M P E J

Mathematical Physics Electronic Journal

ISSN 1086-6655
Volume 3, 1997

Paper 1
Received: Sep 24, 1996, Accepted: Mar 18, 1997
Editor: P. Collet

Dynamics in Binary Neural Networks with a Finite Number of Patterns.

Part 1. General Picture of the Asynchronous Zero Temperature Dynamics.

V.A. Malyshev[*†] F.M. Spieksma[‡§]

Abstract

We rigorously define and study the limiting dynamics for Pastur-Figotin-Hopfield models of neural networks with N nodes and p patterns in the (thermodynamic) limit $N \to \infty$, $p \equiv const.$ We study local and global properties of this limiting dynamics.

[*]This author would like to thank ESI for their hospitality and their financial support

[†]Postal address: INRIA, Domaine de Voluceau, Rocquencourt, B.P. 105, 78153 Le Chesnay, France

[‡]The research of this author has been supported by a fellowship from the Royal Netherlands Academy for Arts and Sciences

[§]Postal address: Dept. of Mathematics and Computer Science, University of Leiden, P.O.Box 9512, 2300 RA Leiden, Holland

1 Introduction

There are sufficiently many mathematical results concerning equilibrium states of Hopfield models, but results on the dynamics of these networks are scattered throughout the physical literature. Moreover, no rigorous exposition nor even a definition of the dynamics in the thermodynamic limit seem to exist. Note that there has been a lot of work on the equilibrium properties of such models, in order of increasing complexity: for $p = 1$ (Curie-Weißmodel), $p = \mathrm{const}$ ([7], [8], [1], [10]), $p = o(N)$ and $p = \alpha N$ for small α ([3], [2]).

As for the dynamics, only the case of $p = 1$ was sufficiently well understood, because in the overlap representation it can be reduced to a one-dimensional random walk (see for example [4] and references therein). Here we get a general picture of the dynamics for $p = \mathrm{const}$.

In Part 1 of this paper we consider asynchronous zero temperature dynamics. We first take the limit for the temperature $T \to 0$ and N fixed and then the thermodynamic limit $N \to \infty$. We use a convenient representation (random walk representation) for the dynamics, which is richer than the overlap representation usually used by physicists. In this representation one can easily pass to the thermodynamic limit. This representation seems to be known and used by physicists by the name “pool dynamics”, cf. [9], p.747.

The simple behaviour in the vicinity of fixed points was of course well understood. There were some results about fixed points themselves. But we know only one paper where a global picture of the limiting dynamics has been established: Procesi and Tirozzi [15] defined some cones in the overlap representation and showed that the dynamics behaves differently in different cones. We make this picture more complete. In particular we study the dynamics on the boundaries of these cones. We indicate different regions with respect to the dynamics behaviour: LLN region (law of large numbers region), zero velocity region, including not only fixed points but also some other points which we call “traps” and scattering (singular) region.

For the limiting dynamics the trajectories of the dynamical system are piecewise linear and they converge to a global or local minimum, or to traps. In the second part of the paper we will obtain estimates for the number of fixed points and for the “transient time”, i.e. the time to reach a neighbourhood of a fixed point or trap.

One of our main concerns in this paper was also to be sufficiently careful with many annoying technical details, usually neglected in physical presentations.

We finally would like to note that for general quadratic (negative definite) energy functions the results will be similar.

2 Definitions and Main Results

2.1 Zero temperature limit

Let $\Lambda = 1, \ldots, N$ be a finite volume. We consider the set $\boldsymbol{S}(\Lambda) = \{-1,1\}^{\Lambda}$ of configurations $\sigma = (\sigma_1, \ldots, \sigma_N)$ on Λ, where $\sigma_i \in \{-1,1\}$ denotes the spin at site i.

On this space we construct discrete time finite Markov chains $\mathcal{L}_{\beta,N} = \{\sigma^{\beta,N}(t)\}$ with the following transition mechanism (asynchronous sequential heat bath dynamics):

i) choose randomly a site i;

ii) erase the spin σ_i;

iii) choose a new spin σ'_i via conditional Gibbs probabilities with inverse temperature β and a quadratic energy function H_Λ given by

$$H_\Lambda(\sigma) = -\frac{1}{2N}\sum_{i,j\in\Lambda} \mathcal{J}_{ij}\sigma_i\sigma_j.$$

The Gibbs distribution on a finite set is thus given by

$$P_\Lambda(\sigma) = (Z_\Lambda)^{-1}\exp(-\beta H_\Lambda(\sigma)).$$

Note that this Gibbs distribution is a stationary distribution for the Markov chain $\mathcal{L}_{\beta,N}$.

Lemma 2.1 *For $\beta \to \infty$, $\mathcal{L}_{\beta,N}$ tends to a finite discrete time Markov chain $\mathcal{L}_N$ with transition probabilities given by i), ii) and the new spin σ'_i is equal to*

$$\sigma'_i = \operatorname{sgn}\big(\sum_{k\neq i} \mathcal{J}_{ik}\sigma_k\big),$$

if $\operatorname{sgn}(\sum_{k\neq i}\mathcal{J}_{ik}\sigma_k) \neq 0$, *and* $\sigma'_i = 1, -1$ *with probability* $1/2$, *if* $\operatorname{sgn}(\sum_{k\neq i}\mathcal{J}_{ik}\sigma_k) = 0$.

Here and in the sequel the sgn-function is defined as follows:

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

The dynamics defined in the lemma is called zero temperature dynamics. We would like to point out that it is not deterministic and it does not coincide on hyperplanes $\sum_{k\neq i}\mathcal{J}_{ik}\sigma_k = 0$ with the dynamics that is usually referred to as zero temperature dynamics. Indeed, the standard definition is not to change the spin in site i, if $\sum_{k\neq i}\mathcal{J}_{ik}\sigma_k = 0$, and site i has been selected, so we stand still. This precludes that stochasticity arises.

Next we consider a special choice of the parameters $\mathcal{J}_{ij}$.

Let ξ^μ, $\mu = 1,\ldots,p$, be given configurations on $\boldsymbol{S}(\Lambda)$. In the first part of the paper we assume p to be fixed as N tends to infinity. The configurations $\pm\xi^\mu$ will be called *patterns*. We choose

$$\mathcal{J}_{ij} = \sum_{\mu,\nu} c_{\mu,\nu}\xi_i^\mu\xi_j^\nu.$$

We get Hopfield networks if we choose

$$\mathcal{J}_{ij} = \sum_{\mu=1}^{p} \xi_i^\mu\xi_j^\mu,$$

according to Hebb's rule. Otherwise speaking

$$H_\Lambda(\sigma) = -\frac{1}{2N}\sum_{\mu=1}^{p}\left(\xi^\mu,\sigma\right)^2.$$

up to an additive constant.

2.2 Random Walk Representation and Scaling

Instead of an N-dimensional state description we can restrict to a 2^p-dimensional state description, which we will describe below. This representation generalises the "overlap representation" widely used in the study of equilibrium distributions [2]. Using this overlap representation of the process, the Markov property gets lost and so the dynamics for this cannot be defined. Our representation still contains all necessary information for defining a dynamics.

Divide Λ into subsets of sites $S_\alpha(N)$, where

$$\alpha = (\alpha_1,\ldots,\alpha_p)$$

with $(\alpha_1,\ldots,\alpha_p)$ a configuration from the set $\{+1,-1\}^p$, such that for all μ $\xi_i^\mu \equiv \alpha_\mu$, for all $i \in S_\alpha(N)$, and the sets $S_\alpha(N)$ are maximal sets with such property. There are at most 2^p such sets. We assume that there are exactly 2^p such sets, and we define $\mathcal{A} = \{1,-1\}^p$. Throughout the paper we will make the following assumption.

Assumption *For*

$$s_\alpha(N) = \#\{i \in S_\alpha(N)\}$$

the sequence $s_\alpha(N)/N$ converges, i.e. there exists $s = (s_\alpha)_\alpha$ with $\sum_\alpha s_\alpha = 1$, such that

$$s_\alpha = \lim_{N\to\infty}\frac{s_\alpha(N)}{N}.$$

Further we assume that $s_\alpha > 0$.

The assumptions are satisfied in the following examples.

i) The ξ^μ, $\mu = 1, \ldots, p$ are orthogonal.

ii) The ξ^μ can be random. In this case one can ask whether the assumption is satisfied a.s. This is the case when ξ_i^μ are mutually independent and take the value 1 with probability $0 < p^\mu < 1$. Hierarchical memory models are also in this class.

Let $Q(N)$ be the parallelepiped

$$Q(N) = \{(r_{\alpha=(\alpha_1,\ldots,\alpha_p)}) : 0 \le r_\alpha \le s_\alpha(N)\} \cap \mathbb{Z}^{2^p}.$$

We define a mapping $r : \{+1,-1\}^\Lambda \to Q(N)$ given by

$$r_\alpha(\sigma) = \#\{i \in S_\alpha(N) \, : \, \sigma_i = 1\}, \tag{1}$$

and so the αth component of the vector r represents the number of sites in $S_\alpha(N)$ where σ has spin 1.

Introduce the space scaling for our reduced representation:

$$x_\alpha = \lim_{N\to\infty} \frac{r_\alpha}{N},$$

provided that this limit exists. Under this scaling $Q(N)$ becomes the parallelepiped

$$Q = \times[0, s_\alpha] \subset \mathbb{R}^{2^p}.$$

We will study the time-space scaling limits for $\mathcal{L}_{\beta,N}$, as β, $N \to \infty$. Denote by $\{\rho^{\beta,N}(t)\}$ and $\{\rho^N(t)\}$ the "reduced state" processes corresponding to $\mathcal{L}_{\beta,N}$ and $\mathcal{L}_N$ respectively, and by $\rho^{\beta,N}(r,t)$, $\rho^N(r,t)$ the corresponding position of the processes at time t, when starting at the point $r \in Q(N)$, i.e.

$$\begin{aligned} \rho^{\beta,N}(r^N,t) &= r\big(\sigma^{\beta,N}(r^{-1}(r^N),t)\big) \\ \rho^N(r^N,t) &= r\big(\sigma^N(r^{-1}(r^N),t)\big). \end{aligned}$$

The mapping r is given in (1) and for $r^{-1}(r^N)$ we can take any σ with $r(\sigma) = r^N$. Note that $\rho^{\beta,N}(t)$ and $\rho^N(t)$ are Markov chains.

2.2.1 Overlap representation

Consider the mapping $m : S(\Lambda) \to \{-1,1\}^p$ with

$$m(\sigma) = \frac{1}{N}\{(\xi^\mu, \sigma)\}_\mu. \tag{2}$$

This representation does not completely respect the Markov property but approximately. We shall see in Section 2.5 that the semigroup property partially survives in the large N limit.

2.3 Finite N dynamics

In the sequel we will restrict to Hopfield models and to the case of zero temperature, i.e. $\beta = \infty$. In Part 2 of this paper we will indicate for what quadratic energy functions our results are still valid.

Definition 2.1 *A point $r^N \in Q(N)$ is called a fixed point (for $\rho^N(t)$) if*

$$\rho^N(r^N, 1) = r^N \quad \text{a.s.}$$

Lemma 2.2 *The following monotonicity property holds*

$$H_\Lambda(\rho^N(t+1)) \leq H_\Lambda(\rho^N(t)), \quad \text{a.s..}$$

Moreover,

$$H_\Lambda(\rho^N(t+2)) < H_\Lambda(\rho^N(t)),$$

with probability at least $1/N^2$, if $\rho^N(t)$ is not a fixed point, and if $s_\alpha(N) \neq 1$ for all α.

The intuition behind this lemma is quite simple. The conditional Gibbs distribution chooses points of smaller energy with big probability that will be 1 for zero temperature. The only problem is that one can move a sufficiently long time in the neighbourhood of for example global maxima, where the decrease in energy is very slow.

Lemma 2.3 *For all sufficiently large N the following statements are true.*

i) *The set of fixed points $\mathcal{P}^N$ is non-empty. They are the global and local minima of H_Λ on $Q(N)$ and they consist of a subset of the set of vertices of the parallelepiped $\times_\alpha[0, s_\alpha(N)] \cap \mathbb{R}^{2^p}$.*

ii) *There exists a time $T = T(N)$, such that for any initial point m^N*

$$\mathsf{P}\{\rho^N(r^N, t) \in \mathcal{P}^N\} = 1, \quad t \geq T.$$

Next we study the limiting dynamics for $N \to \infty$. To this end we need to introduce some new concepts and notation.

Let us write $c^\mu \in \{-1, 1\}^{2^p}$ for the configuration with α-th component $c^\mu_\alpha = \alpha_\mu$ and let $s(N)$ be the vector with α-th component $s_\alpha(N)$. Note that the vectors c^μ do not depend on ξ^μ and they are orthogonal, because all s_α are positive and hence all $s_\alpha(N)$ for N sufficiently large.

In the reduced state representation the energy can be written as the following quadratic polynomial.

Lemma 2.4

$$H_\Lambda(r) = -\frac{1}{2N} \sum_{\mu=1}^{p} \Big(4(c^\mu, r)^2 + (c^\mu, s(N))^2 - 4(c^\mu, r)(c^\mu, s(N)) \Big). \qquad (3)$$

It easily follows that H_Λ has a global maximum at the point

$$\frac{1}{2}(s_1(N), \ldots, s_{2^p}(N)).$$

This maximum is not unique, as we shall see below.

2.4 Limiting Dynamics

2.4.1 Existence

Consider the scaled limit of the energy function

$$\begin{aligned} H(x) &= \lim_{N\to\infty} \frac{1}{N} H_\Lambda(r^N) \\ &= -\frac{1}{2}\sum_{\mu=1}^{p}\Big(4(c^\mu, x)^2 + (c^\mu, s)^2 - 4(c^\mu, x)(c^\mu, s)\Big), \end{aligned} \tag{4}$$

for $r^N/N \to x$. Then

$$\frac{\partial}{\partial x_\alpha} H(x) = -2\sum_\mu c^\mu_\alpha (c^\mu, 2x - s).$$

Only on hyperplanes $(\partial/\partial x_\alpha)H(x) = 0$ the limiting dynamics (see below) may not be deterministic and scattering (cf. [13]) may occur.

For $\alpha \in \{+1, -1\}^p$ define

$$h^k_\alpha = \{x \in Q \,|\, \mathrm{sgn}\Big(\sum_\mu c^\mu_\alpha (c^\mu, 2x - s)\Big) = k\}, \quad k = -1, 0, 1.$$

The planes h^0_α are called *singular hyperplanes.*

Lemma 2.5 *Let $r^N \in Q(N)$ be a sequence with $r^N/N \to x \in Q$, and let $x \notin h^0_\alpha$. Then for all sufficiently small τ the following Euler limit exists*

$$\lim_{N\to\infty} \frac{1}{N} \rho^N_\alpha(r^N, [\tau N]) \overset{\mathfrak{D}}{=} \rho^\infty_\alpha(x, \tau),$$

for some random variable $\rho^\infty_\alpha(x,\tau)$. More precisely, $\rho^\infty_\alpha(x,\tau)$ is deterministic with

$$\frac{\partial}{\partial \tau}\rho^\infty_\alpha(x, \tau) = v_\alpha(\rho^\infty(x, \tau)),$$

for the piecewise linear function $v_\alpha : Q \to \mathbb{R}$ given by

$$v_\alpha(x) = (s_\alpha - x_\alpha)\mathbf{1}_{\{x \in h^1_\alpha\}} - x_\alpha \mathbf{1}_{\{x \in h^{-1}_\alpha\}}.$$

The latter piecewise linear vector function

$$v(x) = \{v_\alpha(x)\} : Q \to \mathbb{R}^{2^p},$$

has discontinuities on the singular hyperplanes.

Remark 2.1 *Instead of time scaling one could consider from the beginning the following asynchronous parallel dynamics. In each site i a Poisson arrival process of rate 1*

$$0 < t_{i,1} < t_{i,2} < \ldots$$

is defined, and all these processes are mutually independent. At time $t_{i,k}$ we change the spin at site i as has been prescribed in the above.

2.4.2 Singular regions and vector field

Let $\mathcal{F} \in \{-1, 0, 1\}^{\mathcal{A}}$, and denote by $\mathcal{F}_\alpha$ the α-component of $\mathcal{F}$. Define the region $B^{\mathcal{F}} \subset Q$ by

$$B^{\mathcal{F}} = \cap_\alpha h_\alpha^{\mathcal{F}_\alpha}.$$

Then $\cup_{\mathcal{F}} B^{\mathcal{F}} = Q$ and $B^{\mathcal{F}} \cap B^{\mathcal{F}'} = \emptyset$ if $\mathcal{F} \neq \mathcal{F}'$. Often we will also denote a region $B^{\mathcal{F}}$ by $\mathcal{F}$ when no confusion can occur. If $B^{\mathcal{F}} \subset \cup_\alpha h_\alpha^0$, then we shall call $B^{\mathcal{F}}$ a *singular region*. By $\dim(B^{\mathcal{F}})$ we shall denote the dimension of the region $B^{\mathcal{F}}$.

Remark 2.2 *The intersection $\cap_\alpha h_\alpha^0$ has dimension $2^p - p$ in Q. More specifically, it is given by*

$$\frac{1}{2}s + \{x \,|\, (c^\mu, x) = 0, \mu = 1, \ldots, p\}.$$

Since the global maxima of the limiting energy are exactly the points x with $(\partial/\partial x_\alpha)H(x) = 0$ for all α, it follows that $\cap_\alpha h_\alpha^0$ is exactly the set of all global maxima, and so this set has dimension $2^p - p$ in Q.

We will now recursively define a vector field on Q.

For $x \in \mathcal{F}$ with $x \notin \cup_\alpha h_\alpha^0$, define

$$V(x) = v(x).$$

We shall also use the system of functions $v^{\mathcal{F}}(x)$ and we shall identify

$$v^{\mathcal{F}}(x) = V(x)$$

iff $x \in \mathcal{F}$. To define the vector field V on the remaining part Q, i.e. on $\cup h_\alpha^0$, we need the following lemma.

Lemma 2.6 *Let α and a region $B^{\mathcal{F}}$ of maximum dimension be given. Define the function $v^{\mathcal{F}}(\cdot)$ on $h^0_\alpha \cap \bar{B}^{\mathcal{F}}$ by continuity. The function*

$$\operatorname{sgn}\Big(\sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}}(x))\Big)$$

is constant on $h^0_\alpha \cap \bar{B}^{\mathcal{F}}$.

Let $B^{\mathcal{F}}$ be contained in at most k singular hyperplanes, say in $h^0_{\alpha(l)}$, $l = 1\ldots,k$. Assume that $v^{\mathcal{F}'}$ has been defined for all $\mathcal{F}'$ that are contained in at most $k-1$ singular hyperplanes. Let $B^{\mathcal{F}'}$ be a region such that $B^{\mathcal{F}} \subset \bar{B}^{\mathcal{F}'}$.

Definition 2.2 *$B^{\mathcal{F}'}$ is called an ingoing region for $B^{\mathcal{F}}$ if for some (and hence for any) $x \in B^{\mathcal{F}}$,*

$$\mathcal{F}'_{\alpha(l)} \sum_\mu c^\mu_{\alpha(l)} (c^\mu, v^{\mathcal{F}'}(x)) \leq 0, l = 1, \ldots, k,$$

where $v^{\mathcal{F}'}(x)$ is defined by continuity; it is outgoing if for some $x \in B^{\mathcal{F}}$,

$$\mathcal{F}'_{\alpha(l)} \sum_\mu c^\mu_{\alpha(l)} (c^\mu, v^{\mathcal{F}'}(x)) \geq 0, l = 1, \ldots, k,$$

with strict inequality for at least one value l. Otherwise it will be called neutral.

The geometrical meaning of this definition is the following. Let $B^{\mathcal{F}'}$ be a face of maximum dimension with $\bar{B}^{\mathcal{F}'} \supset B^{\mathcal{F}}$.

How can one get immediately from $B^{\mathcal{F}'}$ to $B^{\mathcal{F}}$ if the velocity at any point $y \in B^{\mathcal{F}'}$ is $v^{\mathcal{F}'}(y)$? There should exist a point x in $B^{\mathcal{F}'}$, such that if we draw the line $x + \lambda v^{\mathcal{F}'}(x)$, $\lambda \geq 0$, this has a non-empty intersection with $B^{\mathcal{F}}$ (because the velocity is linear). But this only can hold if $B^{\mathcal{F}}$ contains a point y, such that the angle between the velocity vector $v^{\mathcal{F}}(y)$ and the normal on *all* hyperplanes $h^0_{\alpha(l)}$ in the direction of $B^{\mathcal{F}'}$ is at least $\pi/2$. But this is exactly the definition of $B^{\mathcal{F}'}$ being ingoing for $B^{\mathcal{F}}$.

In the same way: how can one immediately get from $B^{\mathcal{F}}$ to $B^{\mathcal{F}'}$? Then there should be a point $y \in B^{\mathcal{F}}$ such that the velocity $v^{\mathcal{F}'}(y)$ at y points inside $B^{\mathcal{F}'}$. But this is only true if the angle between $v^{\mathcal{F}'}(y)$ and the normal on *all* hyperplanes $h^0_{\alpha(l)}$ is at most $\pi/2$. This is the definition of $B^{\mathcal{F}'}$ being outgoing for $B^{\mathcal{F}}$.

Finally neutrality means that both of the above situations do not occur.

We can now complete our definition of the vector field V on Q.

Definition 2.3 *For $x \in B^{\mathcal{F}}$ set $V(x) = v^{\mathcal{F}}(x) = v^{\mathcal{F}'}(x)$, for any outgoing region $B^{\mathcal{F}'}$. If there is no outgoing region we define $V(x) = v^{\mathcal{F}}(x) \equiv 0$.*

Note that in general the vector field on singular regions need not be uniquely defined. The next lemma shows that the vector field can never be 0 on singular hyperplanes.

Lemma 2.7 *Each singular region has at least one outgoing singular region. Morever, $V(x) = 0$ for some $x \in \mathcal{F}$ implies that $\mathcal{F}$ is non-singular.*

We will introduce paths $\Gamma = \Gamma(t)$, which are continuous mappings $\Gamma : [0, T] \to Q$, with T possibly equal to ∞, such that

i) $\Gamma(t)$ belongs to $Q \setminus \cup_\alpha h_\alpha^0$, except possibly for a countable subset $\mathbf{T}(\Gamma)$ of Γ.

ii) in points of the same interval belonging to $[0, T] - \mathbf{T}$ where $\Gamma(t)$ runs through a region $B^{\mathcal{F}} \subset Q \setminus \cup h_\alpha^0$, Γ_x is linear with velocity $(d/dt)\Gamma(t) = v^{\mathcal{F}}(\Gamma(t))$.

The system of paths starting at $x \in Q$ is called an x-bundle of paths and we denote it by V_x; paths starting at x are denoted by Γ_x.

Let x be a point in the non-singular region $\mathcal{F}$. Then for t sufficiently small the path $\Gamma_x(t)$ will not cross any singular hyperplane and so

$$\Gamma_x(t) = s^{\mathcal{F}} - v^{\mathcal{F}}(x) \exp\{-t\} = s^{\mathcal{F}} + (x - s^{\mathcal{F}}) \exp\{-t\} = \rho^\infty(x, t), \qquad (5)$$

where

$$s_\alpha^{\mathcal{F}} = \begin{cases} s_\alpha, & \mathcal{F}_\alpha = 1 \\ 0 & \mathcal{F}_\alpha = -1, \end{cases}$$

by virtue of Lemma 2.5. Therefore, paths starting at $x \in \mathcal{F}$ are headed towards $s^{\mathcal{F}}$ and we shall call $s^{\mathcal{F}}$ the *quasi-attractor* of $\mathcal{F}$.

By virtue of the lemma below, it is sufficient to study the following set of paths: $V_x \subset \boldsymbol{\Gamma}$ if and only if Γ_x crosses only singular regions of co-dimension at most 1 in Q, for all $\Gamma_x \in V_x$.

For such paths, let us consider the velocity vector $V(\Gamma_x(t))$. Then each time Γ_x crosses a singular region, precisely two components of $V(\Gamma_x(t))$ are discontinuous in t (two, since $h_\alpha^0 = h_{-\alpha}^0$).

Lemma 2.8 *The set $\{x \in Q \,|\, V_x \subset \boldsymbol{\Gamma}\}$ contains an open subset dense in Q.*

Following [13] one can give the definition of stability of a path and prove that such paths are stable. One can also define deterministic and essentially deterministic systems and prove that the Hopfield model is essentially deterministic. Note that V_x contains a unique path for x in the above set.

Theorem 2.9 *For all x with $V_x \subset \boldsymbol{\Gamma}$ and all τ, the Euler limit exists, i.e. there is a random vector $\rho^\infty(x, \tau)$, such that*

$$\lim_{N \to \infty} \frac{1}{N} \rho^N(r^N, [\tau N]) \stackrel{\mathfrak{D}}{=} \rho^\infty(x, \tau),$$

for any sequence $r^N \in Q(N)$, with $r^N/N \to x$. More precisely, $\rho^\infty(x, \tau)$ is deterministic and equal to $\Gamma_x(\tau)$.

Remark 2.3 *The limiting dynamics only depends on the patterns ξ^μ by the values s_α.*

This form of the limiting dynamics can be obtained by a "glueing procedure", provided that on each finite time interval the paths Γ_x cross the singular regions only finitely often. The latter is true because of acyclicity of paths, which concept we will now define.

For a given path Γ_x enumerate $\mathbf{T} = \{\cdots, t_i, \cdots\}$. With this path we can associate the sequence of all regions

$$\mathcal{F}_i(\Gamma_x) = \mathcal{F} \qquad \Longleftrightarrow \Gamma_x(t) \in \mathcal{F},\, t \in (t_i, t_{i+1})$$

that Γ_x subsequently visits.

Definition 2.4 *The path Γ_x is called strongly acyclic if the following property holds. The path visits each region at most once, i.e. the sequence $\mathcal{F}_i(\Gamma_x)$ is finite and all $\mathcal{F}_i(\Gamma_x)$ are mutually different. The network is called (strongly) acyclic if for almost all x the paths $\Gamma_x \in V_x$ are (strongly) acyclic.*

Theorem 2.10 *All paths $\Gamma_x \in \boldsymbol{\Gamma}$ are strongly acyclic and hence the network is strongly acyclic.*

Problem 1 *Determine the form of the limiting dynamics for points x with $V_x \not\subset \boldsymbol{\Gamma}$.*

We will comment on this problem. For such points x for which V_x has a unique element, the limiting dynamics will indeed exist and will have the same form as in Theorem 2.9. However, if V_x contains more than one element, then this means that some path Γ_x crosses a singular region having more than one outgoing region, at the point y say. In y scattering will occur (cf. [13]) and with some probability (depending on y) the path continues along one of the outgoing regions. In particular this holds for initial points x where H has a global maximum.

Another problem with respect to such points x is, that the limiting dynamics has no natural unique definition: it depends on the sequence r^N/N. This can be easily understood from the case $p = 1$ and $\xi^1 = (+, +, \ldots, +)$. Then $r^N(\sigma) = \#\{i \,|\, \sigma_i = 1\}$ and $\mathcal{A}$ consists of 1 element, $\alpha = 1$ say. Hence $h^0_\alpha = 1/2$.

Let $r^N/N \to 1/2$, $N \to \infty$. If $r^N/N < 1/2 - 1/2N$ for all N, then with probability 0 there is a jump to a *bigger* state. It means that $r^N(r^N, t)/N \in h^{-1}_\alpha$ for all values t. It easily follows that

$$\lim_{N\to\infty} \frac{1}{N} r^N(r^N, [\tau N]) \stackrel{\mathfrak{D}}{=} \frac{1}{2}\exp\{-\tau\}.$$

Similarly if $r^N/N > 1/2 - 1/2N$ for all N, we have

$$\lim_{N\to\infty} \frac{1}{N} r^N(r^N, [\tau N]) \stackrel{\mathfrak{D}}{=} 1 - \frac{1}{2}\exp\{-\tau\}.$$

Finally, if

$$\frac{1}{2} - \frac{1}{2N} \leq \frac{r^N}{N} \leq \frac{1}{2} + \frac{1}{2N},$$

for all N, then the limiting dynamics is non-deterministic:

$$\lim_{N\to\infty} \frac{1}{N} r^N(r^N, [\tau N]) \stackrel{\mathfrak{D}}{=} X(\tau),$$

where $X(\tau) = (1/2)\exp\{-\tau\}$ or $1-(1/2)\exp\{-\tau\}$ with equal probability. This phenomenon is called "scattering" (cf. [13]).

2.4.3 Fixed points and traps

Next we shall introduce fixed points and study their properties.

Definition 2.5 *x is called a fixed point, if $V(x) \equiv 0$ (in the case that $V(x)$ has multiple values, this means that all values should be 0).*

Lemma 2.11 *On the set $\{y \in Q \,|\, V_x \subset \boldsymbol{\Gamma}\}$ any fixed point $x \in Q$ has the following property: for some sequence $r^N \in Q(N)$ with $r^N/N \to x$*

$$\lim_{N\to\infty} \frac{1}{N} r^N(r^N, \tau N) \stackrel{\mathfrak{D}}{=} x,$$

for all $\tau > 0$. Also vice versa, i.e. if the last property holds for some $x \in \{y \in Q \,|\, V_x \subset \boldsymbol{\Gamma}\}$, then this is a fixed point.

Conjecture 1 *This holds for all $x \in Q$.*

Problem 2 *Prove this conjecture.*

We will explain some ideas on this problem. The problem for proving this for points outside the set $\{x \in Q \,|\, V_x \subset \boldsymbol{\Gamma}\}$ is the following. One should construct a Lyapunov function in the neighbourhood of such points, which has some uniformly negative drift in this neighbourhood. The "bad" points in this set, however, are points with the following properties:

1. it is contained in a singular region with more than one outgoing region;
2. if the point is not an element of h^0_α, then the velocity along the α-direction is 0.

Such "bad" points are local maxima with respect to the coordinates along which the velocity is not equal 0. Therefore the energy cannot be used as a Lyapunov function, since the decrease in energy is arbitrarily close to 0.

Our suggestion is to use the quasi-attractors for the construction of a suitable Lyapunov function, for example the distance to these quasi-attractors, because a whole neighbourhood such points is attracted with some uniformly positive speed towards some set of quasi-attractors (what set will become clearer below).

Remark 2.4 *Under the conditions of the previous lemma, our fixed points are stable fixed points in standard terminology.*

Theorem 2.12 **i)** *The set of fixed points $\mathcal{P}$ is a subset of the vertices of Q. In particular, $\mathcal{P}$ is contained in the regions of maximum dimension and it is precisely the set of local and global minima of the energy H.*

ii) *There exists a set of vertices $\mathcal{T}$ with $\mathcal{P} \cap \mathcal{T} = \emptyset$, such that for any $\epsilon > 0$ there exists a finite time $T(\epsilon)$ with*

$$\mathsf{P}\{\inf_{y \in \mathcal{P} \cup \mathcal{T}} ||\rho^\infty(x,t) - y|| < \epsilon\} = 1$$

for all $t \geq T(\epsilon)$ and all initial points $x \in \{y \,|\, V_y \in \boldsymbol{\Gamma}\}$, where $||\cdot||$ is some norm on $\mathbb{R}^{2^p}$.

Remark 2.5 *As a consequence of this theorem there are no limiting cycles.*

Remark 2.6 *It follows directly from (5) that the time to reach a fixed point is always infinite. This is different from the case of finite N. It means that even though in the case of finite N the fixed points are reached in finite time, this time is larger than linear in N. Additionally, we have convergence to vertices that are no fixed points and hence no local/global minima of the energy on Q.*

Remark 2.7 *It is well-known and easy to see that the only global minima of H_Λ are the patterns $\pm\xi^\mu$, if the ξ^μ are orthogonal. If*

$$\frac{1}{N}(\xi^\mu, \xi^\nu) \to \delta_{\mu\,\nu},$$

then the scaled limiting vectors are the only global minima of the scaled limit of the energy.

From (5) it is clear that $\mathcal{T}$ must consist of quasi-attractors. The function of a quasi-attractor is determined by its location. We can distinguish the following cases.

i) $s^{\mathcal{F}} \in \mathcal{F}$. Then the path starting at $x \in \mathcal{F}$ will converge to $s^{\mathcal{F}}$ (although $s^{\mathcal{F}}$ will be reached only after infinite time). Since $v^{\mathcal{F}}(s^{\mathcal{F}}) = 0$, it follows that $s^{\mathcal{F}}$ is a fixed point of the vector field. Moreover, the cone $\mathcal{F}$ is invariant. We have the following result.

Lemma 2.13 *The set of fixed points for the vector field is given by*

$$\{s^{\mathcal{F}} \,|\, s^{\mathcal{F}} \in \mathcal{F}, \mathcal{F} \text{ non-singular}\}.$$

ii) $s^{\mathcal{F}} \in \overline{\mathcal{F}} \setminus \mathcal{F}$, where $\overline{\mathcal{F}}$ denotes the closure of $\mathcal{F}$, and so $s^{\mathcal{F}}$ is a point in some singular region. Also in this case the path starting at $x \in \mathcal{F}$ will converge to $s^{\mathcal{F}}$ and it will reach it only after infinite time. Clearly

$$\lim_{y \to s^{\mathcal{F}}, y \in \mathcal{F}} v^{\mathcal{F}}(y) = 0,$$

and so $s^{\mathcal{F}}$ can be viewed as a fixed point for the region $\mathcal{F}$. In particular, the closure $\overline{\mathcal{F}}$ of the cone $\mathcal{F}$ is invariant.

By Lemma 2.7 the velocity at $s^{\mathcal{F}}$ cannot be equal to 0. Hence, $s^{\mathcal{F}}$ is not a fixed point and so all paths starting at the point $s^{\mathcal{F}}$ will leave $\overline{\mathcal{F}}$ immediately. Such points $s^{\mathcal{F}}$ will be called *traps*.

Lemma 2.14 *The set $\mathcal{T}$ from Theorem 2.12 is exactly the set of traps, i.e.*

$$\{s^{\mathcal{F}} \mid s^{\mathcal{F}} \in \overline{\mathcal{F}} \setminus \mathcal{F}, \mathcal{F} \text{ non-singular}\}.$$

iii) $s^{\mathcal{F}} \notin \overline{\mathcal{F}}$. Then the path starting at $x \in \mathcal{F}$ will move in the direction of $s^{\mathcal{F}}$ till it hits the boundary $\overline{\mathcal{F}} \setminus \mathcal{F}$. After this, the path will move along some outgoing region of this boundary into the direction of the quasi-attractor corresponding to the outgoing region. Hence, none of $\mathcal{F}$ and $\overline{\mathcal{F}}$ are invariant.

In Section 3 we will give an example of the existence of traps.

Remark 2.8 *In all papers in zero temperature dynamics different conventions are used for the case when $\mathrm{sgn}(\sum_{j\neq i} \mathcal{J}_{ij}\sigma_j) = 0$: the spin in a site is not changed. This leads to a different picture, in particular the number of fixed points increases. The fixed points of the corresponding vector field are not only the fixed points of our vector field, but also the points x with the following properties. If $x \in \mathcal{F}$ then:*

i) *$\mathcal{F}$ is singular;*

ii) *if $\mathcal{F} \not\subset h^0_\alpha$, then $v^{\mathcal{F}}_\alpha(x) = 0$. This means that $x_\alpha = 0$ if $x \in h^{-1}_\alpha$, and $x_\alpha = s_\alpha$, if $x \in h^1_\alpha$.*

Examples of such points are global maxima and traps, but the set of such points can be much bigger.

2.5 Dynamics for the overlap representation for zero temperature

Similarly as for the random walk reduction we can study the processes associated with the overlap representation

$$\begin{aligned} m^{\beta,N}(m^N, t) &= m(\sigma^{\beta,N}(m^{-1}(m^N), t)) \\ m^N(m^N, t) &= m(\sigma^N(m^{-1}(m^N), t)), \end{aligned}$$

where the mapping m is given by (2) and for $m^{-1}(m^N)$ we can take some σ with $m(\sigma) = m^N$. It can be that for different σ with $m(\sigma) = m^N$, we get different associated overlap processes, contrary to the random walk reduction. This is illustrated by the example in this subsection. However, in the thermodynamic limit there is a unique overlap process associated with each initial overlap state, except for at most a set of 0 Lebesgue measure.

We can also study this overlap process as the image of the random walk process under the mapping

$$m^N(r^N) = \frac{1}{N}\{(c^\mu, 2r^N - s(N))\}_\mu,$$

for N finite and under the mapping

$$m(x) = \{(c^\mu, 2x - s)\}_\mu,$$

in the thermodynamic limit. Denote $M = m(Q)$, $m^{\mathcal{F}} = m(s^{\mathcal{F}})$.

For finite N and any β the overlap process is not Markovian. For this process to be Markovian, we need that in the random walk process the transitions from reduced states with the same overlap representation should be the same. At the end of this section we will give an example for zero temperature, where this is not the case. The same example is valid for positive temperature.

Therefore we will only study the dynamics for $N \to \infty$. It can be simply connected to the dynamics for the random walk representation by the representation of the reduced state that we will discuss now.

From Remark 2.2 it follows that the subspace $\cap h^0_\alpha$ is exactly the space $\{s/2 \oplus C^\perp\} \cap Q$, with C the p-dimensional subspace of $\mathbb{R}^{2^p}$ spanned by the vectors c^μ, $\mu = 1, \ldots, p$, and $C^\perp$ the orthogonal ($(2^p - p)$-dimensional) complement of C in $\mathbb{R}^{2^p}$. This space $C^\perp$ is exactly the space of vectors x in $\mathbb{R}^{2^p}$ with overlap vector $m(x) \equiv 0$.

Each vector $x \in Q$ has a unique decomposition

$$x = \frac{1}{2}s + y^\perp + y,$$

with $y, y^\perp \in C, C^\perp$ respectively. Considered as a mapping from $s/2 \oplus C$ to $\mathbb{R}^p$, the mapping m is bijective and so we can interpret $m : Q \to M$ as a projection mapping.

Consider now 2 points x, x' with the same overlap $m(x) = m(x')$. This means that they have decompositions,

$$x = \frac{1}{2}s + y_1^\perp + y, \quad x' = \frac{1}{2}s + y_2^\perp + y,$$

$y_i^\perp \in C^\perp$, $y_i \in C$. Hence $x - x' \in C^\perp$, and they are elements of the same region $\mathcal{F}$. This means that the sets $m(\mathcal{F})$ are mutually disjunct in M. Let us denote $m(\mathcal{F})$ also by $\mathcal{F}$. Clearly $\mathcal{F}$ is either non-singular in both Q and M or singular.

The vector field on Q induces the following vector field on M: for $\hat{x} \in \mathcal{F}$, $\mathcal{F}$ a region of maximum dimension, define

$$V^M(\hat{x}) = v^{M,\mathcal{F}}(\hat{x}) = m(v^{\mathcal{F}}(m^{-1}(\hat{x}))) = m(s^{\mathcal{F}} - m^{-1}(\hat{x})) = m^{\mathcal{F}} - \hat{x}.$$

This relation is equivalent to the following one, used in the physical literature (cf. [1], [15]): on $\mathcal{F}$

$$\dot{m}^{\infty} = m^{\mathcal{F}} - m^{\infty},$$

with $m^N/N \overset{\mathfrak{Q}}{\to} m^{\infty}$.

In turn, this vector field defines a dynamical system Γ^M on M that it is the image under the mapping m of the dynamical system on Q, i.e.

$$\Gamma^M_{\hat{x}}(t) = m(\Gamma_x(t)) \tag{6}$$

for any $x \in m^{-1}(\hat{x})$. Indeed, for small values of t

$$\begin{aligned} \Gamma^M_{\hat{x}}(t) &= m^{\mathcal{F}} - v^{M,\mathcal{F}}(\hat{x}) \exp\{-t\} \\ &= m^{\mathcal{F}} + (\hat{x} - m^{\mathcal{F}}) \exp\{-t\} \\ &= m(s^{\mathcal{F}}) + m(x - s^{\mathcal{F}}) \exp\{-t\} \\ &= m\big(s^{\mathcal{F}} + (x - s^{\mathcal{F}}) \exp\{-t\}\big) \\ &= m\big(s^{\mathcal{F}} + v^{\mathcal{F}} \exp\{-t\}\big) = m(\Gamma_x(t)). \end{aligned}$$

Suppose that at time t_0 the path Γ_x hits a singular region. Since all reduced states with the same overlap belong to the same region, it follows that all paths $\Gamma_{x'}$ hit this singular region at time t_0 for all x' with $m(x') = m(x)$. This shows (6).

We easily obtain the following conclusions. Fixed points of the dynamical system on Q correspond to fixed points of the dynamical system on M. Furthermore, $V_x \in \boldsymbol{\Gamma}$ iff $V_{x'} \in \boldsymbol{\Gamma}$ for x, x' having the same overlap. Hence

$$m(\{x \in Q \,|\, V_x \subset \boldsymbol{\Gamma}\})$$

is well-defined and consists precisely of the points in M, from which emanate only paths crossing at most singular hyperplanes (in M) of co-dimension 1. It contains an open subset dense in M and the dynamical system on this dense subset is acyclic and essentially deterministic. By the form of the dynamics in M also quasi-attractors and traps coincide in Q and M. However, note that vertices of Q are not necessarily vertices of M, although the reverse statement is true.

The following theorem now immediately follows from the previous results. We will only formulate the results in the thermodynamic limit, but for finite N similar results are valid.

Theorem 2.15 i) *For all* $\hat{x}$ *with* $\Gamma_x^M \in m(\boldsymbol{\Gamma})$*, and all* τ *the Euler limit exists, i.e. there exists a random vector* $m^\infty(\hat{x},\tau)$*, such that*

$$\lim_{N\to\infty} \frac{1}{N} m^N(m^N,[\tau N]) \overset{\mathfrak{D}}{=} m^\infty(\hat{x},\tau),$$

for any sequence $m^N \in m(Q(N))$*, with* $m^N/N \to \hat{x}$*. More precisely,* $m^\infty(\hat{x},\tau)$ *is deterministic and equal to* $\Gamma_{\hat{x}}^M(t)$*.*

ii) *The set of fixed points* $\mathcal{P}^M$ *is a subset of the vertices of* M*. They are contained in the non-singular regions in* M *and they are exactly the local/global minima of the energy* H *on* M*.*

iii) *For any* $\epsilon > 0$ *there exists a finite time* $T(\epsilon)$ *with*

$$\mathsf{P}\{\inf_{y\in\mathcal{P}^M\cup m(\mathcal{T})} ||m^\infty(\hat{x},\tau) - y|| < \epsilon\} = 1$$

for all $t \geq T(\epsilon)$ *and all initial points* $\hat{x} \in m(\{y \,|\, V_y \subset \boldsymbol{\Gamma}\})$*.*

Of course, all the above assertions hold on dense subsets of Q and M. For extending these to Q and M one should solve the following problem.

Problem 3 *Consider points* $x \in Q$ *on a singular region with more than one outgoing region. Let* $r^N/N \to x$ *and consider a limiting distribution of* $r^N(r^N,[tN])/N$*. Let* $m(r_1^N) = m(r^N)$*. Will* $r^N(r_1^N,[tN])/N$ *induce the same limiting distribution on* Q*?*

Example Let us consider the case $p = 3$. We take $s_\alpha(N) = N/2^p$ for N a multiple of 2^p, so that the patterns are *completely unbiased*. We have

$$\begin{aligned} c^1 &= (+,+,+,+,-,-,-,-) \\ c^2 &= (+,+,-,-,+,+,-,-) \\ c^3 &= (+,-,+,-,+,-,+,-). \end{aligned}$$

Let us enumerate $\mathcal{A}$ as follows $i \sim (c_i^1, c_i^2, c_i^3)$.

Since $m(s/2) \equiv 0$, we can write each $x \in Q$ uniquely as $x = y + y^\perp$, $y, y^\perp \in C, C^\perp$ respectively. It easily follows that in Q the hyperplane $C^\perp$ is given by

$$\left\{ \lambda_1 \begin{pmatrix} 1\\0\\0\\0\\0\\0\\0\\1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0\\1\\1\\0\\1\\0\\0\\1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0\\0\\0\\1\\1\\0\\0\\0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0\\0\\1\\0\\0\\1\\0\\0 \end{pmatrix} + \lambda_5 \begin{pmatrix} 0\\1\\0\\0\\0\\0\\1\\0 \end{pmatrix} : \lambda_i \in \mathrm{I\!R}, \forall i \right\},$$

and similarly for $Q(N)$. Hence the points

$$(s_1(N), 0, 0, 0, 0, 0, 0, 0)$$

and

$$(s_1(N), 1, 1, 0, 1, 0, 0, 1)$$

have the same overlap representation $(2s_1(N), 2s_1(N), 2s_1(N))/N$.

In the next step a state with overlap representation $(2s_1(N) + 2, 2s_1(N) + 2, 2s_1(N) + 2)/N$ can be reached only if the first coordinate increases by 1 or the last decreases by 1. Clearly, from state $(s_1(N), 0, 0, 0, 0, 0, 0, 0)$ this cannot occur. Let us check state $r^N = (s_1(N), 1, 1, 0, 1, 0, 0, 1)$. For the desired transition to occur the unique $(+1)$-spin in the region $S_8(N)$ should be changed to (-1). For the corresponding site, i say, we should have

$$\begin{aligned} -1 &= \operatorname{sgn}\big(\sum_{j\neq i} \mathcal{J}_{ij}\sigma_j\big) \\ &= \operatorname{sgn}\Big(\sum_{\mu} c_8^{\mu}(c^{\mu}, 2r^N - s(N)) - p\sigma_i\Big) \\ &= \operatorname{sgn}\Big(-3s_1(N) - 3\sigma_i\Big) = \operatorname{sgn}\Big(-3s_1(N)\Big). \end{aligned}$$

This is clearly true.

In the thermodynamic limit we have seen that this problem does not occur: all states with the same overlap still have the same overlap at any time t. The only explanation for this difference seems the following. During large time intervals (of order N) the average change in overlap is the same for all initial (reduced) states with the same overlap. In the thermodynamic limit all coordinates change synchronously, and so the change in overlap is the same. However, for finite N the coordinates of the reduced states change successively, so to say asynchronously, and therefore the change in overlap during small time intervals can be different.

2.6 Singular regions and domains of attraction

We can distinguish two different types of singular regions of co-dimension 1. A similar characterisation of singular regions of higher co-dimension is not necessary, since the set of paths entering such regions have Lebesgue measure 0.

A singular region of co-dimension 1 is called a *transient* region, if it has one ingoing and one outgoing region. It is called a *separatrix*, if it has exactly two outgoing regions. So, a transient region is crossed by paths entering it, and separatrices cannot be reached by any path, unless the path starts on the separatrix itself.

This characterisation is useful for studying domains of attraction of fixed points and traps.

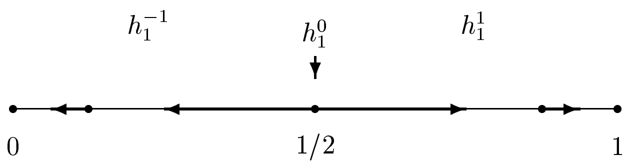

Figure 1 : $p = 1$

The *domain of attraction* of a certain fixed point or trap is the set of points, such that the paths starting at these points will converge to this fixed point or trap.

Then it is clear that the domains of attraction are convex bodies bounded by separatrices, and they form a *tiling* of Q (or M).

3 Examples

In some pictures we will show the dynamical systems for zero temperature for the cases $p = 1$, $p = 2$.

For the case $p = 1$ we will first choose $\xi^1 = (+1, +1, \ldots, +1)$. Then the set $\mathcal{A}$ consists of one element $\{\alpha = 1\}$ and in fact r_1^N denotes the number of sites with $(+1)$ spin. This description is equivalent to using overlap states.

The zero temperature dynamics in the thermodynamic limit is then given by

$$\rho^\infty(x,\tau) = \begin{cases} x\exp\{-\tau\}, & x < \frac{1}{2} \\ 1 + (x-1)\exp\{-\tau\}, & x > \frac{1}{2} \\ \frac{1}{2}\exp\{-\tau\},\ 1 - \frac{1}{2}\exp\{-\tau\}, & \text{with equal probability}, x = \frac{1}{2}. \end{cases}$$

The vector field is shown in Figure 1. We will also give the vector field for $p = 1$, using our formal 2-dimensional state description. Then $\mathcal{A}$ consists of two vectors with one component each, namely $(+)$, $(-)$, enumerated by 1,2. The sets $S_1(N)$, $S_2(N)$ are the blocks of sites, where ξ^1 is $+1$ and -1 respectively. It is seen that the intersection $h_1^0 \cap h_2^0$ has co-dimension 1, and this whole set has overlap 0.

Figure 3 shows the vector field for the case $p = 2$, where we take the following choice of ξ^1, ξ^2:

$$\begin{aligned} \xi^1 &= (+,+,\ldots,+,+,+,+,\ldots,+,+) \\ \xi^2 &= (+,+,\ldots,+,+,-,-,\ldots,-,-). \end{aligned}$$

This gives rise to a 2-dimensional state description. The set $\mathcal{A}$ now consists of the vectors $(+,+)$, $(+,-)$, which we will enumerate by 1 and 2. So $S_1(N)$ is

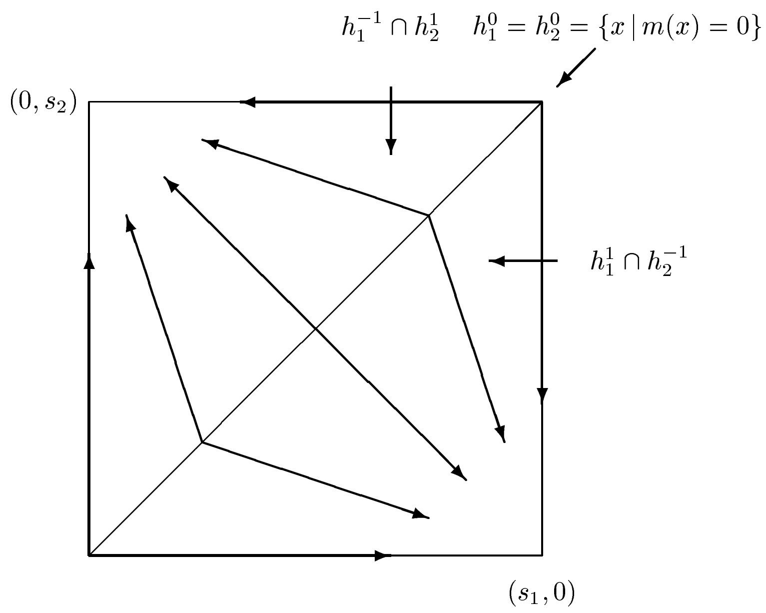

Figure 2: $p = 1$, formal random walk description

the block of sites, where both ξ^1 and ξ^2 are positive, and $S_2(N)$ is the block of sites, where ξ^1 and ξ^2 have opposite signs. Note that the singular hyperplanes are orthogonal. This is only true for $p = 2$, but not for higher dimensions.

Next we will give an example showing that traps exist.

Example Let $p = 4$. We again take the generic construction of *completely unbiased* patterns, i.e. $s_\alpha = s = 1/2^p$ for all α and

$$\begin{array}{rcl} c^1 & = & (+,+,+,+,+,+,+,+,-,-,-,-,-,-,-,-) \\ c^2 & = & (+,+,+,+,-,-,-,-,+,+,+,+,-,-,-,-) \\ c^3 & = & (+,+,-,-,+,+,-,-,+,+,-,-,+,+,-,-) \\ c^4 & = & (+,-,+,-,+,-,+,-,+,-,+,-,+,-,+,-). \end{array}$$

We will enumerate $\mathcal{A}$ as follows: $i \sim (c_i^1, c_i^2, c_i^3, c_i^4)$. Let us consider the quasi-attractor

$$s^{\mathcal{F}} = (0, s, s, s, s, s, s, s, 0, 0, 0, 0, 0, 0, 0, s)$$

of the region

$$\mathcal{F} = \{\cap_{k=1,9,10,11,12,13,14,15} h_k^{-1}\} \cap \{\cap_{l=2,3,4,5,6,7,8,16} h_l^1\}.$$

We should check that $\mathcal{F}$ is non-empty and that $s^{\mathcal{F}} \in \overline{\mathcal{F}} \setminus \mathcal{F}$. To this end, we need to determine the region to which $s^{\mathcal{F}}$ belongs.

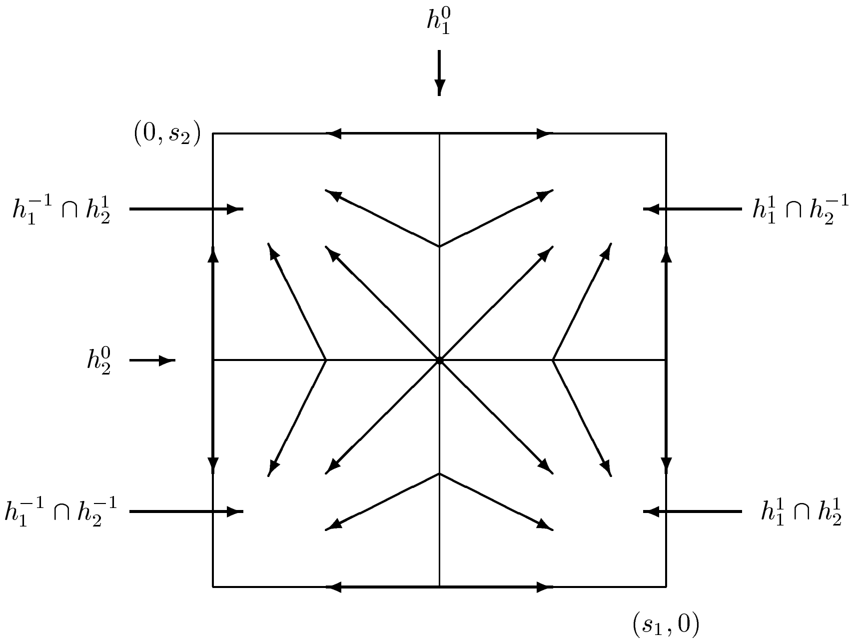

Figure 3: $p = 2$

Since $(c^\mu, s) = 0$ by complete unbiasedness, we have

$$h_\alpha^k = \{x : \mathrm{sgn}(\sum_\mu c_\alpha^\mu (c^\mu, 2x - s)) = k\} = \{x : \mathrm{sgn}(\sum_\mu c_\alpha^\mu (c^\mu, x)) = k\}.$$

We will calculate $(c^\mu, s^{\mathcal{F}})$:

$$(c^1, s^{\mathcal{F}}) = 6s, \quad (c^2, s^{\mathcal{F}}) = -2s, \quad (c^3, s^{\mathcal{F}}) = -2s, \quad (c^4, s^{\mathcal{F}}) = -2s.$$

Hence, we find that

$$\{\mathrm{sgn}(\sum_\mu c_\alpha^\mu (c^\mu, s^{\mathcal{F}}))\}_\alpha = (0, +, +, +, +, +, +, +, -, -, -, -, -, -, -, 0)$$

so that

$$s^{\mathcal{F}} \in \{\cap_{k=9,10,11,12,13,14,15} h_k^{-1}\} \cap \{\cap_{l=1,16} h_l^0\} \cap \{\cap_{m=2,3,4,5,6,7,8} h_m^1\},$$

which is a subset of $\overline{\mathcal{F}} \setminus \mathcal{F}$ if $\mathcal{F}$ is non-empty. But $\mathcal{F}$ is non-empty, since for all sufficiently small ϵ_1, $\epsilon_2 > 0$ with $2\epsilon_1 < \epsilon_2$ the point

$$\hat{s} = (\epsilon_1, s - \epsilon_2, s, s, s, s, s, s, 0, 0, 0, 0, 0, 0, 0, s) \in \mathcal{F}.$$

Indeed, we only have to check that $\hat{s} \in h_1^{-1}$. Since,

$$\begin{array}{ll} (c^1, \hat{s}) = 6s + \epsilon_1 - \epsilon_2, & (c^2, s^{\mathcal{F}}) = -2s + \epsilon_1 - \epsilon_2 \\ (c^3, s^{\mathcal{F}}) = -2s + \epsilon_1 - \epsilon_2, & (c^4, s^{\mathcal{F}}) = -2s + \epsilon_1 + \epsilon_2 \end{array}$$

we find that

$$\sum_\mu c_1^\mu (c^\mu, \hat{s}) = 4\epsilon_1 - 2\epsilon_2 < 0,$$

and so $\mathcal{F}$ is non-empty. Checking that $\mathcal{F}$ is non-empty is crucial. Indeed, the following point

$$s^{\mathcal{F}'} = (s, s, s, 0, 0, s, 0, 0, s, s, 0, s, s, 0, 0, 0)$$

would be a trap for the region

$$\mathcal{F}' = \{\cap_{k=1,2,3,6,9,10,12,13} h_k^1\} \cap \{\cap_{l=4,5,7,8,11,14,15,16} h_l^{-1}\},$$

if $\mathcal{F}'$ would be non-empty. However, it is easily checked that $\mathcal{F}'$ is empty.

It is simple to give an example of a quasi-attractor on a singular hyperplane, which nevertheless is not a trap:

$$s^{\mathcal{F}'} = (s, s, 0, 0, 0, 0, s, 0, s, 0, s, s, s, s, 0, 0)$$

with

$$\mathcal{F}' = \{\cap_{k=1,2,7,9,11,12,13,14} h_k^1\} \cap \{\cap_{l=3,4,5,6,8,10,15,16} h_l^{-1}\}.$$

The overlaps are given by

$$(c^1, s^{\mathcal{F}}) = -2s, \quad (c^2, s^{\mathcal{F}}) = 2s, \quad (c^3, s^{\mathcal{F}}) = 2s, \quad (c^4, s^{\mathcal{F}}) = 2s.$$

So

$$\{\mathrm{sgn}\big(\sum_{\mu} c^{\mu}_{\alpha}(c^{\mu}, s^{\mathcal{F}'})\big)\}_{\alpha} = (+,0,0,-,0,-,-,-,+,+,+,0,+,0,0,-),$$

and

$$s^{\mathcal{F}'} \in \{\cap_{k=1,9,10,11,13} h^1_k\} \cap \{\cap_{l=2,3,5,12,14,15} h^0_l\} \cap \{\cap_{m=4,6,7,8,16} h^{-1}_m\}.$$

Hence $\mathcal{F}' \subset h^1_7$, but $s^{\mathcal{F}'} \in h^{-1}_7$, so that $s^{\mathcal{F}'} \notin \overline{\mathcal{F}'} \setminus \mathcal{F}'$.

4 Proofs

4.1 Proofs for N finite

Proof of Lemma 2.1. The energy difference between spin $+1$ and spin -1 at site i is equal to

$$H_{\Lambda}(\sigma_{\Lambda-\{i\}}, 1) - H_{\Lambda}(\sigma_{\Lambda-\{i\}}, -1) = -\frac{2}{N}\sum_{k\neq i} J_{ki}\sigma_i,$$

where the symmetry of $\mathcal{J}_{ki}$ in i and k is used. It immediately follows for $\sigma'_i = 1, -1$ that

$$\lim_{\beta\to\infty} \mathsf{P}\{\sigma^{\beta,N}(t+1) = (\sigma_{\Lambda-\{i\}}, \sigma'_i) \,|\, \sigma^{\beta,N}(t) = \sigma, \text{ site } i \text{ is chosen}\} =$$
$$= \begin{cases} 1, & \sigma'_i = \mathrm{sgn}\big(\sum_{k\neq i} J_{ki}\sigma_k\big) \\ \frac{1}{2}, & \mathrm{sgn}\big(\sum_{k\neq i} J_{ki}\sigma_k\big) = 0. \end{cases} \tag{7}$$

∎

Hence, given that site i is chosen, the spin value that results in a configuration with the lowest energy is chosen deterministically in the limiting dynamics for $\beta = \infty$ outside hyperplanes where $\mathrm{sgn}\big(\sum_{k\neq i} J_{ki}\sigma_k\big) = 0$. The following lemmas make this more precise.

We need to specify the dynamics for $\beta = \infty$. The transition probabilities of the reduced Markov chain ρ^N are as follows. For $r_\alpha < s_\alpha(N)$

$$\mathsf{P}\{\rho^N(t+1) = r + \delta_\alpha \,|\, \rho^N(t) = r\} =$$
$$= \begin{cases} \frac{s_\alpha(N)-r_\alpha}{N}, & \text{if } \mathrm{sgn}\big(\sum_\mu c^\mu_\alpha(c^\mu, 2r - s(N)) + p\big) = 1 \\ 0, & \text{if } \mathrm{sgn}\big(\sum_\mu c^\mu_\alpha(c^\mu, 2r - s(N)) + p\big) = -1 \\ \frac{s_\alpha(N)-r_\alpha}{2N}, & \text{if } \mathrm{sgn}\big(\sum_\mu c^\mu_\alpha(c^\mu, 2r - s(N)) + p\big) = 0, \end{cases}$$

and for $r_\alpha > 0$

$$\mathsf{P}\{\rho^N(t+1) = r - \delta_\alpha \,|\, \rho^N(t) = r\} =$$
$$= \begin{cases} \frac{r_\alpha}{N}, & \text{if } \operatorname{sgn}\big(\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) - p\big) = -1 \\ 0, & \text{if } \operatorname{sgn}\big(\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) - p\big) = 1 \\ \frac{r_\alpha}{2N}, & \text{if } \operatorname{sgn}\big(\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) - p\big) = 0. \end{cases}$$

Proof of Lemma 2.2. If the spin at site i changes, then in terms of the σ-representation the difference in energy is given by

$$H(\sigma - 2\operatorname{sgn}(\sigma_i)\delta_i) - H(\sigma) = \frac{4}{N}\sum_{k\neq i} \mathcal{J}_{ki}\sigma_k \operatorname{sgn}(\sigma_i).$$

Hence,

$$\operatorname{sgn}\big\{H(\sigma - 2\operatorname{sgn}(\sigma_i)\delta_i) - H(\sigma)\big\} = \begin{cases} -\sigma_i \operatorname{sgn}(\sigma_i) < 0, & \operatorname{sgn}\big(\sum_{k\neq i}\mathcal{J}_{ki}\sigma_k\big) \neq 0 \\ 0, & \operatorname{sgn}\big(\sum_{k\neq i}\mathcal{J}_{ki}\sigma_k\big) = 0. \end{cases}$$

This implies the first monotonicity property in the Lemma.

Translated to the reduced state representation, the energy difference is strictly negative if

$$\operatorname{sgn}\big(\sum_{k\neq i}\mathcal{J}_{ki}\sigma_k\big) = \operatorname{sgn}\big(\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) - p\sigma_i\big) \neq 0, \tag{8}$$

for $r = r(\sigma)$, provided the spin at site i changes.

We will next show that from any point, which is not a fixed point, we jump with positive probability in at most two steps to a point with lower energy, provided that $s_\alpha(N) \neq 1$ for all α.

Clearly, if the spin changes at site $i \in S_\alpha(N)$, and if $0 < m_\alpha < s_\alpha(N)$, then the energy does not decrease only if

$$-p = \sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) = p,$$

yielding a contradiction. Hence, the only possibility for the energy remaining constant with probability 1 after a jump, is when r is a vertex. For this r the following should hold: for all α

$$\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) \leq -p, \quad \text{if } r_\alpha = 0 \tag{9}$$

$$\sum_\mu c^\mu_\alpha (c^\mu, 2r - s(N)) \geq p, \quad \text{if } r_\alpha = s_\alpha(N), \tag{10}$$

and there exists at least one value α for which there is equality in one of (9) and (10), because otherwise r is a fixed point.

Suppose that $r_\alpha = 0$ and (9) holds with equality. Then

$$\rho^N(r, 1) = r + \delta_\alpha$$

with probability $s_\alpha(N)/N$. From this point in one step a state of lower energy can be reached unless $r + \delta_\alpha$ is a vertex as well. This can only happen if $s_\alpha(N) = 1$, contradicting our assumption.

The same conclusion holds for α if $r_\alpha = s_\alpha(N)$, thus proving the second assertion of the Lemma.

Next we should give a lower estimate for the probability of jumping in at most two steps to a state with strictly lower energy. If for the α-coordinate we have that $s_\alpha(N) > r_\alpha > 0$ then it follows immediately that in one step one can step to a state of lower energy with probability at least

$$\frac{1}{2N} \min\{s_\alpha(N) - r_\alpha, r_\alpha\} \geq \frac{1}{2N}.$$

And so with at least this probability one can jump in two steps to a state of lower energy.

Let us assume that (9) and (10) are valid, but for at least one value α we have $r_\alpha = 0$ and there is equality in (9) (the other case is similar). Then with probability

$$\frac{s_\alpha(N)}{2N} \cdot \frac{s_\alpha(N) - 1}{N} \geq \frac{1}{N^2}$$

two successive times a -1-spin in the $S_\alpha(N)$-block is changed into a $+1$-spin, thus yielding a state with lower energy.

$\square$

Proof of Lemma 2.3. First we prove i).

It follows from the above proof that the fixed points should be a subset of the vertices of the parallelepiped $\times_\alpha[0, s_\alpha(N)] \subset \mathbb{R}^{2^p}$.

Clearly by the concavity of the function H_Λ, its local and global minima on the above parallelepiped should be vertices. Therefore they are points of $Q(N)$ and hence fixed points, since the energy cannot strictly increase by a jump.

If the vertex r is a fixed point, then the energy at all neighbouring points on the lattices $Q(N)$ should be strictly larger than the energy in r. If r is not a local/global minimum, then it easily follows that H_Λ has a local minimum in the set

$$\times[0, s_\alpha(N)] \cap \times[r_\alpha - 1, r_\alpha + 1]$$

which is a point in the interior of the parallelepiped. This is a contradiction and thus i) is proved.

We will prove ii). The fixed points are the only closed classes of the Markov chain ρ^N. Since ρ^N is a Markov chain on a finite state space, absorption into the closed classes takes place within finite time, uniformly over all initial states.

$\square$

Proof of Lemma 2.4. Clearly, for $r = r(\sigma)$

$$\begin{aligned}
\sum_{i,j} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j &= \sum_{\alpha,\alpha'} c_\alpha^\mu c_{\alpha'}^\mu \sum_{i \in S_\alpha(N)} \sum_{j \in S_{\alpha'}(N)} \sigma_i \sigma_j \\
&= \sum_{\alpha,\alpha'} c_\alpha^\mu c_{\alpha'}^\mu \Big(r_\alpha r_{\alpha'} - r_\alpha (s_{\alpha'}(N) - r_{\alpha'}) \\
&\qquad - r_{\alpha'}(s_\alpha(N) - r_\alpha) + (s_\alpha(N) - r_\alpha)(s_{\alpha'}(N) - r_{\alpha'}) \\
&= 4(c^\mu, r)^2 + (c^\mu, s(N))^2 - 4(c^\mu, r)(c^\mu, s(N)),
\end{aligned}$$

and the result immediately follows. □

4.2 Proofs for the thermodynamic limit

Euler limits

Proof of Lemma 2.5. Inside regions of maximum dimension the vector of mean drift from a point (and thus the vector field V) is smooth in the thermodynamic limit. The result then easily follows from a straightforward application of [6], Theorem 1.6.5. □

Proof of Theorem 2.9. As in the above proof, Kolmogorov's inequality and Kolmogorov's exponential bounds for sum of i.i.d. random variables is easily generalised to our case, because of smoothness of the drifts inside regions of maximum dimension.

Since we only consider trajectories passing region of co-dimension 0 or 1, the proof of this theorem follows analogously to the proofs for the existence of the Euler limit in the case of random walks on $\mathbb{Z}_2^+$ that are homogeneous on "faces", given in [12], for trajectories of the random walk crossing from a 2-dimensional face to a 1-dimensional face. □

Proof of Lemma 2.11. This follows immediately from Theorem 2.9. □

Construction of the vector field

Proof of Lemma 2.6 Let $x, y \in h_\alpha^0 \cap \bar{B}^{\mathcal{F}}$. From the definition of the vector field on $B^{\mathcal{F}}$ it easily follows that

$$v^{\mathcal{F}}(x) - v^{\mathcal{F}}(y) = y - x.$$

But since $x, y \in h^0_\alpha$, it follows that

$$\sum_\mu c^\mu_\alpha (c^\mu, y - x) = 0$$

and hence

$$\mathrm{sgn}\Big(\sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}}(x))\Big) = \mathrm{sgn}\Big(\sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}}(y))\Big).$$

$\square$

Properties of Paths

First we will show Lemma 2.7, namely the existence of outgoing non-singular regions.

Proof of Lemma 2.7. Since $h^0_\alpha = h^0_{-\alpha}$, $B^{\mathcal{F}}$ can be only contained in minimally 2 singular hyperplanes. Let us first assume that exactly two, $h^0_\alpha = h^0_{-\alpha}$ say, and let $x \in B^{\mathcal{F}}$. Let $\mathcal{F}' = \mathcal{F} + \delta_\alpha - \delta_{-\alpha}$ and $\mathcal{F}'' = \mathcal{F} - \delta_\alpha + \delta_{-\alpha}$ and suppose that both are not outgoing for $\mathcal{F}$. This means that

$$\begin{aligned} \sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}'}(x)) &\leq 0 \\ -\sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}''}(x)) &\leq 0. \end{aligned}$$

However, $v^{\mathcal{F}'}(x)$ and $v^{\mathcal{F}''}(x)$ agree on $\gamma \neq \alpha, -\alpha$ and $v^{\mathcal{F}'}_\alpha(x) = s_\alpha - x_\alpha$, $v^{\mathcal{F}''}_\alpha(x) = -x_\alpha$, $v^{\mathcal{F}'}_{-\alpha}(x) = -x_{-\alpha}$ and $v^{\mathcal{F}''}_{-\alpha}(x) = s_{-\alpha} - x_{-\alpha}$. Adding the two above expressions therefore yields that

$$\sum_\mu (c^\mu_\alpha)^2 s_\alpha + \sum_\mu (c^\mu_{-\alpha})^2 s_{-\alpha} = p(s_\alpha + s_{-\alpha}) \leq 0,$$

which is false. Hence $B^{\mathcal{F}}$ has an outgoing, non-singular region.

We will use induction to the number of singular hyperplanes contained in our singular region divided by 2. So let us assume that all singular regions contained in at most $2(M-1)$ singular hyperplanes have at least one outgoing non-singular region.

Suppose that $B^{\mathcal{F}}$ is singular and contained in $2M$ singular hyperplanes $h^0_{\alpha'(l)} = h^0_{-\alpha'(l)}$, $l = 1, \ldots, M$, but it has no outgoing non-singular region. Then "around" $B^{\mathcal{F}}$ there is a cycle of neutral non-singular regions. More precisely, the following cycle of regions $B^{\mathcal{F}(l)}$, $l = 0, \ldots, 4K$ for some $K \leq M$ can be constructed:

i) $B^{\mathcal{F}(2l)}$ is singular, and $B^{\mathcal{F}(0)} = B^{\mathcal{F}(4K)}$;

ii) $B^{\mathcal{F}(2l+1)}$ is non-singular;

iii) $B^{\mathcal{F}(2l+1)}$ is outgoing for $B^{\mathcal{F}(2l)}$ and ingoing for $B^{\mathcal{F}(2l+2)}$ $(\mathrm{mod}(2K))$;

iv) $\#\{\alpha : \mathcal{F}(2l+1) \neq \mathcal{F}(2l-1)\} = 2$;

v) $\{\alpha : \text{there exists } l \text{ such that } \mathcal{F}_\alpha(2l+1) \neq \mathcal{F}_\alpha(2l-1)\} \subset \{\alpha'(l), -\alpha'(l)\}_{l=1,\ldots,M}$.

Lemma 4.1 *Such cycle of regions cannot exist.*

Proof. Note that each singular hyperplane has to be passed an even number of times. Let us start by considering the following case:

i) $\mathcal{F}(2l-1)$, $\mathcal{F}(2l+2K+1) \subset h^1_{\alpha(l)} = h^{-1}_{-\alpha(l)}$, $\mathcal{F}(2l+1)$, $\mathcal{F}(sl+K-1) \subset h^{-1}_{\alpha(l)} = h^1_{-\alpha(l)}$;

iii) $\alpha(l) \neq \alpha(j)$ for $j \neq l$, $l, j = 0, \ldots, K-1$.

Hence we pass each singular hyperplane exactly twice.

For any region $B^{\mathcal{F}(k)}$ we can define the vector field $v^{\mathcal{F}(k)}$ by continuity in the point $s' = (s_1, \ldots, s_{2^p})/2$. Using similar arguments as in the proof of Lemma 2.6, we only need to consider the point s', since $s' \in \bar{B}^{\mathcal{F}(k)}$ for all $\mathcal{F}(k)$. Write

$$
\begin{aligned}
K_i &= \sum_\mu \sum_{\alpha \neq \alpha(l), l=1,\ldots,k} c^\mu_{\alpha(i)} c^\mu_\alpha v^{\mathcal{F}(0)}_\alpha(s'),\\
K_{-i} &= \sum_\mu \sum_{\alpha \neq -\alpha(l), l=1,\ldots,k} c^\mu_{-\alpha(i)} c^\mu_\alpha v^{\mathcal{F}(0)}_\alpha(s').
\end{aligned}
$$

Let $s'_j = s'_{\alpha(j)}$ and denote $c_{ij} = \sum_\mu c^\mu_{\alpha(i)} c^\mu_{\alpha(j)}$. Note that $\sum_\mu c^\mu_{-\alpha(i)} c^\mu_{\alpha(j)} = -c_{ij}$ and so on.

Then for $B^{\mathcal{F}(2l-1)}$ to be ingoing and $B^{\mathcal{F}(2l+1)}$ to be outgoing for $B^{\mathcal{F}(2l)}$, $l < K$, the following inequalities should hold:

$$
\begin{aligned}
K_l - \sum_{j<l} c_{lj}(s'_j + s'_{-j}) + \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_l &\leq 0; \qquad (11)\\
-K_l + \sum_{j<l} c_{lj}(s'_j + s'_{-j}) - \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_l &> 0; \qquad (12)\\
-K_{-l} - \sum_{j<l} c_{lj}(s'_j + s'_{-j}) + \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_{-l} &\leq 0; \qquad (13)\\
K_{-l} + \sum_{j<l} c_{lj}(s'_j + s'_{-j}) - \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_{-l} &> 0. \qquad (14)
\end{aligned}
$$

For $B^{\mathcal{F}(2l+2K-1)}$ to be ingoing and $B^{\mathcal{F}(2l+2K+1)}$ to be outgoing for $B^{\mathcal{F}(2l+2K)}$, $l < K$, we similarly have the inequalities:

$$\begin{aligned}
-K_l - \sum_{j<l} c_{lj}(s'_j + s'_{-j}) + \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_l &\leq 0; && (15)\\
K_l + \sum_{j<l} c_{lj}(s'_j + s'_{-j}) - \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_l &> 0; && (16)\\
K_{-l} - \sum_{j<l} c_{lj}(s'_j + s'_{-j}) + \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_{-l} &\leq 0; && (17)\\
-K_{-l} + \sum_{j<l} c_{lj}(s'_j + s'_{-j}) - \sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll} s'_{-l} &> 0. && (18)
\end{aligned}$$

Combination of (11), (13), (15) and (17), and similarly of (12), (14), (16) and (18) yields

$$\begin{aligned}
-2\sum_{j<l} c_{lj}(s'_j + s'_{-j}) + 2\sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll}(s'_l + s'_{-l}) &\leq 0; && (19)\\
2\sum_{j<l} c_{lj}(s'_j + s'_{-j}) - 2\sum_{j>l} c_{lj}(s'_j + s'_{-j}) + c_{ll}(s'_l + s'_{-l}) &> 0. && (20)
\end{aligned}$$

Clearly (20) follows from (19), since $c_{ll}(s'_l + s'_{-l}) > 0$, so we will only consider the equations (19). It follows that it suffices to prove that

$$Cx \leq 0, \tag{21}$$

has no non-negative solution $x \neq 0$ for the $K \times K$ matrix C defined by

$$\begin{aligned}
C_{ii} &= c_{ii} = p,\\
C_{ij} = -C_{ji} &= 2c_{ij} = 2c_{ji}, \qquad j > i.
\end{aligned}$$

To this end, we will use the following lemma.

Lemma 4.2 *Let C be a real $N \times N$ matrix with the following properties:*

i) *$C_{ii} \geq 0$ for all i;*

ii) *$C_{ji} = -C_{ij}$ for $j < i$.*

Then there is no non-negative solution $x \in \mathbb{R}^N$ to the inequality $(Cx)_i \leq 0$, $i = 1, \ldots, N$.

Proof. We shall show this in a recursive way. First we shall show that no strictly positive solution $x_i > 0$ exists.

$(Cx)_1 \leq 0$ implies that

$$C_{12} \leq -C_{11}\frac{x_1}{x_2} - \sum_{j>2} C_{1j}\frac{x_j}{x_2}.$$

From $(Cx)_2 \leq 0$, we get

$$C_{12} \geq C_{22}\frac{x_2}{x_1} + \sum_{j>2} C_{2j}\frac{x_j}{x_1},$$

and we find that

$$C_{22}\frac{x_2}{x_1} + \sum_{j>2} C_{2j}\frac{x_j}{x_1} \leq -C_{11}\frac{x_1}{x_2} - \sum_{j>2} C_{1j}\frac{x_j}{x_2}. \tag{22}$$

Next we will estimate C_{13} from the equation $(Cx)_3 \leq 0$:

$$-C_{13} < C_{23}\frac{x_2}{x_1} - C_{33}\frac{x_3}{x_1} - \sum_{j>3} C_{3j}\frac{x_j}{x_1},$$

and together with (22) we find

$$C_{22}\frac{x_2}{x_1} + \sum_{j>3} C_{2j}\frac{x_j}{x_1} \leq -C_{11}\frac{x_1}{x_2} - C_{33}\frac{{x_3}^2}{x_1 x_2} - \sum_{j>3} C_{1j}\frac{x_j}{x_2} - \sum_{j>3} C_{3j}\frac{x_j x_3}{x_1 x_2}. \tag{23}$$

Then using the estimate for $-C_{14}$ obtained from $(Cx)_4 \leq 0$, we similarly get

$$C_{22}\frac{x_2}{x_1} + \sum_{j>4} C_{2j}\frac{x_j}{x_1} \leq -C_{11}\frac{x_1}{x_2} - \sum_{i=3}^{4} C_{ii}\frac{{x_i}^2}{x_1 x_2} - \sum_{j>4} C_{1j}\frac{x_j}{x_2} - \sum_{i=3}^{4}\sum_{j>4} C_{ij}\frac{x_j x_i}{x_1 x_2}.$$

Continuing this procedure till $(Cx)_{N-1} \leq 0$, we get

$$\begin{aligned} C_{22}\frac{x_2}{x_1} + \sum_{j>N-1} C_{2j}\frac{x_j}{x_1} \quad &\leq \quad -C_{11}\frac{x_0}{x_1} - \sum_{i=3}^{N-1} C_{ii}\frac{{x_i}^2}{x_1 x_2} \\ &\qquad - \sum_{j>N-1} C_{1j}\frac{x_j}{x_2} - \sum_{i=3}^{N-1}\sum_{j>N-1} C_{ij}\frac{x_j x_i}{x_1 x_2}, \end{aligned}$$

and using the estimate $(Cx)_N \leq 0$ we finally obtain

$$C_{22}\frac{x_2}{x_1} < -C_{11}\frac{x_1}{x_2} - \sum_{i=3}^{N} C_{ii}\frac{{x_i}^2}{x_1 x_2}.$$

This is a contradiction, since C_{ii} and x_i are all positive. If some of the coordinates of x are allowed to be 0, then similar arguments apply, thus proving the lemma.

$\square$

The proof of Lemma 4.1 follows as a consequence.

$\square$

We will complete the proof of Lemma 2.7. However, it is easily checked that if assertions i) and ii) do not hold, we still get a similar system of inequalities as in the statement of Lemma 4.2. In particular, if assertion ii) does not hold then we should combine the inequalities corresponding to two successive times that we enter the hyperplane $h^0_{\alpha(l)}$. If the hyperplane $h^0_{\alpha(j)}$, $j \neq l$, is not passed between these two times, then this gives (l, j)-th and (j, l)-th entry equal to 0 in the matrix C. This proves that each singular region has at least one outgoing non-singular region.

Suppose that for some $x \in B^{\mathcal{F}}$, $B^{\mathcal{F}}$ singular, $V(x) = 0$ for one of the values of V. Then for some outgoing non-singular region $B^{\mathcal{F}'}$ of $B^{\mathcal{F}}$ we have $v^{\mathcal{F}'}(x) = 0$, where $v^{\mathcal{F}'}(\cdot)$ is defined by continuity on $B^{\mathcal{F}}$. Then for all $y \in B^{\mathcal{F}}$ we have

$$\mathcal{F}'_\alpha \sum_\mu c^\mu_\alpha(c^\mu, v^{\mathcal{F}'}(y)) = 0,$$

for all α with $h^0_\alpha \supset B^{\mathcal{F}}$, by virtue of Lemma 2.6. But then $B^{\mathcal{F}'}$ is not outgoing for $B^{\mathcal{F}}$, yielding a contradiction. $\square$

Note that an immediate consequence of Lemma 4.1 is strong acyclicity of all paths $\Gamma_x \in \boldsymbol{\Gamma}$, since such paths always satisfy the properties i) upto v) mentioned before the formulation of this lemma.

Proof of Theorem 2.10 This follows from Lemma 4.1. $\square$

Characterisation of fixed points for the limiting dynamics

First we will prove Theorem 2.12. To this end we use the following Lemma.

Lemma 4.3 *The energy decreases strictly along paths not indentically equal to a fixed point. More specificly, let x not be a fixed point and let a path Γ_x be given. Then for any $t > 0$ there exists a positive constant $c(t) = c(t, \Gamma_x)$, such that*

$$H(\Gamma_x(t)) \leq H(x) - c(t) \cdot t \cdot \min\big\{ \min_\alpha \{x_\alpha \mid x_\alpha \neq 0\}, \min_\alpha \{s_\alpha - x_\alpha \mid s_\alpha - x_\alpha \neq 0\}\big\}.$$

The problem with the uniform bound in the above lemma is mainly the following. The region $B^{\mathcal{F}}$ for $\mathcal{F} \equiv 0$ corresponds exactly to the global maxima of H. But although this region has an outgoing non-singular region, the decrease in energy is arbitrarily close to 0 as long as the path is close to $B^{\mathcal{F}}$.

Proof of Lemma 4.3. Note that for $x \in \mathcal{F}$

$$\lim_{h \downarrow 0} \frac{1}{h}\big(H(\Gamma_x(h)) - H(\Gamma_x(0))\big) = \sum_\alpha \frac{\partial}{\partial x_\alpha} H(x)\Big((s_\alpha - x_\alpha)\mathbf{1}_{\{\mathcal{F}_\alpha = 1\}} - x_\alpha \mathbf{1}_{\{\mathcal{F}_\alpha = -1\}}\Big). \tag{24}$$

By construction
$$\frac{\partial}{\partial x_\alpha} H(x) < 0$$
is equivalent to $x \in h_\alpha^1$, and
$$\frac{\partial}{\partial x_\alpha} H(x) > 0 \qquad (25)$$
is equivalent to $x \in h_\alpha^{-1}$, and so the left-handside of (24) is strictly negative in all points of non-singular regions except fixed points. There maybe points in singular regions, where the right-handside of (24) also equals 0, for example in the global maxima.

Let x be such point. By Lemma 2.7 each region has an outgoing non-singular region. Hence, it follows that there exists δ, such $\Gamma_x(\delta)$ is a point in a non-singular region. The necessary estimates easily follow. $\square$

Proof of Theorem 2.12 i). We first show that the local/global minima of H on Q are vertices, and they are not elements of the singular hyperplanes.

Suppose that $x \in \mathcal{F}$ is a local/global minimum. Then necessarily $v^{\mathcal{F}}(x) = 0$ by Lemma 4.3. Hence, x is a fixed point and by Lemma 2.7 $\mathcal{F}$ is not a singular region. Suppose that x is not a vertex, then $0 < x_\alpha < s_\alpha$ for some α. Hence $v_\alpha^{\mathcal{F}}(x) \neq 0$, thus yielding a contradiction. Hence x is a vertex.

Suppose next that there is a fixed point x, which is not a global/local minimum of the energy. By definition of the vector field, x must be a vertex. By virtue of Lemma 2.7 x belongs to a non-singular region, $\mathcal{F}$ say. By the form of the vector field, we have for $y \in \mathcal{F}$ that
$$v^{\mathcal{F}}(y) = v^{\mathcal{F}}(x) + x - y = x - y.$$
Further, we have that $v_\alpha^{\mathcal{F}}(y) = x_\alpha - y_\alpha < 0$ implies $\mathcal{F}_\alpha = -1$ and $v_\alpha^{\mathcal{F}}(y) > 0$ implies $\mathcal{F}_\alpha = 1$. But by definition, if $\mathcal{F}_\alpha = -1$ then $v_\alpha^{\mathcal{F}}(y) = -y_\alpha$, and if $\mathcal{F}_\alpha = 1$ then $v_\alpha^{\mathcal{F}}(y) = s_\alpha - y_\alpha$. Hence
$$\Gamma_y(t) = x - (x - y)\exp\{-t\},$$
so that x "attracts" the path starting at $y \in \mathcal{F}$. Since the energy decreases along paths by virtue of Lemma 4.3, necessarily $H(x) \leq H(y)$. As y was arbitrary and $\mathcal{F}$ is open in Q, we have found that x is a local or global minimum. $\square$

We shall now connect the location of fixed points to outgoing non-singular regions for $B^{\mathcal{F}}$, $\mathcal{F} \equiv 0$. Note that this region has no ingoing faces by virtue of Lemma 4.3. Also fixed points are vertices by Theorem 2.12 i) and elements of the set of non-singular regions.

Let $\alpha(l)$ be an enumeration of $\mathcal{A}$. Let us consider the point x defined by $x_{\alpha(l)} = 0$, $l \leq m$, $x_{\alpha(l)} = s_{\alpha(l}$, $l > m$. Then x is the attractor $s^{\mathcal{F}}$ for the region
$$\mathcal{F} = \cap_{l=1}^m h_{\alpha(l)}^{-1} \cap \cap_{l>m} h_{\alpha(l)}^1.$$

Lemma 4.4 *1. The following statements are equivalent.*

i) *The point x is a fixed point.*

ii) $x \in \mathcal{F}$.

iii) *$\mathcal{F}$ is strictly outgoing for the region $\mathcal{F}' \equiv 0$, i.e.*

$$\mathcal{F}_\alpha \sum_\mu c^\mu_\alpha (c^\mu, v^{\mathcal{F}}(s')) > 0,$$

for all $s' \in \cap h^0_\alpha$ and all α.

2. The following statements are equivalent for $\mathcal{F}$ not empty.

i) *x is not a fixed point, but for any $y \in \mathcal{F}$ we have*

$$\Gamma_y(t) \to x, \quad t \to \infty,$$

with $\Gamma_y(t) \in \mathcal{F}$ for all $t \geq 0$.

ii) $x \in \overline{\mathcal{F}} \setminus \mathcal{F}$.

iii) *$\mathcal{F}$ is outgoing but not strictly outgoing for $\mathcal{F}' \equiv 0$.*

This lemma gives an exact characterisation of traps and fixed points.

Proof of Lemma 4.4. We will prove the first equivalences. Assume i). Then x is a point in the non-singular region, $\mathcal{F}'$ say. By continuity of the velocity, we have $v^{\mathcal{F}'}_{\alpha(l)}(y) = -y_{\alpha(l)}$ for $l \leq m$ and $v^{\mathcal{F}'}_{\alpha(l)}(y) = s_{\alpha(l)} - y_{\alpha(l)}$ for $l > m$. It follows that $\mathcal{F}' = \cap_{l \leq m} h^{-1}_{\alpha(l)} \cap \cap_{l > m} h^1_{\alpha(l)}$. This shows ii).

Assume that ii) holds. Let $\hat{s}$ be given by $\hat{s}_{\alpha(l)} = -s_{\alpha(l)}$, $l \leq m$, and $\hat{s}_{\alpha(l)} = s_{\alpha(l)}$, $l > m$. Then at the point x

$$\sum_\mu c^\mu_{\alpha(l)}(c^\mu, 2x - s) = \sum_\mu c^\mu_{\alpha(l)}(c^\mu, \hat{s}) \begin{cases} < 0, & l \leq m \\ > 0, & l > m. \end{cases}$$

For the point $s' = (s_1, \ldots, s_{2^p})/2 \in \cap h^0_\alpha$ we have $v^{\mathcal{F}}(s') = \hat{s}/2$. Hence

$$\mathcal{F}_{\alpha(l)} \sum_\mu c^\mu_{\alpha(l)}(c^\mu, v^{\mathcal{F}}(s')) = \frac{1}{2} \sum_\mu c^\mu_{\alpha(l)}(c^\mu, 2x - s) > 0,$$

for all l, and so $\mathcal{F}$ is strictly outgoing for the region $\cap h^0_\alpha$. This shows that ii) implies iii).

Next assume iii). From the above relations it follows that $x \in \mathcal{F}$. Subsequently, it follows that $v^{\mathcal{F}}(x) = 0$ and so x is a fixed point. Hence iii) implies i).

The proof of the second equivalences is similar. ∎

Proof of Lemma 2.13. This follows immediately from Lemma 4.4. ◻

Proof of Theorem 2.12 ii) and Lemma 2.14. Let $\mathcal{F}$ be a region of maximum dimension, and suppose that its quasi-attractor $s^{\mathcal{F}} \notin \overline{\mathcal{F}}$. Further denote

$$I^{\mathcal{F}} = \mathcal{F} \cup \{\mathcal{F}' \,|\, \mathcal{F} \text{ is ingoing for } \mathcal{F}'\}$$

Then it is easy to see that there exists a finite time $t^{\mathcal{F}} < \infty$, such that for any $x \in I^{\mathcal{F}}$

$$t_x = \inf\{t \,|\, \Gamma_x(t) \notin I^{\mathcal{F}}\} \leq t^{\mathcal{F}}.$$

If $s^{\mathcal{F}} \in \overline{\mathcal{F}}$, then for any $\epsilon > 0$ there exists $t^{\mathcal{F}} = t^{\mathcal{F}}(\epsilon) < \infty$, such that for any $x \in I^{\mathcal{F}}$

$$\inf\{t \,|\, ||\Gamma_x(t) - s^{\mathcal{F}}|| < \epsilon\} \leq t^{\mathcal{F}}.$$

Define $\mathcal{T} = \{s^{\mathcal{F}} \,|\, s^{\mathcal{F}} \in \overline{\mathcal{F}} \setminus \mathcal{F}\}$. Then for the points $\{y \,|\, V_y \in \boldsymbol{\Gamma}\}$ there is a unique, acyclic path. Hence, for all $x \in \{y \,|\, V_y \in \boldsymbol{\Gamma}\}$

$$\inf_{y \in \mathcal{P} \cup \mathcal{T}} ||\Gamma_x(t) - y|| < \epsilon,$$

for $t \geq \sum t^{\mathcal{F}}$, where we take the summation over all $\mathcal{F}$ of maximum dimension. This set $\mathcal{T}$ cannot be chosen smaller because of Lemma 4.4. The assertions now follow immediately from Theorem 2.9. ◻

Proof of Lemma 2.8. Consider a non-singular region $\mathcal{F}$ with $s^{\mathcal{F}} \notin \overline{\mathcal{F}}$. We consider the set $I^{\mathcal{F}}$ defined in the foregoing proof and let for any $x \in I^{\mathcal{F}}$

$$y(x) = \Gamma_x(t_x),$$

with t_x as above. Clearly $y(x)$ is an element of a singular region for which $\mathcal{F}$ is ingoing. It is now simple to see that the set of points $x \in I^{\mathcal{F}}$ with $y(x)$ an element of a singular region of co-dimension at least 2, has co-dimension at least 1, and so this set has 0 Lebesgue measure. Let us denote this set by $O^{\mathcal{F}}$. There are finitely many such regions, and so the set $\cup O^{\mathcal{F}}$, where we take the union over all $\mathcal{F}$ of maximum dimension with $s^{\mathcal{F}} \notin \overline{\mathcal{F}}$, has 0 Lebesgue measure. Clearly, if $s^{\mathcal{F}} \in \overline{\mathcal{F}}$ there is no problem. Hence the set of points, the paths emanating from which hit a singular region after crossing one non-singular region, or which do not leave this singular region in finite time, has 0-Lebesgue measure.

For $\mathcal{F}$, let us now consider the points $x \in I^{\mathcal{F}} \cap \cup h_\alpha^0$, with $y(x)$ an element of a singular region of co-dimension at least 2. This set has co-dimension at least 2. Hence, similarly as in the above we can show that the set of points, the paths emanating from which hit a singular region for the first time after crossing exactly 2 non-singular regions, also has 0 Lebesgue measure.

We can continue this argument, and because of Lemma 4.1 no path can cross more than the total number of non-singular regions, before hitting a singular region of co-dimension at least 2.

Note, that in this recursive argument we did not take into account paths starting at a point in a singular region of co-dimension at least 2. But also this set of points has 0 Lebesgue measure, since it is contained in the union of the singular hyperplanes.

Thus we have a finite union of sets of 0 Lebesgue measure and the proof is completed.

$\square$

References

[1] D.J. Amit (1989) *Modeling Brain Function. The world of attractor neural networks.* Cambridge University Press, Cambridge.

[2] A. Bovier and V. Gayrard (1995) The retrieval phase of the hopfield model: A rigorous analysis of the overlap distribution. Preprint 161, Weierstraß-Institut für Angewandte Analysis und Stochastik.

[3] A. Bovier, V. Gayrard and P. Picco (1994) Large Deviation principles for the Hopfield and the Kac-Hopfield model. Preprint, Weierstraß-Institut für Angewandte Analysis und Stochastik.

[4] M. Cassandro, A. Galves, E. Olivieri and M.E. Vares (1984) Metastable behavior of stochastic dynamics: a pathwise approach. *J. Stat. Physics* **35**, 603–634.

[5] E. Domany, J.L. van Hemmen and K. Schulten (Edts.) (1991) *Models of Neural Networks.* Springer-Verlag, Berlin.

[6] S.N. Ethier and T.G. Kurtz (1986) *Markov Processes. Characterization and Convergence.* J. Wiley, New York.

[7] A.L. Figotin and L.A. Pastur (1977) Exactly solvable model of a spin glass. *Sov. J. Low Temp. Phys.* **3**, 378–383.

[8] A.L. Figotin and L.A. Pastur (1978) On the theory of disordered spin systems. *Theor. Math. Phys.* **35**, 404–414.

[9] W. Gerstner (1995) Time structure of the activity in neural network models. *Physical Review E* **15**, 738–758.

[10] J.J. Hopfield (1982) Neuronal networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558.

[11] W.A. Little and G. Shaw (1978) Analytic study of the memory storage capacity of a neural network. *Math. Biosci.* **39**, 281–290.

[12] V.A. Malyshev (1992) Stabilization laws for processes with a localized interaction. Rapport de Recherche No. 1635, INRIA.

[13] V.A. Malyshev (1993) Networks and dynamical systems. *Adv. Appl. Prob.* **25**, 140–175.

[14] Ch.M. Newman (1988) Memory capacity in Neural Netmwork models: rigorous lower bounds. *Neural Networks* **1**, 223–238.

[15] C. Procesi and B. Tirozzi (1990) Metastable states in the Hopfield model. *Intl. J. of Mod. Physics B* **4**, 143–150.

[16] L. Viana and A.C.C. Coolen (1993) Attraction domains in neural networks. *J. Phys. I France* **3**, 777–786.