Markov Processes Relat. Fields **1**, 141–161 (1995)

Markov Processes and Related Fields

Stochastic Evolution of Neural Networks

F.I. Karpelevich [1,*], **V.A. Malyshev** [2] and **A.N. Rybko** [3,†]

[1] Institute of Transport Engineering, Moscow, Russia
[2] INRIA, Domaine de Voluceau, Rocquencourt, BP105–78153, Le Chesnay Cedex, France
[3] Institute for Problems of Information Transmission, 19, B. Karetnyj per., Moscow, Russia

Received December 6, 1994

Abstract. We study the dynamics of neural networks introduced by M. Cottrell for the case of symmetric connections ($a_{ij} = a_{ji}$). We study the structure of images that can be remembered by these networks and prove convergence to these images starting from approximate images.

KEYWORDS: Neural networks, random walks in the orthant, exit boundaries, Euler limit, processes with local interaction

AMS SUBJECT CLASSIFICATION: 60K35, 60J25

1. Introduction

In this paper we study inhibition models for neural networks. The first paper to consider such models from a mathematical point of view is [1]. This paper is the basis of our analysis. We will not address the biological meaning of the problem (see references in [1]).

Inhibition models are slightly different from the well-known Hopfield–Little models (see [4], [6]) and also, but less, from the model studied by V. Kryukov (see [5]). There is no learning mechanism in the model under consideration, but there is a property that could be called "image restoration from a noisy image" and/or "recalling an object from its approximation".

Let us consider a (non-oriented) graph G with V the set of vertices and L the set of links. There are no loops (i.e. there is no link connecting i with itself). Denote by $D(i) = \{j : (j, i) \in L\}$ the neighbourhood of i. Note that $D(i)$ does not contain i itself.

*This research was partially supported by Grant M5E000 from the International Science Foundation.

†This research was partially supported by Grants: M5E000 from the International Science Foundation and 93-011-1470 from the Russian Foundation of Fundamental Research.

Consider the following Markov chain on the state space $\mathbf{R}_+^V$. At time t there is a potential $s_i(t)$ associated with point $i \in V$. If $s_i(t) > 0$ for all i, then the $s_i(t)$ decrease linearly in time with constant speed v_i until one of the potentials, $s_i(t+t_0)$ say, has become 0. At this moment, $t+t_0$ say, the $s_j(t+t_0+0)$ are all increased by a positive amount $\eta_{ij}(t+t_0+0)$ respectively, which are mutually independent random variables that have the same distribution for fixed i, j. Denote their mean by $a_{ij} = \mathsf{E}\eta_{ij}(t+t_0+0), j \in V$.

We will call this Markov chain $\mathcal{M}$. This chain is non-countable, but in fact the definitions for recurrence and ergodicity do not differ much from those for countable chains (see [10]). Note that many definitions from [7] and [8] can also be easily reformulated for this case. We assume some knowledge of these papers here.

Further we assume the existence of a density and some moments of the random variables $\eta_{ij}(t+t_0+0)$. All this is in no way compulsary and we assume this to avoid some unnecessary complications due to the Markov chain under consideration not being countable: for the difference between ergodicity and positive recurrence for processes on $\mathbf{R}_+^N$ etc. see [10].

For some properties like for example scattering phenomena, the assumption of an everywhere positive density seems necessary.

Note also that in for example [1] the random variables $\eta_{ij}(t+t_0+0)$ are independent and in particular their values are identical for all $j \neq i$. All these modifications are not important because the second vector field can be expressed in terms of the means $\mathsf{E}\eta_{ij}$ only. These modifications, however, do play a role in calculating the scattering probabilities.

The paper is organised as follows. In Section 2 we give the main definitions and we state the main terminological relationships with earlier papers. Also some simple examples are considered here: low dimensions, symmetric case etc.

Section 3.1 contains the main result, which treats the case of a symmetric connection matrix. In particular, we give algebraic sufficient conditions for ergodicity, which are easy to verify. We apply these results to one-dimensional graphs later in Section 3, where the conditions turn out to be necessary and sufficient. In Section 4 we introduce tensor products of networks. Using this technique we are able to get sufficient conditions for ergodicity and transitions and a partial description of the exit boundary (bounded harmonic functions) in the transient case for lattices of dimension at least 2. This gives some justification for the simulation results by M. Cottrell [1]. Section 5 finally considers a simple example of an infinite network.

2. Finite networks

2.1. Basic definitions and classification algorithm

Definition 2.1. By the W-restriction of $\mathcal{M}$ to the subset $W \subset V$ we mean the Markov chain $\mathcal{M}_W$ obtained from $\mathcal{M}$ by deleting $V \backslash W$. By deleting we mean that the states in the points of $V\backslash W$ do not influence the process in W anymore, in other words, we get the W-restriction by taking the potentials in points in $V\backslash W$ to be infinitely large.

We can think of our process $s(t) = (s_1(t), \ldots, s_N(t)) \in \mathbf{R}_+^N$ as a random walk in $\mathbf{R}_+^N$ and we use the terminology from [7]. In this terminology $\mathcal{M}_W$ is the induced chain for the face

$$\Lambda(W) = \{(x_1, \ldots, x_N) : x_i = 0, i \in W, x_j > 0, j \notin W\} \subset \mathbf{R}_+^N.$$

Put $W = W(\Lambda)$ for $\Lambda = \Lambda(W)$. In this way, Λ and W define each other uniquely.

Lemma-Definition 2.1. *Let $\mathcal{M}_W$ be ergodic and let $\pi_j^W, j \in W$, be equal to the mean number of times that the random trajectory $s_j^W(t)$ of the stationary Markov chain $\mathcal{M}_W$ hits 0 in the unit time interval. In other words:*

$$\pi_j^W = \lim_{T\to\infty} \frac{1}{T} \#\{t : s_j^W(t) = 0, t \in [0,T]\},$$

where $s_j^W(t)$ is the j-th coordinate of $\mathcal{M}_W$ at time t. Then the following flow (or traffic) equations hold

$$v_i = \sum_{j\in D(i)\cap W} a_{ji}\pi_j^W + a_{ii}\pi_i^W, \ i \in W, \tag{2.1}$$

or in shorthand notation (setting $\pi_i^W = 0$ for $i \notin W$)

$$P_W A\pi^W = P_W v, \tag{2.2}$$

where A is the transposed matrix (a_{ij}) and P_W the projection operator.

Proof. After a long time, T say, we have

$$s_j^W(T) = -v_i T + \sum_j a_{ji} \#\{t : s_j^W = 0, t \in [0,T]\} + o(T)$$

by the ergodicity assumption. □

Lemma-Definition 2.2. *If $\mathcal{M}_W$ is ergodic, then the second vector field v^W on the face $\Lambda(W)$ is defined by*

$$\begin{aligned} (v^W)_i &= -v_i + \sum_{j\in D(i)\cap W} a_{ji}\pi_j^W, \quad i \notin W, \\ (v^W)_i &= 0, \quad i \in W; \end{aligned} \tag{2.3}$$

in shorthand notation

$$v^W = -P_{V\backslash W}v + P_{V\backslash W}A\pi^W. \tag{2.4}$$

Then in the Euler limit we move locally along the (*ergodic*) *face* $\Lambda(W)$ *with constant speed* v^W, *that is, if at some time,* 0 *say, we are on the face* $\Lambda(W)$, *then for* τ *sufficiently small*

$$\frac{\partial m}{\partial \tau} = v^W,$$

where

$$m(\tau) = \lim_{N\to\infty} \frac{1}{N}\, s(\tau N).$$

The proof is quite standard and so we will not give it here (see [7] and references therein).

The main simplification of this model with respect to the general random walks in $\mathbf{Z}_+^N$ described in [7], is that we can calculate the second vector field from a finite system of linear algebraic equations. The problem is that this system possibly has solutions (spurious solutions) for *non*-ergodic faces as well. This means that we cannot get ergodicity conditions by using purely algebraic methods. One possibility to get the classification is by induction to the dimension. We define the inductive steps as follows.

- Assume that we already classified all induced chains of dimension less or equal to k. Then using formula (2.3) we calculate the second vector field for all ergodic faces of dimension at least $N-k$.
- Take some induced chain of dimension $k+1$. Try to check whether this chain is ergodic or not, by using some general results from [7] or other methods.
- If it is ergodic, then we calculate the second vector field for the corresponding face of dimension $N-k-1$, etc.

Remark 1. *For a given graph denote by* Erg *the subset of the parameter space, for which the chain is ergodic. This set is normally open and* $\det A$ *can be equal to* 0 *only on a subset of lower dimension. For points for which* $\det A = 0$ *and the chain is ergodic* (*do such cases exist??*) *the question arises how to calculate the "physical" solution of the flow equations. The answer is that it can be the limit of solutions for neighbouring points with* $\det A \neq 0$. *This is consistent with the expected continuity of the stationary probabilities with respect to the parameters. This has not been proved in general* (*however, for some general theorems see [2]*), *and hence this question should be considered separately for given examples.*

Definition 2.2. We will call $\mathcal{M}$ completely ergodic, if $\mathcal{M}$ and all $\mathcal{M}_W$ are ergodic (it then follows that all coordinates of the corresponding second vector fields are negative, provided that they are non-zero).

Theorem 2.1. *Inductive ergodicity and transience conditions.*
For the Markov chain $\mathcal{M}$ to be (completely) ergodic it is sufficient that the two following conditions hold:

(i) *All Markov chains $\mathcal{M}_W, \mid W \mid < \mid V \mid$, are completely ergodic;*
(ii) *for all $k \in V$ we have*

$$v_k > \sum_{j \in D(k)} a_{jk} \pi_j^{V \setminus \{k\}}.$$

For the Markov chain $\mathcal{M}$ to be transient it is sufficient that the two following conditions hold:

(i′) *The chain M_W is ergodic for some W;*
(ii′) for all $k \notin W$ we have

$$v_k < \sum_{j \in D(k)} a_{jk} \pi_j^W.$$

Proof. Condition (ii′) means that we go off to ∞ along some $(N-k)$-dimensional face. Under condition (ii) ergodicity follows from the general theory in [7]. $\square$

Remark 2. *An obvious sufficient condition for non-ergodicity in purely algebraic terms is as follows. If $\det A \neq 0$ and the (unique) solution of the flow equations is not positive then the system is not ergodic.*

The examples given below will illustrate the applicability of Theorem 2.1.

Proposition 2.1. (*High temperature region*). *Assume a_{ii}, v to be fixed. Then there exists $a > 0$ such that conditions* (i) *and* (ii) *of Theorem* 2.1 *hold for all $a_{ij} < a, i \neq j$.*

Remark 3. *The constant a will be specified in the corollary to Theorem* 3.1 *below.*

Proof of Theorem 2.1. The validity of the assertion is evident, because the sum in the right-hand side of the formula for the second vector field is small with respect to $\min_i v_i$. $\square$

Proposition 2.2. (*Symmetric case, see* [1] *and* [3]). *Assume that*

$$v_i = v,\ a_{ii} = a,\ a_{ij} = b, \quad i \neq j, a \neq b.$$

The Markov chain is ergodic if and only if $b < a$.

Proof. Under the assumption of the theorem ergodicity of the system implies that the π_is are all equal (start from the symmetric distribution; it will remain symmetric all the time) and for $\mid W \mid = k$ we have

$$\pi_i^W = \pi = \frac{v}{a + b(k-1)}.$$

This follows from the flow equations. Then condition (ii) of Theorem 2.1 is equivalent to

$$\pi bk = \frac{bk}{a + b(k-1)} v < v,$$

which holds for all parameters and all $b < a$. This gives ergodicity for $b < a$ by induction. If $b > a$ then the coordinates of the second vector field on any face are all positive, and so the system is transient by Proposition 1.2.3 of [7]. □

Case $a = b$ is more interesting. We do not consider it here.

The case of symmetric nodes $i = 4, 5, \ldots, N$ and arbitrary parameters for nodes 1,2 and 3 can be considered as well. But we cannot expect complete ergodicity without any further assumptions.

Theorem 2.2. *Under the conditions of Theorem* 2.1 *the Euler limit*

$$m(\tau) = \lim_{N \to \infty} \frac{1}{N} s_{[xN]}([\tau N])$$

exists for all $x \in \Lambda(W)$ *and all* $\tau > 0$*. Moreover, it is equal to* $x + v^W \tau$*, where*

$$v^W \equiv \big(-v_i + \sum_{j \in D(i) \cap W} a_{ji}\, \pi_j^W,\ i \notin W\big)$$

for all $x \in \Lambda(W)$ *and all* $\tau > 0$*, such that* $x + v^W \tau$ *still belongs to the face* $\Lambda(W)$*.*

Remark 4. *It follows that our "random walk" is acyclic (see [7]) under the conditions of Theorem* 2.1*. Moreover, in the ergodic case all vectors of the "second vector field" have negative coordinates only. Note that negativity of the coordinates of the second vector field on all faces often implies negativity of the coordinates of the first vector field (i.e. mean drifts). This allowed Marie Cottrell to obtain sufficient ergodicity conditions (see Proposition 3.2.3 in [1]) that are closely related to Theorem* 3.2*, by only using the first vector field. We refer to the remark following Theorem* 3.2*.*

<u>Dimension 2.</u>

Note that the one-dimensional chain is always ergodic. For two nodes, 1 and 2, and 6 parameters v_i, a_{ij} we have the following result.

Proposition 2.3. *The chain is ergodic if and only if*

$$v_1 \frac{a_{22}}{v_2} > a_{21}, \quad v_2 \frac{a_{11}}{v_1} > a_{12}.$$

The chain is transient if and only if there is at least one node i*, such that*

$$v_i \frac{a_{jj}}{v_j} < a_{ji}, \quad j \neq i.$$

Otherwise the chain is null recurrent.

Proof. Both coordinates of the drift vector are negative inside the quarter plane. When the point reaches axis 1 say (far away from axis 2), it jumps to a distance a_{22} from axis 1. The time to reach axis 1 again along the 2-direction is $t = a_{22}/v_2$. However upon jumping, the point also jumped a distance a_{21} upward along axis 1, and this amount should be less than tv_1 for the point to have a resulting movement towards the origin. This gives one condition. The second condition is argued similarly. The formal proof just repeats the proof for random walks in [2]. $\square$

<u>*Dimension 3.*</u>

We can obtain a complete classification for this case as well. This is analogous to random walks in $\mathbf{Z}_+^N$. However, we will only demonstrate the existence of cases, in which the network is not strongly acyclic (a random walk is called strongly acyclic if the corresponding dynamical system has the following property: starting from an ergodic face we go to 0 without intersecting non-ergodic faces). It is sufficient to construct an example of a negative 1-coordinate of v^{12} and a positive 3-coordinate of v^{123}. As a result we will necessarily hit the one-dimensional axis 2, but we will immediately leave it along the face 23.

We also want to show that a cycle is possible for dimension 3. Consider the second vector field on the two-dimensional faces. On face 23 we go from axis 2 to axis 3, if

$$v_2 \frac{a_{11}}{v_1} > a_{12}, \quad v_3 \frac{a_{11}}{v_1} < a_{13}.$$

Similarly on face 13 we go from axis 3 to 1, and on 12 from 1 to 2, if

$$v_3 \frac{a_{22}}{v_2} > a_{23}, \quad v_1 \frac{a_{22}}{v_2} < a_{21}$$

$$v_1 \frac{a_{33}}{v_3} > a_{31}, \quad v_2 \frac{a_{33}}{v_3} < a_{32}.$$

To construct a cycle we choose v_i, a_{ii} arbitrarily and we adjust the 6 remaining parameters $a_{ij}, i \neq j$, to satisfy all these inequalities.

<u>*Dimension 4 and more.*</u>

We will only mention the existence of parameters for which scattering phenomena (see [7]) occur, as for 4-dimensional random walks. We will see later that usually in the transient case many ways exist for going to infinity. It follows that with some positive probability we go to infinity in any possible way. It is more difficult to show that scattering exists in the ergodic case as well. We will not prove it here.

<u>*Compactification of* $\mathbf{R}_+^N$.</u>

For each W, including $\emptyset$, we add points

$$\Lambda_\infty(W) = \{(x_1, \ldots, x_N) : x_i \geq 0, i \in W, x_j = \infty, j \notin W\}.$$

We denote this compactification by Q_+^N, the convergence notion used being coordinate-wise convergence. The $\Lambda_\infty(W)$ are called faces at ∞. Define the evolution on these faces as follows: the coordinates equal to ∞ are invariant, whereas the evolution of the other coordinates occurs as in the Markov chain $\mathcal{M}_W$.

If the basic chain is transient, then this compactification can be used for a detailed study of its behaviour. In particular, we can define stationary states for this compactification. These are probability measures $\mu_\infty(W)$, with support on the face $\Lambda_\infty(W)$. Stationary states appear, when in the basic chain we go off to infinity along some face.

Scattering.

Start from some initial condition and let the face $\Lambda(V\backslash W)$ be an outgoing face. Then we converge to a convex combination $\sum\limits_W p_W \mu_\infty(W)$. If the p_W with $p_W \neq 0$, depend on the initial conditions then, as in [7], we will say that scattering phenomena occur.

3. Selfadjoint Operators

3.1. General Results.

Scalings.

The behaviour of the system is invariant with respect to the following scalings of the parameters.

- We leave a_{ij} unchanged, but we scale the other parameters

$$v \to v\alpha, \quad t \to \frac{t}{\alpha}.$$

  In this case the dynamical system does not change at all.

- We do not change t and v, but we take

$$a_{ij} \to \alpha a_{ij}$$

  and we scale the initial state $x^{(\alpha)}(0) = \alpha x(0) = \alpha x$. The new dynamical system is given by $m^{(\alpha)}(t) = \alpha m(t/\alpha)$.

Consider some finite graph G and let $a_{ij} = 0$ if i, j are not connected by a link. We assume v_i, a_{ii}, $a_{ij} = a_{ji}$ to be positive, but otherwise arbitrary. The resulting operator A is defined on $\mathbf{R}^N$. We further use the scalar product $(x, y) = \sum\limits_{i=1}^N x_i y_i$ in this space.

Theorem 3.1. *Assume that the operator A satisfies the two following conditions:*

- *A has a positive spectrum;*
- *the equation*
$$A\pi = \vec{v}$$
has a positive solution (*note that this solution is unique, because* $\det A > 0$ *by the first condition*).

Then the Markov chain is ergodic. Moreover, it has a linear Lyapunov function.

Proof. Consider the following linear function on $\mathbf{R}_+^N$:

$$f(x) = \sum_i x_i \pi_i.$$

If we can prove that for any ergodic face W and for $\epsilon > 0$ sufficiently small we have

$$f(x + \epsilon v^W(x)) - f(x) < -\delta,$$

then we can use Theorem 2.1 from [8] to prove that the Markov chain is ergodic. So, for any ergodic W we have to prove that

$$\sum_{i \notin W} \pi_i v_i^W = (\pi, v^W) < -\delta.$$

From (2.4) and (2.2) we get

$$v^W = -\vec{v} + A\pi^W.$$

As $v = A\pi$, we have

$$v^W = -\pi + \pi^W.$$

Hence

$$(\pi, A(-\pi + \pi^W)) = (\pi - \pi^W, A(-\pi + \pi^W)),$$

since v^W and π^W are perpendicular. Then

$$(\pi - \pi^W, A(-\pi + \pi^W)) = -(r, Ar) < 0,$$

where $r = \pi - \pi^W$. Clearly $r \neq 0$, as $\pi_x > 0$ for all x and $\pi_x^W = 0$ for $x \notin W$ by the definition of π^W. This completes the proof of the theorem, as there are only finitely many non-zero vectors r. □

Consider the family of matrices $A_\theta = A + \theta E$ with entries $a_{ij}(\theta) = a_{ij}$, if $i \neq j$, and $a_{ij} + \theta$ otherwise. For θ sufficiently large it is clear (see the high temperature region example above) that our process is ergodic. Define $\theta^* = \inf\{\theta :$ the process is ergodic $\forall\, \theta_1 > \theta\}$.

Next we state a simple corollary providing sufficient conditions for a face to be ergodic.

Corollary 3.1. *Let A be self-adjoint.*

- *If the conditions of Theorem* 3.1 *hold, then for the ergodicity of $\Lambda = \Lambda(W)$ it is sufficient that $\pi_i^W > 0$ for all $i \in W$. We would like to point out that this condition is "almost" necessary: only the case that $\pi_i^W = 0$ for some i and $\pi_j^W > 0$ for all other j is omitted.*
- *If the spectrum of $P_W A P_W$ is positive and Λ is ergodic, then there exists a non-negative linear Lyapunov function $f(y) = \sum_{i \in W} \pi_i^W y_i$ for the induced chain $\mathcal{M}_W$, $y_i \geq 0, i \in W$. This yields the ergodicity of the induced chain $\mathcal{M}_W$.*

Proof. If the spectrum of A is positive, then the spectrum of $P_W A P_W$ is also positive and so the conditions of Theorem 3.1 hold for the induced chain. This proves the first assertion. The two other assertions can be proved similarly to the proof of Theorem 3.1. □

Definition 3.1. The network with a self-adjoint corresponding operator A is said to be *in the generic situation*, if the vector $\vec{v}$ is not perpendicular to the eigenvector ξ_1 corresponding to the minimal eigenvalue of A.

Corollary 3.2. *If A is in the generic situation, then*

$$\theta^* = \inf\{\theta : \forall i,\ \forall \theta' > \theta, \pi_i(\theta') > 0\}$$

where $\pi_i(\theta)$ is the i th coordinate of the solution π_θ to the equation

$$A_\theta \pi_\theta = \vec{v}.$$

Proof. The eigenvectors of A_θ and A are equal and they form an orthonormal basis of $\mathbf{R}^N$. The eigenvalues $\lambda_i(\theta)$ of A_θ are given by

$$\lambda_i(\theta) = \lambda_i + \theta, \quad i = 1, \ldots, N,$$

for λ_i the eigenvalues of A. For $\vec{v}$ and π_θ satisfying

$$A_\theta \pi_\theta = \vec{v},$$

it follows that

$$\pi_\theta = \sum_{i=1}^N \frac{c^i}{\theta + \lambda_i} \xi_i,$$

where $c_i = (\xi_i, \vec{v}),\ i = 1, \ldots, N$, are the coefficients of the expansion of $\vec{v}$ into eigenvectors ξ_i of A:

$$\vec{v} = \sum_{i=1}^N c_i \xi_i.$$

So, if $c_i \neq 0$, then the vector function π_θ has a simple pole at the point $\theta = -\lambda_i$. In particular, π_θ has a simple pole at $\theta = -\lambda_1$. Consequently, at least one of the coordinates of the vector function π_θ changes sign in the neighbourhood of the point $\theta = -\lambda_1$. Let us define

$$\theta_0 = \inf\{\theta : \ \forall\, i, \ \forall\, \theta' > \theta, \ \pi_i(\theta') > 0\}.$$

Then $\theta_0 \geq -\lambda_1$. Futhermore, it is clear that $\theta^* \geq \theta_0$. On the other hand, if $\theta > \theta_0$ then both conditions of Theorem 3.1 hold. Consequently, the process is ergodic. So $\theta^* \leq \theta_0$. This completes the proof. □

Remark 5. *We do not know of any examples where the conditions of Theorem* 3.1 *are not necessary for the ergodicity of $s(t)$. In other words, do $\theta < \theta^*$ exist for which the process corresponding to A_θ is ergodic? Is it true that this is impossible for generic A?*

Consider the following class of graphs G. Let $r_i = \mid D(i) \mid$ and $r = \max_i r_i$. Further let $a_{ii} = a$ and $a_{ij} = 1$ for $i \neq j$ for which i, j are connected by a link: this is the "self-adjoint" case. We denote by $A = D + aE$ the operator defined by the right-hand side of equation (2.1) for $W = V$. Note that A and D are self-adjoint with respect to the scalar product $(f, g) = \sum f_i g_i$ in $\mathbf{R}_+^N$.

In the sequel we will assume that $v_i = v$. Furthermore, we assume that there is a maximal "cross" C in G: it is the set of vertices $D(i)$, for which $r_i = r$ and $j_1 \notin D(j_2)$ for any $j_1, j_2 \in D(i)$. For this class of graphs the following assertion holds.

Theorem 3.2. *For the Markov chain to be completely ergodic it is necessary and sufficient that $a > r$.*

Proof. Let $a \leq r$. We will consider the face Λ having coordinates equal to 0 in the vertices of the maximal cross and positive coordinates in all other vertices, including the centre of the maximal cross. In other words, $W = C$. We will show that the second vector field on this face has a positive coordinate in the centre of the maximal cross. In fact, from the definition of maximal cross it follows that the zero coordinates for this face are all isolated and that $\pi_j^W = va^{-1}$ for $j : x_j = 0$. Hence

$$v_i^W = -v + r\pi_j^W > 0$$

in the centre of the cross.

Next let $a > r$. Choose W arbitrarily and use Theorem 3.1 (we will omit the index W). All eigenvalues of Da^{-1} are inside the unit circle, since the row sums of this matrix are less than 1. Hence the spectrum of $Aa^{-1} = E + Da^{-1}$ is positive. We will prove for any W, that the solution π of the corresponding flow equations, which is given by

$$\begin{aligned}\pi &= va^{-1}(E+D/a)^{-1}\,\vec{1} = \Big[va^{-1}\sum_{1}^{\infty}\frac{(-1)^n}{a^n}D^n\Big]\,\vec{1}\\ &= va^{-1}\Big[\sum_{k=0}^{\infty}\frac{D^{2k}}{a^{2k}}(E-a^{-1}D)\Big]\,\vec{1},\end{aligned}$$

is positive. Indeed, by noting that all components of the vector $(E-a^{-1}D)\vec{1}$ are positive, we find that all terms in the right-hand side of the above equation are positive. □

Remark 6. *In the preceding theorem the condition that the a_{ij} only take the values $0,1$, can be weakened by choosing $r_i = \sum_{j:j\neq i} a_{ij}$ instead.*

Remark 7. *Cottrell [1] notes that the vectors of the first vector field (i.e. the mean drifts) for the embedded chain look inside the octant for each face, if*

$$a_{ij} = \begin{cases} 1, & \text{if } j \in D(i),\\ 0, & \text{if } j \neq i, j \notin D(i),\\ a, & \text{if } j = i,\end{cases}$$

and the number of neighbours in each point is less than a. Consequently the initial process $s(t)$ is ergodic.

3.2. One-dimensional interval

The formulations in this section can also be found in [3].

We will consider intervals $[1,N]$, where links only connect nearest neighbours. We will always assume $N>1$. Further we assume space homogeneity:

$$a_{ii}=a,\ v_i=1,\ a_{i,i+1}=a_{i,i-1}=1,\ i=2,\ldots,N-1.$$

Theorem 3.3.

1. *The system is completely ergodic if and only if $a>2$.*
2. *If N is odd and $a<2$, then the process is transient.*
3. *If N is even and $a<2\cos(\pi/(N+1))$, then the process is transient. If $a>2\cos(\pi/(N+1))$, then the process is ergodic.*
4. *The eigenvalues of the operator D are given by*

$$\lambda_l = 2\cos\frac{\pi l}{N+1},\ l=1,\ldots,N,$$

and the corresponding eigenvectors are

$$\xi_l = \{\sin \frac{\pi l x}{N+1},\ x = 1, \dots, N\},\ l = 1, \dots, N.$$

5. *The solutions of the flow equations are given by* (*here it is convenient to put* $\pi_x = \pi(x)$)

$$\pi(x) = L\Big(1 + \frac{r_1^x - r_2^x + r_1^{N+1-x} - r_2^{N+1-x}}{r_2^{N+1} - r_1^{N+1}}\Big),$$

where $L = (2+a)^{-1}$ *and* r_1, r_2 *are the solutions to the equation*

$$r^2 + ar + 1 = 0.$$

Note that we omit the case $a = 2$ for odd N and $a = 2\cos(\pi/(N+1))$ for N even.

Before passing to the proof, let us note that the theorem gives the classification of all ergodic faces, because any W can be represented as a union of connected intervals.

Proof of Theorem 3.3. The structure of the proof is as follows. We will determine a number a_{cr}, such that the network is ergodic by Theorems 3.1 or 3.2 for $a > a_{\mathrm{cr}}$. For $a < a_{\mathrm{cr}}$ it will appear that either $\pi(x)$ has negative values or there exists an ergodic face with the second vector field having positive coordinates only (in the sequel we will call such faces "traps"). The proof will then follow from Proposition 1.2.3 of [7].

The first assertion follows from Theorem 2.2, because $r = 2$ in the present case. We will prove assertions 4 and 5.

We can write the flow equations for the interval $[1, N] = \{1, 2, \dots, N\}$ in the following way:

$$1 = a\pi(x) + \pi(x+1) + \pi(x-1). \tag{3.1}$$

To solve these equations we put $\pi(0) = \pi(N+1) = 0$ and $\pi(x) = L + p(x)$, where $p(x)$ satisfies the homogeneous equation

$$p(x+1) + p(x-1) + ap(x) = 0,\ x = 1, \dots, N \tag{3.2}$$

with boundary conditions

$$p(0) = p(N+1) = -L. \tag{3.3}$$

Equation (3.1) implies $L = (2+a)^{-1}$. The function $p(x) = c_1 r_1^x + c_2 r_2^x$ with constant c_1, c_2 turns out to satisfy (3.2). Constants c_1, c_2 can be determined from (3.3). This proves assertion 5. Assertion 4 can be proved similarly.

We prove assertion 2 by spotting a trap. Let $N = 2k+1,\ a < 2$. The face $\Lambda(W)$, $W = \{i = 2j+1,\ j = 0, 1, \dots, k\}$, is a trap. In this case all flow equations are one-dimensional. Moreover, $\pi_i = a^{-1}$ for i odd and $v_i^W = 2a^{-1} - 1 > 0$ for i even.

Let us prove assertion 3. First we will show that the process is ergodic, if $a > 2\cos(\pi/(N+1))$. To this end we will check the conditions of Theorem 3.1.

Note that the minimal eigenvalue $s(D)$ of the operator D satisfies

$$s(D) = 2\cos\frac{\pi N}{N+1} = -2\cos\frac{\pi}{N+1} > -a.$$

We only have to verify strict positivity of $\pi(x)$ for all $x = 1, \dots, N$, but this is simple. Indeed, assertion 5 implies that

$$\pi(x) = L\left(1 - (-1)^x \frac{\sin((N+1)/2 - x)\alpha}{\sin(N+1)/2\alpha}\right), \quad x = 1, \dots, N,$$

where $\alpha = \arccos(a/2)$. Since $0 < \alpha < \pi/(N+1)$, this implies the positivity of $\pi(x),\ x = 1, \dots, N$.

Next assume that N is even and

$$2\cos\frac{\pi}{N-1} < a < 2\cos\frac{\pi}{N+1}. \tag{3.4}$$

Consider the face $\Lambda = \{2\}$, and so $W = \{1, 3, 4, \dots, N\}$. It is an ergodic face, since the chains corresponding to $W' = \{1\}$ and to $W'' = \{3, 4, \dots, N\}$ are ergodic (this has been proved in the first part of assertion 3). We will calculate the second vector field. It has only one component $v_2 = v_2^W$ with

$$v_2 = \pi^1(1) + \pi^{N-2}(1) - 1 = \frac{1}{a} + \frac{1}{2+a}\left(1 + \frac{\sin[\alpha(k-1-1+\frac{1}{2})]}{\sin[\alpha(k-1+\frac{1}{2})]}\right) - 1,$$

where we used the fact that the roots of the equation $r^2 + ar + 1 = 0$ are complex conjugate and are equal to $r_1 = -e^{i\alpha}, r_2 = -e^{-i\alpha}$, for $a < 2$. Note also that $a = 2\cos\alpha$, since $r_1 + r_2 = -a$. Consequently we can write

$$v_2 = \frac{1}{2\cos\alpha} + \frac{1}{2(1+\cos\alpha)}\frac{\sin(k-1/2)\alpha + \sin(k-3/2)\alpha}{\sin(k-1/2)\alpha} - 1.$$

Hence

$$v_2 = \frac{1}{2\cos\alpha} + \frac{2\cos(\alpha/2)\cdot\sin(k-1)\alpha}{2\cdot 2\cos^2\alpha/2\sin(k-1/2)\alpha} - 1.$$

Thus

$$v_2 = \frac{1}{2\cos\alpha} + \frac{\sin(k-1)\alpha}{2\cos(\alpha/2)\sin(k-1/2)\alpha} - 1.$$

Note that

$$\begin{aligned}\sin(k-1)\alpha &= \sin[(k-1/2)\alpha-\alpha/2]\\ &= \sin(k-1/2)\alpha\,\cos(\alpha/2)\,-\,\sin(\alpha/2)\cos(k-1/2)\alpha.\end{aligned}$$

Consequently,

$$v_2=\frac{1}{2\cos\alpha}+\frac{1}{2}-\frac{\sin(\alpha/2)\cos(k-1/2)\alpha}{2\cos(\alpha/2)\sin(k-1/2)\alpha}-1.$$

Further we have

$$v_2=\frac{1}{2\cos\alpha}-\frac{1}{2}\left(1+\frac{\sin(\alpha/2)\cos(k-1/2)\alpha}{\cos(\alpha/2)\sin(k-1/2)\alpha}\right).$$

Let us calculate the expression within brackets:

$$1+\frac{\sin(\alpha/2)\cos(k-1/2)\alpha}{\cos(\alpha/2)\sin(k-1/2)\alpha}=\frac{\sin k\alpha}{\cos(\alpha/2)\sin(k-1/2)\alpha}.$$

We get

$$v_2=\frac{1}{2\cos\alpha}-\frac{\sin k\alpha}{2\cos(\alpha/2)\,\sin(k-1/2)\alpha},$$

that is,

$$\begin{aligned}v_2 &= \frac{\cos(\alpha/2)\sin(k-1/2)\alpha-\sin k\alpha\,\cos\alpha}{2\cos\alpha\,\cos(\alpha/2)\,\sin(k-1/2)\alpha}\\ &= \frac{\sin k\alpha\,(\cos^2(\alpha/2)-\cos\alpha)-\cos(\alpha/2)\,\sin(\alpha/2)\,\cos k\alpha}{2\cos\alpha\,\cos(\alpha/2)\,\sin(k-1/2)\alpha}.\end{aligned}$$

Finally we get

$$\begin{aligned}v_2 &= \frac{\sin k\alpha\,\sin^2(\alpha/2)-\cos(\alpha/2)\,\sin(\alpha/2)\,\cos k\alpha}{2\cos\alpha\,\cos(\alpha/2)\,\sin(k-1/2)\alpha}\\ &= \frac{-\sin(\alpha/2)\,\cos(k+1/2)\alpha}{2\cos\alpha\,\cos(\alpha/2)\,\sin(k-1/2)\alpha}.\end{aligned}$$

We have $\alpha>\dfrac{\pi}{N+1}=\dfrac{\pi}{2k+1}$, as $a<2\cos\dfrac{\pi}{N+1}$. Then by (3.4) we find that

$$(k+\frac{1}{2})\alpha>\frac{\pi}{2},\ \ \alpha(k-\frac{1}{2})<\frac{\pi}{2}\ \text{and}\ \cos(k+\frac{1}{2})\alpha<0.$$

It follows that $v_2>0$. □

3.3. Traps for the interval

A face with all coordinates of the corresponding second vector-field positive will be called a *trap*. In other words, a trap is any face along which we can go to infinity.

Assume that the system on the interval is not ergodic. We will describe all traps. Denote an arbitrary face by a sequence of ones (1s correspond to positive coordinates) and zeroes (0s correspond to zero coordinates).

Proposition 3.1.

1. *If $a < 1$, then there is a one-one correspondence between traps and sequences $\{\tau_i, i = 1, \ldots, N\}, \tau_i = 0, 1,$ with the following properties.*

   a) *If $\tau_i = 1$ then either $\tau_{i-1} = 0$ or $\tau_{i+1} = 0$. In other words, more than two ones in a row is impossible.*

   b) *If $\tau_i = 0$ then $\tau_{i-1} = \tau_{i+1} = 1$. In other words, two zeroes in a row is impossible.*

   c) *Two ones in a row at the end-points of the interval is impossible, i.e. any of $\tau_1 = \tau_2 = 1$ and $\tau_{N-1} = \tau_N = 1$ is impossible.*

2. *Let $a^\star = 2$ for N odd and $a^\star = 2\cos(\pi/(N+1))$ for N even.*
   Let $1 \le a < a^\star$ and let $\beta = [1/2(\pi\alpha^{-1} - 1)]$, where $\alpha = \arccos(a2^{-1})$ ($[\cdot]$ stands for "integer part"). Then there is a one-one correspondence between traps and sequences $\{\tau_i, i = 1, \ldots, N\}$, $\tau_i = 0, 1,$ with the following properties.

   a) $\tau_1 = \tau_N = 0$.

   b) *All ones are isolated: if $\tau_i = 1$, then $\tau_{i+1} = \tau_{i-1} = 0$.*

   c) *The distance between the two closest ones is 2 or $2\beta + 1$, i.e. the number of subsequent zeroes is 1 or at least 2β.*

   d) *If there are 2β zeroes in a row, i.e.*

   $$\tau_{i_1} = \tau_{i_1+1} = \ldots = \tau_{i_1+2\beta-1} = 0,$$

   then

   $$\tau_{i_1-3} = \tau_{i_1+2\beta+2} = 1,$$

   if $i_1 - 3 > 1$ and $i_1 + 2\beta + 2 < N$.

   These restrictions constitute all possible restrictions on the sequences $\{\tau_i\}$.

Proof. Theorem 3.3 describes all ergodic faces: if $a < 1$, then the sequence $\{\tau_i\}$ corresponds to an ergodic face if and only if each zero is isolated (condition 1b)). It is easy to see that $v_i^W > 0$ only if 1a) and 1c) are valid.

If $1 < a < a^\star$, then the ergodic faces all correspond to sequences containing isolated zeroes and subsequences of even length L with $2\cos(\pi/L+1) < a$. This can be easily proved by induction.

Let us find necessary and sufficient conditions for all coordinates of the second vector field on an ergodic face to be positive. Note that if $\tau_i = \tau_{i+2} = 0$ and $\tau_{i-1} = \tau_{i+1} = \tau_{i+3} = 1$, then $v_{i+1} = 2/a - 1 > 0$. To this end, it is necessary that either

$$\pi^{(L)}(1) + \pi^{(L')}(1) > 1 \tag{3.5}$$

for L and L' even, or

$$\frac{1}{a} + \pi^{(L)}(1) > 1 \tag{3.6}$$

($\pi^L(1)$ stands for the first coordinate of the vector $\pi(x)$ for $N = L$).

This is because in a trap for $1 < a < a^\star$ there cannot be two ones in a row, otherwise we would have an ergodic odd interval for $a < 2$. Using the calculation of $\pi(x)$ in Theorem 3.3 we get

$$\pi^{(L)}(1) = \frac{1}{2+a}\left(1 + \frac{\sin(\alpha\dfrac{L-1}{2})}{\sin(\alpha\dfrac{L+1}{2})}\right).$$

Finally verifying the equations above explicitly, it is not difficult to see that inequality (3.5) is impossible and inequality (3.6) is only possible when $L/2$ is the integer part of $(\pi/\alpha - 1)/2$. □

3.4. One-dimensional circle

The one-dimensional circle with $N > 1$ points is the interval $[0, N]$, where the points 0 and N are identified.

Theorem 3.4.

1. *If $a > -2\cos\frac{2\pi[N/2]}{N}$, then the process is ergodic. Note that this condition is equivalent to $a > 2$ for N even.*

2. *If $a < -2\cos\frac{2\pi[N/2]}{N}$, then the process is transient. Note that this condition is equivalent to $a < 2$ for N even.*

Proof. The calculations are even easier in this case: $\pi(x) \equiv 1/(2a)$ and $\vec{1}$ is the eigenvector of D. The spectrum is

$$\lambda_l = 2\cos\frac{2\pi l}{N},\ l = 0, 1, \ldots, N-1.$$

So assertion 1 follows from Theorem 2.1.

To prove assertion 2, we first consider the case $N = 2k$ and we prove that the process is transient, if $a < 2$. As the initial state for our dynamical system $y(t)$ we take in this case: $y_i(0) = 0$ for even vertices i and $y_i(0) > 0$ for odd vertices i. Then, as $dy_i/dt = 2/a > 1$ for odd vertices, the proof follows from Proposition 1.2.3 of [7].

If $N = 2k + 1$, we can proceed similarly. □

Traps for the circle.

Let $b^\star = -2\cos\frac{2\pi[N/2]}{N}$. The traps for the circle have the same structure as the traps for the intervals. We only have to delete the boundary conditions: condition 1c) for $a < 1$; condition 2a) for $1 < a < b^\star$; in condition 2d) we do not have to require $i_1 - 3 > 1$ and $i_1 + 2\beta + 2 < N$.

4. Products of networks

4.1. Tensor products of graphs

In this subsection we assume that the network has an arbitrary self-adjoint operator $A = (a_{ij} = a_{ji})$, where $a_{ij} = 0$ for $j \neq i, j \notin D(i)$ and $a_{ij} > 0$ for $j \in D(i)$, and also an arbitrary vector $\vec{v} = \{v_i > 0\}$.

Let two networks be given with parameters $(A, \vec{v})$ and$(B, \vec{u})$. Note that the graphs of the networks are uniquely defined by the corresponding operators.

We define the tensor product $(A \otimes B, \vec{v} \otimes \vec{u})$ of these two networks as follows. The set of vertices of the graph of the product network is the Cartesian product $V = V_1 \times V_2 = \{(i, j)\}, i \in V_1, j \in V_2\}$ of the sets of vertices of the two graphs.

We further define

$$a_{(i,j),(i_1,j_1)} = a_{i,i_1} a_{j,j_1}, \quad v_{ij} = v_i u_j.$$

Thus two different vertices (i, j) and (i_1, j_1) are connected in the new graph iff i is connected with i_1 (or $i = i_1$) and j is connected with j_1 (or $j = j_1$).

Theorem 4.1.

1. *The spectrum of the operator $A \otimes B$ consists of all products $\lambda_i \mu_j$ of eigenvalues of A and B.*

2. *The eigenvectors of $A \otimes B$ are the tensor products of the corresponding eigenvectors of A and B.*

3. *Define the tensor product $\Lambda_1 \otimes \Lambda_2$ of the faces $\Lambda_1 \subset \mathbf{R}_+^{V_1}$ and $\Lambda_2 \subset \mathbf{R}_+^{V_2}$, in the following way. As before we represent the face Λ_1 by a function on the set of vertices V_1 taking values* 0, 1 *only: this function takes the value* 1 *in a vertex if the corresponding coordinate is positive, and it takes the value* 0 *otherwise. The construction for Λ_2 is similar.*

We allocate 1 to vertex $(i,j) \in V$, if 1 was allocated to at least one of i and j. Otherwise we allocate 0. In other words, $W((\Lambda_1 \otimes \Lambda_2) = W_1 \times W_2$.

It then follows that the solutions of the flow equations on the face $\Lambda_1 \otimes \Lambda_2$ are given by

$$\pi^W = \pi^{W_1} \otimes \pi^{W_2}.$$

4. *Denote by π^A and π^B the solutions to the equations*

$$A\pi = \vec{v}, \quad B\pi = \vec{u}$$

respectively. Let $\lambda_1 > 0\,(\mu_1 > 0)$ be the minimal eigenvalue of $A\,(B)$, and let $\pi^A > 0, \pi^B > 0$ (as functions on the sets of vertices). Then the network $(A \otimes B, \vec{v} \otimes \vec{u})$ is ergodic.

5. *Let the networks $(A, \vec{v})$ and $(B, \vec{u})$ be transient, and let the face Λ_1 be a trap for $(A, \vec{v})$, and Λ_2 for $(B, \vec{u})$. Then $\Lambda_1 \otimes \Lambda_2$ a trap for $(A \otimes B, \vec{v} \otimes \vec{u})$, if it is an ergodic face.*

6. *Let the network $(A, \vec{v})$ be transient and let the face Λ_1 be a trap for $(A, \vec{v})$. Let further $(B, \vec{u})$ be ergodic. Then $(A \otimes B, \vec{v} \otimes \vec{u})$ is transient. Moreover, $\Lambda_1 \otimes (0_B)$ is a trap for $(A \otimes B, \vec{v} \otimes \vec{u})$ (where 0_B is the face corresponding to the origin in the network $(B, \vec{u})$), if it is an ergodic face.*

Proof. Assertions 1–3 of the theorem are evident. Assertion 4 follows from assertion 3 and Theorem 3.1. Assertions 5 and 6 follow from 3. □

The following corollary provides a partial explanation of the numerical results in [1].

Corollary 4.1.

- *The tensor product of a number of ergodic one-dimensional networks (intervals or circles) is ergodic.*
- *If the tensor product of a number of one-dimentional networks (intervals or circles) contains at least one transient network, then the tensor product is transient and the traps (<u>not all of them</u>) can be obtained as the tensor products of traps in the transients factors and of origins in the ergodic factors.*

Proof. The first assertion of the corollary follows from Theorems 3.1, 3.3 and 3.4 and assertions 1, 4 of Theorem 4.1. To prove the second assertion it is sufficient to prove (similarly to assertions 5 and 6 of Theorem 4.1), that the tensor product of traps and/or origins is an ergodic face for the tensor product, when all factors in the tensor product are one-dimensional. □

Lemma 4.1. *The spectrum of $P_W A P_W$ is the union of the spectra for all connected components of W, for any W in the one-dimensional interval.*

Proof of Lemma 4.1. This follows from Theorems 3.3 and 3.4 and the fact that the conditions of Theorem 3.1 are necessary and sufficient for ergodicity in the one-dimensional case. Here we use that $P_W A P_W$ is the direct sum of the operators corresponding to the connected components of W, since each W is the union of intervals. □

We will continue the proof of the corollary. Let us consider the induced chain $M^{W_1} \otimes \ldots \otimes M^{W_r}$, where $W^k = W(\Lambda_k)$ is a trap (or origin) for A_k. It is sufficient to show that the spectrum of $P_{W_1 \otimes \ldots \otimes W_r}(A_1 \otimes \ldots \otimes A_r) P_{W_1 \otimes \ldots \otimes W_r}$ is positive, since then we can apply the last lemma and Theorems 3.1 and 4.1. Indeed, the spectrum of some A_k may not be positive, because some of the factors are transient. But we can write

$$P_{W_1 \otimes \ldots \otimes W_r}(A_1 \otimes \ldots \otimes A_r) P_{W_1 \otimes \ldots \otimes W_r} = \otimes_k [P_{W(\Lambda_k)} A_k P_{W(\Lambda_k)}],$$

and so we can use the last lemma, since the spectrum of each factor in the right-hand side is positive. □

5. Infinite Networks

The preceding results can be used to study infinite (countable) graphs as well, although they cannot be automatically transferred to the infinite case. As an example we will consider only a simple infinite graph: the one-dimensional infinite lattice $V = \mathbf{Z}$ with links $L = \{(i, i+1)\}$ and

$$v_i = v, \ a_{ii} = a, \ a_{ij} = 1, \ i \neq j.$$

Assume that the potentials can also take infinite values (see the compactification above), in which case they remain infinite forever. The problem is the following: if we start with some initial configuration (with finite values in each vertex), what is the limiting invariant measure?

Let us define an *inhibition subset* $I \subset \mathbf{Z}$ as the complement of a subset

$$\ldots < i_k < i_{k+1} < \ldots,$$

such that are one or two vertices between each pair i_k, i_{k+1} of consecutive points in this subset. This inhibition subset can be associated with some *IMAGE* and the corresponding dynamics could be interpreted as the convergence to a given image starting from some approximate image. In the following theorem we will assume $a < 1$. In this case the sets I, $\mathbf{Z} \backslash I$ turn out to have the same structure as the traps in the first part of Proposition 3.1.

Theorem 5.1. *Assume $v = 1, \ a < 1$. At time 0 set $s_i(0) = L, \ i \in I$, for some $L > 0$, and $s_i(0) = 0, \ i \notin I$. Call vertex $i \in I$ regular if $i \in I$ and if $s_i(t) \to \infty$ a.s., as $t \to \infty$.*

Then the vertex $i \in I$ is regular with probability $1-\epsilon$, with $\epsilon = \epsilon(L) \to 0$ as $L \to \infty$.

Any vertex $i \notin I$ has the following property: there exists an event $F = F(L)$ such that the conditional expectation $\mathsf{E}(s_i(t) \mid F)$ is uniformly bounded in t with $\mathsf{P}(F) = 1-\epsilon(L)$, $\epsilon(L) \to 0$ as $L \to \infty$.

Proof. Note first that our process is transient for all dimensions $N > 1$, if $a < 1$. We can prove (using standard cluster expansion techniques, see [9]) that for L sufficiently large there is (with probability 1) an infinite number of finite clusters $C_k \subset I$ (take the distance in the definition of the cluster equal to 2) of I, where $s_i(t)$ will not tend to ∞. Any $i \in I$ belongs to the union of these clusters with some probability tending to 0 as $L \to \infty$. □

The theorem shows that in the one-dimensional case this model is rich in phase transitions related to memory. The above theorem shows that there is a continuum of possible stable configurations. In some sense one can get an "arbitrary" image in this case: let us consider the following coding of images. We allocate zeroes to all points of B, we allocate 1s to all separate points of I and we allocate 2 to all coupled vertices of I. Now (forgetting about zeroes) one can get any *arbitrary* sequence of 1s and 2s, constituting an arbitrary one-dimensional image.

References

[1] M. Cottrell (1992) Mathematical analysis of a neural network with inhibitory coupling. *Stoch. Proc. Appl.* **40**, 103–126.

[2] G. Fayolle, V.A. Malyshev and M.V. Menshikov (1994) *Constructive Theory of countable Markov Chains.* Cambridge University Press, Cambridge.

[3] C. Fricker, Ph. Robert, E. Saada and D. Tibi (1993) Analysis of a neural network model. Rapport de recherche, INRIA, No. 1475.

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

[5] V.I. Kryukov (1976) Wald's identity and random walk models for neuron firing. *Adv. Appl. Prob.* **8**, 257–277.

[6] W.A. Little (1974) The existence of persistent states in the brain. *Math. Biosci.* **19**, 66–98.

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

[8] V.A. Malyshev and M.V. Menshikov (1981) Ergodicity, continuity and analyticity of countable Markov chains. *Trans. Moscow Math. Soc.* **1**, 1–48.

[9] V.A. Malyshev and R.A. Minlos (1991) *Gibbs Random Fields.* Kluwer.

[10] D. Revuz (1975) *Markov Chains.* North Holland.