delta\hat{q}_{ab})\sigma^a\sigma^b}\right\rangle_{\boldsymbol{\xi}}\right\rangle_{\boldsymbol{\sigma}}}$$

are

$$(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc})\left\langle\int D\hat{y}\cosh^{-4}(\boldsymbol{\xi}\cdot\boldsymbol{\mu}+\lambda\hat{y})\right\rangle_{\boldsymbol{\xi}}. \tag{B.9}$$

Appendix C

Coefficients in the A-T matrix for the generalised S-K spin-glass

Here we evaluate the coefficients in the AT matrix encountered in the calculation of the evolution equations for the joint spin-field distribution in the generalised S-K model. The determinant of this matrix determines the stability of the replica symmetric solutions to the extremisation of the saddle point exponent (6.9). We have to calculate the nine second derivatives of

$$\log\left[\int\left(\prod_a \frac{dH^a}{\sqrt{2\pi\det(\tilde{J}^2\mathbf{q})}}\right) e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a \hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}\right].$$

Due the symmetry under interchange of the order of differentiation, we need only calculate six quantities.

Since the replicon mode has $\sum_{a\neq b}\delta q_{ab} = \sum_{a\neq b}\delta\mathcal{G}_{ab} = \sum_{a\neq b}\delta\hat{\mathcal{E}}_{ab} = 0$ the only important terms are those containing two delta functions $\delta_{ac}\delta_{bd}$ or $\delta_{ad}\delta_{bc}$, which we evaluate in the RS ansatz using the same notation as Chapter 6.

Calculation of $\frac{\partial^2\Psi}{\partial\mathcal{G}_{ab}\partial\mathcal{G}_{cd}}$

The double derivative with respect to $\mathcal{G}$ can be trivially performed yielding

$$\begin{aligned}
&- \frac{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle \sigma^a\sigma^b\sigma^c\sigma^d\, e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}} \\
&+ \frac{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle \sigma^a\sigma^b\, e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{\int \mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}} \\
&\times \frac{\int \mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle \sigma^c\sigma^d\, e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{\int \mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})} \left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}.
\end{aligned}$$

The only relevant terms are the so called the replicon mode, i.e. those containing exactly two replica space delta functions. The relevant terms evaluated in the replica symmetric ansatz are

therefore

$$-(1-2q+\tau)(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc})$$

where

$$\tau = \left\langle\!\!\left\langle \langle\sigma\rangle_{\mathcal{M}}^4 \right\rangle\!\!\right\rangle_{wy}$$

so that the whole term is

$$-(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc})\left\langle\!\!\left\langle \left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)^2 \right\rangle\!\!\right\rangle_{wy}.$$

Hence the coefficient appearing in th AT matrix is

$$-2\left\langle\!\!\left\langle \left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)^2 \right\rangle\!\!\right\rangle_{wy}. \tag{C.1}$$

Calculation of $\frac{\partial^2\Psi}{\partial\mathcal{G}_{ab}\partial\hat{\mathcal{E}}_{cd}}$

The double derivative with respect to $\mathcal{G}$ and $\hat{\mathcal{E}}$ can be easily evaluated:

$$\begin{aligned}
&-(1-k^2)\frac{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\sigma^a\sigma^b\left(H^c-\tilde{J}_0m^c\right)\sigma^d e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}\cdots}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&+\frac{(1-k^2)\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\sigma^a\sigma^b e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&\times\frac{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\left(H^c-\tilde{J}_0m^c\right)\sigma^d e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}},
\end{aligned}$$

giving

$$-(1-k^2)\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)\left\langle\!\!\left\langle \left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)\right\rangle\!\!\right\rangle_{wy}.$$

Hence the term appearing in the AT matrix is

$$-2(1-k^2)\left\langle\!\!\left\langle \left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)\right\rangle\!\!\right\rangle_{wy}. \tag{C.2}$$

Calculation of $\frac{\partial^2\Psi}{\partial q_{ab}\partial q_{cd}}$

In order to calculate the derivatives with respect to $\mathbf{q}$, we need to go back and represent $e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m}-)(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}$ as an integral over $\hat{\mathbf{h}}$. We then have

$$\log\left\langle\int\left(\prod_a\frac{d\hat{h}^a\,dH^a}{2\pi}\right)e^{-\frac{1}{2}\hat{\mathbf{h}}(\tilde{J}^2\mathbf{q})\hat{\mathbf{h}}+i\hat{\mathbf{h}}(\mathbf{H}-\tilde{J}_0\mathbf{m})-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}.$$

Differentiating with respect to q_{ab}, q_{cd} then gives

$$\frac{\tilde{J}^4\int d\mathbf{H}\,\frac{\partial^4\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}}{\partial H^a\partial H^b\partial H^c\partial H^d}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{4\int d\mathbf{H}\, e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}$$

$$-\frac{\tilde{J}^4\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}}{\partial H^a\partial H^b}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}{4\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}$$
$$\times\frac{\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}}{\partial H^c\partial H^d}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}.$$

We also have a term like $\tilde{J}^2(1-k^2)\left(\hat{\mathcal{E}}_{ac}\hat{\mathcal{E}}_{bd}+\hat{\mathcal{E}}_{ad}\hat{\mathcal{E}}_{bc}\right)$. Again we are only interested in terms containing two delta functions, evaluating in the replica symmetric saddle points, using the fact that $(\mathbf{q})^{-1}_{ab}=\frac{\delta_{ab}}{1-q}-\frac{q}{(1-q)(1-q+nq)}$, this gives

$$\frac{1}{4}\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)\left[-4(1-k^2)\tilde{J}^2R_0^2+\frac{1}{(1-q)^2}\right.$$
$$\left.\left\langle\!\!\left\langle\left(1-\frac{1}{\tilde{J}^2(1-q)}\left[\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle^2_{\mathcal{M}}\right]\right)^2\right\rangle\!\!\right\rangle_{wy}\right].$$

Hence the term in the AT matrix is

$$\frac{1}{2(1-q)^2}\left[\left\langle\!\!\left\langle\left(1-\frac{1}{\tilde{J}^2(1-q)}\left[\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle^2_{\mathcal{M}}\right]\right)^2\right\rangle\!\!\right\rangle_{wy}\right.$$
$$\left.-4(1-k^2)(1-q)^2\tilde{J}^2R_0^2\right]. \tag{C.3}$$

Calculation of $\frac{\partial^2\Psi}{\partial q_{ab}\partial\mathcal{G}_{cd}}$

Again, first using the integral over $\hat{\mathbf{h}}$ we get

$$-\frac{i\frac{\tilde{J}^2}{2}\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}}{\partial H^a\partial H^b}\left\langle\sigma^c\sigma^d\,e^{\cdots-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}$$
$$+i\frac{\tilde{J}^2}{2}\frac{\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}}{\partial H^a\partial H^b}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}$$
$$\times\frac{\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle\sigma^c\sigma^d e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\,e^{-\frac{1}{2}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\left(\tilde{J}^2\mathbf{q}\right)^{-1}\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i\left(1-k^2\right)\left(\mathbf{H}-\tilde{J}_0\mathbf{m}\right)\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}\left(H^a\right)}\right\rangle_{\boldsymbol{\sigma}}},$$

giving

$$-i\frac{\tilde{J}^2}{2}\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)\left(\lambda_0^2+\lambda_1^2\right)^2\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)^2\right\rangle\!\!\right\rangle_{wy}.$$

We also have a term $\frac{i}{2}\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)$ from the $i\sum_{a,b}q_{ab}\mathcal{G}_{ab}$. Therefore the whole term is

$$\frac{i}{2}\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)\left(1-\frac{1}{\tilde{J}^2\left(1-q\right)^2}\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)^2\right\rangle\!\!\right\rangle_{wy}\right).$$

Hence the term appearing in the AT matrix is

$$i\left(1-\frac{1}{\tilde{J}^2\left(1-q\right)^2}\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)^2\right\rangle\!\!\right\rangle_{wy}\right). \tag{C.4}$$

Calculation of $\frac{\partial^2\Psi}{\partial q_{ab}\partial\hat{\mathcal{E}}_{cd}}$

The derivative $\frac{\partial^2\Psi}{\partial q\partial\hat{\mathcal{E}}}$ can similarly be carried out

$$
\begin{aligned}
&-i\frac{\tilde{J}^2(1-k^2)\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}}{\partial H^a\partial H^b}\left\langle\left(H^c-\tilde{J}_0m^c\right)\sigma^d\,e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}\cdots}\right\rangle_{\boldsymbol{\sigma}}}{2\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&+i\frac{\tilde{J}^2(1-k^2)\int d\mathbf{H}\,\frac{\partial^2\ e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}}{\partial H^a\partial H^b}\left\langle e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}{2\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&\times\frac{\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\left(H^c-\tilde{J}_0m^c\right)\sigma^d\,e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\cdots}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle e^{-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}},
\end{aligned}
$$

giving

$$
\begin{aligned}
&-\ i\frac{\tilde{J}^2}{2}(1-k^2)\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)\left(\lambda_0^2+\lambda_1^2\right)^2\\
&\quad\times\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)\left(\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}^2\right)\right\rangle\!\!\right\rangle_{wy}.
\end{aligned}
$$

We also have a term $2i\tilde{J}^2(1-k^2)(1-q)R_0\left(\delta_{ac}\delta_{bd}+\delta_{ad}\delta_{bc}\right)$ from the $\tilde{J}^2(1-k^2)\mathbf{q}\hat{\mathcal{E}}\mathbf{q}\hat{\mathcal{E}}$ term, so that the term appearing in the AT matrix is

$$
\begin{aligned}
i\tilde{J}^2(1-k^2)\Bigg\{&4(1-q)R_0-(\lambda_0^2+\lambda_1^2)^2\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)\right.\right.\\
&\left.\left.\times\left(\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}^2\right)\right\rangle\!\!\right\rangle_{wy}\Bigg\}. \qquad \text{(C.5)}
\end{aligned}
$$

Calculation of $\frac{\partial^2\Psi}{\partial\hat{\mathcal{E}}_{ab}\partial\hat{\mathcal{E}}_{cd}}$

The differential $\frac{\partial^2\Psi}{\partial\hat{\mathcal{E}}\partial\hat{\mathcal{E}}}$ is again similar

$$
\begin{aligned}
&-\ \frac{(1-k^2)^2\int d\mathbf{H}\,e^{\cdots}\left\langle\left(H^a-\tilde{J}_0m^a\right)\sigma^b\left(H^c-\tilde{J}_0m^c\right)\sigma^d\,e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\cdots}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\left\langle e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&+\frac{(1-k^2)^2\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\left(H^a-\tilde{J}_0m^a\right)\sigma^b\ e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}\cdots}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\left\langle e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}}\\
&\times\frac{\int d\mathbf{H}\,e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})}\left\langle\left(H^c-\tilde{J}_0m^c\right)\sigma^d\ e^{\cdots-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}\cdots}\right\rangle_{\boldsymbol{\sigma}}}{\int d\mathbf{H}\left\langle e^{-\frac{1}{2}(\mathbf{H}-\tilde{J}_0\mathbf{m})(\tilde{J}^2\mathbf{q})^{-1}(\mathbf{H}-\tilde{J}_0\mathbf{m})-i\hat{\mathbf{m}}\boldsymbol{\sigma}-i\boldsymbol{\sigma}\mathcal{G}\boldsymbol{\sigma}-i(1-k^2)(\mathbf{H}-\tilde{J}_0\mathbf{m})\hat{\mathcal{E}}\boldsymbol{\sigma}+\sum_a\hat{\rho}^{\sigma^a}(H^a)}\right\rangle_{\boldsymbol{\sigma}}},
\end{aligned}
$$

giving

$$
\begin{aligned}
&-(1-k^2)^2\delta_{ac}\delta_{bd}\left\langle\!\!\left\langle\left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)\left(\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}^2\right)\right\rangle\!\!\right\rangle_{wy}\\
&-(1-k^2)^2\delta_{ad}\delta_{bc}\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)^2\right\rangle\!\!\right\rangle_{wy}.
\end{aligned}
$$

We also have a term like $\tilde{J}^2(1-k^2)\left(q_{ac}q_{bd}+q_{ad}q_{bc}\right)$. Therefore the term appearing in the AT matrix is

$$(1-k^2)\left[2\tilde{J}^2(1-q)^2-(1-k^2)\left\langle\!\!\left\langle\left(1-\langle\sigma\rangle_{\mathcal{M}}^2\right)\left(\left\langle\left(H-\tilde{J}_0m\right)^2\right\rangle_{\mathcal{M}}-\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}^2\right)\right\rangle\!\!\right\rangle_{wy}\right.$$
$$\left.-(1-k^2)\left\langle\!\!\left\langle\left(\left\langle\sigma\left(H-\tilde{J}_0m\right)\right\rangle_{\mathcal{M}}-\langle\sigma\rangle_{\mathcal{M}}\left\langle H-\tilde{J}_0m\right\rangle_{\mathcal{M}}\right)^2\right\rangle\!\!\right\rangle_{wy}\right]. \qquad \text{(C.6)}$$

Appendix D

Code for solving joint field distribution evolution

```
/*************************************************************************/
/************************ Solve evolution of rho (h) ************************/
/*************************************************************************/
/********************* STEPHEN LAUGHTON 13/5/95 ********************/
/*************************** cc code.c -lm -O ***************************/
/*************************************************************************/
/********** integrals performed using Bode's and 6 pt Simpson rules **********/
/*************************************************************************/
```

10

```
#define PI 3.141592654
#define TWOoPI 0.636619772
#define TPI 6.283185307
#define INTERVAL (int)(3*5)                    /****** odd x*5 ******/
#define INTSQUARE (INTERVAL*INTERVAL)
#define PARTITION (int)(6*4)                   /**** even x*4 *******/
#define HMAX 4.5
#define HMIN -4.5
#define HMAX_MIN HMAX-(HMIN)
#define HM_MoP (HMAX_MIN)/PARTITION                                          20
#define K (double)0.0
#define OK (double)(1.0-K*K)
#define KO (double)((1.0-K*K)/(1.0+K*K))
#define NR_END 1
#define ACCURACY (double)0.000001
#define ISTEP (int)6                           /******* even ********/
#define GRAD (double)0.01
#define NSTEPS (int)100000

                                                                             30
#include <stdio.h>
#include <math.h>
#include "extra/extmath.h"
#include "extra/matrices.h"


void calcstuff(),calccoefficients_wy(),calcmmur(),calcrho(),check(),calct();
double saddle(),calcq();

double J_0,J_1,Js,T,beta;                                                    40
double chiplus[PARTITION+1],chiminus[PARTITION+1];
```

```
double rhoplus[PARTITION+1],rhominus[PARTITION+1];
double calcrhoplus[PARTITION+1],calcrhominus[PARTITION+1];
double rhoplussr[PARTITION+1],rhominussr[PARTITION+1];
double rhoplush[PARTITION+1],rhominush[PARTITION+1];
double tanhplus,tanhminus,tanhsigmarho,htanh,h_tanh;
double weights[]={0.311111111,1.422222222,0.5333333333,1.422222222,0.311111111};
double sweight[]={0.329861111,1.302083333,0.868055555,0.868055555,1.302083333,0.329861111};

double ***measureplus,***measureminus;                                    50
double ***f3tensor();

double m,mu,r,Theta_0,Theta_1,E,rterm,tterm,newr1,newt1,newmu;


main()                                                                    main
{
  FILE *fp;
  int iteration,k,l,n;
  fioat in;                                                               60
  double time,deltat,hJm,hmJm,hpJm,thbh,ex,test;
  double Phi_1,Phi_2plus,Phi_2minus,Phi_3plus[PARTITION+1];
  double Phi_3minus[PARTITION+1],Phi_4,interm[2*INTERVAL+2];
  double q,lambda0s,lambda1s;
  double deriv1plus,deriv1minus,deriv2plus,deriv2minus;
  double deltarhoplus[PARTITION+1],deltarhominus[PARTITION+1];
  char flowfile[12],distfile[12],chifile[12];

  measureplus=f3tensor(0,2*INTERVAL+1,0,2*INTERVAL+1,0,PARTITION);
  measureminus=f3tensor(0,2*INTERVAL+1,0,2*INTERVAL+1,0,PARTITION);       70

/************************** input parameters **************************/

  printf("Enter value of J_0 \n");
  scanf("%f",&in);
  J_0=(double)in;
  printf("J = 1 \n");
  J_1=1.0;            Js=J_1*J_1;
  printf("Enter value of T \n");                                          80
  scanf("%f",&in);
  T=(double)in;
  if(T!=0.0)
    beta=1.0/T;
  printf("Enter initial m \n");
  scanf("%f",&in);
  m=(double)in;

  sprintf(flowfile,"flow.%d.%d.%d",(int)(T),(int)(J_0),(int)(m*10.0));
  fp=fopen(flowfile,"w");                                                 90
  fprintf(fp,"INTERVAL %d PARTITION %d\n",INTERVAL,PARTITION);
  fprintf(fp,"ACCURACY %f GRAD %f NSTEPS %d\n",ACCURACY,GRAD,NSTEPS);
  fprintf(fp,"K = %f\n",(float)K);
  fprintf(fp,"J_0 = %f\n",(float)J_0);
  fprintf(fp,"T= %f\n",(float)T);
  fprintf(fp,"m initial = %f\n",(fioat)m);
  fprintf(fp,"t  m  E  dm/dt  dE/dt q AT\n");
  fclose(fp);

  sprintf(distfile,"dist.%d.%d.%d",(int)(T),(int)(J_0),(int)(m*10.0));    100
  fp=fopen(distfile,"w");
```

```
  fprintf(fp,"INTERVAL %d PARTITION %d\n",INTERVAL,PARTITION);
  fprintf(fp,"ACCURACY %f GRAD %f NSTEPS %d\n",ACCURACY,GRAD,NSTEPS);
  fprintf(fp,"K = %f\n",(float)K);
  fprintf(fp,"J_0 = %f\n",(float)J_0);
  fprintf(fp,"T= %f\n",(float)T);
  fprintf(fp,"m initial = %f\n",(float)m);
  fclose(fp);

  sprintf(chifile,"chi.%d.%d.%d",(int)(T),(int)(J_0),(int)(m*10.0));                    110

/*************************************************************************/


/************************* find initial chi ******************************/

  for(k=0;k<=PARTITION;k++) {
    chiplus[k]=0.5*(1.0+m);      chiminus[k]=0.5*(1.0-m);
    hJm=(HMIN+k*HM_MoP)-J_0*m;
    ex=exp(-hJm*hJm/(2.0*Js));                                                          120
    rhoplus[k]=chiplus[k]*ex/sqrt(TPI*Js);
    rhominus[k]=chiminus[k]*ex/sqrt(TPI*Js);
  }

  test=0.0;
  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      test+=weights[n]*(rhoplus[k]+rhominus[k]);
    }                                                                                   130
  }
  test*=HM_MoP;

  for(k=0;k<=PARTITION;k++) {
    rhoplus[k]/=test;
    rhominus[k]/=test;
  }
  q=m*m;    mu=0.0;    r=0.0;    Theta_0=0.0;    Theta_1=0.0;
  rterm=q; tterm=0.0;
  calcstuff(q);                                                                         140

/*************************************************************************/

  time=0.0;
  do {
    if(time!=0.0)  q=saddle(q);
    if((time-(int)time)<=0.001) {
      fp=fopen(distfile,"a");
      fprintf(fp,"t %f\n",time);
      fprintf(fp,"h rhoplus rhominus \n");                                              150
      for(k=0;k<=PARTITION;k++) {
        fprintf(fp,"%f %f %f\n",HMIN+k*HM_MoP,rhoplus[k],rhominus[k]);
      }
      fclose(fp);

      fp=fopen(chifile,"w");
      fprintf(fp,"t %f\n",time);
      fprintf(fp,"h chiplus chiminus \n");
      for(k=0;k<=PARTITION;k++) {
        fprintf(fp,"%f %f %f\n",HMIN+k*HM_MoP,chiplus[k],chiminus[k]);                  160
      }
```

```
    fprintf(fp,"tterm %f rterm %f q %f\n",tterm,rterm,q);
    fclose(fp);
  }

/************** evaluate evolution equation coefficients *****************/

  calccoefficients_wy(q,time,flowfile);

  lambda0s=(1.0-2*q)/(Js*(1.0-q)*(1.0-q));                                    170
  lambda1s=q/(Js*(1.0-q)*(1.0-q));


  Phi_1=1.0-Js*lambda0s*(tanhplus-tanhminus)-Js*lambda1s*tanhsigmarho;
  Phi_2plus=J_0*m-J_0*(tanhplus+tanhminus)-Js*KO*(lambda0s*htanh+lambda1s*h_tanh)
    +Js*2.0*KO*((Theta_0+Theta_1)*(tanhplus-tanhminus)-Theta_1*tanhsigmarho);
  Phi_2minus=J_0*m-J_0*(tanhplus+tanhminus)+Js*KO*(lambda0s*htanh+lambda1s*h_tanh)
    -Js*2.0*KO*((Theta_0+Theta_1)*(tanhplus-tanhminus)-Theta_1*tanhsigmarho);
  for(k=0;k<=PARTITION;k++) {
    Phi_3plus[k]=Js*KO*rhoplussr[k]*((lambda0s+2*lambda1s)*h_tanh-lambda1s*htanh)  180
      +Js*rhoplush[k]*((lambda0s+2*lambda1s)*tanhsigmarho-lambda1s*(tanhplus-tanhminus))
      -Js*2.0*KO*rhoplussr[k]*((Theta_0-Theta_1)*tanhsigmarho+Theta_1*(tanhplus-tanhminus));
    Phi_3minus[k]=Js*KO*rhominussr[k]*((lambda0s+2*lambda1s)*h_tanh-lambda1s*htanh)
      +Js*rhominush[k]*((lambda0s+2*lambda1s)*tanhsigmarho-lambda1s*(tanhplus-tanhminus))
      -Js*2.0*KO*rhominussr[k]*((Theta_0-Theta_1)*tanhsigmarho+Theta_1*(tanhplus-tanhminus));
  }
  Phi_4=Js*(1.0-tanhplus+tanhminus);

/************************** advance solution ************************/
                                                                              190
  deltat=HM_MoP*HM_MoP/(10.0*Phi_4);
  if(((int)time)!=((int)(time+deltat))) deltat=1.0*(int)(time+deltat)-time;
  time+=deltat;

  for(k=0;k<=PARTITION;k++) {
    hJm=HMIN+k*HM_MoP-J_0*m;
    hmJm=hJm-HM_MoP;
    hpJm=hJm+HM_MoP;

/***************************** find terms ******************************/    200

    if(k==0) {
      deriv1plus=(Phi_1*(hpJm*rhoplus[k+1]-hJm*rhoplus[k])
                  +Phi_2plus*(rhoplus[k+1]-rhoplus[k])
                  +(Phi_3plus[k+1]-Phi_3plus[k]))/(HM_MoP);
      deriv1minus=(Phi_1*(hpJm*rhominus[k+1]-hJm*rhominus[k])
                   +Phi_2minus*(rhominus[k+1]-rhominus[k])
                   +(Phi_3minus[k+1]-Phi_3minus[k]))/(HM_MoP);

      deriv2plus=Phi_4*(rhoplus[k+1]-2.0*rhoplus[k])/(HM_MoP*HM_MoP);           210
      deriv2minus=Phi_4*(rhominus[k+1]-2.0*rhominus[k])/(HM_MoP*HM_MoP);
    }

    if(k>0 && k <PARTITION) {
      deriv1plus=0.5*(Phi_1*(hpJm*rhoplus[k+1]-hmJm*rhoplus[k-1])
                      +Phi_2plus*(rhoplus[k+1]-rhoplus[k-1])
                      +(Phi_3plus[k+1]-Phi_3plus[k-1]))/(HM_MoP);
      deriv1minus=0.5*(Phi_1*(hpJm*rhominus[k+1]-hmJm*rhominus[k-1])
                       +Phi_2minus*(rhominus[k+1]-rhominus[k-1])
                       +(Phi_3minus[k+1]-Phi_3minus[k-1]))/(HM_MoP);         220
```

```
      deriv2plus=Phi_4*(rhoplus[k+1]-2.0*rhoplus[k]+rhoplus[k-1])/(HM_MoP*HM_MoP);
      deriv2minus=Phi_4*(rhominus[k+1]-2.0*rhominus[k]+rhominus[k-1])/(HM_MoP*HM_MoP);
    }

    if(k==PARTITION) {
      deriv1plus=(Phi_1*(hJm*rhoplus[k]-hmJm*rhoplus[k-1])
                  +Phi_2plus*(rhoplus[k]-rhoplus[k-1])
                  +(Phi_3plus[k]-Phi_3plus[k-1]))/(HM_MoP);
      deriv1minus=(Phi_1*(hJm*rhominus[k]-hmJm*rhominus[k-1])
                   +Phi_2minus*rhominus[k]-(rhominus[k-1])
                   +(Phi_3minus[k]-Phi_3minus[k-1]))/(HM_MoP);

      deriv2plus=Phi_4*(-2.0*rhoplus[k]+rhoplus[k-1])/(HM_MoP*HM_MoP);
      deriv2minus=Phi_4*(-2.0*rhominus[k]+rhominus[k-1])/(HM_MoP*HM_MoP);
    }

/*************************************************************************/

    if(T!=0.0)
      thbh=tanh(beta*(HMIN+k*HM_MoP));
    else {
      if(HMIN+k*HM_MoP>0.0) thbh=1.0;
      if(HMIN+k*HM_MoP<=0.0) thbh=-1.0;
      if(HMIN+k*HM_MoP==0.0) thbh=0.0;
    }
    deltarhoplus[k]=deltat*(0.5*((1.0+thbh)*rhominus[k]-(1.0-thbh)*rhoplus[k])
                            +deriv1plus+deriv2plus);
    deltarhominus[k]=deltat*(0.5*((1.0-thbh)*rhoplus[k]-(1.0+thbh)*rhominus[k])
                             +deriv1minus+deriv2minus);
  }

  for(k=0;k<=PARTITION;k++) {
    rhoplus[k]+=deltarhoplus[k];
    rhominus[k]+=deltarhominus[k];
    if(rhoplus[k]<0.0) rhoplus[k]=0.0;
    if(rhominus[k]<0.0) rhominus[k]=0.0;
  }

  test=0.0;
  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      test+=weights[n]*(rhoplus[k]+rhominus[k]);
    }
  }
  test*=HM_MoP;

  for(k=0;k<=PARTITION;k++) {
    rhoplus[k]/=test;
    rhominus[k]/=test;
  }

/*************************************************************************/

 }                                                    /**** sweeps loop ****/
 while(time<=10.0);

 q=saddle(q);
 fp=fopen(distfile,"a");
 fprintf(fp,"t %f\n",time);
```

```
  fprintf(fp,"h rhoplus rhominus \n");
  for(k=0;k<=PARTITION;k++) {
    fprintf(fp,"%f %f %f\n",HMIN+k*HM_MoP,rhoplus[k],rhominus[k]);
  }
  fclose(fp);
}                                                    /******* main *******/
/*************************************************************************/

                                                                          290

/*************************************************************************/
/****************************** subroutines ******************************/
/*************************************************************************/



/***************************** allocate 3tensor *************************/

double ***f3tensor(nrl,nrh,ncl,nch,ndl,ndh)                               300
     int nrl,nrh,ncl,nch,ndl,ndh;
{
  int i,j,nrow=nrh-nrl+1,ncol=nch-ncl+1,ndep=ndh-ndl+1;
  double ***t;

  t=(double ***) malloc((unsigned)((nrow+NR_END)*sizeof(double**)));
  if (!t) nrerror("allocation failure 1 in f3tensor()");
  t+=NR_END;      t-=nrl;

  t[nrl]=(double **) malloc((unsigned)((nrow*ncol+NR_END)*sizeof(double*)));  310
  if (!t[nrl]) nrerror("allocation failure 2 in f3tensor()");
  t[nrl]+=NR_END;      t[nrl]-=ncl;

  t[nrl][ncl]=(double *) malloc((unsigned)((nrow*ncol*ndep+NR_END)*sizeof(double)));
  if (!t[nrl][ncl]) nrerror("allocation failure 3 in f3tensor()");
  t[nrl][ncl]+=NR_END;      t[nrl][ncl]-=ndl;

  for(j=ncl+1;j<=nch;j++)     t[nrl][j]=t[nrl][j-1]+ndep;
  for(i=nrl+1;i<=nrh;i++) {
    t[i]=t[i-1]+ncol;          t[i][ncl]=t[i-1][ncl]+ncol*ndep;             320
    for(j=ncl+1;j<=nch;j++) t[i][j]=t[i][j-1]+ndep;
  }
  return t;
}
/*************************************************************************/


/***************************** calculate stuff **************************/

void calcstuff(q)                                                         330
     double q;
{
  int i,j,k,ip,ipIp1,ip1pIp1,jp,jpIp1,jp1pIp1,l,n;
  double hJm,istep,jstep,explus[4],exminus[4];
  double sqrtq,lambda_1,Tlambda01,thetaoverr,rtl,rtlsqrt;
  double denominatorplus[2*INTERVAL+2][2*INTERVAL+2];
  double denominatorminus[2*INTERVAL+2][2*INTERVAL+2];

  sqrtq=sqrt(q);
  Tlambda01=2*Js*(1.0-q);                                                 340
  lambda_1=sqrtq/(J_1*(1.0-q));
```

```
if(r!=0.0)
  thetaoverr=OK*Theta_1/r;
else
  thetaoverr=0.0;
rtl=lambda_1*lambda_1-thetaoverr*thetaoverr;
if(rtl>0.0) rtlsqrt=sqrt(lambda_1*lambda_1-thetaoverr*thetaoverr);
else rtlsqrt=lambda_1;

for(i=0;i<=INTERVAL;i++) {
  ipIp1=i+INTERVAL+1;
  istep=-1.0+2.0*(double)i/(double)INTERVAL;
  for(j=0;j<=INTERVAL;j++) {
    jpIp1=j+INTERVAL+1;
    jstep=-1.0+2.0*(double)j/(double)INTERVAL;

/************************* calculate denominator ************************/

      denominatorplus[i][j]=0.0;
      denominatorplus[ipIp1][j]=0.0;
      denominatorplus[i][jpIp1]=0.0;
      denominatorplus[ipIp1][jpIp1]=0.0;

      denominatorminus[i][j]=0.0;
      denominatorminus[ipIp1][j]=0.0;
      denominatorminus[i][jpIp1]=0.0;
      denominatorminus[ipIp1][jpIp1]=0.0;

      for(l=0;l<(PARTITION/4);l++) {
for(n=0;n<5;n++) {
  k=l*4+n;
  hJm=(HMIN+k*HM_MoP)-J_0*m;

  explus[0]=exp(-hJm*(hJm/Tlambda01-OK*Theta_0-thetaoverr*jstep)+r*jstep);
  explus[1]=exp(mu+hJm*rtlsqrt*istep);
  explus[2]=exp(-hJm*(hJm/Tlambda01-OK*Theta_0-thetaoverr/jstep)+r/jstep);
  explus[3]=exp(mu+hJm*rtlsqrt/istep);
  if(explus[0]>1.0E300) explus[0]=1.0E300;
  if(explus[1]>1.0E300) explus[1]=1.0E300;
  if(explus[2]>1.0E300) explus[2]=1.0E300;
  if(explus[3]>1.0E300) explus[3]=1.0E300;

  exminus[0]=exp(-hJm*(hJm/Tlambda01+OK*Theta_0-thetaoverr*jstep)-r*jstep);
  exminus[1]=exp(-mu+hJm*rtlsqrt*istep);
  exminus[2]=exp(-hJm*(hJm/Tlambda01+OK*Theta_0-thetaoverr/jstep)-r/jstep);
  exminus[3]=exp(-mu+hJm*rtlsqrt/istep);
  if(exminus[0]>1.0E300) exminus[0]=1.0E300;
  if(exminus[1]>1.0E300) exminus[1]=1.0E300;
  if(exminus[2]>1.0E300) exminus[2]=1.0E300;
  if(exminus[3]>1.0E300) exminus[3]=1.0E300;

  denominatorplus[i][j]+=weights[n]*chiplus[k]*explus[0]*explus[1];
  denominatorplus[ipIp1][j]+=weights[n]*chiplus[k]*explus[0]*explus[3];
  denominatorplus[i][jpIp1]+=weights[n]*chiplus[k]*explus[2]*explus[1];
  denominatorplus[ipIp1][jpIp1]+=weights[n]*chiplus[k]*explus[2]*explus[3];

  denominatorminus[i][j]+=weights[n]*chiminus[k]*exminus[0]*exminus[1];
  denominatorminus[ipIp1][j]+=weights[n]*chiminus[k]*exminus[0]*exminus[3];
  denominatorminus[i][jpIp1]+=weights[n]*chiminus[k]*exminus[2]*exminus[1];
  denominatorminus[ipIp1][jpIp1]+=weights[n]*chiminus[k]*exminus[2]*exminus[3];
}
```

```
    }                                                    /****** l loop ******/

    denominatorplus[i][j]*=HM_MoP;
    denominatorplus[ipIp1][j]*=HM_MoP;
    denominatorplus[i][jpIp1]*=HM_MoP;
    denominatorplus[ipIp1][jpIp1]*=HM_MoP;

    denominatorminus[i][j]*=HM_MoP;
    denominatorminus[ipIp1][j]*=HM_MoP;                                                    410
    denominatorminus[i][jpIp1]*=HM_MoP;
    denominatorminus[ipIp1][jpIp1]*=HM_MoP;

/************************** calculate measure ****************************/

    for(k=0;k<=PARTITION;k++) {
      hJm=(HMIN+k*HM_MoP)-J_0*m;

      explus[0]=exp(-hJm*(hJm/Tlambda01-OK*Theta_0-thetaoverr*jstep)+r*jstep);
      explus[1]=exp(mu+hJm*rtlsqrt*istep);                                                 420
      explus[2]=exp(-hJm*(hJm/Tlambda01-OK*Theta_0-thetaoverr/jstep)+r/jstep);
      explus[3]=exp(mu+hJm*rtlsqrt/istep);
      if(explus[0]>1.0E300) explus[0]=1.0E300;
      if(explus[1]>1.0E300) explus[1]=1.0E300;
      if(explus[2]>1.0E300) explus[2]=1.0E300;
      if(explus[3]>1.0E300) explus[3]=1.0E300;

      exminus[0]=exp(-hJm*(hJm/Tlambda01+OK*Theta_0-thetaoverr*jstep)-r*jstep);
      exminus[1]=exp(-mu+hJm*rtlsqrt*istep);
      exminus[2]=exp(-hJm*(hJm/Tlambda01+OK*Theta_0-thetaoverr/jstep)-r/jstep);             430
      exminus[3]=exp(-mu+hJm*rtlsqrt/istep);
      if(exminus[0]>1.0E300) exminus[0]=1.0E300;
      if(exminus[1]>1.0E300) exminus[1]=1.0E300;
      if(exminus[2]>1.0E300) exminus[2]=1.0E300;
      if(exminus[3]>1.0E300) exminus[3]=1.0E300;

      measureplus[i][j][k]=chiplus[k]*explus[0]*explus[1]/
        (denominatorplus[i][j]+denominatorminus[i][j]);
      measureplus[ipIp1][j][k]=chiplus[k]*explus[0]*explus[3]/
        (denominatorplus[ipIp1][j]+denominatorminus[ipIp1][j]);                            440
      measureplus[i][jpIp1][k]=chiplus[k]*explus[2]*explus[1]/
        (denominatorplus[i][jpIp1]+denominatorminus[i][jpIp1]);
      measureplus[ipIp1][jpIp1][k]=chiplus[k]*explus[2]*explus[3]/
        (denominatorplus[ipIp1][jpIp1]+denominatorminus[ipIp1][jpIp1]);

      measureminus[i][j][k]=chiminus[k]*exminus[0]*exminus[1]/
        (denominatorplus[i][j]+denominatorminus[i][j]);
      measureminus[ipIp1][j][k]=chiminus[k]*exminus[0]*exminus[3]/
        (denominatorplus[ipIp1][j]+denominatorminus[ipIp1][j]);
      measureminus[i][jpIp1][k]=chiminus[k]*exminus[2]*exminus[1]/                         450
        (denominatorplus[i][jpIp1]+denominatorminus[i][jpIp1]);
      measureminus[ipIp1][jpIp1][k]=chiminus[k]*exminus[2]*exminus[3]/
        (denominatorplus[ipIp1][jpIp1]+denominatorminus[ipIp1][jpIp1]);
    }                                                    /****** k loop ******/
  }                                                      /****** j loop ******/
}                                                        /****** i loop ******/

  check();
  calcrho();
}                                                                                      460
/************************************************************************/
```

```
/************************* check \int \sum rho = 1 *********************/

void check()
{
  int i,j,k,ipIp1,jpIp1,l,n,g,h;
  double w,y,ww,yy,ex1,ex3;
  double checkplus[2*INTERVAL+2][2*INTERVAL+2];                                    470
  double checkminus[2*INTERVAL+2][2*INTERVAL+2];
  double checkintermp[2*INTERVAL+2],checkintermm[2*INTERVAL+2];
  double chk,chkp,chkm;


  for(i=0;i<=INTERVAL;i++) {
    ipIp1=i+INTERVAL+1;
    for(j=0;j<=INTERVAL;j++) {
      jpIp1=j+INTERVAL+1;
                                                                                   480
      checkplus[i][j]=0.0;              checkplus[ipIp1][j]=0.0;
      checkplus[i][jpIp1]=0.0;          checkplus[ipIp1][jpIp1]=0.0;
      checkminus[i][j]=0.0;             checkminus[ipIp1][j]=0.0;
      checkminus[i][jpIp1]=0.0;         checkminus[ipIp1][jpIp1]=0.0;

      for(l=0;l<(PARTITION/4);l++) {
        for(n=0;n<5;n++) {
          k=l*4+n;
          checkplus[i][j]+=weights[n]*measureplus[i][j][k];                        490
          checkplus[ipIp1][j]+=weights[n]*measureplus[ipIp1][j][k];
          checkplus[i][jpIp1]+=weights[n]*measureplus[i][jpIp1][k];
          checkplus[ipIp1][jpIp1]+=weights[n]*measureplus[ipIp1][jpIp1][k];

          checkminus[i][j]+=weights[n]*measureminus[i][j][k];
          checkminus[ipIp1][j]+=weights[n]*measureminus[ipIp1][j][k];
          checkminus[i][jpIp1]+=weights[n]*measureminus[i][jpIp1][k];
          checkminus[ipIp1][jpIp1]+=weights[n]*measureminus[ipIp1][jpIp1][k];
        }
      }                                                                            500
      checkplus[i][j]*=HM_MoP;          checkplus[ipIp1][j]*=HM_MoP;
      checkplus[i][jpIp1]*=HM_MoP;      checkplus[ipIp1][jpIp1]*=HM_MoP;
      checkminus[i][j]*=HM_MoP;         checkminus[ipIp1][j]*=HM_MoP;
      checkminus[i][jpIp1]*=HM_MoP;     checkminus[ipIp1][jpIp1]*=HM_MoP;
    }
  }

  for(i=0;i<=INTERVAL;i++) {
    ipIp1=i+INTERVAL+1;
    checkintermp[i]=0.0;                checkintermp[ipIp1]=0.0;                 510
    checkintermm[i]=0.0;                checkintermm[ipIp1]=0.0;

    for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
      for(h=0;h<ISTEP;h++) {
        j=g*(ISTEP-1)+h;
        jpIp1=j+INTERVAL+1;
        w=-1.0+2.0*(double)j/(double)INTERVAL;      ww=w*w;
        ex1=exp(-0.5*ww);                           ex3=exp(-0.5/ww)/ww;

        checkintermp[i]+=sweight[h]*(ex1*checkplus[i][j]+ex3*checkplus[i][jpIp1]);  520
        checkintermp[ipIp1]+=sweight[h]*(ex1*checkplus[ipIp1][j]+ex3*checkplus[ipIp1][jpIp1]);
```

```
        checkintermm[i]+=sweight[h]*(ex1*checkminus[i][j]+ex3*checkminus[i][jpIp1]);
        checkintermm[ipIp1]+=sweight[h]*(ex1*checkminus[ipIp1][j]+ex3*checkminus[ipIp1][jpIp1]);
      }
    }
  }

  chkp=0.0;
  chkm=0.0;                                                                    530
  for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
    for(h=0;h<ISTEP;h++) {
      i=g*(ISTEP-1)+h;
      ipIp1=i+INTERVAL+1;
      y=-1.0+2.0*(double)i/(double)INTERVAL;        yy=y*y;
      ex1=exp(-0.5*yy);                             ex3=exp(-0.5/yy)/yy;

      chkp+=sweight[h]*(ex1*checkintermp[i]+ex3*checkintermp[ipIp1]);
      chkm+=sweight[h]*(ex1*checkintermm[i]+ex3*checkintermm[ipIp1]);
    }                                                                          540
  }
  chkp*=TWOoPI/INTSQUARE;
  chkm*=TWOoPI/INTSQUARE;
  chk=chkp+chkm;

  if(fabs(chk-1.0)>0.01)
    printf("check error one %f plus %f minus %f\n",chk,chkp,chkm);

  for(i=0;i<=INTERVAL;i++) {                                                   550
    ipIp1=i+INTERVAL+1;
    for(j=0;j<=INTERVAL;j++) {
      jpIp1=j+INTERVAL+1;
      for(k=0;k<=PARTITION;k++) {
        measureplus[i][j][k]/=chk;          measureminus[i][j][k]/=chk;
        measureplus[i][jpIp1][k]/=chk;      measureminus[i][jpIp1][k]/=chk;
        measureplus[ipIp1][j][k]/=chk;      measureminus[ipIp1][j][k]/=chk;
        measureplus[ipIp1][jpIp1][k]/=chk; measureminus[ipIp1][jpIp1][k]/=chk;
      }
    }
  }                                                                            560
}
/*************************************************************************/


/********************* solve saddle point equations ************************/

double saddle(q)
     double q;
{
  int k,l,n,ti,i;                                                              570
  double oldq,dq,rp,hJm;
  double step;

  m=0.0;        newmu=0.0;        newr1=0.0;        newt1=0.0;

  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      m+=weights[n]*(rhoplus[k]-rhominus[k]);
    }                                                                          580
  }
```

```
m*=HM_MoP;

for(l=0;l<(PARTITION/4);l++) {
  for(n=0;n<5;n++) {
    k=l*4+n;
    hJm=(HMIN+k*HM_MoP)-J_0*m;
    newmu+=weights[n]*hJm*(rhoplus[k]+rhominus[k]);
    newr1+=weights[n]*hJm*hJm*(rhoplus[k]+rhominus[k]);
    newt1+=weights[n]*hJm*(rhoplus[k]-rhominus[k]);
  }
}
newmu*=HM_MoP;          newr1*=HM_MoP;          newt1*=HM_MoP;

calcmmur(q);

for(step=3.0;step>=1.0;step-=1.0) {
  ti=0;
  do {
    oldq=q;

    calcstuff(q);  calct();
    rp=0.0;
    for(k=0;k<=PARTITION;k++) {
      chiplus[k]-=step*GRAD*(calcrhoplus[k]-rhoplus[k])/chiplus[k];
      chiminus[k]-=step*GRAD*(calcrhominus[k]-rhominus[k])/chiminus[k];

      rp+=(rhoplus[k]-calcrhoplus[k])*(rhoplus[k]-calcrhoplus[k]);
      rp+=(rhominus[k]-calcrhominus[k])*(rhominus[k]-calcrhominus[k]);
    }
    rp/=2.0*(PARTITION+1.0);
    rp=sqrt(rp);

    calcmmur(q);    calcstuff(q);    calct();    q=calcq(q);
    calcmmur(q);

    dq=(q-oldq);
    ti++;
   }
   while(fabs(dq)>step*step*ACCURACY || rp>0.0003 && ti<=NSTEPS);
  }

  return(q);
}
/*************************************************************************/


/********************** calculate saddle point equations ********************/

void calcmmur(q)
     double q;
{
  double omq,omqs,rsquared,oldr;

  oldr=r;
  omq=1.0-q;
  omqs=Js*(1.0-q)*(1.0-q);

  mu=J_0/(Js*omq)*newmu;
  Theta_0=1.0/(2.0*omqs)*(newt1-tterm);
  mu-=OK*J_0*m*Theta_0;
```

```
  Theta_1=-1.0/(2.0*omqs)*(2.0*Js*(1.0-q*q)*Theta_0-newt1);
  rsquared=Js*q+(-2.0*q*newr1+(1.0+q)*rterm)/omq
    -2.0*Js*OK*Theta_0*(-3.0*q*newt1+(2.0+q)*tterm);

  if(rsquared>=0.0) r=sqrt(rsquared)/(J_1*omq);
  else r=oldr;
}
/*************************************************************************/

/*************************** calculate q ********************************/

double calcq(q)
     double q;
{
  int k,l,n;

  q=0.0;
  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      q+=weights[n]*(rhoplussr[k]-rhominussr[k]);
    }
  }
  q*=HM_MoP;

  return(q);
}
/*************************************************************************/

/************************** evaluate coefficients ************************/

void calccoefficients_wy(q,time,flowfile)
     double q,time;
     char flowfile[10];
{
  FILE *fp;
  int l,n,k;
  double hJm,th,at;

  calcstuff(q);

  tanhplus=0.0;
  tanhminus=0.0;
  tanhsigmarho=0.0;
  htanh=0.0;
  h_tanh=0.0;
  E=0.0;

  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      hJm=(HMIN+k*HM_MoP)-J_0*m;

      if(T!=0.0) th=tanh(beta*(HMIN+k*HM_MoP));
      else {
        if(HMIN+k*HM_MoP>0.0)   th=1.0;
        if(HMIN+k*HM_MoP<0.0)   th=-1.0;
        if(HMIN+k*HM_MoP==0)    th=0.0;
```

```
      }

      tanhplus+=weights[n]*th*rhoplus[k];
      tanhminus+=weights[n]*th*rhominus[k];
      tanhsigmarho+=weights[n]*th*(rhoplussr[k]+rhominussr[k]);
      htanh+=weights[n]*hJm*th*(rhoplus[k]+rhominus[k]);
      h_tanh+=weights[n]*th*(rhoplush[k]+rhominush[k]);
      E+=weights[n]*(HMIN+k*HM_MoP)*(rhoplus[k]-rhominus[k]);
    }                                                                    710
  }

  tanhplus*=HM_MoP;
  tanhminus*=HM_MoP;
  tanhsigmarho*=HM_MoP;
  htanh*=HM_MoP;
  h_tanh*=HM_MoP;
  E*=-0.5*HM_MoP;

/************************** EXPORT  dm/dt dE/dt ************************/    720

  fp=fopen(flowfile,"a");
  fprintf(fp,"%f %f %f %f %f %f\n",(float)time,(float)m,(float)E,(float)(tanhplus+tanhminus-m),
          (float)(-htanh-J_0*m*(tanhplus+tanhminus)-2.0*E),q);
  fclose(fp);
}
/*********************************************************************/

/************************** calculate rho ****************************/    730

void calcrho()
{
  int i,j,k,kprime,l,n,ipIp1,jpIp1,g,h;
  double y,yy,w,ww,ex1,ex3,hJm,test,oldrterm,oldtterm;
  double intermplus[2*INTERVAL+2],intermminus[2*INTERVAL+2];
  double intermpp[2*INTERVAL+2],intermpm[2*INTERVAL+2];
  double intermmp[2*INTERVAL+2],intermmm[2*INTERVAL+2];
  double rhopp[PARTITION+1][PARTITION+1],rhopm[PARTITION+1][PARTITION+1];
  double rhomp[PARTITION+1][PARTITION+1],rhomm[PARTITION+1][PARTITION+1];  740

  for(k=0;k<=PARTITION;k++) {

/************************** calculate rhorho ************************/

    for(kprime=0;kprime<=PARTITION;kprime++) {
      rhopp[k][kprime]=0.0;                    rhopm[k][kprime]=0.0;
      rhomp[k][kprime]=0.0;                    rhomm[k][kprime]=0.0;

      for(i=0;i<=INTERVAL;i++) {                                         750
        ipIp1=i+INTERVAL+1;
        intermpp[i]=0.0;                       intermpp[ipIp1]=0.0;
        intermpm[i]=0.0;                       intermpm[ipIp1]=0.0;
        intermmp[i]=0.0;                       intermmp[ipIp1]=0.0;
        intermmm[i]=0.0;                       intermmm[ipIp1]=0.0;

        for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
          for(h=0;h<ISTEP;h++) {
            j=g*(ISTEP-1)+h;
            jpIp1=j+INTERVAL+1;                                          760
            w=-1.0+2.0*(double)j/(double)INTERVAL; ww=w*w;
```

```
        ex1=exp(-0.5*ww);                    ex3=exp(-0.5/ww)/ww;

        intermpp[i]+=sweight[hi]*(ex1*measureplus[i][j][k]*measureplus[i][j][kprime]
          +ex3*measureplus[i][jpIp1][k]*measureplus[i][jpIp1][kprime]);
        intermpp[ipIp1]+=sweight[hi]*(ex1*measureplus[ipIp1][j][k]*measureplus[ipIp1][j][kprime]
          +ex3*measureplus[ipIp1][jpIp1][k]*measureplus[ipIp1][jpIp1][kprime]);

        intermpm[i]+=sweight[hi]*(ex1*measureplus[i][j][k]*measureminus[i][j][kprime]
          +ex3*measureplus[i][jpIp1][k]*measureminus[i][jpIp1][kprime]);
        intermpm[ipIp1]+=sweight[hi]*(ex1*measureplus[ipIp1][j][k]*measureminus[ipIp1][j][kprime]
          +ex3*measureplus[ipIp1][jpIp1][k]*measureminus[ipIp1][jpIp1][kprime]);

        intermmp[i]+=sweight[hi]*(ex1*measureminus[i][j][k]*measureplus[i][j][kprime]
          +ex3*measureminus[i][jpIp1][k]*measureplus[i][jpIp1][kprime]);
        intermmp[ipIp1]+=sweight[hi]*(ex1*measureminus[ipIp1][j][k]*measureplus[ipIp1][j][kprime]
          +ex3*measureminus[ipIp1][jpIp1][k]*measureplus[ipIp1][jpIp1][kprime]);

        intermmm[i]+=sweight[hi]*(ex1*measureminus[i][j][k]*measureminus[i][j][kprime]
          +ex3*measureminus[i][jpIp1][k]*measureminus[i][jpIp1][kprime]);
        intermmm[ipIp1]+=sweight[hi]*(ex1*measureminus[ipIp1][j][k]*measureminus[ipIp1][j][kprime]
          +ex3*measureminus[ipIp1][jpIp1][k]*measureminus[ipIp1][jpIp1][kprime]);
      }
    }
  }

  for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
    for(h=0;h<ISTEP;h++) {
      i=g*(ISTEP-1)+h;
      ipIp1=i+INTERVAL+1;
      y=-1.0+2.0*(double)i/(double)INTERVAL;      yy=y*y;
      ex1=exp(-0.5*yy);                           ex3=exp(-0.5/yy)/yy;

      rhopp[k][kprime]+=sweight[hi]*(ex1*intermpp[i]+ex3*intermpp[ipIp1]);
      rhopm[k][kprime]+=sweight[hi]*(ex1*intermpm[i]+ex3*intermpm[ipIp1]);
      rhomp[k][kprime]+=sweight[hi]*(ex1*intermmp[i]+ex3*intermmp[ipIp1]);
      rhomm[k][kprime]+=sweight[hi]*(ex1*intermmm[i]+ex3*intermmm[ipIp1]);
    }
  }
  rhopp[k][kprime]*=TWOoPI/INTSQUARE;
  rhopm[k][kprime]*=TWOoPI/INTSQUARE;
  rhomp[k][kprime]*=TWOoPI/INTSQUARE;
  rhomm[k][kprime]*=TWOoPI/INTSQUARE;
}                                              /*** kprime loop ****/

/************ calculate rhoplus-rhominus and h*(rhoplus+rhominus) ********/

  rhoplussr[k]=0.0;                rhominussr[k]=0.0;
  rhoplush[k]=0.0;                 rhominush[k]=0.0;

  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      kprime=l*4+n;
      hJm=(HMIN+kprime*HM_MoP)-J_0*m;

      rhoplussr[k]+=weights[n]*(rhopp[k][kprime]-rhopm[k][kprime]);
      rhominussr[k]+=weights[n]*(rhomp[k][kprime]-rhomm[k][kprime]);

      rhoplush[k]+=weights[n]*hJm*(rhopp[k][kprime]+rhopm[k][kprime]);
      rhominush[k]+=weights[n]*hJm*(rhomp[k][kprime]+rhomm[k][kprime]);
    }
```

```
    }
    rhoplussr[k]*=HM_MoP;              rhominussr[k]*=HM_MoP;
    rhoplush[k]*=HM_MoP;               rhominush[k]*=HM_MoP;


/************************ calculate rhoplus and rhominus ****************/

    calcrhoplus[k]=0.0;                    calcrhominus[k]=0.0;
                                                                                    830
    for(i=0;i<=INTERVAL;i++) {
      ipIp1=i+INTERVAL+1;
      intermplus[i]=0.0;                intermplus[ipIp1]=0.0;
      intermminus[i]=0.0;               intermminus[ipIp1]=0.0;

      for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
        for(h=0;h<ISTEP;h++) {
          j=g*(ISTEP-1)+h;
          jpIp1=j+INTERVAL+1;
          w=-1.0+2.0*(double)j/(double)INTERVAL;      ww=w*w;                       840
          ex1=exp(-0.5*ww);                           ex3=exp(-0.5/ww)/ww;

          intermplus[i]+=sweight[hi]*(ex1*measureplus[i][j][k]+ex3*measureplus[i][jpIp1][k]);
          intermplus[ipIp1]+=sweight[hi]*(ex1*measureplus[ipIp1][j][k]+ex3*measureplus[ipIp1][jpIp1][k]);

          intermminus[i]+=sweight[hi]*(ex1*measureminus[i][j][k]+ex3*measureminus[i][jpIp1][k]);
          intermminus[ipIp1]+=sweight[hi]*(ex1*measureminus[ipIp1][j][k]+ex3*measureminus[ipIp1][jpIp1][k]);
        }
      }
    }                                                                               850

      for(g=0;g<(INTERVAL/(ISTEP-1));g++) {
        for(h=0;h<ISTEP;h++) {
          i=g*(ISTEP-1)+h;
          ipIp1=i+INTERVAL+1;
          y=-1.0+2.0*(double)i/(double)INTERVAL;      yy=y*y;
          ex1=exp(-0.5*yy);                           ex3=exp(-0.5/yy)/yy;

          calcrhoplus[k]+=sweight[hi]*(ex1*intermplus[i]+ex3*intermplus[ipIp1]);
          calcrhominus[k]+=sweight[hi]*(ex1*intermminus[i]+ex3*intermminus[ipIp1]);  860
        }
      }
      calcrhoplus[k]*=TWOoPI/INTSQUARE;
      calcrhominus[k]*=TWOoPI/INTSQUARE;
    }                                                  /****** k loop ******/

    test=0.0;
    for(l=0;l<(PARTITION/4);l++) {
      for(n=0;n<5;n++) {
        k=l*4+n;                                                                    870
        test+=weights[n]*(calcrhoplus[k]+calcrhominus[k]);
      }
    }
    test*=HM_MoP;

    if(fabs(test-1.0)>0.05) printf("calcrho test error %f\n",test);

    for(k=0;k<=PARTITION;k++) {
      calcrhoplus[k]/=test;
      calcrhominus[k]/=test;                                                        880
    }
```

```
}
/*************************************************************************/


/************* calculate non-trivial terms in saddle point equations ***********/

void calct()
{
  int l,n,k;                                                              890
  double hJm;

  rterm=0.0;                    tterm=0.0;

  for(l=0;l<(PARTITION/4);l++) {
    for(n=0;n<5;n++) {
      k=l*4+n;
      hJm=(HMIN+k*HM_MoP)-J_0*m;
      rterm+=weights[n]*hJm*(rhoplush[k]+rhominush[k]);
      tterm+=weights[n]*(rhoplush[k]-rhominush[k]);                       900
    }
  }
  rterm*=HM_MoP;
  tterm*=HM_MoP;
}
/*************************************************************************/
```

Appendix E

List of Symbols

$\ln[x]$	$0 \leq x \leq \infty$	natural logarithm of x
$\lim_{x \to y} f(x)$		limiting value of the function $f(x)$ as x approaches y
N	$\to \infty$	number of spins (neurons)
i, j, k	$\in \{1, \ldots, N\}$	spin (neuron) labels
n	$\to 0$	number of replicas
a, b, c, d	$\in \{1, \ldots, n\}$	replica labels
σ_i^a	$\in \{-1, +1\}$	i^{th} spin (neuron) in a^{th} replica
J_{ij}	$\in \Re$	exchange interaction (synaptic efficacy) between i^{th} and j^{th} spins (neurons)
$P(J_{ij})$	$\in [0, 1]$	probability distribution of J_{ij}'s
$\langle f(J_{ij}) \rangle_{J_{ij}}$	$=$	$\int dJ_{ij}\, P(J_{ij})\, f(J_{ij})$ - average of f over $P(J_{ij})$
p	$\in \Re^+$	number of patterns embedded in neural network
α	$= \frac{p}{N} \geq 0$	- number of patterns per spin embedded in neural network
μ, ν	$\in \{1, \ldots, p\}$	pattern labels
ξ_i^μ	$\in \{-1, +1\}$	i^{th} bit of μ^{th} pattern embedded in a neural network
$\langle f(\boldsymbol{\xi}) \rangle_{\boldsymbol{\xi}}$	$=$	$\frac{1}{2^p} \sum_{\boldsymbol{\xi} \in \{-1,+1\}^p} f(\boldsymbol{\xi})$ - average of f over $\boldsymbol{\xi}$
β	$=$	$\frac{1}{T}$ - inverse temperature (measure of stochasticity)
$\mathcal{H}(\boldsymbol{\sigma})$	$\in \Re$	Hamiltonian for microstates $\boldsymbol{\sigma}$
h_i	$=$	$\sum_{i \neq j} J_{ij} \sigma_j$ - local magnetic field (post synaptic potential) experienced by i^{th} spin (neuron)
θ_i	$\in \Re$	external field (threshold potential) of i^{th} spin (neuron)
$p_t(\boldsymbol{\sigma})$	$\in [0, 1]$	probability distribution of microstates $\boldsymbol{\sigma}$
$W(\boldsymbol{\sigma}' \to \boldsymbol{\sigma})$	$\in [0, 1]$	transition probability per unit time for the process $\boldsymbol{\sigma}' \to \boldsymbol{\sigma}$
F_i		spin flip operator

		$F_i\Phi(\sigma_1,\ldots,\sigma_i,\ldots,\sigma_N) = \Phi(\sigma_1,\ldots,-\sigma_i,\ldots,\sigma_N)$
$w_i(\boldsymbol{\sigma})$	$\in [0,1]$	transition rate for the process $\boldsymbol{\sigma} \to F_i\boldsymbol{\sigma}$
Z	$=$	$\sum_{\boldsymbol{\sigma}} e^{-\beta\mathcal{H}(\boldsymbol{\sigma})}$ - canonical partition function
f	$=$	$-\frac{1}{\beta N}\ln[Z]$ - free energy per spin (neuron)
m^a	$=$	$\frac{1}{N}\sum_i \sigma_i^a$ - magnetisation in a^{th} replica
$\hat{m}^a$		conjugate parameter of m_a
m_a^μ	$=$	$\frac{1}{N}\sum_i \xi_i^\mu \sigma_i^a$ - overlap with μ^{th} pattern in a^{th} replica
q_{ab}	$=$	$\frac{1}{N}\sum_i \sigma_i^a \sigma_i^b$ - spin glass order parameter
$\hat{q}_{ab}$		conjugate parameter to q_{ab}
$\boldsymbol{\Omega}(\boldsymbol{\sigma})$	$=$	$(\Omega^1,\ldots,\Omega^c)$ general set of order parameters
$\mathcal{P}_t(\boldsymbol{\Omega})$	$\in [0,1]$	macroscopic probability distribution
$\langle f(\boldsymbol{\sigma})\rangle_{\boldsymbol{\Omega},t}$	$=$	$\frac{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\,\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))\,f(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma})\,\delta(\boldsymbol{\Omega}-\boldsymbol{\Omega}(\boldsymbol{\sigma}))}$ - sub-shell average over states corresponding to macroscopic state $\boldsymbol{\Omega}$
$\Delta_j^\mu(\sigma)$	$=$	$\Omega^\mu(F_j\boldsymbol{\sigma}) - \Omega^\mu(\boldsymbol{\sigma})$ - change in μ^{th} order parameter by flipping j^{th} spin
$\tilde{J}_0$	$\in \Re$	rescaled ferromagnetic interaction in S-K model
$\tilde{J}$	$\in \Re$	rescaled disordered interaction in S-K model
k	$\in \Re$	degree of asymmetry in the generalised S-K model
$A_{\mu\nu}$		embedding matrix in generalised Hopfield model
$\alpha r(\boldsymbol{\sigma})$	$=$	$\sum_{[\mu,\nu>c \mid \mu>c,\ \nu\leq c \mid \mu\leq c,\ \nu>c]} \left[\frac{1}{N}\sum_j \xi_j^\mu\sigma_j\right] A_{\mu\nu} \left[\frac{1}{N}\sum_j \xi_j^\nu\sigma_j\right]$ - uncondensed contribution to the 'energy' in the generalised Hopfield model
$\mathcal{D}_{\boldsymbol{\zeta}}[z^s,z^a]$	$=$	$2^c\left\langle \frac{1}{N}\sum_i \delta(z^s - z_i^\sigma(\boldsymbol{\sigma}))\,\delta(z^a - z_i^a(\boldsymbol{\sigma}))\,\delta_{\boldsymbol{\zeta}\boldsymbol{\xi}_i}\right\rangle_{\mathbf{m},r;t}$ - intrinsic noise distribution in the generalised Hopfield model
$\mathcal{L}(\mathbf{m})$	$\in \Re$	Lyapunov function of macrostate $\mathbf{m}$
$c^*(\mathbf{m})$	$=$	$\sup_{\mathbf{x}}(\mathbf{m}\cdot\mathbf{x} - \langle\log\cosh(\mathbf{x}\cdot\boldsymbol{\xi})\rangle)$ - Legendre transform of the cumulant generating function
Dy	$=$	$e^{-\frac{1}{2}y^2}\frac{dy}{\sqrt{2\pi}}$ - Gaussian integration measure
$\boldsymbol{\nabla}_{\mathbf{m}}$	$=$	$\left(\frac{\partial}{\partial m^1},\ldots\frac{\partial}{\partial m^\mu},\ldots,\frac{\partial}{\partial m^p}\right)$ - gradient differential operator
$\delta_{a,b}$	$=$	1 if $a=b$; 0 if $a\neq b$ — Kronecker delta function

$\delta(x-y)$	$= \begin{matrix}\infty \text{ if } x=y \\ 0 \text{ if } x\neq y\end{matrix}$	such that $\int dx\, f(x)\,\delta(x-y) = f(y)$ - Dirac delta function
$\Re(x)$	$=$	Real part of complex number x
$\Im(x)$	$=$	Imaginary part of complex number x
$\arg(x)$	$= \tan^{-1}\left(\frac{\Im(x)}{\Re(x)}\right)$	argument of complex number x
$\rho_i^{\varsigma}(h)$	$=$	$\frac{1}{N}\sum_j \delta_{\varsigma,\sigma_j}\,\delta(h-h_j(\boldsymbol{\sigma}))$ - joint distribution of fields and spins
Ψ		exponent in extremaly dominated integral
K	$\to \infty$	number of levels in ultrametric structure of replica matrices

E seppure volta è necessario nascondere
con le parole una cosa, bisogna farlo
in modo o che non appaia, o apparendo,
sia parata e prestala difesa

-*Niccolò Machiavelli*

Bibliography

[1] S. N. Laughton and A. C. C. Coolen. *Lyapunov function for separable stochastic neural networks with detailed balance.* Journal of Statistical Physics, **80**:375, 1995.

[2] S. N. Laughton and A. C. C. Coolen. *Quasi-periodicity and bifurcation phenomena in asymmetric neural networks.* Journal of Physics **A**, **27**:8011, 1994.

[3] S. N. Laughton and A. C. C. Coolen. *Order-Parameter flow in symmetric and non-symmetric fully connected attractor neural networks near saturation.* Physical Review **E**, **51**:2581, 1995.

[4] S. N. Laughton, A. C. C. Coolen, and D. Sherrington. *Order-parameter flow in the SK spin-glass II: inclusion of microscopic memory effects.* Oxford University preprint OUTP-95-44S, 1995.

[5] S. N. Laughton, A. C. C. Coolen, and D. Sherrington. *Dynamic replica theory for disordered spin systems.* Oxford University preprint OUTP-95-43S, 1995.

[6] L. F. Cugliandolo and M. V. Tsodyks. *Capacity of networks with correlated attractors.* Journal of Physics **A**, **27**:741, 1994.

[7] H. Haken. *Co-operative phenomena in systems far from thermal equilibrium and in non-physical systems.* Reviews of Modern Physics, **47**:67, 1975.

[8] B. Derrida and M. R. Evans. *The asymmetric exclusion model: Exact results through a matrix approach.* In V. Privman, editor, *Nonequilibrium Statistical Mechanics in One Dimension.* CUP, Cambridge, 1995. To be published.

[9] L. E. Reichl. *A Modern Course in Statistical Physics.* Edward Arnold, London, 1980.

[10] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry.* North Holland, Amsterdam, 1992.

[11] C. W. Gardiner. *Handbook of Stochastic Methods.* Springer, Berlin, 1983.

[12] E. Ising. *Beitrag zur theorie des ferromagnetismus.* Zeitschrift für Physik, **31**:253, 1925.

[13] R. J. Glauber. *Time dependent statistics of the Ising model.* Journal of Mathematical Physics, **4**:294, 1963.

[14] S. F. Edwards and P. W. Anderson. *Theory of spin glasses.* J. Phys. F., **5**:965, 1975.

[15] D. Sherrington and S. Kirkpatrick. *Solvable model of a spin glass.* Physical Review Letters, **35**:1792, 1975.

[16] G. Parisi. *Infinite number of order parameters for spin-glasses.* Physical Review Letters, **43**:1754, 1979.

[17] G. Parisi. *A sequence of approximated solutions to the S-K model for spin-glasses.* Journal of Physics **A**, **13**:L115, 1980.

[18] G. Parisi. *The order parameter for spin-glasses: A function of the interval 0-1.* Journal of Physics **A**, **13**:1101, 1980.

[19] G. Parisi. *Order parameter for spin-glasses.* Physical Review Letters, **50**:1946, 1983.

[20] D. J. Amit. *Modelling Brain Function.* CUP, Cambridge, 1989.

[21] B. Müller and J. Reinhardt. *Neural Networks - An Introduction.* Springer - Verlag, Berlin, 1991.

[22] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation.* Addison - Wesley, Redwood City, 1991.

[23] L. F. Abbott. *Learning in neural network memories.* Network, **1**:105, 1990.

[24] D. Sherrington. *Neural Networks: the spin glass approach.* In J. G. Taylor, editor, *Mathematical Approaches to Neural Networks*, chapter 8, page 261. North - Holland, Amsterdam, 1993.

[25] J. L. van Hemmen and R. Kühn. *Collective phenomena in neural networks.* In E. Domany, J. L. van Hemmen, and K. Schulten, editors, *Models of Neural Networks*, chapter 1. Springer - Verlag, Berlin, 1991.

[26] A. C. C. Coolen and D. Sherrington. *Dynamics of attractor neural networks.* In J. G. Taylor, editor, *Mathematical Approaches to Neural Networks*, chapter 9, page 293. North - Holland, Amsterdam, 1993.

[27] A. L. Hodgkin and A. F. Huxley. *A quantitative description of current and its application to conduction and excitation in nerve.* Journal of Physiology, **117**:500, 1952.

[28] L. F. Abbott and T. B. Kepler. *Model neurons: from Hodgkin-Huxley to Hopfield.* In L. Garrido, editor, *368 - Statistical Mechanics of Neural Networks*, Lecture Notes in Physics, chapter 1, page 5. Spriger-Verlag, Berlin, 1990.

[29] L. F. Abbott. *Decoding neuronal firing and modeling neural networks.* Quarterly Reviews of Biophysics, **27**:291, 1994.

[30] L. F. Abbott. *A network of oscillators.* Journal of Physics **A**, **23**:3835, 1990.

[31] R. Kühn and S. Bos. *Statistical mechanics for neural networks with continuous-time dynamics.* Journal of Physics **A**, **26**:831, 1993.

[32] J. J. Hopfield. *Neurons with graded response have collective computational properties like those of two-state neurons.* Proceedings of the National Academy of Science, USA, **81**:3088, 1984.

[33] J. J. Hopfield. *Neural networks and physical systems with emergent collective computational abilities.* Proceedings of the National Academy of Science, USA, **79**:2554, 1982.

[34] D. O. Hebb. *The Organisation of Behaviour: A Neurophysiological Theory.* Wiley, New York, 1949.

[35] D. J. Amit, H. Gutfreund, and H. Sompolinsky. *Spin-glass models of neural networks.* Physical Review **A**, **32**:1007, 1985.

[36] D. J. Amit, H. Gutfreund, and H. Sompolinsky. *Storing infinte numbers of patterns in spin-glass models of neural networks.* Physical Review Letters, **55**:1530, 1985.

[37] D. J. Amit, H. Gutfreund, and H. Sompolinsky. *Statistical mechanics of neural networks near saturation.* Annals of Physics, **173**:30, 1987.

[38] A. C. C. Coolen and Th. W. Ruijgrok. *Image evolution in Hopfield networks.* Physical Review **A**, **38**:4253, 1988.

[39] H.-J. Sommers. *Path-Integral approach to Ising spin-glass dynamics.* Physical Review Letters, **58**:1268, 1987.

[40] H. Horner, D. Bormann, M. Frick, H. Kinzelbach, and A. Schmidt. *Transients and basins of attraction in neural network models.* Zeitschrift für Physik **B**, **76**:381, 1989.

[41] O. Bernier. *Stochastic analysis of synchronous neural networks with asymmetric weights.* Europhysics Letters, **16**:531, 1991.

[42] U. Riedel, R. Kühn and J. L. van Hemmen. *Temporal sequences and chaos in neural nets.* Physical Review **A**, **38**:1105, 1988.

[43] B. Derrida, E. Gardner, and A. Zippelius. *An exactly solvable asymmetric neural network model.* Europhysics Letters, **4**:167, 1987.

[44] A. C. C. Coolen and D. Sherrington. *Order paramter flow in the fully connected Hopfield model near saturation.* Physical Review **E**, **36**:1921, 1994.

[45] A. C. C. Coolen and D. Sherrington. *Order parameter flow in the SK spin-Glass I: Replica symmetry.* Journal of Physics **A**, **27**:7687, 1994.

[46] A. C. C. Coolen and S. Franz. *Closure of macroscopic laws in disordered spin systems: A toy model.* Journal of Physics **A**, **27**:6947, 1994.

[47] K. H. Fisher and J. A. Hertz. *Spin Glasses.* CUP, Cambridge, 1991.

[48] M. Mezard, G. Parisi, and Virasoro. *Spin-glass Theory and Beyond.* World Scientific, Singapore, 1987.

[49] S. Kirkpatrick and D. Sherrington. *Infinite-ranged models of spin-glasses.* Physical Review **B**, **17**:4384, 1978.

[50] J. R. L. de Almeida and D. J. Thouless. *Stability of the Sherrington-Kirkpatrick solution of a spin-glass model.* Journal of Physics **A**, **11**:983, 1978.

[51] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. A. Virasoro. *Replica symmetry breaking and the nature of the spin glass phase.* Journal de Physique, **45**:843, 1984.

[52] A. Crisanti and H. Sompolinsky. *Dynamics of spin systems with asymmetric bonds: Langevin dynamics and a spherical model.* Physical Review **A**, **36**:4922, 1987.

[53] A. Crisanti and H. Sompolinsky. *Dynamics of spin systems with randomly asymmetric bonds: Ising spins and Glauber dynamics.* Physical Review **A**, **37**:4865, 1988.

[54] P. Spitzner and W. Kinzel. *Freezing transition in asymmetric random neural networks with deterministic dynamics.* Zeitschrift für Physik **B**, **77**:511, 1989.

[55] H. Rieger, M. Schreckenberg, and J. Zittartz. *Glauber dynamics of the Little-Hopfield model.* Zeitschrift für Physik **B**, **72**:523, 1988.

[56] H. Rieger, M. Schreckenberg, P. Spitzner, and W. Kinzel. *Alignment in the fully asymmetric SK model.* Journal of Physics **A**, **24**:3399, 1991.

[57] H. Eissfeller and M. Opper. *Mean-field Monte Carlo approach to the Sherrington-Kirkpatrick model with asymmetric couplings.* Physical Review **E**, **50**:709, 1994.

[58] A. C. C. Coolen and D. Sherrington. *Dynamics of fully connected attractor neural networks near saturation.* Physical Review Letters, **71**:3886, 1993.

[59] D. Sherrington and A. C. C. Coolen. *Evolution of order parameters in disordered spin systems - a closure procedure.* In L. Garrido, editor, *Proceedings of the Sitges Conference on '25 Years of Non-Equlibrium Statistical Mechanics'*, Berlin, 1995. Springer-Verlag.

[60] P. Peretto. *Collective properties of neural networks: a statistical physics approach.* Biological Cybernetics, **50**:51, 1984.

[61] D. Bollé, P. Dupont and B. Vinck. *On the overlap dynamics of multi-state neural networks with a finite number of patterns.* Journal of Physics **A**, **25**:2859, 1992.

[62] A. V. M. Herz, Z. Li, and J. L. van Hemmen. *Statistical mechanics of temporal association in neural networks with transmission delays.* Physical Review Letters, **66**:1370, 1991.

[63] H. Sompolinsky, A. Crisanti, and H. J. Sommers. *Chaos in random neural networks.* Physical Review Letters, **61**:259, 1988.

[64] M. Bauer and W. Martienssen. *Quasi-periodicity route to chaos in neural networks.* Europhysics Letters, **10**:427, 1989.

[65] K. Aihara, T. Takabe, and M. Toyoda. *Chaotic neural networks.* Physics Letters **A**, **144**:333, 1990.

[66] M. Bauer and W. Martienssen. *Lyapunov exponents and dimensions of chaotic neural networks.* Journal of Physics **A**, **24**:4557, 1991.

[67] B. Tirozzi and M. Tsodyks. *Chaos in highly diluted neural networks.* Europhysics Letters, **14**:727, 1991.

[68] L. Wang and J. Ross. *Chaos, multiplicity and synchronisity in higher-oder neural networks.* Physical Review **A**, **44**:R2259, 1991.

[69] D. Bollé, G. M. Shim, B. Vinck, and V. A. Zagrebnov. *Retrieval and chaos in extremely diluted Q-Ising neural networks.* Journal of Statistical Physics, **74**:565, 1994.

[70] D. Bollé, G. M. Shim, and B. Vinck. *Retrieval and chaos in layered Q-Ising neural networks.* Journal of Statistical Physics, **74**:583, 1994.

[71] A. Crisanti, M. Falcioni, and A. Vulpiani. *Transition from regular to complex behaviour in a discrete deterministic asymmetric neural network model.* Journal of Physics **A**, **26**:3441, 1993.

[72] K. Nutzel. *The length of attractors in asymmetric random neural networks with deterministic dynamics.* Journal of Physics **A**, **24**:L151, 1991.

[73] J. D. Crawford. *An introduction to bifurcation theory.* Reviews of Modern Physics, **63**:991, 1991.

[74] H. K. Khalil. *Nonlinear Systems.* Macmillan, New York, 1992.

[75] P. Hartman. *Ordinary Differential Equations.* Birkhauser, Boston, 1982.

[76] M. J. Feigenbaum. *Quantitative universality for a class of nonlinear transformations.* Journal of Statistical Physics, **19**:25, 1978.

[77] M. J. Feigenbaum. *The universal properties of nonlinear transformations.* Journal of Statistical Physics, **21**:669, 1979.

[78] V. I. Arnold. *Ordinary Differential Equations.* Springer, Berlin, 1978.

[79] A. C. C. Coolen and D. Sherrington. *Competition between pattern reconstruction and sequence processing in non-symmetric neural networks.* Journal of Physics **A**, **25**:5493, 1992.

[80] H. Nishimori and T. Ozeki. *Retrieval dynamics of associative memory of the Hopfield type.* Journal of Physics **A**, **26**:859, 1993.

[81] S. Amari and K. Maginu. *Statistical neurodynamics of assiciative memory.* Neural Networks, **1**:63, 1988.

[82] G. A. Kohring. *Convergence time and finite size effects in neural networks.* Journal of Physics **A**, **23**:2237, 1990.

[83] T. Ozeki and H. Nishimori. *Noise distributions in retrieval dynamics of the Hopfield model.* Journal of Physics **A**, **27**:7061, 1994.

[84] A. C. C. Coolen, S. N. Laughton, and D. Sherrington. *Order parameter flow in the SK spin-glass III: Aspects Replica symmetry breaking.* in preparation, 1995.

[85] H. Horner. *Time dependent local field distribution and remanent magnetization in the S-K spin glass.* Zeitschrift für Physik **B**, **78**:27, 1990.

[86] M. Thomsen, M. F. Thorpe, T. C. Choy, D. Sherrington, and H.-J. Sommers. *Local magnetic field distributions. III. Disordered systems.* Physical Review **B**, **33**:1931, 1986.

[87] Wolfgang Kinzel. *Remanent magnetisation of the infinite-range Ising spin glass.* Physical Review **B**, **33**:5086, 1986.

[88] N. D. Mackenzie and A. P. Young. *Lack of ergodicity in the infinite-range Ising spin-glass.* Physical Review Letters, **49**:301, 1982.

[89] K. Nemoto. *A numerical study of the pure states of the Sherrington-Kirkpatrick spin-glass model - a comparison with Parisi's replica-symmetry breaking solution.* Journal of Physics **C**, **20**:1325, 1987.

[90] B. Duplantier. *Comment on Parisi's equation for the SK model for spin glasses.* Journal of Physics **A**, **14**:283, 1981.

[91] A. V. Goltsev. *Equations of state for an infinte-ranged Ising spin glass in the framework of the Parisi solution.* Journal of Physics **A**, **17**:237, 1984.

[92] H. Sompolinsky. *Time-dependent order parameters in spin-glasses.* Physical Review Letters, **47**:935, 1981.

[93] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. A. Virasoro. *Nature of the spin glass phase.* Physical Review Letters, **52**:1156, 1984.

[94] J. R. L. de Almdeida and J. S. Lage. *Internal field distribution in the infinite-range Ising spin glass.* Journal of Physics **C**, **16**:939, 1983.

[95] R. W. Penney. *The Statistical Mechanics of Neural Networks and Spin Glasses.* PhD thesis, Oxford University Department of Physics - Theoretical Physics, 1994.

[96] H.-J. Sommers and W. Dupont. *Distribution of frozen field in the mean-field theory of spin glasses.* Journal of Physics **C**, **17**:5785, 1984.

[97] A. N. Shoshitaishvili. *Bifurcation of topological type at singular points of parametrised vector fields.* Funct. Anal. Appl., **6**:169, 1972.

[98] A. N. Shoshitaishvili. *The bifurcation of the topological type of the singular points of vector fields that depend on parameters.* Trudy Seminara I. G. Petrovskogo, **1**:279, 1975.

[99] M. H. Jensen, P. Bak, and T. Bohr. *Complete devil's staircase, fractal dimension and universality of mode-locking structure in the circle map.* Physical Review Letters, **50**:1637, 1983.

[100] M. H. Jensen, P. Bak, and T. Bohr. *Transition to chaos by interaction of resonances in dissipative systems. I. Circle maps.* Physical Review **A**, **30**:1960, 1984.

[101] P. Bak, T. Bohr, and M. H. Jensen. *Mode-locking and the transition to chaos in dissipative systems.* Physica Scripta, **T9**:50, 1985.

[102] P. Cvitanovic, B. Shraiman, and B. Soderberg. *Scaling laws for mode locking in circle maps.* Physica Scripta, **36**:263, 1985.

[103] P. Cvitanovic, M. H. Jensen, L. P. Kadanoff, and I. Procaccia. *Renormalisation, unstable manifolds and fractal structure of mode locking.* Physical Review Letters, **55**:343, 1985.

[104] M. J. Feigenbaum, L. P. Kadanoff, and S. J. Shenker. *Quasiperiodicity in disspiative systems: A renormalisation group analysis.* Physica **D**, **5**:370, 1982.

[105] S. J. Shenker. *Scaling behaviour in a map of a circle onto itself: Empirical results.* Physica **D**, **5**:405, 1982.

[106] D. Rand, S. Ostlund, J. Sethna, and E. D. Siggia. *Universal transistion from quasiperiodicity to chaos in dissipative systems.* Physical Review Letters, **49**:132, 1982.

[107] S. Newhouse, D. Ruelle, and F. Takens. *Occurence of strange axiom-A attractors near quasiperiodic flow on* $T^m, m \leq 3$. Communications in Mathematical Physics, **64**:35, 1978.

[108] N. S. Rasband. *Chaotic Dynamics of Non-linear Systems.* Wiley, New York, 1990.