Dual processes in neural network models. II. Analysis of zero-temperature fixed-point equations

1992 J. Phys. A: Math. Gen. 25 2593

(http://iopscience.iop.org/0305-4470/25/9/029)

J. Phys. A: Math. Gen. 25 (1992) 2593–2606. Printed in the UK

Dual processes in neural network models: II. Analysis of zero-temperature fixed-point equations

A C C Coolen†‡§ and L G V M Lenders†

† Department of Medical Physics and Biophysics, University of Nijmegen Geert Grooteplein Noord 21, NL-6525 EZ Nijmegen, The Netherlands
‡ Utrecht Biophysics Research Institute, University of Utrecht Princetonplein 5, NL-3584 CC Utrecht, The Netherlands

Received 21 August 1991

Abstract. The dynamics of learning in an unsupervised formulation of the Kohonen model has been shown to be equivalent to the dynamics of local order parameters in an attractor neural network with short-range Hebbian interactions and long-range anti-Hebbian interactions. In this paper we analyse the zero-temperature fixed-point equations of these systems. For the special case where the distribution of p-dimensional input vectors has rotational symmetry, the problem of finding the ground state is equivalent to finding the ground state of a system of p-dimensional Heisenberg spins with short-range ferromagnetic couplings and long-range anti-ferromagnetic couplings.

1. Introduction

In a previous paper [1] we have shown that the dynamics of local order parameters in attractor neural networks with short-range Hebbian interactions and long-range anti-Hebbian interactions is equivalent to the dynamics of learning in an unsupervised formulation of the Kohonen [2] model. Both processes are described by the same partial differential equation. This duality enables us to study the creation of topology conserving maps, using the mathematical tools and intuition developed in the field of attractor networks. We refer to [1] for a more general introduction into the fields of attractor neural networks and learning.

In this paper we will study in more detail the zero-temperature fixed-point equations of the above-mentioned dual models. We restrict ourselves to the case where the input vectors of the Kohonen model are drawn from a spherically symmetric distribution. In this case calculating the ground state is found to be equivalent to calculating the ground state of a continuous system of p-dimensional Heisenberg spins with short-range ferromagnetic interactions, in combination with long-range anti-ferromagnetic interactions. In the attractor model the variable p represents the number of vectors stored, whereas in the Kohonen model p stands for the number of input channels.

Since the Heisenberg spins are defined on a finite set D, and since we want to avoid all technical problems concerned with the system behaviour near the boundary ∂D, we expand the Hamiltonian and the fixed-point equations into first order in the width σ of the short-range interaction. In this case the contribution of the boundary

§ Present address: Department of Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK.

0305-4470/92/092593+14$04.50

region to the total energy can be neglected as far as the first order in σ is concerned. Imposing periodic boundary conditions allows one easily to calculate the ground state for any p. However, the maps formed are found to be of too low dimension (they are not topology conserving, in contrast to the outcome of the numerical experiments). Calculating the ground state with free boundary conditions generally requires solving a set of coupled nonlinear partial differential equations. For $p=2$ one finds the sine-Gordon equation. Here we focus on the case $p=3$ and calculate solutions of the fixed-point equations with free boundary conditions; these solutions have a significantly lower energy than the ground state corresponding to periodic boundary conditions. Although we are unable to prove this solution to be the configuration with the lowest energy, it qualitatively describes the results of iterating the zero-temperature equations numerically. Finally appendix A contains an overview of the notation introduced and a brief description of the main variables.

2. Sperical symmetry: Heisenberg spins

The dynamic equations describing the evolution of either interactions (Kohonen model) or local order parameters (attractor model) are [1]:

$$\partial_t \boldsymbol{\psi} = \langle \boldsymbol{\xi} \tanh(\beta \boldsymbol{\xi} \cdot n \otimes \boldsymbol{\psi}) \rangle_{\xi} - \boldsymbol{\psi} \tag{1}$$

where $\boldsymbol{\psi}(x, t) \in \mathbb{R}^p$, $x \in D \subset \mathbb{R}^m$ (bounded) and

$$(n \otimes f)(x) \equiv \int_D \mathrm{d}y \, n(x, y) f(y)$$

$$\langle \Phi(\boldsymbol{\xi}) \rangle_{\xi} = \int \mathrm{d}\boldsymbol{\xi} \, \rho(\boldsymbol{\xi}) \Phi(\boldsymbol{\xi}).$$

We will write spatial averages over D as

$$\langle f \rangle_D \equiv |D|^{-1} \int_D \mathrm{d}x \, f(x) \qquad |D| = \int_D \mathrm{d}x.$$

The operator n is assumed to be symmetric and semipositive definite. Furthermore the second moment of the probability distribution ρ is assumed to be finite. If the temperature $T \equiv \beta^{-1}$ is zero the (rescaled) Hamiltonian H of the attractor model is a Liapunov functional [1]

$$H = -\tfrac{1}{2} \langle \boldsymbol{\psi} \cdot n \otimes \boldsymbol{\psi} \rangle_D. \tag{2}$$

For the dual Kohonen model H is a Liapunov functional as soon as

$$(\forall x \in D)(\exists \gamma_{\xi}(x) \in [-1, 1]): \qquad \boldsymbol{\psi}(x) = \langle \boldsymbol{\xi} \gamma_{\xi}(x) \rangle_{\xi}$$

which, if not true at $t=0$, will certainly be true near equilibrium. Once the above condition is satisfied we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} H[\boldsymbol{\psi}] &= -\langle \partial_t \boldsymbol{\psi} \cdot n \otimes \boldsymbol{\psi} \rangle_D \\ &= -\langle\!\langle |\boldsymbol{\xi} \cdot n \otimes \boldsymbol{\psi}| (1 - \gamma_{\xi} \operatorname{sgn}[\boldsymbol{\xi} \cdot n \otimes \boldsymbol{\psi}]) \rangle\!\rangle_{\xi D} \leqslant 0. \end{aligned}$$

We will now take for ρ a distribution with spherical symmetry: $\rho(\boldsymbol{\xi})=\rho(|\boldsymbol{\xi}|)$. The distribution ρ is normalized by setting

$$\langle|\xi_\mu|\rangle_{\xi}=1.$$

In this case one can work out the average in (1): the zero-temperature equation is found to be

$$\partial_t\boldsymbol{\psi}=\frac{n\otimes\boldsymbol{\psi}}{|n\otimes\boldsymbol{\psi}|}-\boldsymbol{\psi}. \tag{3}$$

Equation (3) shows that in equilibrium the fields $\boldsymbol{\psi}$ are Heisenberg spins: $|\boldsymbol{\psi}(\boldsymbol{x})|=1$ for all $\boldsymbol{x}$. In terms of the formation of topology conserving maps this means that only representations of the *direction* of input vectors $\boldsymbol{\xi}$ can be formed; all information regarding $|\boldsymbol{\xi}|$ will be lost. This might have been expected since in the original Kohonen model [2] the same happens.

Our problem, which is to find the lowest energy fixed-point solution of (3), is equivalent to minimizing the Hamiltonian (2) for a field of Heisenberg spins $\boldsymbol{S}(\boldsymbol{x})$, since the energy minima of the Heisenberg system follow from

$$\frac{\delta}{\delta S_\mu(\boldsymbol{x})}\left(H+\int_D \mathrm{d}\boldsymbol{y}\,\Lambda(\boldsymbol{y})C_{\boldsymbol{y}}[\boldsymbol{S}]\right)=0 \qquad C_{\boldsymbol{y}}[\boldsymbol{S}]\equiv\tfrac{1}{2}\boldsymbol{S}^2(\boldsymbol{y})-\tfrac{1}{2} \tag{4}$$

which, in combination with the constraint equations $C_{\boldsymbol{y}}[\boldsymbol{S}]=0$, leads directly to

$$(n\otimes\boldsymbol{S})(\boldsymbol{x})=\Lambda(\boldsymbol{x})\boldsymbol{S}(\boldsymbol{x}) \qquad \boldsymbol{S}^2(\boldsymbol{x})=1. \tag{5}$$

Equation (5) is equivalent to the fixed-point equation of (3). The multiplier field $\Lambda(\boldsymbol{x})$ is found to be proportional to the energy density. Note that for the *spherical* spin model, where the local constraints $\boldsymbol{S}^2(\boldsymbol{x})=1$ are replaced by the global constraint $\langle\boldsymbol{S}^2\rangle_D=1$, one would have found a similar equation, where Λ would have been a constant. For positive semidefinite kernels n equation (5) can also be seen as the result of requiring $\boldsymbol{S}$ to be a critical point of the zero-temperature Liapunov functional F [1] of the dynamics (3):

$$F[\boldsymbol{S}]\equiv\tfrac{1}{2}\langle\boldsymbol{S}\cdot n\otimes\boldsymbol{S}\rangle_D-\langle|n\otimes\boldsymbol{S}|\rangle_D$$

since the requirement $\delta F=0$ (for all $\delta\boldsymbol{S}$ with $\delta\boldsymbol{S}|_{\partial D}=\boldsymbol{0}$) yields equation (5). From now on we will work with expression (5).

In [1] we have shown that topology conserving maps can be expected to emerge if the operator n has the form

$$\begin{aligned} &n(\boldsymbol{x},\boldsymbol{y})=n_+(|\boldsymbol{x}-\boldsymbol{y}|)-J|D|^{-1}\\ &0<J<J^*\equiv|D|^{-1}\int_D\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\,n_+(|\boldsymbol{x}-\boldsymbol{y}|)\end{aligned} \tag{6}$$

where n_+ stands for a positive short-range interaction. In order to gain insight into which types of solutions of (5) are low-energy configurations, we have iterated equation (1) numerically for $T=0$, $p=3$, $D\equiv[-\frac{1}{2},\frac{1}{2}]^2$ and

$$\rho(\boldsymbol{\xi})=\frac{3}{4\pi}\left(\frac{3}{8}\right)^3\theta\left(\frac{8}{3}-|\boldsymbol{\xi}|\right) \qquad n_+(z)=(2\pi\sigma^2)^{-1}\exp[-\tfrac{1}{2}z^2/\sigma^2].$$

Since, by virtue of the above choice for $\rho(\boldsymbol{\xi})$, the normalization $\langle|\xi_\mu|\rangle=1$ is built in, one must expect the vectors $\boldsymbol{\psi}(\boldsymbol{x})$ to be located on the surface of the unit sphere in

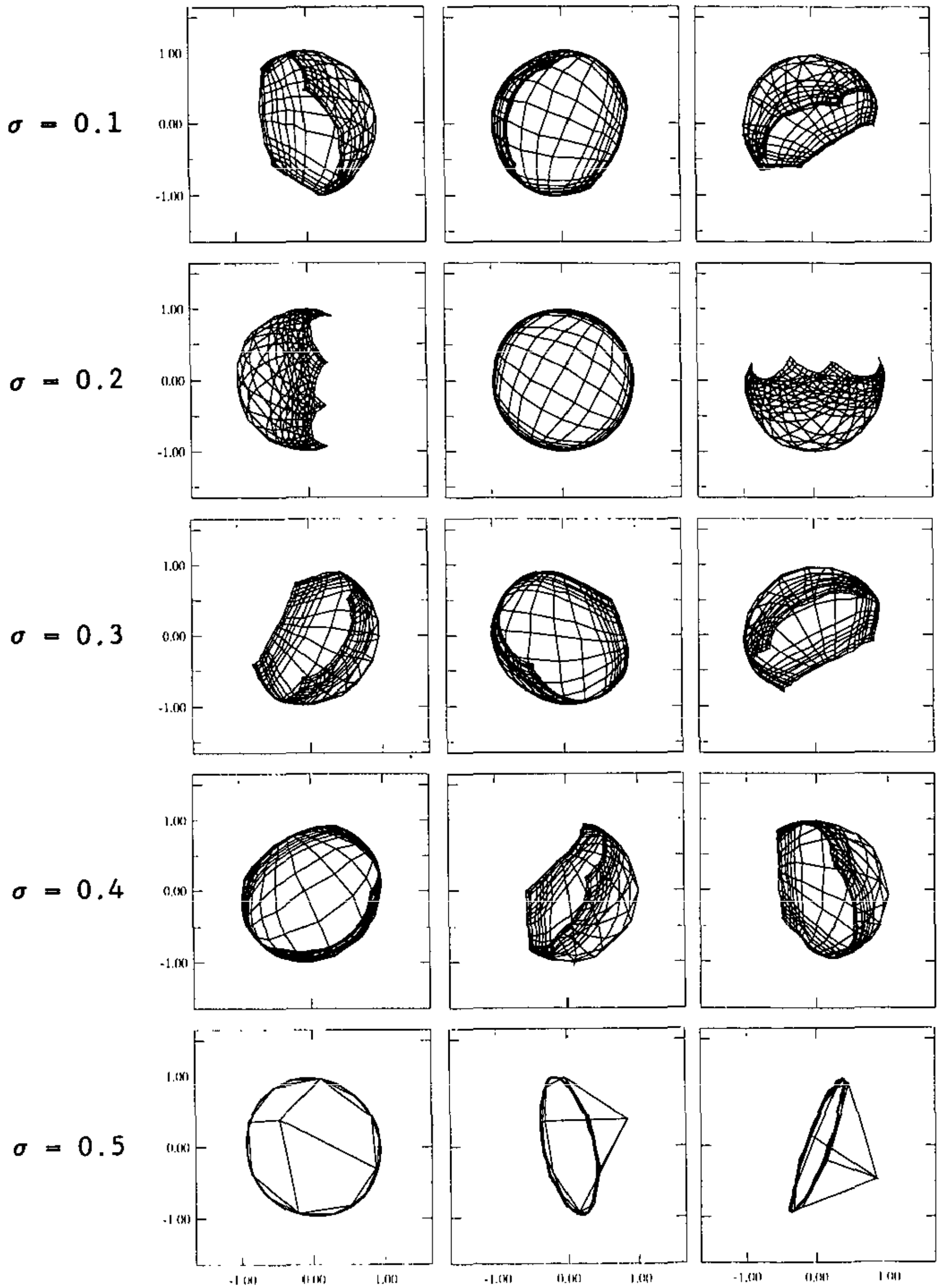


Figure 1. Projections of the equilibrium state $\psi(x, \infty)$ (obtained by numerical iteration of the dynamics (1)) onto the planes $\psi_1 = 0$ (left), $\psi_2 = 0$ (middle) and $\psi_3 = 0$ (right), for $J = 0.5$ and σ ranging from 0.1 to 0.5.

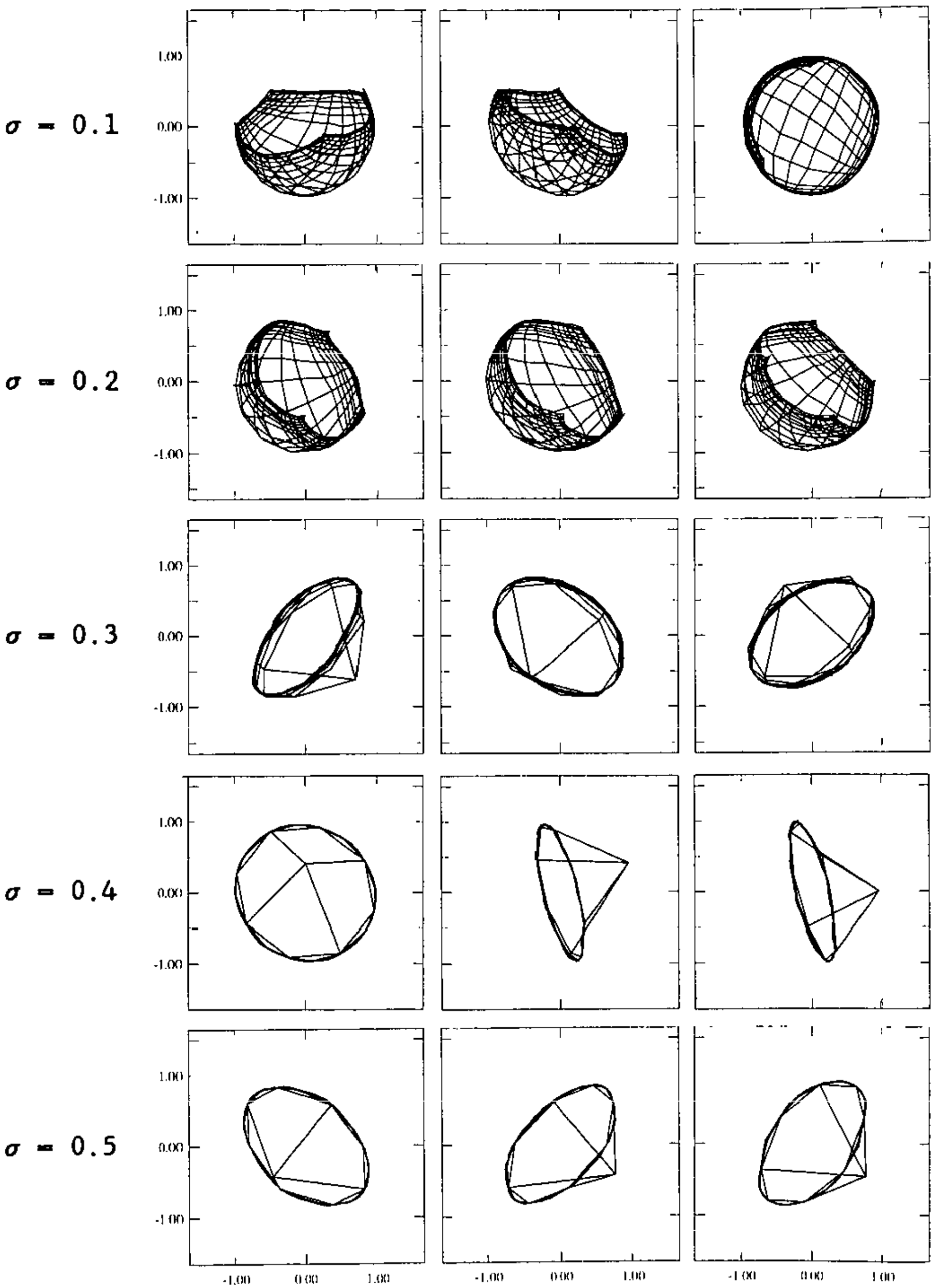


Figure 2. Projections of the equilibrium state $\psi(x, \infty)$ (obtained by numerical iteration of the dynamics (1)) onto the planes $\psi_1 = 0$ (left), $\psi_2 = 0$ (middle) and $\psi_3 = 0$ (right), for $J = 1.0$ and σ ranging from 0.1 to 0.5.

$\mathbb{R}^3$. As in [1], we have chosen (following Kohonen [2]) as a graphical representation of the equilibrium state the projections of $\boldsymbol{\psi}(\boldsymbol{x}, \infty)$ onto the planes $\psi_1 = 0$, $\psi_2 = 0$ and $\psi_3 = 0$. Figures 1 and 2 show the equilibrium configurations reached after numerical iteration of (1) for $J = 0.5$ and $J = 1.0$ (respectively). The width σ was varied from 0.1 to 0.5 ($\Delta\sigma = 0.1$). These results show that, for small width σ of the positive interaction, toplogy conserving maps are formed of the orientations of the input vectors $\boldsymbol{\xi}$. If, however, the width σ is too large compared to the strength J of the negative long-range interaction, the number of degrees of freedom of the ground state is found to be reduced by one. This reduction can be expected to be related to the appearance of negative eigenvalues of the operator n, since for large σ the condition $J > J^*$ in (6) will be violated.

3. Expansion in the width of the ferromagnetic interaction

Since the numerical experiments showed topology conserving maps of the distribution $\rho(\boldsymbol{\xi})$ to be formed for sufficiently small values of the range σ of the ferromagnetic interaction n_+, we proceed by making expansions in powers of σ. From now on we will assume that the neurons are located on a finite 2D array: $D \subset \mathbb{R}^2$. First we write n_+ as

$$n_+(z) \equiv \sigma^{-2} f(z/\sigma) \qquad \int \mathrm{d}\boldsymbol{x}\, f(|\boldsymbol{x}|) = \int \mathrm{d}\boldsymbol{x}\, \boldsymbol{x}^2 f(|\boldsymbol{x}|) = 1. \tag{7}$$

Using (7) we can write for the operation of the kernel n:

$$(n \otimes g)(\boldsymbol{x}) = g(\boldsymbol{x}) \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) - J\langle g\rangle_D + \sigma \boldsymbol{\nabla} g(\boldsymbol{x}) \cdot \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z$$
$$+ \frac{1}{2}\sigma^2 \sum_{\mu\nu} \partial^2_{\mu\nu} g(\boldsymbol{x}) \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z_\mu z_\nu + \mathrm{O}\left(\sigma^3 \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z^3\right).$$

As a consequence we find

$$\langle \boldsymbol{S} \cdot n \otimes \boldsymbol{S}\rangle_D = |D|^{-1} \int_D \mathrm{d}\boldsymbol{x} \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) - J\langle \boldsymbol{S}\rangle^2_D$$
$$+ \frac{1}{2}\sigma^2 |D|^{-1} \sum_{\mu\nu} \left(\int_D \mathrm{d}\boldsymbol{x}\, \boldsymbol{S}(\boldsymbol{x}) \cdot \partial^2_{\mu\nu} \boldsymbol{S}(\boldsymbol{x}) \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z_\mu z_\nu \right)$$
$$+ \mathrm{O}\left(\sigma^3 |D|^{-1} \int_D \mathrm{d}\boldsymbol{x} \int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z^3\right)$$

(where we have used $\boldsymbol{S} \cdot \partial_\lambda \boldsymbol{S} = 0$). The above expression shows that, apart from a constant, the contribution of the boundary region ∂D (of width σ) to the energy (2) is of order σ^3, in contrast to the surface energy which is of order σ^2. Therefore, as far as the first non-trivial order in σ is concerned, we may forget about the boundary ∂D. Since away from the boundary region we have

$$\int_{(D-\boldsymbol{x})/\sigma} \mathrm{d}z\, f(|z|) z_\mu z_\nu = \frac{1}{2} \delta_{\mu\nu}$$

the fixed-point equation (5) and the corresponding energy (2) become

$$S+\varepsilon\Delta S-J\langle S\rangle+\mathrm{O}(\varepsilon^2)=\Lambda S \tag{8}$$

$$E=-\tfrac{1}{2}+\tfrac{1}{2}J\langle S\rangle_D^2+\tfrac{1}{2}\varepsilon\sum_{\lambda}\langle(\nabla S_\lambda)^2\rangle_D+\mathrm{O}(\varepsilon^2) \tag{9}$$

where $\varepsilon\equiv\frac{1}{4}\sigma^2$ and $\Lambda=\Lambda(x)$. If J is non-zero there is a competition between the surface-tension term in (9) (which favours configurations with $S=\langle S\rangle_D$) and the term proportional to J (which favours configurations with $\langle S\rangle_D=0$). Since $S^2(x)=1$ for all x, the two terms cannot be minimized simultaneously. From now on we will choose for D a unit square: $D\equiv[-\frac{1}{2},\frac{1}{2}]^2$.

We find that the calculation of the energy of the ground state becomes a trivial problem if we assume the solution to be periodic, or if we impose periodic boundary conditions. One can now expand the solution S in a discrete Fourier series:

$$S(x)=\sum_{k\in\mathbb{Z}^2}\hat{S}_k\,\mathrm{e}^{2\pi i k\cdot x}\qquad \hat{S}_k^*=\hat{S}_{-k} \tag{10}$$

$$E=-\tfrac{1}{2}+\tfrac{1}{2}J\hat{S}_0^2+2\pi^2\varepsilon\sum_{k\in\mathbb{Z}^2}k^2|\hat{S}_k|^2+\mathrm{O}(\varepsilon^2). \tag{11}$$

We now replace the local constraint $S^2(x)=1$ of the Heisenberg system by the global constraint $\langle S^2\rangle_D$ of the spherical model, which is justified *a posteriori* if the ground state of the spherical model obeys the local constraint as well. In terms of the Fourier coefficients this means

$$\sum_{k\in\mathbb{Z}^2}|\hat{S}_k|^2=1. \tag{12}$$

Using (12) we can rewrite the energy (11):

$$E=-\tfrac{1}{2}+\tfrac{1}{2}J+\sum_{k\in\mathbb{Z}^2,\neq 0}|\hat{S}_k|^2(2\pi^2\varepsilon k^2-\tfrac{1}{2}J)+\mathrm{O}(\varepsilon^2). \tag{13}$$

Expression (13) allows us to identify the ground state $S_g(x)$ of the system with periodic boundary conditions (apart from the pathological case $p=1$):

$$\begin{aligned}
&J<4\pi^2\varepsilon: && E_g=-\tfrac{1}{2}+\tfrac{1}{2}J+\mathrm{O}(\varepsilon^2)\\
&S_g(x)=\hat{e} && \text{(a constant field).}
\end{aligned} \tag{14}$$

For $J>4\pi^2\varepsilon$ we have to distinguish the situations $p\leqslant 3$ and $p>3$. In the first case there are two distinct components (separated by energy barriers), in the second case there is one:

$$\begin{aligned}
&J>4\pi^2\varepsilon: && E_g=-\tfrac{1}{2}+2\pi^2\varepsilon+\mathrm{O}(\varepsilon^2)\\
&p=2,3: && S_{g1}(x)=\hat{e}\cos(2\pi x_1)+\hat{f}\sin(2\pi x_1)\\
& && S_{g2}(x)=\hat{e}\cos(2\pi x_2)+\hat{g}\sin(2\pi x_2)\\
& && \hat{e}\cdot\hat{f}=0\qquad \hat{e}^2=\hat{f}^2=1\\
&p\geqslant 4: && S_g(x)=\cos\gamma(\hat{e}\cos(2\pi x_1)+\hat{f}\sin(2\pi x_1))\\
& && \quad+\sin\gamma(\hat{g}\cos(2\pi x_2)+\hat{h}\sin(2\pi x_2))\\
& && \gamma\in[0,2\pi]\\
& && \hat{e}^2=\hat{f}^2=\hat{g}^2=\hat{h}^2=1\\
& && \hat{e}\cdot\hat{f}=\hat{g}\cdot\hat{h}=\hat{e}\cdot\hat{g}=\hat{e}\cdot\hat{h}=\hat{f}\cdot\hat{g}=\hat{f}\cdot\hat{h}=0.
\end{aligned} \tag{15}$$

Clearly, the ground state configurations found for $p=3$, assuming periodicity (circles), are different from the outcome of the numerical experiments described in section 2; apparently the true ground state of the system is not periodic. However, the ground state energy of the periodic system may serve as a guideline, since the non-periodic configurations that we will try to calculate must have an energy which is significantly lower than (15).

Let us now expand both the solution $\boldsymbol{S}$ of the fixed-point equation (8) and the multiplier field Λ in powers of ε:

$$\boldsymbol{S}=\boldsymbol{S}_0+\varepsilon^{\eta}\boldsymbol{S}_1+\mathrm{O}(\varepsilon^{2\eta}) \qquad \boldsymbol{S}_0^2=1 \qquad \boldsymbol{S}_0\cdot\boldsymbol{S}_1=0$$

$$\Lambda=\Lambda_0+\varepsilon^{\xi}\Lambda_1+\mathrm{O}(\varepsilon^{2\xi}).$$

$(\xi,\eta\geqslant\frac{1}{2})$ in terms of which the fixed-point equation (8) and the energy (9) become

$$\boldsymbol{S}_0(1-\Lambda_0)+\varepsilon\Delta\boldsymbol{S}_0+\varepsilon^{\eta}\boldsymbol{S}_1(1-\Lambda_0)-\varepsilon^{\xi}\Lambda_1\boldsymbol{S}_0-J(\langle\boldsymbol{S}_0\rangle_D+\varepsilon^{\eta}\langle\boldsymbol{S}_1\rangle_D)=\mathrm{O}(\varepsilon^2,\varepsilon^{1+\eta}) \tag{16}$$

$$E=-\tfrac{1}{2}+\tfrac{1}{2}J\langle\boldsymbol{S}_0\rangle_D^2+\varepsilon^{\eta}J\langle\boldsymbol{S}_0\rangle_D\cdot\langle\boldsymbol{S}_1\rangle_D+\tfrac{1}{2}\varepsilon\sum_{\lambda}\langle(\boldsymbol{\nabla}S_{0\lambda})^2\rangle_D+\mathrm{O}(\varepsilon^2,\varepsilon^{1+\eta}). \tag{17}$$

The trivial solution $\boldsymbol{S}_0(\boldsymbol{x})=\hat{\boldsymbol{e}}$ (a constant field) was already discussed while periodic solutions were being studied; from now on we will assume $\boldsymbol{S}_0\neq\langle\boldsymbol{S}_0\rangle_D$. We must now specify the order of the long-range interaction parameter J. If $J\equiv\tilde{J}\varepsilon^{\rho}$ $(\rho>0)$ the zeroth order of (16) immediately yields $\Lambda_0=1$, therefore

$$\varepsilon\Delta\boldsymbol{S}_0-\varepsilon^{\xi}\Lambda_1\boldsymbol{S}_0-\tilde{J}\varepsilon^{\rho}(\langle\boldsymbol{S}_0\rangle_D+\varepsilon^{\eta}\langle\boldsymbol{S}_1\rangle_D)=\mathrm{O}(\varepsilon^2,\varepsilon^{1+\eta}).$$

The case $\rho>1$ is equivalent to retaining only the short-range interaction; therefore it is easy to calculate the ground state in lowest order in ε:

$$\rho>1: \qquad \boldsymbol{S}_{0\mathrm{g}}=\hat{\boldsymbol{e}} \qquad \text{(a constant field)}$$

$$E_{\mathrm{g}}=-\tfrac{1}{2}+\mathrm{O}(\varepsilon^2,\varepsilon^{\rho}).$$

Clearly for $J\ll\varepsilon$ topology conserving maps will not be formed. The situation becomes more interesting as soon as J really enters the problem. Therefore we now consider $0\leqslant\rho\leqslant1$. We can already derive from $\Delta\boldsymbol{S}_0^2=0$ that the equation $\Delta\boldsymbol{S}_0=\boldsymbol{0}$ implies $\boldsymbol{S}_0=\hat{\boldsymbol{e}}$ (the constant field), a case which we will not consider. For $\rho\in[0,1]$ we must therefore (according to the fixed-point equation) require $\xi=1$ or $\rho=1$ or $\eta+\rho=1$. One can now work out the various possibilities for the exponents of ε; the result is

$$\rho=1: \qquad \boldsymbol{S}_0^2=1 \qquad \Delta\boldsymbol{S}_0=\lambda\boldsymbol{S}_0+\tilde{J}\langle\boldsymbol{S}_0\rangle_D \tag{18}$$

$$\rho<1: \qquad \boldsymbol{S}_0^2=1 \qquad \langle\boldsymbol{S}_0\rangle_D=\boldsymbol{0} \qquad \Delta\boldsymbol{S}_0=\lambda\boldsymbol{S}_0+\boldsymbol{c}. \tag{19}$$

Equations (18) and (19) are to be solved for the fields $\boldsymbol{S}(\boldsymbol{x})$ and $\lambda(\boldsymbol{x})$ and the constant vector $\boldsymbol{c}$ simultaneously (the field $\lambda(\boldsymbol{x})$ can be eliminated by using $\boldsymbol{S}^2(\boldsymbol{x})=1$). The energy corresponding to the solutions of (18) and (19) is

$$E_{\rho=1}=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon\left(\tilde{J}\langle\boldsymbol{S}_0\rangle_D^2+\sum_{\lambda}\langle(\boldsymbol{\nabla}S_{0\lambda})^2\rangle_D\right)+\text{higher orders} \tag{20}$$

$$E_{\rho>1}=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon\sum_{\lambda}\langle(\boldsymbol{\nabla}S_{0\lambda})^2\rangle_D+\text{higher orders}. \tag{21}$$

4. Free boundary conditions

In this section we will construct solutions of (18) and (19) by writing $\boldsymbol{S}_0$ in polar coordinates (in order to build in the local constraints $\boldsymbol{S}_0^2(\boldsymbol{x})=1$). The fixed-point equations will then be transformed into nonlinear partial differential equations for the polar angles. For $p=2$ only one field ϕ remains:

$$\boldsymbol{S}_0=(\cos\phi, \sin\phi) \qquad \phi=\phi(\boldsymbol{x}).$$

The result of working out $\Delta\boldsymbol{S}_0=\lambda\boldsymbol{S}_0+\boldsymbol{c}$ in terms of the field ϕ is (after elimination of the field λ)

$$\Delta\phi=\boldsymbol{c}\cdot\begin{pmatrix}-\sin\phi\\ \cos\phi\end{pmatrix}.$$

Upon choosing a convenient basis in $\boldsymbol{S}_0$-space, $\boldsymbol{c}\equiv c(-1,0)$, we find the sine-Gordon equation. Taking the constraints on $\langle\boldsymbol{S}_0\rangle_D$ into account, the $p=2$ problem and the corresponding energies are found to be

$$\Delta\phi=c\sin\phi$$

$$\rho=1: \qquad \langle\cos\phi\rangle_D=-c\tilde{J}^{-1} \qquad \langle\sin\phi\rangle_D=0$$

$$\rho<1: \qquad \langle\cos\phi\rangle_D=0 \qquad \langle\sin\phi\rangle_D=0$$

$$E_{\rho=1}=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon(\tilde{J}^{-1}c^2+\langle(\nabla\phi)^2\rangle_D)+\ldots$$

$$E_{\rho<1}=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon\langle(\nabla\phi)^2\rangle_D+\ldots$$

Here we will concentrate on the case $p=3$:

$$\boldsymbol{S}_0=(\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta) \qquad \phi=\phi(\boldsymbol{x}) \qquad \theta=\theta(\boldsymbol{x}). \tag{22}$$

The equation to be solved is

$$\Delta\boldsymbol{S}_0=\lambda\boldsymbol{S}_0+\boldsymbol{c}. \tag{23}$$

If we choose the z-axis in the direction of $\boldsymbol{c}$, the result of working out (23) in terms of the polar angles is

$$\begin{aligned}&\sin\theta\Delta\theta+\cos\theta(\nabla\theta)^2+c+\lambda\cos\theta=0\\ &\sin\theta\Delta\phi+2\cos\theta(\nabla\theta\cdot\nabla\phi)=0\\ &\cos\theta\Delta\theta=\sin\theta(\nabla\phi)^2+\sin\theta(\nabla\theta)^2+\lambda\sin\theta\end{aligned} \tag{24}$$

(where $c\equiv|\boldsymbol{c}|$). Elimination of the field λ gives

$$\begin{aligned}&\sin\theta\Delta\phi+2\cos\theta(\nabla\theta\cdot\nabla\phi)=0\\ &\Delta\theta-\sin\theta\cos\theta(\nabla\phi)^2+c\sin\theta=0.\end{aligned} \tag{25}$$

The constraints to be met by the solutions of (25) are (according to (18) and (19)):

$$\rho=1: \qquad \langle\cos\phi\sin\theta\rangle_D=\langle\sin\phi\sin\theta\rangle_D=0 \qquad \langle\cos\theta\rangle_D=c\tilde{J}^{-1} \tag{26}$$

$$\rho<1: \qquad \langle\cos\phi\sin\theta\rangle_D=\langle\sin\phi\sin\theta\rangle_D=\langle\cos\theta\rangle_D=0. \tag{27}$$

The corresponding energies are found to be

$$\rho=1: \qquad E=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon(\tilde{J}^{-1}c^2+\langle\sin^2\theta(\nabla\phi)^2+(\nabla\theta)^2\rangle_D)+\ldots \tag{28}$$

$$\rho<1: \qquad E=-\tfrac{1}{2}+\tfrac{1}{2}\varepsilon\langle\sin^2\theta(\nabla\phi)^2+(\nabla\theta)^2\rangle_D+\ldots. \tag{29}$$

Clearly for $c=0$ the cases $\rho=1$ and $\rho<1$ are identical.

In order to proceed we will have to make use of symmetries of the problem, since the fixed-point equations (25) are not easily solved directly. Inspired by the numerical results in [1] we make an ansatz for the field ϕ:

$$(\cos\phi, \sin\phi) = |x|^{-1}(x_1, x_2). \tag{30}$$

The above ansatz implies $\nabla\phi = r^{-2}(-x_2, x_1)$ (where $r \equiv |x|$), which transforms the fixed-point equations (25) into

$$\nabla\theta\cdot(-x_2, x_1) = 0 \qquad \Delta\theta - r^{-2}\sin\theta\cos\theta + c\sin\theta = 0. \tag{31}$$

The energy becomes

$$\rho = 1: \qquad E = -\tfrac{1}{2} + \tfrac{1}{2}\varepsilon(\tilde{J}^{-1}c^2 + \langle r^{-2}\sin^2\theta + (\nabla\theta)^2\rangle_D) + \ldots \tag{32}$$

$$\rho < 1: \qquad E = -\tfrac{1}{2} + \tfrac{1}{2}\varepsilon\langle r^{-2}\sin^2\theta + (\nabla\theta)^2\rangle_D + \ldots \tag{33}$$

We may now conclude from (31) that $\theta = \theta(r)$ and we find our problem of solving (31) reduced to solving an ordinary (nonlinear) differential equation:

$$r^2\theta''(r) + r\theta'(r) - \sin\theta\cos\theta + cr^2\sin\theta = 0. \tag{34}$$

Upon making the substitution $\theta(r) \equiv \pi/2 + \frac{1}{2}G(u)$ $(u \equiv \log r)$ equation (34) acquires the form

$$G'' + \sin G + 2c\,\mathrm{e}^{2u}\cos(\tfrac{1}{2}G) = 0. \tag{35}$$

If $c \neq 0$ we cannot solve this equation analytically. We will study the case $c = 0$ and try to analyse the effect of assuming $c > 0$ as a perturbation on the $c = 0$ solution. For $c = 0$ equation (35) simply describes the motion of a pendulum; there are two types of solutions [3, 4], which correspond to oscillatory motion and libration, respectively:

$$G_k^+(u) = 2\arcsin(k\,\mathrm{sn}[u+q, k]) \qquad |k| \leqslant 1 \qquad q \in \mathbb{R}$$

$$G_k^-(u) = 2\arcsin(\mathrm{sn}[k(u+q), k^{-1}]) \qquad |k| \geqslant 1 \qquad q \in \mathbb{R}$$

or

$$\theta_k^+(r) = \frac{\pi}{2} + \arcsin(k\,\mathrm{sn}[\log(r)+q, k]) \qquad |k| \leqslant 1 \qquad q \in \mathbb{R} \tag{36}$$

$$\theta_k^-(r) = \frac{\pi}{2} + \arcsin(\mathrm{sn}[k\log(r)+kq, k^{-1}]) \qquad |k| \geqslant 1 \qquad q \in \mathbb{R}. \tag{37}$$

Here the function sn[.,.] is the elliptic function of Jacobi [5], defined by the relations

$$\mathrm{sn}[u, k] \equiv \sin\phi \qquad u = \int_0^\phi \mathrm{d}\psi(1 - k^2\sin^2\psi)^{-1/2}.$$

The corresponding energies can now be calculated from (32) and (33):

$$|k| \leqslant 1: \qquad \langle r^{-2}\sin^2\theta + (\nabla\theta)^2\rangle_D = \langle r^{-2}(k^2 + 1 - 2k^2\,\mathrm{sn}^2[\log(r)+q, k])\rangle_D$$

$$|k| \geqslant 1: \qquad \langle r^{-2}\sin^2\theta + (\nabla\theta)^2\rangle_D = \langle r^{-2}(k^2 + 1 - 2\,\mathrm{sn}^2[k\log(r)+kq, k^{-1}])\rangle_D.$$

Since $\langle r^{-2}\rangle_D = \infty$ we can infer from the above relations, using $\mathrm{sn}^2[\,] \leqslant 1$, that for $k \neq 1$ the first non-trivial contribution to the energy diverges. The only energically acceptable solution is found for $k = 1$. Since $\mathrm{sn}[x, 1] = \tanh(x)$ we find

$$\theta(r) = \frac{\pi}{2} + \arcsin\left(\frac{r^2 - e^{-2q}}{r^2 + e^{-2q}}\right) \tag{38}$$

$$S_0 = (r^2 + e^{-2q})^{-1}\begin{pmatrix} 2x_1\, e^{-q} \\ 2x_2\, e^{-q} \\ e^{-2q} - r^2 \end{pmatrix}. \tag{39}$$

The value of the remaining parameter q is fixed by the constraint $\langle S_0\rangle_D = \mathbf{0}$:

$$A(q) \equiv \left\langle \frac{e^{-2q} - r^2}{e^{-2q} + r^2} \right\rangle_D = 0.$$

Since $\mathrm{d}A/\mathrm{d}q < 0$ for all q and $A(-\infty) = -A(\infty) = 1$, the value of q is well defined and can be calculated numerically. One finds

$$q \approx 1.023.$$

The corresponding energy turns out to be

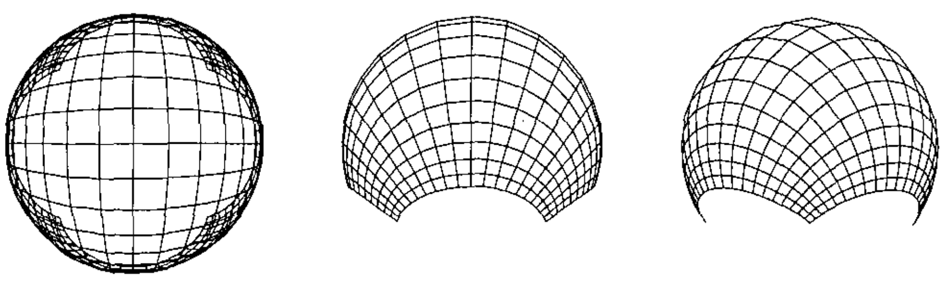
$$\begin{aligned} E &= -\tfrac{1}{2} + 4\varepsilon\, e^{-2q}\langle (r^2 + e^{-2q})^{-2}\rangle + \ldots \\ &\approx -\tfrac{1}{2} + 8.857\varepsilon + \ldots . \end{aligned} \tag{40}$$

Figure 3 shows the projections of the solution (39) (which is a mapping from the square D to the surface of the unit sphere in $\mathbb{R}^3$) onto the planes $S_3 = 0$, $S_2 = 0$ and $S_1 = S_2$. There is a qualitative agreement between configuration (39) and the numerical results obtained in section 2 (for small σ). For $\rho < 1$ the energy (40) is clearly below the ground state energy of the periodic fixed points, as calculated in section 3 ($E_{\mathrm{g,per}} = -\frac{1}{2} + 2\pi^2\varepsilon + \ldots$); for $\rho = 1$, however, the energy (40) will no longer be below the periodic ground state energy as soon as $\tilde{J} < 17.714$ (where the constant solution $S_0 = \hat{e}$ will have a lower energy).

Finally we will study the consequences of allowing for c not to be zero ($c \ll 1$). We will expand the solution of the full equation (35) in powers of c and take for the zeroth order the solution corresponding to (39):

$$G(u) = G_0(u) + cG_1(u) + O(c^2)$$

$$G_0(u) = 2\arcsin(\tanh(u + q)).$$



(a) (b) (c)

Figure 3. Projections of the fixed-point solution (39) onto the planes (a) $S_3 = 0$, (b) $S_2 = 0$ and (c) $S_1 = S_2$.

Insertion into (35) gives a linear problem for G_1:

$$G_1''+G_1(2\cosh^{-2}(u+q)-1)+2\,\mathrm{e}^{2u}\cosh^{-1}(u+q)=0.$$

This equation can be cast into a more convenient form by the transformation

$$u\equiv -q-z \qquad G_1(u)\equiv\phi(z).$$

In terms of these new variables we have

$$\left(\frac{\mathrm{d}}{\mathrm{d}z}-\tanh(z)\right)\left(\frac{\mathrm{d}}{\mathrm{d}z}+\tanh(z)\right)\phi=f \tag{41}$$

where

$$f(z)\equiv -2\,\mathrm{e}^{-2q}\,\mathrm{e}^{-2z}\cosh^{-1}(z).$$

The solution of (41) can be calculated easily, since one now has to solve a first-order differential equation twice. The solution is

$$\begin{aligned}\phi(z)=A\cosh^{-1}(z)&+B(z\cosh^{-1}(z)+\sinh(z))\\&+\cosh^{-1}(z)\int_0^z\mathrm{d}s\,\cosh^2(s)\int_0^s\mathrm{d}t\,\cosh^{-1}(t)f(t).\end{aligned} \tag{42}$$

We can now substitute the function f into (42), perform some integrations and ignore all homogeneous terms resulting from these integrations (since they are accounted for by the first part of (42)). The final result is

$$\begin{aligned}\phi(z)=A\cosh^{-1}(z)&+B(z\cosh^{-1}(z)+\sinh(z))\\&+2\,\mathrm{e}^{-2q}\bigg(\log(1+\mathrm{e}^{-2z})(\sinh(z)+z\cosh^{-1}(z))\\&+\mathrm{e}^{-z}-\cosh^{-1}(z)\int_0^z\mathrm{d}s\,s(\tanh(s)-1)\bigg).\end{aligned} \tag{43}$$

The energy shift resulting from the introduction of $c\neq 0$ can be written as

$$\Delta E=-\tfrac{1}{2}\varepsilon\left(c\,\mathrm{e}^{2q}\left\langle\left(\mathrm{e}^{2z}\cosh^{-1}(z)\left(\frac{\mathrm{d}}{\mathrm{d}z}-\tanh(z)\right)\phi(z)\right)_{z=-q-\log(r)}\right\rangle_D+\mathrm{O}(c^2)\right).$$

Tedious but straightforward integration shows that the energy contributions of both the homogeneous and the inhomogeneous parts of (43) are finite. Therefore we can write

$$-2\varepsilon^{-1}\Delta E=c(AE_1+BE_2+E_{\text{inhom}})+\mathrm{O}(c^2). \tag{44}$$

The constraint to be met can, in terms of ϕ, be written as

$\rho<1$: $\qquad \langle\cosh^{-1}[q+\log(r)]\phi[-q-\log(r)]\rangle_D=0$

$\rho=1$: $\qquad \langle\cosh^{-1}[q+\log(r)]\phi[-q-\log(r)]\rangle_D=\tilde{J}^{-1}$.

Again both the homogeneous and the inhomogeneous parts of ϕ give finite contributions to the integrals appearing in the above averages. This results in a linear relation between the constants A and B(which allows for elimination of one of the two). One must now conclude from (44) that the energy shift for $c\neq 0$ is ill-defined: by varying the remaining constant A the energy shift can have any value. The explanation is that the energy depends non-analytically on the fixed-point configurations S_0 (which can also be inferred from the divergence of the $k \neq 1$ energy of the $c=0$ problem). Apparently, a perturbational approach to the problem of calculating S_0 for $c \neq 0$ is inadequate. The physical meaning of restricting ourselves to $c=0$ is that we assume $\langle S_1 \rangle_D = \mathbf{0}$. For $\rho < 1$ this does not seem unrealistic; however, for $\rho = 1$ one does not even know whether $\langle S_0 \rangle_D = \mathbf{0}$ for the ground state, so the assumption $c=0$ may be too crude.

5. Discussion

In this paper we have tried to calculate the zero-temperature fixed points of the dual models introduced in [1], for the case where the distribution of input vectors has spherical symmetry. Finding the ground state was shown to be equivalent to finding the ground state of a field of Heisenberg spins having short-range ferromagnetic interactions in combination with long-range anti-ferromagnetic interactions. We have shown, by solving the periodic case and by showing that non-periodic fixed points exist with lower energy, that in general the ground state will not be periodic. The consequence is that one may not impose periodic boundary conditions, which makes the problem mathematically more difficult to solve: one is faced with coupled nonlinear partial differential equations.

In studying the problem with free boundary conditions we had to restrict our analysis in several ways. We expanded the short-range interaction in powers of the width and retained only the first non-trivial contribution. We studied only the case $p=3$ and made a symmetry ansatz for the solution. The fixed point obtained was found to have an energy which is significantly below the energetically most favourable periodic state and qualitatively describes the result of iterating the evolution equations of the system numerically.

The purpose of this paper was to show that the duality introduced in [1] can be used to study analytically the creation of topology conserving maps; it does not contain a full analysis of the physics contained in the corresponding equations. A next step might be to calculate the full phase diagram in terms of the parameters T, J and σ. It might also be interesting to drop the restriction to spherically symmetric distributions of the input vectors. In the phase diagram one might expect to find transitions between phases with different intrinsic dimensions of the field $\boldsymbol{\psi}(\boldsymbol{x})$.

Acknowledgments

We would like to thank Theo Ruijgrok and Harm Jonker for interesting discussions on this subject.

Appendix A. Notation

This appendix contains an overview of the notation introduced and a brief description of the main variables.

The field equations

$\boldsymbol{\psi}(\boldsymbol{x}, t) \in \mathbb{R}^p$	p dynamic fields ($\boldsymbol{x} \in D \subset \mathbb{R}^n$), bounded)
$\langle f \rangle_D \equiv \|D\|^{-1} \int_D \mathrm{d}\boldsymbol{x}\, f(\boldsymbol{x})$	spatial average
$(n \otimes f)(\boldsymbol{x})$	integral operator n operating on function f (eigenvalues of n: $\lambda \leqslant \lambda_{\max}$)
n_+	positive short-range part of n (of width σ)
J	strength of background inhibition in n
$\boldsymbol{\xi} \in \mathbb{R}^p$	randomly drawn vectors (probability distribution: ρ)
$C_{\mu\nu} \equiv \langle \xi_\mu \xi_\nu \rangle_{\boldsymbol{\xi}}$	covariance matrix (eigenvalues: $c_\mu \in [0, c_{\max}]$)
$F[\boldsymbol{\psi}]$	Liapunov function (free energy)
$H[\boldsymbol{\psi}]$	Hamiltonian (fixed-point energy: E)
$\boldsymbol{S}(\boldsymbol{x}) \equiv \lim_{t\to\infty} \boldsymbol{\psi}(\boldsymbol{x}, t)$	field of Heisenberg spins ($\|\boldsymbol{S}(\boldsymbol{x})\| = 1\ \forall \boldsymbol{x}$)

Expansion of field equations

$\varepsilon \equiv \frac{1}{4}\sigma^2$	expansion parameter: $\boldsymbol{S} \equiv \boldsymbol{S}_0 + \varepsilon^\eta \boldsymbol{S}_1 + \mathrm{O}(\varepsilon^{2\eta})$
$J \equiv \tilde{J}\varepsilon^\rho$ ($\rho > 0$)	scaling of background inhibition
$\phi(\boldsymbol{x}), \theta(\boldsymbol{x})$	polar angle fields:
	$p = 2$: $\boldsymbol{S}_0 = (\cos\phi, \sin\phi)$
	$p = 3$: $\boldsymbol{S}_0 = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)$

References

[1] Coolen A C C and Lenders L G V M 1992 Dual processes in neural network models. I. Neural dynamics versus dynamics of learning *J. Phys. A: Math. Gen.* **25** 2577

[2] Kohonen T 1982 *Biol. Cybern.* **44** 135

[3] Corben H C and Stehle P 1950 *Classical Mechanics* (New York: Wiley)

[4] Davis H T 1962 *Introduction to Nonlinear Differential and Integral Equations* (New York: Dover)

[5] Abramowitz M and Stegun I 1965 *Handbook of Mathematical Functions* (New York: Dover)